

This is the peer reviewed version of the following article:

Linguaggi disciplinari, analisi dei corpora e verifica del vocabolario/Subject lexicons, corpora analyses and vocabulary assessment / Zini, Andrea. - (2015), pp. 49-60. (Intervento presentato al convegno am-Learning Individualizzazione del messaggio di apprendimento in ambiente adattivo tenutosi a Laboratorio di Pedagogia Sperimentale, Università Roma Tre nel 6 giugno 2014).

Franco Angeli  
*Terms of use:*

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

23/03/2025 09:49

(Article begins on next page)

In Poce, A., (ed.), *Individualizzazione del messaggio di apprendimento in ambiente adattivo / Individualisation of the learning message in adaptive environments*, Milano: FrancoAngeli, IT 49 – 60 / EN 47 – 56, ISBN 978-88-917-0994-3

## ***Linguaggi disciplinari, analisi dei corpora e verifica del vocabolario***

di Andrea Zini<sup>1</sup>

### **Introduzione**

Il progetto Firb “Adaptive-Message Learning” (2009-2014) si è posto l’obiettivo scientifico primario di sviluppare un modello didattico di individualizzazione che consiste nell’adattamento del «messaggio di istruzione» alla «competenza verbale» di chi lo riceve (Vertecchi, 2010).

Secondo Chris Whetton (2011, 33) il costrutto generale alla base del progetto può essere formulato in modo molto semplice. Ogni studente ha un suo personale lessico, che può essere misurato. Questa informazione può essere utilizzata per costruire materiali di apprendimento individualizzati, tali che l’apprendimento per quella persona sia sostenuto e potenziato. Per il gruppo nel suo insieme, in termini di misurazione «questo dovrebbe comportare una riduzione della varianza dei punteggi e un incremento generale delle medie dei punteggi». All’interno del modello am-learning la valutazione ha tempi, contenuti ed obiettivi diversi che Whetton (2011, 32) riconduce a due elementi fondamentali: «la misurazione iniziale delle abilità del discente di comprendere, e la misurazione finale dei risultati dell’apprendimento, misurati alla fine del corso o del modulo didattico. Per la prima parte, il progetto ha concepito una misura del lessico personale che utilizza i cloze test per valutare la disponibilità di vocabolario rilevante».

La valutazione iniziale e continua della competenza lessicale degli studenti, affidata al software LexMeter (Agrusti, 2010), è un elemento centrale nel modello “am-learning”. Essa richiede da un lato l’analisi lessicale dell’input al quale gli studenti sono esposti in un determinato contesto di apprendimento formale, dall’altro la rilevazione di misure sul lettore in relazione al vocabolario dei testi e al linguaggio specifico di un settore disciplinare. Più in dettaglio, prove di cloze di tipo “mirato” con risposte a scelta multipla generate dal software sono utilizzate per misurare fino a che punto gli studenti sappiano attualmente maneggiare il vocabolario rilevante in testi di una specifica disciplina, al fine di stimare quanto efficacemente essi potrebbero affrontare i materiali standard del corso e quindi utilizzare questa informazione per adattare la formulazione del testo di studio per il singolo studente.

La validazione della valutazione diagnostica di am-learning ha richiesto particolare attenzione. In am-learning, l’analisi quantitativa delle parole in un corpus di testi di un linguaggio specialistico viene utilizzata per identificare il vocabolario rilevante, cioè per definire con criteri statistici l’elenco delle parole da prendere in esame (quali parole possono essere cancellate da un dato testo). La verifica del vocabolario individuale, quindi la costruzione e la valutazione informatizzata delle prove, sono fondate sulla frequenza di occorrenze delle parole lessicali in un linguaggio specifico. In estrema sintesi, il programma:

- produce un lessico di frequenza specifico attraverso l’analisi lessico-statistica di un campione rappresentativo del linguaggio del settore disciplinare;
- dato un testo (analogo alle letture curricolari in programma), costruisce, somministra e corregge prove di cloze di tipo mirato con risposte a scelta multipla. Le parole grammaticali sono escluse

---

<sup>1</sup> Andrea Zini, Unità di Ricerca Università degli Studi di Modena e Reggio Emilia, andreazini3@gmail.com

dall'esame. Una volta stabilito quali classi grammaticali di parole siano eleggibili per la cancellazione (nel caso delle sperimentazioni condotte: nome e aggettivo), le parole nascoste devono ricadere all'interno di un intervallo di frequenze determinato dall'autore della prova. Si fornisce all'esaminato un testo che presenta un determinato numero di lacune, accompagnato da una lista di alternative tra le quali egli deve scegliere quella da reintegrare. La lista comprende tutte le parole del testo originale nascoste più altre eventualmente aggiunte dall'autore della prova.

Le indagini condotte dal gruppo di ricerca nazionale hanno consentito la sperimentazione e la messa a punto di questo modello all'interno dell'ambiente OrbisDictus.

Inoltre, è possibile fondare un sistema di valutazione automatico su aspetti diversi del vocabolario, purché possano essere descritti con riferimento a dati espliciti e quantificabili del linguaggio. In uno degli studi esplorativi condotti a Reggio Emilia la valutazione è riferita alla stratificazione del vocabolario di un corpus specifico non in termini di frequenza ma di peculiarità dell'uso.

### **Sommario delle prove sul campo**

1. La fase preliminare del lavoro è stata dedicata alla validazione dello strumento valutativo automatico ed in particolare:

- al collaudo di prove di cloze generate automaticamente dal programma LexMeter;
- all'esame della difficoltà di diverse tipologie di quesiti;
- alla verifica della validità della frequenza d'uso delle parole nascoste come criterio per differenziare la difficoltà delle prove.

Le prove somministrate all'Università di Modena e Reggio Emilia hanno coinvolto

- 28 studenti iscritti al CdL in Fisioterapia e 120 studenti del CdL in Infermieristica (I anno);
- un gruppo di 100 iscritti al II anno del CdL in Infermieristica.

La frequenza d'uso delle parole risulta sino dai primi studi ottocenteschi un importante criterio quantitativo utile a predire la leggibilità dei testi scritti per una determinata categoria di lettori (Amizzoni, Mastidoro, 1993). L'aspetto della frequenza d'uso è considerato influente sia sulle prestazioni recettive (decodifica e comprensione) e sia produttive (Colombo, 1993), in quanto ha effetto sulla forza e sulla disponibilità della rappresentazione di una parola nel lessico mentale. Confermando questi assunti, la frequenza d'uso delle parole nascoste ha dimostrato di essere un criterio efficace per predeterminare la difficoltà del test.

2. In seguito alla definizione dell'ambiente OrbisDictus, ciascuna unità di ricerca ne ha condotto un collaudo per verificare la funzionalità del sistema ed avviare la messa a punto di soluzioni tecniche.

Il tryout a Reggio Emilia è stato ospitato dal CdL in Fisioterapia (Facoltà di Medicina e Chirurgia) all'interno del Corso integrato di Biomeccanica e Cinesiologia applicata ed ha coinvolto una coorte di 30 studenti (I anno). Di questi, 26 hanno completato tutte le valutazioni previste: due prove di cloze mirato (entrambe somministrate all'ingresso e all'uscita) e due test di profitto relativi al contenuto delle due unità didattiche che compongono il corso. I cloze test, costruiti su due campioni di circa 750 parole di un manuale di Cinesiologia, contenevano 15 lacune e offrivano 15 alternative di risposta. I due passaggi apparivano equivalenti in termini di profilo lessicale e in termini di leggibilità (confronto con il Vocabolario di base De Mauro, 1980; indice di facilità di lettura Gulpease). Un raffronto delle parole nascoste nelle due prove con il Dizionario italiano De Mauro (2000) ha dimostrato che per quasi tutte è possibile trovare una definizione specifica nel campo delle scienze mediche (principalmente anatomia e fisiologia).

In termini di misurazione, il cloze test dove le parole nascoste sono state scelte all'interno delle fasce di frequenze medie e basse (esclusi gli hapax) in un vocabolario di frequenza estratto da un corpus

specifico di Medicina riabilitativa, si è rivelato allo stesso tempo, più omogeneo, difficile e discriminativo dell'altro, in cui la fascia di basse frequenze era esclusa dall'esame (maggiori dettagli si trovano in Zini, 2011).

3. Una diversa esperienza sul campo ha coinvolto un gruppo di 74 studenti del secondo anno del Corso di laurea in Scienze dell'Educazione (Facoltà di Scienze della Formazione).

La variante di prova d'ingresso che abbiamo studiato in questa occasione è stata concepita come un saggio di lettura di una specifica monografia (considerata come mini-corpus), indicata nella bibliografia del corso (Cecconi, 2002). I due campioni testuali (di circa 600 parole ciascuno) selezionati per produrre i cloze sono stati estratti dal primo capitolo del medesimo testo.

Rispetto all'esperienza precedente, la differenza più importante risiede nel diverso grado di formalizzazione dei linguaggi usati per scrivere i testi: quelli sulla ricerca educativa sono scritti in un registro accademico dell'italiano standard, mentre in quelli di cinesiologia si utilizza prevalentemente il linguaggio speciale della medicina riabilitativa.

Le parole bersaglio sono state scelte dentro la lista delle duecento parole piene che ricorrono con più frequenza nel libro. La lista comprende il 3% delle forme grafiche attestate (200 su circa 7300), alle quali corrisponde il 21% circa delle occorrenze totali nel mini-corpus. Quasi nessuna delle 200 forme grafiche è originale di un vocabolario tecnico-specialistico (uno dei rari esempi è "postulato"). Consultando il Dizionario italiano De Mauro (2000), si trova per la metà delle parole nascoste un'accezione tecnico-scientifica pertinente all'uso che della parola si fa nel testo considerato. Raramente questa accezione ricade nell'uso specifico di scienze sociali ed umane o della filosofia.

Il punto critico dei cloze si è dimostrato essere la scarsa coerenza interna. Nella discussione dei risultati questo fatto è stato collegato al grado di ambiguità del linguaggio del testo e in particolare delle parole-bersaglio.

La somministrazione del cloze affiancato ad altre prove che misurano costrutti noti (padronanza della terminologia; determinati aspetti della comprensione del testo) ha permesso di osservare una significativa correlazione per ranghi ( $\rho$  di Spearman) tra i risultati ottenuti dagli stessi soggetti al cloze e alla prova di vocabolario specialistico.

Inoltre, ai 15 item che compongono i cloze test se ne è aggiunto (e considerato a parte) uno "erratico", ottenuto cancellando l'avverbio connettivo "Di conseguenza" che ricorre in entrambi i testi e collega in modo coerente due periodi sintatticamente indipendenti. È una richiesta di completamento che intende sollecitare direttamente la competenza testuale del lettore. Per questa lacuna sono offerte quattro alternative di completamento specifiche. Oltre a quella corretta, compaiono i distrattori "Tuttavia", "Allo stesso modo", "Invece". Gli esiti delle prove suggeriscono che questo tipo di item, che ha una natura diversa dagli altri e con ogni evidenza non contribuisce alla stima del lessico disponibile, può tuttavia dimostrare un alto indice di discriminatività all'interno del cloze-test e una associazione significativa con le risposte a quesiti di comprensione del testo che richiedono di produrre inferenze connettive. Questa prima esperienza ha aperto la strada ad ulteriori sviluppi: come abbiamo sperimentato anche in seguito, esiste, per quanto appaia limitata, la possibilità di costruire questo tipo di item a partire da dati espliciti del linguaggio, quindi rispettando i vincoli propri di un trattamento automatizzato.

4. L'attività dell'Unità di ricerca dell'Università di Modena e Reggio Emilia si è ulteriormente concentrata sui contenuti e sugli strumenti della valutazione iniziale. Un nuovo esame della letteratura ha accompagnato l'analisi dei risultati parziali ottenuti, consentendo di disegnare un nuovo esperimento mirato alla verifica del vocabolario peculiare di un linguaggio disciplinare nel contesto dei materiali di studio. La successiva prova sul campo ha coinvolto una nuova coorte di 30 studenti del primo anno del

Corso di laurea in Fisioterapia. Nei prossimi paragrafi discuteremo di alcuni aspetti di questo lavoro (per maggiori dettagli, si veda Zini, 2012, 2014).

### **Punti salienti dell'esame della letteratura**

La letteratura attesta il ruolo cruciale della conoscenza del vocabolario nella comprensione della lettura, benché questo ruolo non sia da intendersi ingenuamente.

Infatti, la conoscenza del vocabolario e la padronanza delle strutture morfosintattiche della lingua sono immediatamente richieste a chi voglia comprendere anche un singolo enunciato, ma non assicurano necessariamente la comprensione dei contenuti (espressi o impliciti) di un testo, inteso come discorso coeso e coerente, che implica anche altri processi. La stessa corretta attribuzione di significato alle parole, sia note e sia nuove, incontrate nella lettura di un testo, è normalmente influenzata o permessa dal contesto e dal cotesto nel quale sono collocate. Di conseguenza, “per comprendere un testo è necessario non solo possedere un ampio patrimonio lessicale, ma soprattutto capire che significato assumono le parole quando si usano, come si modifica il loro significato in un co-testo” (Salerni, 1989).

La competenza linguistica di un individuo e l'ampiezza del suo vocabolario sono strettamente associate ai processi d'apprendimento. Va fatta attenzione che la “conoscenza del significato” può avere varie accezioni, e può essere valutata in modo diverso. Per valutare il vocabolario si possono infatti usare diverse procedure (Passolunghi, De Beni, 2001).

La differenza più evidente separa una verifica esplicita della conoscenza del lessico, che si ottiene “utilizzando un test formato da item a punti discreti, ciascuno dei quali va a testare un singolo componente lessicale”, da una verifica implicita, “come in un test integrato, che presuppone la conoscenza di un particolare insieme di elementi lessicali al fine di portare a termine il compito che è stato richiesto” (Davies, 2008).

Chiedendo il significato di una parola isolata, otteniamo informazioni solo sull'estensione del patrimonio lessicale del soggetto, mentre rimaniamo all'oscuro circa il grado di organizzazione del suo sistema lessicale, cioè della capacità del soggetto di padroneggiare le regole di combinazione delle parole (Boschi, Aprile, Scibetta, 1992).

La revisione della letteratura ha confermato che la procedura cloze di tipo mirato è un formato di prova idoneo a integrare le misure di estensione del vocabolario con misure più contestualizzate all'uso del vocabolario, che tengano conto della profondità di conoscenza del vocabolario in contesti d'uso specifici, come ad esempio le discipline accademiche (Read, 2000, 2004; Salerni, 1989).

Come ha affermato Salerni (1989):

Certamente l'uso di un test è in funzione dell'obiettivo che si vuole verificare. Se infatti lo scopo è misurare l'abilità lessicale degli studenti in relazione al processo totale di lettura, sembra essenziale che il test misuri la conoscenza del lessico nel contesto. Il test dovrebbe misurare fino a quando uno studente è in grado di utilizzare le indicazioni sintattiche, semantiche e morfologiche per ricavare il significato specifico di una parola così come essa è usata in un particolare contesto. A questo scopo le prove di cloze mirato o di buco potranno dare un'utile indicazione.

Secondo la nostra analisi del compito, le risposte a quesiti di cloze mirato con risposte a scelta multipla possono dar conto di tre abilità operative, la seconda delle quali si applica alla lettura del testo lacunoso, le altre due alla lettura dell'elenco delle parole offerte per colmare le lacune:

a) ricordare il significato di parole note (cioè delle quali sia già noto il significato pertinente);

- b) inferire il significato che le parole parzialmente note o sconosciute assumono in base al contesto;
- c) anticipare correttamente il significato delle singole parole mancanti utilizzando gli indizi presenti nel testo e le proprie conoscenze pregresse (generali e specifiche).

Le tre abilità operative potrebbero essere ricondotte ad altrettante componenti del processo di comprensione del testo: vocabolario, inferenza lessicale e inferenza semantica (come definite in De Beni, R., Cornoldi, C., gruppo MT, 2004).

In Read (2007) si trovano riferimenti al contributo attuale e potenziale dell'analisi dei corpora per estendere la nostra comprensione delle caratteristiche lessicali dei registri accademici e per l'applicazione di liste di parole nella valutazione delle abilità lessicali dei discenti: "Per la verifica del vocabolario i corpora possono fornire la base per liste di parole più precise dalle quali le parole-bersaglio possono essere campionate, tenendo conto della frequenza, del rango e di altri criteri" (Read, 2007: 106).

Abbiamo bisogno di completare i test di vocabolario per punti discreti con misure integrate dell'abilità degli studenti di maneggiare elementi lessicali in contesto. Tradizionalmente, il contesto è stato concepito, in termini linguistici, come una frase od un più largo co-testo in cui un vocabolo compare. L'analisi dei corpora ora ci fornisce potenti strumenti per arricchire la nostra comprensione del contesto, fornendo descrizioni dettagliate dell'uso del vocabolario in registri specifici o generi disciplinari. Le applicazioni di queste ricche descrizioni alla valutazione del vocabolario non sono ancora del tutto chiare, ma rappresentano un terreno fertile per ulteriori indagini.

Fra le esperienze condotte da Read, quella denominata "Words in Context test" (2004) ha particolare attinenza con l'oggetto di questo studio anche perchè il formato adottato è la prova di cloze di tipo mirato con risposte a scelta multipla. Nei termini del modello di Read (2000), si tratta dell'esplorazione di un tipo di prova discrete (che misura la conoscenza e l'uso del vocabolario come costruito indipendente), selective (mirata ad un vocabolario specifico: le parole che appartengono alla Academic Word List) e al tempo stesso context-dependent.

La questione principale nella scrittura del test è come rendere i quesiti dipendenti dal contesto. L'intenzione iniziale era quella di avere, come opzioni per ciascun item, frasi che esprimessero significati possibili della parola bersaglio in modo che gli esaminandi avessero bisogno di elaborare le informazioni contestuali per poter scegliere l'opzione corretta. [...] Tuttavia, nel test nel suo complesso, si è rivelato difficile mantenere coerentemente questo approccio. Sembra che le parole AWL non siano così polisemiche come previsto, o almeno non fino al punto che si possano comporre tre frasi che rappresentino significati effettivi della parola che siano plausibili nel particolare contesto della parola bersaglio. [...] Per la maggior parte degli item, allora, ho invertito la logica in un certo senso. [...] le opzioni tendono a rappresentare significati che siano plausibili nel contesto e gli esaminandi hanno bisogno di avere qualche conoscenza specifica della parola bersaglio per poter scegliere la risposta corretta. (Read, 2004: 25, 27).

Questioni molto simili si sono presentate anche nella nostra esperienza, dove la disponibilità di alternative di risposta plausibili dipende direttamente della scelta di quali parole nascondere.

### **Parole di uso comune e specialistico**

Uno studio recente (Truffelli, 2010) sulla comprensione della lettura di testi accademici condotto su studenti in Scienze dell'Educazione all'Università di Bologna, descrive un quadro complesso delle diverse componenti del processo, dal quale risulta che la conoscenza del significato di specifici termini o sintagmi e la capacità di cercare e identificare informazioni specifiche contenute nel testo sono le più critiche per l'intero gruppo campione. "In particolare lo scarto maggiore tra intero campione e lettori

“poveri” si registra sia rispetto alla conoscenza dei termini, in coerenza con quanto riscontrato in letteratura [...], sia rispetto alla comprensione globale del testo”.

Assunta la distinzione tra termini e parole, «i termini dotati di significati determinati, le parole dotate di significati indeterminati e dilatabili nell'uso per estensioni e metafore», si può affermare che «in un campo di discorsi quanto maggiore è il ricorso a termini, tanto maggiore è la quota di tecnicità e scientificità» (De Mauro, 1988, 14). Ciò non comporta né che i termini siano «parole di fisionomia diversa dall'ordinario», né che il repertorio sia esteso, ma «comporta solo che molta parte dei discorsi deve essere costruita con termini ben definiti, anche tratti da un repertorio assai ristretto». L'incidenza dei termini all'interno dei testi varia nei diversi ambiti di discorsi.

Nel nostro studio, assumono una particolare rilevanza la precisazione e l'estrazione automatica di una categoria di termini da prendere in esame. Secondo Adriano Colombo (2002: 33), «Quando i termini tecnici sono stati creati appositamente, non costituiscono in genere difficoltà: saranno per lo più parole nuove per gli allievi, ma il loro (unico) significato sarà appreso insieme ai concetti disciplinari». Tuttavia molti termini, «lungi dall'essere conati ex-novo, non sono che parole della lingua comune che subiscono una rideterminazione e precisazione di significato nel momento in cui vengono assunte entro un dato lessico specialistico» (Lavinio, 2004: 101). Questo tipo di termini può offrire agli studenti una particolare difficoltà di comprensione, perché «le parole che diventano termini si portano inevitabilmente dietro l'alone dei significati che continuano a possedere nella lingua comune» (*Ibid.*) e «passare dalle accezioni comuni a quelle dei lessici specifici può non essere facile» (Colombo, 2002: 33).

### **Un cloze test sul linguaggio peculiare della Medicina Riabilitativa**

Nell'ultimo degli esperimenti condotti, le parole da nascondere nei cloze mirati sono state scelte all'interno di un particolare repertorio, cioè il lessico peculiare della medicina riabilitativa.

È stato allestito un corpus del linguaggio del settore della riabilitazione (Corpus di Medicina Riabilitativa, CMR) che rappresenta prevalentemente il linguaggio della medicina riabilitativa intesa come settore scientifico disciplinare accademico (una parte preponderante dei testi raccolti è costituita da manuali, lezioni e tesi di laurea). Il corpus, stratificato in base a tre variabili (autore, destinatario e tipologia testuale) che possono assumere diverse modalità, conta circa 1.100.000 occorrenze (token) di 64.000 parole (intese come forme grafiche diverse, type).

L'estrazione del linguaggio peculiare consiste nel «selezionare le parole da analizzare non in funzione della loro frequenza assoluta, bensì del loro sovra-sotto uso rispetto all'uso medio in un linguaggio di riferimento» (Bolasco, 2008). A tale scopo si è fatto ricorso non solo a risorse endogene, ma anche a risorse esogene al corpus specifico. La risorsa statistico-linguistica utilizzata come modello di riferimento per estrarre informazione utile dal corpus specifico è la lista di frequenza denominata italiano standard (Bolasco, 2008).

Confrontando il vocabolario del corpus CMR con il lessico di frequenza dell'italiano standard POLIF2002 contenuto nel database di TalTac2, si è definito l'insieme delle forme grafiche che sono rappresentate sia nell'uso standard dell'italiano, sia nell'uso della medicina riabilitativa (intersezione) e che all'interno di quest'ultimo sono sovra-rappresentate (con uno scarto sulle occorrenze di segno positivo e di valore significativo).

La lista definitiva delle forme peculiari comprende 1230 elementi, pari al 2% circa delle forme grafiche diverse (type) presenti nel CMR, alle quali corrisponde il 16% circa delle occorrenze (token) totali nello stesso corpus.

Quali parole appartengono a questa lista e quali invece ne sono escluse?

In primo luogo, rimangono escluse le parole “banali”, che presentano una frequenza di occorrenze equivalente in CMR e in Polif. Inoltre, sono escluse tutte le parole “originali” del corpus di medicina riabilitativa, cioè tutte quelle forme grafiche non attestate nel lessico dell’italiano standard Polif.

Quelle incluse sono in molti casi termini il cui significato tecnico è legato alla disciplina, in altri casi vocaboli accademici oppure comuni (se non fondamentali), che presentano un legame semantico più lasso col campo di discorsi specifico. Le parole che ricadono nel repertorio considerato sembrano quindi abbastanza differenziate da consentire una rilevazione della competenza lessicale che sia al tempo stesso estesa e specifica.

## **Risultati**

Nelle prove d'ingresso, che in questo caso coincidono con l'oggetto dello studio e si sono svolte in presenza, ai due cloze test di tipo mirato si affiancano prove di altro tipo che riguardano la comprensione del testo, il vocabolario tecnico e le conoscenze pregresse.

Per capire come hanno funzionato i cloze-test sul gruppo di studenti che ha completato tutte le prove (28 soggetti su 30), utilizziamo alcuni strumenti dell'item analisi classica. Gli item che compongono i cloze test mirati coprono livelli di difficoltà ben differenziati. Nel primo test, tre item su quindici risultano critici per eccessiva facilità e scarsa discriminatività, mentre nel secondo test si nota un solo item critico su venti. L'affidabilità delle prove di cloze misurata in termini di correlazione fra componenti parallele, definite come item pari e item dispari, corretta con la formula di Spearman-Brown, assume nel caso del primo test il valore 0,72 e nel caso del secondo test il valore 0,82.

Un esame della relazione fra i punteggi ai cloze test e quelli alle altre prove condotto utilizzando l'indice di correlazione per ranghi di Spearman ( $\rho$ ) permette di osservare una correlazione significativa ( $r_s=0.52$ ,  $p 0.01$ ) tra i punteggi ottenuti dagli stessi soggetti al primo cloze mirato e alla prova di lessico specialistico (prova di definizione di 22 termini fuori contesto). Nel caso del secondo cloze test si trova una relazione positiva, più debole ( $r_s=0.38$ ,  $p 0.05$ ), coi risultati alla prova di valutazione delle conoscenze disciplinari pregresse.

## **Conclusioni**

La relazione riscontrata con misurazioni concomitanti dell'estensione del patrimonio terminologico dei lettori e con prove che riguardano l'apprendimento dei concetti disciplinari spiega in parte la natura di queste prove di cloze mirato come prove di vocabolario specifico e in parte come prove di conoscenza. Nel contesto del progetto “am-learning” e rispetto alla finalità che ha al suo interno l'attività valutativa di cui ci si è occupati, lo strumento risulta adeguato.

L'esperienza condotta suggerisce anche la possibilità di arricchire di ulteriori strumenti un apparato valutativo informatizzato che utilizzi sia indicatori lessicali, sia linguistici in senso lato, per intervenire su diverse difficoltà di comprensione. In particolare, si è esplorata la possibilità di istruire una procedura automatica per generare quesiti a completamento cancellando alcuni connettivi da un testo, in modo sollecitare la sensibilità alla coesione e alla coerenza del testo.

## **Riferimenti**

- Agrusti, F. (2010). From LexMeter to Adapter. Towards a match up between the Virtual and the Real Reader, *Cadmo. Giornale Italiano di Pedagogia sperimentale. An International Journal of Educational Research*, 1, 97-108.
- Amizzoni, M, Mastidoro, N. (1993). Linguistica applicata alla leggibilità: considerazioni teoriche e applicazioni, *Bollettino della Società Filosofica Italiana*, 149, 49 – 63.



- Bolasco, S. (2008). *Corpora e liste di frequenza d'uso: criteri e tecniche per l'analisi automatica dei testi*, in Barni, M., Troncarelli, D., Bagna C. (eds.), *Lessico e apprendimenti*, Milano: Franco Angeli: 113 – 142.
- Boschi, F., Aprile, L., Scibetta, I. (1992). *Le parole e la mente. Studi e ricerche sullo sviluppo della competenza lessicale nel ragazzo*, Firenze: Giunti, 69.
- Cecconi, L. (ed.) (2002). *La ricerca qualitativa in educazione*, Milano: Franco Angeli.
- Colombo, A. (2002). *Leggere. Capire e non capire*, Zanichelli: Bologna.
- Colombo, L. (1993), *Locus o loci dell'effetto frequenza?*, in Laudanna, A. e Burani, C. (eds.), *Il Lessico: processi e rappresentazioni*, Roma: NIS.
- Davies, A. (2008). *Che cosa ci dicono i test lessicali riguardo alla padronanza linguistica?*, in Barni, M., Troncarelli, D., Bagna C. (eds.), *Lessico e apprendimenti*, Milano: Franco Angeli: 104.
- De Beni, R., Cornoldi, C., gruppo MT (2004). *Nuova guida alla comprensione del testo*. Trento: Erickson.
- De Mauro, T. (1980). *Guida all'uso delle parole*, Roma: Editori Riuniti.
- De Mauro, T. (1988). *Linguaggi scientifici e lingue storiche*, in Guerriero, A.R. (ed.), *L'educazione linguistica e i linguaggi delle scienze*, Firenze: La Nuova Italia, 9 - 19.
- De Mauro, T. (2000). *Dizionario italiano*, Milano: Paravia.
- Lavinio, C. (2004). *Comunicazione e linguaggi disciplinari. Per un'educazione linguistica trasversale*, Carocci: Roma.
- Passolunghi, M.C., De Beni, R. (2001). *I test per la scuola*, Bologna: Il Mulino: 95.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. (2004). *Second Language Vocabulary Testing: Taking a Broader Perspective*, 2004 *International Conference on English Instruction and Assessment*.
- Read, J. (2007). *Second Language Vocabulary Assessment: Current Practices and New Directions*, *International Journal of English Studies IJES*, 2, 105 - 125.
- Salerni, A. (1989). *La verifica della competenza lessicale*, in Lucisano, P., Salerni, A., Benvenuto, G., Siniscalco, M.T. (a cura di), *Lettura e comprensione*, Torino: Loescher, 117 - 151.
- Truffelli, E. (2010). *Comprendere per migliorare lo studio: analisi e riflessioni a partire da un'esperienza biennale di sostegno alle matricole universitarie*, *Giornale italiano della ricerca educativa*, giugno 2010, pp. 91 - 104.
- Vertecchi, B. (2010). *Nuove ipotesi per lo sviluppo della didattica in linea*, *Journal of e-Learning and Knowledge Society*, 6(1), 31 – 42.
- Whetton in Poce, A., Angelini, C. (2011). *Per una nuova cultura didattica. Riflessioni sul progetto am-learning*, FrancoAngeli: Milano.
- Zini, A. (2011). *Analisi del lessico e individualizzazione del messaggio di apprendimento. La preparazione di un corso nell'ambito del progetto "adaptive message learning"*, in Minerva T., Colazzo L. (eds.), *Connessi! Scenari di Innovazione nella Formazione e nella Comunicazione. VIII Congresso Nazionale della Società Italiana di e-Learning*, Milano: LediPublishing, 961-970.
- Zini, A. (2012). *Misurare la competenza lessicale in contesto specifico attraverso prove di cloze*, *Giornale Italiano della Ricerca Educativa - Italian Journal of Educational Research*, 9, 108 - 119.
- Zini, A. (2014). *Una prova di cloze sul lessico peculiare di un linguaggio scientifico*, in Colombo, A., Pallotti, G. (eds.), *L'italiano per capire, "I Quaderni del GISCEL"*, Roma: Aracne, 223 - 239.