



UNIVERSITA' DEGLI STUDI DI PAVIA  
DIPARTIMENTO DI SCIENZE DEL SISTEMA NERVOSO E DEL  
COMPORTAMENTO

CORSO DI LAUREA MAGISTRALE IN PSICOLOGIA

DIFFERENZE INDIVIDUALI  
NELLA DECISIONE LESSICALE  
LUNGO LA DISTRIBUZIONE  
DELLE ABILITÀ DI LETTURA

RELATORE:  
Prof.ssa Mascheretti Sara

CORRELATORE:  
Dr. Gatti Daniele

Tesi di Laurea di  
Loredana Tonali  
522498

Anno Accademico 2023/2024

*“Leggere è andare incontro  
a qualcosa che sta per essere  
e ancora nessuno sa cosa sarà.”  
(Italo Calvino)*

# Indice

Abstract .....	5
Introduzione .....	7
<b>Capitolo 1: Il Linguaggio.....</b>	<b>14</b>
1.1. Che cos'è il linguaggio .....	14
1.2. Definire il linguaggio .....	18
1.3. Le componenti del linguaggio.....	21
1.4. Linguaggio scritto e capacità di lettura.....	24
<b>Capitolo 2: La capacità di lettura .....</b>	<b>28</b>
2.1. I modelli di lettura .....	28
2.1.1. Modello di lettura a due vie .....	28
2.1.2. Modelli semantici distribuzionali .....	32
2.2. Le variabili che influenzano la capacità di lettura.....	37
2.2.1. Variabili base.....	38
2.2.2. Variabili semantiche .....	39
2.2.3. Variabili lessicali .....	40
<b>Capitolo 3: La Dislessia .....</b>	<b>43</b>
3.1. I Disturbi Specifici dell'Apprendimento: DSA.....	43
3.2. Dislessia Evolutiva .....	49
3.2.1. Inquadramento diagnostico e incidenza.....	49
3.2.2. Sottotipi di Dislessia evolutiva.....	51
3.3. La Dislessia nell'adulto e la compensazione delle difficoltà di lettura .....	55

Capitolo 4: Lo studio .....	58
4.1. Il progetto.....	58
4.2. Metodi.....	59
4.2.1. Partecipanti .....	59
4.2.2. Strumenti.....	59
4.2.3. Procedura .....	66
4.3. Analisi dei dati e risultati .....	67
 Capitolo 5: Discussione .....	 71
5.1. Discussione .....	71
5.1.1. Densità di vicinato semantico .....	72
5.1.2. Lunghezza.....	73
5.1.3. Frequenza.....	75
5.1.4. Differenze individuali.....	77
5.2. Limiti e punti di forza.....	79
5.3. Riflessioni personali e prospettive future .....	80
 Conclusione .....	 82
 Bibliografia .....	 85

## Abstract

La lettura può essere definita come il processo che consente l'elaborazione e la comprensione del linguaggio scritto; generalmente viene acquisita durante i primi due anni di scuola primaria. È opportuno distinguere tra: difficoltà di apprendimento (vari gradi di difficoltà che si possono manifestare in ambito scolastico) e compromissione della capacità di lettura, che fa riferimento ad una precisa categoria diagnostica che prende il nome di Dislessia Evolutiva (DE). La DE è un disturbo del neurosviluppo, appartenente all'area clinica dei Disturbi Specifici dell'Apprendimento, caratterizzato da difficoltà nell'acquisizione e automatizzazione dei processi di lettura, in presenza di un adeguato funzionamento neurologico/sensoriale, opportunità educative e capacità cognitive nella media. Le prove di decisione lessicale rappresentano un valido strumento che può essere impiegato ai fini della valutazione delle capacità di lettura. Nella presente tesi viene presentato il progetto di ricerca condotto presso l'Università di Pavia, finalizzato ad indagare se differenti componenti semantiche ed ortografiche siano in grado di predire le performance di lettura in un campione di partecipanti (N=53, range di età: 19-29 anni) con diverse competenze di lettura, tramite l'utilizzo di un compito di decisione lessicale. L'ipotesi di partenza presume la presenza di una maggiore sensibilità ai fattori semantici nei partecipanti con peggiori abilità di lettura. Per l'analisi dei dati è stato utilizzato un Modello Semantico Distribuzionale. I risultati dimostrano la presenza di una correlazione negativa tra velocità di lettura (tempo di lettura espresso in secondi;  $b=-.02$ ,  $p\text{-value}=.42$ ) e densità di vicinato semantico ( $b=-.29$ ,  $p\text{-value}<.001$ ) e frequenza ( $b=-.03$ ,  $p\text{-value}<.001$ ), e di una correlazione positiva tra velocità di lettura e la variabile lunghezza ( $b=.04$ ,  $p\text{-value}<.001$ ). In definitiva si osserva che gli individui tendono a leggere più rapidamente parole con un elevato numero di vicini semantici ed elevata frequenza, mentre la velocità diminuisce in relazione all'aumento della lunghezza della parola stessa.

**Parole chiave:** capacità di lettura, Dislessia Evolutiva, decisione lessicale, componenti semantiche, modelli semantici distribuzionali.

## Abstract

Reading can be defined as the process that allows the processing and comprehension of written language; generally acquired during the first two years of elementary school. It's necessary to differentiate between: learning disabilities (various degrees of difficulty that can manifest in the school context) and reading impairment, which refers to a precise diagnostic category named Developmental Dyslexia (ED). ED is a neurodevelopmental disorder, belonging to the clinical area of Specific Learning Disorders, characterized by difficulties in the acquisition and automation of reading processes, in the presence of adequate neurological/sensory functioning, educational opportunities and average cognitive abilities. Lexical decision tasks are a valuable tool that can be used to assessing reading skills. This thesis presents the research project conducted at the University of Pavia, aimed at investigating if semantic and ortographic components are able to predict the performance of a sample of participants (N=53, age range: 19-29 years) with different reading skills, in a lexical decision task. The starting hypothesis assumes the presence of higher sensitivity to semantic factors in participants with worse reading skills. A Distributional Semantic Model was used to analyze the data. The results demonstrate the presence of a negative correlation between reading (speed-reading time expressed in seconds;  $b=-.02$ ,  $p\text{-value}=.42$ ) and semantic neighborhood density ( $b=-.29$ ,  $p\text{-value}=<.001$ ) and frequency ( $b=-.03$ ,  $p\text{-value}=<.001$ ), and a positive correlation between reading speed and the length variable ( $b=.04$ ,  $p\text{-value}=<.001$ ). Ultimately, it is observed that individuals tend to read words with a high number of semantic neighbors and high frequency more quickly, while speed decreases with increasing word length.

**Keywords:** reading ability, Developmental Dyslexia, lexical decision making, semantic components, Distributional Semantic Models.

# Introduzione

La lettura può essere definita come il processo che consente l'elaborazione e la comprensione del linguaggio scritto. Le informazioni testuali che il lettore decodifica implicano la capacità di riconoscere la sequenza di lettere/parole e individuare le relazioni esistenti tra esse (Rayner, Schotter, Masson, Potter & Treiman, 2016).

Il processo di lettura comprende una serie di aspetti fondamentali, tra i quali: consapevolezza fonologica, vocabolario e comprensione. La consapevolezza fonologica fa riferimento alla capacità di elaborare e manipolare in modo consapevole sequenze fonologiche. Il concetto di vocabolario comprende la conoscenza delle parole e del loro significato. La comprensione riguarda la capacità dell'individuo di acquisire il significato di ciò che sta leggendo; è un aspetto che coinvolge le conoscenze di base del lettore, le informazioni testuali lette e il contesto socioculturale in cui avviene la lettura (Birsh, 2018, Honig et al., 2018, Snow, 2010).

La lettura e la scrittura costituiscono il prodotto di un'evoluzione culturale e sociale e per poter essere apprese dagli individui necessitano di istruzioni esplicite. Pinker la definisce come *“un accessorio opzionale che deve essere faticosamente montato”* (Pinker, 1997).

L'acquisizione della capacità lettura inizia generalmente intorno ai 6 anni di età e perdura per i primi due anni di scuola primaria. Una volta consolidati i meccanismi

alla base, il processo è automatizzato, richiedendo un minimo dispendio di risorse per essere attuato.

I bambini possono migliorare gradualmente la loro capacità di leggere, fino a diventare lettori più o meno esperti. Le differenze individuali nella capacità di lettura possono essere rappresentate mediante la curva di distribuzione normale, che racchiude al suo interno lettori più o meno abili. Ciò che determina la posizione che i singoli individui occupano lungo la curva è il livello corrispondente alla loro capacità di leggere accuratamente e rapidamente, che può essere influenzato da aspetti genetici e ambientali (McGowan et al., 2019).

Uno studio condotto da Logan e colleghi ha esaminato l'impatto che ereditabilità e ambiente condiviso possono esercitare sulla distribuzione della capacità di lettura. I risultati hanno mostrato che la genetica esercita maggiore influenza sui soggetti con scarse abilità di lettura, mentre l'ambiente sembra impattare maggiormente sui lettori più abili (Logan et al., 2012).

Questi risultati sono stati confermati da uno studio più recente, condotto da McGowan e colleghi. Anche in questo caso sono emerse ridotte influenze ambientali condivise per i lettori con capacità inferiori, in relazione ad una maggiore influenza esercitata da aspetti genetici. Questo consente di ipotizzare che i bambini in età prescolare abbiano difficoltà a imparare a leggere nonostante la presenza di un adeguato supporto ambientale, a causa del forte impatto esercitato dalle influenze genetiche sulla capacità di lettura (McGowan et al., 2019).

Sulla base di queste premesse è opportuno precisare una distinzione tra la più generica espressione “difficoltà di apprendimento”, che riguarda vari gradi di difficoltà che si possono manifestare in ambito scolastico, e “compromissione della capacità di lettura”, che invece fa riferimento ad una specifica categoria diagnostica contraddistinta da criteri oggettivi e valutabili, che prende il nome di Dislessia Evolutiva (DE).

La DE rappresenta un disturbo del neurosviluppo caratterizzato da difficoltà nell’acquisizione e automatizzazione dei processi di lettura, in presenza di un adeguato funzionamento neurologico/sensoriale, opportunità educative e capacità cognitive nella media (American Psychiatric Association, 2013). Implica una ridotta velocità o fluidità nella lettura, difficoltà marcate nel processo di decodifica e compromissione della capacità di comprensione del testo. La DE appartiene alla vasta area dei Disturbi Specifici dell’Apprendimento (DSA), disturbi di origine neurobiologica che determinano una compromissione delle capacità scolastiche di lettura, scrittura e calcolo. I DSA presentano un andamento cronico, ma evolutivo: mutano in relazione all’età, all’esperienza dell’individuo e alle richieste dell’ambiente. I fattori ambientali (quali contesto familiare e sociale) svolgono dunque un ruolo fondamentale, esercitando la loro influenza in modo diretto e continuativo.

Trattandosi di un disturbo con insorgenza durante l’età scolare (diagnosticabile a partire dalla fine della classe seconda della scuola primaria), la letteratura scientifica di riferimento, si concentra prevalentemente sulla popolazione infantile;

gli studi e le ricerche dedicate alla popolazione adulta risultano essere di numerosità nettamente inferiore. Tuttavia non si tratta di un disturbo circoscritto all'infanzia, bensì coinvolge un'ampia fetta di popolazione costituita da adolescenti e giovani adulti, sia per quanto concerne l'elevato numero di soggetti diagnosticati tardivamente, sia per le implicazioni a lungo termine delle difficoltà caratterizzanti la diagnosi (Wilson et al., 2015). L'interesse abbastanza recente da parte della letteratura scientifica nei confronti della condizione di DSA nell'adulto ha contribuito a spiegarne le caratteristiche peculiari e confermano non solo la loro base costituzionale, ma soprattutto la persistenza di tali difficoltà che possono esercitare influenza sulle scelte e sul successo scolastico, accademico e lavorativo (Kemp, Parrila e Kirby, 2009).

La valutazione delle abilità di lettura in individui con DE avviene tradizionalmente mediante batterie standardizzate che indagano i parametri di rapidità e correttezza (Cornoldi, Colpo, 1995, 1998), permettendo di individuare le eventuali prestazioni deficitarie del soggetto. Recentemente, sono stati proposti nuovi parametri che possono essere utilizzati nel processo diagnostico, tra i quali le prove di decisione lessicale (Caldarola, Perini e Cornoldi, 2012). Si tratta di processi che prevedono la somministrazione, computerizzata o non, di prove decisionali, durante le quali il soggetto è chiamato ad effettuare una scelta dicotomica relativamente ad una stringa di lettere presentata, indicando se essa corrisponde ad una parola o meno. La letteratura scientifica offre un vasto repertorio di applicazione di prove di decisione lessicale in differenti campi della ricerca (psicologia

dell'apprendimento, psicolinguistica, neuropsicologia). Uno studio condotto da Palumbo (2008) conferma che l'utilizzo di prove di decisione lessicale costituisce un valido indicatore della capacità di riconoscimento delle parole, al fine di determinare la fluidità di lettura in giovani lettori e studenti universitari. I risultati ottenuti hanno evidenziato una correlazione significativa tra i punteggi conseguiti al compito di decisione lessicale e le prestazioni ottenute in un test di lettura, confermando l'efficacia della prova a livello metodologico.

Una ricerca spagnola condotta da Acha e Perea (2008) prevede l'impiego di una prova di decisione lessicale (lo spagnolo è una lingua a ortografia trasparente e molto regolare) in un campione di soggetti di differenti età (rispettivamente 7, 11 e 22 anni), al fine di individuare eventuali differenze a livello delle strategie di lettura utilizzate. I risultati ottenuti mostrano una progressiva evoluzione delle strategie impiegate dai lettori: i bambini prediligono la decodifica lettera per lettera, mentre con l'aumentare dell'età si assiste ad un accesso diretto lessicale. Questo conferma che l'utilizzo di compiti di decisione lessicale risulta efficace nell'individuare i cambiamenti nelle strategie di lettura utilizzate a partire dai lettori iniziali fino a quelli esperti.

Uno studio condotto Paizi et al. (2013) dimostra che i soggetti con DE sottoposti a compiti di decisione lessicale, subiscono una significativa influenza da parte delle variabili semantiche nei processi di lettura. Nello specifico i risultati dello studio hanno confermato un'elevata sensibilità alla lunghezza dello stimolo, sia nella lettura ad alta voce che nelle prove di decisione lessicale. La variabile

lunghezza costituisce un indicatore di DE in italiano (Zoccolotti et al., 1999) così come in altre ortografie trasparenti. In generale, è possibile affermare che i soggetti con DE ottengono performance peggiori in compiti di decisione lessicale rispetto a soggetti normo-lettori.

Altri fattori che influenzano i processi di lettura sono: la frequenza di utilizzo, la classe, l'età di acquisizione, l'immaginabilità, la concretezza ed infine le variabili semantiche relative al ruolo svolto dal vicinato ortografico. Le variabili semantiche vengono generalmente indagate mediante l'utilizzo di modelli distribuzionali, in grado di stimare il significato di una parola in base al suo utilizzo all'interno del linguaggio (Günther et al., 2019). In breve, tramite questi modelli è possibile estrarre una stima numerica di quanto l'utilizzo nel linguaggio di una determinata parola è simile a quello di una o più altre parole. Si tratta di modelli fondati sulla teoria definita "Ipotesi distribuzionale", elaborata negli anni '70 dal linguista statunitense Harris. Secondo l'autore due parole sono considerate semanticamente simili in base a quanto entrambe ricorrono in contesti linguistici simili. I significati delle parole derivano dunque dalla loro distribuzione in diversi testi linguistici e tale distribuzione può essere rappresentata da un vettore che descrive la frequenza di occorrenza della parola con altre. Semplificando: quanto più due parole ricorrono in contesti linguistici simili, tanto più sono simili a livello semantico.

Il filo conduttore di questa tesi è rappresentato dal concetto di "lettura", intesa come capacità influenzata da una serie di variabili e potenzialmente soggetta a compromissioni o difficoltà. Nel corso dei successivi capitoli verranno espresse,

all'interno di differenti cornici teoriche, le principali caratteristiche che consentono di definire il linguaggio e le rispettive componenti. L'elaborato proseguirà con la spiegazione dei meccanismi sottostanti alla capacità di lettura, le rispettive variabili che la influenzano e i differenti modelli teorici esplicativi del processo di lettura. Infine, il focus descrittivo verrà orientato verso una direzione prettamente clinica, trattando ampiamente i disturbi correlati alla lettura quali la dislessia evolutiva e i relativi sottotipi.

Nella seconda parte del lavoro (vedi capitolo 4) verrà presentato il progetto di ricerca sperimentale realizzato presso l'Università degli studi di Pavia, condotto al fine di indagare la relazione tra fattori semantici e ortografici e processi di lettura. Nello specifico, lo studio si occupa di analizzare le componenti semantiche alla base della prestazione in partecipanti con diverse abilità di lettura in un compito di decisione lessicale.

# Capitolo 1: Il Linguaggio

## 1.1. Che cos'è il linguaggio

Il linguaggio è una funzione cognitiva specifica dell'essere umano, si tratta di una complessa forma di comunicazione, parlata, scritta o gestuale, basata su un apparato simbolico.

Rappresenta il sistema di comunicazione di cui l'uomo si avvale al fine di esprimere, comprendere e rappresentare il pensiero, mediante un codice condiviso.

L'unicità del linguaggio rispetto ad ogni altro sistema di comunicazione è determinata da una serie di caratteristiche specifiche che lo differenziano dalle altre modalità di espressione verbale non umane, animali o artificiali.

La principale funzione del linguaggio è dunque la funzione comunicativa, che consente la trasmissione di informazioni e l'interazione sociale. Accanto ad essa, occupa senza dubbio un posto di rilievo la funzione conoscitiva, che si concretizza nella capacità di descrivere gli eventi senza che l'individuo ne abbia avuto un'esperienza diretta.

La comunicazione umana si fonda principalmente sul canale vocale-uditivo: esso svolge un ruolo fondamentale, in quanto garantisce l'efficacia del processo comunicativo, impiegando una minima quantità di energia. In tal modo l'organismo ha la possibilità di svolgere altre attività mentre è in atto il processo di comunicazione.

Il concetto di linguaggio viene spesso confuso con quello di lingua, ma si tratta di due costrutti differenti, separati da un netto confine. Con il termine “lingua” si intende infatti un sistema simbolico, astratto, un codice specifico, comune a tutti gli individui che appartengono alla stessa comunità e socialmente determinato. La lingua rappresenta una possibile concretizzazione della facoltà del linguaggio all’interno di una comunità.

Per distinguere il linguaggio umano dalle altre forme di comunicazione animale o artificiale, sono state individuate una serie di proprietà uniche ed esclusive delle lingue storico-naturali, ad opera del linguista americano Charles F. Hockett (1916-2000). La presenza globale di tali proprietà è riservata esclusivamente alla comunicazione umana.

Una delle proprietà più significative del linguaggio è la dualità di strutturazione, *Duality of patterning*. Fa riferimento alla natura articolata del linguaggio, che ne consente la scomponibilità in unità minori. La comunicazione animale possiede una struttura costituita da forme inarticolate e difficilmente analizzabili, mentre il linguaggio verbale è un sistema riducibile a dei costituenti minimi.

La seconda fondamentale proprietà che contraddistingue il linguaggio umano è la Biplanarità. Con questo termine l’autore definisce la duplice struttura che costituisce le singole unità della lingua. Ogni segno linguistico è costituito da due dimensioni interdipendenti: l’espressione, che fa riferimento al significante, e il contenuto, che corrisponde al significato.

Altrettanto importante appare il concetto di discretezza con il quale si intende la proprietà delle unità linguistiche di opporsi reciprocamente, di distinguersi tra loro in modo dicotomico, senza la possibilità di gradazioni. Si tratta dunque di presenza o assenza di una determinata unità linguistica, non presenza maggiore o minore di esse.

La Produttività rappresenta una delle caratteristiche fondamentali che contribuiscono all'unicità del linguaggio umano. Tale proprietà consente ai parlanti di produrre un numero potenzialmente illimitato di messaggi, partendo da un numero definito di unità di base (fonemi e parole) e di regole di combinazione. Il linguaggio umano è dotato inoltre di Innovatività permanente. Esso è soggetto ad una continua trasformazione, un costante mutamento che dipende da fattori sociali e culturali (*“Le lingue possono e devono cambiare. L'innovatività permanente è una necessità vitale nel loro funzionamento”* Tullio De Mauro).

Hockett attribuisce fondamentale importanza alla ricorsività: con tale espressione fa riferimento alla prerogativa esclusiva del linguaggio umano, che consente di applicare illimitatamente la stessa regola, definita appunto ricorsiva. Si tratta dunque della possibilità di combinare tra loro gli elementi, senza limiti, mediante l'applicazione dello stesso meccanismo ricorsivo. A titolo di esempio, sono ricorsive le regole che permettono l'aggiunta di un aggettivo al nome (è possibile dire: “un amico leale, generoso, cordiale, allegro ...”).

Una delle proprietà fondamentali che contraddistinguono il linguaggio umano è la rapida evanescenza, che fa riferimento alla breve durata dei suoni che costituiscono la lingua parlata, a differenza della comunicazione scritta.

Il linguaggio è dotato altresì di intercambiabilità; questo concetto indica la possibilità dei due attori della comunicazione (parlante e ascoltatore) di riproporre quanto già espresso e udito, essi presentano appunto ruoli intercambiabili.

Altrettanto importante è la retroazione completa: il parlante ha la possibilità di ascoltare sé stesso, monitorando così la propria produzione verbale.

Lo scopo primario del linguaggio umano è quello di trasmettere informazioni; tale proprietà viene definita specializzazione.

Hockett attribuisce fondamentale importanza al costrutto di semanticità: il linguaggio umano veicola un contenuto semantico, correlato con un determinato referente.

L'autore si sofferma inoltre sul concetto di distanziamento o libertà dallo stimolo: il linguaggio umano può essere impiegato per parlare di eventi lontani nel tempo e nello spazio, oggetti e persone non presenti fisicamente, concetti astratti e accadimenti passati o futuri. L'uomo è in grado di svincolarsi dal "qui ed ora", a differenza della comunicazione animale, che risulta invece ancorata al presente e al concreto. Infine, ultima ma non per importanza, la tradizione, definita da Hockett *traditional transmission* che riguarda la prerogativa del linguaggio umano di essere trasmesso da una generazione all'altra.

## 1.2. Definire il linguaggio

La definizione di linguaggio costituisce da tempo un tema assai controverso. Molti autori si sono interrogati sulle origini e sullo sviluppo del linguaggio, ma rimane tutt'ora complesso definire in modo univoco questo dominio.

Il linguaggio rappresenta una delle capacità cognitive biologicamente determinate, una facoltà secondo Chomsky innata (Teoria Innatista), che si inserisce all'interno del vasto repertorio di abilità che contraddistinguono la specie umana.

I primi studi riguardanti il linguaggio, che segnano l'avvento della psicolinguistica, risalgono ai primi anni Cinquanta del Novecento, periodo in cui la cornice teorica di riferimento è prettamente di stampo comportamentista. Skinner definisce il linguaggio come un "Comportamento Verbale" (*Verbal Behaviour*): "*Comportamento rinforzato attraverso la mediazione di altre persone*" (Skinner, 1957).

Viene dunque visto come il risultato di apprendimento, associazione e imitazione. Analogamente ai processi di condizionamento operante attuati su animali da laboratorio, basati sull'associazione stimolo-risposta e incentivati dai rinforzi positivi, l'adulto modella il comportamento verbale del bambino mediante gratificazioni, sorrisi e altri cenni di approvazione.

In questo contesto teorico, appare evidente come, secondo Skinner, il contributo del bambino al processo di sviluppo, risulta pressoché inesistente. Alla fine degli anni cinquanta si assiste a quella che viene definita "la svolta cognitiva", ad opera del linguista americano Noam Chomsky. L'autore cognitivista disapprova

l'operato di Skinner, pubblicando un'aspra critica relativa al saggio del collega "*Verbal Behaviour*" (Chomsky, 1959). Chomsky smentisce l'idea che la competenza linguistica sia spiegabile esclusivamente mediante manifestazioni comportamentali prodotte in risposta a stimoli ricevuti durante i processi di apprendimento. La tematica trattata da Chomsky viene definita come "Ipotesi della povertà dello stimolo".

Il termine "povertà" sottolinea proprio come l'input linguistico che il bambino riceve dall'ambiente esterno, sia insufficiente per potergli permettere di sviluppare competenze linguistiche mediante la semplice associazione stimolo-risposta. Gli stimoli linguistici che il neonato riceve dal *caregiver*, sono costituiti da frasi incomplete, interrotte e grammaticalmente imprecise. Risulta quindi difficile spiegare come il bambino possa costruire la propria competenza linguistica a partire da stimoli ambientali, come postulato dall'approccio comportamentista.

Il nuovo clima culturale determinato dalla teoria linguistica chomskiana abbraccia una nuova concezione di linguaggio, inteso come "fatto mentale", ponendo l'attenzione sui processi psicologici sottostanti l'esecuzione e la comprensione del linguaggio, la formazione e la comprensione delle frasi.

In particolare, Chomsky postula l'esistenza di una Grammatica Universale. Si tratta di un dispositivo mentale innato, una dotazione biologica esclusiva della specie umana, in grado di guidare il bambino nell'apprendimento della lingua.

L'intervento di Chomsky innesca un dibattito che pervaderà il panorama psicologico e linguistico per diversi anni, fino alla contemporaneità.

La dicotomia tra innatismo ed evoluzione relativa alle origini del linguaggio rappresenta dunque un tema assai controverso, da un punto di vista filogenetico. La teoria di Chomsky è in grado di spiegare a livello ontogenetico come l'individuo umano possa acquisire il linguaggio mediante la grammatica universale, tuttavia non fornisce spiegazioni relative all'origine del linguaggio in un'ottica filogenetica. Chomsky rifiuta l'ipotesi darwiniana secondo cui il linguaggio è frutto di processi di adattamento e selezione naturale (Aroldi, 2010). A partire da questo presupposto si sono dunque delineate due differenti visioni, entrambe accumulate dall'adesione all'ipotesi innatista chomskiana, ma che si distinguono per il ruolo attribuito alla selezione naturale. La Teoria Evoluzionista, rappresentata dai neodarwiniani, concepisce il linguaggio umano come il prodotto dell'adattamento biologico degli individui alle pressioni ambientali, sulla base di un processo di selezione naturale.

In contrapposizione si schiera la Teoria Discontinuista, la quale definisce il linguaggio umano come un aspetto estremamente complesso, al punto tale da considerare riduttiva una spiegazione fondata sulla semplice selezione naturale. I sostenitori di questo approccio affermano che lo sviluppo del linguaggio umano sia determinato da leggi di crescita indipendenti dalla selezione naturale. Essi abbracciano altresì una visione esaptazionista, sostenendo che lo sviluppo del linguaggio sia stato determinato dall'evoluzione e lo sviluppo di abilità e fattori biologici inizialmente non linguistici (Aroldi, 2010).

Attualmente la ricerca psicolinguistica è ancora parzialmente ispirata alle teorie di Chomsky, ma sta attraversando la fase post-chomskiana, che attribuisce maggiore rilievo agli aspetti semantici e pragmatici del linguaggio. In questo periodo si manifesta inoltre la tendenza a contestualizzare il linguaggio in un più ampio contesto cognitivo (Cacciari, Canal, 2023).

### **1.3. Le componenti del linguaggio**

Il linguaggio è un sistema gerarchico costituito da differenti livelli interdipendenti. Le abilità linguistico-comunicative presentano un'organizzazione bottom-up, che procede dal basso verso l'alto, dunque i livelli inferiori esercitano influenza su quelli superiori. Complessivamente il sistema è governato da principi di auto-organizzazione, che consentono il funzionamento ottimale delle capacità linguistico-comunicative, mediante l'integrazione equilibrata di aspetti individuali, sociali, neurali e adattivi. Tale processo è definito emergente, in quanto imprevedibile e di difficile spiegazione (Cangelosi e Turner, 2002).

La struttura del linguaggio, inteso come sistema gerarchico, presenta diversi livelli di analisi; in particolare, vi sono componenti: fonetiche, fonologiche, morfofonologiche, morfologiche, morfosintattiche, semantiche, sintattiche, pragmatiche ed infine testuali/discorsive.

Assumendo una visione prettamente cognitiva, il focus di interesse si sposta dall'analisi dei singoli livelli che costituiscono il linguaggio, alla comprensione

del funzionamento complessivo delle dimensioni elaborative, responsabili di specifiche competenze.

In quest'ottica è dunque possibile distinguere tra una prima dimensione microelaborativa del linguaggio, alla quale si accosta la dimensione macroelaborativa (Marini, 2021).

La dimensione microelaborativa del linguaggio comprende i sistemi di elaborazione fonetica, fonologica, morfofonologica, morfologica, semantico-lessicale, morfosintattica, sintattica e semantica.

Il sistema di elaborazione fonetica è responsabile dell'elaborazione e articolazione dei foni, ovvero i suoni linguistici emessi e percepiti (Albano Leoni e Maturi, 1995). I foni che vengono usati in modo sistematico in una lingua vengono definiti fonemi.

Il sistema di elaborazione fonologica, invece, si occupa di classificare foni diversi come rappresentazioni differenti di un medesimo fonema (Nespor, 1994).

Un ruolo fondamentale è svolto dal sistema di elaborazione morfofonologica, che consente di organizzare foni e fonemi in sillabe.

Fonemi e sillabe vengono così organizzati in sequenze fonologiche definite morfi, che corrispondono a specifici significati astratti, noti come morfemi. Questa funzione viene svolta dal sistema di elaborazione morfologica.

Il sistema di elaborazione semantico-lessicale si occupa di organizzare il significato letterale (inteso come possibili significati di una specifica parola indipendentemente dal contesto) delle parole prodotte o percepite.

Il sistema di elaborazione morfosintattica opera con aspetti prettamente grammaticali, esso è il responsabile della corretta costruzione di frasi.

In stretta relazione con tale livello, è possibile individuare il sistema di elaborazione sintattica, che si occupa di generare gruppi ordinati di parole, definiti sintagmi, i quali andranno a costituire le frasi.

Infine, risulta fondamentale il ruolo svolto dal sistema di elaborazione semantico-frasale, che consente di associare a una frase un significato letterale, ovvero una proposizione.

La dimensione macroelaborativa del linguaggio, è costituita dal sistema di elaborazione pragmatica e dal sistema di elaborazione testuale/ discorsiva.

Il primo si occupa di contestualizzare i significati veicolati da singole parole e da intere frasi mediante la generazione di inferenze. L'operato di tale sistema è reso possibile grazie alle intenzioni comunicative dei parlanti e alla loro capacità di riconoscere quelle dei loro interlocutori.

Fondamentale in questa dimensione è il ruolo del contesto. E' possibile distinguere tra un contesto linguistico, inteso come l'insieme degli argomenti toccati all'interno di un discorso, una conversazione o un libro; e un contesto situazionale (o extralinguistico), che è invece costituito da informazioni di varia natura che riguardano le coordinate spazio/temporali della situazione comunicativa, la natura dell'evento comunicativo, le conoscenze implicite, il canale comunicativo e il codice utilizzato (Marini, 2021).

La seconda componente della dimensione macroelaborativa è costituita dal sistema di elaborazione testuale/discorsiva. Esso permette di generare o comprendere testi, discorsi o conversazioni, unità linguistiche di variabile lunghezza e complessità. La capacità di produzione implica la generazione di campioni di linguaggio percepiti come coerenti da chi legge o ascolta. In merito alla comprensione invece, consente di collegare tra loro i significati contestualizzati delle unità linguistiche lette o ascoltate, instaurando legami di coesione linguistica e coerenza concettuale.

#### **1.4. Linguaggio scritto e capacità di lettura**

Fino ad ora la presente tesi si è occupata di esaminare i principali aspetti strutturali e processuali che contraddistinguono il linguaggio parlato, sottolineando la sua natura multicomponentiale che contraddistingue le due dimensioni elaborative individuate (Cangelosi e Turner, 2002). In questa sezione verrà posta attenzione agli aspetti relativi all'elaborazione e comprensione del linguaggio scritto.

*“L'obbiettivo del linguaggio scritto è quello di rappresentare il linguaggio parlato mediante un codice visivo, la sfida per il sistema visivo umano è tradurre questa complessa serie di caratteristiche della parola scritta in un suono dotato di significato”* (Yeatman & White, 2021).

Il successo della lettura dipende da alcuni aspetti principali: la decodifica fonologica, che implica l'applicazione di abilità fonologiche di conversione grafema-fonema, e le abilità linguistiche orali, vocabolario e conoscenza grammaticale (Gough & Tunmer, 1986).

Durante le prime fasi di acquisizione della capacità di lettura per i bambini è fondamentale una decodifica accurata al fine di garantire una corretta comprensione di ciò che leggono (Castles et al., 2018). Quando i bambini imparano a leggere, inizialmente tendono ad usare le proprie capacità di elaborazione fonologica per decodificare la parola scritta e accedere successivamente alla rappresentazione corrispondente nel lessico (Hulme & Snowling, 2013).

Con l'automatizzazione dei processi di decodifica, che vengono consolidati con la pratica, le abilità linguistiche orali sembrano diventare le principali responsabili della capacità di comprensione del contenuto letto (Gough et al., 1996).

Capacità di lettura e abilità fonologiche si influenzano reciprocamente: le abilità fonologiche che l'individuo possiede sono determinanti per una lettura accurata e, allo stesso tempo, l'accuratezza di lettura si riflette positivamente sui processi di sviluppo e consolidamento delle abilità fonologiche dell'individuo (Cunningham et al., 2021; Wagner et al., 1994). Nello specifico, è stato dimostrato che la consapevolezza fonologica, intesa come la capacità di elaborare e manipolare in modo consapevole sequenze fonologiche, costituisce un valido predittore per l'accuratezza nella lettura precoce (Cunningham et al., 2021; Hulme & Snowling, 2013). A supporto di ciò, alcuni studi longitudinali effettuati su bambini hanno dimostrato che la consapevolezza fonologica all'inizio dell'apprendimento delle abilità di lettura è positivamente correlata all'accuratezza della lettura (Cunningham et al., 2021; Muter et al., 2004).

Lo sviluppo dell'abilità di lettura rappresenta un aspetto fondamentale per lo sviluppo della competenza lessicale, intesa come l'abilità di un individuo di riconoscere, percepire e comprendere le unità lessicali e saperle combinare al fine di costruire frasi (Gairns e Redman, 2009). L'arricchimento del lessico è una condizione fondamentale per migliorare l'abilità di lettura e la *reading comprehension* e, a sua volta, il progresso nelle capacità di lettura incentiva l'espansione lessicale; si tratta dunque di una relazione bidirezionale (Ellis, 1994). La conoscenza del lessico è multidimensionale e può essere acquisita in modo incrementale attraverso l'esposizione ripetuta. Si ipotizza che siano le esposizioni nel tempo e in contesti diversi a consentire il perfezionamento e la differenziazione della conoscenza delle parole (Stahl & Fairbanks, 1986). Quando i bambini sono esposti a parole in contesti diversi attraverso incontri ripetuti, ogni incontro fornisce nuove informazioni sulla parola; l'interazione attiva con le parole favorisce l'acquisizione del vocabolario (Stahl, 2004).

Durante i primi due anni di scuola il patrimonio lessicale del bambino cresce di circa mille parole all'anno (Nagy 1997). Un fattore determinante per l'espansione lessicale è rappresentato dal numero di esposizioni a una determinata porzione di lessico cui lo studente è esposto (Saragi et al., 1978).

Il ruolo del contesto in relazione all'arricchimento del vocabolario viene preso in esame da Smith, secondo cui la quantità di linguaggio scritto a cui gli individui sono quotidianamente esposti (ad esempio cartelli, insegne, etichette, annunci, ...) viene percepita in modo differente da bambini e adulti. Mentre i lettori adulti hanno

sviluppato la capacità di ignorare gli stimoli considerati non rilevanti, per i bambini tali aspetti possono costituire vere e proprie opportunità di apprendimento e sperimentazione. L'individuo si trova immerso in una realtà molto stimolante da un punto di vista linguistico: sia in merito al linguaggio parlato, al quale è costantemente esposto quotidianamente, che in riferimento a stimoli linguistici scritti, che il bambino può interpretare come significativi, rendendoli una valida occasione di apprendimento. Secondo Smith è proprio questa costante stimolazione ad esercitare influenza sull'apprendimento del linguaggio (Smith 1976).

# Capitolo 2: La capacità di lettura

## 2.1. I modelli di lettura

L'analisi dei modelli di lettura permette di approfondire diverse prospettive che mirano a spiegare il funzionamento sottostante ai processi di lettura.

Nei seguenti paragrafi verranno descritti due tra i più importanti modelli di lettura, che presentano approcci teorici differenti: il Modello di lettura a due vie e i Modelli semantici distribuzionali.

### 2.1.1. Modello di lettura a due vie

Il modello maggiormente considerato e al quale si fa frequentemente riferimento nella prassi diagnostica e clinica, è il Modello di lettura a due vie, *Dual Route Model*, (Coltheart et al., 1993) di impostazione non connessionista.

L'idea su cui si fonda tale modello parte dal presupposto che le parole scritte possono essere elaborate attraverso due vie: una via fonologica o sub-lessicale e una via diretta o lessicale, che a sua volta può seguire una linea semantica o non semantica.

Il punto di partenza, che sancisce l'avvio del processo di lettura, consiste nel riconoscimento visivo della parola. Attraverso l'attivazione della via lessicale, il lettore è in grado di riconoscere globalmente la parola e individuare, all'interno del lessico mentale, la rispettiva forma scritta e la pronuncia.

A questo punto, la via lessicale, può seguire due percorsi: uno semantico, che prevede l'attivazione del lessico mentale semantico, consentendo la rievocazione del significato della parola, e uno non semantico, che invece non coinvolge il significato. L'accesso alla forma fonologica della parola scritta è consentito dal passaggio dell'informazione dal lessico ortografico di input al lessico fonologico di output. La connessione diretta tra questi due distretti consente una rapida trasformazione della parola dalla sua forma scritta a quella fonologica.

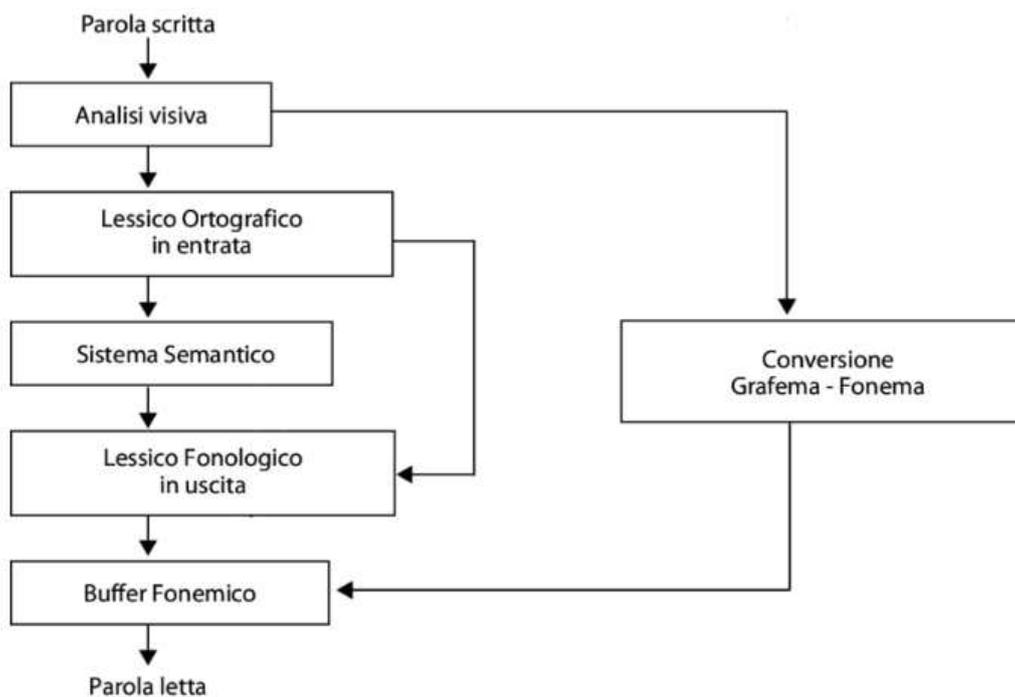


Figura 1: Modello standard di lettura (Coltheart et al., 1993)

Il tratto finale attraversato da questa via di lettura è condiviso con il percorso sublessicale: il buffer di output fonologico. Si tratta di una componente fonologica della memoria di lavoro, che svolge due funzioni principali: trattiene l'informazione fonologica fino alla sua completa produzione e, successivamente,

assembla le varie unità linguistiche al fine di formare parole e frasi. In tal modo l'individuo è in grado di leggere le parole note, delle quali conosce il significato; tuttavia risulta impossibile leggere sia le unità linguistiche che non ha mai avuto modo di sperimentare, che le stringhe di lettere non dotate di significato (definite non-parole).

La lettura di queste ultime è resa possibile grazie all'attivazione della via fonologica (o sub-lessicale). Questo processo prevede la scomposizione della parola in grafemi e la successiva conversione in fonemi, mediante l'applicazione delle regole grafema-fonema in modo sequenziale. Infine, grazie all'intervento della memoria fonologica di lavoro, la via di lettura sub-lessicale si ricongiunge con quella lessicale, permettendo la conversione corretta dei grafemi in fonemi.

I lettori esperti leggono attraverso entrambe le vie: se si tratta di una parola esistente e nota all'individuo, risulta più rapida una lettura mediante la via lessicale. Di fronte ad una parola non esistente o non presente nel lessico del lettore, può entrare in funzione unicamente la via sub-lessicale. (Friedmann & Coltheart, 2018).

Il modello a due vie riprende l'idea che già Saussure portava avanti sostenendo: *“Noi leggiamo in due modi: la parola nuova o sconosciuta viene letta lettera dopo lettera, mentre la parola usuale e familiare s'abbraccia d'un sol colpo d'occhio, indipendentemente dalle lettere che la compongono, sicché l'immagine di parole del genere acquista per noi un valore idiografico”* (Saussure, 1922).

Il modello a due vie di Coltheart, si contrappone a quelli di stampo connessionista, che implicano l'interazione della dimensione fonologica e semantica durante processo di lettura (Plaut et al., 1996; Harm e Seidenberg, 1999, 2004).

Questa contrapposizione fa perno sulle tre principali differenze esistenti tra i due approcci.

In primo luogo, i due modelli si differenziano per la natura della rappresentazione: il modello di lettura a due vie rappresenta localmente le parole, in termini di unità lessicali. Per i modelli connessionisti invece la rappresentazione è distribuita, dunque ogni parola è rappresentata mediante l'attivazione di diverse unità nel sistema, e ogni unità nel sistema ha un ruolo nella rappresentazione di parole diverse.

I due modelli si trovano in disaccordo anche rispetto alla natura del processamento: il modello a due vie opera in modo seriale e il processo di traduzione grafema-fonema avviene una lettera alla volta, da sinistra a destra. I modelli connessionisti, invece, prediligono un'elaborazione in parallelo, ovvero in ogni parola ciascuna lettera è processata contemporaneamente dal modello.

Infine, ciò che li differenzia è il concetto di apprendimento: secondo i modelli di stampo connessionista, la lettura ad alta voce costituisce un'abilità che si sviluppa mediante la continua esposizione all'ortografia delle parole e alla loro corretta pronuncia. Questa competenza si ottiene sotto il controllo di un algoritmo di apprendimento che, con lo sviluppo, diventa progressivamente più accurato. Il modello a due vie invece, definisce l'acquisizione della capacità di lettura come il

risultato di un processo di apprendimento circoscritto al sistema di elaborazione delle informazioni.

### **2.1.2. Modelli semantici distribuzionali**

Accanto ai modelli tradizionali di stampo non connessionista, si contrappone un approccio differente, che abbraccia la visione della semantica distribuzionale.

Secondo la visione della semantica distribuzionale il lessico costituisce uno spazio metrico i cui elementi – le parole – sono separati da distanze che dipendono dal loro grado di similarità semantica (Miller & Charles, 1991).

Il concetto di similarità semantica fa riferimento alla misura delle distribuzioni statistiche di co-occorrenza delle parole nei testi. Grandi corpi testuali rappresentano un buon indicatore del linguaggio a cui gli individui sono esposti nella vita quotidiana, dunque risultano fondamentali per cogliere i pattern di co-occorrenza di parole che compaiono spesso negli stessi contesti per poter inferire le associazioni tra tali parole.

Il principio epistemologico alla base di questo approccio è stato elaborato da Harris nel 1970 e viene definito “Ipotesi distribuzionale”. Secondo tale teoria due parole sono considerate semanticamente simili, in relazione a quanto entrambe ricorrono in contesti linguistici simili.

I significati delle parole derivano dunque dalla loro distribuzione in diversi testi linguistici e tale distribuzione può essere rappresentata da un vettore che descrive la frequenza di occorrenza della parola con altre. Semplificando: quanto più due

parole ricorrono in contesti linguistici simili, tanto più sono simili a livello semantico.

Questa cornice teorica di riferimento trova l'applicazione pratica corrispondente nei modelli distribuzionali: "*Distributional Semantic Models*" (DSM).

I DSM costituiscono una classe di modelli mediante i quali il significato di una parola viene appreso estraendo regolarità statistiche dal linguaggio naturale. Forniscono meccanismi espliciti su come le parole o le caratteristiche di un concetto possono essere apprese dall'ambiente naturale. (Kumar 2021).

Questa classe di modelli include al suo interno una distinzione fondata sulla presenza di due differenti meccanismi alla base dell'apprendimento associativo: uno error-free e uno error-driven. Il primo sostiene che l'apprendimento avvenga attraverso l'identificazione di cluster di eventi che tendono a co-occorrere con vicinanza temporale. I meccanismi error-driven, invece, teorizzano che l'apprendimento avvenga quando si applicano metodi di correzione degli errori alle predizioni di eventi in risposta a stimoli. (Kumar 2021).

Tra i primi e più noti DSM proposti, fondati sul concetto di error-free, è possibile individuare l'"*Hyperspace Analogue to Language*" (HAL) ad opera di Lund e Burgess (1996).

Le rappresentazioni vengono costruite sulla base dell'occorrenza di parole in un range di 5-10 parole. Il grado di co-occorrenza tra due termini è inversamente proporzionale alla distanza dei due all'interno della finestra: le parole

rappresentate più vicine nello spazio sono maggiormente relate, viceversa, parole lontane spazialmente risultano avere un minor grado di relazione.

Di fondamentale importanza è il “*Latent Semantic Analysis*” (LSA) elaborato da Landauer e Dumais (1997). Secondo tale modello la rappresentazione semantica di una parola è concepita come un pattern distribuito su più dimensioni. Prevede la costruzione di una matrice delle parole di un corpus testuale, in cui ogni riga rappresenta la frequenza di una parola nel testo. Tali valori vengono trasformati in frequenze logaritmiche ponderate in relazione all’importanza complessiva della parola nel testo, al fine di ridurre l’influenza di parole frequenti nel corpus, ma non importanti. Infine, questa matrice trasformata viene sottoposta ad analisi fattoriale. Un ulteriore modello appartenente ai DSM è il “*Bound Encoding of the Aggregate Language Enviroment - BEAGLE*”, di Jones e Mewhort (2007).

Esso si fonda sul presupposto per cui le rappresentazioni semantiche si costruiscono gradualmente proseguendo con il processamento del testo contestuale.

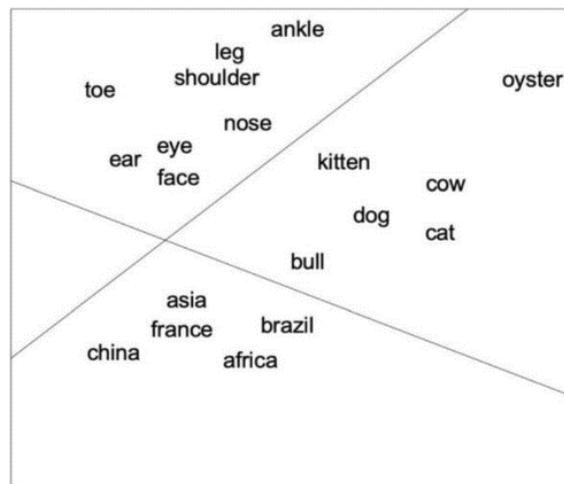


Figura 2: “*Bound Encoding of the Aggregate Language Enviroment*” BEAGLE (Jones e Mewhort, 2007)

I meccanismi error-driven fanno invece riferimento ai principi ipotizzati da Rescorla e Wagner (1972) e suggeriscono che l'apprendimento potrebbe essere guidato dagli errori di aspettativa.

Modelli che seguono questo tipo di meccanismo vengono definiti “connessionisti” e generalmente comprendono un livello di input e uno di output, separati da una o più unità intermedie, definite livelli nascosti, ognuno dei quali contiene uno o più "nodi". Con l'arrivo di uno stimolo esterno si attivano i nodi di input portando all'attivazione o soppressione delle rispettive unità collegate, a seconda della forza delle connessioni tra di esse. Questo processo continua progressivamente fino a raggiungere le unità di output.

Attualmente il modello più popolare in quest'ambito è il “word2vec” (Mikolov & Chen, 2013). E' possibile suddividere il modello in due forme distinte: il “*Continuous bag of words - CBOW*” e lo “*Skip gram*”.

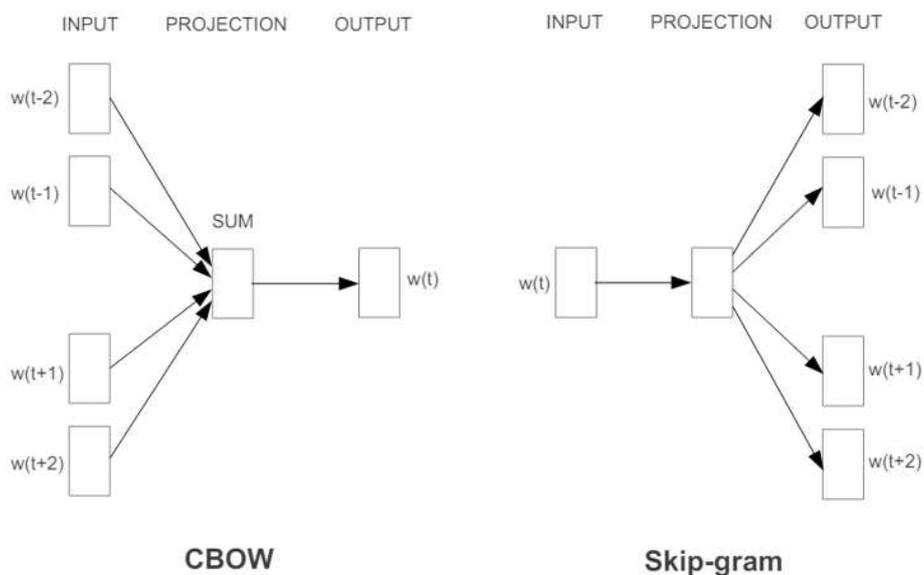


Figura 3: “Word2vec” Mikolov, Chen, et al., 2013.

Nel primo caso si può predire una parola target avendo a disposizione quattro parole di contesto prima e dopo il target, mentre la seconda versione esegue il compito inverso ovvero predire le parole contestuali avendo a disposizione una parola input.

Il “word2vec” è in grado di minimizzare il ruolo delle parole meno frequenti, scartando oltre una determinata soglia.

Un modello successivo, elaborato da Pennington, Socher e Manning (2014), prende il nome di “*Global Vectors - GloVe*”. Tale modello si occupa di stimare il rapporto delle probabilità di co-occorrenze tra parole, allo scopo di prevedere il logaritmo di questi rapporti di co-occorrenza attraverso un modello di regressione. Questo modello utilizza informazioni globali nel processo di apprendimento, seguendo un meccanismo error-free, ma fa altresì riferimento a meccanismi error-driven che consentono di minimizzare la funzione di costo nel modello di regressione. Combina dunque i vantaggi delle due principali famiglie di modelli presenti in letteratura: la fattorizzazione globale e i metodi locali. Si pone inoltre l'obiettivo di ridurre l'influenza eccessiva di parole frequenti e rare e sfrutta in modo efficiente l'informazione statistica escludendo termini con frequenza estremamente alta o bassa.

Il modello GloVe ha dimostrato ottimo funzionamento in compiti di analogia, giudizi di somiglianza delle parole e riconoscimento di entità nominative.

In entrambi i modelli descritti (*word2vec* e *GloVe*) ogni parola è rappresentata da un vettore distinto, tuttavia non viene considerata la struttura morfologica delle

singole parole. Questo può rivelarsi limitante soprattutto per lingue morfologicamente più complesse. A partire da questo presupposto è stato proposto un nuovo approccio, ideato sulla base del modello *Skip-gram*, che prende il nome di fastText (Joulin et al., 2016).

Questo modello è in grado di calcolare rappresentazioni semantiche che includono anche informazioni sulle sottoparole, mediante la somma dei vettori degli n-grammi di lettere associati a ciascuna parola. (Bojanowskiy, Gravey, Joulin e Mikolov, 2017).

## **2.2. Le variabili che influenzano la capacità di lettura**

La capacità di lettura può essere influenzata da una serie di variabili, che sono in grado di modellare la prestazione del lettore intervenendo sui processi di attivazione lessicale.

E' possibile suddividere i vari fattori in tre distinte sezioni, ciascuna delle quali comprende concetti differenti. In primo luogo si fa riferimento alle variabili base, che includono lunghezza, frequenza e classe; a seguire, le variabili semantiche (età di acquisizione, immaginabilità e concretezza); infine, di notevole importanza, vi sono le variabili lessicali, che riguardano il concetto di vicinato ortografico (frequenza media del vicinato, ampiezza del vicinato, distanza media dal vicinato e densità del vicinato).

### **2.2.1. Variabili base**

Le variabili di base comprendono: lunghezza, frequenza, familiarità e classe. La lunghezza di una parola è definibile come il numero dei simboli ortografici che la compongono, intesi come lettere o sillabe, entrambe potenzialmente influenti sul processo di lettura e riconoscimento di parole: la velocità di lettura diminuisce all'aumentare della lunghezza (Weekes, 1997).

La frequenza, invece, è ulteriormente suddivisibile in: frequenza nello scritto, che fa riferimento alle parole che vengono lette e scritte più frequentemente; frequenza per la lingua parlata, che indica le parole che compaiono più frequentemente durante gli scambi comunicativi; frequenza dei bigrammi, che corrisponde alla frequenza con cui coppie di lettere adiacenti compaiono nei testi stampati. In quest'ottica, assume un ruolo altrettanto importante, la variabile di familiarità: si tratta di una misura di frequenza soggettiva, che valuta "quanto una parola è presente nella vita di una persona" (Noble, 1953). Indica dunque, la frequenza con cui una parola viene utilizzata nella vita quotidiana. Lunghezza e frequenza si dimostrano essere i principali fattori che condizionano l'apprendimento della lettura: parole molto lunghe e poco frequenti richiedono un maggiore sforzo cognitivo, soprattutto per i neo-lettori. La progressiva esposizione alle parole scritte tende a determinare una riduzione della differenza di velocità di lettura tra parole ad alta e bassa frequenza (Zoccolotti et al., 2009).

La classe, infine, fa riferimento al livello di scolarità del bambino. Generalmente, ad un aumento livello di scolarità, dunque dell'età, corrisponde una diminuzione

del tempo di lettura per parola. Con l'esperienza maturata durante gli anni di scolarità, i bambini imparano a sviluppare strategie sempre più efficaci che permettono loro di migliorare la rapidità e l'accuratezza nella lettura (Cornoldi e Candela, 2015; Cornoldi e Carretti, 2016).

### **2.2.2. Variabili semantiche**

Le variabili semantiche fanno riferimento all'età di acquisizione, all'immaginabilità e alla concretezza della parola. L'età di acquisizione corrisponde all'età in cui una parola e il suo corrispondente semantico, sono stati appresi, in forma parlata o in forma scritta. Si tratta di una variabile strettamente connessa al linguaggio parlato, dunque esercita maggiore influenza nei compiti in cui è previsto un output verbale, come la lettura ad alta voce e la denominazione di oggetti (Morrison e Ellis, 1995).

L'immaginabilità può essere definita come "la facilità e rapidità di una parola ad evocare un'immagine mentale, una rappresentazione visiva, un suono o qualche altra esperienza sensoriale" (Paivio, Yuille, e Madigan, 1968; Cornoldi, 1974; Roncato, 1974).

La concretezza, infine, riguarda la capacità di una parola di fare riferimento a un concetto concreto, ovvero percepibile tramite i sensi. All'aumentare della concretezza della parola, aumenta anche la velocità di lettura. (Paivio et al., 1968; Roncato, 1974).

### 2.2.3. Variabili lessicali

Le variabili lessicali riguardano il concetto di vicinato ortografico di una parola, in termini di frequenza media, ampiezza e distanza da esso. Per vicinato ortografico di una parola (ON: *Orthographic Neighborhood*), si intende l'insieme delle parole che differiscono da essa per un carattere omesso, inserito o sostituito. I vicini ortografici di una parola target corrispondono dunque alle parole che possono essere ottenute cambiando una lettera e preservando la posizione delle altre lettere. (Coltheart, Davelaar, Jonasson e Besner, 1977).

Alcuni autori (Peereman & Content, 1995) ipotizzano che l'effetto di vicinato ortografico sia dovuto alla frequenza delle unità sub-lessicali (grafemi o bigrammi) nella lingua. Assumendo questa prospettiva, è possibile affermare che le parole con molti vicini ortografici dovrebbero essere lette più velocemente, in quanto contengono corrispondenze grafema-fonema più comuni.

La frequenza media del vicinato indica la frequenza dei vicini ortografici di una parola target: le parole che presentano vicini ad alta frequenza tendono ad essere lette più velocemente.

La grandezza del vicinato corrisponde al numero dei vicini ortografici della parola target. Generalmente le parole con quartieri ampi (parole che hanno molti vicini ortografici) vengono lette più velocemente rispetto alle parole con ridotto vicinato.

La distanza media dal vicinato indica la misura di quanto la parola target si discosta dai propri vicini ortografici. All'aumentare della distanza media dal vicinato, il tempo di lettura di una parola target aumenta.

Il concetto di ON è tuttavia oggetto di critiche, in quanto alcuni autori sostengono che costituisca una misura eccessivamente restrittiva (Davis & Bowers, 2006). Uno dei limiti principali di questa variabile è rappresentato dal fatto che la somiglianza semantica tra due parole è rappresentabile mediante valori di variabilità graduale: due parole possono essere più o meno vicine ortograficamente, tuttavia ON costituisce una misura binaria, che consente unicamente di affermare se due parole sono o non sono vicini ortografici. Inoltre ON definisce due parole come ortograficamente vicine quando differiscono esclusivamente per una singola lettera, nonostante le operazioni di inserimento, cancellazione o trasposizione possano anche risultare in parole molto simili, mostrando effetti di attivazione che non sono riscontrabili utilizzando l'indice ON, ad esempio: TRIAL-TRAIL, WIDOW-WINDOW, PLANE-PLANET (Davis 2006). Queste limitazioni hanno determinato la necessità di individuare misurazioni alternative, che siano in grado di colmare i limiti della variabile di Coltheart. A tal proposito, Yarkoni, Balota e Yap hanno individuato una nuova misura di somiglianza ortografica, definita distanza ortografica di Levenshtein 20 (OLD20), graduale e maggiormente flessibile.

Si basa sulla *distanza di Lev-enshtein* (LD; Levenshtein, 1966), una metrica informatica standard della distanza di modifica delle stringhe. L'LD tra due parole corrisponde al numero minimo di operazioni di sostituzione, inserimento o cancellazione richieste per trasformare una parola in un'altra (Yarkoni, Balota & Yap, 2008).

Infine una delle variabili che esercitano influenza sui processi di lettura è la densità di vicinato semantico: *semantic neighborhood density SND*, intesa come il numero di parole che sono semanticamente simili a una parola o non parola (Hendrix & Sun, 2021). Il quartiere semantico di una parola corrisponde al numero di parole semanticamente vicine ad essa.

Nel quartiere la distanza di una parola rispetto ad altre varia, questa varianza si riflette nel livello di densità semantica. Quanto maggiore è il numero di parole presenti nello spazio, tanto più denso sarà il quartiere (Buchanan, Westbury e Burgess, 2001).

## **Capitolo 3: La Dislessia**

### **3.1. I Disturbi Specifici dell'Apprendimento: DSA**

I Disturbi Evolutivi Specifici dell'Apprendimento (DSA) sono entrati a far parte della vasta area di interesse clinico che riguarda l'età evolutiva, con diagnosi sempre più frequenti ed accurate, grazie ai preziosi contributi della ricerca scientifica e all'affinamento delle tecniche diagnostiche (Cornoldi, Pra Baldi e Rizzo, 1991; Sartori, Job e Tressoldi 1995; Cornoldi e Colpo, 1995,1998).

Si tratta di una specifica categoria diagnostica contraddistinta da criteri oggettivi e valutabili, aspetto cruciale che li separa nettamente dalla più generica espressione "difficoltà di apprendimento", che include differenti gradi e tipologie di difficoltà che si possono manifestare in ambito scolastico.

Generalmente, l'apprendimento e lo sviluppo delle capacità linguistiche, di lettura, di scrittura e di calcolo, avviene durante i primi due anni di scuola primaria, tali abilità diventano automatizzate, richiedendo un minimo dispendio di risorse per essere attuate.

La situazione si presenta differente in caso di presenza di Disturbi Specifici di Apprendimento (DSA). Gli individui ai quali viene rilasciata tale diagnosi, o potenzialmente diagnosticabili, non sono in grado di automatizzare le capacità di lettura, scrittura e calcolo.

I DSA rappresentano disturbi di origine neurobiologica che presentano un andamento cronico, ma evolutivo: mutano in relazione all'età, all'esperienza

dell'individuo e alle richieste dell'ambiente. I fattori ambientali (contesto familiare e sociale), svolgono dunque un ruolo fondamentale, esercitando la loro influenza in modo diretto e continuativo.

Il crescente interesse in ambito clinico, relativo a questa categoria diagnostica, ha determinato la necessità di definire degli standard clinici condivisi per la diagnosi e la riabilitazione di tali disturbi. La Consensus Conference del 2007, promossa dall'Associazione Italiana Dislessia, specifica le principali caratteristiche da considerare a fini diagnostici e riabilitativi. Recentemente questi criteri sono stati aggiornati e arricchiti nella nuova “Linea Guida sulla gestione dei Disturbi Specifici dell'Apprendimento”, pubblicata nel 2022 dall'Istituto Superiore di Sanità (ISS). Le principali e necessarie condizioni per poter porre diagnosi sono: la “discrepanza” e la “specificità”. Tali termini mirano a sottolineare che questa categoria di disturbi è caratterizzata da un buon funzionamento cognitivo generale, essenzialmente discrepante dal funzionamento di alcune abilità specifiche riguardanti le capacità di apprendimento scolastiche (lettura, scrittura e calcolo) che risultano inferiori alla media per età e classe frequentata. L'acquisizione delle competenze richieste, nonostante abbia margine di miglioramento, raramente raggiunge i livelli attesi per età e/o scolarità.

Vengono inoltre definiti “specifici” poiché interessano uno specifico dominio di abilità, in modo significativo ma circoscritto, senza intaccare il funzionamento intellettuale generale.

Più in generale, è possibile porre diagnosi solo se in assenza di deficit cognitivi, problemi ambientali e psicologici, deficit sensoriali o neurologici.

Secondo l'ICD-10 *"... I DSA sono disturbi nei quali le modalità normali di acquisizione delle abilità scolastiche sono alterate già dalle fasi iniziali dello sviluppo. Essi non sono semplicemente una conseguenza di una mancanza delle opportunità di apprendere e non sono dovuti a un trauma o a una malattia cerebrale acquisita ..."* (ICD-10: International Statistical Classification of Diseases and Related Health Problems 10th Revision).

I Disturbi Specifici dell'Apprendimento possono riguardare un ambito circoscritto, come lettura, scrittura o calcolo; tuttavia nella pratica clinica è più frequente incontrare l'associazione di più deficit.

Oltre alle difficoltà negli apprendimenti scolastici, questi disturbi possono riguardare anche altri ambiti cognitivi come la memoria, l'attenzione, il processamento visivo e le funzioni esecutive.

Generalmente la scuola costituisce l'ambiente in cui i disturbi si manifestano e dove, se non riconosciuti e ben gestiti, possono essere responsabili dell'insuccesso scolastico e avere conseguenze su comportamento e sviluppo psicologico dell'individuo, determinando ad esempio insicurezza, scarsa autostima e un maggiore livello di ansia, che si manifesta soprattutto nelle situazioni di verifica, interrogazione o lettura ad alta voce (Rodriguez, 2012).

In quest'ottica diventa fondamentale la loro identificazione precoce. La tempestività e l'accuratezza dell'intervento sono determinanti per l'evoluzione

positiva dei DSA. E' importante riconoscere i primi segnali, al fine di intervenire fin da subito adottando le misure adeguate. L'applicazione di misure compensative e l'utilizzo di strumenti dispensativi nell'ambito del percorso scolastico, consentono di creare le condizioni per garantire un apprendimento efficace agli individui con tale diagnosi.

A sostegno di tali considerazioni, nel 2010 è stata approvata una legge che tutela gli individui con diagnosi di Disturbi Specifici dell'Apprendimento: la Legge 170. Si tratta di un decreto che, coerentemente con i sistemi internazionali di classificazione clinica (DSM e ICD), definisce i Disturbi Specifici dell'Apprendimento e sancisce il diritto degli alunni con tale diagnosi di poter usufruire di misure educative e didattiche idonee lungo il percorso scolastico (compresa l'Università), e il dovere delle istituzioni scolastiche e accademiche, di rispettarli.

Tratto dalla Legge n. 170:

*“... è compito delle scuole di ogni ordine e grado, comprese le scuole dell'infanzia, attivare, previa apposita comunicazione alle famiglie interessate, interventi tempestivi, idonei a individuare i casi sospetti di DSA degli studenti...”* (art. 3).

Questa Legge ha dunque fornito gli strumenti per contrastare l'esclusione e l'emarginazione a livello educativo, modellando il sistema scolastico e orientandolo verso una direzione più inclusiva.

Negli ultimi anni si è assistito ad un aumento delle diagnosi di DSA, in quanto le nuove misure introdotte dal decreto legge, hanno consentito una migliore gestione

dei disturbi e incentivato una maggiore attenzione e scrupolosità nell'individuare i segnali precoci, rendendo possibile un intervento tempestivo.

Tuttavia ancora oggi un numero significativo di bambini con DSA può rimanere non riconosciuto come tale fino all'età adulta (Fenzi e Cornoldi, 2015). Questo è possibile in quanto le difficoltà di apprendimento possono rimanere latenti fino a quando le richieste dell'ambiente non superano le capacità del soggetto.

Il decreto dedica spazio anche all'analisi dei DSA nell'età adulta; le ricerche in quest'ambito sono poche, in quanto tale categoria diagnostica è generalmente associata all'età evolutiva. Viene sottolineato il diritto degli studenti universitari con DSA, di poter usufruire dei necessari strumenti di supporto e misure compensative, al fine di garantire un efficace apprendimento e rendimento accademico sia relativamente ai test di ammissione all'università che per gli esami universitari (Legge n. 170/2010, art. 5, comma 4).

Questo ha determinato un progressivo incremento delle richieste di valutazione da parte di individui che richiedono un aggiornamento della diagnosi (redatta durante la scuola primaria) al fine di poter beneficiare delle misure di supporto previste dalla normativa. Molte delle richieste provengono inoltre da studenti delle scuole secondarie di secondo grado o universitari, che sono stati ingiustamente costretti a dover affrontare l'intero percorso di studi in assenza di un adeguato supporto, costantemente alla ricerca di una spiegazione valida relativa alla natura delle loro difficoltà scolastiche, che non sono mai riusciti a trovare (Cornoldi e De Carli, 2015).

In quest'ottica risulta dunque fondamentale individuare le caratteristiche dei DSA in età adulta, conoscerne la loro evoluzione e i mutamenti. Conoscere i punti di forza e debolezza permette di individuare, le misure dispensative e gli strumenti compensativi ideali per il percorso scolastico e accademico di ogni individuo, in modo personalizzato.

Al fine di poter identificare la presenza di DSA in età adulta è necessario prendere in considerazione una serie di fattori durante la valutazione diagnostica. Molto spesso, raggiunta l'età adulta i soggetti con DSA non mostrano evidenti o profonde difficoltà negli apprendimenti, nonostante abbiano alle spalle una storia caratterizzata da difficoltà e insuccessi scolastici, (Beaton, McDougall e Singleton, 1997). La maggior parte tende a sviluppare strategie per compensare le difficoltà. Adulti con tali caratteristiche vengono descritti da Singleton, Horne e Simmons con l'accezione di "falsi negativi": si tratta di individui che possono facilmente passare inosservati alla diagnosi, tuttavia risentono comunque profondamente delle loro difficoltà.

In quest'ottica diventa quindi fondamentale conoscere le caratteristiche dei DSA in età adulta e disporre di strumenti diagnostici finalizzati a individuarne i punti di forza e debolezza. Questo, al fine anche di valutare, in modo personalizzato, le misure dispensative e gli strumenti compensativi più idonei per il percorso universitario di ogni singolo soggetto.

## **3.2. Dislessia Evolutiva**

### **3.2.1. Inquadramento diagnostico e incidenza**

Con il termine Dislessia evolutiva (DE) si fa riferimento al Disturbo specifico dell'apprendimento con compromissione della lettura. I manuali diagnostici DSM-IV e ICD-10 definiscono i criteri necessari per porre diagnosi, i quali devono includere difficoltà o compromissione relativamente ai seguenti domini:

- a) velocità o fluidità nella lettura;
- b) accuratezza nel processo di decodifica;
- c) comprensione del testo.

La diagnosi di DE può essere redatta a partire dalla fine della seconda classe della scuola primaria. Gli alunni che presentano tale disturbo mostrano due principali tipologie di compromissioni: le difficoltà fonologiche che riguardano compromettono il processo di conversione grafema-fonema; le difficoltà lessicali o ortografiche che implicano un deficit nel meccanismo che consente l'accesso e il recupero della forma ortografica e fonologica della parola dal lessico mentale.

Porre diagnosi di DE è un processo definito da due fasi distinte: in primo luogo avviene la diagnosi clinica, che è costituita dall'analisi di criteri di inclusione (sintomi presenti) e di esclusione (di patologie o anomalie sensoriali, neurologiche, cognitive e/o psicopatologiche). In seconda battuta viene redatta la diagnosi funzionale, che descrive il quadro clinico del soggetto. Si tratta di un approfondimento del profilo del disturbo, mediante la valutazione delle componenti dell'apprendimento e altre abilità complementari (linguistiche,

percettive, prassiche, visuo-motorie, attentive, mnestiche), unitamente all'analisi dei fattori ambientali e delle condizioni emotive e relazionali. Per completare il quadro diagnostico è possibile inoltre considerare la presenza di eventuali disturbi in comorbilità; in bambini con DSA è frequente la co-occorrenza di altri disturbi specifici dell'apprendimento, o la compresenza di altri disturbi evolutivi (come ad esempio *Attention Deficit Hyperactivity Disorder- ADHD*, *Oppositional Defiant Disorder- ODD*). Predisporre il profilo funzionale è un passo essenziale per la presa in carico e per un eventuale progetto riabilitativo.

L'incidenza della DE in popolazione è definibile sulla base di alcuni studi condotti in tempi relativamente recenti. Secondo il Manuale diagnostico e statistico dei disturbi mentali (DSM-5) la prevalenza dei DSA oscilla tra il 5 e il 15% in età scolare, trasversalmente alle diverse culture (APA, 2013). Il disturbo risulta essere più frequente nella popolazione maschile, rispetto a quella femminile, con un rapporto M/F stimato intorno a 2-3:1 (APA, 2013). Al fine di ottenere dati più attendibili e stimare la prevalenza della DE con maggiore accuratezza, nel 2008 è stato istituito il Comitato Epidemiologico Nazionale sulla Dislessia (CENDi). Si tratta di un organo che si occupa di definire i metodi e gli strumenti per la realizzazione di uno studio sulla prevalenza della DE nella popolazione scolastica italiana, facendo riferimento alle indicazioni espresse dai documenti Consensus e in particolare delle precisazioni contenute nell'ultimo documento PARCC (2011).

Il primo studio condotto con la metodologia proposta dal CENDi, è stato realizzato in Friuli Venezia Giulia, su un campione di 1774 alunni frequentanti la classe quarta della scuola primaria (Barbiero e colleghi, 2012). I risultati hanno dimostrato che la percentuale di studenti con DE è compresa tra il 3,1% e il 3,2%. Successivamente, una ricerca condotta nella Regione Veneto (Tucci e Vio, 2016), utilizzando la stessa metodologia proposta dal CENDi, ha riportato risultati pressoché analoghi allo studio precedentemente citato. Più in generale, è possibile affermare che l'incidenza della DE, nella scuola primaria, corrisponde a un valore percentuale compreso tra il 3 e il 5%.

### **3.2.2. Sottotipi di Dislessia evolutiva**

La clinica diagnostica della DE elabora una classificazione dei sottotipi del disturbo, considerando come quadro teorico di riferimento il Modello a due vie di Coltheart. Attualmente è possibile distinguere 3 forme di DE: Dislessia Superficiale, Dislessia Fonologica e Dislessia Profonda, caratterizzate rispettivamente da deficit nella procedura di lettura di tipo lessicale, sub-lessicale o in entrambe.

Viene definito Dislessia Superficiale, il disturbo che implica la compromissione della via lessicale. Il bambino non ha accesso al magazzino semantico del lessico, dunque non è in grado di riconoscere parole regolari o irregolari note. Questo lo costringe ad un iper utilizzo della via sub-lessicale, applicando la conversione grafema-fonema anche con parole regolari o irregolari note. In definitiva, il

paziente con diagnosi di dislessia superficiale è in grado di leggere correttamente le parole regolari, mentre commette errori nella lettura di quelle irregolari. Proprio per questa sua caratteristica, questo sottotipo di dislessia tende ad essere meno diagnosticato nella popolazione appartenente alle lingue trasparenti (come l'italiano) e più facilmente diagnosticabile in pazienti la cui lingua madre è opaca, dunque contiene parole con ortografia irregolare o con eccezioni di pronuncia (come l'inglese). Alcuni errori che i bambini tendono a commettere sono: generalizzazioni nella lettura di parole irregolari (la parola viene letta come se fosse regolare), errori relativi alla posizione dell'accento nelle parole polisillabiche, errori ortografici come ad esempio sostituzioni, eliminazioni o aggiunte di lettere.

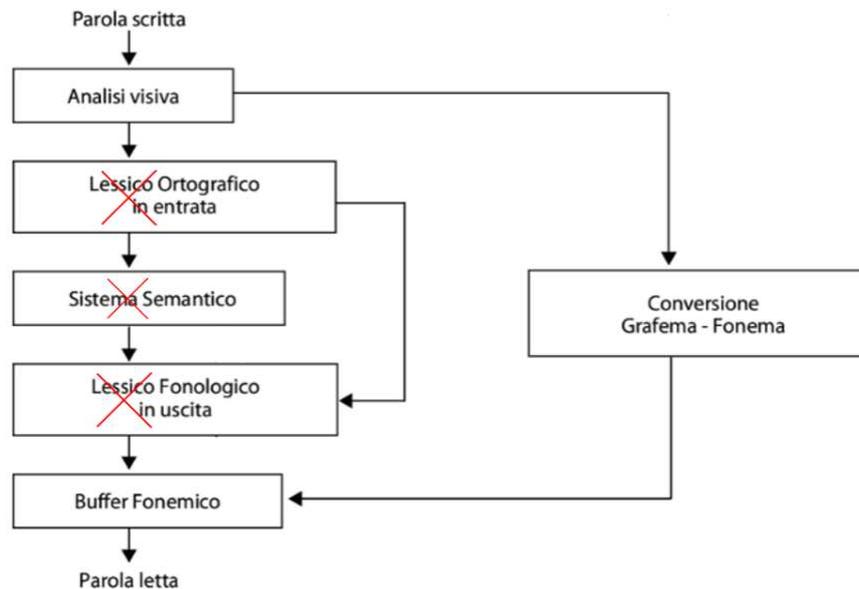


Figura 4: Componenti compromesse nella Dislessia Superficiale secondo il Modello a due vie di Coltheart

Mantenendo come riferimento teorico il Modello a due vie di Coltheart, è possibile individuare, in contrapposizione alla Dislessia Superficiale, la Dislessia Fonologica, descritta per la prima volta da Temple e Marshall (1983). Si tratta di un disturbo caratterizzato da una compromissione del percorso di lettura fonologica. Il paziente mostra una difficoltà sproporzionata nella lettura di non parole e di nuove parole non familiari, interpretate come un riflesso dell'incapacità di utilizzare la via sub-lessicale di conversione. Non essendo in grado di operare il processo di conversione grafema-fonema, l'individuo adotta un meccanismo di lettura che segue esclusivamente la via lessicale.

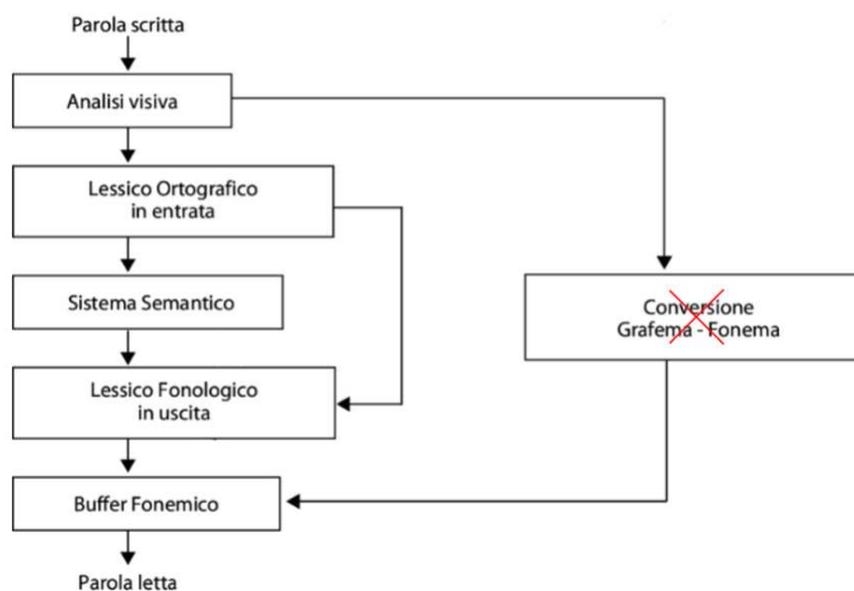


Figura 5: Componenti compromesse nella Dislessia Fonologica secondo il Modello a due vie di Coltheart

Il bambino non è in grado di applicare le regole di conversione grafema-fonema, dunque il processo di lettura è possibile esclusivamente mediante la via lessicale diretta. L'individuo risulta dunque impossibilitato nella lettura di non-parole o di parole sconosciute, mentre sembra rimanere conservata la capacità di leggere le

parole che fanno parte del vocabolario usuale del paziente. Tra gli errori più comunemente riscontrabili nei pazienti con tale diagnosi vi sono: errori visivi, soprattutto con parole visivamente simili, errori morfologici, incapacità di leggere le parole funzione (come congiunzioni e preposizioni) in quanto prive di significato e non immaginabili visivamente.

Infine, quando la compromissione riguarda entrambi i meccanismi di lettura presenti nel modello, è possibile una diagnosi di Dislessia Profonda, un disturbo multi componenziale che determina un'interruzione della via sub-lessicale di conversione, della via lessicale diretta e un deficit a carico del sistema semantico.

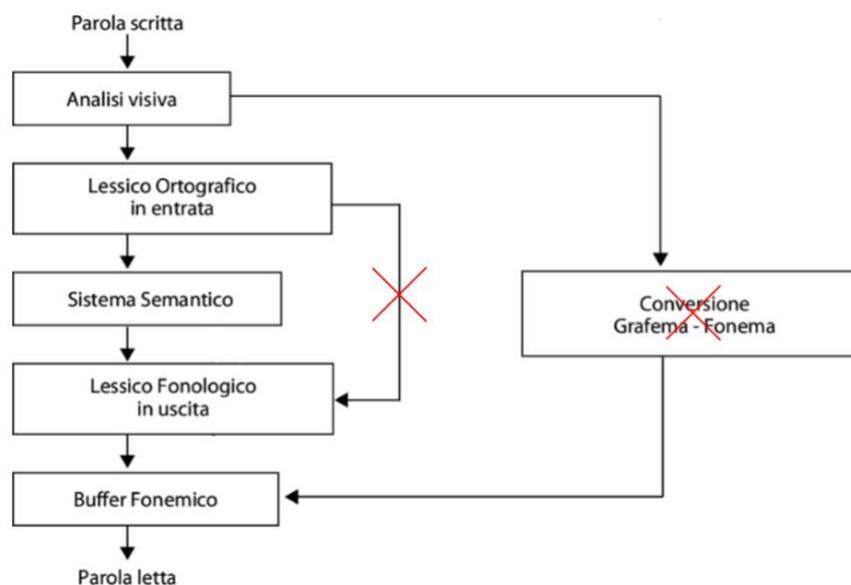


Figura 6: Componenti compromesse nella Dislessia Profonda secondo il Modello a due vie di Coltheart

Il bambino con tale diagnosi commette errori di varia natura: errori di tipo semantico, ovvero quando vengono presentate parole singole i pazienti non leggono la parola effettivamente scritta ma una di significato correlato a questa,

ad esempio viene presentata la parola “cane” e il soggetto dice la parola “lupo”; errori visivi: ad esempio “letto” al posto di “metto”, cioè quando la parola letta ha molte lettere in comune con quella che è realmente scritta; sostituzione di parole funzione, ad esempio “quando” viene letto “quanto”; errori morfologici, che riguardano la tendenza del lettore a modificare il prefisso o il suffisso di una parola, ad esempio “andavo” viene letto “andato”. Si assiste inoltre a quello che viene definito “effetto concretezza”: le parole concrete (facilmente immaginabili) sono lette meglio rispetto alle parole astratte.

### **3.3. La Dislessia nell’adulto e la compensazione delle difficoltà di lettura**

La DE costituisce un disturbo di origine neurobiologica che presenta un andamento cronico, ma evolutivo: pur manifestandosi con estrema variabilità nei soggetti, la sua presenza sembra rimanere costante nel tempo.

Alcune evidenze scientifiche confermano che i bambini con diagnosi di DE tendono a presentare difficoltà nei processi lettura anche in età adulta (Beaton, McDougall & Singleton, 1997). Il Connecticut Longitudinal Study, condotto da Shaywitz e colleghi nel 1999, ha fornito un prezioso contributo a supporto di queste premesse. Si tratta di uno studio longitudinale che prende in esame una popolazione di bambini con diagnosi di dislessia evolutiva, dai primi anni di scuola fino all’età adulta. Dai risultati si evince che i deficit nella codifica fonologica sembrano rimanere costanti durante la crescita (Shaywitz et al., 1999).

Accanto a queste evidenze, il panorama della letteratura scientifica offre inoltre risultati incoraggianti. Sembrerebbe che gli individui con DE siano in grado di compensare le abilità carenti mediante l'impiego di risorse personali che possono rappresentare dei punti di forza.

Ad esempio, uno studio condotto da Wiseheart e Altmann ha dimostrato che adulti con diagnosi di dislessia evolutiva che presentano un vocabolario più ampio, sembrano in grado di compensare meglio le proprie difficoltà rispetto a quelli con un vocabolario ridotto. Nello specifico, questi individui hanno mostrato punteggi migliori nelle prove di vocabolario, risultato che si riflette positivamente sulle loro prestazioni nel compito di produzione di frasi. Questi aspetti potrebbero rappresentare un fattore protettivo in grado di compensare le carenze a livello di memoria di lavoro, spesso evidenti in relazione alle difficoltà di lettura (Wiseheart e Altmann, 2017).

Questi dati sembrano essere in linea con l'idea proposta da Lefly e Pennington negli anni Novanta. Gli autori elaborano il concetto di "dislessia compensata", facendo riferimento ad una popolazione di individui che mostrano evidenti disturbi di lettura durante l'infanzia, ma che in età adulta non risultano diagnosticabili come dislessici (Lefly e Pennington, 1991).

Non si tratta di una completa scomparsa dei parametri deficitari; tuttavia sembra che i cosiddetti "dislessici compensati" siano in grado di leggere in modo accurato mantenendo tuttavia un scarso livello di fluency (Shaywitz et al., 2003).

In linea generale è possibile affermare che esistono possibilità di compensazione delle difficoltà di lettura che, seppur apportando un miglioramento alle prestazioni dell'individuo, tuttavia non consentono una risoluzione totale delle difficoltà, mostrando effetti che tendono a perdurare nel tempo (Kemp, Parrila e Kirby, 2009).

## Capitolo 4: Lo studio

### 4.1. Il progetto

Nella sezione seguente viene descritto il progetto di ricerca condotto presso l'Università degli studi di Pavia, finalizzato ad osservare la relazione tra fattori semantici e ortografici e processi di lettura. Nello specifico, il presente studio si occupa di indagare le componenti semantiche e ortografiche in grado di predire la prestazione in partecipanti con diverse abilità di lettura in un compito di decisione lessicale.

L'obiettivo è dunque quello di individuare eventuali differenze, in termini di prestazione, tra soggetti con peggiori o migliori abilità di lettura. In riferimento alle premesse teoriche menzionate nei precedenti capitoli, i risultati attesi dovrebbero confermare una maggiore sensibilità agli effetti semantici da parte dei partecipanti con peggiori capacità di lettura, rispetto ai lettori più abili. I partecipanti sono stati sottoposti ad un compito di decisione lessicale, seguito dalla somministrazione di una batteria di test costituita da prove riguardanti: lettura ad alta voce, ragionamento logico non verbale, memoria di lavoro, fluenza fonemica e semantica ed infine capacità di denominazione.

Le procedure sperimentali, le metodologie utilizzate e i risultati ottenuti verranno trattati nel dettaglio nei paragrafi successivi.

## **4.2. Metodi**

### **4.2.1. Partecipanti**

Hanno partecipato allo studio 53 studenti (20 maschi, 33= femmine, età media = 22 anni, SD = 2.16, range di età = 19 - 29). Tutti i partecipanti erano di madrelingua italiana, avevano una visione normale o corretta e non erano a conoscenza dello scopo dello studio. Prima dell'esperimento è stato ottenuto il consenso informato da tutti i partecipanti. Il protocollo è stato approvato dal comitato etico psicologico dell'Università di Pavia e i partecipanti sono stati trattati in conformità alla Dichiarazione di Helsinki.

### **4.2.2. Strumenti**

#### **Compito di decisione lessicale**

Le parole incluse nello studio sono state selezionate tra le 20.000 parole italiane più frequenti secondo SUBTLEX-it (<http://crr.ugent.be/subtlex-it/>), mentre le pseudoparole sono state selezionate tra un ampio set di circa 100.000 stimoli costruiti utilizzando il modulo ortografico italiano di Wuggy (Keuleers, & Brysbaert, 2010) a partire dalle parole incluse nel database italiano ANEW (Montefinese et al., 2014). Il set finale comprendeva 200 stimoli (100 parole e 100 pseudoparole). Le parole stimolo sono state selezionate in modo da avere una distribuzione quasi uniforme del nostro principale predittore semantico di interesse (vedi sotto) e, allo stesso tempo, per non avere distribuzioni sbilanciate degli altri

predittori inclusi. Abbiamo calcolato i seguenti predittori: frequenza, densità di vicinato semantico (SND), densità di vicinato ortografico (OND) e lunghezza.

Il nostro principale predittore semantico è stato SND, calcolato secondo Hendrix & Sun (2021; ma si veda anche Anceresi et al., 2024) come la similarità semantica media tra ogni pseudoparola e i suoi  $k$  vicini più prossimi (con  $k = 5$ ) all'interno delle 20.000 parole più frequenti in SUBTLEX-it. Valori più alti indicano quindi che un certo stimolo è più inserito nello spazio semantico. Per maggiori informazioni sul calcolo della similarità semantica, si veda la sezione Modello semantico distribuzionale.

La OND è stata calcolata come la distanza di Levenshtein media (che misura la distanza ortografica tra due stringhe di simboli quantificando il numero minimo di modifiche di un singolo carattere necessarie per cambiare un elemento nell'altro) tra la stringa di lettere e i suoi 20 vicini più prossimi (cioè OLD20; Yarkoni et al., 2008) tra le 20.000 parole più frequenti in SUBTLEX-it.

La lunghezza è stata calcolata come numero di lettere nella stringa di lettere e la frequenza delle parole è stata recuperata da SUBTLEX-it e log-trasformata per tener conto della sua distribuzione non simmetrica (come è comune in questo tipo di compiti; Hendrix & Sun, 2021).

Il compito è stato realizzato con Psychopy (Peirce et al., 2019). Ai partecipanti veniva mostrata una parola alla volta e veniva chiesto di indicare se lo stimolo fosse o meno una parola italiana nel modo più veloce e accurato possibile, premendo il tasto A o L rispettivamente con l'indice sinistro o destro. Le prove

sono state presentate in ordine casuale. Ogni prova iniziava con una croce di fissazione centrale (presentata per 500 ms) ed era seguita da una stringa di lettere (presentata per un massimo di 5000 ms). La risposta dei partecipanti terminava la prova e passava a una schermata vuota (presentata per 1000 ms), che precedeva la croce di fissazione della prova successiva.

### **Misure di lettura**

I partecipanti sono stati sottoposti alle prove di lettura appartenenti alla “LSC-SUA: Batteria per la valutazione dei DSA e altri disturbi in studenti universitari e adulti” (Montesano, Valenti e Cornoldi, 2020).

La batteria è composta dal Questionario di autovalutazione Vinegrad Plus e da 17 prove, finalizzate a valutare gli aspetti fondamentali di lettura, scrittura, comprensione del testo e calcolo.

Il questionario di autovalutazione Vinegrad Plus è costituito da 26 item a risposta dicotomica positiva o negativa, che indagano principalmente la presenza di difficoltà in compiti quotidiani in cui è richiesto l’automatismo del processo di letto-scrittura. Vi sono inoltre, item specifici che indagano: difficoltà nell’apprendimento delle lingue straniere, difficoltà di calcolo, infine è presente un item che valuta la correttezza di decodifica.

Le ultime due prove previste, appartenenti alla medesima batteria, sono: la lettura di parole e la lettura di non parole (stringhe di lettere pronunciabili ma inesistenti nella lingua italiana). La prima consiste nella lettura ad alta voce di quattro liste di

parole di diversa lunghezza e frequenza d'uso, delle quali vengono valutate: accuratezza (numero di errori) e velocità (tempo di esecuzione). La prova di lettura di non parole, prevede la lettura ad alta voce di due liste di pseudo-parole di diversa lunghezza; le variabili indagate sono analoghe alla prova precedente.

### **Altre misure impiegate**

In aggiunta alle prove appartenenti alla batteria LSC-SUA, la procedura prevede la somministrazione del *Cattell's Fluid Intelligence Test* (Cattell, 1949).

Si tratta di una prova costituita da 4 subtest denominati rispettivamente: Serie, Classificazione, Matrici e Analogie. Il subtest Serie è composto da una serie incompleta di figure, il partecipante deve scegliere tra 5 opzioni alternative quale inserire per completare la serie; nella prova di Classificazione l'obiettivo è di riuscire ad individuare la figura estranea all'interno di un gruppo di figure; il terzo subtest prevede il completamento di una matrice che presenta una parte mancante, scegliendo tra 5 opzioni; infine, nella prova di Analogie, il partecipante deve scegliere, tra le 5 opzioni, la figura in cui poter inserire il punto, secondo le stesse modalità della figura di riferimento. Poiché si tratta di un test di intelligenza non verbale, è indicato per misurare l'intelligenza indipendentemente dalla fluency verbale, dal contesto culturale e dal livello educativo del soggetto.

Inoltre è stata somministrata la prova di Memoria di cifre, appartenente al test WAIS-IV (Wechsler, 2008). La prova è suddivisa in 3 parti: Ripetizione diretta, Ripetizione inversa e Riordinamento di cifre. Ciascuna delle parti è composta da

due serie di 8 item di difficoltà crescente (per la Ripetizione diretta e il Riordinamento sono previste stringhe numeriche da 2 a 9 cifre, per quella inversa da 2 a 8 cifre). In ciascun item, al partecipante viene letta la serie numerica e viene chiesto di ripeterla nello stesso ordine (ripetizione diretta), in ordine inverso (ripetizione inversa), o pronunciando le cifre in ordine crescente (riordinamento). La procedura è continuata con la somministrazione di: “Tre test clinici di ricerca e produzione lessicale”, a cura di Novelli, Papagno, Capitani, Laiacona, Vallar e Cappa (Novelli, Papagno, Capitani & Laiacona, 1986). Nello specifico, è stata presa in considerazione esclusivamente la prova di Fluenza verbale.

Si tratta di un test di associazione controllata di parole, suddiviso in due parti: fluenza per categorie fonemiche e fluenza per categorie semantiche. Nel primo caso si chiede al partecipante di dire, in un minuto di tempo, il maggior numero di parole possibili, che iniziano con una determinata lettera dell’alfabeto. Le lettere prese in considerazione sono F, P, L.

La seconda prova prevede invece la produzione, da parte del partecipante, di quante più parole possibili appartenenti ad una determinata categoria semantica, nel medesimo arco di tempo (1 minuto). Le categorie considerate sono: marche d’auto, frutti e animali. Il punteggio è dato dalla somma totale delle parole per ciascuna prova.

L’ultima prova prevista consiste in un test di denominazione rapida: *Rapid Automatization Naming*, o RAN (Denckla e Rudel, 1976). Nello specifico abbiamo utilizzato il “Test di denominazione rapida e ricerca visiva di colori, figure e

numeri” a cura di Maria De Luca, Gloria Di Filippo, Anna Judica, Donatella Spinelli, Pierluigi Zoccolotti (De Luca, Di Filippo, Judica, Spinelli & Zoccolotti, 2005). La prova è composta da due parti: RAN di Colori e RAN di Figure, entrambe costituite da una matrice 4x5 di prova preliminare e due matrici 10x5 di test.

L’obiettivo è quello di denominare gli stimoli presenti sulla tavola, svolgendo il compito più accuratamente e più rapidamente possibile. Il punteggio viene calcolato sulla base di due variabili: accuratezza (numero di errori) e velocità (tempo impiegato).

### **Modello semantico distribuzionale**

Il DSM utilizzato è fastText (Joulin et al., 2016). In particolare, abbiamo addestrato un DSM fastText su itWaC (<http://wacky.sslmit.unibo.it>): un corpus di 2 miliardi di parole costruito dal Web limitando il crawl al dominio .it e utilizzando parole a media frequenza del corpus di riviste italiane di Repubblica e liste di vocabolari italiani di base (Baroni et al., 2009).

Il modello è stato addestrato utilizzando il metodo *Continuous Bag of Words* (CBoW), un approccio originariamente proposto da Mikolov e colleghi (2013), con 300 dimensioni e una finestra di co-occorrenza di 5 parole. Quando si utilizza il CBoW, le dimensioni del vettore ottenuto catturano la misura in cui un elemento target è previsto in modo affidabile dai contesti linguistici in cui appare, dove il "contesto" è rappresentato dalle parole contenute in una finestra di dimensioni fisse

intorno alla parola target. In particolare, il modello CBoW indurrà una rappresentazione per un dato target  $w_0$  basata sulle parole di contesto  $w_{-n}, \dots, w_{-1}, w_1, \dots, w_n$ . Abbiamo utilizzato fastText per la sua capacità di calcolare rappresentazioni semantiche che includono anche informazioni sulle sottoparole. Infatti, fastText si basa sull'idea (originariamente proposta da Schütze, 1993; e realizzata computazionalmente da Bojanowski et al., 2017) di prendere in considerazione le informazioni sulle sottoparole inducendo rappresentazioni semantiche come somma dei vettori degli n-grammi di lettere associati a ciascuna parola. In altre parole, fastText calcola la rappresentazione semantica di ogni stringa di lettere come la somma dei vettori dell'intera stringa più tutti i vettori degli n-grammi che la compongono. Il DSM di fastText qui impiegato è stato addestrato includendo n-grammi che vanno da 3 a 6 (Bojanowski et al., 2017). A titolo di esempio, si consideri la parola *<lettura>*, composta da n-grammi di caratteri di lunghezza diversa, come riportato nella Tabella 1.

*Tabella 1: Esempio di vettori di sottoparole recuperati per rappresentare il vettore della parola "lettura".*

Word	Length(n)	Character n-grams
lettura	3	<le, let, ett, ttu, tur, ura, ra>
lettura	4	<let, lett, ettu, ttur, tura, ura>
lettura	5	<lett, lettu, ettur, ttura, tura>
lettura	6	<lettu, lettur, ettura, ttura>

La rappresentazione indotta da fastText sarà quella di un n-gramma di lunghezza diversa. La rappresentazione indotta da fastText sarà la somma del vettore della parola <in memoria> con i vettori di tutti gli elementi riportati nella Tabella 1.

Utilizzando fastText, abbiamo quindi ottenuto le rappresentazioni semantiche per le parole incluse in questo studio e per le 20.000 parole più frequenti incluse nel SUBTLEX-it. Per ogni coppia abbiamo calcolato un indice di similarità semantica (SRel) basato sul coseno dell'angolo formato dai vettori che rappresentano i significati delle stringhe corrispondenti. Più alto è il valore di SRel, più ci si aspetta che le stringhe di lettere siano semanticamente correlate, come stimato dal modello.

### **4.2.3. Procedura**

I partecipanti sono stati testati individualmente nel Laboratorio di Psicologia Cognitiva dell'Università di Pavia.

La prima prova somministrata era il compito di decisione lessicale (vedi paragrafo 4.2.2). A seguire venivano somministrate le prove appartenenti alla Batteria LSC-SUA, nell'ordine: questionario di autovalutazione Vinegrad Plus, lettura di parole e lettura di non parole. Infine veniva somministrata una serie di prove finalizzate ad indagare differenti aspetti relativi all'ambito cognitivo e linguistico: *Cattell's Fluid Intelligence Test*, Memoria di cifre, Tre test clinici di ricerca e produzione lessicale e *Rapid Automatization Naming*, ampiamente descritte nella sezione "Altre misure impiegate" (vedi paragrafo 4.2.2).

### 4.3. Analisi dei dati e risultati

Tutte le analisi sono state effettuate utilizzando R-Studio (RStudio Team, 2015). I dati sono stati analizzati attraverso un approccio a effetti misti, che incorpora sia effetti fissi che effetti casuali (associati ai partecipanti e agli item) e consente di gestire la non indipendenza delle osservazioni sia a livello di partecipanti che di item (Baayen et al., 2008). È stato utilizzato un Modello Misto Lineare (LMM) utilizzando il pacchetto R lme4 (Bates, et al., 2015).

La variabile dipendente era costituita dalle latenze di risposta corretta (RT) dei partecipanti per le parole incluse, che sono state trasformate in log- per tener conto della loro naturale distribuzione non simmetrica (vedi: Gatti et al., 2023). Il LMM includeva la velocità di lettura a livello individuale come predittore continuo e, a livello di prova, SND, OND, frequenza e lunghezza, nonché l'interazione tra la velocità di lettura e questi ultimi predittori. I partecipanti e gli stimoli sono stati inclusi come intercetta random. Più precisamente, il modello stimato è stato:

$$\log RTs \sim (\text{SND} + \text{OND} + \text{frequenza} + \text{lunghezza}) \times \text{velocità di lettura} + (1 | \text{ID}) + (1 | \text{trial})$$

Abbiamo quindi effettuato una selezione del modello utilizzando il pacchetto MuMIn R, con la funzione dredge (Bartoń, 2020). Questa procedura seleziona il modello più adatto (con minor criterio di informazioni di Akaike, che fornisce una stima della qualità del modello, AIC; Akaike, 1973) che si adatta a tutte le possibili combinazioni degli effetti fissi inclusi. Per verificare la coerenza degli effetti,

abbiamo stimato anche tutti i modelli con  $\Delta AIC < 2$ , poiché i modelli in questo intervallo di AIC possono essere considerati equivalenti nello spiegare le prestazioni dei partecipanti (Hilbe, 2011). I trial in cui i partecipanti hanno classificato erroneamente la parola come pseudoparola o in cui i RT erano più brevi di 300 ms o più lunghi di 3000 ms sono stati rimossi dall'analisi (2.2% dei trial rimossi). Il modello migliore comprendeva la velocità di lettura, l'SND, la frequenza, la lunghezza e l'interazione tra velocità di lettura e lunghezza, Pseudo- $R^2$  (marginale) = .22, Pseudo- $R^2$  (totale) = .52. Tutti gli altri fattori fissi sono stati eliminati. Nessun modello aveva un  $\Delta AIC < 2$  rispetto al modello migliore.

*Tabella 2: Risultati ottenuti*

FIXED EFFECT	F-value	NumDF, DenDF	p-value	b
SND	15.59	1,93	< .001	-.29
Frequency	28.45	1,93	< .001	-.03
Length	56.19	1,98	< .001	.04
reading speed	.66	1,77	.42	-.02
reading speed : length	56.30	1,5031	< .001	-.01

I risultati ottenuti, riportati nella Tabella 2, mostrano che l'effetto negativo della SND indica che più alta è la densità del vicinato semantico, più veloci sono i partecipanti, come rappresentato nella figura 7.

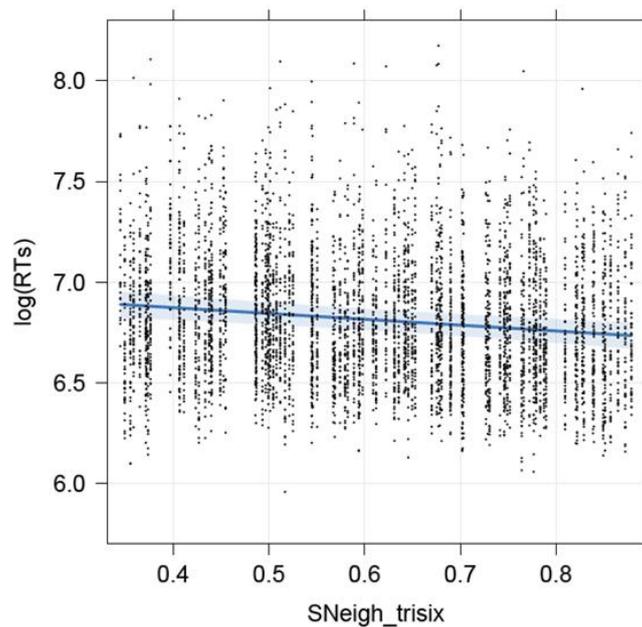


Figura 7: Effetto della variabile SND

La figura 8 mostra che l'effetto negativo della frequenza indica che più alta è la frequenza della parola, più veloci sono i partecipanti. L'effetto positivo della lunghezza indica che più lunga è la parola, più lenti sono i partecipanti.

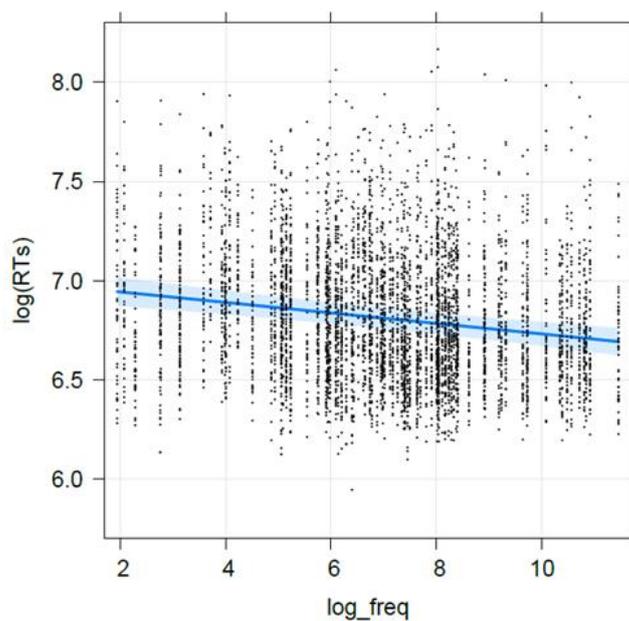


Figura 8: Effetto della variabile frequenza

Infine, da un'ispezione visiva della Figura 9 possiamo dedurre che l'interazione significativa tra velocità di lettura e lunghezza indica che più lunga è la parola, più lenti sono i partecipanti, ma anche che i partecipanti che mostrano tempi di lettura più rapidi sono meno influenzati da questa componente.

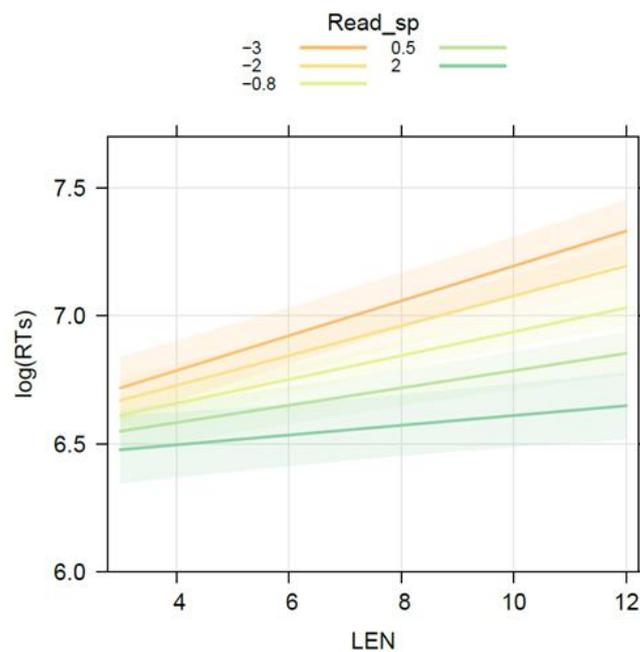


Figura 9: Interazione tra le variabili velocità di lettura e lunghezza

# Capitolo 5: Discussione

## 5.1. Discussione

La capacità di lettura è contraddistinta da una serie di proprietà di varia natura (semantiche, lessicali, ortografiche) che si dimostrano in grado di modellare la prestazione del lettore, influenzandone la velocità e l'accuratezza (Seidenberg, 1990). Il presente progetto di ricerca si è posto come obiettivo principale quello di individuare la presenza di eventuali differenze individuali nelle prestazioni di partecipanti impiegati in una prova di decisione lessicale, indagando la relazione tra fattori semantici e ortografici e i processi di lettura.

La scelta di utilizzare la prova di decisione lessicale come strumento di misurazione è motivata dal fatto che questa tipologia di compito permette di mettere in luce i meccanismi coinvolti durante il processo di lettura, consentendo dunque di indagare le variabili che influenzano la risposta del lettore (Seidenberg, 1990). Come spiegato nei precedenti capitoli, il meccanismo di decisione lessicale coinvolge processi di natura fonologica e lessicale. Considerando la questione in riferimento al modello teorico proposto da Coltheart (*Dual Route Model*, Coltheart et al., 1993), il lettore sembrerebbe percorrere la via lessicale quando lo stimolo verbale presentato corrisponde ad una parola - elaborando le stringhe di lettere in parallelo - mentre la decodifica delle non parole dovrebbe coinvolgere la via sublessicale - adottando un'elaborazione in sequenza mediante il meccanismo di conversione grafema-fonema (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001).

Le variabili prese in considerazione sono state: la velocità di lettura a livello individuale come predittore continuo e, a livello di prova, densità di vicinato semantico (SND), densità di vicinato ortografico (OND), frequenza e lunghezza. Sono state inoltre indagate le relazioni presenti tra la velocità di lettura e i quattro predittori. Nei paragrafi seguenti verranno discussi i risultati emersi dal nostro studio e confrontati con alcune evidenze significative presenti nel panorama scientifico.

### **5.1.1. Densità di vicinato semantico**

La prima variabile presa in considerazione è la densità di vicinato semantico. I risultati ottenuti nel nostro studio hanno mostrato la presenza di una relazione negativa tra SND e le prestazioni dei partecipanti al compito di decisione lessicale (rapidità di classificazione): quanto maggiore è la densità del vicinato semantico delle parole, tanto inferiore sarà il tempo necessario a classificare lo stimolo come parola.

Il contributo della letteratura scientifica (Locker, Simpson & Yates, 2003) offre supporto a questi dati, indicando che tanto più si arricchisce la rappresentazione semantica associata ad una data parola, quanto più rapidamente la parola viene elaborata.

Le evidenze emerse dal nostro studio, trovano ulteriore conferma in una ricerca precedente, condotta da Pexman & Lupker. Gli autori hanno valutato l'impatto esercitato dalla variabile SND sulla prestazione dei partecipanti, in prove di

decisione lessicale e di riconoscimento. Da un'analisi dei risultati ottenuti è stato possibile riscontrare in entrambe le prove una riduzione dei tempi di latenza dei partecipanti nell'elaborazione di parole che presentavano un vicinato semantico più denso, rispetto a quelle con vicinato meno denso. Gli autori, coerentemente ai nostri risultati, affermano dunque che la velocità di riconoscimento ed elaborazione di parole tende ad aumentare in relazione alla densità del loro vicinato semantico (Pexman & Lupker, 1999). I ricercatori hanno inoltre individuato che l'effetto della variabile SND sembra esercitare un impatto significativo anche in relazione alle parole polisemiche (parole che presentano più di un significato). Nello specifico, è stato ipotizzato che le parole ambigue siano dotate di "rappresentazioni semantiche multiple", dunque si presume che dispongano di un vicinato semantico più denso rispetto a parole univoche. Questo potrebbe avvantaggiare il lettore nei compiti di riconoscimento o decisione lessicale (Pexman & Lupker, 1999).

### **5.1.2. Lunghezza**

Il secondo predittore che abbiamo analizzato è la variabile lunghezza. L'effetto positivo rilevato in relazione alla velocità di lettura indica che più lunga è la parola, più lenti sono i partecipanti. Una spiegazione coerente con i risultati da noi ottenuti viene fornita dal Modello a due vie di Coltheart (Coltheart, 1978). I presupposti teorici espressi dal modello definiscono che il processo di lettura implica il coinvolgimento di due meccanismi differenti: uno di tipo sequenziale, fondato

sulla conversione grafema-fonema (impiegato per l'elaborazione di parole non note e non parole) e un meccanismo di elaborazione in parallelo, che prevede l'accesso diretto alla rappresentazione semantica della parola (impiegato per l'elaborazione di parole note al lettore). Dunque, quando viene attuata la procedura sublessicale, il tempo necessario al processo di conversione dovrebbe aumentare in relazione al numero di lettere che compongono la parola (Coltheart et al., 2001). Anche in questo caso, da un confronto con le evidenze presenti nel panorama scientifico, emerge una similarità con i nostri risultati. Uno studio condotto da Balota e colleghi volto ad indagare gli effetti semantici e lessicali sul riconoscimento di parole e non parole, conferma che i partecipanti tendevano ad impiegare più tempo per riconoscere e discriminare le parole e le non parole costituite da più lettere (Balota et al., 2004), coerentemente con la spiegazione fornita dal modello di Coltheart (Coltheart et al., 2001).

In uno studio successivo, Yap e colleghi hanno indagato gli effetti delle caratteristiche semantiche e lessicali che influenzano l'elaborazione di parole e non parole. Analogamente gli autori hanno riscontrato tempi di risposta maggiori per le non parole più lunghe, rispetto a quelle più corte, confermando un'influenza della variabile lunghezza (Yap et al., 2015).

La letteratura scientifica relativa a questo tema (Hendrix e Sun, 2020) offre inoltre spunti che non confermano totalmente i risultati ottenuti nel nostro studio.

Le analisi condotte da Hendrix e Sun mettono in luce un effetto interessante: sembra che inizialmente (appena viene presentato lo stimolo) le parole più lunghe

necessitano di un tempo di elaborazione maggiore rispetto a quelle corte, mentre successivamente (a distanza di alcuni ms dalla presentazione dello stimolo) la direzione dell'effetto si inverte, mostrando una riduzione del tempo di risposta in relazione alla lunghezza della parola. Questo effetto (*U-shaped effect of word length*) elaborato da Baayen e New, sembrerebbe essere motivato dal fatto che le parole lunghe contengono più informazioni sublessicali, che aiutano il lettore a identificare la parola, rispetto alle parole più brevi (Baayen, 2005; New et al., 2006).

### **5.1.3. Frequenza**

Tra i predittori indagati nel nostro studio, effetti significativi sono emersi dai valori riportati dalla variabile frequenza. Nello specifico, la relazione negativa individuata tra frequenza e velocità di lettura indica che quanto maggiore è la frequenza della parola, tanto minore sarà il tempo necessario per elaborarla. I partecipanti del nostro studio si sono dimostrati molto rapidi nell'individuazione di parole ad alta frequenza, confermando le assunzioni di Forster e Chambers, secondo cui i lettori, in compiti di decisione lessicale, elaborano più velocemente parole ad alta frequenza rispetto a quelle a bassa frequenza (Forster & Chambers, 1973).

Uno studio condotto da Keuleers e colleghi fornisce un ulteriore supporto ai risultati ottenuti. Gli autori hanno elaborato un database di tempi di decisione lessicale per parole e non parole inglesi, volto a individuare effetti semantici e

lessicali sottostanti alle procedure di elaborazione degli stimoli, in relazione ai tempi di risposta impiegati dai partecipanti. I dati ottenuti hanno confermato quanto emerso dalle analisi da noi condotte: i partecipanti hanno mostrato risposte più rapide a parole ad alta frequenza, rispetto a quelle a bassa frequenza (Keuleers, et al., 2012).

La variabile frequenza, come si evince dai risultati ottenuti nel nostro studio, si rivela dunque un predittore efficace in relazione alla rapidità di elaborazione delle parole.

Uno studio condotto da Andrews approfondisce questo aspetto, individuando come questa relazione potrebbe non essere valida per le non parole. Gli effetti differenti esercitati dal predittore frequenza sulle parole e sulle non parole sembrano dipendere dalla natura del compito di decisionale lessicale: maggiore è la somiglianza di una non parola ad una parola reale, maggiore è la difficoltà nella decisione (Andrews, 1989).

Risultati differenti relativi a questo aspetto sono emersi in uno studio di Yap e colleghi, i quali hanno individuato come la frequenza delle parole esercitasse un effetto facilitatore sul riconoscimento delle non parole. Nello specifico sembra che le non parole derivate da parole reali con frequenza più elevata vengano riconosciute più rapidamente e accuratamente, rispetto a quelle derivate da parole reali a bassa frequenza (Yap, Balota, Sibley & Ratcliff, 2012).

#### **5.1.4. Differenze individuali**

Ciò che consente di definire un lettore come abile o meno abile, dipende da fattori di varia natura (per una descrizione approfondita vedi De Beni e Pazzaglia 1995).

Nel nostro progetto di ricerca abbiamo scelto di prendere in considerazione la capacità di lettura in un'ottica continua, definendo le differenze individuali tra le abilità dei partecipanti sulla base delle loro prestazioni in relazione ai parametri di rapidità e accuratezza.

La nostra ipotesi di partenza presupponeva che i lettori considerati meno abili si dimostrassero essere più sensibili all'effetto delle variabili semantiche indagate.

I risultati ottenuti mostrano che, mentre alcuni partecipanti si sono rivelati più lenti nel riconoscimento di stimoli lunghi confermando quanto precedentemente esposto (Balota et al., 2004), altri hanno invece mantenuto una buona rapidità di lettura, dimostrandosi meno influenzati da questa componente. Alla luce di questa evidenza, è interessante indagare le potenziali motivazioni sottostanti alle differenze individuali riscontrate.

Studi precedenti hanno dimostrato che sembrerebbe esistere una relazione positiva tra vocabolario (inteso come conoscenza delle forme delle parole e del relativo significato) e abilità di lettura: individui con una maggiore conoscenza del vocabolario si dimostrano essere lettori più veloci e accurati (Butler & Hains, 1979). Gli studi hanno esplorato come la conoscenza del vocabolario possa riflettersi positivamente sull'abilità di lettura, interagendo con il modo in cui le caratteristiche dello stimolo influenzano il riconoscimento (Yap, Balota, Sibley &

Ratcliff, 2012). Analizzando la questione in riferimento alla cornice teorica espressa da Coltheart (Coltheart et al., 2001), Yap e colleghi hanno ipotizzato che i percorsi di lettura che costituiscono il modello a due vie sembrerebbero essere sensibili a proprietà differenti delle parole. Nello specifico, la via lessicale tende ad essere maggiormente influenzata dalla frequenza delle parole, ma non dalla lunghezza, viceversa, la via sublessicale di conversione si rivela maggiormente sensibile alla lunghezza della parola e indipendente dalla frequenza (Yap, Balota, Sibley & Ratcliff, 2012). A partire da questo presupposto e considerando il fatto che il lettore esperto tende ad impiegare sempre più frequentemente meccanismi di lettura lessicali automatici (LaBerge & Samuels, 1974; Stanovich, 1980), gli autori hanno ipotizzato che una maggiore conoscenza del vocabolario sembrerebbe facilitare il riconoscimento delle parole mediante la via lessicale, rendendo quindi il lettore meno sensibile all'influenza della variabile lunghezza. Questo potrebbe rivelarsi coerente con i risultati ottenuti nel nostro progetto relativi alle differenze individuali riscontrate tra lettori in relazione all'effetto lunghezza. Assumendo questo punto di vista, i lettori che si sono rivelati meno sensibili alla lunghezza degli stimoli verbali presentati potrebbero essere dunque considerati lettori più esperti.

## **5.2. Limiti e punti di forza**

Il presente progetto di ricerca ha dimostrato risultati coerenti con le ipotesi di partenza. Le variabili indagate (lunghezza, frequenza, densità di vicinato semantico) si sono dimostrate predittive delle prestazioni dei partecipanti.

Uno dei principali limiti dello studio è rappresentato dall'esiguità numerica del campione coinvolto. Un gruppo limitato di partecipanti coinvolti rende difficile poter generalizzare le conclusioni a una popolazione più ampia.

L'ipotesi iniziale è stata solo parzialmente confermata, in quanto unicamente la variabile lunghezza si è rivelata essere un predittore significativo delle differenze individuali, mentre le altre variabili prese in considerazione non hanno prodotto evidenze significative da questo punto di vista.

Osservando la questione da una prospettiva differente, la scelta di utilizzare come campione sperimentale un gruppo di partecipanti di età compresa tra 19 e 33 anni, costituisce un punto di forza della ricerca in questione. Le indagini e gli studi presenti in letteratura relativi alle capacità di lettura sembrano essere per la maggior parte orientati verso l'età evolutiva, con uno specifico interesse per l'area relativa ai disturbi dell'apprendimento (Fenzi e Cornoldi, 2015). È opportuno precisare che negli ultimi anni si è assistito ad un incremento delle ricerche relative alla valutazione delle difficoltà di apprendimento in età adulta, con particolare attenzione alla fascia d'età che interessa gli studenti universitari (Angelini et al., 2017; Cornoldi e Montesano, 2020). Per questo motivo, la scelta di prendere in esame di un campione costituito da una popolazione di giovani adulti, consente di

rendere il nostro studio un piccolo ma prezioso contributo volto a rafforzare la ricerca scientifica in questa nuova direzione.

Un ulteriore punto di forza è rappresentato dal fatto che, nonostante l'esiguità numerica del campione, i risultati ottenuti si sono dimostrati essere in linea con quelli presenti nel panorama scientifico, confermando che le proprietà semantiche e lessicali sottostanti all'elaborazione delle parole possono modellare la prestazione dei lettori.

### **5.3. Riflessioni personali e prospettive future**

I dati raccolti e analizzati hanno permesso di giungere a delle conclusioni preliminari che meritano di essere ulteriormente approfondite e ampliate.

Si tratta di un progetto potenzialmente in grado di arricchire il panorama della ricerca scientifica, a tal proposito si propone di proseguire l'indagine durante il prossimo anno accademico, al fine di ampliare il numero di partecipanti coinvolti e garantire una maggiore disponibilità di materiale utile al lavoro.

A fronte dei limiti presentati nella sezione precedente, le prospettive future potrebbero dunque includere: il coinvolgimento di un campione di partecipanti più numeroso e omogeneo e l'analisi di alcune variabili che non sono state prese in esame, come ad esempio l'immaginabilità, la concretezza, la coerenza tra ortografia e semantica, al fine di individuare eventuali indici significativi.

Potrebbe inoltre essere auspicabile indagare separatamente l'effetto delle variabili in relazione alle parole e alle non parole, al fine di porre l'attenzione sulle eventuali differenze presenti.

Infine potrebbe rivelarsi interessante riuscire ad estendere il progetto allo scopo di indagare la presenza di eventuali difficoltà specifiche delle abilità di lettura, mettendo a confronto lettori tipici e individui con difficoltà diagnosticate, in un'ottica dicotomica.

Il coinvolgimento di lettori con difficoltà specifiche, a mio avviso, potrebbe rappresentare una potenziale occasione per individuare i punti di forza e le risorse che questi individui possiedono, in vista di fornire loro strategie mirate a compensare le difficoltà.

## Conclusion

La capacità di lettura costituisce un'abilità distribuita in modo eterogeneo all'interno della popolazione. È possibile individuare differenti livelli di competenza lungo la normale curva di distribuzione, definendo così lettori più o meno abili. La posizione occupata da ciascun individuo lungo l'arco di distribuzione riflette le rispettive abilità di lettura, in termini di rapidità e accuratezza (McGowan et al., 2019).

Il filo conduttore della presente tesi è rappresentato dal concetto di "lettura", intesa come capacità acquisita, potenzialmente soggetta a compromissioni o difficoltà e influenzata da una serie di variabili che sono in grado di modellare la prestazione del lettore intervenendo sui processi di varia natura.

È proprio quest'ultimo aspetto che costituisce il focus centrale del progetto di ricerca che abbiamo condotto.

Lo studio si è occupato di indagare la relazione tra fattori semantici e ortografici e processi di lettura. Nello specifico, l'obiettivo primario del lavoro è stato quello di analizzare il ruolo delle componenti semantiche alla base della prestazione in partecipanti con differenti abilità di lettura, impegnati in un compito di decisione lessicale. La nostra ipotesi di partenza presumeva la presenza di una maggiore sensibilità a queste variabili negli individui con minori capacità di lettura.

I risultati che abbiamo ottenuto si sono dimostrati coerenti con alcuni dei dati presenti nel panorama scientifico, confermando che le prestazioni dei lettori sono modellate dagli effetti delle componenti semantiche delle parole. Nello specifico,

abbiamo riscontrato che la densità del vicinato semantico si riflette positivamente sulla capacità di lettura in termini di rapidità, quanto più denso è il vicinato semantico della parola, tanto maggiore si rivela essere la velocità con la quale uno stimolo viene letto .

Anche la frequenza di una parola è correlata positivamente alla velocità di lettura: parole ad alta frequenza tendono ad essere lette più velocemente di quelle a bassa frequenza.

Infine, la lunghezza della parola incide negativamente sulla velocità di lettura: le parole più lunghe richiedono più tempo per essere lette. Tuttavia, non tutti i dati raccolti si sono rivelati conformi a questo risultato in quanto alcuni partecipanti non hanno mostrato un aumento significativo dei tempi di lettura in relazione a parole lunghe, dimostrandosi meno sensibili all'effetto di questa variabile.

Questi valori contribuiscono a fornire una parziale conferma all'ipotesi iniziale: i lettori meno rapidi si sono rivelati essere maggiormente esposti all'effetto della variabile lunghezza, rispetto a quelli più abili che hanno invece dimostrato una minore sensibilità.

Assumendo un punto di vista più ampio, la coerenza con i presupposti di partenza e la conformità alle evidenze scientifiche già presenti in letteratura, consentono di rendere questo studio un piccolo ma prezioso contributo potenzialmente utile alla ricerca scientifica in questa direzione. Si tratta di un progetto destinato a continuare, che merita di essere ampliato e approfondito, in vista di poter ottenere

risultati utili a rafforzare le ipotesi presentate ed eventualmente individuare ed indagare nuovi spunti potenzialmente interessanti.

## Bibliografia

- Acha J. e Perea M. (2008), The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers, *British Journal of Psychology*, vol. 99, pp. 245-264. (s.d.).
- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 60(2), 255-265. (s.d.).
- Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 802.
- Aroldi. (2010). Elementi di psicologia del linguaggio. Edizioni Simone.
- Baayen, R. H. (2005). Data mining at the intersection of psychology and linguistics. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 69– 83). Hillsdale, New Jersey: Erlbaum.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4), 390-412. (s.d.).
- Balota DA, Cortese MJ, Sergent-Marshall S, Spieler DH, Yap MJ. (2004). Visual word recognition of single syllable words. *Journal of Experimental Psychology: General.*; 133:336–345.
- Barca, L., Burani, C. & Arduino, L.S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers* 34, 424–434 (s.d.).
- Baroni, M., Pirrelli, V., & Guevara, E. (2009). Sulla tipologia dei composti N+ N in italiano: Principi categoriali ed evidenza distribuzionale a confronto, 1000-1023. (s.d.).
- Barto'n K. MuMIn: Multi-model inference. (2020) v.1.43.17. (s.d.).
- Basile, G. (2001). *Le parole nella mente: Relazioni semantiche e struttura del lessico*. FrancoAngeli.

- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, 12(1), 2. (s.d.).
- Beaton, A., McDougall, S., & Singleton, C. (1997). Editorial Humpty Dumpty Grows Up?—Diagnosing Dyslexia in Adulthood. *Journal of Research in Reading*, 20(1).
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5, 135-146. (s.d.).
- Bonk, W. J. (2000). Second Language Lexical Knowledge and Listening Comprehension. *International Journal of Listening*, 14(1), 14–31. (s.d.).
- Brizzolaro, D., Gasperini, F., & Mazzotti, S. (2007). Modelli neuropsicologici della dislessia evolutiva. *Giornale di Neuropsichiatria dell'Età Evolutiva*, 27, 229-243. (s.d.).
- Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Ageof-acquisition effects in semantic processing tasks. *Acta Psychologica*, 104(2), 215-226. (s.d.).
- Buchanan, Lori & Westbury, Chris & Burgess, Curt. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic bulletin & review*. 8. 531-44. 10.3758/BF03196189. (s.d.).
- Burani, C., Barca L., & Arduino L. S. (2001). Una base di dati sui valori di età di acquisizione, frequenza, familiarità, immaginabilità, concretezza, e altre variabili lessicali e sublessicali per 626 nomi dell'italiano. *Giornale Italiano di Psicologia* 28:839–56
- Butler B, Hains S. (1979). Individual differences in word recognition latency. *Memory & Cognition*. 7:68–76.
- Cardona, M. (2008). L'abilità di lettura e lo sviluppo della competenza lessicale. *Studi di Glottodidattica*, 2, 10-36.
- Caldarola, N., Perini, N., & Cornoldi, C. (2012). DLC. Una prova di decisione lessicale per la valutazione collettiva delle abilità di lettura. *Dislessia*, 9(1), 89-104. (s.d.).

- Camaioni L., *Psicologia dello sviluppo del linguaggio*, Bologna, Il Mulino (2001). (s.d.).
- Cangelosi, A., & Turner, H. (2002). *L'emergere del linguaggio*. Scienze della Mente.
- Catherine E. Snow (1991) *The Theoretical Basis For Relationships Between Language and Literacy in Development*, *Journal of Research in Childhood Education*, 6:1, 5-10.
- Cattell, R. B. (1949). Culture free intelligence test, scale 1, handbook. Champaign, IL: *Institute of Personality and Abilit.*
- Chomsky, N. (1959). Review of Skinner's Verbal Behavior. *Language*, 35, 26-58. (s.d.).
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, 108(1), 204–256. (s.d.).
- Daloiso, M. (2009). La dislessia evolutiva: Un quadro linguistico, psicolinguistico e glottodidattico. *Studi di glottodidattica*, 3(3), 25-43. (s.d.).
- D'Amico, S., Devescovi, A., & Bates, E. (2001). Picture naming and lexical access in Italian children and adults. *Journal of Cognition and Development*, 2(1), 71–105.
- Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology*, 41 (1), 60–74.
- De Beni R., & Pazzaglia F. (1995). *La comprensione del testo. Modelli teorici e programmi di intervento*. Torino: UTET.
- De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Official journal of the Society for Behavioral and Cognitive Neurology*, 21(4), 227–235.
- De Luca, M., Di Filippo, G., Judica, A., Spinelli, D., & Zoccolotti, P. (2005). Test di denominazione rapida e ricerca visiva di colori, figure e numeri. *Roma: IRCCS Fondazione Santa Lucia*.

- De Martino, M., Mancuso, A., & Laudanna, A. (2017). Variabili Rilevanti nella Rappresentazione delle Parole nel Lessico Mentale: Dati Psicolinguistici da una Banca-Dati di Nomi e Verbi Italiani. *CLiC-it 2017 11-12 December 2017, Rome*, 125.
- Di Vincenzo F. & Manzi G. (2013). L'origine darwiniana del linguaggio. In *Sull'origine del linguaggio e delle lingue storico-naturali*, ed. E. Banfi, 217-247. Roma: Bulzoni. (s.d.).
- Cortese, M. J., Toppi, S., Khanna, M. M., & Santo, J. B. (2020). AoA effects in reading aloud and lexical decision: Locating the (semantic) locus in terms of the number of backward semantic associations. *Quarterly Journal of Experimental Psychology*, 73(11), 2036-2044.
- Fantoni, P. E. (2016). I disturbi specifici dell'apprendimento.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Memory and Language*, 12 (6), 627–635.
- Gatti, D., Anceresi, G., Marelli, M., Vecchi, T., & Rinaldi, L. (2024). Decomposing geographical judgments into spatial, temporal and linguistic components. *Psychological Research*, 1-12. (s.d.).
- Gatti, D., Marelli, M., & Rinaldi, L. (2023). Out-of-vocabulary but not meaningless: Evidence for semantic-priming effects in pseudoword processing. *Journal of Experimental Psychology: General*, 152(3), 851. (s.d.).
- Hendrix, P., & Sun, C. C. (2021). A word or two about nonwords: Frequency, semantic neighborhood density, and orthography-to-semantics consistency effects for nonwords in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(1), 157. (S.d.). (s.d.).
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press. (s.d.).
- Jasińska, K. K., Shuai, L., Lau, A. N. L., Frost, S., Landi, N., & Pugh, K. R. (2021). Functional connectivity in the developing language network in 4-year-old children predicts future reading ability. *Developmental science*, 24(2), e13041. (s.d.).

- Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., & Mikolov, T. (2016). Fasttext. Zip: Compressing text classification models. arXiv preprint arXiv:1612.03651. (s.d.).
- Kemp, N., Parrila, R. K., & Kirby, J. R. (2009). Phonological and orthographic spelling in high-functioning adult dyslexics. *Dyslexia*, *15*(2), 105-128.
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, *42*, 627-633. (s.d.).
- Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior research methods*, *44*, 287-304.
- Kumar A. A. (2021). Semantic memory: A review of methods, models, and current challenges. *Psychonomic bulletin & review*, *28*(1), 40–80.://doi.org/10.3758/s13423-020-01792-x. (s.d.).
- Kuo, N. C., Wahl, H., McKinney, W., Thompson, M., Roberson, N., Mattison, T., & Ray, M. (2024). Assessing students with dyslexia and other reading difficulties through multiple diagnostic assessments. *Dyslexia* (Chichester, England), *30*(3), e1777.
- LaBerge D, Samuels SJ. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*.; *6*:293–323.
- Lefly, D. L., & Pennington, B. F. (1991). Spelling errors and reading fluency in compensated adult dyslexics. *Annals of dyslexia*, *41*, 141-162.
- Legum, S. E., & Cronnell, B. A. (1973). A Psycholinguistic Analysis of Reading and Learning to Read, by F. Smith. *Review of Understanding Reading*, *49*(2), 523–528.
- Lena L., Pinton A., Trombetti B., (2004). Valutare linguaggio e comunicazione, Manuale per logopedisti e psicologi. Roma, Carocci (s.d.).
- Locker, L., Simpson, G., & Yates, M. (2003). Semantic neighborhood effects on the recognition of ambiguous words. *Memory & cognition*, *31*, 505–515. <https://doi.org/10.3758/BF03196092>
- Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, *4*, 5–17.

- Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of Orthography-Semantics Consistency on word recognition. *Quarterly Journal of Experimental Psychology*, 68(8), 1571-1583.
- Marinelli, C. V., Traficante, D., Zoccolotti, P., & Burani, C. (2013). Orthographic Neighborhood-Size Effects on the Reading Aloud of Italian Children With and Without Dyslexia. *Scientific Studies of Reading*, 17(5), 333–349.
- Marini, A. (2021). La comunicazione e il linguaggio: Una panoramica di strutture e processi. In Marotta, L. e Paloscia, C. (a cura di). *Il disturbo socio-pragmatico comunicativo*. Erickson, Trento (pagg. 13-22). (s.d.).
- McCarthy, K. M., & Skoruppa, K. (2023). Language-specific phonological skills and the relationship with reading accuracy in Sylheti-English sequential bilinguals. *Child development*, 94(2), e85–e102.
- McGowan, D., Little, C., Coventry, W., Corley, R., Olson, R., Samuelsson, S., & Byrne, B. (2019). Differential Influences of Genes and Environment Across the Distribution of Reading Ability. *Behavior Genetics*, 49.
- Melero Rodriguez, C. (2012). Dislessia evolutiva. *EL.LE*, 1, 133–165.
- Mikolov, T., Chen, K., Corrado, G.S., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *International Conference on Learning Representations*. (s.d.).
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior research methods*, 46, 887-903. (s.d.).
- Montesano L., Casadio M., Fresu E., Frinco M., De Marco E., Palmieri R. & Cornoldi C. (2023). I DSA nell’adulto: nuove analisi su un campione di 519 adulti non universitari di età compresa fra i 18 ei 66 anni.
- New B, Ferrand L, Pallier C, Brysbaert M. (2006). Re-examining word length effects in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*. 13:45–52.
- Novelli, G., Papagno, C., Capitani, E., & Laiacona, M. (1986). Tre test clinici di ricerca e produzione lessicale. Taratura su sogetti normali. *Archivio di psicologia, neurologia e psichiatria*.

- Paizi, D., De Luca, M., Zoccolotti, P. and Burani, C. (2013), A comprehensive evaluation of lexical reading in Italian developmental dyslexics. *Journal of Research in Reading*, 36: 303-329.
- Palumbo T.J. (2008), Unit of word recognition as an indicator of reading fluency: Use of a lexical decision task, Dissertation abstracts international, section B, *The sciences and engineering*, vol. 68(7-B), p. 4859. (s.d.).
- Pearson P., & Cervetti G. (2012). *The Psychology and Pedagogy of Reading Processes* (pp. 507–554).
- Peirce J., Gray J. R., Simpson S., MacAskill M., Höchenberger R., Sogo H. & Lindeløv J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, 51, 195-203. (s.d.).
- Pellizzari, M. (2009). *Disturbi Evolutivi Specifici di Apprendimento*. Erickson.
- Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).
- Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 53(4), 323.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it?. *Cognition*, 95(2), 201–236. 2004.08.004. (s.d.).
- Rayner K, Clifton C Jr. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biol Psychol.* 80(1):4-9.
- Rayner, K., Schotter, E. R., Masson, M. E., Potter, M. C., & Treiman, R. (2016). So Much to Read, So Little Time: How Do We Read, and Can Speed Reading Help?. *Psychological science in the public interest: A journal of the American Psychological Society*, 17(1), 4–34.
- Schindler A., Ginocchio D., (2014). *Elementi di Fonetica e Linguistica*, Da Ambrosetti U., Di Bernardino F., Del Bo L., Torino, Minerva Medica (s.d.).

- Schütze, H., & Pedersen, J. (1993). A vector model for syntagmatic and paradigmatic relatedness. In Proceedings of the 9th Annual Conference of the UW Centre for the New OED and Text Research (pp. 104-113). (s.d.).
- Sears CR, Siakaluk PD, Chow VC, Buchanan L. (2008). Is there an effect of print exposure on the word frequency effect and the neighborhood size effect? *Journal of Psycholinguistic Research.*; 37:269–291.
- Seidenberg, MS. (1990) Lexical access: Another theoretical soupstone?. In: Balota, D.; Flores D’arcais, G.; Rayner, K., editors. Comprehension processes in reading. Hillsdale, NJ: Erlbaum;
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *The New England journal of medicine*, 326(3), 145–150.
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., ... & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics*, 104(6), 1351-1359.
- Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., ... & Gore, J. C. (2003). Neural systems for compensation and persistence: young adult outcome of childhood reading disability. *Biological psychiatry*, 54(1), 25-33.
- Stanovich KE. (1980) Towards an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*. 16:32–71.
- Temple C. M. (2006). Developmental and acquired dyslexias. *Cortex; a journal devoted to the study of the nervous system and behavior*, 42(6), 898–910.
- Vallar, G., & Papagno, C. (2018). *Manuale di neuropsicologia. Clinica ed elementi di riabilitazione*. Il mulino.
- Wechsler, D. (2008). Wechsler Adult Intelligence Scale—Fourth Edition Administration and Scoring Manual. San Antonio, TX: Pearson.

- Wiseheart, R., & Altmann, L. J. (2018). Spoken sentence production in college students with dyslexia: Working memory and vocabulary effects. *International journal of language & communication disorders, 53*(2), 355-369.
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic bulletin & review, 15*(5), 971-979. (s.d.)
- Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 53.
- Yeatman, J. D., & White, A. L. (2021). Reading: The Confluence of Vision and Language. *Annual review of vision science, 7*, 487–517. (s.d.).
- Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Spinelli, D. (2005). Prova di lettura di parole e non parole. *Rome, Italy: IRCCS Fondazione Santa Lucia.*