



Il corpus KIParla. Una risorsa per la didattica dell'oralità

Silvia Ballarè

Alma Mater Studiorum – Università di Bologna, Italia

Claudia Borghetti

Alma Mater Studiorum – Università di Bologna, Italia

Paolo Della Putta

Università degli Studi di Torino, Italia

Caterina Mauri

Alma Mater Studiorum – Università di Bologna, Italia

Eleonora Zucchini

Alma Mater Studiorum – Università di Bologna, Italia

Abstract This paper presents KIParla, a *corpus* of spoken Italian, and highlights its potential for language teaching professionals. KIParla aims to represent the variation of spoken Italian and consists of modules featuring conversations in different settings (formal and informal, at university and at home) among speakers with diverse sociodemographic backgrounds (origins, age, education, profession, and native language). As will be shown through activities designed for learners with different profiles, this resource can support teachers in addressing language variation and promoting awareness of the oral dimension of Italian.

Keywords Corpora. Spoken language. Language variation. Data-driven learning. Language education.

Sommario 1 Corpora e didattica. – 2 L'oralità e la didattica dell'italiano. – 3 Il *corpus* KIParla. – 4 KIParla in classe. – 5 Riflessioni conclusive.



Peer review

Submitted 2025-03-17
Accepted 2025-07-25
Published 2025-08-29

Open access

© 2025 Ballarè, Borghetti, Della Putta, Mauri, Zucchini | 4.0



Citation Ballarè, S.; Borghetti, C.; Della Putta, P.; Mauri, C.; Zucchini, E. (2025). "Il corpus KIParla. Una risorsa per la didattica dell'oralità". *EL.LE*, 14(2), 213-230.

1 Corpora e didattica

Un *corpus* è definibile come «a collection of naturally-occurring language text, chosen to characterize a state or variety of language» (Sinclair 1991, 171; per approfondire vedi, ad esempio, Lenci et al. 2005); i testi, tipicamente, sono digitalizzati e consultabili. L'uso dei *corpora* nella didattica delle lingue non è molto diffuso, sebbene esista un'evidenza sperimentale a supporto della loro importanza sia per la didattica in sé sia per la formazione dei docenti (Corino 2019; Chen et al. 2019).¹ In particolare, vari autori (cf. discussione in Gablasova, Bottini 2022) suggeriscono l'importanza per la glottodidattica dei *corpora* di lingua parlata, che permettono ad apprendenti e a insegnanti un confronto con l'uso reale della lingua, dunque con la sua autenticità, la sua variazione e la sua dinamica discorsivo/interazionale.

Tradizionalmente, l'applicazione dei *corpora* alla pedagogia linguistica viene divisa in due macrocategorie, individuate sulla base dei fini e dei destinatari d'uso (McEnery, Xiao 2011): da una parte, le applicazioni 'dirette', che vedono i *corpora* come strumenti utili all'attività di classe o allo studio linguistico individuale; dall'altra, le applicazioni 'indirette', nelle quali è il docente a fare uso dei *corpora* per la preparazione dei materiali didattici, per la risoluzione di dubbi linguistici o per la propria formazione professionale.

Secondo molti studiosi (ad esempio, Boulton, Cobb 2017; Corino, Onesti 2019; Lusta et al. 2023), l'uso diretto dei *corpora* è vantaggioso per l'acquisizione, perché permette di affrontare efficacemente alcuni aspetti linguistici 'non canonici', come l'analisi delle collocazioni e della variazione linguistica o, più in generale, di far prendere coscienza agli studenti della differenza fra norma e uso. Vi sono, inoltre, vantaggi emotivi e cognitivi nell'introduzione dei *corpora* nella pratica didattica: gli studenti sono spinti a cercare informazioni linguistiche, a generalizzarle e a costruire attivamente il proprio sapere, attivando così meccanismi di apprendimento proattivi e inferenziali, come l'induzione, che possono facilitare la ritenzione mnemonica, sostenere la motivazione e l'impiego di strategie autonome di apprendimento.

Si parla di applicazioni indirette quando i *corpora* sono impiegati nella formazione professionale dei docenti di lingua, che possono, a loro volta, essere istruiti a usare i *corpora* in classe, a farvi riferimento per risolvere i propri dubbi o per costruire materiali didattici più

1 Questa ricerca è stata condotta all'interno del Progetto *DiverSIta - Diversity in Spoken Italian*, finanziato dall'Unione Europea - NextGenerationEU a valere sul Piano Nazionale di Ripresa e Resilienza (PNRR) - Missione 4 Istruzione e ricerca. Codice progetto MUR [P2022RFR8T] - CUP [J53D23017320001].

aderenti all'effettivo uso della lingua che insegnano. Anche in questo caso, gli studi (Zareva 2017; Chen et al. 2019) mostrano che aiutare gli insegnanti a comprendere più a fondo le potenzialità dei *corpora* per l'insegnamento, specialmente grazie a un'esperienza diretta del loro utilizzo, offre diversi vantaggi. Aiuta ad esempio a ridurre pregiudizi negativi verso alcuni aspetti della lingua considerati aprioristicamente errati, a sviluppare un atteggiamento più proattivo verso l'insegnamento (senza limitarsi, ad esempio, all'uso di materiali già pronti), a valutare le produzioni degli studenti basandosi su dati linguistici reali e, infine, a sperimentare direttamente l'uso dei *corpora*, come spiegato in precedenza.

Nonostante queste ricadute positive, la ricerca sulle pratiche d'uso dei *corpora* rivela uno scenario poco in linea con le scoperte scientifiche: l'uso diretto dei *corpora* è molto poco praticato, per mancanza di conoscenza dell'esistenza stessa dei *corpora*, per la sensazione che il loro uso richieda molto tempo e per una più generale riluttanza verso l'applicazione della tecnologia nella didattica (Boulton, Vyatkina 2021; Lenko-Szymanska 2014). Anche la formazione o l'aggiornamento dei docenti attraverso e/o verso l'uso dei *corpora* è pratica molto rara, sovente in conflitto con un atteggiamento sospettoso degli insegnanti stessi verso l'uso della tecnologia nella formazione e nella didattica.

Tuttavia, come accennato, prove che dimostrano i buoni esiti dell'implementazione dell'uso dei *corpora* in glottodidattica esistono, e riguardano sia il loro uso diretto sia quello indiretto: gli studi di Lin e Lee (2015) e Ma et al. (2022), per esempio, hanno rilevato attitudini positive verso l'uso dei *corpora* da parte di docenti con più o meno esperienza, al netto di difficoltà operative di vario ordine, fra cui la difficoltà di accesso all'interfaccia informatica e l'assenza, in aula, di adeguate strumentazioni; numerosi studi sull'applicazione dei *corpora* all'insegnamento linguistico evidenziano effetti particolarmente positivi nell'apprendimento di aspetti lessicali e grammaticali. L'esposizione a dati autentici e la loro analisi consente agli apprendenti di osservare in modo diretto pattern d'uso ricorrenti, favorendo una maggiore consapevolezza nell'impiego delle collocazioni e delle combinazioni lessicali frequenti. Parallelamente, l'analisi di esempi reali supporta l'interiorizzazione di strutture grammaticali attraverso l'osservazione di regolarità morfologiche e sintattiche difficilmente accessibili nei materiali didattici tradizionali. In particolare, l'uso dei *corpora* permette di sviluppare competenze linguistiche più aderenti all'uso effettivo della lingua, superando una visione normativa e decontestualizzata delle regole grammaticali. (cf., per una revisione, Boulton 2021).

In questo solco di ottimismo, e seguendo anche le indicazioni di Lenko-Szymanska (2022) su come offrire alla comunità docente *corpora* sempre più in linea con le loro possibili richieste, in questo

contributo presentiamo il *corpus* modulare KIParla (Mauri et al. 2019), che mira a rappresentare la variazione interna all'italiano parlato. Il *corpus* è composto da moduli diversificati per tipo di situazione comunicativa (formale e informale, in contesto universitario e familiare, interviste semi-strutturate e conversazioni a tavola) e caratteristiche sociodemografiche dei parlanti (provenienza, età, titolo di studio, professione, lingue materne); inoltre, saranno pubblicati a breve anche due moduli (StraParla-BO e StraParla-TO) che si concentrano sul parlato spontaneo di persone con background migratorio, siano esse di prima o seconda generazione, un tipo di dato nuovo nel panorama dei *corpora* di apprendenti di italiano. Come vedremo, le caratteristiche di KIParla lo rendono particolarmente adeguato ad applicazioni pedagogiche nella scuola italiana d'oggi che, sempre più multiculturale e multilingue, è spesso prima e principale testimone dello sfumare dei confini fra chi è parlante nativo e chi non nativo dell'italiano come lingua di scolarizzazione.

Nelle pagine che seguono, dopo aver inquadrato il ruolo dell'oralità nella didattica dell'italiano (§ 2), presentiamo il *corpus* nelle sue caratteristiche principali (§ 3). Nella sezione successiva (§ 4), ne esemplifichiamo l'utilizzo, commentando tre esempi di altrettante attività didattiche incentrate sulla scoperta attiva della variazione della lingua italiana parlata, prima di giungere a qualche riflessione conclusiva (§ 5).

2 L'oralità e la didattica dell'italiano

La letteratura sull'educazione linguistica rivela in maniera frequente che la dimensione orale riveste un ruolo marginale nei libri di testo e nelle pratiche didattiche, e che il parlato viene descritto dai manuali come secondario rispetto allo scritto.² L'inclusione dell'oralità nella pedagogia linguistica è, tuttavia, descritta come auspicabile sin dalle *Dieci tesi per l'educazione linguistica democratica* (GISCEL 1975) ed è inoltre prevista dalle Indicazioni nazionali in adozione fino a tempi recentissimi (2012).

Se si osservano i libri di testo per la scuola secondaria, emerge chiaramente che la lingua presa a modello per la riflessione grammaticale è la varietà scritta e, in particolare, lo scritto sorvegliato (Calaresu 2016; Dal Negro 2016); inoltre, l'osservazione della dimensione diamesica e l'esercizio delle competenze orali, se contemplati, non varcano i confini dei fascicoli integrativi. È solo qui, infatti, che vengono rilevate timide aperture all'oralità, tramite la presenza di testi dialogici con tratti del parlato (Ferrerri 2024;

² Voghera 2017; Bachis 2019; Pani 2019; Voghera, Sammarco 2021; Dimalta et al. 2024.

Pani et al. 2024); si registra dunque ancora, come già osservato da Sobrero (1997), la mancanza di dialogo fra i volumetti aggiuntivi e il cuore della grammatica, cioè l'analisi logica e grammaticale. Si fatica a introdurre la dimensione orale nelle pratiche didattiche della lingua materna, anche perché si stenta a riconoscere la necessità di far riflettere gli alunni sul parlato, specialmente se informale, credendoli già competenti e considerando certi fenomeni del parlato (si pensi a 'a me mi') come marginali o errati (Voghera 2017; Voghera, Sammarco 2021).

Non includere la lingua parlata nell'educazione linguistica, tuttavia, significa escludere dall'educazione formale una parte sostanziale dell'identità degli alunni, non rispettare l'ordine naturale di acquisizione, che vede lo sviluppo delle abilità orali precedere quello delle capacità di lettura e scrittura (Voghera 2023) e, più in generale, ridurre le opportunità di apprendimento linguistico (Voghera 2017; Pani 2019). Inoltre, sensibilizzare gli alunni alla dimensione orale significa metterli nelle condizioni di distinguerla dalla modalità scritta e, quindi, di acquisire con più facilità anche le abilità di scrittura (Voghera 2017). Infine, è da notare che varie esperienze didattiche dedicate alla riflessione sull'oralità hanno avuto esiti positivi, perché inclusive, pluriprospectiche ed efficaci nel sensibilizzare gli alunni all'osservazione del dato linguistico (Dal Negro 2016; Pani 2019; Voghera et al. 2020).

Per quanto riguarda la didattica dell'italiano come L2 o LS, lo scenario è certamente diverso; è sufficiente considerare il *Quadro comune di riferimento per la conoscenza delle lingue* (Consiglio d'Europa 2002) per rendersi conto che la didattica delle lingue straniere dedica uno spazio apposito alle abilità orali. Tuttavia, molti studi sono concordi nel sostenere che, nei materiali didattici, la lingua parlata viene rappresentata spesso in maniera poco realistica e senza fare ricorso effettivo a materiali autentici. Le fonti di input orali sono povere dal punto di vista pragmatico (Nuzzo 2016) e gli esempi di dialogo, sia stampati sia contenuti negli inserti multimediali, non presentano tratti tipici del parlato spontaneo, ad esempio i fenomeni di negoziazione, e, pertanto, non offrono spunti per riflettere sulle caratteristiche tipiche dell'interazione orale (Gillani Pernas 2013; Pauletto 2021). Infine, la varietà di lingua esemplificata è spesso sociolinguisticamente uniforme (De Roberto, Venco 2024).

Prima di concludere, occorre ammettere che l'uso di materiali orali pone numerose difficoltà pratiche aggiuntive rispetto all'impiego di testi scritti, i quali sono di gran lunga più accessibili e gestibili di quelli orali. Per quanto riguarda sia l'educazione linguistica che la didattica dell'italiano L2/LS, la difficoltà di reperimento e di gestione di fonti audio non facilita l'integrazione della dimensione orale nelle pratiche didattiche. In secondo luogo, i testi autentici, vale a dire prodotti a prescindere dalla finalità didattiche, sia

orali che scritti, risultano spesso complessi da adattare alle esigenze dell'apprendimento e dell'insegnamento, anche a causa dell'approccio, assai diffuso nella scuola, di partire dalle forme e non dai testi, quando si fa riflessione metalinguistica; inoltre, per quanto riguarda riflessione grammaticale a scuola, ad esempio, i testi confezionati *ad hoc* permettono facilmente di mettere in rilievo i fenomeni oggetto di insegnamento (Lo Duca 1997; 2003). In questo senso, l'impiego di *corpora* può sia rendere l'educazione all'oralità di agile gestione che costituire una mediazione tra l'orientamento alle forme e quello ai testi, grazie alle varie modalità di ricerca messe a disposizione dalle interfacce.

3 Il corpus KIParla

Il *corpus* KIParla è una risorsa per lo studio dell'italiano parlato (Mauri et al. 2019; Ballarè et al. 2022). Il *corpus* è ad oggi costituito da 2.326.171 tokens ed è liberamente consultabile attraverso la piattaforma NoSketch Engine.³

Al momento, i dati linguistici non sono ancora stati né lemmatizzati né annotati, ma è possibile ricercare concordanze e ottenere i risultati con la tradizionale visualizzazione keyword in context. I dati sono accompagnati da un vasto corredo di metadati relativi ai partecipanti e alla conversazione, a cui si può accedere tramite la seguente procedura: nell'esito della ricerca, cliccando sul codice della conversazione, è possibile avere informazioni riguardo al tipo di contesto interazionale in cui è stato prodotto l'elemento linguistico ricercato, la fascia d'età e la regione di provenienza del parlante, e così via. I medesimi metadati, inoltre, sono impiegabili a monte come filtri di ricerca (selezionando 'text types'). Dalla schermata dell'esito della ricerca, poi, è possibile avere accesso anche alla trascrizione dell'intera conversazione – sia in versione ortografica, sia con alcuni simboli impiegati in analisi della conversazione (una versione semplificata di quanto proposto in Jefferson 2004) – e, previa registrazione, all'ascolto della traccia audio della conversazione stessa.

Caratteristica peculiare del *corpus* KIParla è la sua struttura modulare. La risorsa è costituita infatti da una serie di moduli (ad oggi ne sono stati pubblicati quattro) che possono essere consultati singolarmente oppure in modalità congiunta. Il numero dei moduli è destinato ad aumentare e, auspicabilmente, il *corpus* permetterà di monitorare la variazione dell'italiano parlato nel corso del tempo.

³ La descrizione della risorsa e l'accesso all'interfaccia di ricerca sono accessibili sul sito del *corpus* (www.kiparla.it).

Nella Tabella 1, si riportano le principali caratteristiche dei quattro moduli pubblicati.

Tabella 1 Caratteristiche dei moduli del *corpus* KIParla

Nome modulo	Anno di pubblicazione	Nr. tokens	Luogo di raccolta	Tipologia di conversazione	Parlanti
KIP	2019	668.581	Bologna e Torino	conversazioni spontanee, interviste semi-strutturate, esami, ricevimenti e lezioni (registrate in ambiente universitario)	Colti, variazione in base all'età e alla regione di nascita
ParlaTO	2019	561.388	Torino	interviste semi-strutturate	Diversa caratterizzazione sociale
KIPasti	2024	482.887	13 regioni italiane	conversazioni a tavola	Diversa caratterizzazione sociale
ParlaBO	2024	701.354	Bologna	Interviste semi-strutturate	Diversa caratterizzazione sociale

Attualmente, presso le università di Bologna e di Torino, sono in fase di allestimento due nuovi moduli, ovvero lo StraParla-BO e lo StraParla-TO in cui sono raccolte interazioni che coinvolgono parlanti con background migratorio e repertori plurilingui, residenti nelle città di Bologna e Torino; i due nuovi moduli, che includono sia interviste semi-strutturate che conversazioni libere registrate in diversi contesti, coinvolgono parlanti che hanno appreso l'italiano come L2 in età adulta, nati all'estero ma scolarizzati in Italia o nati qui da famiglie plurilingui.

Inoltre, presto saranno resi accessibili moduli realizzati in collaborazione con l'università di Napoli Federico II e la Libera università di Bolzano.

4 KIParla in classe

A partire dalle considerazioni precedenti (§ 1 e § 2) e facendo leva sulle caratteristiche della risorsa (§ 3), nelle pagine successive, si introducono e commentano tre attività didattiche condotte su altrettanti testi estratti appositamente dal modulo ParlaBO del

corpus KIParla. Si tratta di applicazioni indirette del *corpus*, ossia simulazioni di come insegnanti e formatori potrebbero impiegare i campioni di lingua messi a disposizione dal *corpus* per pianificare i propri interventi. Nel primo caso (§ 4.1), l'ambito è quello dell'educazione linguistica nella scuola secondaria di primo grado; il secondo esempio (§ 4.2) è invece pensato per i contesti in cui l'italiano costituisce una lingua aggiuntiva, a scuola come altrove; infine, il terzo caso (§ 4.3) volutamente non distingue tra L1 e L2 quando illustra una modalità di impiego dei dati del KIParla in situazioni didattiche in cui è necessario rispondere ai bisogni linguistici di tutti, esperti e meno esperti di italiano.

4.1 Un esempio per l'educazione linguistica

In questa sezione, proponiamo un'attività che, elaborata a partire da un estratto del modulo ParlaBO, ha come tema centrale il testo narrativo. La proposta pedagogica mira a trattare con gli alunni le caratteristiche tipiche di questo genere testuale e a renderli capaci di gestirle nella scrittura (competenze previste dalle Indicazioni Nazionali, DM n. 254 del 16/11/2012). In secondo luogo, si pone l'obiettivo di sensibilizzare gli apprendenti ai tratti caratteristici del parlato informale e spontaneo, in funzione anche di una successiva attività di riscrittura.

Nell'esempio (1) si riporta per motivi di spazio, solo una parte dell'intervista PBC007, che può però essere consultata in maniera integrale sul sito del KIParla.⁴

⁴ Gli estratti proposti nell'articolo sono riproduzioni fedeli degli originali reperibili nel *corpus*, di cui mantengono anche le convenzioni di trascrizione dell'oralità. Ci si è limitati a sostituire con degli pseudonimi i codici identificativi dei parlanti impiegati in KIParla, per aumentare la leggibilità del testo. Quanto alle convenzioni di trascrizione, il lettore troverà in particolare le seguenti: '(.)' segnala una pausa breve; la tilde e i due punti indicano rispettivamente la troncatura e l'allungamento di un suono; le parentesi uncinate sono usate per trascrivere una porzione di enunciato pronunciata a un ritmo più lento ('<testo>') o veloce ('>testo<'); le doppie parentesi riportano annotazioni relative ai comportamenti non verbali, come ((ride)); le parentesi singole indicano un'ipotesi di trascrizione avanzata quando l'audio riportava problemi di comprensibilità. Il testo va a capo alla fine di un'unità di trascrizione (ritagliate in base a criteri di tipo prosodico); per questioni di leggibilità, alcune unità intonative sono state riportate sulla stessa riga e separate attraverso l'uso di //.

Estratto dell'intervista PBC007

1. *un racconto e::hm::*
2. *che noi la casa era stata occupata prima dai tedeschi e poi dopo occupata*
3. *dalle truppe m::h alleate*
4. *e::hm: i magazzini del legno del del mio papà diventarono // dei depositi alimentari*
5. *per per le truppe alleate // e:: la comandante era una donna polacca*
6. *e noi sentivamo // e in u~ in un in un magazzino c'era // c'erano i::l pane, i dolci,*
7. *sentivamo questi profumi noi bambini, lì davanti*
8. *arrivava la polacca // con un bache~ // raus rau au au au*
9. *vabbè non ci ha mai dato niente*
10. *m::h delle volte c'era la frutta // era natale*
11. *era natale (.) e riempirono questo magazzino di paste*
12. *di e da due tre giorni facevano queste paste // e da da dai fornai e:: portavano*
13. *d~ d~ dei dei grandi canestri di queste pa~*
14. *la la polacca // chiudeva dentro (.) il magazzino // due ragazze*
15. *che erano (.) del nostro paese // che dovevano sistemare da questi cesti*
16. *de~ del degli scatoloni d~ (a)desso non mi ricordo se eran di legno o se erano di carta*
17. *col numero, mettere, contare e mettere il numero // di queste paste, di queste*
18. *mio padre scoprì // che // di fianco al mag~ a questo magazzino // c'era*
19. *una porta chiusa inverniciata di bianco che sembrava // un un muro*
20. *ma era una porta di legno*
21. *allora con una raspa // in fondo*
22. *d'accordo con le due ragazzine che erano // del nostro paese*
23. *con una raspi~ in fondo fece una specie di buco // dove si entrava con una mano*
24. *è il natale del millenovecentoquarantaquattro*
25. *la polacca che non aveva solo (.) il nostro magazzino ma ne aveva degli altri chiudeva queste ragazze // dentro*
26. *io ho la chiave l'ho tenuta come ricordo*
27. *chiudeva queste ragazze dentro // e poi andava negli altri luoghi*
28. *parlando con le ragazze che pativano come noi perchè*
29. *perchè lei non non non non dava non dava niente // era:: era:: rigida*
30. *e (.) questi erano // depositi di alimenti // che venivano poi portati al fronte // al fronte*
31. *che era fermo sul lamone // di faenza // d'accordo?*
32. *quando la polacca se ne era andata le abbiamo chiuse dentro*
33. *entrando in questo corridoio davanti a questa porticina col buco*
34. *io ((ride)) io ((ride)) // mettevo la manina*
35. *e quella che era dentro mi dava*
36. *le paste fu fu fu il natale dolce il dolce natale del quarantaquattro*
37. *mi mi mi dava que~ questi queste paste*
38. *mi ricordo che riempimmo diverse borse // proprio diverse borse*
39. *eran due le ragazze una che faceva questo lavoro, quell'altra che stava sulla porta*
40. *si arrivava eventualmente e vedeva l'arrivo della macchina del // del della co~ della jeep*
41. *d~ d~ della polacca*
42. *e io che ero dentro ((ride)) c:h // riempimmo mi ricordo // ta~ (.) tante borse*

L'attività ha inizio con l'ascolto dell'estratto sul sito del *corpus* (accessibile previa registrazione), a cui segue una fase di lettura e di osservazione guidata: tramite la predisposizione di attività specifiche, si può ad esempio proporre agli alunni di distinguere fra gli avvenimenti di primo piano e lo sfondo della narrazione (Weinrich 2004) e, contestualmente, di individuare e – per quanto possibile – caratterizzare i personaggi. In seguito, a gruppi, gli studenti possono cercare tutte le forme verbali e osservare come la parlante le utilizza o per descrivere la scena (trapassato e imperfetto) o per portare avanti la narrazione (presente storico, passato prossimo e remoto). Si può, infine, tracciare le catene referenziali relative a ciascun personaggio e osservare con quali strategie linguistiche vengono costruite (nomi, sintagmi nominali, pronomi, ecc.); un esempio potrebbe essere la catena anaforica relativa al personaggio della comandante polacca, che ricorre per tutto il testo, prima con un sintagma nominale pieno specificato da una predicazione nominale («la comandante era una donna polacca» r. 5), poi come soggetto zero («non ci ha mai dato niente» oppure «chiudeva queste ragazze dentro» r. 9 e 27), e con un aggettivo nominalizzato («la polacca»), in numerose occasioni (r. 8, 14, 25...).

L'attività potrebbe poi proseguire con una riflessione sui tratti tipici dell'oralità presenti nel testo, anche in vista di un compito di riscrittura dell'estratto in forma di testo narrativo: false partenze («d~ d~ dei dei grandi canestri di queste pa~», r. 13), ripetizioni («era natale // era natale», r. 10 e 11, oppure «perchè lei non non non non dava non dava niente // era:: era:: rigida», r. 29), autocorrezioni («in un magazzino c'era // c'erano i::l pane, i dolci» r. 6), e così via.

Prima che ciascuno proceda con la scrittura, in plenaria, si potrebbero elencare le operazioni necessarie a compiere questa trasformazione del testo: ad esempio, rimuovere i tratti dell'oralità, uniformare le forme verbali ed eventualmente renderle alla terza persona, rielaborare le frasi in periodi più complessi e riordinare lo svolgimento del racconto in base alla strategia narrativa scelta. Si potrà poi lasciare agli alunni il tempo necessario a progettare, stendere e revisionare il proprio testo.

4.2 Un'attività per l'insegnamento dell'italiano L2

Come anticipato (§ 2), nella classe di lingua seconda, raramente l'input orale presenta le caratteristiche della lingua realmente parlata dalle persone; di conseguenza, è anche raro che gli apprendenti siano coinvolti in momenti di riflessione sulla grammatica degli usi parlati della lingua: che cosa fanno le persone quando parlano, e in quanti modi diversi lo fanno? In base a che cosa diversificano le proprie scelte linguistiche?

L'attività presentata in questa sezione mira a portare queste domande nella classe di italiano L2, perché gli studenti possano sì imparare le regolarità degli usi linguistici, ma – proprio grazie all'analisi dell'oralità – scoprire nella lingua italiana anche contingenze e imprevedibilità. Quanto agli obiettivi linguistico-comunicativi più tipici della didattica linguistica, immaginiamo che, in una classe di livello B1 o B1+, si intenda esercitare l'abilità di raccontare eventi passati di carattere personale, sia oralmente che per iscritto. Sappiamo che, a questo livello di competenza, gli studenti sono in genere capaci di impiegare una pluralità di tempi al passato (passato prossimo, imperfetto e trapassato prossimo) quando raccontano; non usano forse il passato remoto, ma ne conoscono l'esistenza e di solito sono pronti a esplorarne forme e usi con una certa sistematicità. Infine, questo è spesso il momento in cui gli apprendenti alternano discorso diretto e indiretto e cominciano a sperimentare alcune costruzioni comparative come «(tanto) più [...] (quanto) più». In merito alla consapevolezza della variazione, la proposta operativa intende guidare lo sguardo degli studenti su una serie di fenomeni che distinguono la modalità scritta da quella orale, il monologo dal parlato interazionale e, infine, il registro formale da quello informale.

L'insieme di questi obiettivi può essere perseguito a partire dall'estratto riportato in (2), che contiene una sequenza interazionale dell'intervista PBA026:

Estratto dell'intervista PBA026

- 1 Enrica *ma ti sei fatto male?*
2 Simone *una volta sì // una volta sì*
3 *stavo provando un'asta nuova // che: mh lunga quattro metri e*
cinquanta
4 *quindi fatta per saltare (.) oltre i cinque metri*
5 *stavo: stavo provando quest'asta qui*
6 *e: praticamente // più l'asta è lunga // più*
7 *tu ci devi mettere forza e (.) più devi correre veloce*
8 *per riuscire (.) a piegare quest'asta*
9 *io la riu- la riuscii a piegare // andai su in altissimo*
10 *non abituato al: m:h // alla forza:*
11 *al ritorno dell- dell'asta c(io)è*
12 *quando quando si stende poi // l'asta ti spara in su*
13 *non abituato a questa spinta qui // e:hm: // mollai la presa*
14 *pensando di andare dall'altra parte dell'asticella*
15 *e invece cadetti giù come una pera*
16 *dove si va ad imbucare // ehm: l'asta*
17 *E IO lì (.) ringraziai*
18 *tutti gli anni prima (.) di ginnastica artistica*
19 *perché // ehm*
20 *son stato sveglio ho detto okay sto cadendo di schiena*
21 *sto cadendo dove si va ad imbucare l'asta*
22 *se cado di schiena lì mi faccio un male cane*
23 *alla spina dorsale ries- rischio di andare a romperla*
24 *e per fortuna sono riuscito a fare mezzo giro // e e:hm*
25 *ad atterrare prima un pochettino con i piedi per a- per ammortizzare*
poi cadere
26 *e: diciamo*
27 Enrica *c(io)è [come un gatto]*
28 Simone *[avanti]*
29 Enrica *come un gatto*
30 Simone *so- sono [stato] molto fortunato:*
31 Enrica *[mamma] // non ci posso pensare ((ride)) c(io)è*
32 Simone *(stare) restare cosciente durante tutta*
33 Enrica *mamma*
34 Simone *la caduta*
35 Enrica *ma infatti*

Specie se impiegato come testo centrale di una più ampia unità didattica, l'estratto può servire a raggiungere una pluralità di obiettivi in italiano L2. A livello morfosintattico, si presta a guidare l'attenzione degli studenti sulle forme in via di consolidamento o di

sviluppo a cui accennavamo (es. i tempi verbali) così come su altre, tra cui avverbi alterati come («in altissimo» r. 9, «un pochettino» r. 25) e costruzioni riflessive particolari («dove si va a imbucare l'asta» r. 16 e 21, o «quando si stende» r. 12). Queste stesse forme, insieme al lessico specialistico del salto con l'asta e degli sport, possono poi essere esercitate tramite varie attività orientate alla produzione. È specialmente in questa fase che è possibile sfruttare a fini didattici le peculiarità del testo. Possiamo ad esempio chiedere agli studenti di rielaborare il racconto di Simone, immaginando che lo abbia condiviso per iscritto prima con un amico e dopo con il professore di educazione fisica. Le due riscritture del testo, precedute dall'analisi attenta – in plenaria o a gruppi – dell'originale trascritto, aiuta a notare varie specificità del parlato, anche in relazione alle differenze tra formato monologico e interazionale: pause piene e vuote, false partenze, ripetizioni, frasi interrotte ed enunciati sintatticamente collegati distribuiti su più turni, e così via (cf. Voghera 1992). Se, per lavorare sullo scritto sia informale (all'amico) che formale (al professore), gli studenti dovranno pianificare la disposizione degli eventi, ridurre la densità informativa e diluire la sintassi, la variazione di registro richiesta dalla due riscritture potrà aiutarli a notare e manipolare altre caratteristiche del testo iniziale. Innanzitutto, molte delle espressioni colloquiali («ti spara su» r. 12, «mi faccio un male cane», r. 22), così come alcune forme poco controllate (ad esempio, quella analogica di «cadetti» r. 15), potrebbero essere mantenute efficacemente nel messaggio informale, ma risulterebbero probabilmente inappropriate se impiegate in quello formale al docente.

Prima di concludere quello che, per ragioni di spazio, può essere solo un accenno a uno dei molteplici usi didattici dell'estratto in (2), precisiamo che queste prassi di 'grammatica pedagogica' (Ciliberti 2013) degli usi orali dell'italiano non necessariamente devono realizzarsi interamente in lingua target, né a questa sono vincolate; si possono, al contrario, proporre interessanti confronti interlinguistici sulle (ir)regolarità delle lingue orali.

4.3 Un caso per l'italiano senza etichette'

A differenza dei due casi precedenti, in questa sezione, simuliamo un intervento didattico, senza distinguere a priori tra chi sta imparando l'italiano come lingua prima e chi come lingua aggiuntiva. Quella della classe multilingue ad abilità differenziate in italiano costituisce infatti la realtà della scuola, un contesto in cui la varietà e le complessità dei repertori linguistici individuali rendono spesso inutili – se non anche discriminatorie – categorie rigide come quella di 'nativo' e 'non nativo'. A partire da queste considerazioni, formuliamo quindi

una proposta di impiego del *corpus* KIParla che mira a insegnare l'italiano a tutti o, appunto, l'italiano senza etichette'.

Partiamo quindi dal testo in (3), il racconto di una vacanza costruito dialogicamente da tre parlanti (Giulia, Elena e Andrea nomi di fantasia):

Estratto dell'intervista PBA027

- 1 Andrea *sì ma la cosa più bella è che dopo siamo andati a fare un giro in un deserto ma*
2 *a mille e passa chilometri da questo posto*
3 *iniziamo a fare il il giro in cammello in mezzo al deserto con due:*
 m:h
4 *cio[è: due raga]zzi due beduini che ci accompagnavano*
5 Giulia *[e due beduini]*
6 Andrea *a una certa ci lasciano in mezzo alla duna ((ride))*
7 *e ci di[cono] ((ride)) passerete la notte qua ci vediamo domattina*
8 Giulia *[°voi passerete la notte qua°] ((ride))*
9 *su una duna*
10 Elena *co[sì]*
11 Giulia *[ba]sta*
12 Andrea *(no) non è non era non era questo il patto*
13 *[...]*
14 Giulia *e:h m:h e loro (alla fine) se la cantavano con questo deserto quando*
15 *semplicemente avevano una chiazza di di sabbia che era*
 veramente
16 *cioè (.) duecento metri per trecento*
17 *cioè oltretutto pieno di immondizia cioè e non potevi starci scalzo*
18 Andrea *sì probabilmente tipo non so la spiaggia di: della riviera dopo il il*
 jova beach
19 *party*
20 *[(una roba)]*
21 Giulia *[dopo la notte rosa dopo la] notte rosa*
22 Elena *[sì]*
23 Andrea *(sì)*

L'estratto riportato in (3) sarebbe stato adatto anche per raggiungere molti degli obiettivi descritti in precedenza (§ 4.1 e § 4.2). Tuttavia, il formato dialogico che lo caratterizza si presta particolarmente bene a stimolare negli studenti una riflessione sulla differenza tra norma e uso. In quest'ottica, potrebbe essere ancora una volta utile proporre attività di riscrittura, pensate appositamente per mettere in evidenza differenze e similitudini tra generi e tipi testuali

diversi. Ad esempio, a gruppi, gli alunni potrebbero essere invitati a scrivere una lettera formale di reclamo indirizzata a un'agenzia di viaggi, una recensione del servizio ricevuto da pubblicare su un blog per viaggiatori, oppure un racconto realizzato dialogicamente sempre tra amici, ma via Whatsapp. L'obiettivo è in ciascun caso valutare quali forme linguistiche presenti nel dialogo originale sia opportuno adottare o, al contrario, evitare, in funzione degli scopi, dei destinatari e dei mezzi di comunicazione. In quale tipo testuale risulterebbe adeguato o addirittura efficace mantenere espressioni come «a mille e passa chilometri» (r. 2), «a una certa» (r. 6), «se la cantavano» (r. 14)? E in quale, invece, tali scelte rischierebbero di generare un senso indesiderato di confidenza con il destinatario o di trasmettere un'impressione di approssimazione, giovanilismo o persino ignoranza?

Grazie al suo formato dialogico, l'estratto può inoltre essere efficacemente impiegato per invitare gli studenti a osservare la dimensione interazionale del parlato e, di riflesso, a considerare la complessità della scrittura. Per iniziare, l'insegnante potrebbe eliminare i turni di Giulia per mostrare che, in assenza di una delle parti, la narrazione non funziona e che, più in generale, ciascuno di noi, in interazione, è abituato a collaborare con gli interlocutori nella costruzione dei significati propri e altrui. Si potrebbe poi andare oltre e stimolare un confronto con la scrittura, facendo sperimentare agli alunni come, in maniera molto diversa da quanto accade con il parlato dialogico, i discorsi scritti siano da pianificare e costruire con attenzione, perché staccati dalla situazione dell'enunciazione e privi del contributo altrui; per questo motivo, quando si scrive, è fondamentale codificare minuziosamente a livello linguistico (testuale, sintattico e lessicale) significati che nell'oralità sono generalmente desumibili dal contesto fisico o, appunto, costruiti in interazione con gli altri (Pallotti et al. 2021).

È nel contesto di queste sperimentazioni sociolinguistiche, attivate dall'analisi dell'oralità, che l'insegnante può, all'occorrenza, proporre interventi mirati sulle forme. Tali interventi possono servire a rendere il testo più comprensibile a chi abbia bisogno di decodificare i riferimenti culturali presenti nel testo («riviera», «Jova Beach Party», «Notte Rosa», r. 18 e 21), ad arricchire di nuove parole («beduini», r. 4 e 5, «chiazza», r. 15, «cantarsela», r. 14) ed espressioni fisse («mille e passa chilometri», r. 2, «a una certa», r. 6) il vocabolario di altri o, infine, a far consolidare alcune regolarità morfosintattiche (il superlativo relativo o le forme passate dei verbi riflessivi).

5 Riflessioni conclusive

Come illustrato in apertura, l'impiego dei *corpora* nella prassi didattica è scarsamente diffuso, malgrado questi presentino numerosi vantaggi. Abbiamo anche visto come raramente esempi di parlato spontaneo e autentico trovino spazio tanto nell'educazione linguistica quanto nella didattica dell'italiano L2, specialmente se si osservano i materiali didattici oggi a disposizione.

Tuttavia, trascurare la dimensione orale – o trattarla solo a partire da dialoghi confezionati *ad hoc* – significa, nei fatti, non fornire agli alunni gli strumenti necessari a capire a fondo il funzionamento di questa modalità di comunicazione, passo fondamentale per comprendere anche le specificità della scrittura. A partire da queste considerazioni, abbiamo proposto tre attività didattiche elaborate a partire da altrettanti estratti dal *corpus* KIParla, che raccoglie campioni di lingua italiana autentica, prodotta da una varietà di parlanti, in contesti diversi e su argomenti vari (§ 3). È proprio grazie alla ricchezza sociolinguistica della risorsa che, come esemplificato, è possibile progettare attività didattiche di scoperta e manipolazione della lingua italiana orale in tutte le sue forme. Attività di esplorazione e ricerca come quelle proposte si adattano a una pluralità di contesti; in ciascuno, aiutano gli studenti a maturare la consapevolezza che l'italiano, come qualsiasi altra lingua, è tutt'altro che unitario e omogeneo e, al contempo, a potenziare le loro competenze di comprensione e produzione di una varietà di testi, sia orali che scritti.

Bibliografia

- Bachis, D. (2019). *Le grammatiche scolastiche dell'italiano edite dal 1919 al 2018*. Firenze: Accademia della Crusca.
- Ballarè, S.; Gorla, E.; Mauri, C. (2022). *Italiano parlato e variazione linguistica. Teoria e prassi nella costruzione del corpus KIParla*. Bologna: Pàtron Editore.
- Boulton, A. (2021). «Research in Data-Driven Learning». Pérez-Paredes, P.; Mark, G. (eds), *Beyond the Concordance: Corpora in Language Education*. Amsterdam: John Benjamins, 9-34.
- Boulton, A.; Cobb, T. (2017). «Corpus Use in Language Learning: A Meta-Analysis». *Language Learning*, 67(2), 156-77.
- Calaresu, E. (2016). «Dialogicità e grammatica». Andorno, C.; Grassi, R. (a cura di), *Le dinamiche dell'interazione*. Milano: AltLA, 13-27.
- Chen, M.; Flowerdew, J.; Anthony, L. (2019). «Introducing In-Service English Language Teachers to Data Driven Learning for Academic Writing». *System*, 87, 102-48.
- Corino, E. (2019). «Data-driven Learning: tra lingue straniere e CLIL, tra ricerca e didattica». *EL.LE*, 8(2), 271-85.
- Corino, E.; Onesti, C. (2019). «Data-Driven Learning: A Scaffolding Methodology for CLIL and LSP Teaching and Learning». *Frontiers in Education*, 4(7), 1-12.
- Ciliberti, A. (2013). «La nozione di Grammatica e l'insegnamento di L2». *Italiano LinguaDue*, 1, 1-14.
- Dal Negro, S. (2016). «Il dialogo nella riflessione grammaticale esplicita». Andorno, C.; Grassi, R. (a cura di), *Le dinamiche dell'interazione*. Milano: AltLA, 45-57.
- De Roberto, E.; Venco, C. (2024). «Una ricognizione in chiave discorsiva dei dialoghi artificiali scritti presenti nei manuali di italiano L2». Diadori, P.; Troncarelli, D. (a cura di), *Il dialogo nei manuali didattici di italiano L2 di ieri e di oggi*. Firenze: Franco Cesati.
- Dimalta, C.; Gasparetto, F.; Masillo, P.; Peri, G. (2024). «Spunti per un'analisi comparativa dei modelli di educazione linguistica dei libri di testo di italiano nella scuola primaria e secondaria di primo grado». Amenta, L.; Loiero, S. (a cura di), *Fare scuola con i libri di testo*. Firenze: Franco Cesati, 215-28.
- Ferreri, S. (2024). «Il libro di testo: una storia infinita». Amenta, L.; Loiero, S., *Fare scuola con i libri di testo*. Firenze: Franco Cesati, 23-42.
- Gablasova, D.; Bottini, R. (2022). «Spoken Learner Corpora for Language Teaching». Jablonkay, R.; Eneka, C. (eds), *The Routledge Handbook of Corpora and English Language Teaching and Learning*. London: Routledge, 296-311.
- Gillani, E.; Pernas, P. (2013). «Input orale e competenza dialogica: le scelte dei manuali di italiano LS». *SILTA*, 1, 61-100.
- Jefferson, G. (2004). «Glossary of Transcript Symbols with an Introduction». Lerner, G.H. (ed.), *Conversation Analysis: Studies from the First Generation*. Amsterdam: John Benjamins, 13-31.
- Lenci, A.; Montemagni, S.; Pirrelli, V. (2005). *Testo e computer*. Roma: Carocci.
- Leńko-Szymańska, A. (2014). «Is This Enough? A Qualitative Evaluation of the Effectiveness of a Teacher Training Course on the Use of Corpora in Language Education». *ReCALL*, 26(2), 260-78.
- Leńko-Szymańska, A. (2022). «Training Teachers and Learners to Use Corpora». Jablonkay, R.; Eneka, C. (eds), *The Routledge Handbook of Corpora and English Language Teaching and Learning*. London: Routledge, 509-24.
- Lin, M.H.; Lee, J.Y. (2015). «Data-Driven Learning: Changing the Teaching of Grammar in EFL Classes». *English Language Teaching Journal*, 69(3), 264-74.
- Lo Duca, M.G. (1997). *Esperimenti grammaticali*. Roma: Carocci.

- Lo Duca, M.G. (2003). *Lingua italiana ed educazione linguistica*. Roma: Carocci.
- Lusta, A.; Demirel, O.; Mohammadzadeh, B. (2023). «Language Corpus and Data Driven Learning (DDL) in Language Classrooms: A Systematic Review». *Helion*, 9, e22731.
- Ma, Q.; Tang, J.; Lin, S. (2022). «The Development of Corpus-Based Language Pedagogy for TESOL Teachers: A Two-Step Training Approach Facilitated by Online Collaboration». *Computer Assisted Language Learning*, 35(9), 2731-60. <https://doi.org/10.1080/09588221.2021.1895225>
- Mauri, C.; Ballarè, S.; Goria, E.; Cerruti, M. (2019). «KIParla Corpus: a New Resource for Spoken Italian». Bernardi, R.; Navigli, R.; Semeraro, G. (eds), *Proceedings of the 6th Italian Conference on Computational Linguistics*. <https://ceur-ws.org/Vol-2481/paper45.pdf>
- Mcenery, T.; Xiao, R. (2011). «What Corpora Can Offer in Language Teaching and Learning». Hinkel, E. (ed.), *Handbook of Research in Second Language Teaching and Learning*. London: Routledge, 140-58.
- Nuzzo, E. (2016). «Fonti di input per l'insegnamento della pragmatica in italiano L2: riflessioni a partire dal confronto tra manuali didattici, serie televisive e parlato spontaneo». Santoro, E.; Vedder, I. (a cura di), *Pragmatica e interculturalità in italiano lingua seconda*. Firenze: Franco Cesati, 15-27.
- Pallotti, G.; Borghetti, C.; Rosi, F. (2021). *Insegnare a scrivere nella scuola primaria: il progetto Osservare l'interlingua*. Cesena; Bologna: Caissa Italia.
- Pani, G. (2019). «Il parlato come strategia inclusiva nella didattica della grammatica». Ianes, D. (a cura di), *Didattica e inclusione scolastica*. Firenze: FrancoAngeli, 120-33.
- Pauletto, F. (2021). «Un genere a sé. L'interazione verbale nei videocorsi dei manuali di italiano LS/L2». *Italiano LinguaDue*, 13(1), 194-221.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sobrero, A.A. (1997). «Il peso della grammatica». Calò, R.; Ferreri, S. (a cura di), *Il testo fa scuola. Libri di testo, linguaggi ed educazione linguistica*. Firenze: La Nuova Italia, 489-502.
- Voghera, M. (1992). *Sintassi e intonazione nell'italiano parlato*. Bologna: il Mulino.
- Voghera, M. (2017). *Dal parlato alla grammatica*. Roma: Carocci.
- Voghera, M. (2023). «Scritto-parlato e altri modi nell'educazione linguistica». *Italiano LinguaDue*, 14(2), 1-18.
- Voghera, M.; Mayrhofer, V.; Ricci, M.; Rosi, F.; Sammarco, C. (2020). «Il Modokit: un identikit modale delle produzioni parlate e scritte dalle medie al biennio». Voghera, M.; Maturi, P.; Rosi, F. (a cura di), *Orale e scritto, verbale e non verbale: la multimodalità nell'ora di lezione*. Firenze: Franco Cesati, 225-45.
- Voghera, M.; Sammarco, C. (2021). *Ascoltare e parlare*. Firenze: Franco Cesati.
- Weinrich, H. (2004). *Tempus. Le funzioni dei tempi nel testo*. Bologna: il Mulino.
- Zareva, A. (2017). «Incorporating Corpus Literacy Skills into TESOL Teacher Training». *English Language Teaching Journal*, 71(1), 69-79.