

ANALISI DEL LINGUAGGIO UTILIZZATO DAGLI STUDENTI DURANTE UN'ATTIVITÀ DI PROBLEM SOLVING DI MATEMATICA IN UN AMBIENTE DIGITALE DI APPRENDIMENTO

Cecilia Fissore¹, Francesco Floris¹, Marina Marchisio Conte¹, Sergio Rabellino²

¹Dipartimento di Biotecnologie Molecolari e Scienze per la Salute, Università di Torino
{cecilia.fissore, francesco.floris, marina.marchisio}@unito.it

²Dipartimento di Informatica, Università di Torino
sergio.rabellino@unito.it

-- FULL PAPER --

ARGOMENTO: Istruzione secondaria

Abstract

L'utilizzo di un Ambiente Digitale di Apprendimento, un ecosistema di apprendimento in cui insegnare, apprendere e sviluppare competenze, per lo svolgimento di attività di problem solving consente la raccolta di una grande quantità di dati, tra cui informazioni sul linguaggio utilizzato dagli studenti. Molte ricerche in didattica della matematica hanno mostrato come tante difficoltà nella risoluzione dei problemi siano legate a componenti linguistiche che influenzano la rappresentazione, la comprensione e l'interpretazione della situazione. Risulta quindi rilevante lo studio del linguaggio utilizzato dagli studenti, considerando tutti i registri che vengono utilizzati nel fare e comunicare la matematica. La domanda di ricerca che guida questa ricerca è: il linguaggio che utilizzano gli studenti nella risoluzione di un problema matematico quando si confrontano con i loro pari è il medesimo che utilizzano nella consegna della loro risoluzione finale o come si differenzia? Il contesto della ricerca è un'attività di problem solving in piattaforma, con elementi di gamification, che ha coinvolto circa 400 studenti di grado 11. Gli studenti avevano la possibilità di collaborare con i loro pari alla risoluzione del problema utilizzando un forum in piattaforma, dopodiché ognuno di loro consegnava la propria risoluzione e riceveva una valutazione da parte di un tutor. Per rispondere alla domanda di ricerca è stato analizzato il linguaggio nei 382 elaborati finali degli studenti e quello in 620 interventi nel forum. La metodologia di ricerca prevede la progettazione e l'analisi di due corpora (uno per le consegne e uno per i forum) e tecniche di Natural Language Processing per il confronto dei linguaggi. I risultati ottenuti potranno anche essere utilizzati per l'implementazione di nuove forme di tutoraggio (anche basate su LLM) per aiutare gli studenti a sviluppare competenze matematiche e di problem solving. Questa ricerca può avere diverse applicazioni nell'ambito dello sviluppo di competenze matematiche, di problem solving, linguistiche e nell'ambito dell'apprendimento personalizzato.

Keywords: Ambiente Digitale di Apprendimento, Analisi del linguaggio, Matematica, Problem Solving.

1 INTRODUZIONE

Negli ultimi tempi si è consolidata la consapevolezza dell'estrema importanza della competenza linguistica nell'apprendimento della matematica. Da un lato, l'esigenza crescente di insegnare matematica in comunità multilingue o comunque composte da individui con competenze linguistiche eterogenee, ha fatto emergere nuovi problemi; dall'altro il diffondersi di pratiche didattiche basate su diverse forme di comunicazione richiede più attenzione verso i linguaggi utilizzati e le funzioni che assumono [1]. L'importanza dei linguaggi è stata valorizzata in modo particolare da Sfard [2] che interpreta il pensiero come forma di comunicazione e considera i linguaggi non come veicoli di significati

preesistenti ma come costruttori dei significati stessi. Sotto questa prospettiva il linguaggio non è solo uno strumento di comunicazione ma influenza il pensiero in modo determinante [1].

Numerosi studi in didattica della matematica hanno dimostrato che le difficoltà di apprendimento in questa disciplina possono derivare da difficoltà linguistiche, come l'acquisizione, la comprensione e la gestione del suo linguaggio specialistico [3, 4, 5]. Come sottolineano le Indicazioni Nazionali [6], infatti, padroneggiare una disciplina significa non solo sviluppare la conoscenza dei contenuti, ma anche essere consapevoli dei mezzi linguistici che una lingua richiede per esprimerla, ovvero le caratteristiche lessico grammaticali, i modelli di testualizzazione, nonché l'adeguatezza e l'accettabilità.

Le difficoltà linguistiche degli studenti possono influenzare anche le competenze di problem solving, ad esempio per quanto riguarda la rappresentazione, la comprensione e l'interpretazione della situazione problematica [7, 8]. Secondo Ferreri [9], le difficoltà di carattere linguistico possono riguardare sia il riconoscimento delle situazioni e la loro classificazione come problemi sia la loro risoluzione. Tra queste incidono anche gli atteggiamenti che gli studenti hanno elaborato nel tempo nei confronti del testo di un problema, e, in particolare, come ci si approccia alla lettura di un problema [10]. Altre difficoltà linguistiche possono riguardare l'interpretazione e la produzione dei testi in conformità ai modi di espressione tipici del linguaggio orale quotidiano piuttosto che a quelli specifici dei linguaggi scientifici e l'incapacità di utilizzare in ambito scientifico le competenze linguistiche acquisite in ambito linguistico-letterario [11].

La linguistica dei corpora può supportare in modo efficace lo studio dei linguaggi specialistici e lo sviluppo di competenze linguistiche e disciplinari [12, 13, 14]. Con il termine "corpus" si intende una raccolta di testi (scritti, orali o multimediali) o parti di essi in numero finito in formato elettronico trattati in modo uniforme così da essere gestibili ed interrogabili informaticamente, dove i testi sono scelti in modo da essere autentici e rappresentativi [15]. Attraverso specifici software è impossibile creare un corpus e taggarlo automaticamente per parti del discorso (in inglese "PoS – Part of Speech"), per avere la possibilità di scegliere solo una parte del discorso da interrogare all'interno del corpus.

L'obiettivo di questa ricerca preliminare è quello di indagare il linguaggio utilizzato da 400 studenti di grado 11 durante un'attività di problem solving in matematica all'interno di un Ambiente Digitale di Apprendimento. L'utilizzo di una piattaforma Moodle per lo svolgimento di attività di matematica ma anche per far parlare gli studenti di matematica permette, tra i numerosi vantaggi, la raccolta di una grande quantità di dati. Da un lato questo si traduce nello studio e nell'applicazione dei Learning Analytics (LA) per analizzare grandi moli di dati sull'apprendimento e dall'altro nello studio di tecniche di Natural Language Processing (NLP) per lo studio del linguaggio. Gli studenti avevano la possibilità di collaborare con i loro pari alla risoluzione del problema utilizzando un forum in piattaforma, dopodiché ognuno di loro consegnava la propria risoluzione e riceveva una valutazione da parte di un tutor. Si è voluto quindi studiare il linguaggio degli studenti analizzando due registri diversi: quello più formale della consegna in piattaforma (per i quali gli studenti ricevevano una valutazione) e quello più informale del forum. La domanda di ricerca che guida questa ricerca è: il linguaggio che utilizzano gli studenti nella risoluzione di un problema matematico quando si confrontano con i loro pari è il medesimo che utilizzano nella consegna della loro risoluzione finale o come si differenzia? Per rispondere alla domanda di ricerca è stato analizzato il linguaggio nei 382 elaborati finali degli studenti e quello in 620 interventi nel forum, attraverso la creazione e la consultazione di due corpora.

1.1 Il contesto della ricerca

Il contesto della ricerca è un'attività di problem solving asincrona online che si è svolta all'interno del progetto Digital Math Training (DMT), sulla piattaforma Moodle ad esso dedicata (raggiungibile al link <https://digitalmatetraining.i-learn.unito.it>). Il DMT è un percorso di eccellenza finalizzato a far sviluppare agli studenti della scuola secondaria di secondo grado competenze matematiche, digitali, di problem solving e di lavoro collaborativo e in ogni edizione coinvolge studenti provenienti da scuole secondarie di secondo grado del Piemonte e Valle d'Aosta. Il progetto si è concluso nel 2023 e ha visto in totale nove edizioni, a partire dall'anno scolastico 2014/15.

Ogni anno, tra i 350 e 750 studenti partecipano volontariamente a un training online della durata di tre mesi sulla piattaforma Moodle dedicata al progetto, allestita e gestita dal Servizio ICT del Dipartimento di Informatica dell'Università di Torino. Quest'ultima è integrata con un Ambiente di Calcolo Evoluto (ACE), uno strumento ideato per l'apprendimento delle discipline STEM e il sistema di servizio di Web

Conference Adobe Connect [16]. Il training online prevede la risoluzione, facoltativa, di 8 problemi contestualizzati nella realtà, uno ogni dieci giorni circa, utilizzando l'ACE Maple (<https://www.maplesoft.com/>). Gli studenti possono collaborare con i loro pari, dello stesso ordine e grado, alla risoluzione del problema attraverso un forum specifico per ogni problema, ma ogni studente deve consegnare in piattaforma la propria risoluzione [17, 18]. Sono inoltre formati e accompagnati nel loro percorso da tutor universitari che possono intervenire nei forum e ogni settimana tengono un tutorato online incentrato sulle funzionalità dell'ACE. Durante tali appuntamenti, i tutor possono condividere il proprio schermo e l'audio con i partecipanti e utilizzare una chat per comunicare con gli studenti che non utilizzano un microfono.

Per motivare gli studenti a partecipare attivamente, misurare con opportuni indicatori il livello che raggiungono, e avere informazioni relative alla loro partecipazione, è stato implementato un sistema di valutazione formativa che fa uso di opportuni punteggi, chiamati "Digital Math Coins" (DMC). Grazie alla potenzialità di Moodle di poter valutare tutte le attività e di tenere traccia del completamento dell'attività vengono assegnati dei punteggi agli studenti per ogni attività online, sfruttando elementi di gamification [19]. Nei forum vengono assegnati 3 DMC per ogni intervento pertinente e tale da permettere uno sviluppo ulteriore della discussione, al fine di evitare l'abuso degli stessi con il solo fine di guadagnare punti [16]. Per la partecipazione ai tutorati online vengono assegnati 20 DMC. La consegna del problema viene valutata dal tutor da 1 a 100 DMC, secondo un'apposita griglia di valutazione, per dare un feedback agli studenti sulle competenze digitali e di problem solving acquisite. L'attività Compito viene quindi utilizzata sia per la consegna degli elaborati degli studenti (con una specifica e rigorosa finestra di disponibilità) sia per restituire loro le valutazioni ed eventuali commenti.

Per questa ricerca è stato scelto il problema dal titolo "Il negozio online" che è stato proposto, a studenti di grado 11, in tre edizioni del progetto: 2017/18, 2018/19, 2021/22. Nella prima edizione è stato proposto come primo problema del training, nella seconda come quarto problema del training e nella terza come quinto problema. Il problema appartiene all'ambito "relazioni e funzioni" e ha quattro richieste. Il testo è il seguente:

Claudio, che di professione fa il commerciante di articoli sportivi e vuole essere al passo con i tempi, apre un sito web in cui è possibile acquistare i suoi prodotti online. Vorrebbe massimizzare il suo guadagno sulla vendita di Fasce da Braccio Porta cellulari. Il suo fornitore gli vende ogni singolo pezzo a 2.50€, mentre lui lo rivende a 5.10€. Mediamente ogni mese 350 clienti acquistano una Fascia da Braccio Portacellulare. Statisticamente, ogni volta che applica al singolo pezzo una riduzione del prezzo di 0.20€, Claudio nota che il numero di acquirenti al mese aumenta di 80 unità.

1) Supponendo che all'aumentare del numero di riduzioni di prezzo applicate corrisponda un costante aumento di acquirenti, a quanto deve vendere le Fasce da Braccio per massimizzare il suo guadagno?

2) Rappresentare graficamente il guadagno totale in funzione del numero di riduzioni effettuate.

Claudio vende anche cerotti tape per sportivi, ogni rotolo a un prezzo di 1.85€. Facendo un rapido giro su internet, vede che anche un altro rivenditore concorrente li vende scontati del 13.8% sul suo stesso prezzo. Claudio dunque sconta anche i sui cerotti.

3) Qual è lo sconto massimo che Claudio può applicare per rivendere i cerotti allo stesso prezzo dell'altro rivenditore? (Esprimere il risultato con non più di una cifra decimale)

4) Costruire un sistema di componenti interattive che generalizzino il punto (1). Dato il prezzo di vendita, il costo dal fornitore e il valore di riduzione del prezzo di un articolo, insieme al numero di persone che acquistano al mese e l'aumento causato da una riduzione, mi restituisca il guadagno ottimale.

L'ultima richiesta, caratteristica di ogni problema, chiede di costruire un sistema di componenti interattive che generalizzi il processo risolutivo del problema, in modo da visualizzare come variano i risultati di un problema quando vengono modificati i dati iniziali.

Per ogni edizione del training online il problema aveva un apposito Forum in piattaforma come quello in Figura 1. Gli studenti potevano creare una nuova discussione e inserire un intervento nel forum, rispondere a un intervento di un loro pari o semplicemente leggere gli interventi senza intervenire.

Discussioni sul problema "Il negozio online"

Da fare: Effettuare interventi: 1 Da fare: Ricevere una valutazione

Utilizza questo forum per confrontarti con i compagni sulla soluzione di questo problema.

[Aggiungi un argomento di discussione](#)

Discussione	Iniziato da	Ultimo intervento ↓	Repliche
Maggiori informazioni	 13 dic 2017	 3 gen 2018	7
punto 3 del problema	 20 dic 2017	 1 gen 2018	9
Equazione della parabola	 16 dic 2017	 22 dic 2017	26
prima richiesta del il negozio online	 14 dic 2017	 22 dic 2017	7
vertice della parabola	 22 dic 2017	 22 dic 2017	5

Fig. 1: Esempio di forum in cui gli studenti si confrontavano sulla risoluzione del problema

2 METODOLOGIA DI RICERCA

Per ogni edizione del progetto sono state considerate tutte le consegne finali e tutti gli interventi nel forum degli studenti. Il numero totale di studenti partecipanti delle tre edizioni è 790, di questi 382 hanno consegnato la propria risoluzione in piattaforma e 130 hanno inserito almeno un intervento nel forum. In totale sono state create 87 discussioni e sono stati inseriti 620 interventi nei vari forum. La seguente tabella (Tab. 1) riassume tutti i dati considerati per l'analisi, divisi per le varie edizioni.

Edizione	Numero problema nel training	Numero di discussioni aperte	Numero di interventi forum	Numero di consegne	Numero di partecipanti al training	Numero studenti che hanno scritto nel forum
2017/18	1	47	354	152	289	67
2018/19	4	21	173	111	265	41
2021/22	5	19	93	119	236	22
Totale		87	620	382	790	130

Tab.1: Riepilogo dei dati considerati per l'analisi suddivisi per edizioni del training online

I tre forum sono stati esportati in formato XLSX dalla piattaforma. In questo modo si è ottenuto per ogni forum un file con molte informazioni, tra le quali:

- id dell'intervento;
- id della discussione;
- informazioni sull'intervento genitore;
- nome completo e id dell'utente che ha inserito l'intervento;
- informazioni su quando è stato inserito l'intervento;
- oggetto dell'intervento;
- messaggio dell'intervento.

Il numero di interventi inseriti corrisponde al numero di righe del file, il numero di discussioni sono i valori univoci degli id delle discussioni e il numero di studenti che ha inserito almeno un messaggio nel forum sono i valori univoci degli id dell'utente. Per la creazione del corpus sono stati considerati esclusivamente tutti i messaggi degli studenti, inseriti in un unico file txt.

Le consegne sono state scaricate dalla piattaforma ed è stato creato uno script in Python che navigando la cartella delle consegne dei problemi, leggesse ciascun problema (i file Maple che hanno estensione mw possono essere letti come xml) e salvasse in un file txt esclusivamente il testo scritto dagli studenti, escludendo formule matematiche e comandi del software. Ne consegue che ad essere particolarmente influenti nella creazione del corpus sono state le risoluzioni in cui gli studenti hanno argomentato maggiormente il processo risolutivo e i risultati ottenuti.

2.1 Creazione e analisi dei due corpora

Per la creazione dei corpora è stato utilizzato il software Sketch Engine (<https://www.sketchengine.eu/>). Sketch Engine è un software di gestione dei corpora e di analisi del testo. Il suo scopo è consentire alle persone che studiano il comportamento linguistico di effettuare ricerche in grandi raccolte di testi in base a query complesse e motivate linguisticamente. Un corpus creato su questo software è taggato per parti del discorso e annotato per essere pronto per ricerche complesse di frasi e strutture linguistiche.

Il medesimo software è stato utilizzato per la consultazione e la comparazione dei due corpora. Il confronto è stato effettuato con la funzione "compare corpora", che permette di confrontare corpora nella stessa lingua confrontando le loro parole. Viene calcolato un punteggio che indica in quale misura i corpora sono simili o diversi: un punteggio di 1 indica corpora identici, più alto è il punteggio più i corpora sono diversi. Poiché il confronto viene effettuato sulle parole, il punteggio non è influenzato dalla lunghezza della frase, dal numero di documenti, dalla dimensione del corpus o dalle caratteristiche grammaticali. Per effettuare il confronto, Sketch Engine identifica le 5.000 parole più frequenti nei due corpora; le due liste vengono combinate in una sola e i duplicati vengono rimossi. Per ogni parola viene calcolato il "keyness score" che è una misura del rapporto delle frequenze nel corpus principale (con una frequenza relativa più alta) e nel corpus di riferimento (con una frequenza relativa più bassa). A questo punto vengono identificate le 500 parole con il "keyness score" più alto e viene restituita la media aritmetica di questi valori. In seguito, è stata svolta un'analisi delle parole più frequenti nelle due wordlist per studiare più nel dettaglio similitudini e divergenze.

Il secondo strumento utilizzato per confrontare i due corpora è il metodo "Similarity.calculate" di Python che effettua un confronto tra le parole più frequenti utilizzando il coefficiente di correlazione di Spearman. In questo caso il valore di somiglianza è compreso tra 0 (molto diversi) e 1 (molto simili). Questo metodo si basa sulle ricerche di Dunn ([20]) che dimostrano che tra i metodi di calcolo della similarità tra corpora basate sulla frequenza dei termini, quelli più consistenti sono quelli che funzionano classificando le frequenze di parole o caratteri e utilizzando il coefficiente di correlazione di Spearman o il χ^2 tra i ranghi di frequenza. Il χ^2 è leggermente migliore per i campioni con lo stesso numero di caratteristiche, ma il coefficiente di Spearman ha una maggiore flessibilità con un impatto minimo sulle prestazioni nel caso di corpora di grandezza differente (come in questo caso).

Nell'ottica di studiare in quale dei due corpora gli studenti avessero utilizzato un linguaggio più specialistico, i due corpora sono stati confrontati (con gli stessi metodi spiegati in precedenza) con un corpus di linguaggio specialistico dell'ambito "relazioni e funzioni". Questo corpus (intitolato "Corpus_Matematica_funzioni") è stato creato selezionando testi relativi al concetto di funzione da libri di testo per la scuola secondaria di secondo grado.

3 RISULTATI

Sono stati creati due corpora, uno contenente i txt di tutte le consegne degli studenti (intitolato "DMT_Moot24_consegne") e uno contenente tutti i txt degli interventi nel forum (intitolato "DMT_Moot24_forum"). Il primo corpus è composto da 104846 tokens (conteggio delle occorrenze di una stessa parola nel testo), 86544 types (parole) e 2433 frasi. Il secondo corpus è composto da 16687 tokens, 13914 types e 526 frasi.

La ricchezza del vocabolario, che si trova dividendo l'insieme dei types per l'insieme dei tokens, è 0,83 per entrambi i corpora. Il rapporto type/token indica la varietà di parole differenti contenute in un testo e un valore superiore al 20% caratterizza un corpus testuale non sufficientemente ricco dal punto di vista lessicale. L'alto valore ottenuto è in linea con le aspettative, dal momento che c'è un solo argomento molto specifico, e dimostra il fatto che (soprattutto nell'attività Forum) gli studenti si siano concentrati su questo argomento e non siano andati fuori tema.

Nella tabella 2 sono riportati i primi quindici valori delle wordlist (nomi) per entrambi i corpora, in cui sono stati evidenziati in blu i nomi inerenti al problema, in arancione i nomi inerenti all'ACE e in verde i nomi legati al linguaggio specialistico della matematica.

DMT_Moot24_consegne		DMT_Moot24_forum	
Nome	Frequenza	Nome	Frequenza
1 prezzo	2,201	prezzo	138
2 guadagno	1,553	problema	118
3 riduzione	1,306	comando	85
4 claudio	1,303	punto	83
5 numero	1,217	sconto	73
6 sconto	999	claudio	70
7 punto	519	numero	60
8 cerotto	499	maple	59
9 valore	453	parabola	59
10 parabola	388	guadagno	56
11 cliente	381	valore	54
12 mese	372	equazione	51
13 funzione	370	ciao	47
14 vendita	354	funzione	45
15 rivenditore	353	grafico	41

Tab. 2: Primi quindici valori delle wordlist (nomi) per entrambi i corpora

I nomi presenti nelle due wordlist sono molto simili ma cambia la loro posizione in termini di frequenza all'interno del corpus. In entrambe le wordlist, come prevedibile, i nomi più frequenti sono quelli inerenti al problema (prezzo, guadagno, sconto, claudio, ...). È interessante osservare come la parola "punto" che in matematica ha un significato specifico, in questo caso sia presente nei corpora solo nell'accezione "punto del problema". Nel corpus "DMT_Moot24_forum" sembrano essere di più le parole inerenti al linguaggio matematico. Questo risultato potrebbe essere dato dal fatto che, per farsi capire meglio dai loro pari e fornire spiegazioni più chiare, gli studenti argomentano i loro procedimenti e i loro risultati con spiegazioni matematiche (che potrebbero dare per scontate nell'argomentazione della risoluzione del problema consegnato). Ad esempio, la parola "equazione" è presente al 12esimo posto nella wordlist del corpus dei forum e al 27esimo nel corpus delle consegne. Gli studenti hanno utilizzato questo termine per fornire spiegazioni matematiche ("l'equazione iniziale di una parabola è $ax^2+bx+c=0$ ", "e poi sostituire la x nell'equazione e trovare così il guadagno massimo") ma anche per spiegare delle procedure dell'ACE per ottenere un risultato ("Per fare il grafico su Maple devi andare sull'equazione della parabola", "ho assegnato all'equazione un nome con(=)").

Un'altra differenza tra i due corpora è la presenza di termini inerenti all'ACE nel corpus "DMT_Moot24_forum". Questo risultato è dato dal fatto che gli studenti potevano utilizzare i forum anche per aiutarsi a vicenda nell'utilizzo del software per la risoluzione del problema. Questo aspetto condiziona anche l'accezione con cui sono stati utilizzati alcuni termini specialistici del linguaggio matematico. Ad esempio, il termine "funzione" nel corpus dei forum viene utilizzato anche come sinonimo di "comando" dell'ACE ("Ci sono apposite funzioni per modificare il risultato (ad esempio `trunc()`"), "basta cercare la funzione su 'Maple help' e potrai ricavare il codice"). Nel corpus delle consegne questo non succede e il termine viene utilizzato con il suo significato specialistico ("E'

possibile sostituire i dati alla funzione $G=R-S$ (con G =guadagno, R =ricavo e S =spesa), "La formula che descrive il guadagno G in funzione del Ricavo R e del costo C ").

Un'ultima differenza è la presenza in questo corpus di termini del linguaggio colloquiale, come ad esempio "ciao", tipico del linguaggio dei forum e utilizzato sempre per introdurre un intervento.

Nella tabella 3 sono riportati i primi quindici valori delle wordlist (verbi) per entrambi i corpora, in cui sono stati evidenziati dello stesso colore i verbi presenti in entrambi i corpora. Nel corpus delle consegne sono presenti molti più verbi inerenti al problema, come ad esempio "vendere", "applicare", "rivendere". Questo risultato può essere dato dal fatto che gli studenti argomentavano i risultati ottenuti contestualizzandoli all'interno del problema ("Il prezzo al quale dovrà vendere le fasce sarà, pertanto di", "Claudio per ottenere il maggior guadagno dovrebbe vendere le Fasce da Braccio a"). Questo aspetto è molto meno presente nei forum dal momento che l'attenzione è maggiormente rivolta al procedimento rispetto ai risultati ottenuti (che gli studenti non sempre rivelano dal momento che si tratta sempre di una competizione).

Nel corpus dei forum sono presenti quasi esclusivamente verbi del linguaggio comune mentre nel corpus delle consegne sono presenti due verbi inerenti al linguaggio matematico "calcolare" e "rappresentare" (presenti anche nell'altro corpus ma rispettivamente alla 27esima e 49esima posizione). Quest'ultimo verbo, tuttavia, è utilizzato dagli studenti sia nell'accezione specialistica ("Rappresentando in un grafico (x =numero di riduzioni, y =guadagno) la funzione") sia nell'accezione comune come sinonimo di "esprime" ("l'equazione della curva che rappresenta le vendite di Claudio"). È importante osservare che verbi del linguaggio colloquiale sono presenti anche nel corpus delle consegne, dove si presuppone che gli studenti utilizzino un registro più formale. Ad esempio, è presente alla 24esima posizione il verbo "andare", scelto dagli studenti per spiegare il procedimento risolutivo ("L'idea quindi è quella di andare a creare una formula", "andremo a lavorare su numeri davvero piccoli") adottando anche costruzioni e verbi non corrette dal punto vista matematico ("andremo a mettere la funzione che ci troveremo su un grafico per vedere meglio la situazione") o da quello linguistico ("andiamo a graficare la funzione che ci siamo trovati").

DMT_Moot24_consegne		DMT_Moot24_forum	
Verbo	Frequenza	Verbo	Frequenza
1 vendere	694	essere	428
2 applicare	632	avere	244
3 avere	466	potere	168
4 essere	456	fare	143
5 trovare	403	dovere	101
6 dovere	389	usare	84
7 ottenere	358	trovare	63
8 effettuare	331	dare	55
9 fare	319	andare	48
10 rivendere	306	provare	47
11 acquistare	289	sapere	47
12 potere	247	dire	45
13 aumentare	241	pensare	42
14 rappresentare	222	venire	40
15 calcolare	214	intendere	39

Tab. 3: Primi quindici valori delle wordlist (verbi) per entrambi i corpora

Attraverso la funzione "compare corpora" è possibile constatare che i due corpora sono notevolmente diversi (Fig. 2). Un risultato simile si ottiene con il metodo di Python che restituisce un valore di 0,49.

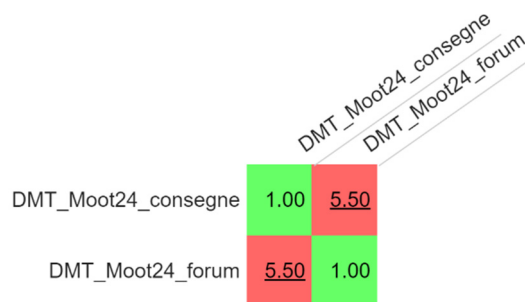


Figura 2: Confronto dei due corpora su Sketch Engine

Questo risultato può essere dato dal fatto che, come abbiamo visto negli esempi precedenti, gli studenti utilizzano un linguaggio più colloquiale nel forum. Inoltre, anche se i due corpora sono entrambi inerenti alla risoluzione del problema, gli studenti nel forum trattano anche argomenti diversi (come l'utilizzo dell'ACE) e si concentrano maggiormente sulla risoluzione del problema. In entrambe le attività di Moodle è fondamentale l'argomentazione ma nel forum ci studenti argomentano maggiormente i procedimenti risolutivi, gli aspetti matematici o le loro decisioni, mentre nelle consegne argomentano maggiormente i risultati ottenuti contestualizzandoli all'interno del problema.

Confrontando su Sketch Engine i due corpora con il corpus "Corpus_Matematica_funzioni" si ottiene il risultato in Figura 3.

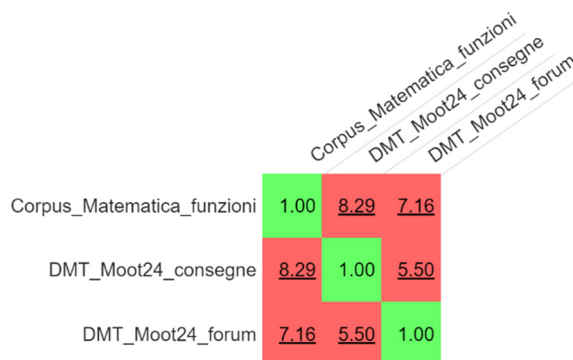


Figura 3: Confronto dei tre corpora su Sketch Engine

Entrambi i corpora "DMT_Moot24_consegne" e "DMT_Moot24_forum" sono molto diversi dal corpus "Corpus_Matematica_funzioni". Questo risultato è dato dal fatto che i due corpora sono inerenti alla risoluzione del problema e, pur utilizzando il concetto di funzione, non si riferiscono alla spiegazione del concetto di funzione in generale. In questo caso il corpus dei forum è più simile al corpus teorico rispetto al corpus delle consegne. Il risultato opposto si ottiene utilizzando il metodo di Python, che restituisce 0,42 come valore di similarità tra "Corpus_Matematica_funzioni" e il corpus "DMT_Moot24_forum" e 0,5 nel caso del corpus "DMT_Moot24_consegne". Anche in questo caso, tuttavia, i due corpus sembrano abbastanza diversi dal corpus di riferimento "Corpus_Matematica_funzioni".

4 APPLICAZIONI, CONCLUSIONI E SVILUPPI FUTURI

In questa ricerca preliminare abbiamo studiato il linguaggio utilizzato da circa 400 studenti di grado 11 durante un'attività di problem solving in matematica all'interno di un Ambiente Digitale di Apprendimento. Per rispondere alla domanda di ricerca è stato analizzato il linguaggio nei 382 elaborati finali degli studenti e quello in 620 interventi nel forum, utilizzando la comparazione di corpora e tecniche di Natural Language Processing. I risultati ottenuti dimostrano che il linguaggio che utilizzano gli studenti nella risoluzione di un problema matematico quando si confrontano con i loro pari non è lo stesso che utilizzano nella consegna della loro risoluzione finale. Questa ricerca ha evidenziato come gli studenti argomentino maggiormente i loro procedimenti e i loro risultati nel forum, per farsi capire meglio dai loro

pari e per difendere le loro idee e i loro punti di vista. Le argomentazioni nel forum sono caratterizzate anche da spiegazioni matematiche che, invece, sono meno presenti nelle consegne. Gli studenti, tuttavia, dovrebbero argomentare le loro risoluzioni anche nella consegna finale con il tutor (o in generale gli educatori). Questo risultato può essere uno spunto per i docenti di matematica a fornire una spiegazione più dettagliata nel testo del problema, inserendo richieste più esplicite come “motiva”, “spiega”, etc.

Questa ricerca costituisce la base dello studio dell'utilizzo del linguaggio specialistico in due contesti diversi: uno più formale, nell'elaborato finale che consegnano in piattaforma, e uno meno formale, nei forum dedicati alla risoluzione del problema. Ad esempio, l'analisi proposta in questa ricerca può essere approfondita confrontando i corpora dei forum e delle consegne con un corpus di linguaggio specialistico creato ad hoc inerente la risoluzione del problema.

Un naturale sviluppo di questo studio consiste nell'analizzare se e quanto l'utilizzo di un linguaggio specialistico (utilizzato dagli studenti nei due contesti o solo nell'elaborato finale) influisca sulla valutazione finale ottenuta per la risoluzione del problema, e di conseguenza sullo sviluppo delle competenze matematiche e di problem solving.

I risultati ottenuti in questa ricerca mostrano l'importanza di analizzare il linguaggio degli studenti per capire l'influenza delle competenze linguistiche sulle competenze matematiche e di problem solving e per studiare come gli studenti sviluppino tutte queste competenze. I risultati aprono anche a riflessioni sull'importanza di considerare, ed eventualmente valutare, le argomentazioni prodotte dagli studenti non solo dal punto di vista matematico ma anche dal punto di vista linguistico. Questo studio suggerisce di intraprendere azioni per sensibilizzare e formare i tutor del progetto e, in generale i docenti di matematica, sul tema del linguaggio specialistico [21]. Se il docente o il facilitatore è in grado di porre maggiore attenzione al linguaggio utilizzato (nelle attività di problem solving ma anche nelle spiegazioni teoriche orali o scritte), ne beneficiano anche gli studenti perché li guida nella comprensione ed elaborazione dei testi matematici e li aiuta a fare attenzione al linguaggio che utilizzano.

Infine, i risultati ottenuti da queste analisi possono essere utilizzati per l'implementazione di nuove forme di tutoraggio, anche automatiche, per aiutare gli studenti a sviluppare competenze matematiche e di problem solving, ma anche competenze linguistiche. Tool di intelligenza artificiale integrati con la piattaforma Moodle possono dare vita a strumenti utili per classificare gli interventi degli studenti all'interno dei forum (spiegazione matematica, spiegazione sull'utilizzo dell'ACE, etc.) e per fornire feedback personalizzati.

Dal momento che un Ambiente Digitale di Apprendimento permette di raccogliere molti dati, si può inoltre estendere lo studio per applicare l'analisi (anche in maniera automatizzata) a differenti problemi dello stesso ambito o di ambito diverso.

Riferimenti bibliografici

- [1] Ferrari, P. L. (2003). Costruzione di competenze linguistiche appropriate per la matematica a partire dalla media inferiore. *L'insegnamento della matematica e delle scienze integrate A*, 26, 469-496.
- [2] Sfard, A. (2000), *Symbolizing Mathematical Reality Into Being - Or How Mathematical Discourse and Mathematical Objects Create Each Other*, in Cobb, P., E.Yackel and K.McClain (eds.), *Symbolizing and Communicating in Mathematics Classrooms*, Lawrence Erlbaum Associates, 37-98, Routledge.
- [3] Ferrari, P. L. (2021). *Educazione matematica, lingua, linguaggi. Costruire, condividere e comunicare matematica in classe*. Torino: Utet.
- [4] Prediger, S., Erath, K., & Moser Opitz, E. (2019). The language dimension of mathematical difficulties. In A. Fritz, V. Haase, & P. Räsänen, *International Handbook of math learning difficulties: From the laboratory to the classroom*, 437–455. Cham: Springer. doi:10.1007/978-3-319-97148-3_27

- [5] Viale, M. (2019). I fondamenti linguistici delle discipline scientifiche. L'italiano per la matematica e le scienze a scuola. Padova: Cleup.
- [6] MIUR. (2010). Indicazioni nazionali riguardanti gli obiettivi specifici di apprendimento concernenti le attività e gli insegnamenti compresi nei piani degli studi previsti per i percorsi liceali.
- [7] Verschaffel, L., Greer, B., & De Corte, E. (2000). Making sense of word problems. Lisse, The Netherlands : Swets & Zeitlinger.
- [8] Zan, R. (2016). I problemi di matematica. Difficoltà di comprensione e formulazione del testo. Roma: Carocci.
- [9] Ferreri, S. (1998). Il problema di matematica: un problema linguistico. (Guerriero, Ed.) *L'educazione linguistica e i linguaggi delle scienze*, 317-329.
- [10] Demartini, S., & Sbaragli, S. (2019). Le parole che "ingannano". La componente lessicale nell'insegnamento e nell'apprendimento della matematica. *Quaderni di Ricerca in Didattica (Mathematics)*, n.2 Numero speciale n.5, pp. 19-25.
- [11] Ferrari, P. L. (2004). Matematica e linguaggio: quadro teorico e idee per la didattica. Pitagora.
- [12] Bernardini, S. (2004). Corpora in the classroom. In J. Sinclair, *How to Use Corpora in Language Teaching* (pp. 15-36). Amsterdam: John Benjamins Publishing Company.
- [13] Corino, E., Fissore, C., & Marchisio, M. (2022). Data Driven Learning activities within a Digital Learning Environment to study the specialized language of Mathematics. *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*. 167-176. IEEE.
- [14] Fissore, C. (2022). Data Driven Learning activities in English on the specialized language of mathematics in secondary school. In *EDULEARN22 Proceedings, IATED*, 9326-9333.
- [15] Barbera, E. F. (2013). *Linguistica dei corpora e linguistica dei corpora italiana. Un'introduzione*. Milano: Qu. ASAR.
- [16] Barana A., Fissore C., Marchisio M., Rabellino S., Roman F. (2019). Comunità di Moodle per incentivare la collaborazione nelle attività di problem solving. *Atti di MoodleMoot2019*, 39-53.
- [17] Barana A., Boetti G., Marchisio, M. (2022) Self-Assessment in the Development of Mathematical Problem-Solving Skills. *Education Sciences* 12(81).
- [18] Fissore, C., Floris, F., Marchisio, M., & Rabellino, S. (2023). Learning analytics to monitor and predict student learning processes in problem solving activities during an online training. *IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*. 481-489. IEEE. doi:10.1109/COMPSAC57700.2023.00070
- [19] Floris, F., Fradiante, V., Marchisio Conte, M., & Rabellino, S. (2024). Strategie di gamification con Moodle per lo sviluppo di competenze di problem solving e per uno sviluppo sostenibile. *BRICKS*, 2, 25-35.
- [20] Dunn, J. (2021). Representations of language varieties are reliable given corpus similarity measures. *arXiv preprint arXiv:2104.01294*.
- [21] Corino, E., Fissore, C., Marchisio Conte, M. Training disciplinary teachers: the specialised language of Mathematics, in press.