

---

# Condizioni di coerenza e procedure di coesione nella scrittura scolastica: Il caso dei connettivi

**Lorenzo Zanasi**

*Eurac Research*

lorenzo.zanasi@eurac.edu

**Arianna Bienati**

*Eurac Research/Università degli studi di Modena e Reggio Emilia*

arianna.bienati@eurac.edu

**Jennifer-Carmen Frey**

*Eurac Research*

jennifer-carmen.frey@eurac.edu

**Chiara Vettori**

*Eurac Research*

chiara.vettori@eurac.edu

## Abstract

In questo contributo si analizzano l'uso e lo sviluppo di determinati dispositivi di coesione nel contesto della scrittura scolastica, sullo sfondo del rapporto che intercorre tra le condizioni di coerenza testuale e le procedure di coesione. I dati raccolti provengono dal progetto *ITACA - Coerenza nell'ITALiano Accademico*, un'indagine dedicata all'analisi e alla valutazione della coerenza testuale nelle produzioni scritte di studenti delle scuole secondarie di secondo grado in Alto Adige. Il tema viene affrontato a partire dalle più importanti riflessioni sulla definizione di coerenza che la linguistica testuale ha maturato in decenni di studi e dibattiti, per poi approdare alla discussione di alcune analisi condotte, nell'ambito della stessa indagine, sull'uso dei connettivi nella scrittura scolastica. I risultati convergono sul fatto che scriventi più esperti ricorrono a un minore uso dei connettivi e, al contempo, si avvalgono di forme connettive più ricercate. Se e come questo correli con il grado di coerenza dei testi sarà oggetto di indagini future.

## 1. Introduzione

Il titolo di questo articolo è un chiaro riferimento alle riflessioni teoriche di Maria Elisabeth Conte e di Angela Ferrari (Conte 1988; Ferrari 2010a), che, in tempi diversi ed entrambe con un'attività di ricerca pluridecennale, hanno ragionato sulla testualità e in particolare sulla nozione di coerenza. Di tali riflessioni si riprendono qui due concetti cardine, ovvero quello di condizione (di coerenza) e quello di procedura (di coesione). Il

primo riguarda la possibilità di esistenza di un testo, mentre il secondo riguarda il modo in cui il testo è costruito e costituisce “il rispecchiamento” del precedente. La relazione tra i due concetti è, come noto, asimmetrica, poiché la condizione è *sine qua non* rispetto alla procedura (Andorno 2003: 17). Questa relazione è una squadra utile per tracciare e mettere nella giusta prospettiva percorsi di analisi e ricerca in ambito testuale.

Nelle prossime pagine si mostrerà come si sviluppa l’uso di determinati dispositivi di coesione in un contesto come quello della scrittura scolastica che, rispetto alle condizioni di coerenza, possiede caratteristiche proprie in virtù delle motivazioni degli scriventi e dell’interpretazione dei lettori. In particolare, ci si concentrerà su una classe di elementi coesivi che dovrebbero rispecchiare la coerenza, ovvero i connettivi. Gli studi di pragmatica sperimentale, infatti, hanno spesso testato gli effetti della presenza di questi dispositivi di coesione e verificato la loro influenza positiva sui tempi di lettura e il recupero di informazioni dagli stimoli letti (cfr. Kleijn et al. 2019). Localmente, dunque, i connettivi sembrano aiutare la comprensione e facilitare il processing di frasi. D’altro canto, quando si prendono in considerazione le procedure di coesione nella loro globalità, la quantità di connettivi usati non sembra correlare né con la qualità testuale o l’organizzazione del testo (cfr. Crossley & McNamara 2011), né con lo sviluppo delle abilità di scrittura (cfr. Crossley et al. 2011; Bienati & Frey, in stampa). Questo paradosso apparente è stato uno dei punti di partenza che ha ispirato le analisi presentate in questo contributo, le quali impiegano i dati di un corpus di testi argomentativi raccolto dall’Istituto di linguistica applicata di Eurac Research nel corso del progetto *ITACA – Coerenza nell’ITALiano accademico*,<sup>1</sup> di cui saranno descritti i presupposti, le finalità e le caratteristiche.

A questa introduzione segue una prima parte composta da una sezione (2) dedicata alla nozione di coerenza, articolata in una sintetica disamina sul rapporto che intercorre tra coerenza e coesione (paragrafo 2.1) e in un’analisi della specificità di tale rapporto nel contesto della scrittura scolastica (paragrafo 2.2). Una successiva sezione (3) si sofferma sulle indagini empiriche e valutative nell’ambito della testualità e di alcuni suoi aspetti. In essa è presente una ricognizione sugli studi italiani degli ultimi anni (paragrafo 3.1) e una descrizione del progetto ITACA, di cui si riportano i presupposti, gli obiettivi e i metodi di analisi (paragrafo 3.2 e ss.). A partire dai dati raccolti in questo progetto, nella seconda parte del contributo verrà illustrato il modello teorico della ricerca (paragrafo 4.1) e saranno presentate alcune analisi sullo sviluppo dell’uso dei connettivi, soffermandosi in particolare sul confronto tra gli usi degli studenti e quello di scriventi esperti (paragrafo 4.2), per giungere infine alle conclusioni (sezione 5).

## 2. Sulla nozione di coerenza

In questa sezione si discute la nozione di coerenza testuale in una prospettiva storica e situazionale. Nel primo paragrafo si richiamano alcuni temi del vivace dibattito che ha avuto luogo nella linguistica testuale tra gli anni Settanta e gli anni Novanta del secolo scorso. In particolare, la coerenza viene osservata nel suo rapporto con la coesione. Nel secondo paragrafo si avanzano riflessioni sul concetto di coerenza nella scrittura scolastica.

---

<sup>1</sup> Progetto finanziato dalla Provincia Autonoma di Bolzano (Ripartizione “Diritto allo studio, università e ricerca scientifica”. Legge provinciale 13 dicembre 2006, n.14 “Ricerca e innovazione”).

## 2.1 Sul rapporto tra coerenza e coesione

In numerose lingue moderne, la coppia terminologica coerenza/coesione è stata (e nel linguaggio non specialistico lo è tuttora) semanticamente correlata e, a volte, considerata equivalente (cfr. Zanasi et al. 2024).

La relazione è indubbia, poiché il valore, in senso saussuriano, della testualità, ovvero la sua condizione di esistenza rispetto ad altre forme non testuali, è in molti casi collegato a tratti linguistici concreti che tessono il discorso attraverso rinvii e dipendenze interne ed esterne. Ciò, tuttavia, non è sempre verificato, ovvero la possibilità di relazione tra coerenza e coesione esiste, ma non si manifesta obbligatoriamente in tutti i casi. Come anticipato, infatti, la condizione di coerenza è indipendente da esplicite procedure di coesione, giacché la testualità può emergere attraverso rapporti semantici associativi e inferenziali non segnalati in superficie.

L'equivalenza tra le due nozioni è stata invece dibattuta in linguistica testuale almeno fin dagli anni Settanta del secolo scorso e la modellizzazione del rapporto coerenza/coesione si è articolata lungo un *continuum* che va da un polo in cui l'equivalenza è totale a uno opposto in cui essa è messa in forte discussione. Alla prima concezione si rifanno essenzialmente gli studi che, sulla scia di Halliday & Hasan (1976), postulano la coerenza come *prodotto* e di cui la coesione è parte integrante.<sup>2</sup> Alla seconda concezione appartengono invece quegli studi che presentano una definizione di coerenza fortemente polarizzata sul *processo*, in cui la coesione ha un peso, ma è predominante l'enfasi sull'aspetto dell'interpretazione.<sup>3</sup>

Il dibattito si è articolato in diverse posizioni intermedie che hanno arricchito la riflessione sulle proprietà fondamentali del testo e che sono state raccolte per il pubblico italiano nelle pubblicazioni di Conte, con una rielaborazione personale dell'autrice. La studiosa si è concentrata sulla nozione di coerenza, esaminandone le varie accezioni possibili e domandandosi a) che cosa renda tale un testo e b) che relazioni intrattenga il testo con la lingua e con il mondo extralinguistico. A questi interrogativi Conte risponde distinguendo le proprietà che esprimono la *qualitas* del testo, da quella che cattura la sua *quidditas*, ovvero l'essenza.

Una qualità "privativa" o negativa è espressa dalla coerenza intesa come assenza di contraddizione. Quando questa caratteristica viene meno, emerge un'incoerenza tra determinate relazioni semantiche all'interno del testo. Tuttavia, la contraddittorietà agisce a livello locale; essa, infatti, può essere presente all'interno del testo senza per questo snaturarlo e trasformarlo in un non-testo.

Una seconda qualità della coerenza risiede nel suo manifestarsi esplicitamente nel testo attraverso "mezzi linguistici" che intrinsecamente la segnalano (coesione semantico-tematica e connessità formale). A rendere un testo un insieme coerente di enunciati con un'architettura riconoscibile è una proprietà in questo caso positiva: l'elemento semantico della continuità, la connessione delle parti in un tutto, che Conte (1988) tematizza in comunanza con la *Textlinguistik* di area continentale. La continuità del senso non è, tuttavia, esclusivamente inerente ai mezzi linguistici di cui si compone il testo (la cosiddetta coerenza *a parte obiecti*). In massima parte, essa è insita nell'interpretazione

---

<sup>2</sup> Cfr. anche Daneš (1974), Reinhard (1980).

<sup>3</sup> Cfr. le posizioni di Kintsch & Van Dijk (1978), Kehler (2002) e, soprattutto, Brown & Yule (1983: 199) per cui la coerenza non è qualcosa che pertiene al testo, bensì a chi lo interpreta: "texts are what hearers and readers treat as texts". Gli autori sottolineano quindi l'importanza delle conoscenze pregresse dei destinatari nell'interpretazione del messaggio.

che ne dà il destinatario del testo che farà di tutto per trovarvi quella che Hörmann (1976) definisce una *Sinnkostanz*, ossia una ‘costanza del senso’ (coerenza *a parte subiecti*).

Per Conte (1993), dunque, la coerenza è anche il risultato dell’attività costruttiva dell’interprete, il quale “trae inferenze, costruisce anelli mancanti e, nel processo interpretativo, fa intervenire le sue conoscenze enciclopediche, le sue credenze e i suoi atteggiamenti valutativi”.<sup>4</sup>

## 2.2 Coerenza e coesione nella scrittura scolastica

La possibilità di esistenza di testi nella scuola, e quindi la loro coerenza, è determinata da un contesto in cui l’esercizio della scrittura ha una funzione solo indirettamente comunicativa. Gli scopi principali sono, infatti, di natura dimostrativa: dar prova di aver acquisito competenze disciplinari, di gestire l’uso del vocabolario e l’appropriatezza dei registri, di possedere la capacità di strutturare la concatenazione dei pensieri per iscritto. Inoltre, la scrittura scolastica è per lo più programmata e, potremmo dire, prevista “su commissione”. Se da un lato queste caratteristiche la rendono scarsamente spontanea e motivante, dall’altro l’allenamento fin da bambini alla scrittura per scopi specifici fa sì che gli elaborati prodotti siano testi formalmente organizzati e che, salvo casi eccezionali, non ci sia deviazione da un modello normativo di produzione testuale. Accertata, rispetto alla coerenza, la quiddità dello scritto scolastico, sarà allora fondamentale esaminarne le qualità: si possono rinvenire globalmente o localmente delle violazioni del principio di non contraddizione in questi testi? Come sono gestite dagli scriventi e come maturano nel corso del percorso educativo le procedure di coesione, ovvero i mezzi linguistici che sostanziano la coerenza? Con quali risultati? E ancora: sul fronte dell’interpretazione del testo, che atteggiamento dovrà assumere il destinatario sapendo di essere in una posizione fortemente asimmetrica rispetto al produttore sia per ruolo che per età?

Questi motivi rendono difficile centrare e valutare il grado di coerenza testuale nell’ambito dell’istruzione. Si tratta di un tema su cui, per mancanza di strumenti dedicati, non si sono condotti, nel contesto italiano, molti studi empirici. Numerose indagini sono, invece, state compiute nel settore più ampio della descrizione della valutazione di fenomeni testuali eterogenei.

## 3. Ricerche empiriche su testualità e coerenza

In questa sezione si richiamano le principali ricerche e indagini empiriche di ambito italiano che hanno avuto come oggetto di studio la testualità, le competenze di scrittura e alcuni elementi linguistici di natura testuale. Ci si concentra quindi sul progetto *ITACA – Coerenza nell’ITAliano Accademico*, di cui si ripercorrono motivazioni e procedimento, per poi riflettere sugli esiti di alcune analisi condotte su una particolare classe di dispositivi di coesione: i connettivi.

---

<sup>4</sup> Idea maturata a seguito della pubblicazione di De Beaugrande & Dressler (1981) e condivisa in quel periodo anche da Charolles (1983).

### 3.1 Alcuni risultati degli studi sulla testualità in Italia

Per quanto concerne la realtà italiana, le ricerche più estese e standardizzate riguardano la lettoscrittura e i risultati sono mediocri. Nell'analisi dei primi dieci anni di rilevazioni dei test PISA/OCSE (Asquini 2011) si osserva, infatti, come i punteggi medi in lettura collochino l'Italia sistematicamente al di sotto della media internazionale, situazione che non migliora nelle rilevazioni del 2012 e del 2015.

Anche i rapporti INVALSI (2010-2018) denunciano una scarsa dimestichezza degli studenti con la comprensione di testi espositivi e argomentativi e, dal punto di vista della produzione scritta, con aspetti testuali ben definiti, fra cui quello della coerenza. Il dato è confermato dalle valutazioni della prima prova d'italiano agli esami di Stato a cura degli accademici della Crusca (INVALSI 2012). Le analisi indicano, infatti, che il 29,4% degli errori riscontrati è costituito da errori di tipo testuale. In quest'area di competenza 436 prove su 499 hanno evidenziato la presenza di errori in ambito testuale: nel 71% dei casi l'errore è relativo al descrittore "enunciazioni coerenti in sé", mentre nel 40,8% dei casi l'errore riguarda il descrittore "enunciazioni che contribuiscono alla costruzione di un testo complessivamente coerente".

Al di là degli allarmismi mediatici che puntualmente si fanno strada dopo la pubblicazione dei risultati ora dell'una ora dell'altra indagine (cfr. Balboni 2017: 7), vi è un'effettiva fragilità nelle competenze testuali degli studenti italiani di età preadolescenziale e adolescenziale, segnalata, in realtà, da tempo dai settori della linguistica più sensibili ai temi educativi (cfr. Boscolo & Zuin 2015; De Mauro 2010 per un excursus delle maggiori ricerche; Lavinio & Sobrero 1991; Colombo 1992 e Corno 2002 per il periodo antecedente all'avvento dei nuovi media digitali; Pallotti 1998 per le considerazioni sui processi di alfabetizzazione del pensiero).

Proprio per questo, anche a livello nazionale cominciano ad affacciarsi studi sulla coerenza testuale: ci si riferisce qui in particolare a Tarallo (2019) che ha testato, sull'italiano scritto di studenti della scuola secondaria di I grado, la scala di valutazione della coerenza in testi argomentativi sviluppata da Kuiken & Vedder (2017) in uno studio sull'adeguatezza funzionale. Esistono poi studi empirici eterogenei per scopi (accuratezza linguistica/qualità del testo), destinatari (ciclo scolastico primario/secondario di I e II grado) e metodi di analisi (manuali/computazionali). Fra le ricerche degli ultimi anni che includono anche un focus sulla testualità, si ricorda Ruggiano (2011) che esamina 380 temi raccolti nel ciclo triennale della scuola secondaria di I grado, De Santis & Gatta (2013) che sondano le competenze linguistico-grammaticali, lessicali e testuali di 153 matricole universitarie, Cisotto & Novello (2012) che esplorano le "componenti di elaborazione concettuale della scrittura di sintesi" di 150 studenti del 1° anno di Scienze della Formazione Primaria e Boscolo & Zuin (2015) e Ruele & Zuin (2020) che hanno esplorato la scrittura degli studenti trentini, nel primo caso con un'estesa indagine sulla qualità di testi argomentativi e sintesi raccolti nelle scuole superiori, nel secondo con un'analisi linguistica di circa tremila testi della prova di italiano dell'Esame di Stato. Ancora, sulla scrittura degli universitari sono imminenti i risultati del progetto UniverS-Ita coordinato da Nicola Grandi (<https://site.unibo.it/univers-ita/it>, consultato il 23/11/2024).

Un filone di studi assai recente è quello che tenta di valutare le componenti della testualità tramite approcci di NLP (*Natural Language Processing*). In questo settore, gli sforzi per migliorare la misurazione della coerenza si concentrano ultimamente sull'identificazione dei coesivi e di altri indici di coesione del testo (cfr. Crossley et al. 2016), sul riconoscimento e la classificazione di relazioni discorsive esplicite e implicite

(cfr. Lin et al. 2009; Xue et al. 2016; Shi et al. 2017) e sulla modellazione automatica della coerenza, ovvero la classificazione del grado di coerenza o incoerenza di un testo.

Tuttavia, come spesso accade in ambito NLP, la maggior parte delle ricerche finora è stata svolta sull'inglese. Alcune ricerche di ambito italiano impiegano metodi automatici di analisi, ma non sono specificamente indirizzate al comparto testuale, come ad esempio Barbagli et al. (2016) in cui si presenta il corpus CltA, parte di uno studio dell'IEA (*International Association for the Evaluation of Educational Achievement*) e dedicato all'analisi di 1352 testi di vario tipo fra cui anche quelli argomentativi ed espositivi (con un focus sull'uso dei connettivi), Borghi (2013), tesi di dottorato sull'analisi e la valutazione di 1474 testi di un campione di studenti degli istituti secondari di II grado di Roma con misure automatizzate di carattere lessicale e morfosintattico, Sprugnoli et al. (2018) con le prime analisi di un'osservazione longitudinale sulle abilità di scrittura dei maturandi trentini, con particolare focus sui tratti dell'italiano neostandard.

### 3.2 Il progetto ITACA: obiettivi e presupposti

In questa generosa messe di studi si inserisce il progetto *ITACA - Coerenza nell'ITALiano Accademico*, svolto tra il 2020 e il 2023 dall'Istituto di linguistica applicata di Eurac Research, con il duplice obiettivo di:

- 1) esaminare e valutare la coerenza sulla base di un corpus di testi di tipo argomentativo scritti da un campione statisticamente rappresentativo<sup>5</sup> di studenti del IV anno della scuola secondaria di II grado, in lingua italiana, della Provincia di Bolzano.
- 2) verificare se e in che modo una serie di fattori motivazionali, sociolinguistici e comportamentali influenzino la scrittura accademica di tipo argomentativo del campione.

Una serie di riflessioni ha motivato la ricerca. Innanzitutto, la consapevolezza che studiare le competenze di scrittura dei cittadini di domani, le quali presuppongono quella competenza alfabetica funzionale necessaria per “afferinarsi attraverso la lingua in modo convincente, critico e sicuro” (cfr. ANP 2018), è un'urgenza avvertita da più parti, proprio per il ruolo fondamentale che la scrittura svolge nella vita, professionale e non, di ciascuno (cfr. Light 2001; McNamara et al. 2010). Ciononostante, la ricerca sulla scrittura scolastica è in ritardo rispetto a quella sulle competenze ricettive (cfr. Hodges et al. 2016: 10), in parte per ragioni sociali (è prioritario saper leggere un testo, rispetto al saperlo produrre), in parte per ragioni metodologiche relative alla complessità della scrittura come oggetto di studio (cfr. Boscolo & Zuin 2015: 5).

Inoltre, la non lusinghiera infrequenza di indagini empiriche sullo stato della lingua scritta nella scuola secondaria di secondo grado italiana è stata messa in luce (peraltro poco prima della stesura del progetto ITACA) in una lettera sottoscritta da numerosi linguisti (cfr. Lo Duca 2017) e concepita in risposta a un appello siglato da altri accademici (cfr. Gruppo di Firenze 2017), preoccupati per lo stato di salute dell'italiano tra gli universitari. A questo quadro, si aggiunge la specificità del contesto altoatesino in

---

<sup>5</sup> Da un punto di vista procedurale, a partire da una popolazione di 1308 studenti iscritti al quarto anno di uno dei licei o istituti tecnici in lingua italiana della Provincia di Bolzano, è stato effettuato un campionamento a *cluster* o “grappoli” stratificati proporzionalmente per tipo di scuola (licei/istituti). Le classi (*cluster*) sono state selezionate casualmente fino a raggiungere il numero teorico di soggetti previsti per ogni strato.

cui le rilevazioni linguistiche più articolate hanno – in parte comprensibilmente – per lo più guardato alla relazione tra il tedesco e l’italiano e il plurilinguismo (cfr. Vettori 2021; Zanasi 2021).

Infine, la ricerca ha preso le mosse anche dalla constatazione che le valutazioni della coerenza testuale in ambito scolastico derivano in molti casi da griglie in cui coesione e coerenza sono unificate e graduate da indicatori di scarsa granularità (cfr. Bienati et al. 2022).

### 3.2.1 Il design della ricerca

Per rispondere ai due obiettivi del progetto, dalla popolazione scolastica complessiva delle scuole in lingua italiana dell’Alto Adige nell’A.S. 2021/2022 è stato estratto un campione di studenti iscritti alle classi quarte che fosse statisticamente rappresentativo.

Al campione sono state somministrate due differenti prove e un questionario. Innanzitutto, è stato chiesto di scrivere un testo argomentativo sotto forma di lettera, connotata da un registro formale,<sup>6</sup> sul tema della DDI (Didattica Digitale Integrata). Al compito sono stati posti vincoli di parole (minimo 600) e di tempo (massimo 100 minuti).

In secondo luogo, ai partecipanti è stato somministrato un test di competenza generale dell’italiano L1, individuato nel Test di Verifica Iniziale (da ora anche TVI) dell’Università di Bergamo (cfr. Grassi & Nuzzo 2012). Il test, progettato per le matricole universitarie, si è rivelato adeguato anche al target di età del campione, composto nel 70% dei casi da diciottenni.

Infine, agli studenti è stato richiesto di compilare un questionario di 45 domande, sia per sondarne gli atteggiamenti nei confronti della lettura (cfr. Schutte & Malouff 2007) e della scrittura (cfr. Troia et al. 2012), le loro abitudini di lettura sia in ambito scolastico e soprattutto extrascolastico (cfr. ISTAT – Indagine multiscopo sulle famiglie. Aspetti della vita quotidiana anno 2019/Sezione: Letture; I cittadini e il tempo libero 2015), sia per ottenere informazioni sociodemografiche e sociolinguistiche (domande adattate da Vettori & Abel 2017).

Tutti i dati sono stati raccolti in forma digitale e anonima e sono stati sottoposti a successive procedure di controllo e di pulizia, in parte manuali, per assicurarne l’integrità e per garantire l’anonimato dei partecipanti. Dopo la chiusura del progetto, i dati saranno accessibili alla comunità scientifica tramite la piattaforma PORTA di Eurac Research (<https://www.porta.eurac.edu/>) e tramite l’apposito research data repository, Eurac Research Clarin Center (ERCC), che fa parte della rete CLARIN. Tutti i dati saranno consultabili e scaricabili con una licenza che permette l’uso per finalità di ricerca (CLARIN ACADEMIC END-USER LICENCE ACA-BY-NC-NORED 1.0). Il deposito dei dati sul repository permette, inoltre, l’archiviazione e preservazione dei dati digitali per garantire la riproducibilità a lungo termine dei risultati.

---

<sup>6</sup> Il destinatario della missiva era il Ministro dell’Istruzione.

### 3.2.2 L'analisi dei dati

La documentazione raccolta per ognuna delle tre prove è stata analizzata in modi diversi. I 635 testi argomentativi sono stati valutati in doppio cieco da un gruppo di sei esperti che hanno applicato una scala creata *ad hoc* per valutare la coerenza testuale.

Un sottoinsieme di testi (388) è stato, inoltre, analizzato e annotato relativamente a una serie di tratti testuali, tra cui punteggiatura, accordi, anafore, contraddizioni e connettivi, oggetto di questo contributo. Per l'annotazione è stato impiegato INCEPTION (cfr. Klie et al. 2018), uno strumento multilivello (ovvero che permette di marcare vari tratti testuali su diversi *layers*), implementato per gli scopi del progetto grazie a una collaborazione con la Fondazione Bruno Kessler. La totalità dei testi raccolti costituisce il corpus ITACA (424.693 token).

I questionari sono stati analizzati mediante il software statistico SPSS e hanno restituito diverse informazioni per descrivere sia le caratteristiche del campione (genere, età, indice di status socioeconomico-culturale, tipo di scuola frequentata, ecc.) sia quei fattori extralinguistici che potrebbero influenzare le produzioni scritte raccolte.

Questi ultimi dati, infatti, sono stati incrociati con le valutazioni dei testi prodotti dal campione, in modo da metterne in luce eventuali rapporti di correlazione e da evidenziare differenze statisticamente significative nella performance degli studenti riferibili alle variabili prese in esame.

Per quanto riguarda i TVI, i loro risultati sono stati messi in relazione sia con le dimensioni valutate nei testi, sia con le caratteristiche sociodemografiche del campione, come il genere e il tipo di scuola frequentata (liceo versus istituto tecnico/professionale).

## 4. Esempi di analisi sui connettivi nella scrittura scolastica

In questa sezione si introduce il modello teorico su cui si basa l'indagine empirica di ITACA, soffermandosi poi su una specifica categoria di coesivi, i connettivi, che sono stati oggetto di alcune delle analisi svolte nell'ambito del progetto.

### 4.1 Il modello basilese della coerenza

Nel progetto ITACA, le principali urgenze teoriche-metodologiche sono state a) quella di identificare un modello teorico che fosse sufficientemente duttile per coprire le condizioni di coerenza e le procedure di coesione tipiche del testo scolastico e b) quella di stabilire una procedura che permettesse di passare dal modello teorico della coerenza a uno strumento di valutazione della stessa.

Rispondere alla prima urgenza ha significato trovare un modello capace di "misurare" la gradualità della coerenza testuale attraverso una descrizione dettagliata di questa nozione. L'obiettivo è stato raggiunto adottando un approccio teorico che ha origine nella tradizione contiana degli studi sulla coerenza: il cosiddetto "modello basilese" (cfr. Ferrari 2014; Ferrari et al. 2021). Il modello poggia le sue basi sui concetti semantici dell'unità, della continuità e della progressione testuale (cosiddetta "coerenza di fondo"), intorno ai quali, sorrette da una struttura che prende il nome di segmentazione, si sviluppano le dimensioni in cui il discorso si produce: quella logico-argomentativa, quella tematico-referenziale e quella chiamata enunciativo-polifonica.

Rispondere alla seconda urgenza è stato più complesso perché ha significato operazionalizzare la teoria, cioè renderla applicabile ai casi concreti (selezionando alcuni elementi e tralasciandone altri) al fine di ottenere una scala che fornisse ai valutatori la possibilità di esprimere giudizi sul testo e di trasformarne le valutazioni in un punteggio assegnabile, per ciascuna dimensione presa in esame, alle produzioni scritte raccolte. Inoltre, la scala doveva integrare gli aspetti linguistici, descritti minuziosamente dal modello basilese, con aspetti procedurali che facessero riferimento alle contingenze pragmatiche tipiche del contesto scolastico (il genere testuale, l'occasione, il target di riferimento) nel quale i testi in esame sono stati prodotti.

Lasciando ad altre future pubblicazioni il compito di raccontare nel dettaglio il processo di costruzione della scala, in questa sede vale la pena richiamare da Zanasi et al. (2024) la versione applicata del modello basilese, utile a introdurre l'analisi dei connettivi causali.

L'architettura complessiva del modello è rappresentabile come un castello (Figura 1), alle cui fondamenta si colloca la coerenza di fondo che, insieme alla segmentazione che gerarchizza le unità comunicative del testo, sostiene i piani superiori (tematico-referenziale, logico-argomentativo ed enunciativo-polifonico), rappresentati come ali del castello. Ogni ala ha stanze "dedicate" a fenomeni micro (ovvero gli elementi coesivi puntuali come i connettivi e le anafore) e a fenomeni macro (gli elementi coesivi globali come la progressione *topic-comment* e le relazioni discorsive). Il modello basilese applicato permette dunque di leggere un testo identificando i livelli che lo compongono e di verificare se esso è semanticamente continuo, unitario e progressivo, se la punteggiatura e la paragrafazione che lo accompagnano segnalano correttamente la segmentazione degli enunciati, se gli elementi coesivi sono utilizzati in modo opportuno.

All'interno del castello della coerenza, i connettivi occupano il piano logico-argomentativo. Per Ferrari (2021: 145), infatti, "il termine connettivo indica ciascuna delle forme linguistiche morfologicamente invariabili che offrono istruzioni su come legare gli eventi evocati dal testo o gli atti linguistici di composizione testuale attraverso relazioni logico-argomentative quali la causa, la consecuzione, la riformulazione, l'esemplificazione, la concessione, l'opposizione ecc."

In questa definizione di natura formale e funzionale sono contenute due restrizioni in rapporto tra loro: non vanno considerati connettivi "elementi che pur essendo parole grammaticali invariabili, non indicano relazioni logico-argomentative" né, viceversa, quelle espressioni che, "pur essendo associate a una relazione logico-argomentativa, sono morfologicamente variabili" (cfr. Ferrari 2021: 146). A questa definizione se ne affiancano altre,<sup>7</sup> a testimonianza del fatto che la categoria dei connettivi è assai complessa da descrivere e delimitare.

---

<sup>7</sup> In Palermo (2013: 191-192), per esempio, un connettivo è "qualsiasi elemento di connessione, indipendentemente dalla portata e dalla funzione del collegamento".

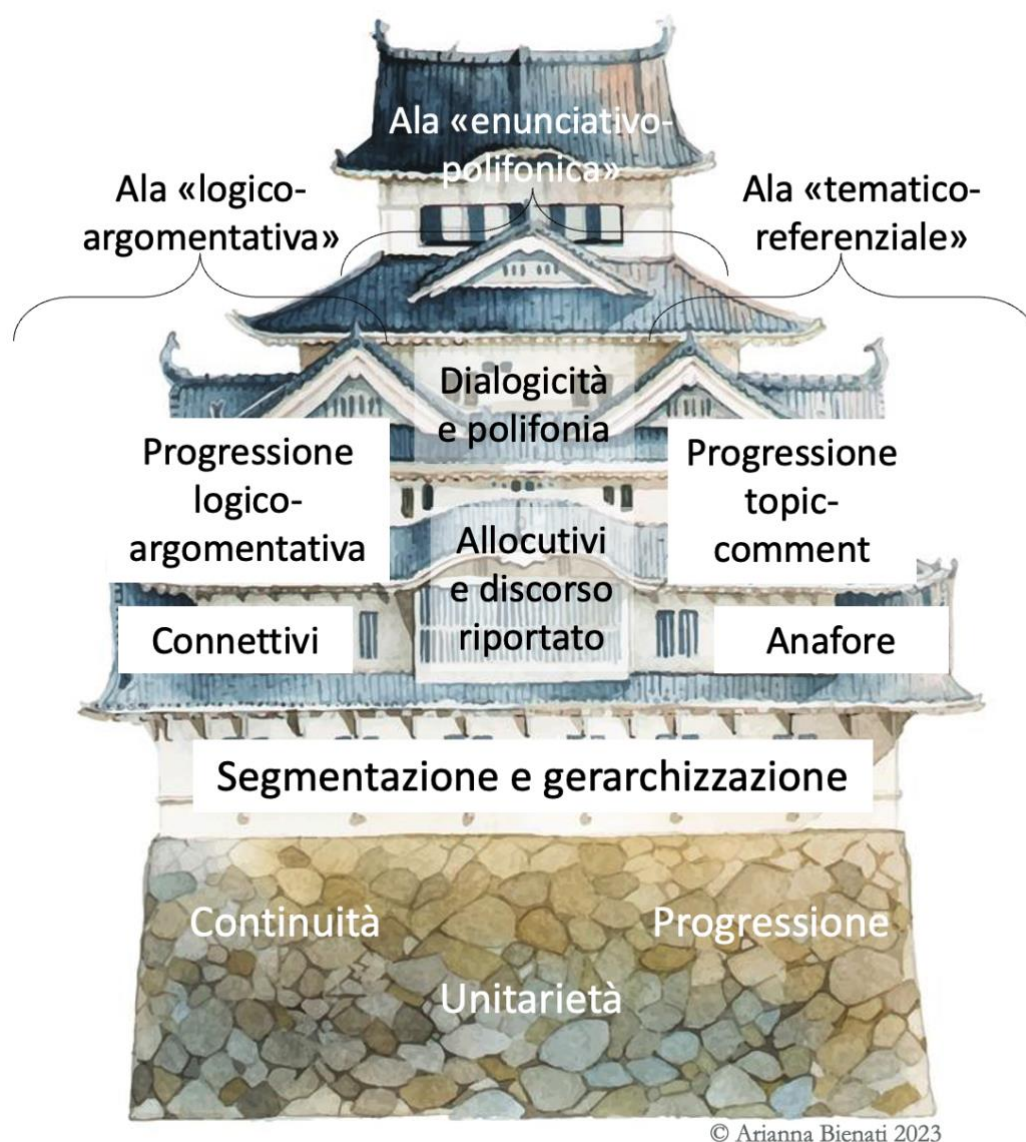


Figura 1. Il castello della coerenza

#### 4.2 I connettivi nella scrittura scolastica

Proprio il tentativo di rendere operativa e automatizzabile la definizione di connettivo proposta da Ferrari (2021) guida la prima delle tre analisi che presentiamo. Come mostrato nella rappresentazione grafica, tutti i livelli superiori del castello poggiano sulla segmentazione e gerarchizzazione delle unità discorsive. Anche l'identificazione dei connettivi poggia su queste fondamenta, dal momento che l'atto stesso di segmentare e gerarchizzare il testo individua le unità semantiche tra cui possono sussistere relazioni logico-argomentative, le quali, a loro volta, possono essere segnalate esplicitamente da connettivi. Nel modello basilense, le unità semantiche di base sono le unità informative, gli enunciati e i movimenti testuali; le relazioni discorsive possono esistere tra unità dello stesso livello o tra unità di livelli diversi. Un approccio molto simile è adottato dalle

maggiori risorse computazionali annotate per relazioni discorsive. Un esempio preminente è la PDTB, ovvero la *Penn Discourse Treebank*, (cfr. Prasad et al. 2008; Webber et al. 2019) che adotta un’annotazione del discorso basata su “a variety of lexical elements that take clausal elements as arguments and convey discourse-level predication on those arguments” (cfr. Webber 2004: 3). Tuttavia, anche tra risorse costruite con lo stesso approccio ci possono essere discrepanze tra i criteri di segmentazione<sup>8</sup> e, a seconda di come si decide di segmentare, l’identificazione degli elementi di superficie che segnalano esplicitamente le relazioni discorsive e la loro progressione sul piano logico-argomentativo cambia.

Al contempo, pur rimanendo fedeli alla definizione di connettivo proposta da Ferrari, sembrerebbe non essere univoco ciò che viene considerato “morfologicamente invariabile” dalle diverse risorse disponibili per l’annotazione dei connettivi in italiano. In Bienati et al. (2023) si utilizza per l’identificazione dei connettivi il *Lexicon of Italian Connectives* (LICO, cfr. Feltracco et al. 2016). La risorsa adotta la definizione di connettivo di Ferrari (2010b), allargando leggermente le maglie del criterio di invariabilità morfologica. Nella risorsa, infatti, si possono trovare sia connettivi che mostrano una variazione sull’asse paradigmatico, come nel caso di ‘per esempio’ e ‘ad esempio’, in cui le due espressioni multi-parola differiscono per la diversa preposizione usata, sia connettivi che mostrano una variazione sull’asse sintagmatico (come ‘per questo motivo’) e quindi permetterebbero potenzialmente sia la flessione al plurale (‘per questi motivi’), sia l’inserzione di altro materiale linguistico (‘per tutti questi motivi’), varianti che però non sono presenti nel lessico.

Nello studio si testano le performance di un metodo automatico molto semplice, basato sul *match* di stringhe tra le forme connettive presenti nel LICO e quelle presenti in un campione di testi tratti dal corpus ITACA, descritti in questo contributo, e dal corpus LEONIDE (cfr. Glaznieks et al. 2022), precedentemente annotati da umani. I risultati suggeriscono che un sistema così semplice ha sì un buon *recall* (0.91), segno che la risorsa lessicale utilizzata è molto completa e sicuramente sufficiente per l’identificazione dei connettivi in testi scolastici, ma restituisce una scarsa *precision* (0.72),<sup>9</sup> principalmente per due motivi. In primis, la semplicità del sistema automatico non permette di gestire la multicategorialità di alcune forme connettive che possono avere diverse funzioni sintattiche in relazione al contesto in cui appaiono; in secondo luogo, alcuni connettivi possono essere attivi su più livelli, sia a livello testuale/discorsivo (rientrando quindi a pieno titolo tra i connettivi) sia a livello di frase (si pensi, ad esempio, alla ‘e’ che connette elementi nello stesso sintagma nominale che, seguendo le linee guida di annotazione della PDTB 3.0, non è stata segnalata come connettivo nell’annotazione manuale). Un *match* di stringhe automatico non è in grado di distinguere questi due livelli e, infatti, ha i tassi di errore più alti su quelle forme connettive che operano a vario titolo su più livelli, come ‘e’, ‘ma’ o ‘o’.

Un altro studio (cfr. Bienati & Frey 2023), condotto a partire dai risultati emersi da questo primo esperimento, si concentra proprio sul connettivo che ha registrato i più alti

---

<sup>8</sup> Si vedano le discrepanze relative all’annotazione delle congiunzioni che coordinano sintagmi verbali o introducono frasi infinitive nelle due versioni della PDTB (cfr. Prasad et al. 2008; Webber et al. 2019), così come nelle risorse ispirate alla PDTB, come il LUNA corpus (cfr. Tonelli et al. 2010)

<sup>9</sup> In questo specifico contesto, con *precision* si intende il rapporto fra i connettivi identificati correttamente dal modello e tutti i connettivi identificati dal modello (correttamente e non); *recall*, invece, è il rapporto fra i connettivi correttamente identificati dal modello e il numero totale dei connettivi, sia quelli identificati automaticamente sia quelli ignorati dal modello e individuati dall’annotazione manuale.

tassi di errore di identificazione da parte del sistema automatico, ovvero la congiunzione ‘e’. La ‘e’ è uno dei connettivi con la più alta polivalenza funzionale, tanto che il DISC online riporta almeno 6 relazioni discorsive potenzialmente introdotte da esso (vedi tabella 1).

Nomenclatura dizionari	Nomenclatura PDTB	Treccani	Sabatini-Coletti	De Mauro	LICO
Contemporaneità	Temporal: Synchronous		x		
Contiguità/aggiunta	Expansion: Conjunction	x	x	x	x
Contrasto	Comparison: Contrast	x		x	
Successione temporale	Temporal: Asynchronous: Precedence		x		x
Condizione	Contingency: Condition		x		
Conseguenza	Contingency: Cause: Result		x		
Scopo	Contingency: Cause: Result + goal		x		

Tabella 1. Relazioni discorsive che possono essere segnalate dalla congiunzione ‘e’ secondo i dizionari Treccani (online), Sabatini-Coletti (online), De Mauro (online) e secondo il *Lexicon of Italian Connectives* (LICO, cfr. Feltracco et al. 2016).

Nei testi di ITACA, e più in generale nei testi prodotti da studenti, sembra particolarmente interessante indagare gli usi di questa congiunzione, innanzitutto perché dalla lettura dei testi degli studenti essa emerge, a livello qualitativo, come tratto saliente e potenzialmente peculiare delle loro produzioni. Queste prime rilevazioni sono suffragate da precedenti studi sull’uso di connettivi semanticamente generici da parte di studenti del ciclo superiore e universitari. Si veda, ad esempio, Calamai (2012: 86), che nota come nei testi da lei analizzati

i connettivi semanticamente più poveri hanno le frequenze più elevate: le preferenze degli studenti mostrano una variazione lessicale molto ridotta e si avvicinano, anche da questo punto di vista, a una caratteristica tipica del parlato, in cui c’è poca dispersione tra i diversi tipi, cioè la presenza di molte occorrenze di pochi tipi.

Che i connettivi “semanticamente più poveri” (ovvero quelli più polifunzionali) registrino le frequenze più alte non sorprende: più un connettivo è “tuttofare” e più occorrerà in diversi contesti d’uso, determinando così la sua alta frequenza in ogni tipo di corpus che

si analizzerà. Quello che invece potrebbe sorprendere è l'uso di questo connettivo per veicolare relazioni discorsive più specifiche di quelle normalmente ad esso associate, così come la presenza di eventuali differenze tra l'uso che ne fanno gli studenti rispetto a quello degli adulti. Nello studio di Bienati & Frey (2023), l'obiettivo è stato dunque verificare se l'uso della 'e' per introdurre una relazione causale fosse una strategia effettivamente più presente nei testi di studenti rispetto a testi scritti da adulti, definibili come scrittori esperti.

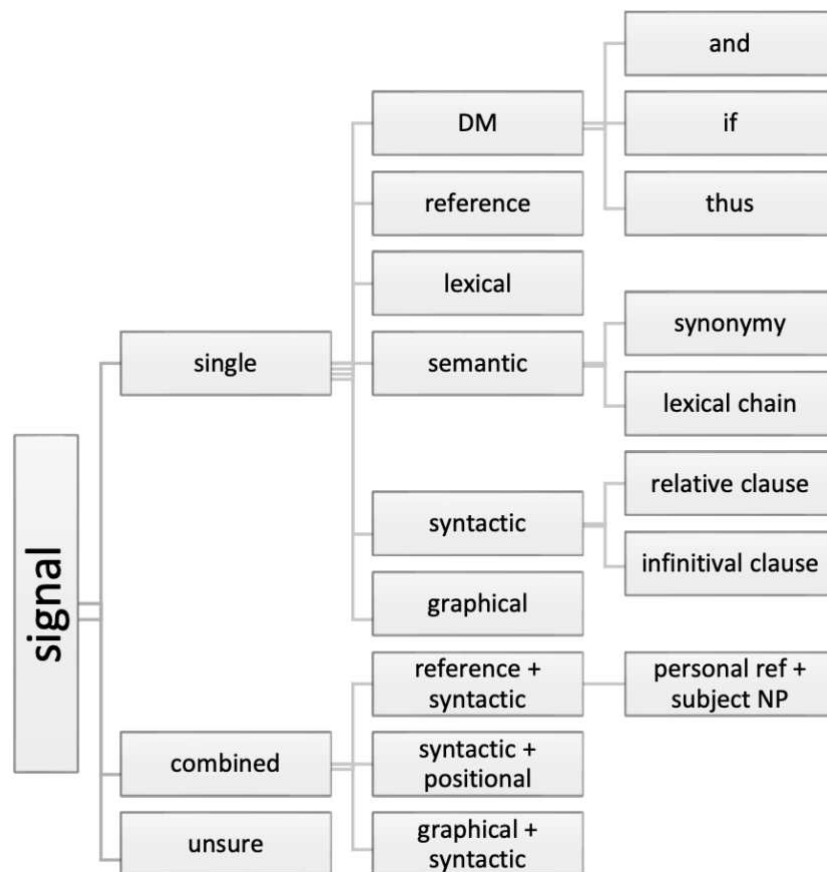


Figura 2. Tassonomia dei marcatori che segnalano relazioni discorsive in modo alternativo rispetto ai connettivi (in figura 'DM'). Visualizzazione tratta dallo schema di annotazione di Das & Taboada (2014: 6).

In seconda battuta, lo studio ha preso in considerazione eventuali altri elementi con cui si può segnalare la presenza delle relazioni discorsive oltre ai connettivi. Le tassonomie di Das & Taboada (2014: 6) qui riportate in Figura 2, per quanto si appoggino alla sistematizzazione di Halliday & Hasan (1976) e mostrino dunque alcune differenze sostanziali rispetto al modello basilese, sono un ottimo esempio di come esistano svariate strategie (lessicali, sintattiche, referenziali, grafiche) per segnalare le medesime relazioni discorsive. Ferrari stessa (2014: 282 e ss.) discute estensivamente la possibilità di segnalare le relazioni discorsive anche in modi alternativi (es. con verbi e nomi), diversi dai connettivi. Queste strategie alternative possono occorrere sia isolatamente sia in combinazione con i connettivi, secondo Hoek et al. (2019: 3), seguendo un meccanismo di compensazione:

If another element within a discourse segment already signals or partly signals how that segment should be related to another segment from the discourse, this would eliminate or reduce the amount of information a connective would contribute. Other linguistic elements that convey information or raise expectations about the type of coherence relation that should be constructed are thus expected to influence the marking of coherence relations by connectives.

A partire da queste indagini, si è voluto comprendere se l'utilizzo di un connettivo così semanticamente generico come 'e' in funzione di marcatore di relazione discorsiva causale si accompagni ad altri marcatori che si aggiungono per specificare ulteriormente la funzione causale che sussiste tra le due unità discorsive e se ci siano differenze nella frequenza con cui questi elementi aggiuntivi sono usati da apprendenti e da scrittori esperti.

I risultati del confronto mostrano come effettivamente, all'interno del campione preso in esame, gli studenti sembrano usare con più frequenza il connettivo generico 'e' in concomitanza con una relazione discorsiva più specifica, come quella causale (su 1000 occorrenze di 'e' analizzate, 101 sono causali), rispetto agli esperti (su 1000 occorrenze di 'e' analizzate, 68 sono causali). Tuttavia, questo risultato va letto in relazione agli altri elementi che appaiono insieme alla 'e'. Gli studenti, infatti, sembrerebbero avere la tendenza a non lasciare sottospecificata la relazione discorsiva causale, ma a specificarla ulteriormente nella maggioranza dei casi (53%); gli esperti, invece, la specificano meno frequentemente (38% dei casi). Nella maggioranza dei casi osservati, quindi, essi usano unicamente la 'e' per segnalare la relazione discorsiva causale. Le strategie con cui la 'e' co-occorre in contesti di relazioni causali sono: altri connettivi (singole parole o espressioni multiparola morfologicamente invariabili) (1), espressioni multiparola (variabili o invariabili) che contengono incapsulatori che riprendono o l'intero contenuto proposizionale dell'unità discorsiva precedente (2a), o un singolo elemento dell'unità discorsiva precedente (2b); strategie lessicali (3); strategie referenziali e lessicali (4); e, infine, sintattiche (5).

- (1) a. *Si tratta di questioni troppo ampie per essere trattate qui e, dunque, non proverò neppure a rispondere.* (esperto)  
b. *a distanza il professore non può avere il pieno controllo degli studenti, e quindi non può verificare se essi stiano copiando oppure no.* (studente)
- (2) a. *Il nostro futuro si concentra sulle decisioni che prendiamo in questo momento e per questo motivo dobbiamo dare una possibilità a tutti gli studenti di avere una formazione piena di conoscenza (...).* (studente)  
b. *è lapalissiano anche solo intuire che in DAD i professori sommergano di compiti gli studenti e proprio a causa dei troppi compiti essi di tempo libero non ne avranno mai.* (studente)
- (3) *Che se i sordi provavano a entrare nel mondo degli udenti era spesso un fallimento, e il motivo è che non conoscevano bene l'italiano e gli udenti non li capivano e incontravano delle difficoltà.* (esperto)

- (4) a. *A differenza di altre specie di montagna, il camoscio appenninico non mostra adattamenti compensativi e questo avrà conseguenze nefaste in particolare per le popolazioni più di bassa quota (...).* (esperto)  
 b. *la partecipazione degli studenti cala notevolmente durante la Didattica Digitale Integrata e ciò ha, ancora una volta, un'influenza negativa sul rendimento degli studenti e sulla buona uscita dall'Esame di Stato.* (studente)
- (5) *In conclusione, la Didattica Digitale Integrata ha alcuni aspetti positivi ma molti negativi, e per evitare tali complicazioni credo che la scelta migliore sia quella di non estendere la DDI a partire dal prossimo anno.* (studente)

Il minore utilizzo di strategie esplicite per segnalare le relazioni discorsive da parte di scrittori più esperti è un risultato emerso anche in un altro studio condotto da Bienati & Frey (in stampa) che indaga le variazioni d'uso dei connettivi in generale e di connettivi causali in particolare in testi di studenti di italiano L1 e L2 nel passaggio dalla scuola secondaria di primo a quella di secondo grado. I risultati dello studio mostrano una decrescita del numero di connettivi utilizzati tra i due cicli.<sup>10</sup> Inquadrandolo il risultato nella vasta letteratura a disposizione sull'inglese, che mostra come ci sia un'effettiva crescita del numero di connettivi usati almeno fino al 7° anno di istruzione (cfr. Andreev & Uccelli 2023) e poi una graduale decrescita nei gradi successivi fino all'università (cfr. Crossley et al. 2011), le autrici concludono che il minore utilizzo di connettivi in generale e di connettivi causali in particolare nei testi da loro analizzati:

could mean that – until a certain grade – students use connectives as scaffolds to properly manage and structure the longer and more complex writing tasks, while contemporarily signaling explicitly to the assumed audience (the teacher) that the main function of this text genre is indeed performed. Then, right at the end of lower secondary education, around 8th grade, an inversion of the tendency might happen. Around 12th grade, i.e. after four more years of training, students have learned how to manage long and complex texts and gained experience in the subtleties of tailoring the information expressed for an assumed audience other than the teacher, thus starting to remove these scaffolding and ordering devices, probably replacing them by other, more advanced structures, such as alternative lexicalizations. (Bienati & Frey, in stampa).

Dunque, se non è la maggiore quantità di coesivi a fare la differenza tra scrittori con più o meno formazione alle spalle, quali sono gli altri assi entro cui si vede una variazione?

Dallo studio di Bienati & Frey (in stampa) si constata come la ricercatezza lessicale<sup>11</sup> dei connettivi causali utilizzati aumenti dalla scuola secondaria di primo a quella di secondo grado. Gli studenti più grandi userebbero quindi, in media, connettivi causali meno frequenti rispetto ai loro compagni più giovani. Questo abbassamento della media dei punteggi di frequenza, tuttavia, non si verifica perché gli studenti smettono di

<sup>10</sup> Si ricorda che lo studio si configura come pseudolongitudinale: i testi raccolti alla secondaria di primo grado non sono stati scritti dagli stessi autori dei testi raccolti alla secondaria di secondo grado.

<sup>11</sup> Per rendere operativo il concetto di ricercatezza lessicale, nella letteratura anglofona *sophistication*, le autrici hanno innanzitutto determinato, per ogni connettivo presente nei testi degli studenti, la frequenza di quel connettivo in un corpus di riferimento (CORIS, Rossini Favretti et al. 2002). Successivamente hanno creato un punteggio medio di frequenza per ogni testo dividendo la somma di tutti i punteggi di frequenza per il numero di connettivi presenti nel testo.

utilizzare connettivi causali comuni, come ‘perché’, ma poiché, accanto ai connettivi causali frequenti, ormai pienamente acquisiti nel primo ciclo, essi affiancano connettivi molto meno frequenti, come ‘per via di’ o ‘siccome’.

Per avere un'idea più precisa dei connettivi che sono maggiormente associati a un certo gruppo di studenti, è stata impiegata la tecnica della *correspondence analysis* (Greenacre, 2010), che permette di ridurre la complessità del dataset di partenza a una rappresentazione bidimensionale della variazione interna dello stesso. Come mostrato nella figura 3, la variazione interna ai dati può essere spiegata in maniera soddisfacente e senza la perdita di troppe informazioni da sole due dimensioni, per un totale di 93.4% di variazione spiegata, rappresentate nel piano cartesiano dall'asse delle ascisse e delle ordinate. La dimensione rappresentata dall'asse delle ascisse spiega già una percentuale molto alta (76.4%) della variazione interna ai dati e separa le produzioni degli studenti di italiano L1 della scuola secondaria di secondo grado (etichetta *L1 upper*) dalle produzioni di studenti di italiano L2 dello stesso grado d'istruzione (etichetta *L2 upper*), così come dagli studenti della secondaria di primo grado (sia L1 sia L2, etichetta *L1 lower* e *L2 lower*). La seconda dimensione, sull'asse delle ordinate, spiega un ulteriore 17% di variazione e separa chiaramente le produzioni degli studenti della secondaria di primo grado da quelle della secondaria di secondo grado. La posizione dei connettivi è indicata dalle frecce che dipartono dall'origine: più i connettivi sono distanti dal centro del grafico, maggiore è il loro potere esplicativo in termini di differenziazione dei gruppi. Inoltre, più l'angolo che si forma tra la posizione del gruppo nel grafico e la posizione del connettivo è piccolo, più il gruppo e il connettivo risultano associati. Ad esempio, nonostante occorra molto frequentemente in tutti i gruppi all'interno del dataset, l'uso del connettivo ‘perché’ sembra essere particolarmente caratteristico degli studenti della scuola secondaria di primo grado; mentre i profili degli studenti di italiano L1 della scuola secondaria di secondo grado sono caratterizzati da una maggiore varietà di connettivi causali (piuttosto rari) che non sono utilizzati o lo sono solo marginalmente negli altri gruppi. Un limite importante di questa analisi risiede nel fatto che ogni gruppo di studenti ha scritto i testi oggetto dell'analisi a partire da task diversi, specificamente progettati per il livello di istruzione (primo/secondo ciclo) e la formazione linguistica (italiano L1/L2) di riferimento. Le differenze lessicali che si notano tra i gruppi, quindi, potrebbero anche essere dovute a specifici elementi lessicali presenti nei prompt, che potrebbero aver avuto un effetto di priming sulle scelte lessicali degli studenti dei diversi gruppi.

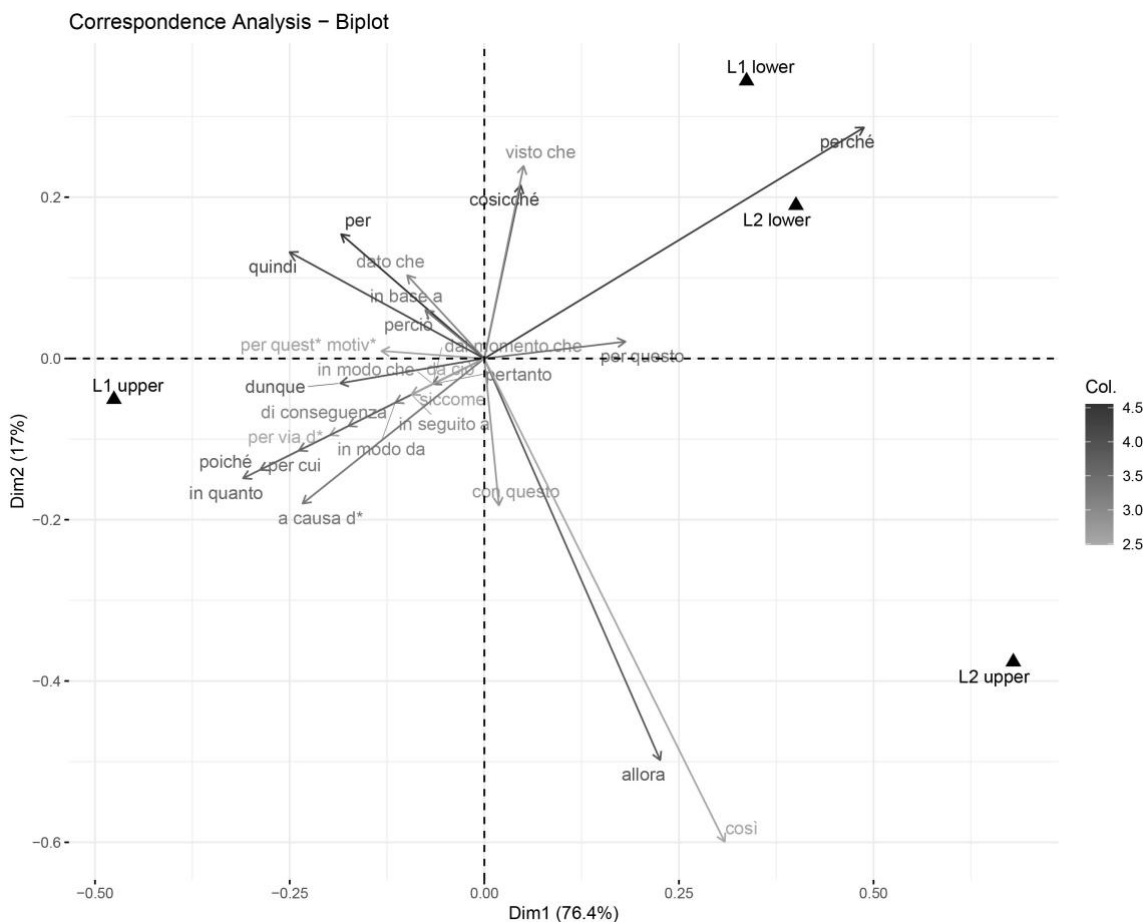


Figura 3. *Correspondence analysis* dei connettivi causali utilizzati nei diversi gruppi. I colori più scuri indicano i connettivi comuni, quelli più chiari quelli più infrequenti. (Tratto da Bienati & Frey, in stampa).

## 5. Conclusioni

All'interno della relazione tra le condizioni di coerenza, che definiscono un testo in quanto tale, e le procedure di coesione, che riguardano la sua struttura linguistica, le analisi qui presentate hanno esplorato i pattern di utilizzo di una specifica categoria di elementi coesivi: i connettivi. Il percorso formativo che gli studenti intraprendono a scuola, così come la formazione professionale successiva, sembrano avere un grande impatto sui modi in cui le procedure di coesione vengono realizzate, fermo restando che i testi analizzati obbediscono alle condizioni di coerenza che li rendono testi.

I risultati delle indagini svolte hanno permesso di comprendere che lo sviluppo delle procedure di coesione che riguardano i connettivi non è sempre lineare. Al contrario, esso segue traiettorie complesse che coinvolgono non solo la quantità, ma anche il tipo e la varietà dei coesivi. Emerge, infatti, che gli scriventi più esperti, siano essi studenti più maturi o professionisti della lingua, prediligono un equilibrio dinamico tra la segnalazione esplicita – che può assumere la forma di connettivo generico con ulteriori specificazioni oppure di connettivo funzionalmente più specializzato e meno frequente – che secondo studi psicolinguistici facilita il processo di comprensione, e la segnalazione

implicita delle relazioni discorsive, che lascia al lettore il compito di “trarre inferenze e costruire anelli mancanti” (cfr. Conte 1993) rendendo così la lettura sfidante.

Se le nostre ricerche hanno finora esplorato la variazione nel rapporto tra procedure di coesione e condizioni di coerenza lungo l’asse temporale, tenendo fisse queste ultime, una possibile direzione futura sarà quella di esaminare la variazione nelle procedure di coesione in testi con diversi gradi di coerenza all’interno dello stesso gruppo di studenti omogeneo per grado di istruzione. In tal senso, faremo affidamento sulle valutazioni prodotte per il corpus.

## Bibliografia

- Andorno, Cecilia. 2003. *Linguistica testuale. Un’introduzione*. Roma: Carocci.
- Andreev, Linda & Uccelli, Paola. 2023. The secret life of connectives: A taxonomy to study individual differences in mid-adolescents’ use of connectives in writing to persuade. *Reading and Writing* 37(3). 1–32.
- ANP. 2018. *Raccomandazione del Consiglio dell’Unione Europea relativa alle competenze chiave per l’apprendimento permanente. Approfondimenti e riflessioni*. ([https://www.anp.it/wp-content/uploads/2018/10/2018\\_10\\_30-nuove-competenze-chiave-UE\\_2018\\_riflessioni-ANP.pdf](https://www.anp.it/wp-content/uploads/2018/10/2018_10_30-nuove-competenze-chiave-UE_2018_riflessioni-ANP.pdf)) (Consultato il 24.01.2024.)
- Asquini, Giorgio. 2011. Dieci anni di PISA: Primi bilanci e nuove prospettive. *Giornale Italiano della Ricerca Educativa* IV(7). 71–83.
- Balboni, Paolo. 2017. *Perché insegnare l’italiano ai ragazzi italiani. E come*. Venezia: Marsilio.
- Barbagli, Alessia & Lucisano, Pietro & Dell’Orletta, Felice & Montemagni, Simonetta & Venturi, Giulia. 2016. CltA: An Italian learner corpus to study the development of writing competence. In Calzolari, Nicoletta & Choukri, Khalid & Declerck, Thierry & Goggi, Sara & Grobelnik, Marko & Maegaard, Bente & Mariani, Joseph & Mazo, Helene & Moreno, Asuncion & Odijk, Jan & Piperidis, Stelios (a cura di), *Proceedings of 10th International Conference on Language Resources and Evaluation (LREC 2016)*, 88–95. Portorose: European Language Resources Association (ELRA).
- Beaugrande, Robert-Alain de & Dressler, Wolfgang Ulrich. 1984. *Introduzione alla linguistica testuale*. Bologna: Il Mulino.
- Bienati, Arianna & Frey, Jennifer C. In stampa. Development of causal connectives in Italian L1 and L2 student writing: A comparison of argumentative texts from lower and upper secondary school. In Ackerley, Katherine & Castello, Erik (a cura di), *Continuing learner corpus research: Challenges and opportunities*. Louvain: Presses universitaires de Louvain.
- Bienati, Arianna & Frey, Jennifer-Carmen & Palmero Aprosio, Alessio & Facchinelli, Noemi. 2023. Applicazione delle risorse disponibili per l’italiano all’annotazione automatica delle relazioni discorsive in testi scolastici: alcune implicazioni teoriche. Presentazione al *LVI congresso internazionale della Società di Linguistica Italiana – WS3: Linguistica teorica e trattamento automatico delle lingue: verso nuove sinergie*, 15 settembre 2023, Torino.
- Bienati, Arianna, & Frey, Jennifer-Carmen. 2023. Defining a connective by the company it keeps: A corpus-based study on students’ (creative) ways of expressing causality using polyfunctional ‘e’. Presentazione al *45 Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft*, 7 marzo 2023, Colonia.

- Bienati, Arianna & Vettori, Chiara & Zanasi, Lorenzo. 2022. In viaggio verso ITACA: La coerenza testuale come meta della scrittura scolastica. Proposta di una griglia di valutazione. *Italiano a scuola* 4(1). 55–70.
- Borghini, Carlotta Caterina. 2013. *Analisi di produzioni scritte. Valutazioni e misure automatizzate di elaborati scolastici*. Roma: Università di Roma La Sapienza. (Tesi di dottorato.)
- Boscolo, Pietro & Zuin, Elvira. 2015. *Come scrivono gli adolescenti. Un'indagine sulla scrittura scolastica e sulla didattica della scrittura*. Bologna: Il Mulino.
- Brown, Gillian & Yule, George. 1983. *Discourse analysis*. Cambridge: Cambridge University Press.
- Calamai, Silvia. 2012. Per un'analisi quantitativa delle competenze negli studenti universitari. In Bernini, Giuliano & Lavinio, Cristina & Valentini, Ada & Voghera, Miriam (a cura di), *Competenze e formazione linguistiche. In memoria di Monica Berretta. Atti dell'11. Congresso dell'Associazione italiana di linguistica applicata*, 77–99. Perugia: Guerra Edizioni.
- Charolles, Michel. 1983. Coherence as a principle in the interpretation of discourse. *Text – Interdisciplinary Journal for the Study of Discourse* 3(1). 71–98.
- Cisotto, Lerida & Novello, Nazzarena. 2012. La scrittura di sintesi di studenti del primo anno di Scienze della Formazione Primaria. *Italian Journal of Educational Research* 8. 41–57.
- Colombo, Adriano (a cura di). 1992. *I pro e i contro. Teoria e didattica dei testi argomentativi*. Scandicci: La Nuova Italia.
- Conte, Maria Elisabeth. 1988. *Condizioni di coerenza. Ricerche di linguistica testuale*. Scandicci: La Nuova Italia.
- Conte, Maria Elisabeth. 1993. Linguistica testuale. *Enciclopedia dell'italiano*. ([https://www.treccani.it/enciclopedia/linguistica-testuale\\_\(Enciclopedia-Italiana\)/](https://www.treccani.it/enciclopedia/linguistica-testuale_(Enciclopedia-Italiana)/)) (Consultato il 11.02.2024.)
- Corno, Dario. 2002. *Scrivere e comunicare. Teoria e pratica della scrittura in lingua italiana*. Milano: Mondadori.
- Crossley, Scott A. & McNamara, Danielle S. 2011. Text coherence and judgments of essay quality: models of quality and coherence. In Carlson, Laura & Hoelscher, Christoph & Shipley, Thomas F. (a cura di), *Proceedings of the 33th Annual Meeting of the Cognitive Science Society*, 1236–1241. Boston, MA: Cognitive Science Society.
- Crossley, Scott A. & Weston, Jennifer L. & McLain Sullivan, Susan T. & McNamara, Danielle S. 2011. The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication* 28(3). 282–311.
- Crossley, Scott A. & Kyle, Kristopher & McNamara, Danielle S. 2016. The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing* 32. 1–16.
- Daneš, Frantisek. 1974. Functional sentence perspective and the organization of the text. In Daneš, Frantisek (a cura di), *Papers on functional sentence perspective*, 106–128. Berlin, Boston: De Gruyter Mouton.
- Das, Depobam & Taboada, Maite. 2014. *RST signalling corpus annotation manual*. Burnaby, BC: Simon Fraser University.
- De Mauro online = Il nuovo De Mauro (<https://dizionario.internazionale.it/>) (Consultato il 29.02.2024.)
- De Mauro, Tullio. 2010. Misurare e valutare le abilità linguistiche. In Lugarini, Edoardo (a cura di), *Valutare le competenze linguistiche*, 17–22. Milano: Franco Angeli.

- De Santis, Cristiana & Gatta, Francesca. 2013. Notizie dalla scuola. Le competenze grammaticali e testuali degli studenti madrelingua all'uscita dalla scuola secondaria. I risultati di un'indagine. *Studi di Grammatica Italiana* XXXI-XXXII. 411–441.
- DISC online = *Dizionario Italiano Sabatini Coletti*. ([https://dizionari.corriere.it/dizionario\\_italiano/Y](https://dizionari.corriere.it/dizionario_italiano/Y)) (Consultato il 28.02.2024.)
- Feltracco, Anna & Jezek, Elisabetta & Magnini, Bernardo & Stede, Manfred. 2016. LICO: A lexicon of Italian connectives. In Corazza, Anna & Montemagni, Simonetta & Semeraro, Giovanni (a cura di), *Proceedings of the Third Italian Conference on Computational Linguistics (CLiC-It 2016)*, 141–145. Napoli: Accademia University Press.
- Ferrari, Angela. 2010a. Connettivi. *Enciclopedia dell'italiano*. ([https://www.treccani.it/enciclopedia/connettivi\\_\(Enciclopedia-dell'Italiano\)](https://www.treccani.it/enciclopedia/connettivi_(Enciclopedia-dell'Italiano))) (Consultato il 27.02.2024.)
- Ferrari, Angela. 2010b. Coesione, procedure di. *Enciclopedia dell'italiano*. ([https://www.treccani.it/enciclopedia/procedure-di-coesione\\_\(Enciclopedia-dell'Italiano\)/](https://www.treccani.it/enciclopedia/procedure-di-coesione_(Enciclopedia-dell'Italiano)/)) (Consultato il 11.02.24.)
- Ferrari, Angela. 2014. *Linguistica del testo. Principi, fenomeni, strutture*. Roma: Carocci.
- Ferrari, Angela. 2021. Segnali discorsivi e connettivi. *Lingua e Stile* LVI(1). 143–150.
- Ferrari, Angela & Lala, Letizia & Zampese, Luciano. 2021. *Le strutture del testo scritto. Teoria e esercizi*. Roma: Carocci.
- Glaznieks, Aivars & Frey, Jennifer-Carmen & Stopfner, Maria & Zanasi, Lorenzo & Nicolas, Lionel. 2022. Leonide: A longitudinal trilingual corpus of young learners of Italian, German and English. *International Journal of Learner Corpus Research* 8(1). 97–120.
- Grassi, Roberta & Nuzzo, Elena. 2012. Analizzare le (in)competenze di scrittura all'università: Evidenze dai test di valutazione iniziale. In Bernini, Giuliano (a cura di), *Competenze e formazione linguistiche. In memoria di Monica Berretta. Atti dell'11° Congresso dell'Associazione Italiana di Linguistica Applicata*, 101–118. Perugia: Guerra Edizioni.
- Greenacre, Michael. J. 2010. Correspondence analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(5). 613–619.
- Gruppo di Firenze. 2017. *Contro il declino dell'italiano a scuola*. (<https://gruppodifirenze.blogspot.com/2017/02/contro-il-declino-dellitaliano-scuola.html>) (Consultato il 28.02.2024.)
- Halliday, Michael & Hasan, Ruqaiya. 1976. *Cohesion in English*. Londra: Longman.
- Hodges, Tracey S. & Feng, Luxi & Kuo, Li Jen & McTigue, Erin. 2016. Discovering the literacy gap: A systematic review of reading and writing theories in research. *Cogent Education* 3. 1–13.
- Hoek, Jet & Zufferey, Sandrine & Evers-Vermeul, Jacqueline & Sanders, Ted J. M. 2019. The linguistic marking of coherence relations: Interactions between connectives and segment-internal elements. *Pragmatics & Cognition* 25(2). 276–309.
- Hörmann, Hans. 1976. *Psicolinguistica*. Bologna: il Mulino.
- INVALSI. 2012. Esame di stato conclusivo dei percorsi di istruzione secondaria superiore. Rilevazione degli errori più diffusi nella padronanza della lingua italiana nella prima prova di italiano. ([https://www.invalsi.it/download/rapporti/es2\\_0312/RAPPORTO\\_ITALIANO\\_prov\\_e\\_2010.pdf](https://www.invalsi.it/download/rapporti/es2_0312/RAPPORTO_ITALIANO_prov_e_2010.pdf)) (Consultato il 24.01.2024.)

- Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.
- Kintsch, Walter, & van Dijk, Teun A. 1978. Toward a model of text comprehension and production. *Psychological Review* 85(5). 363–394.
- Kleijn, Suzanne & Pander Maat, Henk L. W. & Sanders, Ted J. M. 2019. Comprehension effects of connectives across texts, readers, and coherence relations. *Discourse Processes* 56. 447–464.
- Klie, Jan-Christoph & Bugert, Michael & Boullosa, Beto & Eckart de Castilho, Richard & Gurevych, Iryna. 2018. The INCEption platform: Machine-assisted and knowledge-oriented interactive annotation. In Zhao, Dongyan (a cura di), *Proceedings of System Demonstrations of the 27th International Conference on Computational Linguistics (COLING 2018)*, 5–9.
- Kuiken, Folkert & Vedder, Ineke. 2017. Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing* 34(3). 321–336.
- Lavinio, Cristina & Sobrero, Alberto (a cura di). 1991. *La lingua degli studenti universitari. Quaderni del Giscel* 7. Firenze: La Nuova Italia.
- Light, Richard. 2001. *Making the most of college*. Cambridge, MA: Harvard University press.
- Lin, Ziheng & Kann, Min-Yen & Ng, Hwee Tou. 2009. Recognizing implicit discourse relations in the Penn Discourse Treebank. In Koehn, Philipp & Mihalcea, Rada (a cura di), *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*. 343–351.
- Lo Duca, Maria G. 2017. Lettera di M. G. Lo Duca in risposta alla “proposta dei Seicento” sul declino della lingua italiana. ([http://www.societadilinguisticaitaliana.net/wp-content/uploads/2017/04/Testo-Lo-Duca-lungo-e-firme\\_28aprile.pdf](http://www.societadilinguisticaitaliana.net/wp-content/uploads/2017/04/Testo-Lo-Duca-lungo-e-firme_28aprile.pdf)) (Consultato il 28.02.2024.)
- McNamara, Danielle S. & Crossley, Scott A. & McCarthy Philip M. 2010. Linguistic features of writing quality. *Written Communication* 27(1). 57–86.
- Palermo, Massimo. 2013. *Linguistica testuale dell’italiano*. Bologna: il Mulino.
- Pallotti, Gabriele. 1998. I ragazzi di oggi non sanno pensare. Alcune riflessioni di antropologia della scrittura. *Riforma e Didattica* II(2). 1–38.
- Prasad, Rashmi & Dinesh, Nikhil & Lee, Alan & Miltsakaki, Eleni & Robaldo, Livio & Joshi, Aravind & Webber, Bonnie. 2008. The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference of Language Resources and Evaluation (LREC 2008)*, 2961–2968.
- Reinhard, Tanya. 1980. Conditions for text coherence. *Poetics Today* 1(4). 161–180.
- Rossini Favretti, Rema & Tamburini, Fabio & De Santis, Cristiana. 2002. CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model. In Wilson, Andrew & Rayson, Paul & McEnery, Tony (a cura di), *A rainbow of corpora: Corpus linguistics and the languages of the world*, 27–38. Monaco: Lincom-Europa.
- Ruele, Michele & Zuin, Elvira. 2020. *Come cambia la scrittura a scuola*. Trento: IPRASE.
- Ruggiano, Fabio. 2011. *L’italiano scritto a scuola*. Roma: Aracne.
- Schutte, Nicola S. & Malouff, John M. 2007. Dimensions of reading motivation: Development of an adult reading motivation scale. *Reading Psychology: An International Quarterly* 28. 469–489.

- Shi, Wei & Yung, Frances & Rubino, Raphael & Demberg, Vera. 2017. Using explicit discourse connectives in translation for implicit discourse relation classification. In Kondrak, Greg & Watanabe, Taro (a cura di), *Proceedings of the Eighth International Joint Conference on Natural Language Processing*, 484–495.
- Sprugnoli, Rachele & Tonelli, Sara & Palmero Aprosio, Alessio & Moretti, Giovanni. 2018. Analysing the evolution of students' writing skills and the impact of neo-standard Italian with the help of computational linguistics. In Cabrio, Elena & Mazzei, Alessandro & Tamburini, Fabio (a cura di), *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*, 354–359.
- Tarallo, Claudia. 2019. Distribuzione pentenaria e analisi della competenza testuale su gruppi differenziali di studenti. In Palermo, Massimo & Salvatore, Eugenio (a cura di), *Scrivere nella scuola oggi. Obiettivi, metodi, esperienze*, 81–91. Firenze: Cesati.
- Tonelli, Sara & Riccardi, Giuseppe & Prasad, Rashmi & Joshi, Aravind. 2010. Annotation of discourse relations for conversational spoken dialogs. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, 2084–2090.
- Treccani online = Il vocabolario Treccani (<https://www.treccani.it/vocabolario/>) (Consultato il 28.02.2024.)
- Troia, Gary A. & Harbaugh, Allen G. & Shankland, Rebecca K. & Wolbers, Kimberly A. W & Lawrence, Ann M. 2012. Relationships between writing motivation, writing activity, and writing performance: effects of grade, sex, and ability. *Reading and Writing* 26(1). 7–44.
- Vettori, Chiara. 2021. La scuola come unico contesto di apprendimento della seconda lingua? Il caso dell'Alto Adige. In Favilla, Maria Elena & Machetti, Sabrina (a cura di), *Lingue in contatto e linguistica applicata: individui e società*, 239–254. Milano: AItLA.
- Vettori, Chiara & Abel, Andrea. 2017. *KOLIPSI II: gli studenti altoatesini e la seconda lingua; indagine linguistica e psicosociale*. Bolzano: Eurac Research.
- Webber, Bonnie. 2004. D-LTAG: Extending lexicalized TAG to discourse. *Cognitive Science* 28(5). 751–779.
- Webber, Bonnie & Prasad, Rashmi & Lee, Alan & Joshi, Aravind. 2019. *The Penn Discourse Treebank 3.0 Annotation Manual*.
- Xue, Nianwen & Ng, Hewe Tou & Pradhan, Sameer & Rutherford, Attapol & Webber, Bonnie & Wang, Chuan & Wang, Hongmin. 2016. CoNLL 2016 shared task on multilingual shallow discourse parsing. In Xue, Nianwen (a cura di), *Proceedings of the CoNLL-16 shared task*, 1–19.
- Zanasi, Lorenzo (a cura di) 2021. Plurilinguismo: nuovi scenari teorici e didattici. *Lingua e Nuova Didattica – LEND 3* (numero monografico).
- Zanasi, Lorenzo & Bienati, Arianna & Vettori, Chiara. 2024. Sulla definizione di coerenza testuale. Dizionari e ricerca a confronto verso un modello operativo per le classi del ciclo superiore. *Ladinia XLVIII*.