



Studi e Ricerche



# Tracciare la comprensione

*L'eye tracking nella ricerca linguistica*

Maria Roccaforte e Veronica D'Alesio



University Press



SAPIENZA  
UNIVERSITÀ EDITRICE



Collana Studi e Ricerche 173



# Tracciare la comprensione

*L'eye tracking nella ricerca linguistica*

Maria Roccaforte e Veronica D'Alesio



SAPIENZA  
UNIVERSITÀ EDITRICE

2025

Volume pubblicato grazie a un finanziamento della Sapienza  
Università di Roma (Progetti di ricerca medi 2018-RM11816436393D70),  
erogato dal Dipartimento di Lettere e culture moderne.

Copyright © 2025

**Sapienza Università Editrice**  
Piazzale Aldo Moro 5 – 00185 Roma  
www.editricesapienza.it  
editrice.sapienza@uniroma1.it

Iscrizione Registro Operatori Comunicazione n. 11420  
*Registry of Communication Workers registration n. 11420*

ISBN 978-88-9377-387-4

DOI 10.13133/9788893773874

Pubblicato nel mese di luglio 2025 | *Published in July 2025*



Opera distribuita con licenza Creative Commons Attribuzione –  
Non commerciale – Non opere derivate 3.0 Italia e diffusa in modalità  
open access (CC BY-NC-ND 3.0 IT)

*Work published in open access form and licensed under Creative Commons Attribution – NonCommercial –  
NoDerivatives 3.0 Italy (CC BY-NC-ND 3.0 IT)*

Impaginazione a cura di | *Layout by:* Daniele Priori

In copertina | *Cover image:* immagine rielaborata con AI (ChatGPT Images Generator) e Photoshop,  
a partire da materiale di pubblico dominio.

# Indice

Introduzione	7
1. Cosa è l'eye tracker	11
1.1. Hardware	11
1.2. Software	14
1.2.1. Experiment Builder e Data Viewer	14
1.2.2. Tobii Pro Lab (Tobii)	17
1.2.3. Be Gaze (SMI)	18
1.2.4. Software open source	19
1.3. Condizioni sperimentali nella ricerca con eye tracker	20
1.3.1. Luce	21
1.3.2. Spazio	22
1.3.3. Tempo	23
2. Cosa misura un eye tracker	25
2.1. Le metriche	25
2.1.1. Fissazioni	26
2.1.2. Saccadi	27
2.2. Gli strumenti di visualizzazione	28
2.3. Le aree di interesse	29
2.4. Early e late measures	30
2.4.1. Early Measures (Misure Precoci)	31
2.4.2. Late Measures (Misure Tardive)	32
2.5. Metriche e strumenti di analisi negli studi su letteratura e comprensione	34

3. Reading	37
3.1. Introduzione ai modelli di lettura	37
3.1.1. E-Z Reader	39
3.1.2. SWIFT	42
3.1.3. Glenmore	44
3.2. Variabili linguistiche	45
3.3. Elementi di base di un esperimento di reading	56
4. Scene perception e visual search	61
4.1. Scene perception	61
4.2. Visual search	64
4.2.1. Movimenti oculari nella visual search	67
5. Visual world paradigm	69
5.1. Integrazione di audio e video	69
5.2. Studi principali	71
5.3. Elementi di base dell'esperimento in visual world	78
6. L'eye tracking nella Second language research	83
6.1. Eye tracking e meccanismi cognitivi dell'apprendimento linguistico	83
6.2. Aree di ricerca e stimoli nella L2 con eye tracker	85
6.3. L'uso degli eye tracker nello studio dei manuali di lingua: stato dell'arte e prospettive future	87
6.4. Comprensione multimodale e linguaggio bilingue	88
6.5. Conclusioni e direzioni future	88
7. La metodologia eye tracking nella ricerca linguistica in italiano	91
7.1. Studi sul lessico	91
7.2. Studi di stampo pragmatico	93
7.3. Studi in ambito sintattico	96
7.4. Studi con popolazioni non tipiche	97
7.5. Processing bilingue in età adulta	97
7.6. Studi sul processing in età evolutiva	98
Appendice I	101
Appendice II	103
Bibliografia	107

# Introduzione

Molti dei volumi e degli articoli prodotti sull'eye tracking introducono l'argomento evocando una metaforica e suggestiva "finestra sulla mente". In effetti, l'eye tracker è un dispositivo che permette di monitorare e registrare i movimenti degli occhi mentre osservano un'immagine, leggono o svolgono un'attività di *problem solving*, dando così l'impressione di spiare, attraverso la "finestra" degli occhi, il percorso che la mente suggerisce allo sguardo.

A partire da questo presupposto, lo strumento è stato largamente usato, anche in ambito scientifico, per studiare il comportamento visivo e la lettura e ha avuto nel corso degli ultimi trent'anni applicazioni pratiche nei più vari settori dell'industria, del commercio e della comunicazione: dalla progettazione di interfacce uomo-macchina alla pubblicità, dalla sicurezza allo sport, dal design all'accessibilità, solo per citarne alcuni.

Già dal diciannovesimo secolo si sono fatti strada studi sul movimento oculare che prevedevano osservazioni dirette. È stato per primo Louis Émile Javal nel 1879 a sostenere che la lettura non prevedesse un movimento lineare e continuo degli occhi lungo il segmento del testo, ma una serie di brevi interruzioni e rapidi salti.

Edmund Huey, all'inizio del ventesimo secolo costruisce quello che potrebbe essere considerato il primo eye tracker della storia, usando una lente a contatto con un foro per la pupilla. L'obiettivo della lente era collegato a un puntatore di alluminio che si muoveva in risposta al movimento dell'occhio; in questo modo Huey aveva potuto quantificare empiricamente i movimenti oculari progressivi e regressivi dimostrando, per la prima volta, la non linearità del *pattern* di lettura.

I primi eye tracker non intrusivi sono nati poco più tardi a Chicago e utilizzavano fasci di luce che, riflettendo sull'occhio, registravano il loro percorso su una pellicola. Buswell è ancora oggi citato per i suoi studi sistematici sulla lettura e sulla esplorazione di immagini<sup>1</sup>.

Negli anni Cinquanta del secolo scorso Alfred L. Yarbus era riuscito a dimostrare come il compito assegnato a un soggetto abbia un'influenza considerevole sul movimento degli occhi del soggetto. Qualche anno dopo, la ricerca sul tracciamento oculare, e in particolare sulla lettura, aveva definitivamente preso il via con i lavori, sempre più dettagliati e accurati, di Rayner prima e di Just & Carpenter poco dopo. È proprio in questi anni che viene elaborata la *Eye-mind link theory* (Just and Carpenter 1980) secondo cui "c'è netta corrispondenza tra ciò che è fissato e ciò che è elaborato a livello cognitivo"; in altre parole, quando un soggetto guarda una parola o un oggetto, ci pensa anche (processo cognitivo) e quel processo è misurabile nella "finestra" (temporale e metaforica) della fissazione dell'occhio. L'ipotesi era sempre stata data per scontata da coloro che avevano utilizzato il tracciamento oculare come tecnica per le indagini sperimentali; tuttavia, dagli anni Ottanta è stata messa in discussione alla luce della *Covert Attention Hypothesis* (Teoria dell'attenzione implicita), secondo la quale è possibile volgere l'attenzione a qualcosa che non si sta guardando e anche guardare qualcosa senza prestarle effettivamente attenzione (Reichle 2006). Se l'attenzione nascosta è comune durante le registrazioni del tracciamento oculare, il risultante percorso di scansione e i modelli di fissazione mostrerebbero allora non dove è stata l'attenzione, ma solo dove l'occhio ha guardato, non riuscendo a indicare in alcun modo l'elaborazione cognitiva. La finestra sulla mente citata poc'anzi, in questo caso, somiglierebbe a una "finestra cieca".

Per tutte queste ragioni il tracciamento oculare è spesso accompagnato, nei disegni sperimentali, da altre metodologie di ricerca, come ad esempio i protocolli verbali introspettivi o le indagini di comprensione del testo, di cui torneremo a parlare più avanti, quando affronteremo il tema della costruzione di un disegno sperimentale nel campo delle ricerche linguistiche e glottodidattiche.

---

<sup>1</sup> How People Look at Pictures, by Guy Thomas Buswell, in *Parnassus*, vol. 7, n. 7, 1<sup>o</sup> dicembre 1935, pp. 32-32, DOI:10.1080/15436314.1935.11467398. URL consultato il 3 marzo 2023. G.T. Buswell, *How people look at pictures: a study of the psychology and perception in art*, University of Chicago Press, 1935, Trove 12223957.

Per il momento ci limiteremo a descrivere come sono fatti i moderni eye tracker dal punto di vista strutturale, quali sono i modelli ad oggi disponibili e le funzionalità dei diversi software a seconda delle varie condizioni sperimentali che possono essere attuate nell'ambito di un laboratorio di neuro-psicolinguistica.



# 1. Cosa è l'eye tracker

*Maria Roccaforte*

## 1.1. Hardware

Per comprendere come è fatto un eye tracker occorre prima di tutto passare in rassegna i principali tipi di hardware che possono essere impiegati nell'ambito di ricerche di matrice linguistica. Saranno classificati prima sulla base della tecnologia che sfruttano e poi sulla base della loro forma e funzionalità.

*Tecnologia Infra-red/Near-Infra-red:* Questo tipo di eye tracker utilizza una tecnologia che emette una leggera luce infrarossa invisibile all'occhio umano che colpisce l'occhio dell'utente e registra il riflesso di questa luce sulla superficie dell'occhio (la cornea o il cristallino). Misurando il punto esatto in cui il raggio di luce infrarossa ha colpito l'occhio è possibile ricostruire precisamente il percorso compiuto dallo sguardo con un livello di accuratezza che dipende dalla potenza e dalla qualità dello strumento e che si misura in Hertz. L'"Hertz" (Hz) è un'unità di misura che indica la frequenza di campionamento dell'occhio. Rappresenta, cioè, il numero di volte al secondo in cui l'eye tracker ha acquisito e registrato dati rispetto alla posizione dello sguardo dell'utente. Ad esempio, un eye tracker con una frequenza di campionamento di 300 Hz acquisisce e registra dati sulla posizione dello sguardo 300 volte al secondo. La potenza e precisione di un eye tracker *near-infrared*, però, può essere misurata anche con altre unità di misura, a seconda del parametro specifico che si sta valutando. Tra le unità di misura utilizzate per decretare la precisione di un eye tracker ricordiamo anche quelle in 1) gradi visivi ( $^{\circ}$ ), che rappresentano l'ampiezza del campo visivo in cui l'eye tracker è in grado di individuare lo sguardo dell'utente. Una precisione più elevata corrisponde a un angolo visivo più stretto; 2) in

Pixel sullo schermo (Px): questa unità dà conto della precisione con cui l'eye tracker può determinare la posizione esatta dello sguardo dell'utente rispetto ai pixel sullo schermo del dispositivo; 3) in Millimetri (mm), che rappresentano la distanza lineare tra la posizione registrata dall'eye tracker e la posizione reale dello sguardo dell'utente.

Nelle ricerche di area linguistica quella in Hertz è la misura più significativa, perché una frequenza di campionamento più elevata permette una maggiore precisione nella registrazione dei movimenti oculari entro porzioni limitate di input, e perché consente di catturare variazioni rapide nella direzione dello sguardo e nei continui cambiamenti di posizione degli occhi che caratterizzano la lettura di un testo. Questo è particolarmente importante in applicazioni che richiedono una precisione elevata, come gli studi sulla lettura, la navigazione visiva o la ricerca nella percezione visiva.

*Telecamera ad alta velocità e sensori elettro-oculografici:* Oltre agli hardware di tipo near-Infrared, vale la pena menzionare anche gli altri modelli di eye tracker, anche se nella linguistica applicata sono stati usati assai più raramente; ad esempio, la tecnologia con telecamera ad alta velocità e quella a sensori elettro-oculografici. Il primo approccio non richiede l'uso di luce infrarossa e, a differenza dei *near-infrared*, può funzionare con una grande varietà di illuminazioni ambientali. La seconda tecnologia si basa sulla misurazione dei cambiamenti nel campo elettrico generato dagli occhi durante i movimenti. Sebbene non richieda l'uso di luce infrarossa o telecamere ad alta velocità, può essere meno precisa rispetto alle altre tecnologie e più soggetta a interferenze da altri segnali elettrici. Si tratta di una scelta conveniente nelle situazioni in cui la precisione non è fondamentale e si desidera un sistema più semplice e meno costoso.

Come dicevamo, negli ultimi anni, la ricerca ha sfruttato assai più largamente i lettori near-Infrared perché garantiscono un'accuratezza più alta rispetto a tutti gli altri e perché sono meno invasivi e dunque permettono una più agile impostazione del protocollo sperimentale.

Passando dalla tecnologia alla conformazione strutturale dell'hardware, i modelli di eye tracker attualmente in commercio sono classificabili in quattro categorie, con qualche differenza tra modelli e case produttrici.

*Screen-based:* I primi, e i più diffusi nella ricerca sulla lettura a schermo, sono gli eye tracker di tipo *screen-based*, ovvero montati su schermo. Monitorano il movimento degli occhi utilizzando i sensori *near-infrared* posizionati sopra o sotto al display. Possono funzionare

utilizzando diverse tecniche, tra cui il tracciamento della pupilla o il tracciamento della cornea.

*Glasses/Head-Mounted:* Questi eye tracker sono integrati in occhiali o dispositivi indossabili, come caschi. Solitamente includono telecamere o sensori di tracciamento degli occhi posizionati strategicamente per monitorare il movimento degli occhi dell'utente. Consentono il monitoraggio del comportamento visivo degli utenti in modo più naturale e in condizioni più realistiche, poiché gli utenti possono muoversi liberamente mentre indossano gli occhiali o il casco. Questi dispositivi sono spesso utilizzati in ricerche sul campo dove è necessario garantire la maggiore mobilità e naturalezza possibili. Trovano quindi applicazione in indagini di tipo commerciale, sportivo, e sociologiche.

*Barre eye tracking:* Il terzo tipo si presenta come una barra che, oltre a garantire la raccolta dei dati che compaiono su uno schermo digitale, consente di monitorare la lettura su carta in ambienti controllati. Può essere posizionato su un tavolo o su un supporto per registrare i movimenti oculari mentre un individuo legge un documento cartaceo o mentre esplora un qualunque contenuto analogico o digitale. Si tratta di dispositivi leggeri e facili da trasportare, dunque molto adatti per condurre ricerche su larga scala.

*Visore per realtà virtuale e aumentata:* È la quarta e più recente frontiera dell'hardware eye tracker. Si tratta di strumenti in grado di combinare il tracciamento oculare con le potenzialità del visore.

Anche se in questo volume non li tratteremo, esistono ovviamente altri hardware che sono usati in ambito clinico, riabilitativo e nel gaming. Ciascuno dei principali produttori di eye tracker a livello mondiale ha sviluppato expertise molto specifiche nei vari settori dell'industria e della ricerca. Allo stato attuale, nella ricerca di stampo linguistico (come riportano i dati della rassegna pubblicata su *Educational sciences* (Hu, Aryadoust, 2024) oggi sono soprattutto tre: "SR Research EyeLink"; "Tobii"; "SensoMotoric Instruments (SMI)" le principali aziende che sviluppano eye tracker nel campo della ricerca psicolinguistica, linguistica e glottodidattica e che forniscono i laboratori di ricerca e le università di tutto il mondo che si occupano di questo settore di studi<sup>2</sup>.

---

<sup>2</sup> Tra le principali case produttrici di eye tracker ricordiamo anche EyeTech Digital Systems, Gazepoint, Smart Eye, ISCAN, EyeNTNU, FaceLAB, LC Technologies e Pupil Dev.

## 1.2. Software

Numerosi sono i software che possono essere usati quando si conduce una ricerca di impianto linguistico con eye tracker, tuttavia, in questo paragrafo passeremo in rassegna i tre principali software che accompagnano le strumentazioni delle tre più grandi case produttrici di hardware (SR Research Eye Link, Tobii e SMI) maggiormente implicate negli studi di matrice linguistica; una sezione a parte sarà dedicata ai software non a pagamento (open source) del tutto sganciati dalle case produttrici di eye tracker<sup>3</sup>.

### 1.2.1. Experiment Builder e Data Viewer

**Experiment Builder** serve a creare gli esperimenti, mentre **Data Viewer** ad analizzare i dati raccolti. Entrambi i software sono sviluppati dalla stessa azienda e sono progettati per lavorare insieme.

Experiment Builder è un software che permette di progettare esperimenti di ricerca con l'EyeLink, che è l'hardware della SR Research Lab. Offre un'interfaccia utente intuitiva per la creazione di paradigmi sperimentali complessi, in cui è possibile programmare sequenze di eventi, registrare input dei partecipanti e controllare vari dispositivi contemporaneamente. Lo schema dell'esperimento si può disegnare tramite un'interfaccia "drag and drop" con cui, attraverso l'aggiunta di azioni e trigger, si progetta un diagramma di flusso che piloterà la prova. Dentro ciascun blocco di questo diagramma devono essere inclusi a mano, dal ricercatore, tutti gli elementi necessari per l'esecuzione dello studio. All'interno di ogni blocco, i "nodi" rappresentano ciascuna delle funzioni che si prevede di includere, ovvero ogni azione che vogliamo far eseguire al software durante l'esperimento.

Se è necessaria una domanda di comprensione nel corso di un esperimento, quindi, si dovrà aggiungere una schermata di visualizzazione entro i nodi del blocco di registrazione, tra uno stimolo e l'altro. In questo caso si renderà necessaria una seconda serie di trigger, con il trigger della tastiera specificato per consentire solo determinate risposte (ad esempio, i tasti "y" e "n" per consentire ai partecipanti di rispondere a una domanda "sì/no").

---

<sup>3</sup> Pare doveroso segnalare che non sono infrequenti i casi di ricerche nelle quali il team stesso ha sviluppato un software per l'analisi in modo autonomo (user-developed software).

Qualsiasi azione aggiuntiva viene aggiunta alla struttura nello stesso modo. Ad esempio, per aggiungere la procedura di impostazione e calibrazione dell'eye tracker è sufficiente aggiungere un nodo "CAMERA SETUP" nel punto appropriato. Questo nodo di impostazione deve essere aggiunto all'inizio di ogni esperimento che costruiamo e, inserendolo in questo punto piuttosto che nel blocco "Esperimento", la procedura di calibrazione sarà automaticamente inclusa prima di ogni blocco di prove.

Ciò significa che, se nel corso di un esperimento linguistico si dividono gli stimoli in blocchi separati, ognuno di essi includerà automaticamente una procedura di impostazione e calibrazione prima di mostrare la serie successiva di stimoli.

Gli stimoli testuali vengono aggiunti utilizzando lo strumento *Text Resource* (per la lettura su una sola riga) o *Multiline Text Resource* (per la lettura su più righe). L'editor di testo permette di digitare il testo e di impostare proprietà come il tipo di carattere, la dimensione e la posizione sullo schermo. Gli studi basati sulle immagini vengono creati utilizzando lo strumento "Risorsa immagine", selezionando il file desiderato dalla libreria e impostando poi le proprietà come larghezza, altezza e posizione sullo schermo.

In alternativa, il testo (con o senza immagini) può essere preparato e salvato come file immagine in anticipo. Per qualsiasi studio che richieda la riproduzione di una traccia audio in concomitanza con il testo (ad esempio, uno studio di lettura-ascolto) o con le immagini (ad esempio, uno studio sul *Visual world paradigm*), la casella di spunta "Sincronizza audio" deve essere selezionata nell'elenco delle proprietà del nodo.

Experiment Builder non supporta l'inclusione di stimoli dinamici, come la navigazione su pagine web o applicazioni software esterne, ma supporta la registrazione della voce del partecipante durante la visione di uno stimolo. Questo può essere utile negli studi sull'endofasia, o per mettere in relazione il parlato e la lettura di apprendenti L2.

Dal momento che la maggior parte degli studi di linguistica applicata richiede l'aggiunta di numerose frasi-stimolo<sup>4</sup>, procedere manualmente potrebbe risultare molto dispendioso in termini di tempo; pertanto, il software della SR Research consente di aggiungere una "fonte di dati", che funziona come un foglio di calcolo e che può essere

---

<sup>4</sup> In media gli studi linguistici condotti con eye tracker contano dalle 30 alle 150 frasi-stimolo per ciascun disegno sperimentale

popolata manualmente o attraverso l'importazione dei dati da un file di testo. Se si vogliono aggiungere domande di comprensione, è sufficiente aggiungere una colonna nella tabella che contiene gli stimoli e digitare la domanda corrispondente a ciascun elemento dello stimolo.

Per impostazione predefinita, Experiment Builder presenterà gli elementi nell'ordine specificato nell'origine dati. Se necessario, gli elementi possono essere randomizzati o pseudo-randomizzati in anticipo nel file di input. In alternativa, vengono forniti anche metodi per la randomizzazione interna ed esterna, il blocco e il controbilanciamento degli stimoli.

Infine, gli esperimenti con EyeLink devono essere convertiti in file .exe, oppure .app per gli utenti Mac, prima di poter essere utilizzati con i partecipanti.

Experiment builder funziona con Windows e con IOS, anche se il software EyeLink Display è progettato principalmente per il sistema operativo Windows e non è nativamente compatibile con macOS. Nonostante ciò, è possibile eseguire EyeLink Display su un Mac utilizzando software di virtualizzazione come Parallels Desktop o VMware Fusion, che consentono di eseguire Windows all'interno di un ambiente virtuale sul Mac.

**Data Viewer**, invece, è un software per la visualizzazione e l'elaborazione dei dati dello sguardo registrati con i tracciatori oculari EyeLink, disponibile per Windows, macOS e Linux. Si presenta con un'interfaccia semplice, con diversi pannelli e finestre configurabili, che consente di visualizzare, raggruppare, elaborare e riportare i dati dello sguardo provenienti da più file di dati EyeLink.

La vista di sovrapposizione spaziale consente di sovrapporre le saccadi e le fissazioni ai grafici delle prove e di utilizzarle per la creazione di mappe di calore e aree di interesse (cfr. §2.3). I dati dello sguardo e della pupilla possono essere tracciati nel tempo con la vista "Grafico temporale" e i singoli eventi possono essere selezionati e regolati.

Come dicevamo, Data Viewer è progettato per elaborare più file di dati contemporaneamente. I dati possono essere poi raggruppati in base a una o più variabili sperimentali e, a loro volta, i dati di ciascun gruppo possono essere codificati individualmente con un colore, per consentire un confronto diretto tra le condizioni sperimentali. La modalità "Aggregate Data" consente agli utenti di visualizzare simultaneamente i dati di eye tracking di più partecipanti/trials. È possibile creare periodi di interesse per isolare e allineare finestre temporali

specifiche tra le prove. Infine, il raggruppamento delle prove consente di filtrare i dati in base alle variabili della prova, come l'accuratezza, e di includerli o escluderli dalle elaborazioni e dai report successivi. Il visualizzatore di dati consente di generare report a livello di prova, area di interesse, fissazione, saccade e persino campione. I report delle serie temporali (*binning*), i report "Aggregate Event" e "Aggregate Interest Area" consentono di ottenere una rapida panoramica delle principali metriche (vd. §2.1) dello sguardo, suddivise in base alle variabili di raggruppamento delle prove, che possono essere selezionate individualmente, comprese molte "misure di lettura" standard come il numero di regressioni, il tempo di permanenza in un'area di interesse e il numero e la durata delle fissazioni. Tutti i report del Data Viewer possono essere importati direttamente in Excel o in pacchetti di analisi statistica come SPSS e R per un'ulteriore elaborazione.

### 1.2.2. Tobii Pro Lab (Tobii)

**Tobii Pro Lab** è un software per l'analisi dei dati eye tracker che consente agli utenti di esaminare e interpretare i dati raccolti con gli hardware Tobii. Si presenta con un'interfaccia estremamente semplice e chiara e consente di importare dati già raccolti o caricare file di dati dalla sessione di registrazione nell'interfaccia del software.

Una volta importati i dati, Tobii Pro Lab fornisce una varietà di strumenti di visualizzazione per esaminare i movimenti oculari registrati: *gaze plot*, *heat map* e video registrazioni in tempo reale che consentono agli utenti di osservare i pattern di fissazione.

Il software offre una vasta gamma di strumenti analitici per esaminare i dati e identificare pattern significativi. Gli utenti possono confrontare i risultati tra i partecipanti, le condizioni sperimentali e le diverse aree di interesse individuate sul testo così da identificare eventuali correlazioni o differenze significative tra le variabili, dipendenti e indipendenti, dell'esperimento. Tobii Pro Lab consente agli utenti di generare report dettagliati sui risultati dell'analisi ed è possibile esportare i dati nei formati .tvs, .csv o .xls, così da poterli sottoporre ad analisi di tipo statistico attraverso applicativi esterni al software.

A partire dal 2022 Tobii pro Lab ha lanciato uno strumento specifico per le sperimentazioni di stampo linguistico, TextTool, uno strumento progettato per consentire agli sviluppatori di integrare facilmente il tracciamento oculare con input di tipo testuale. Questo significa che gli

utenti possono caricare un testo o digitarlo direttamente nel software, laddove nelle precedenti versioni ogni stimolo doveva essere caricato come immagine. Al contrario, con lo strumento di testo è possibile determinare lo stile del carattere, il colore, e la dimensione e definire lo sfondo su cui proiettare il testo.

Tobii presenta gli input nell'ordine di caricamento, ma gli elementi possono essere randomizzati attraverso una funzione ad hoc e la creazione di cartelle e sottocartelle. Ad oggi Tobii Pro Lab non prevede una versione per IOS, tuttavia, è possibile eseguire Tobii Pro Lab su un Mac utilizzando software di virtualizzazione, che consentono di eseguire Windows all'interno di un ambiente virtuale sul Mac.

### 1.2.3. Be Gaze (SMI)

**BeGaze** è un software di analisi dei dati di tracciamento oculare sviluppato da SensoMotoric Instruments (SMI), progettato per esplorare e interpretare i dati raccolti durante le sessioni di registrazione con gli eye tracker SMI. Questo software fornisce agli utenti una serie di strumenti avanzati per visualizzare, analizzare e interpretare i pattern di fissazione, gli sguardi e altri comportamenti oculari registrati durante gli esperimenti di ricerca.

Come per gli altri programmi, il funzionamento di BeGaze si articola in varie fasi che includono l'importazione dei dati, la visualizzazione sommaria, l'analisi quantitativa e la generazione di report. BeGaze consente agli utenti di importare facilmente i file di dati che contengono informazioni dettagliate sui movimenti oculari dei partecipanti durante l'esperimento, inclusi *timestamp*, coordinate dello sguardo e altre metriche. Una volta incamerati, i dati possono essere selezionati per periodi o aree di interesse e suddivisi sulla base di specifiche condizioni sperimentali o dei gruppi di partecipanti.

Nella fase di visualizzazione e analisi dei dati BeGaze offre la possibilità di visualizzare tracciati oculari, mappe di calore, grafici temporali e altre rappresentazioni grafiche dei pattern di fissazione e sguardo. Queste visualizzazioni consentono agli utenti di identificare i pattern significativi nei dati e di esaminare la distribuzione dello sguardo durante l'esperimento. BeGaze include inoltre una serie di strumenti di analisi quantitativa per confrontare i pattern oculari tra condizioni sperimentali o gruppi di partecipanti, individuare correlazioni tra lo

sguardo e altre variabili e identificare eventuali differenze significative nei comportamenti oculari.

Infine, BeGaze consente agli utenti di generare report dettagliati sui risultati dell'analisi. Questi report possono includere grafici, tabelle e altre visualizzazioni dei dati, nonché una documentazione completa delle procedure di analisi e delle conclusioni dell'esperimento. I report possono essere esportati in vari formati per la presentazione o la pubblicazione.

#### 1.2.4. Software open source

A partire dagli anni Duemila sono stati lanciati sul web diversi software open source per l'analisi dei dati di tracciamento oculare.

Questi primi programmi sono stati creati da sviluppatori indipendenti con il fine di promuovere un principio di accessibilità e collaborazione all'interno della comunità scientifica. Negli anni successivi, con l'aumentare dell'interesse sull'uso degli eye tracker nella ricerca accademica e industriale, sono stati rilasciati software open source con una sempre maggiore varietà di funzionalità e supporto. Oggi, esistono numerosi software open source per eye tracker disponibili, ciascuno con le proprie caratteristiche e funzionalità uniche. Qui di seguito menzioneremo solo alcuni tra i più diffusi.

**OpenGaze** è un applicativo open source per il tracciamento oculare scritto in Python. Questo framework fornisce una piattaforma per l'acquisizione e l'analisi dei dati di tracciamento oculare che include le funzionalità di base per il controllo dell'hardware, per la registrazione dei dati e per l'analisi.

**iView X SDK** è un kit di sviluppo software open source fornito da SensoMotoric Instruments (SMI) per l'integrazione dei loro sistemi di eye tracking. Questo SDK offre una serie di funzionalità per l'acquisizione e l'analisi dei dati di tracciamento oculare, nonché per la creazione di applicazioni personalizzate utilizzando i dati di tracciamento oculare acquisiti.

**OpenEyes** è un software open source sviluppato per l'analisi dei dati di tracciamento oculare. Offre strumenti per l'importazione e la visualizzazione dei dati, nonché per l'analisi statistica e la generazione

di report. È progettato per essere flessibile e personalizzabile, consentendo agli utenti di adattare il software alle proprie esigenze di ricerca.

**GazeParser** è un altro software open source per l'analisi dei dati oculari. È progettato per essere facile da usare e offre funzionalità per l'importazione e la visualizzazione dei dati. GazeParser supporta una varietà di formati di dati ed è adatto per una vasta gamma di applicazioni di ricerca.

**PyGaze** è una libreria Python open source per il tracciamento oculare. Offre un'interfaccia di programmazione altamente flessibile che può essere integrata facilmente in workflow esistenti. È particolarmente adatto per lo sviluppo di applicazioni di tracciamento oculare personalizzate e sperimentali.

**EyeRecToo** è un software open source per l'acquisizione e l'analisi dei dati oculari. Offre funzionalità avanzate per l'analisi dei dati oculari, tra cui l'identificazione delle fissazioni, delle saccadi e la generazione di mappe di calore. EyeRecToo è personalizzabile e può essere adattato alle esigenze specifiche di ricerca.

Questi sono solo alcuni esempi di software open source gratuiti per eye tracker. È importante notare che la disponibilità e le funzionalità dei software possono variare a seconda del tipo di eye tracker e delle specifiche esigenze di ricerca. Indubbiamente, ad oggi, i software open access si presentano più adatti a una utenza con esperienza medio-alta perché se da una parte risultano più flessibili e adattabili, dall'altra, l'interfaccia è meno user-friendly e qualche competenza di programmazione è comunque richiesta per poterli sfruttare al meglio.

### 1.3. Condizioni sperimentali nella ricerca con eye tracker

Dopo aver sommariamente introdotto le parti che costituiscono lo strumento di ricerca e le sue principali funzioni, in questo paragrafo ci occuperemo delle condizioni sperimentali utili a condurre uno studio di impianto linguistico. Ben lungi dal pensare di poter prevedere ogni variabile di ogni disegno sperimentale (che ovviamente dipende dal tipo di studio e dagli obiettivi della ricerca stessa), ci accingiamo in questo paragrafo ad elencare alcune condizioni generali di natura

tecnica e ambientale e altre più specifiche delle ricerche di stampo linguistico e glottodidattico. Tra le condizioni sperimentali comuni a tutte le ricerche condotte con eye tracker, le tre principali sono senz'altro legate a tre fattori: 1) l'illuminazione; 2) lo spazio di raccolta; 3) la durata del protocollo sperimentale

### 1.3.1. Luce

Quando si allestisce un laboratorio o un ambiente di raccolta dati attraverso il tracciamento oculare, in primo luogo è bene assicurarsi che l'ambiente in cui si svolge la raccolta dati con eye tracker abbia una luce uniforme, preferibilmente proveniente dall'alto per evitare riflessi e ombre che possano interferire con il tracciamento degli occhi. Non c'è una direttiva riguardo al tipo di luce. Una luce fredda, che ha una temperatura di colore più alta, risulterà intensa e abbagliante ed è da preferire in laboratori ampi e luminosi o durante le ore del giorno quando la luce naturale è predominante. Tuttavia, una luce fredda troppo intensa potrebbe causare affaticamento visivo ai partecipanti. La luce calda, invece, avendo una temperatura di colore più bassa, tende ad essere più rilassante per gli occhi e può essere preferibile in ambienti più bui o durante le ore della sera. La luce calda può creare un'atmosfera più confortevole per i partecipanti durante le sessioni di ricerca più lunghe.

Se la calibrazione non riesce, oltre a un problema di luce, potrebbe esserci un problema oculistico o una criticità nel modo in cui i near-infra-red colpiscono l'occhio. In alcuni casi accade che il rilevatore oculare scambi qualcos'altro per la pupilla. Le fonti di confusione più comuni sono il trucco degli occhi, le ciglia rivolte verso il basso, le sopracciglia scure o gli occhiali. In tal caso è consigliabile rimuovere il trucco e nei casi in cui le ciglia tendano verso il basso procedere a piegarle con un piegaciglia.

Se insorgono problemi nella calibrazione di partecipanti con gli occhiali, controllare che gli occhiali non siano troppo distanti dalla fronte: devono essere ben sollevati sul ponte del naso e del tutto puliti. Gli occhiali bifocali non possono essere calibrati nella maggior parte dei casi, così come quelli che presentano lenti progressive o scure e quelli con montatura molto evidente e scura. Se il partecipante è d'accordo, si può chiedere di coprire la montatura con un nastro adesivo chiaro leggermente appoggiato sulla montatura.

Può accadere che, nonostante i molti tentativi, la calibrazione non vada a buon fine e che il partecipante non sia idoneo alla raccolta, in questi casi non avrebbe molto senso farlo procedere senza occhiali: l'affaticamento nella lettura, o peggio, l'impossibilità di vedere nitidamente il testo, altererebbe le modalità di lettura in maniera irreversibile. In questi casi può essere opportuno comunque evitare di far percepire come un fallimento quell'appuntamento e, ove possibile, raccogliere ugualmente i dati, magari in misura ridotta, e poi scartarli.

### 1.3.2. Spazio

Lo spazio di raccolta deve essere predisposto con cura e attenzione. L'eye tracker e il soggetto devono essere stabilmente posizionati prima di iniziare la calibrazione<sup>5</sup>. Solitamente il soggetto è posizionato di fronte a uno schermo, preferibilmente su una sedia con braccioli, senza rotelle ma regolabile in altezza. Nei casi di esperimenti di lettura con eye tracker screen-based, è raccomandabile l'utilizzo di una mentoniera (*chin-rest*). Alcuni eye tracker riescono a mantenere una buona stabilità di raccolta anche senza il ricorso alla mentoniera. La decisione di usare o no la mentoniera va ben ponderata perché se da una parte assicura una più accurata raccolta del *sample* e inibisce i movimenti accidentali, dall'altra toglie naturalezza all'attività che si sta chiedendo di svolgere al partecipante.

Per gli eye tracker screen-based la distanza tra il partecipante e l'eye tracker è importante. Una volta scoperta la distanza ottimale tra la sedia e lo schermo è bene segnalarla sul tavolo apponendo un nastro adesivo di carta. Il tavolo dovrebbe essere spazioso abbastanza da contenere la tastiera e il mouse e ove possibile un computer per il monitoraggio on line del tracciamento oculare da parte dello sperimentatore.

Al di là dei fattori logistici e tecnici, garantire uno spazio di raccolta accogliente è un aspetto da non sottovalutare. Partecipanti a proprio agio tendono a essere buoni partecipanti. Come suggerisce Godfroid (2020), spuntini e conversazioni prima dell'inizio della sessione aiutano a creare una buona relazione tra sperimentatore e partecipante. Qualora ce ne fosse la possibilità, è bene chiarire fin da subito se c'è un

---

<sup>5</sup> La calibrazione serve ad adattare i parametri del tracciamento alle caratteristiche oculari del soggetto. Questo processo coinvolge la presentazione di punti di calibrazione sullo schermo mentre il soggetto li fissa.

*cadeau* di ringraziamento o una ricompensa. Nella raccolta in contesto universitario, ad esempio, gli studenti apprezzano di solito gadget di cancelleria con il marchio della propria università o buoni regalo spendibili in libreria o ancora carnet per il cinema, il teatro, ecc.

L'ambiente ottimale non è solo amichevole e rilassato, ma anche silenzioso. In questo modo i partecipanti possono concentrarsi sul loro compito. Se il laboratorio si trova in una zona trafficata del campus o dell'edificio, il rumore può essere un problema. L'affissione di cartelli "Silenzio per favore - Raccolta dati in corso dalle ore X alle ore Y del giorno Z" nel corridoio è una valida ed economica alternativa all'insonorizzazione del laboratorio.

Durante il periodo della raccolta, il laboratorio ha bisogno di alcuni prodotti: fazzoletti di carta, igienizzante spray per mentoniera (o in alternativa dei fogli monouso da applicare sulla parte che entra in contatto con il mento dei partecipanti), salviette struccanti, specchio, piegaciglia, nastro adesivo di carta di colore chiaro.

Quanto allo spazio di lavoro, è bene predisporre al meglio anche quello digitale oltre a quello fisico, ovvero un luogo virtuale in cui eseguire i back-up dei dati raccolti, organizzare il calendario degli appuntamenti, condividere i progressi della ricerca.

### 1.3.3. Tempo

Le sessioni di tracciamento dovrebbero avere una durata tale da poter garantire una certa significatività del campionamento senza affaticare troppo il partecipante e in ogni caso la durata della prova dovrebbe essere adattata a seconda delle capacità e delle caratteristiche del partecipante. La raccolta dei dati può richiedere molto tempo, pertanto può essere utile organizzare una prova tecnica con tre partecipanti prima di prendere decisioni sul tempo da dedicare a ciascun partecipante e sul numero di sperimentatori ottimale per la raccolta (di solito due ricercatori sono sufficienti). Il giorno dell'appuntamento è bene arrivare con largo anticipo e controllare di avere a disposizione tutto il materiale (moduli di consenso informato, questionari di verifica della comprensione, questionari anamnestici, eventuali compensi o *cadeau* di ringraziamento). È buona norma tenere in laboratorio cartelle o file separati ed etichettati per ogni progetto e disporre di un protocollo sperimentale scritto, chiaro e dettagliato per ciascuna ricerca. Un buon protocollo deve specificare, passo dopo passo, i vari momenti

della ricerca: dall'incontro con il partecipante, all'impostazione dell'esperimento, alle istruzioni e alla raccolta dei dati, fino al debriefing del partecipante, all'archiviazione dei dati e al back-up dei dati.

## 2. Cosa misura un eye tracker

*Maria Roccaforte*

### 2.1. Le metriche

Con il termine “metriche”, nei lavori di ricerca che impiegano l’eye tracker, si fa riferimento alle misurazioni e ai dati che l’hardware è in grado di raccogliere e il software di elaborare nel corso della registrazione del comportamento degli occhi. La maggior parte degli studi sulla lettura condotti con l’eye tracker si basa su dati relativi alle fissazioni: numero, durata e presenza (o assenza) all’interno di una determinata area di interesse costituiscono da sempre il dato principale su cui i ricercatori hanno basato la maggior parte delle indagini sperimentali (Godfroid, 2020).

Tuttavia, i sistemi di eye tracking forniscono anche la posizione, la sequenza e la durata di tutti i movimenti che gli occhi compiono su un testo o all’interno di aree di interesse (d’ora in avanti AoI) stabilite dagli sperimentatori, nonché le informazioni sui cambiamenti della pupilla e sui battiti delle ciglia (Holmqvist et al., 2011).

Esistono diversi modi per concettualizzare le misure di eye tracking, in questo capitolo faremo nostra la tassonomia di Lai et al. (2013) che classificano le misure di movimento oculare in base a una scala di misurazione (temporale, spaziale, quantitativa) e al tipo movimento oculare (fissazione, saccade, mista). Le misure temporali quantificano infatti la durata dei comportamenti dello sguardo per permettere di fare inferenze sulla qualità dell’elaborazione cognitiva. La scala spaziale rappresenta invece i movimenti oculari all’interno di aree o porzioni dello stimolo ed è quindi collegata alle posizioni, alle distanze alle direzioni o alle sequenze dello sguardo. La scala del conteggio infine quantifica i comportamenti oculari (ivi comprese le dilatazioni

della pupilla e i battiti delle ciglia) in termini di numero, proporzione e frequenza.

### 2.1.1. Fissazioni

Cominciamo dalla fissazione, che è la metrica più importante negli studi sulla lettura. Esistono quattro sottotipi di misure correlate alla fissazione: (1) il conteggio, (2) la durata della fissazione, (3) la latenza della fissazione (ovvero il tempo impiegato da un partecipante prima di fissare una particolare area di interesse) e (4) la posizione della fissazione in relazione allo stimolo.

**Il numero di fissazioni** (*Fixations Count*) indica il numero di volte in cui l'occhio si ferma su un punto del testo. Questa metrica è espressa con un numero cardinale dentro un'area di interesse preventivamente disegnata dallo sperimentatore oppure sull'intero stimolo. Le aree in cui si verificano le fissazioni possono suggerire quali elementi dell'immagine o dell'interfaccia hanno catturato di più l'attenzione dell'utente.

**La durata della fissazione** (*Fixation Duration*) indica quanto tempo una persona ha mantenuto lo sguardo fisso su uno stimolo. Il tempo di fissazione può essere utilizzato per valutare l'attenzione dedicata a particolari elementi visivi. Esistono due varianti di questa metrica: la durata totale in termini assoluti e la durata media dentro una specifica area di interesse disegnata dallo sperimentatore. Di solito il software restituisce il dato in millisecondi (es. 320 ms) o in secondi (0,32 s).

**La latenza della fissazione** (*Fixation Latency*) è il tempo che intercorre tra la presentazione di uno stimolo visivo e il momento in cui l'occhio si fissa su di esso. In altre parole, è il tempo necessario affinché una persona sposti il suo sguardo verso un nuovo oggetto o punto di interesse dopo che questo è apparso nel suo campo visivo. Anche in questo caso, l'unità di misura è il secondo o il millisecondo.

**Posizione della fissazione** (*Fixation Point*) si riferisce al punto specifico nello spazio visivo su cui l'occhio si è concentrato in un dato momento. Quando utilizziamo un eye tracker per monitorare i movimenti

oculari, la posizione di fissazione rappresenta il luogo preciso in cui lo sguardo è fermo e sta elaborando le informazioni visive. Viene reso attraverso delle coordinate spaziali o indicato sullo stimolo attraverso un punto o un cerchietto.

### 2.1.2. Saccadi

L'altra misura largamente utilizzata nella ricerca con eye tracker riguarda i movimenti saccadici. Il termine "saccade" è un prestito dal francese che significa "scatto": le saccadi, infatti, sono i movimenti più rapidi che il corpo umano sia in grado di eseguire. Questi spostamenti hanno lo scopo di rendere l'immagine nitida e hanno una durata che va dai 10 ai 100 millisecondi, durante i quali l'occhio non è in grado di vedere nulla. Le saccadi vengono registrate in base alla loro velocità, ampiezza, durata, accelerazione e decelerazione, ma è bene precisare che non tutti gli eye tracker oggi in commercio sono in grado di calcolare tutti questi aspetti del movimento saccadico.

La **velocità** è una caratteristica cruciale della saccade e si esprime in gradi di angolo visivo per secondo ( $^{\circ}/s$ ). Poiché le saccadi sono molto brevi, coprono solitamente un raggio di  $1/10^{\circ}$  per secondo; possono essere effettuati movimenti anche più ampi, ma ciò comporterebbe un movimento della testa che negli studi con eye tracker a schermo (i più impiegati nella ricerca di matrice linguistica e cognitiva) è sempre preferibile evitare. La velocità delle saccadi non è costante: si articola in un primo periodo di accelerazione a cui ne segue uno di decelerazione.

Un altro ordine di movimento saccadico è quello definito "di **inseguimento lento**": rispetto a quelli visti finora, prevede che il target osservato sia in movimento e seguito dall'occhio durante lo spostamento; la velocità del movimento è ridotta rispetto a quella delle saccadi tipiche della lettura e dipende dalla velocità del target. In campo linguistico questa misura è usata quasi esclusivamente negli studi sulla sottotitolazione.

Infine, tra i movimenti saccadici troviamo i movimenti di **nistagmo** che sono per lo più involontari, rapidissimi e ripetitivi e identificabili come leggeri tremori, detti anche "movimenti a dente di sega"; si tratta di un disturbo provocato principalmente da una disfunzione delle aree del cervello che controllano i movimenti oculari e pertanto, ad oggi, questo tipo di saccade non ha mai interessato studi e ricerche in campo linguistico.

Le saccadi possono sembrare imprevedibili e di grandezze casuali, tuttavia gli occhi si muovono in punti precisi; ad esempio, possono fermarsi in un punto chiamato *preferred viewing location* (PVL), collocato tra l'inizio di una parola e la sua metà oppure in un'altra zona chiamata *optimal viewing position* (OVP), più vicina al centro della parola. Quando gli occhi si fermano sul primo punto, anziché sul secondo, è più probabile ci sia una nuova fissazione sulla parola in questione e che il tempo di lettura aumenti; ciò avviene per più motivi: ad influire può essere la lunghezza della parola, eventuali disturbi oculomotori e la posizione stessa degli occhi durante la fissazione. Il PVL prevede infatti uno spostamento sulla parte anteriore rispetto all'OVP e spesso le deviazioni da quest'ultimo avvengono quando si legge un testo rispetto a delle parole isolate – in questo caso però la durata diminuisce.

## 2.2. Gli strumenti di visualizzazione

Gli strumenti di visualizzazione consentono ai ricercatori di comprendere come le persone interagiscono con i contenuti e sono in grado di restituire dati cumulativi in modo rapido e di grande impatto visivo. I principali strumenti di visualizzazione utilizzati nell'eye tracking sono i seguenti:

**Gaze Plot:** Il gaze plot mostra i vari percorsi del tracciamento oculare facendo corrispondere ad ogni fissazione un cerchio numerato e ad ogni saccade un segmento che unisce i cerchi. La numerazione dei cerchi corrisponde all'ordine di occorrenza delle varie fissazioni; il loro diametro, invece, è proporzionale alla durata della fissazione. Se sono state effettuate più registrazioni di uno stesso partecipante o di partecipanti diversi, è possibile visualizzarle tutte contemporaneamente sulla stessa scena, in modo da poter effettuare confronti.

A volte, i gaze plot utilizzano i colori per indicare particolari aspetti, come la durata della fissazione (utilizzando una scala cromatica) o la sequenza temporale (con colori diversi per diverse fasi del tempo).

**Heat map:** L' Heat map è una rappresentazione visiva dell'intensità o della durata delle fissazioni in un'area specifica di interesse. La mappa mostra le aree "calde" in cui gli utenti hanno concentrato la loro

attenzione in termini quantitativi (numero di fissazioni) o qualitativi (durata della fissazione) in modo più significativo.

La mappa si costruisce (per lo più in trasparenza) sullo stimolo per evidenziare le zone di fissazione. Questo strumento risulta particolarmente interessante perché permette di selezionare e anche di sovrapporre le mappe dei partecipanti così da restituire l'andamento generale delle registrazioni ed evidenziare "a colpo d'occhio" le aree in cui i partecipanti si sono soffermati di più o di meno. Attraverso i software di analisi è possibile impostare i criteri della visualizzazione sul numero di fissazioni (a prescindere dalla durata), sulla durata assoluta delle fissazioni (la somma delle durate di tutte le fissazioni relative a tutte le registrazioni, a prescindere dal numero di fissazioni) oppure sulla lunghezza della durata relativa delle fissazioni (la somma delle durate di tutte le fissazioni di una singola registrazione).

**Cluster:** Il cluster è uno strumento che accorpa le zone con più ampia frequenza di fissazioni. È possibile specificare un valore di soglia, in Pixel, che rappresenta la distanza massima che le fissazioni devono avere per poter essere raggruppate entro lo stesso cluster.

**Bee swarm:** Il Bee swarm è un video che permette di mostrare la registrazione del tracciato oculare. Questo strumento si può sfruttare in diretta, durante la registrazione, oppure a posteriori, consultando le registrazioni degli andamenti oculari relativi a una o a più registrazioni contemporaneamente.

**La traiettoria oculare** (*Gaze Path o Scan Path*): Rappresenta la sequenza e quindi l'ordine dei punti di fissazione e dei movimenti saccadici che gli occhi eseguono mentre esplorano uno stimolo. Questa metrica può fornire un'immagine completa del percorso visivo di un individuo elaborata dopo la registrazione della prova sperimentale.

### 2.3. Le Aree di interesse (AoI)

Un aspetto fondamentale dell'analisi condotta con eye tracking è l'identificazione delle aree di interesse all'interno di uno stimolo (Lowry et al., 1951). Per area di interesse si intende una regione specifica dello stimolo che è oggetto di attenzione per il ricercatore. Negli studi

linguistici, queste aree possono corrispondere a figure, parole, frasi, o altri elementi linguistici importanti ai fini della ricerca.

Il tracciamento delle AOI nei software di eye tracking è un passo cruciale per analizzare il comportamento visivo. Dopo aver importato lo stimolo nel software, lo sperimentatore disegna le AOI. Il software di eye tracking consente di disegnare le AOI direttamente sopra lo stimolo visivo attraverso strumenti come: 1) la funzione rettangolo, ideale per aree rettangolari e per tenere sotto controllo la dimensione dell'area dell'AOI; 2) la funzione cerchi o ellissi, utili per aree circolari o ovali; 3) la funzione poligoni liberi, per aree irregolari che non si adattano a forme geometriche standard, utile nei casi in cui il testo va a capo ma debba essere inserito dentro una medesima area di interesse. La funzione "modifica e regola" permette, dopo aver disegnato un'area, di modificarla, ridimensionarla, riposizionarla e rinominarla per una più agevole identificazione. È possibile anche copiare e incollare la stessa forma di AOI così da avere la certezza di riportare le stesse dimensioni su stimoli diversi.

In alcuni software, è anche possibile specificare ulteriori parametri, come l'importanza relativa delle AOI o l'intervallo temporale per cui queste sono attive (ad esempio, in un video).

Una volta che le AOI sono state definite, il software può calcolare diverse metriche di eye tracking all'interno di queste aree, come ad esempio il tempo di fissazione (quanto tempo l'occhio si ferma su un'AOI); il numero di fissazioni (quante volte l'occhio ritorna su un'AOI); il tempo fino al primo sguardo (quanto tempo impiega l'utente a guardare un'AOI dalla comparsa dello stimolo); la sequenza (l'ordine in cui le AOI vengono osservate).

## 2.4. Early e late measures

Come da una immaginaria cassetta degli attrezzi, lo sperimentatore dovrà a questo punto scegliere gli strumenti giusti per misurare il comportamento degli occhi nel corso dell'esperimento.

Sebbene l'estrazione dei dati sulla base delle metriche venga fatta a posteriori dal ricercatore (perché la macchina raccoglie comunque tutti i dati e in un secondo momento è il ricercatore a decidere di interrogarli estraendo le metriche di suo interesse) un buon disegno sperimentale è costruito già sulla base della scelta delle metriche da estrarre, prima ancora che la raccolta venga compiuta.

Talvolta però, la stessa metrica viene denominata nei software in modo diverso e, d'altro canto, a nomi apparentemente simili corrispondono spesso metriche diverse.

Riportiamo qui di seguito, in forma di elenco, una serie di definizioni di metriche tratta dai principali software di analisi (§1.4) che saranno suddivise in due categorie: le *early measures*, ovvero le reazioni visive iniziali degli utenti quando guardano uno stimolo visivo e *late measures*, che si riferiscono al comportamento visivo degli utenti dopo le prime reazioni iniziali, utili per comprendere come le persone esplorano e interagiscono con uno stimolo visivo nel tempo, piuttosto che alla loro reazione immediata (Clifton, Staub, Rayner, 2007, p. 349).

### 2.4.1. Early Measures (Misure Precoci)

Con l'etichetta di *early measures* facciamo riferimento alle metriche che vengono registrate all'inizio del processo di percezione visiva, quando gli occhi stanno iniziando a esplorare uno stimolo.

Queste misure includono:

**Il tempo fino alla prima fissazione** (*Time to first fixation*): Misura quanto tempo impiega un utente a fissare per la prima volta una specifica Area di Interesse (AOI) dopo che lo stimolo è stato presentato, con l'obiettivo di indicare che l'elemento ha catturato immediatamente l'attenzione del lettore.

**La posizione della prima fissazione** (*First fixation position*): Registra la posizione della prima fissazione sullo stimolo. Questo aiuta a identificare quale parte dello stimolo è stata osservata per prima e quali elementi visivi sono più salienti o accattivanti.

**La durata della Prima Fissazione** (*First fixation duration*): Indica quanto tempo dura la prima fissazione su un'AOI. Una durata lunga può suggerire che l'utente ha trovato l'elemento immediatamente interessante o complesso, mentre una durata breve può indicare un interesse temporaneo o una facile comprensione.

**La durata della fissazione** (*Fixation duration*): Indica quanto tempo gli occhi rimangono fermi su un punto specifico. Misurare il tempo di fissazione può rivelare l'attenzione dedicata a un elemento visivo.

La **dilatazione pupillare** (*Pupil dilation*): La variazione nella dilatazione pupillare durante l'esplorazione visiva può riflettere variazioni nell'attenzione e nell'interesse.

I **movimenti saccadici** (*Saccades*): Questa metrica misura l'ampiezza dei movimenti degli occhi tra due punti di fissazione.

La **velocità saccadica** (*Saccadic Velocity*): Questa metrica indica la velocità con cui avvengono i movimenti degli occhi tra una fissazione e l'altra. La velocità delle saccadi viene misurata in ms (tra i 120 e i 350 ms).

Il **tempo di lettura al primo passaggio** (*First Pass Reading Time* o *Gaze duration*): È una metrica particolarmente utile nel contesto della lettura e della comprensione del testo. Questa misura si riferisce al tempo che un lettore impiega a leggere una specifica porzione di testo (come una parola, una frase o una zona di interesse) durante il primo passaggio o visita su quella porzione, senza tornare indietro su di essa.

#### 2.4.2. Late Measures (Misure Tardive)

Le *late measures* si riferiscono alle metriche che vengono registrate dopo che gli occhi hanno esplorato uno stimolo visivo per un certo periodo di tempo e possono riflettere l'elaborazione cognitiva successiva. Queste misure includono:

Il **tempo totale di fissazione** (*Total Fixation Time* o *Total Reading Time*): rappresenta la somma di tutti i tempi di fissazione su un determinato elemento visivo durante l'esplorazione. Può indicare quanto tempo una persona dedica alla comprensione o all'analisi di un elemento.

Il **numero totale di fissazioni** (*Total Number of Fixations*): questa metrica conta il numero complessivo di fissazioni su un elemento. Un numero elevato di fissazioni può suggerire una maggiore complessità o difficoltà nell'elaborazione.

I **ritardi tra fissazioni** (*Intervals Between Fixations*): misura il tempo trascorso tra una fissazione e la successiva. Può riflettere il tempo necessario per elaborare un elemento visivo.

La **durata Media di Fissazione** (*Fixation duration mean*): calcola il tempo medio delle singole fissazioni su una AOI. Una durata maggiore potrebbe indicare che l'utente sta cercando di elaborare o comprendere un contenuto complesso.

Il **tempo fino alla Prima Fissazione** (*Time to first fixation*): indica quanto tempo impiega un utente a fissare una AOI dopo la prima fissazione. Un ritardo lungo potrebbe suggerire che l'elemento non è stato percepito come immediatamente rilevante, ma ha comunque catturato l'attenzione successivamente.

La **sequenza di Fissazioni** (*Fixation sequence*): analizza l'ordine in cui l'utente osserva le AOI, fornendo informazioni su come l'occhio si muove attraverso lo stimolo. Questo può rivelare il percorso visivo preferito e come le informazioni sono processate in sequenza.

Le **rivisitazioni** (*Revisitations*): misura quante volte un utente torna a una determinata AOI dopo averla già osservata. Questo può indicare un alto livello di interesse o difficoltà nel comprendere il contenuto di quell'area.

Il **numero di Saccadi prima della Prima Fissazione** (*Number of Saccades to First Fixation*): conta quante saccadi avvengono prima che l'occhio si fermi su una AOI. Un numero elevato potrebbe indicare che l'utente ha esplorato più aree prima di trovare un punto di interesse.

La **dilatazione pupillare** (*Pupil Dilation*): questa metrica misura la variazione del diametro della pupilla durante l'esplorazione visiva. La dilatazione pupillare può riflettere le variazioni dell'attenzione o dell'emozione.

I **battiti di ciglia** (*Blinks*): si tratta di un riflesso muscolare che è collegato anche alla carica emotiva associata a un compito o a uno stimolo visivo.

## 2.5. Metriche e strumenti di analisi negli studi su lettura e comprensione

Negli studi di stampo linguistico le metriche di fissazione sono quelle principalmente usate da coloro che effettuano ricerche sulla L2 perché permettono di valutare gli effetti dell'attenzione e del *noticing*. Per studiare questi fenomeni i ricercatori tendono a paragonare due o più fissazioni in condizioni diverse (una marcata e una no, ad esempio grammaticale vs agrammaticale); in questo modo è possibile considerare come l'informazione viene elaborata e manipolata dall'individuo (Conklin, Pellicer-Sanchez, Carrol, 2018; Indrarathne & Kormos, 2017).

Rayner (1998, 2010), psicologo e pioniere del moderno eye tracking, ritiene che nei processi complessi, come quelli della lettura, i movimenti oculari possano rivelare a cosa si sta facendo attenzione e quanto sforzo richiede l'azione. Per esempio, la difficoltà e la complessità di ciò che viene letto o osservato influenza saccadi, fissazioni e regressioni; nel caso della lettura di un testo 'difficile' si osserverà l'aumento della durata di fissazioni e regressioni, mentre diminuirà la lunghezza delle saccadi. Come abbiamo visto, durante le fissazioni gli occhi si fermano e il sistema cognitivo percepisce e processa gli input e, allo stesso tempo, pianifica quanto e come far muovere gli occhi subito dopo, compiendo un movimento saccadico durante il quale non vengono elaborati nuovi input ma è possibile proseguire la processazioni di quelli decodificati durante la fissazione precedente.

Le fissazioni, le saccadi e le regressioni presentano misure standard, ma in base al compito richiesto si adottano delle metriche differenti; infatti, è stato osservato che i movimenti oculari sono differenti in base fattori diversi come la competenza del lettore, al tipo di lettura richiesta e al tipo di input; Rayner (2010) ha anche sottolineato come questi elementi richiedano un diverso coinvolgimento del sistema cognitivo che si manifesta nei movimenti oculari. Un esempio è quello della lettura a voce alta e della lettura silenziosa: nel primo caso le fissazioni sono più lunghe e le saccadi più brevi, mentre avviene il contrario nel caso della lettura silenziosa.

Secondo la maggior parte degli studi linguistici condotti finora i fattori principali che possono predire la durata di una fissazione sono la frequenza, il legame con il contesto, la prevedibilità e la lunghezza della parola; a questi contribuiscono anche l'età di acquisizione della parola e la familiarità con essa, la concretezza e la parte del discorso

interessata (Kliegl, Nuthmann, Engbert, 2006). Queste elencate sono tutte variabili che possono influenzare la durata delle fissazioni.

Per studiare le variabili elencate, nello specifico la correlazione tra frequenza, lunghezza e prevedibilità, gli studiosi scelgono di solito tra due tipi di approcci: il primo è quello del controllo sperimentale, mentre il secondo è quello del controllo statistico (Kliegl, Grabner, Rolfs, Engbert, 2004). Il controllo sperimentale è quello più comune e più datato e consiste nella manipolazione sperimentale delle parole target; infatti, in questo tipo di approccio le parole devono diversificarsi per una variabile, come per esempio quella della frequenza, ma essere uguale sulle altre dimensioni (prevedibilità, lunghezza ecc.). Più recentemente gli studiosi hanno adottato il secondo tipo di approccio che si basa su un modello di studio delle variabili preso in prestito alla statistica; è un metodo basato sulle regressioni multiple e permette di studiare una determinata variabile tramite lo studio delle altre variabili.

La conoscenza profonda delle metriche e delle loro implicazioni dal punto di vista cognitivo è fondamentale per l'elaborazione e l'impostazione di ogni disegno sperimentale; tuttavia, dal momento che la loro definizione può variare a seconda del tipo di eye tracker, è bene approfondire sul manuale d'uso di ciascuna macchina quali siano i nomi che il software attribuisce alla metrica prima di progettare l'esperimento.



## 3. Reading

*Veronica D'Alesio*

### 3.1. Introduzione ai modelli di lettura

Come abbiamo visto, l'occhio si muove sul testo scritto procedendo "a salti", definiti da saccadi e fissazioni. La lettura è un'abilità complessa, in cui l'attenzione visiva è guidata e influenzata da una molteplicità di fattori: ad un livello più profondo, anche l'elaborazione di quanto letto avviene su diversi piani, partendo dalla decodifica ortografica, fonologica e lessicale e integrando via via ulteriori informazioni (sintattiche, pragmatiche...) per arrivare alla comprensione del messaggio. Utilizzare l'eye tracking per capire come avviene l'elaborazione del linguaggio attraverso la lettura significa progettare di volta in volta dei "test", ovvero degli esperimenti, per capire quali sono i fattori linguistici che guidano l'occhio lungo il testo. Prima di andare avanti, è necessaria un'ulteriore precisazione: negli studi di eye tracking basati su testo, l'unità di misura di base è generalmente la parola grafica, intesa, quindi, come la sequenza di caratteri compresa tra due spazi bianchi. Anche quando si parla, ad esempio, di espressioni polirematiche o collocazioni, ci si riferisce a *multiword units*, letteralmente unità multiparola. Ciò non equivale a dire che questa metodologia indaga solo ed esclusivamente il piano lessicale, anzi: l'eye tracking è molto usato nello studio della struttura sintattica, così come in semantica e pragmatica sperimentale. È bene sottolineare, però, che non si può mai prescindere dalle caratteristiche ortografiche, fonologiche e lessicali di ogni parola. Ogni volta che sottoponiamo un testo a un lettore, dobbiamo tenere a mente che parametri come numero di caratteri, frequenza, significati associati, status grammaticale e così via, giocano un ruolo fondamentale nelle prime fasi di elaborazione del testo scritto. In que-

sto capitolo, analizzeremo i principali studi che hanno gettato le basi per l'implementazione di un possibile modello di lettura, ovvero un sistema che, dato un testo, sia capace di predire il più accuratamente possibile dove si posa il nostro occhio, dopo quanto tempo si stacca da quel punto per proseguire, e in che direzione avviene il salto. Nella vastissima letteratura a riguardo, i diversi approcci metodologici si posizionano lungo due assi, che possiamo sintetizzare in due domande: 1) cosa muove l'occhio? e 2) quanto è accessibile all'occhio in un dato momento?

Il dibattito attorno alle due questioni è molto acceso, principalmente perché, come abbiamo visto finora, lo studio dei movimenti oculari coinvolge discipline diverse tra loro, ognuna con le proprie metodologie e i propri assunti di base. In risposta alla prima domanda, la maggior parte dei modelli si divide tra modelli cognitivi e modelli oculomotori. Da un lato, i modelli cognitivi danno grandissimo rilievo alla durata e al posizionamento delle fissazioni, in quanto principali correlati dell'elaborazione linguistica. I proponenti di questi modelli, quindi, assumono un legame molto forte tra l'occhio e la mente, argomentando che i movimenti dell'occhio durante la lettura siano guidati in gran parte da fattori cognitivi. Dall'altro lato, i modelli a base oculomotoria fanno perno maggiormente su principi motori, mettendo in risalto meccanismi fisiologici (movimenti saccadici, acuità visiva) e informazioni periferiche (numero di caratteri in oggetto) alla base del movimento degli occhi. Il secondo punto controverso riguarda l'elaborazione in sé, che divide gli approcci tra sequenziali (o "seriali", dalla denominazione originale *serial*) e paralleli. Come si evince, nei modelli sequenziali l'elaborazione cognitiva avviene su una parola per volta, ovvero solo sull'elemento fissato in quel momento. Al contrario, i modelli paralleli, prevedono un raggio più ampio, arrivando a concepire un processing esteso fino a 4 parole. I modelli oculomotori derivano dai lavori di O'Regan (1990;1992) e McConkie (McConkie et al., 1989; Radach e McConkie, 1998; per una rassegna sui vari modelli si rimanda a Reichle, Rayner e Pollatsek, 2003). Qui ci concentreremo sui modelli cognitivi, in quanto maggiormente rappresentativi dei fenomeni linguistici operanti in lettura e documentati da una vasta letteratura. In particolare, ci soffermeremo sull'E-Z Reader, a elaborazione sequenziale (o seriale), per poi passare a SWIFT e Glennmore, a elaborazione parallela.

### 3.1.1. E-Z Reader

Uno dei modelli più conosciuti e utilizzati in letteratura scientifica è quello che va sotto il nome di E-Z Reader, un modello cognitivo a elaborazione sequenziale. Per evitare ogni confusione, precisiamo che il suo nome non ha niente a che vedere con quanto descrive: infatti, E-Z Reader è preso da un personaggio di un programma per bambini ed era, a sua volta, un giocoso riferimento al film *Easy Rider*. Il perno del modello consiste nella scissione in due fasi dell'elaborazione lessicale, una iniziale (*early* o  $L_1$ ) e una successiva (*late* o  $L_2$ ). In un primo momento -che inizialmente veniva chiamato dagli autori *familiarity check*<sup>6</sup>- avviene infatti l'identificazione ortografica della parola, mentre nella fase successiva si ha l'elaborazione fonologica e semantica e, per questo, ci si riferisce a questa fase come all'elaborazione lessicale completa e corrisponde al *lexical access* in letteratura psicolinguistica<sup>7</sup>. Entrambe le fasi sono funzioni di due proprietà della parola, di rilevanza più statistica che linguistica: la frequenza (calcolata secondo Francis e Kučera, 1982) e la predicibilità, ovvero la probabilità di una parola di apparire in un dato contesto frasale. Nella versione aggiornata del modello (Reichle, Pollatsek e Rayner, 2006), viene specificato che, mentre la fase  $L_1$  può ridursi fino ad azzerarsi, la fase  $L_2$ , al contrario, è obbligatoria: ciò vuol dire che, anche nel caso di parole altamente frequenti e totalmente predicibili dal contesto, è sempre richiesto un tempo minimo per "comprendere" una parola, e questo tempo minimo è attribuito all'ultima fase di elaborazione, ritenuta responsabile dell'accesso al significato. Negli anni diversi studiosi, tra cui Andrews (2003), e Huestegge et al., (2003) hanno osservato come questa divisione in due stadi possa essere con-

<sup>6</sup> Il termine *familiarity check* viene reintrodotta in Reichle, Pollatsek e Rayner (2006) come alternativa al termine  $L_1$ . È importante notare che i tempi previsti per questa fase dell'elaborazione lessicale sono dipendenti dall'acuità visiva (cfr. supra): l'effetto lunghezza sull'elaborazione lessicale -già attivo in visione parafoveale- è una conseguenza della distanza dei confini di parola dal centro foveale, ovvero dal numero di caratteri, a destra e sinistra, che si allontanano dal punto di fissazione.

<sup>7</sup> Notiamo che in Reichle, Pollatsek e Rayner (2006) gli autori aprono alla possibilità di identificare  $L_1$  come *lexical access* e  $L_2$  come una fase di integrazione post-lessicale: questa osservazione, che non appare nel numero speciale di *Behavioural and Brain Sciences* (2003) dedicato al modello, non è ulteriormente motivata e sono gli autori stessi a definirsi "agnostici" rispetto all'interpretazione dei due livelli di elaborazione lessicale.

siderata in parte arbitraria<sup>8</sup>, soprattutto per quanto riguarda la specifica ripartizione sequenziale tra decodifica ortografica, fonologica e lessicale; inoltre, la relazione tra questi piani e le variabili di frequenza e predicibilità non sono specificati. Occorre ricordare, però, che un modello, in quanto tale, è una computazione matematica che mira a rappresentare un quadro ideale della realtà: come gli stessi autori ribadiscono, ciò che conferisce chiarezza e trasparenza al modello è spesso una ipersemplicificazione della moltitudine dei fenomeni cognitivi e comportamentali che interagiscono in un'abilità complessa come la lettura. A questo proposito, un ulteriore limite dell'E-Z Reader è quello di non considerare le regressioni: questo modello ripercorre il tracciato di una lettura senza intoppi di un lettore abile con un'ampia conoscenza lessicale e una competenza linguistica sempre adeguata al testo. Conseguentemente, si assume che l'identificazione della parola si esaurisca in un solo passaggio e, verosimilmente, con l'apporto del solo co-testo precedente: la fase L<sub>2</sub>, quindi, non coincide con le *late measures* che abbiamo visto in precedenza (§2.4); o meglio, le *late measures* misurano l'elaborazione lessicale completa solo nel caso di una lettura che procede di pari passo con il testo senza salti tra parole. Nell'esperienza reale, però, sappiamo che le *late measures* inglobano tutte le successive riletture di una data porzione di testo, mentre la misura della prima lettura di una parola è data dalla *gaze duration* (§2.4), che è tradizionalmente considerata una *early measure*. Al di là dell'apparente incongruenza terminologica che deriva dalla sostanziale -e spesso necessaria- differenza tra dati sperimentali e uniformità computazionale, è bene notare che non ci sono evidenze inconfutabili dell'esatta scansione temporale o logica dell'elaborazione dei diversi piani linguistici durante la lettura, o del fatto stesso che avvengano in momenti separati. Come gli stessi autori sostengono, l'E-Z Reader non vuole essere una spiegazione profonda dell'elaborazione del linguaggio, in quanto non tiene conto dei processi cognitivi di livello superiore: tuttavia, l'attribuzione dell'elaborazione lessicale ad un piano più "basso" o più "immediato" rispetto a un ordine

---

<sup>8</sup> La stessa ripartizione temporale tra L<sub>1</sub> e L<sub>2</sub> cambia notevolmente tra l'E-Z Reader 7 (Reichle, Rayner e Pollatsek, 2003) e il suo aggiornamento (Reichle, Pollatsek e Rayner, 2006): nella prima versione, la predicibilità influenzava maggiormente la fase L<sub>2</sub> che, di conseguenza, poteva teoricamente azzerarsi; nella versione successiva, frequenza e predicibilità sono processate nella L<sub>1</sub>, mentre la fase L<sub>2</sub> è maggiormente centrata sull'elaborazione del significato e, quindi, risulta necessaria.

superiore di cognizione (a cui presumibilmente appartengono integrazioni semantiche, pragmatiche e sintattiche) distinto e, pertanto, superfluo alla riuscita ottimale e completa dell'identificazione della parola, è da ritenersi in sé un approccio metodologico con una precisa affiliazione teorica. Proseguendo con la descrizione del modello, vediamo che la definizione di due stadi di elaborazione lessicale è funzionale a un determinato meccanismo saccadico: alla fine della prima fase di elaborazione lessicale, infatti, si inizia a programmare la successiva saccade e a computare la distanza del salto; a elaborazione lessicale completa, il comando di spostamento viene effettuato, passando così alla lettura della parola successiva. Questo meccanismo, ancora una volta in due fasi, tiene conto di tre fenomeni diversi: lo *skipping* (ovvero, il "saltare" una parola), le rifissazioni (ovvero, fissazioni immediatamente successive alla prima) e l'effetto *spillover* (che comporta un aumento dei tempi di lettura nella parola immediatamente successiva a un target laborioso in termini di elaborazione). Essendo la prima fase della programmazione "labile", questa può essere cancellata in virtù di alcune proprietà della parola successiva già accessibile nella visione parafoveale<sup>9</sup> Per quanto questo possa sembrare in contraddizione con la stretta sequenzialità prevista dal modello, la visione parafoveale è definita come pre-attentiva: è la fase in cui l'informazione dalla retina viene trasmessa al cervello quando, presumibilmente, l'occhio già vede i confini di parola e il suo contorno. Queste informazioni si traducono, perciò, in una prima approssimazione della lunghezza e della frequenza della parola successiva: nei casi in cui questa sia molto corta ed estremamente frequente (come accade per le parole funzione), la programmazione della saccade viene annullata, passando direttamente all'esecuzione del salto verso la parola ancora successiva. Per quanto riguarda le rifissazioni, per la maggior parte vengono interpretate come prodotti di errori, che riflettono lo scarto tra la programmazione saccadica e l'effettiva esecuzione. Gli errori possono essere standard -ovvero sistematicamente riconducibili a cause prevedibili- o random (McConkie, 1988;1991, in Reichle, Rayner, Pollatsek, 2003). Un tipico errore sistematico risulta, ad esempio, dal tentativo mancato di centrare l'OVP

---

<sup>9</sup> Si chiama visione 'foveale' quella che avviene nella fovea, corrispondente a 1° di angolo visivo a ogni lato del centro della fissazione; procedendo verso l'esterno, si ha poi la visione 'parafoveale' (5°) e, infine, la visione 'periferica'.

(*Optimal Viewing Position*, vedi §2.1): dato che la lunghezza ottimale di una saccade -almeno per l'inglese- copre sette caratteri, quando una parola è più lunga o più corta, capita che la prima fissazione cada in un punto non ottimale e, di conseguenza, l'occhio programma una serie di rifissazioni tendenti al centro della parola. Infine, l'effetto *spillover* viene spiegato dal fatto che la difficoltà nel processare una parola inibisce la contemporanea elaborazione parafoveale della seguente: in questo modo, ai tempi minimi di programmazione e esecuzione saccadica (175 ms in media), si sommeranno le informazioni che non sono state processate in fase pre-attentiva, facendo aumentare i tempi di elaborazione lessicale a carico della parola seguente. In quest'ottica, quindi, lo *spillover* si configura come l'effetto dell'assenza di un beneficio che normalmente ha luogo nei casi di lettura senza difficoltà, piuttosto che come l'aggravamento dei costi di elaborazione di una parola non ancora nella sfera attentiva. Come abbiamo visto, la stretta sequenzialità che il modello sostiene è frutto di una triplice distinzione dell'attenzione: da una parte c'è la pre-attentività, che permette all'occhio una prima e grossolana rappresentazione della parola successiva; dall'altra parte, c'è l'attenzione vera e propria, definita sequenziale perché passa da una parola all'altra in maniera lineare. A sua volta l'attenzione, come è di unanime accordo in letteratura, si divide in implicita ed esplicita: l'attenzione implicita si "sposta" durante la programmazione della saccade, ovvero alla fine della prima fase di elaborazione lessicale della parola; il cambio di attenzione esplicita si manifesta, invece, con l'effettivo movimento dell'occhio, che si posa infine sulla parola successiva. È bene ricordare che la sequenzialità dell'elaborazione cognitiva in lettura è uno dei punti più dibattuti in letteratura e, nonostante ne faccia un suo pilastro, l'E-Z Reader prevede che una parola sia processata parafovealmente. Questo approccio viene identificato dall'acronimo SAS (Sequential Attention Shifts) e rappresenta un punto di raccordo, o perlomeno, non di mutua esclusione, con i modelli a elaborazione parallela.

### 3.1.2. SWIFT

*Saccade-generation With Inhibition by Foveal Targets* (d'ora in avanti, SWIFT; Engbert et al., 2005) è un altro modello cognitivo che presenta diverse analogie con l'E-Z Reader: entrambi i modelli presentano

due livelli di elaborazione lessicale e di programmazione saccadica; in entrambi i casi le saccadi risultano in parte da errori random e sistematici e, infine, la difficoltà associata a una parola è strettamente collegata alla sua frequenza e predicibilità in contesto. Al contrario dell'E-Z Reader, però, lo SWIFT è un modello basato su gradienti attenzionali (*Guidance by Attentional Gradients*, o modelli GAG), dove l'elaborazione linguistica è distribuita su una stringa di parole adiacenti: nello SWIFT in particolare, il campo di attivazione dell'elaborazione comprende la parola fissata, la parola prima, quella dopo e, in maniera residuale, quella dopo ancora. La distanza dal centro di fissazione, inoltre, influenza la velocità di elaborazione, assumendo che questa sia limitata principalmente dall'acuità visiva. Nello specifico, è la prima fase di elaborazione lessicale a essere influenzata dallo stimolo visivo. In questo stadio, detto pre-processamento, si esaminano i possibili candidati per l'obiettivo della prossima saccade e il parametro maggiormente legato alla difficoltà della parola è la frequenza della parola della scritta. La predicibilità, invece, ha un effetto inibitorio sul pre-processamento della parola in visione parafoveale (ovvero, quella dopo): gli autori sostengono, infatti, che questo parametro sia per definizione svincolato dallo stimolo visivo, in quanto si può dire che una parola è predicibile solo se può essere "indovinata" senza essere vista. Questo rallentamento del primo stadio lessicale comporta una minore attivazione sulla parola in questione per cui, con ogni probabilità, la parola altamente prevedibile non sarà obiettivo della prossima saccade, dando luogo allo *skipping*. La fase successiva di elaborazione lessicale, quella del completamento, è la risultante dei tempi di recupero della parola dalla memoria e, secondo questo modello, non è legata ad altre proprietà lessicali. Come abbiamo detto, anche nello SWIFT i movimenti saccadici si dividono in programmazione (labile, quindi cancellabile) ed esecuzione (non labile). Dal punto di vista dell'elaborazione lessicale, durante la saccade si arresta il pre-processamento in quanto questo, lo ricordiamo, è strettamente dipendente dall'input visivo; il completamento dell'elaborazione lessicale può invece proseguire anche durante i movimenti oculari. Al centro del modello troviamo la netta separazione tra programmazione del movimento (che inizia dopo un intervallo di tempo casuale) e la direzione della saccade (che, ricordiamo, deriva dalla maggiore attivazione di un elemento sugli altri in fase di *pre-processing*). La programmazione saccadica, pur essendo totalmente casuale, può essere ritardata dall'elaborazione lessicale in corso, nel caso in

cui la parola che si sta attualmente fissando presenti delle difficoltà per il lettore: il nome del modello SWIFT descrive proprio l'occhio che si sofferma su una parola difficile, perché la generazione saccadica (ovvero, il salto successivo) è stata inibita dal target foveale.

### 3.1.3. Glenmore

Il modello Glenmore<sup>10</sup> (Reilly, Radach, 2003) è un largamente ispirato ai modelli connettivisti, in special modo quelli ad attivazione interattiva (*Interactive Activation models*): alla base del funzionamento di Glenmore vi è, infatti, una rete di unità interconnesse che si attivano una dopo l'altra fino al raggiungimento di un culmine; arrivata a questa soglia massima, l'attivazione in quel punto decresce e nuovi input competono tra di loro per riattivare di nuovo lo stesso circuito in un altro punto. Gli elementi essenziali -o moduli- di Glenmore, lungo i quali scorre questa catena di attivazione, sono i moduli di input visivo, elaborazione della parola, il centro di fissazione e il generatore delle saccadi. L'input visivo è costituito dalle parole sul testo (Glenmore opera su una larghezza di trenta caratteri), che insieme si configurano come una "mappa di salienza" (*saliency map*): ogni parola presente in quello spazio è un potenziale target per una fissazione, ma non tutte lo diventano effettivamente. In altre parole, le parole in fila sul testo competono tra di loro, ma solo quelle che raggiungono un livello alto di attivazione vengono poi fissate ed elaborate. Per questo gli autori non definiscono il *word skipping*, o meglio non lo concepiscono in termini di "salto" ma piuttosto come un'implicazione normale nel processo di lettura: infatti, le parole che non vengono fissate sono semplicemente parole a bassa attivazione che, quindi, in un certo senso non hanno bisogno di essere fissate per essere elaborate. L'attivazione si basa sull'elaborazione di lettere e parole in contemporanea, e i due livelli si scambiano stimoli reciprocamente rafforzandosi a vicenda. Dal punto di vista grafico, anche qui la lunghezza gioca un ruolo in quanto misura di eccentricità, per cui le lettere (e le parole a queste collegate) più lontane dal centro focale si attivano di meno. A livello di parola, invece, l'unico fattore linguistico che interviene nella competizione è la frequenza, che è direttamente proporzionale al grado di attivazione. In

<sup>10</sup> Anche in questo caso, il nome del modello è legato alla sua architettura, ma è ripreso dal nome della località dove gli autori l'hanno ideato.

pratica, una parola molto frequente viene attivata moltissimo ma, allo stesso modo, raggiunge in poco tempo il picco massimo, il “tetto” a cui l’attivazione cade. Una volta “spente”, queste parole non competono più: l’obiettivo finale della fissazione, quindi, saranno le ultime parole rimaste, ovvero quelle con un’attivazione lenta e, perciò, duratura, come le parole lunghe e poco frequenti. Glenmore non implementa il processo di elaborazione linguistica vera e propria, per cui lo scopo principale del modello rimane la determinazione del luogo delle fissazioni: in altre parole, spiega principalmente “dove” si muove l’occhio, ma non “quando”. Gli autori lo definiscono un modello, in quanto tiene conto di fattori sia linguistici che motori, prevedendo che queste due componenti interagiscano continuamente in tutte le fasi di lettura. Infine, come abbiamo visto, il meccanismo di competizione permette di prescindere dal principio di spostamento dell’attenzione, poiché realizza l’elaborazione parallela e contemporanea di più parole. Secondo questa modalità, inoltre, l’occhio non è automaticamente spinto in una direzione, ma va semplicemente dove l’attivazione lo attira: esattamente come il *word skipping*, quindi, rifissazioni e regressioni non sono eccezioni, ma la norma, purché queste rientrino nella mappa di salienza.

### 3.2. Variabili linguistiche

Durante la lettura, i movimenti oculari e la durata delle fissazioni sono influenzati dalle caratteristiche delle singole parole e, in alcuni casi, delle parole in combinazione tra loro (Rayner, 1998; Clifton et al., 2007). Tra le caratteristiche visive, la più studiata è la **lunghezza**, considerata come il numero di caratteri di cui è composta la parola. Se consideriamo un sistema di scrittura alfabetico che procede da sinistra verso destra, lo *span* visivo di un lettore esperto copre circa 14-15 caratteri a destra della fissazione. In questo modo, il lettore recupera un’informazione visiva fondamentale, ovvero quella dei confini di parola: individuando lo spazio bianco successivo già in lettura parafoveale, durante una fissazione l’occhio già ha un’idea della lunghezza approssimativa della parola successiva. Inoltre, come diversi studi di Inhoff (1989a,b; 1990) suggeriscono, in un raggio più limitato -di circa tre lettere a destra della fissazione- è possibile anche un parziale recupero lessicale. La conseguenza più evidente del cosiddetto “effetto lunghezza” (dall’inglese *length effect*) è, quindi, il salto (*skipping*) di

parole molto brevi per una duplice ragione: da una parte, il *processing* di una parola di circa tre caratteri può avvenire già in visione parafoveale, quindi quando l'occhio sta ancora fissando la parola precedente; dall'altra, le parole molto corte sono spesso anche le più frequenti (in inglese, ma anche in italiano: pensiamo, ad esempio, a tutte le parole funzione). Il parametro della frequenza, infatti, influisce fortemente sull'elaborazione di una parola in lettura e, come è stato spesso notato in letteratura, non è facile controllare e discernere l'effetto frequenza dall'effetto lunghezza, proprio perché le due caratteristiche spesso risultano essere direttamente proporzionali tra di loro, rendendo di fatto le due variabili sovrapponibili. Un'ulteriore conseguenza dell'effetto lunghezza, come è facilmente immaginabile, è l'incremento della durata totale delle fissazioni (in millisecondi), dovuto al fatto che più è lunga la parola, maggiore sarà la probabilità che venga fissata più volte in punti diversi (vedi le rifissazioni nel paragrafo precedente). Passando adesso alla **frequenza**, già nel 1986 Inhoff e Rayner avevano dimostrato come l'alta frequenza di una parola avesse un effetto facilitatorio nell'accesso lessicale. Gli autori sottopongono a 24 partecipanti una lista di 60 frasi in cui si alternano una parola ad alta frequenza e una a bassa frequenza, di uguale lunghezza e inserite esattamente nello stesso contesto:

1. His painful *accident* caused him trouble.  
'Il suo doloroso *incidente* gli ha causato disagi.'
2. His painful *fracture* caused him trouble.  
'La sua dolorosa *frattura* gli ha causato disagi.'

Le parole critiche sono state lette in tre condizioni diverse, a seconda che il limite destro della frase fosse posto subito dopo la parola critica, una parola dopo o che la frase fosse presentata per intero.

La durata della prima fissazione subisce solo minime variazioni nella prima condizione, ma l'effetto frequenza è notevolmente evidenziato nelle altre due condizioni, suggerendo che il primo accesso lessicale sia in qualche modo favorito dalla visione parafoveale. I cambiamenti maggiori si verificano nel *gaze duration*, ovvero la durata del totale delle fissazioni su una parola in prima lettura: indipendentemente dalla condizione e a parità di lunghezza, le parole a bassa frequenza vengono rifissate più volte rispetto a quelle ad alta frequenza, portando a un aumento totale del tempo di lettura. In uno studio riguardante lunghezza,

frequenza e predicibilità, Kliegl et al. (2004) hanno, inoltre, osservato che la frequenza di una parola influisce largamente sulle sue probabilità di essere saltata in lettura (aumenta, cioè, il suo *skipping rate*). A questo proposito, lo studio di White (2008) tenta di fare luce sull'intricato rapporto tra frequenza e familiarità ortografica. Come l'autrice sottolinea, le stringhe di caratteri appartenenti a parole molto frequenti in una lingua risultano necessariamente più familiari al lettore; tuttavia, non tutte le combinazioni di lettere molto frequenti sono sempre associate a parole frequenti. La familiarità ortografica, dunque, è diversa dalla frequenza di una stringa di lettere perché, a differenza di quest'ultima, può presentare gradi diversi di informatività. L'esempio, preso dal suo studio ma facilmente trasponibile in italiano, è quello delle sequenze "pn" e "irr": per quanto "pn" sia molto poco frequente in inglese, è al contempo una stringa molto riconoscibile, nel senso in cui "riduce" la scelta lessicale a poche unità (*pneumonia, pneumatic...*); nel caso di "irr", invece, che è un prefisso molto comune (anche in italiano), la sua alta frequenza quasi ostacola l'immediato recupero lessicale visti i numerosi concorrenti contenenti la stessa sequenza. La familiarità ortografica, quindi, deve essere distinta dalla frequenza di parola proprio perché le due operano in maniera diversa e pertengono a due piani separati: quello visivo da un lato e quello lessicale dall'altro. Nel caso particolare dello *skipping*, ci si aspetta che la parola saltata venga in qualche modo riconosciuta ed elaborata mentre l'occhio fissa la parola precedente: come abbiamo visto, sono spesso parole molto corte ed estremamente frequenti, che l'occhio "riconosce" in anticipo e che, quindi, non ha bisogno di fissare per poter proseguire con la comprensione del testo. Lo studio di White (2008) propone 39 frasi in cui la parola critica si presenta in tre condizioni diverse: 3) una parola frequente e con un'ortografia familiare; 4) una parola infrequente, ma familiare dal punto di vista ortografico; 5) una parola a bassa frequenza e non familiare.

3. He examined the small *group* using a basic written test for each child.

'Esaminò il piccolo *gruppo* usando un test scritto elementare per ogni bambino.'

4. He examined the small *louse* using a special microscope.

'Esaminò il piccolo *pidocchio* usando un microscopio speciale.'

5. He examined the small *ulcer* using key-hole surgery.

'Esaminò la piccola *ulcera* usando la laparoscopia.'

Come si nota, le parole critiche sono controllate per lunghezza e sia il contesto precedente che la parola successiva rimangono invariati in tutte e tre le condizioni, per valutare eventuali effetti parafoveali e di *spillover*. Se da un lato i risultati mostrano chiaramente che sono i fattori lessicali (in questo caso, la frequenza di parola) a influenzare il *processing* sia in termini di tempi di lettura che di programmazione saccadica per saltare la parola successiva (*skipping probability*), la familiarità ortografica sembra avere un ruolo nei cosiddetti effetti parafoveali sull'elaborazione foveale (*parafoveal on foveal effects*). Infatti, è stato riscontrato un effetto, seppur limitato, sulla durata della fissazione precedente alla parola critica. Gli autori suggeriscono, quindi, che i fattori puramente visivi e, in qualche modo, "pre-lessicali" possano essere processati in parallelo (come nel modello SWIFT) oppure, ammettendo un modello seriale, in fase pre-attentiva (come abbiamo visto, un punto cardine del modello EZ-Reader). Il concetto di familiarità ortografica non va confuso con quello di **familiarità** della parola, esplorato in particolare in due studi (Chaffin et al., 2001; Williams, Morris 2004). Nel primo studio, Chaffin e colleghi (2001) selezionano una lista di parole (già usata in Chaffin, 1997) in base a un giudizio soggettivo di familiarità, ovvero chiedendo a 18 partecipanti esterni all'esperimento principale di giudicare la familiarità della parola su una scala da 1 a 5. Le parole target sono poi organizzate in triplette: una altamente familiare, una a bassa familiarità e una pseudo-parola fonotatticamente plausibile sia rispetto alla lingua target che alla categoria grammaticale delle sue controparti. La parola è inserita in un contesto di due frasi, dove la prima frase contiene degli disambiguanti e la seconda frase contiene un sinonimo o un iperonimo della parola target. Si opera così una sorta di triangolazione, per cui il carico cognitivo misurato sull'ultima area di interesse non dovrebbe variare in funzione della familiarità della parola critica nel caso in cui il contesto serva a fare luce sul suo significato.

6. Joe picked up the *guitar/zither/asdor* and began to *strum* a *tune*.

'Joe prese la *chitarra/cetra/asdor* e iniziò a strimpellare una melodia.'

He played the *instrument* to relax.

'Suona lo *strumento* per rilassarsi.'

Come si nota già dall'impostazione, gli obiettivi della ricerca sono molteplici e toccano processi di altra natura, oltre a quella puramente

lessicale, come quelli di implicazione e di recupero di informazioni dal contesto, che si intreccia a sua volta con giudizi di ordine semantico e di predicibilità. Per quel che riguarda strettamente l'opposizione di familiarità, basti osservare qui che le parole altamente familiari registrano durate più basse per prima fissazione, *gaze duration* e nello *spillover* rispetto alle due controparti. Non si notano, invece, differenze significative tra parole a bassa familiarità e pseudo-parole, per cui gli autori ipotizzano che il *familiarity check* contemplato nell'E-Z Reader possa aver interpretato le due parole in maniera simile data la uguale plausibilità ortografico-visiva. Un altro dato interessante riguarda, infine, le *late measures*: le pseudo-parole attraggono più regressioni rispetto a entrambe le classi di parole familiari e i tempi di lettura e il numero di visite sono maggiori per le parole a bassa familiarità rispetto a quelle altamente familiari. Per indagare più dettagliatamente la distinzione tra familiarità e frequenza, Williams e Morris (2004) propongono un nuovo studio di lettura di frasi in quattro condizioni, secondo l'esempio:

7. Kate chose the *knife* as her weapon.  
Alta frequenza – molto familiare  
'Kate scelse il *coltello* come arma.'
8. Kate chose the *rifle* as her weapon.  
Alta frequenza – familiare  
'Kate scelse il *fulmine* come arma.'
9. Kate chose the *dagger* as her weapon.  
Bassa frequenza – familiare  
'Kate scelse il *pugnale* come arma.'
10. Kate chose the *lance* as her weapon.  
Bassa frequenza – meno familiare  
'Kate scelse la *lancia* come arma.'

Le parole critiche sono state selezionate sulla base di tre parametri: la frequenza, estratta dalle liste di Francis e Kučera (1982); un questionario di familiarità, in cui veniva chiesto quanto spesso avessero visto una determinata parola indicando un punto da 1 a 7 su una scala Likert; infine, un test lessicale per verificare l'effettiva conoscenza della parola. Uno dei dubbi lasciati in sospeso dallo studio precedente (Chaffin et al., 2001), infatti, era proprio la mancata differenza nelle misure di *processing* iniziale, tra parole a bassa familiarità e pseudo-parole, per cui gli autori ipotizzano che in entrambi i casi le parole

fossero totalmente sconosciute ai partecipanti. Per questo, Williams e Morris (2004) scartano le parole con i giudizi di familiarità più bassi (il primo quartile), assicurandosi in questo modo che le parole fossero effettivamente conosciute dai partecipanti. Infine, come si nota, per non influenzare le misure di primo accesso, le frasi di supporto sono identiche nella parte precedente l'elemento critico, mentre nella parte successiva si rivelano molto informative riguardo alla parola critica, proprio per esaminare i rapporti tra familiarità e co-testo. I risultati mostrano che sia nelle misure di primo accesso (durata della prima fissazione, numero e durata delle fissazioni in prima lettura, *spillover*) che nelle *late measures* (tempi di rilettura, numero di regressioni verso la parola target) la familiarità sembra essere un fattore discriminante solo all'interno delle parole a bassa frequenza, mostrando un incremento nella condizione di bassa familiarità. Nessuna differenza, invece, si registra tra parole ad alta e bassa frequenza nel caso in cui la familiarità venga mantenuta costante: questo dato è forse il più interessante, poiché suggerisce che l'effetto frequenza può essere inibito dalla familiarità. In altre parole, il dato soggettivo (il giudizio di familiarità del singolo parlante) può prevalere sul dato oggettivo (la frequenza di una parola nella lingua) in termini di sforzi di elaborazione in lettura. Un'ultima considerazione va fatta, poi, sul rapporto tra parola e co-testo: come gli autori notano, infatti, le parole a bassa familiarità incoraggiano il lettore a raccogliere indizi dal contesto linguistico, influenzando in questo modo non solo le regressioni sulla parola target ma anche i tempi di lettura e rilettura delle altre porzioni della frase che possono essere informative per chiarirne il significato. Seguendo la direzione opposta, ovvero gli effetti che il contesto ha su una parola, un'altra caratteristica cruciale in lettura è la **predicibilità** (Ehrlich e Rayner, 1981; Balota et al., 1985; Rayner e Well, 1996, Frisson et al. 2005). Questa viene solitamente misurata attraverso dei test di completamento (cloze), per cui tanto più è predicibile una parola quanto più verrà indicata da diversi partecipanti per completare una frase data. Si nota, quindi, che la predicibilità è strettamente dipendente dal contesto. Rayner e Well (1996) utilizzano i cloze test per elaborare tre fasce di predicibilità a seconda della percentuale associata a una parola di apparire in un dato contesto. Ad esempio, in una frase come "Ha spedito la lettera senza un \_\_\_\_\_ quindi non è arrivata" (eng. *He mailed the letter without a \_\_\_\_\_ so it didn't arrive*), la parola "francobollo" (eng. *stamp*) è stata indicata il 98% delle volte, mentre la parola "indirizzo"

molto meno (20%), collocandosi in una fascia di predicibilità media. Gli autori osservano che le parole fortemente condizionate dal contesto, ovvero altamente predicibili, vengono fissate per meno tempo -sia in prima lettura (*gaze duration*) che in totale- e hanno più probabilità di essere saltate (ricordiamo che le parole scelte in questo studio rimangono tra i 5 e i 7 caratteri di lunghezza). McDonald e Shillcock (2003a) adottano una prospettiva diversa e indagano la predicibilità non come fattore dipendente dal contesto ma come conseguenza della forza di associazione tra le parole. Questo tipo di dato non è estratto da giudizi soggettivi, come sono quelli di un cloze test, ma da misure oggettive di co-occorrenza nei corpora linguistici (McDonald e Shillcock, 2003b). Per riferirsi a questo tipo di predicibilità, i due autori parlano di *transitional probability*, che può essere sia progressiva (come la probabilità di *wreck* di essere seguito da *havoc*) che regressiva (seguendo il loro esempio, tra *year* e *picture*, la prima ha più probabilità di essere preceduta da *last*). Come detto, gli autori si basano su corpora, nel loro caso specifico estratti da articoli di giornale, e si concentrano in particolare su combinazioni verbo-nome. I risultati mostrano effetti facilitatori soprattutto nelle misure di primo accesso (durata della prima fissazione e *gaze duration*), ma nessuna significatività si riscontra nella *skipping rate*. Per spiegare ciò, McDonald e Shillcock (2003b) distinguono la *transitional probability* -che viene descritta come una predicibilità di livello superficiale, largamente data da fattori visivi e statistici- dalla predicibilità data dal testo che, in quanto risultante da integrazioni di significato post-lessicali, si colloca ad un livello di elaborazione più profondo. Nonostante questa visione sia dibattuta (cfr. Frisson et al., 2005), una fetta cospicua di studi è stata poi dedicata alle cosiddette **unità multiparola** (*multi-words units*). Con questo termine ombrello ci si riferisce in genere a una classe di unità costituite da più parole che occorrono generalmente insieme, come ad esempio modi di dire (tagliare la corda), collocazioni (bandire un concorso), binomi (bianco e nero), formule rituali (buon appetito) e proverbi. Diversi studi hanno identificato effetti facilitatori nell'elaborazione di queste unità (per una rassegna, Conklin e Schmitt, 2012; Siyanova-Chanturia, 2013; Carrol e Conklin, 2014; Carrol e Conklin, 2020), utilizzando metodologie sia offline che online. Benché in questo campo le ricerche con eye tracker si siano concentrate soprattutto sul confronto tra lingua materna e lingua seconda (che noi affronteremo nel dettaglio nel capitolo 6), sappiamo che i movimenti oculari in lettura sono influenzati dalle unità

fraseologiche in almeno tre modi. Underwood et al. (2004), i primi a indagare la fraseologia con questa metodologia, prendono in esame 20 sequenze lunghe dalle 4 alle 8 parole, il cui significato complessivo non è deducibile dalla somma dei significati dei singoli componenti (*idioms*). Partendo dagli studi sulla predicibilità, la loro domanda di ricerca interroga l'unità fraseologica come contesto, ipotizzando che i suoi componenti (in questo caso specifico, l'ultimo di una stringa di 4-8 parole) siano elaborati più facilmente rispetto a quando si trovano in contesti non fraseologici (e hanno, quindi, un significato a sé e non come parte di un'unità multiparola). Osserviamo l'esempio seguente, preso da Underwood et al. (2004):

11. You've been talking in circles for 30 minutes trying to tell me something. Please don't beat around the *bush* for another half an hour, but just get to the point and tell me! If it was you who dropped my flower *basket*, don't worry because I won't be angry with you.  
'Sono 30 minuti che parli a vuoto cercando di dirmi qualcosa. Per favore, non menare il can per l'aia per un'altra mezz'ora, ma vai al punto e dimmelo! Se sei stato tu a far cadere il mio cesto di fiori, non preoccuparti perché non mi arrabbio con te.'

Qui le aree di interesse sono due: la parola *bush* come parte finale del modo di dire *beat around the bush*; la parola *basket* in un contesto non fraseologico, confrontata con la stessa parola in *put all your eggs in one basket* contenuta in un altro passaggio (analogamente alla parola *bush* che sarà inserita in un altro brano e confrontata con quelle facente parte dell'unità fraseologica). I risultati mostrano un minor numero di fissazioni e tempi di lettura ridotti per le parole inserite in un contesto fraseologico, confermando il ruolo della predicibilità anche in sequenze corte ma altamente convenzionali come gli *idioms*. Un secondo modo di indagare le unità fraseologiche consiste nel considerarle nel loro insieme, studiandone le proprietà proprio come viene fatto con le parole singole. Infatti, oltre alle caratteristiche dei singoli componenti dell'unità, è stato dimostrato come anche le proprietà globali dell'intera sequenza abbiano degli effetti sulla lettura. Siyanova-Chanturia, Conklin e Van Heuven (2011) considerano la frequenza dell'unità fraseologica e per farlo scelgono il binomio, un tipo di unità formato da due parole che mantengono il loro significato all'interno del binomio.

Quest'unità, infatti, è semanticamente trasparente, nel senso che il suo significato è dato dalla somma dei significati dei suoi componenti o, più precisamente, anche quando il significato globale è figurato, rimane comunque deducibile dai significati delle singole parole ("cane e gatto")<sup>11</sup>. In più, questi componenti nella lingua non solo sono frequentemente accoppiati (è un binomio "bianco e nero", ma non "rosso e viola") ma appaiono più spesso in un determinato ordine ("sale e pepe", "burro e marmellata"). Invertendo i termini, quindi, è possibile ottenere un binomio dallo stesso significato ma con frequenza molto minore. Comparando forme come *bride and groom* e *groom and bride*, gli autori osservano che la frequenza globale della sequenza influisce significativamente sulla durata delle fissazioni, sia in prima lettura (*first-pass reading time*) che in totale (*total reading time*). Un'altra caratteristica indagata nelle unità multiparola è la pluralità di significati, nello specifico quelle sequenze che possono avere un'accezione letterale e una figurata (come in italiano potrebbe essere il caso di "gettare la spugna" o "tagliare la corda" a seconda del contesto). Siyanova-Chanturia, Conklin e Schmitt (2011) comparano tre tipi di sequenze: idiomatica con significato figurato (*at the end of the day* con il significato di "alla fine"); idiomatica con significato letterale (*at the end of the day* con il significato di "alla fine della giornata") e non idiomatica (*at the end of the war*). Si nota un incremento delle *late measures* (tempo di lettura totale e numero di fissazioni) nel caso delle espressioni non idiomatiche, che indica un minore sforzo cognitivo nonché un minor bisogno di riletture e regressioni del lettore in presenza di formule fisse. Questi stessi effetti non si ritrovano, però, né nelle *early measures* (tempo della prima lettura) né nella distinzione tra senso letterale e figurato. Gli autori ipotizzano che nel caso specifico di questo esperimento, il contesto fortemente disambiguante permette al lettore di prevedere già a quale significato si fa riferimento nel testo. Nel caso di contesto neutro, però, potrebbe esserci un effetto frequenza con specifico riferimento a ciascun significato della sequenza, ma questa ipotesi non è stata finora ulteriormente esplorata. Una questione ancora aperta in questo ambito riguarda il grado di lessicalizzazione delle unità multiparola, ovvero se queste siano effettivamente immagazzinate -e quindi processate- come un'unica parola. A questo proposito, Cutter et al. (2014)

---

<sup>11</sup> Gibbs e Nayak (1989) ipotizzano che vi siano diversi gradi di fissità sintattica a seconda del grado di trasparenza semantica di ciascun idiom.

indagano quelle combinazioni che in letteratura anglofona vengono chiamati *spaced compounds*, ovvero delle parole composte formate da due unità separate da uno spazio allo scritto (ad esempio, *bathroom* è una parola composta come ‘portaombrelli’, mentre *teddy bear* è uno *spaced compound* come ‘frutti rossi’). Per farlo, usano il *boundary paradigm* originariamente proposto da Rayner et al. (1975) per verificare i *preview benefit*, ovvero eventuali effetti facilitatori dovuti alla visione parafoveale della parola successiva. In questa impostazione, la parola target ( $n+1$ ) viene preceduta da una linea di confine invisibile e, in una delle due condizioni sperimentali, viene modificata e rimpiazzata con una parola graficamente simile (ad esempio, invece di ‘palazzo’ viene messo ‘gafoure’: non è necessario, quindi, che sia fonotatticamente consona alla lingua target né che sia foneticamente simile). Quando l’occhio varca quella soglia invisibile, la parola target torna a essere quella corretta. Come si è visto, questa viene letta più velocemente nella condizione non modificata, benché l’occhio abbia accesso alla parola ‘scorretta’ solo in posizione di anteprima, ovvero quando fissa  $n$ , la parola pre-target. Utilizzando questo stesso paradigma, Cutter et al. (2014) tentano di dimostrare come alcune unità composte, sebbene costituite da due parole grafiche, siano in realtà lessicalizzate come un’unità (in basso riportiamo l’esempio delle quattro condizioni sperimentali per *teddy bear*). I due autori ipotizzano, quindi, che in questi casi la lettura benefici dell’anteprima corretta non di una, ma di due parole grafiche alla destra della parola fissata ( $n+2$ ).

The small child gently cuddled his fluffy	teddy bear while trying to get to sleep.
The small child gently cuddled his fluffy	teddy hocu while trying to get to sleep.
The small child gently cuddled his fluffy	fohbg bear while trying to get to sleep.
The small child gently cuddled his fluffy	fohbg hocu while trying to get to sleep.

Fig. 3.1. La figura è tratta da Cutter et al. (2014)

Confermando le loro ipotesi, lo studio dimostra che  $n+2$  induce degli effetti facilitatori nell’elaborazione di  $n+1$  solo nel caso in cui questa sia disponibile e segnali, perciò, l’inizio di un’unità multiparola, seppur in anteprima: in altre parole, quando l’occhio arriva alla parola target  $n+1$ , questa verrà letta più velocemente se la parola seguente  $n+2$  è corretta, nonostante l’occhio abbia avuto informazioni sulla correttezza/incorrettezza di  $n+1$  e  $n+2$  solo in visione parafoveale (cioè prima che l’occhio sorpassi la linea). Visti i risultati talvolta contrastanti

e incerti ottenuti dalle ricerche sul beneficio da anteprema esteso a due parole, gli autori interpretano questi dati come una conferma del fatto che gli *spaced compound* sono in effetti assimilabili ad un'unica parola.

Al confine tra analisi della parola singola e integrazione del contesto (linguistico e non), troviamo infine la plausibilità e l'ambiguità semantica. Lo studio forse più noto sulla **plausibilità** è quello di Rayner et al. (2004), poi seguito da diversi altri, tra cui Warren, McConnell (2007), Matsuki et al. (2012) e Abbot, Staub (2015). Rayner e colleghi hanno indagato la plausibilità in termini di improbabilità (13) e anomalia (14), esemplificato nella tripletta che segue:

12. John used *a knife to chop* the large *carrots* for dinner.
13. John used *an axe to chop* the large *carrots* for dinner.
14. John used *a pump to inflate* the large *carrots* for dinner.

Come spiegano gli autori, la condizione di implausibilità prevede un legame coerente tra il predicato e l'oggetto (ovvero, una carota che viene affettata) ma sproorzionata tra l'oggetto e il paziente (l'ascia e la carota), data la nostra conoscenza del mondo. La condizione di anomalia, invece, investe il legame tra predicato e paziente (la carota che viene gonfiata), rendendo l'evento impossibile nella realtà (ammettendo che l'evento non sia una cena a tema ortofrutta con dei palloncini a forma di carota!). Sebbene i risultati non mostrino una variazione nel primo accesso lessicale (corrispondente generalmente alla durata della prima fissazione), gli effetti dell'anomalia sono in realtà visibili abbastanza precocemente (rilevabili, infatti, dalla *gaze duration*), entro circa 300 ms dal primo incontro con la parola target (i.e., *carrot*) in contesto anomalo; la plausibilità, invece, sembra avere effetti solo dopo, causando un incremento dei tempi di lettura totali sulla parola target in (14). L'**ambiguità semantica** viene inizialmente studiata da Rayner, Duffy (1986; 1987) che osservano un incremento della durata delle fissazioni in prima lettura (*gaze duration*) sulle parole ambigue rispetto alle parole controllo. Gli autori operano un'ulteriore distinzione tra parole ambigue con due significati ugualmente frequenti nella lingua e parole ambigue con un significato dominante. Nel loro esperimento, il contesto informativo (disambiguante) segue la parola target, mentre il contesto precedente è neutro. Siccome i tempi di lettura sono aumentati nel caso di parole ambigue con due significati ugualmente frequenti, ma non differiscono tra controlli e parole ambigue con un significato domi-

nante, i risultati suggeriscono che in contesto neutro i diversi significati di una parola vengono selezionati simultaneamente, risultando in un maggiore sforzo cognitivo (e, quindi, tempi di lettura allungati). In questo quadro, la parola ambigua con un significato dominante non differisce dalla parola controllo e non presenta alcuna difficoltà di accesso. Questi risultati sono stati confermati e ampliati da Duffy, Morris e Rayner (1988) che aggiungono un'ulteriore condizione all'esperimento, ovvero l'opposizione tra contesto neutro e contesto disambiguante prima della parola target. D'accordo con le ipotesi avanzate, in questo caso sono le parole ambigue con un significato dominante a mostrare tempi di lettura più lunghi, dato che il contesto seleziona espressamente il significato meno frequente (e, quindi, meno atteso in termini di accesso lessicale). Questo fenomeno in particolare viene chiamato da Rayner et al. (1994) *subordinate bias effect* e sta a dimostrare, secondo gli autori, che l'accesso al significato di una parola non avviene in maniera selettiva (partendo, cioè, dal contesto), ma in maniera automatica e indipendente per cui, anche dopo aver letto un contesto fortemente informativo, il costo cognitivo nel selezionare il significato meno frequente resta comunque alto. La risoluzione dell'ambiguità semantica è stata estensivamente studiata nella ricerca con eye tracking, per indagare sia le vie di accesso lessicale (cfr. Sereno, Pacht, Rayner, 1992; Sereno, 1995) sia la modulazione dei significati delle singole entrate lessicali (tra gli altri, cfr. Pickering et al. 2004; Frisson e Pickering (1999) sulla metonimia; Foraker, Murphy (2012) sulla polisemia).

In questa breve rassegna, si è voluto dare uno sguardo alle proprietà lessicali che influenzano i movimenti oculari e in particolare i tempi di fissazione associati ai costi di elaborazione. Avere una comprensione globale di queste proprietà è di cruciale importanza per padroneggiare l'eye tracking come metodologia di indagine di linguaggio e cognizione. L'esplorazione rigorosa dell'elaborazione linguistica parte innanzitutto da una calibrazione cosciente dello stimolo sperimentale e da un'attenta selezione delle parole, anche di quelle non propriamente oggetto di indagine.

### 3.3. Elementi di base di un esperimento di reading

Nel progettare gli stimoli di testo ci sono alcune misure da adottare per evitare perdita di dati e garantire la massima leggibilità. Se lo studio non è incentrato sul livello testuale, ma è teso a indagare parole

singole nel loro prossimo contesto o la relazione tra parole adiacenti, è sempre preferibile mantenere l'unica frase su un solo rigo. Quando, invece, è necessario prendere in considerazione brani più lunghi, occorre prestare particolare attenzione alla spaziatura per impedire che le lunghe saccadi di ritorno a capo si sovrappongano al testo. In alcuni casi, poi, può capitare che le fissazioni che si susseguono sul testo vengano registrate sopra o sotto l'effettiva linea su cui avvengono: avere un margine di spazio tra due linee di testo, in questi casi, permette al ricercatore di assegnare il tracciato oculare al testo corrispondente senza errori di interpretazione. Che sia su una o più righe, il testo dovrebbe essere sempre centrato in verticale e le aree di interesse non dovrebbero presentarsi né all'inizio né alla fine di frase o prima di punteggiatura per evitare che il cosiddetto *wrap-up effect* (Rayner et al., 2000; Hirotsu et al., 2006) induca fissazioni più lunghe. Inoltre, in linea generale, è bene prevedere dei margini vuoti, poiché alle estremità dello schermo la registrazione dei movimenti oculari tende a essere meno accurata. Per i caratteri, usare un font monospaced (come il Courier New) è sempre l'opzione migliore, perché ogni lettera ha la stessa larghezza. Infine, per evitare che l'occhio si stanchi, è preferibile evitare schermi bianchi e usare toni più riposanti, come uno schermo grigio con scritte nere o uno schermo nero con scritte gialle. Accanto al controllo di parametri più strettamente visivi, di primaria importanza è anche il bilanciamento tra le variabili linguistiche. Come abbiamo visto nel paragrafo precedente, infatti, i tempi di lettura sono fortemente influenzati da tre fattori fondamentali: lunghezza, frequenza e predicibilità. Nel paragonare condizioni diverse, dobbiamo essere sicuri che in uno stesso *item* le parole nelle nostre aree di interesse siano paragonabili sotto ogni punto di vista eccetto la variabile manipolata. A seconda di quale sarà la nostra domanda di ricerca, quindi, le parole o le combinazioni che andremo a comparare dovranno essere comparabili in lunghezza, frequenza, categoria lessicale, astrattezza e così via. Se, ad esempio, vogliamo studiare la diversa predicibilità di uno delle accezioni di una stessa parola, sarà solo il contesto a cambiare. Se, invece, pensiamo a uno studio sui falsi amici in una L1 e una L2, dobbiamo inderogabilmente tener conto della variazione in frequenza, contesto d'uso e altre variabili linguistiche potenzialmente pertinenti (e.g., astrattezza, età di acquisizione) nel progettare le nostre coppie di stimoli. Ove questo non fosse possibile, lo sperimentatore può effettuare un controllo a posteriori, inserendo le possibili variabili confondenti

nel modello statistico per discernere gli effetti dalla nostra variabile indipendente<sup>12</sup>. Per l'italiano, ci sono varie risorse a cui possiamo fare riferimento per l'estrazione di queste misure. Il valore di frequenza generalmente si ricava da un corpus, ovvero una raccolta di testi formata per essere rappresentativa di una varietà linguistica e, attualmente, i corpora disponibili per l'italiano sono molteplici<sup>13</sup>. Una volta decisa la struttura degli *item* adatti al nostro oggetto di indagine, il numero di *item* da inserire dipenderà anche dal numero di partecipanti previsti. Per incrementare la potenza statistica, infatti, è necessario ingrandire il proprio dataset, ovvero la raccolta dei punti di osservazione di un dato esperimento, derivante dall'incrocio del numero dei partecipanti per il numero di *item*. Brysbaert e Stevens (2018) raccomandano, per ricerche su tempi di reazione, un dataset di 1600 osservazioni, mentre Godfroid (2019) osserva che la dimensione media degli studi con eye tracking in L2 è di 300 osservazioni. Date queste premesse, la scelta del numero di *item* risponde necessariamente anche ad esigenze diverse, come l'adozione di un disegno tra gruppi (*between*) o entro gruppi (*within*)<sup>14</sup> e la popolazione target (ad esempio con partecipanti L2 di livello intermedio o con bambini, è sempre meglio non eccedere nei tempi di durata dell'esperimento o di prevedere pause tra blocchi di *item*). Ogni *item* è contenuto in un *trial*, ovvero una sequenza di eventi che si ripete sempre uguale, come ad esempio la lettura di una frase con una domanda di comprensione subito dopo. Questa articolazione è dovuta al fatto che un esperimento presenta generalmente due compiti, uno primario e uno secondario. Il primo compito è quello che il partecipante deve fare per restituirci i dati di nostro interesse: guardare delle immagini su uno schermo mentre ascolta delle frasi registrate,

---

<sup>12</sup> La variabile manipolata dallo sperimentatore è detta variabile indipendente (ad esempio frequenza o astrattezza in coppie di parole), mentre le misure che cambiano in conseguenza di questa manipolazione sono dette variabili dipendenti (ad esempio, la durata delle fissazioni).

<sup>13</sup> Con specifico riferimento ai corpora monolingui di italiano scritto, attualmente il corpus esteso è il TenTen20: formato da 12 miliardi di parole e raccolto a partire da testi tratti dal web, è disponibile su SketchEngine. Come corpora liberamente accessibili, troviamo invece LaRepubblica (<http://sslmit.unibo.it/repubblica>), costituito da circa 380 milioni di token, e Paisà (<https://www.corpusitaliano.it/>), con circa 250 milioni di token.

<sup>14</sup> Un disegno sperimentale tra gruppi (*between-subjects*) prevede che ogni gruppo sia sottoposto a una sola condizione sperimentale. Al contrario, in un disegno entro gruppi (*within-subjects*) ogni gruppo osserva tutte le condizioni. Un esperimento che utilizza entrambe le modalità ha un design misto (*mixed*).

leggere una frase o un brano. A volte, per esigenze diverse, questa consegna può essere affiancata da un'altra, ad esempio rispondere a delle domande di comprensione su quanto letto o dare dei giudizi di grammaticalità/accettabilità. Se l'obiettivo dello sperimentatore è assicurarsi che l'attenzione del partecipante si mantenga a un livello alto durante tutto l'esperimento, il *task* più semplice ed efficace sono le domande di comprensione con risposta sì/no. In questo caso, la percentuale di risposte esatte, proprio perché queste misurano l'attenzione nella lettura e il coinvolgimento nell'esperimento, può anche valere come criterio di esclusione dei partecipanti (ad esempio, fissando una soglia del 75% di risposte esatte). Altri compiti di comprensione possono essere il *re-telling*, in cui si chiede al partecipante di riformulare quanto letto a parole proprie, o la traduzione. Quest'ultima, però, come anche avviene nel caso dei giudizi di grammaticalità, può guidare l'attenzione del lettore verso strutture e punti precisi del testo. Bisogna, quindi, valutare attentamente il tipo di lettura che si vuole elicitare (naturale, focalizzata sulla morfosintassi, mirata a svelare la conoscenza esplicita di una L2...) e, in base all'obiettivo della nostra ricerca, scegliere il *task* più appropriato. L'esperimento generalmente si apre con un piccolo numero (meno di 10) di *item* di prova, chiamati *practice trial*, il cui obiettivo è di far familiarizzare il partecipante con l'esperimento, sia con la modalità che con la strumentazione. In un esperimento con eye tracker, ad esempio, è particolarmente importante la posizione assunta dal partecipante, poiché è auspicabile che resti fermo (per quanto possibile) per tutta la durata dell'esperimento a partire dalla calibrazione. In questa prima parte di prova, quindi, è possibile che il partecipante non si senta ancora del tutto a proprio agio nella posizione, a volte un po' innaturale, data la presenza della mentoniera o di altri supporti e si trovi ancora a fare piccoli movimenti di assestamento. In questa fase, è anche possibile che chieda chiarimenti, ad esempio sui tasti da premere per andare avanti o per rispondere alle domande, trovandosi non solo a falsare i tempi di lettura ma anche a rivolgere lo sguardo involontariamente al di fuori dello schermo per cercare lo sperimentatore. Una volta terminata questa sezione, si può inserire una schermata di testo in cui si avvisa il partecipante della fine della prova e l'inizio dell'esperimento vero e proprio. I *practice trial* vengono scartati in fase di analisi dati. L'esperimento a sua volta consiste in una commistione bilanciata di *trial* sperimentali (a partire dagli stimoli creati e manipolati dallo sperimentatore sulla base delle domande di ricerca) e *filler*.

Questi ultimi sono anche *trial* “di scarto” che servono principalmente a mascherare l’obiettivo dell’esperimento, di modo che il partecipante non si accorga della variabile manipolata negli stimoli critici. È molto importante, infatti, che il partecipante sia all’oscuro dei dettagli riguardanti lo specifico oggetto di indagine, per evitare che la sua attenzione venga rivolta alle strutture di interesse falsando, in questo modo, i tempi di lettura. Per questo motivo, i *filler* dovrebbero essere simili agli stimoli critici ma non contenere le manipolazioni di interesse. Generalmente, la percentuale di *filler* dovrebbe essere compresa tra il 50% e il 75% dei *trial* totali dell’esperimento, tenendo sempre conto, come osservato in precedenza, di non gravare troppo sulla lunghezza della sessione, soprattutto in caso di popolazioni di lettori meno esperti.

## 4. Scene perception e visual search

*Veronica D'Alesio*

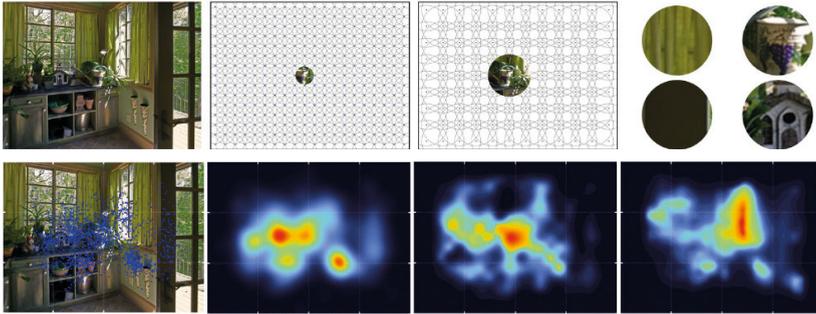
### 4.1. Scene perception

In psicolinguistica non si lavora solo con parole, lettere e frasi ma anche con immagini e oggetti del mondo reale. Di particolare rilevanza per l'eye tracking è il vasto campo di studi della percezione visiva, di cui qui faremo dei brevi accenni riguardanti l'esplorazione di scene e il riconoscimento di oggetti. Benché esistano ricerche su *task* naturali che implicano manipolazioni sul mondo reale (come lavarsi le mani, fare un panino o guidare), con 'scena' indicheremo qui una rappresentazione del mondo. Nello specifico, usiamo la definizione di Henderson e Ferreira (2004) secondo cui una scena è una rappresentazione semanticamente coerente del mondo reale in scala umana, comprendente elementi di sfondo e oggetti discreti. Introducendo la nozione di 'scala umana', gli autori limitano questa definizione di scena alla rappresentazione di un ambiente dove una persona potrebbe muoversi e agire, essendovi immersa: in questo modo, rappresentazioni più piccole, come una scrivania o l'interno di un cassetto, non si configurano come vere e proprie scene, ma come punti di vista dai quali vedere uno o più oggetti. Gli elementi di sfondo sono generalmente entità amovibili, come pareti, pavimenti o elementi paesaggistici, come il cielo o una vallata. Al contrario, gli oggetti sono entità discrete e mobili che, per far parte di una scena reale che possa definirsi tale, devono essere presentati in maniera plausibile in base alle loro caratteristiche sia fisiche (ad esempio, peso o grandezza, per cui è inverosimile trovare un'incudine sorretta da una piuma) che semantiche (è plausibile vedere una rondine in cielo, ma non un asino). Come abbiamo detto, inoltre, una caratteristica definitoria della scena è la coerenza seman-

tica, intesa come una sorta di collegamento, un *fil rouge*, che accomuna tutti gli elementi come appartenenti a quel contesto. È ormai indiscusso, infatti, che la prima cosa che viene catturata dallo sguardo è il cosiddetto *gist*, che in letteratura indica il concetto generale, l'idea di fondo (vedi anche Oliva, Torralba, 2006). La categorizzazione semantica di una scena avviene molto presto, sicuramente entro i limiti della prima fissazione, in un tempo stimato tra i 30 e i 50 ms (Henderson e Ferreira, 2004). È facile immaginare come quest'essenza generale dell'ambientazione non si ritrovi, invece, nelle scene forse più usate in psicolinguistica, ovvero la disposizione di oggetti isolati all'interno di una matrice neutra. Henderson e Ferreira (2004) parlano a questo proposito di *ersatz scene*, un termine che indica tutte quelle scene senza coerenza interna (perché senza sfondo, o i cui oggetti non sono in proporzione tra loro, ad esempio). Tra una scena vera e propria e una semplice sequenza di oggetti sconnessi tra loro, ci sono varie gradazioni, tra cui foto o disegni più o meno stilizzati di un ambiente (*sketch*) e anche gruppi di oggetti che, seppur collocati in una matrice neutra, sono in relazione tra loro, ad esempio, sono tutti giocattoli o tutti oggetti d'arredo di un salotto (*interpretable arrays*). Forse la domanda più importante che ci dobbiamo porre di fronte a qualsiasi tipo di scena è: sono i fattori visivi o quelli cognitivi a guidare l'occhio? Che l'occhio non si muovesse in maniera casuale sulle immagini si sa dagli albori della disciplina (anzi, da ancora prima che la tecnologia di eye tracking attuale esistesse): già Buswell (1935), Yarbus (1967) e Mackworth e Morandi (1967) – poi replicato da Antes (1974) – osservarono che le fissazioni tendevano a posarsi sulle parti dell'immagine considerate "informative", ovvero il cui contenuto fosse in qualche modo utile o di interesse per la comprensione della scena. Con l'avvento dell'eye tracking come lo conosciamo oggi, l'ipotesi forse più esplorata e indagata è quella della salienza visiva. Nelle cosiddette *saliency maps* (Itti e Koch, 2000), la scena viene computata in una mappa generata a partire dalle caratteristiche fisiche degli oggetti presenti (orientamento spaziale, luminanza, colore). Tuttavia, questo approccio strettamente *bottom-up* (ovvero basato solamente su dettagli visivi e non anche su fattori cognitivi) risulta avere diversi punti dubbi, tra cui il non tenere conto della durata delle fissazioni (che secondo alcuni è strettamente legata all'elaborazione dell'informazione portata dall'oggetto) e la non adattabilità ad ogni tipo di *task* (con riferimento ad entrambi i punti, si vedano gli studi su *task* reali in Land et al., 1999; Hayhoe et al., 2003).

Nell'ambito della percezione della scena, di cui ci occupiamo in questa sede, la durata delle fissazioni in relazione al significato degli oggetti è stata indagata particolarmente tramite lo studio dell'anomalia semantica. Questo paradigma opera su coppie di scene identiche, tranne che per un oggetto, che da una parte è semanticamente coerente con l'ambientazione, dall'altra è palesemente fuori posto. Loftus e Mackworth (1978) osservano che oggetti non predicibili a partire dal contesto (e.g., un polpo in una fattoria, confrontato con un trattore in una fattoria) vengono guardati prima e più a lungo. De Graef et al. (1990) indagano diversi tipi di violazione dell'oggetto nella scena (dimensione, base di appoggio, posizione e plausibilità con il contesto), trovando degli effetti sulla durata totale delle fissazioni nelle violazioni di posizione e plausibilità, ma non nelle fasi iniziali. Analogamente, Henderson et al. (1999) riferiscono maggiori durate in *first gaze duration*, *second gaze duration*, tempo totale delle fissazioni nonché un numero significativamente più alto di entrate nell'area critica su oggetti semanticamente incongruenti con il contesto, mentre i risultati sulle prime fasi di osservazione sono più sfumati, suggerendo che gli oggetti plausibili siano addirittura trovati prima di quelli implausibili in un compito di *visual search* (nel secondo esperimento dello studio). Per quanto riguarda l'influenza del task, invece, come Tatler e colleghi (2011) osservano, è il concetto stesso di 'osservazione libera' (*free-viewing*) a sembrare controverso: così come sembra implausibile che una persona muovendosi e agendo nella vita reale guardi 'per guardare' e che il suo occhio sia guidato solo dalle caratteristiche esterne degli oggetti e non dalle proprie intenzioni (magari anche future, individuando spazi o oggetti che, benché poco salienti, saranno destinazione di un movimento o di una manipolazione, cfr. Pelz e Canosa, 2001), allo stesso modo il partecipante a cui viene chiesto di guardare senza alcuno scopo dichiarato o istruzione da seguire, potrebbe in realtà essere intento a fare collegamenti per conto suo, o a memorizzare gli elementi della scena o la loro disposizione. Henderson e Ferreira (2004) parlano di quattro principali fattori cognitivi che influenzano i movimenti oculari: memoria a breve termine, memoria a lungo termine, conoscenza generale di una scena (ad esempio, come è fatto un soggiorno) e conoscenza generale dello svolgimento di un dato compito. Infine, come si è visto da tutta la letteratura sul *visual world paradigm* (vedi capitolo 5), è ormai ben noto che lo sguardo risponda anche a stimoli di natura linguistica. Sulla base di queste osservazioni, diversi approcci tengono ora conto di una commi-

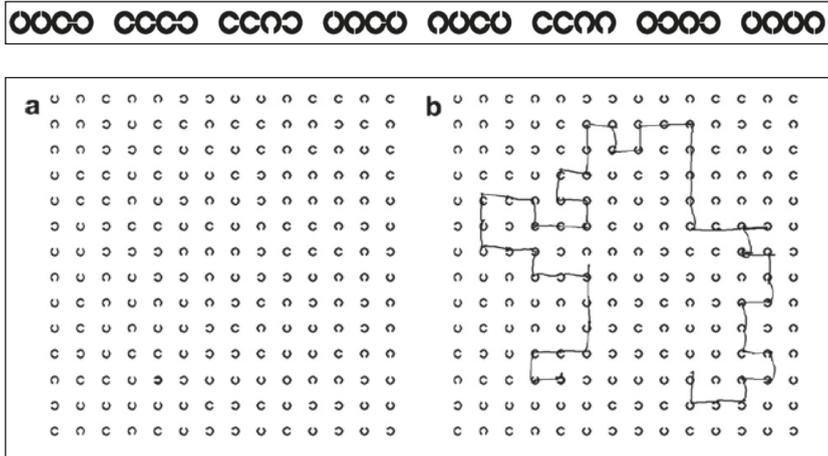
stione di fattori di basso e di alto ordine, non solo nei compiti di vita reale ma anche nell'esplorazione di scene (in particolare, Henderson e Hayes (2017) hanno argomentato approfonditamente la distinzione tra *saliency* e *meaning*).



**Fig. 4.1.** Henderson e Hayes (2017) pongono a confronto salienza e significato. Nella prima riga, la costituzione di una *meaning map*: l'immagine viene scomposta in piccole aree circolari parzialmente sovrapposte, e sottoposte al giudizio di 165 soggetti. Nell'immagine 1d, esempi di patch non significativi e significativi. Nella seconda riga, la comparazione tra le fissazioni di 65 soggetti (1a, 1b) con la *meaning map* (1c) e la *saliency map* (1d) estratte dalla stessa immagine.

## 4.2. Visual search

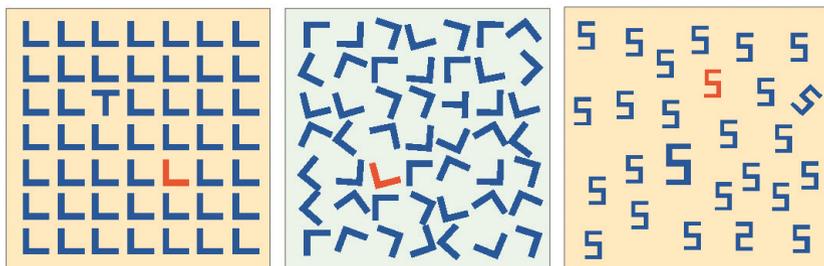
Esattamente come la lettura, la ricerca visiva è un compito quasi onnipresente nella nostra vita quotidiana. Ogni volta che cerchiamo un vestito nell'armadio, quell'ingrediente che non possiamo mangiare dietro a un pacco di biscotti, o il nome di una strada mentre guidiamo, ci troviamo nella situazione di avere una scena di fronte probabilmente ricolma di oggetti. Generalmente, i compiti di ricerca visiva in laboratorio sono tre: al partecipante viene chiesto di trovare un oggetto, di dire se quell'oggetto è presente o meno, o di seguire un percorso guidato da un punto a un altro della scena (cfr. fig. 4.2). Da un punto di vista oculomotorio, sappiamo che le saccadi sono più lunghe rispetto alla lettura ma meno ampie rispetto alla percezione della scena (Rayner, 2009). Similmente alla lettura, la fissazione tende a posarsi sul centro dell'oggetto (Findlay, 1997; Henderson, 1993; Trukenbrod e Engberd, 2007) e la loro durata (benché molto più variabile) rispecchia un tempo di elaborazione (Trukenbrod e Engber, 2007; Williams, Polatsek, 2007).



**Fig. 4.2.** Nella stringa in alto vediamo un esempio degli stimoli usati da Williams e Polatsek (2007), dove al partecipante viene chiesto di identificare l'unica O in una fila di C di Landolt (nell'intenzione degli autori, gli item disposti su una riga permettono la comparazione con la linearità della lettura). Nella immagine più sotto, vediamo gli stimoli utilizzati da Trukenbrod e Engberd (2007) in un compito di ricerca sequenziale: ai partecipanti viene chiesto di seguire con gli occhi il percorso che dall'unica C in grassetto porta all'unica O, seguendo la direzione dell'apertura della C. Entrambi gli studi hanno indagato il rapporto tra carico di elaborazione e durata della fissazione. Modificando la grandezza dello spazio nella C di Landolt, è stato osservato che alla riduzione di questo spazio (e quindi, alla difficoltà) corrispondono tempi di fissazione maggiori.

Oltre allo specifico tipo di task (van Zoest et al., 2004), nella ricerca l'occhio viene guidato principalmente da due aspetti: le caratteristiche dell'oggetto che si sta cercando e le caratteristiche della scena in cui è inserito. A partire dagli anni '80, il lavoro di Anne Treisman ha dato un contributo fondamentale agli studi sull'attenzione nella ricerca visiva. L'idea che alcune caratteristiche dell'oggetto guidino la ricerca è alla base della sua nota *Feature Integration Theory* (Treisman e Gelade, 1980). Nella sua prima versione, in seguito rivista, gli autori postulavano l'esistenza di due tipi di ricerca. Nel primo modello, l'oggetto salta all'occhio (*pop-out*) per una caratteristica fondamentale che lo distingue dagli altri oggetti presenti (chiamati comunemente in letteratura 'distrattori'): in questa modalità, la ricerca è agevolata, al punto da essere 'pre-attentiva', ovvero coinvolgendo solo l'attenzione implicita (*covert attention*). Nel secondo caso, invece, la ricerca è complicata dal fatto che le caratteristiche che fanno risaltare l'oggetto sono più di una per cui l'occhio dovrà 'scansionare' la scena (*visual scanning*) oggetto per oggetto. Anche questa modalità, chiamata *seriale* (in contrapposizione

alla prima, definita *parallela*) era pensata come guidata implicitamente. Tuttavia, ulteriori ricerche hanno perlopiù confutato sia la dominanza dell'attenzione implicita, che l'esistenza di due livelli distinti di ricerca (Findlay, 2004). Wolfe e Horowitz (2004; 2017) operano una rassegna esaustiva delle caratteristiche oggettuali capaci di guidare l'attenzione (da loro definite *guiding features*): mentre su colore, taglia, orientamento spaziale e movimento c'è accordo unanime in letteratura, tante altre rimangono ancora in dubbio (come la forma, la categoria semantica o i tratti del viso e delle espressioni facciali).



**Fig. 4.3.** Esempi di *guiding features* (Wolfe, Horowitz, 2004). Nel primo riquadro, è facile notare il 5 rosso (colore), così come quello obliquo (orientamento) e quello più grande (taglia); è più difficile, invece, trovare il 2 (numero). Nel secondo riquadro, è facile notare sia la L rossa (colore) che la T (lettera o differente tipo di incrocio tra due segmenti perpendicolari). Tuttavia, quest'ultimo criterio non regge alla variazione del contesto: nel terzo riquadro si nota ancora chiaramente la L rossa, mentre la T si perde, segno che la caratteristica 'lettera' o 'incrocio di segmenti' è solo una conseguenza dell'omogeneità dei distrattori.

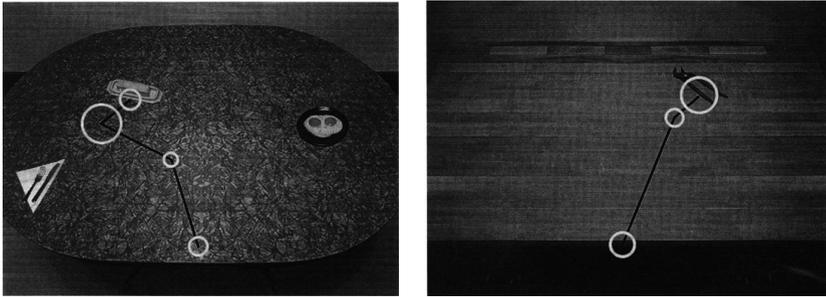
Come si vede dalla figura 4.3, però, talvolta non è la caratteristica in sé ad attirare l'attenzione, ma solo il suo confronto con la scena circostante. Si reintroduce qui il concetto di salienza, integrato non solo in alcuni modelli di percezione della scena (cfr. supra) ma anche propriamente di ricerca visiva (Rao et. al., 2002). A questo proposito, Duncan e Humphreys (1989) parlano di due condizioni specifiche che incrementano la difficoltà nel trovare un oggetto target: da un lato, la somiglianza tra il target e i distrattori, dall'altro l'eterogeneità dei distrattori. I due effetti, secondo gli studiosi, interagiscono tra loro creando un continuum di difficoltà più che una reale distinzione binaria tra ricerca seriale o parallela: ad esempio, la difficoltà derivante dalla similarità tra target e non target può essere compensata dalla somiglianza tra tutti i distrattori e viceversa. Bisogna ricordare che la maggior parte degli studi agli albori della disciplina tendevano a utilizzare come misura

principale i tempi di risposta (RT) in relazione al numero dei distrattori (*sample size*) e principalmente su queste misure si fondano i modelli sequenziali o seriali. Un esempio ne è la *Guided Search* di Wolfe (2007) che, come abbiamo appena visto, è basata sulla salienza del target rispetto ai distrattori. Altri studiosi, invece, hanno messo in discussione quest'impostazione, sostenendo che l'attenzione sia allocata in maniera parallela, ovvero simultaneamente su tutti gli oggetti della scena (ad esempio, Thornton e Gilden, 2007, Townsend 1972, 1990). Il cuore del dibattito (ben approfondito in Moran et al., 2016) risiede proprio nella misura prediletta dai modelli seriali, in quanto i tempi di reazione non sono ritenuti unanimemente adeguati alla dimostrazione di due livelli di ricerca (tra gli altri, Palmer, 1995).

#### 4.2.1. Movimenti oculari nella visual search

Più prettamente incentrati sull'osservazione dei movimenti oculari sono gli studi, tra gli altri, di Rayner e colleghi. Le caratteristiche della scena, e in particolare la natura e la distribuzione dei distrattori in rapporto al target, influenzano anche la durata delle fissazioni, la lunghezza delle saccadi e l'ampiezza del campo di percezione (*perceptual span*). Già Rayner e Fisher (1987a,b), in un compito di ricerca su stringhe di lettere disposte in righe, avevano notato come sia il campo di percezione che quello di previsione variassero in funzione della 'difficoltà' delle lettere non-target (ovvero della loro similarità). In uno studio successivo, Bertera e Rayner (2000) confermano questo risultato in un esperimento con lettere e numeri in ordine sparso e mostrano che, oltre ai tempi di risposta, anche la durata delle fissazioni aumenta al variare del numero dei distrattori, segnalando una ricerca più difficile nelle scene più affollate. Allo stesso modo, Vlaskamp e Hooge (2006) osservano anche la riduzione della lunghezza delle saccadi in condizione di affollamento visivo e un notevole aumento dei tempi di ricerca. Inoltre, Findlay (2004) in una serie di esperimenti rileva una maggiore accuratezza nella direzione della saccade verso l'oggetto in presenza di più distrattori (15 invece di 2), corrispondente a maggiori tempi di latenza. Infine, uno dei risultati più rilevanti è la scoperta di un effetto 'gravità' (*center-of-gravity effect*), secondo cui la ricerca procederebbe non item per item, ma seguendo un modello 'a lente di ingrandimento' (*zoom-lens model*). Predisponendo sei oggetti a semicerchio, Zelinsky e colleghi (1997) hanno notato che i partecipanti fissano

prima il centro della scena, poi il centro del quadrante (destro o sinistro) dove è collocato il target, fino ad avvicinarsi progressivamente al target. Oltre alla necessità di integrare misure *online* di diversa natura (non solo tempi di risposta ma anche il tracciamento oculare), questo studio dimostra che è l'intera scena a essere presa in considerazione durante un compito di ricerca visiva; per di più, se le caratteristiche degli oggetti possono indubbiamente giocare un ruolo (cfr. supra), l'occhio si orienta sicuramente anche in base alla loro disposizione.



**Fig. 4.4.** La figura è tratta da Zelinsky et al (1997). Alla scena sono sovrapposte graficamente le fissazioni, rappresentate in cerchi bianchi dal diametro proporzionale alla loro durata.

## 5. Visual world paradigm

*Veronica D'Alesio*

### 5.1. Integrazione di audio e video

Con *Visual World Paradigm* si fa riferimento a un paradigma sperimentale che integra un input uditivo linguistico e stimoli visivi, che a loro volta possono essere extra-linguistici (una serie di immagini, una scena reale o una scena stilizzata) o linguistici (ad esempio, parole isolate). La paternità di questo design è certamente di Cooper (1974), che per primo fece ascoltare una storia a un gruppo di partecipanti mentre osservavano un riquadro con 9 immagini. L'autore scoprì che lo stimolo uditivo (ad esempio, la parola Africa, o il riferimento a un safari fotografico) spingeva l'occhio a soffermarsi su immagini ad esso collegato (una zebra o una macchina fotografica). Benché possa sembrare un risultato se non ovvio, quantomeno atteso, lo studio di Cooper dimostrò l'esistenza di un legame tra l'elaborazione dell'input acustico e lo spostamento dell'attenzione visiva e, per questo tramite, il disvelamento della rappresentazione mentale di uno stimolo linguistico. Torniamo qui a un concetto fondamentale già accennato nei capitoli dedicati alla lettura: ogni pratica sperimentale misura dei parametri e, per avere fondamento, deve essere assunto che questi siano indice di un processo cognitivo o di un suo funzionamento. Nel caso dell'eye tracker, cioè, l'occhio deve essere "legato" alla mente. Per la lettura, abbiamo visto che l'*eye-mind link*, così è chiamato in letteratura, è stato posto per la prima volta da Just e Carpenter (1980). Nel *visual world paradigm*, però, i movimenti oculari non registrano un carico cognitivo come nel *reading*, ma svelano una rappresentazione. In altre parole, questo paradigma si fonda sul legame tra linguaggio e mondo reale, inteso non solo come realtà oggettiva ma anche, e soprattutto, come rappresen-

tazioni mentali e conoscenze ad esso legate, che siano esse soggettive o condivise. La misura privilegiata, di conseguenza, non sarà più la durata delle fissazioni, ma quanto spesso sull'intero *trial* una data area viene fissata. Ci sono due misure che esprimono questa evenienza: la *fixation proportion*, ovvero la proporzione di partecipanti che fissano un dato *item*, e il numero di saccadi dirette verso la regione di interesse. Inoltre, essendo il movimento oculare interpretabile solo a partire dallo stimolo uditivo, queste misure vengono ricercate non solo in termini spaziali (quindi, in una data area), ma anche temporali, restringendo l'estrazione delle metriche a una finestra di tempo circoscritta (misurata in millisecondi dall'*onset* di una determinata parola o di un indizio fonologico rilevante per il compito). Questa doppia caratterizzazione è un punto cruciale dell'utilità del *visual world paradigm* in psicolinguistica poiché, a differenza della lettura, questo paradigma sperimentale permette di risalire allo sviluppo temporale del *processing*, permettendo di vedere in tempo reale gli stadi di accesso alla parola. Nonostante queste premesse, lo studio di Cooper (1974) passò inosservato per circa vent'anni, fino a quando la stessa tecnica non venne utilizzata da Tenenhaus e colleghi in un celebre articolo pubblicato su *Science* nel 1995. Tenenhaus et al. (1995) utilizzano il paradigma in un'indagine sull'elaborazione frasale, dando il via a un fiorente campo di ricerca oggi diventato uno dei pilastri della linguistica sperimentale. Con il diffondersi delle applicazioni del paradigma, anche la *linking hypothesis* si evolve e estende gli orizzonti della rappresentazione del linguaggio (per approfondimenti, Tanenhaus et al., 2000; Magnuson, 2019). Allopenna et al. (1998), il cui oggetto di ricerca è il riconoscimento di parola, pongono le basi della *linking hypothesis* che per prima mappa i movimenti oculari all'accesso lessicale in questo tipo di esperimenti. Come approfondiremo più avanti, gli studiosi tracciano lo sviluppo temporale dell'accesso lessicale guidato da input fonologico, mostrando che l'occhio fissa oggetti i cui referenti presentano sillabe identiche al nome udito. Quasi un decennio dopo, Altmann e Kamide (2007), i cui nomi sono indissolubilmente legati alla ricerca dell'anticipazione (cfr. infra), dimostrano come l'attenzione visiva sia guidata non solo da parole udite, ma anche da cose non dette o, per meglio dire, inferite. Prendendo ad esempio uno dei loro stimoli, all'ascoltare delle frasi 'l'uomo berrà la birra' o 'l'uomo ha bevuto il vino', i partecipanti tendono a dirigere lo sguardo verso un calice di vino vuoto o un boccale da birra pieno nel momento in cui iniziano a elaborare l'informazione

data dal tempo verbale; quindi, ben prima che né vino né birra siano nominati. Questi risultati portano gli autori a suggerire che l'attenzione non cambia a partire da quanto ci è presentato visivamente, ma da quanto è rappresentato mentalmente: in questo senso, la *mental world hypothesis* supera la semplice *visual world hypothesis* arrivando a postulare che l'elaborazione del linguaggio si fonda non solo sulla realtà oggettuale ma anche sulla nostra conoscenza del mondo e su come ne rappresentiamo gli eventi.

## 5.2. Studi principali

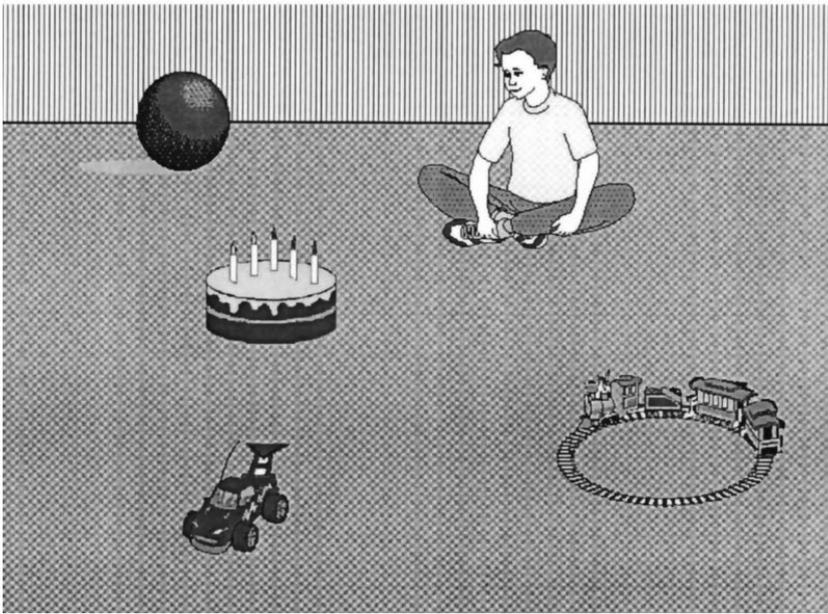
Gli studi di **comprensione** condotti con il *visual world paradigm* ci aiutano a capire cosa interviene nell'interpretazione di un messaggio linguistico. Tenenhaus et al. (1995), nello studio già citato, espongono una serie di esperimenti volti a dimostrare l'influenza del campo visivo nell'elaborazione linguistica. Descrivono come i partecipanti, ascoltando l'indicazione "*Touch the starred yellow square*", volgano gli occhi di volta in volta verso l'oggetto rispondente alla caratteristica appena nominata: dopo circa 250 ms al termine della parola *starred*, i partecipanti guardano l'unico oggetto a forma di stella, così come dopo *yellow* puntano all'unico oggetto giallo, nel caso ci fossero due referenti a forma di stella. Inoltre, contrariamente a quanto si professava all'epoca sull'autonomia della sintassi<sup>15</sup>, gli autori mostrano come anche qualcosa di esterno alla lingua, come determinati oggetti da manipolare, guidino la comprensione. Nell'esempio diventato ormai classico '*Put the apple on the towel in the box*', *on the towel* può essere interpretato sia come modificatore del nome ("la mela sull'asciugamano") che come destinazione del movimento ("sull'asciugamano che è nella scatola"). In un primo momento, vengono mostrati quattro oggetti: una mela su un asciugamano, un asciugamano, una matita e una scatola. In questa condizione, i partecipanti dirigono lo sguardo verso l'unica mela della scena (ovvero, quella sull'asciugamano) circa 500 ms dopo aver sentito la parola *apple*. Nella condizione critica, invece, il quadro cambia: posti davanti a due mele, una su un asciugamano e una isolata, i partecipanti

---

<sup>15</sup> Alla fine degli anni Ottanta, in particolare gli studi di Frazier suggerivano il primato della sintassi nell'elaborazione. A partire dai primi anni Novanta, invece, il contributo della ricerca con paradigmi *visual world* (rivisti in questo paragrafo) ha aperto la via a modelli integrati di elaborazione, in cui grande importanza riveste anche il piano lessicale.

esitano tra le due, finché l'asciugamano non viene nominato. L'esperimento dimostra che l'elemento ambiguo è interpretato una volta come destinazione, e un'altra come modificatore, a seconda esclusivamente di ciò che ci si presenta davanti nella realtà (vd. anche Trueswell et al. 1999; Snedeker and Trueswell, 2004). Più tardi, Chambers et al. (2004) osservano che non è solo la presenza di determinati oggetti a influenzare l'elaborazione cognitiva, ma sono anche le loro caratteristiche fisiche. Lo studio si inserisce nel filone della ricerca delle *affordances*, ovvero le proprietà fisiche di un oggetto che ne determinano la manipolazione (cfr. Gibson, 1977). Già in una serie di esperimenti (1998; 2002) Chambers e colleghi avevano osservato che la preposizione *inside* 'dentro' (contrapposta a *below*, 'sotto') guidava lo sguardo verso dei 'contenitori' e, per di più, quando veniva chiesto ai partecipanti di spostare oggetti più grandi, in presenza di più contenitori le fissazioni si dirigevano maggiormente su quelli grandi abbastanza. Il messaggio linguistico viene computato in tempo reale anche attraverso l'analisi di ciò che abbiamo a disposizione nel mondo reale. In un altro esperimento, Chambers et al. (2004) tornano a indagare la natura e la forma degli oggetti nella risoluzione dell'ambiguità sintattica. Anche qui, il segmento ambiguo poteva essere interpretato sia come modificatore del nome che come direzione finale dell'azione. Dati quattro referenti – un uovo allo stato liquido in una ciotola, un uovo allo stato liquido in un bicchiere, una ciotola vuota e una montagnella di farina- i partecipanti tendono a ritardare le fissazioni verso l'obiettivo finale (la ciotola vuota o la farina) quando ascoltano l'istruzione '*pour the egg in the bowl on the flour*'. Al contrario, le fissazioni verso il falso obiettivo (la ciotola vuota) aumentano in concomitanza della fine della parola '*bowl*' quando al partecipante vengono mostrati un uovo allo stato liquido in una ciotola e un uovo solido nel bicchiere (cioè, ancora nel guscio). Gli autori sostengono che in presenza di due referenti compatibili (in questo caso, con il verbo *pour*, 'versa'), i partecipanti esitano tra i due; nella condizione opposta, invece, *in the bowl* viene interpretato subito come destinazione, come dimostra l'incremento di fissazioni verso la ciotola vuota in concomitanza della fine del sostantivo. Se l'elaborazione linguistica fosse guidata in prevalenza dall'ordine sintattico, infatti, *in the bowl* verrebbe analizzato in prima istanza come direzione del movimento; tuttavia, cambiando il campo visivo, si notano maggiormente gli effetti dell'ambiguità. Secondo gli autori, questo avviene a causa della presenza di un referente non compatibile con il verbo

*pour* (l'uovo intatto) che, restringendo i possibili referenti a uno, fa prevalere l'interpretazione della ciotola come destinazione e non come possibile modificatore del referente. Strettamente legata alla nozione di *affordance* è la nozione di anticipazione, investigata soprattutto da Gerry Altmann e Yuki Kamide. In un loro articolo del 1999 (da cui è stata tratta l'immagine sotto), gli autori osservarono che determinati input linguistici portavano a dei movimenti anticipatori: l'occhio, cioè, si spostava verso referenti non ancora nominati, come se il partecipante sfruttasse indizi a vari livelli (contestuali, semantici...) per costruire delle deduzioni sulla continuazione del messaggio.



**Fig. 5.1.** Figura tratta da Altmann, Kamide (1999). Lo stimolo viene visualizzato in corrispondenza della frase stimolo "il ragazzo mangerà la torta" o "il ragazzo sposterà la torta".

In presenza di uno stimolo uditivo del tipo *the boy will move the cake* o *the boy will eat the cake* ('il bambino sposterà/mangerà la torta'), i partecipanti tendono a guardare l'oggetto target (la torta) prima che questo venga nominato nel caso in cui il verbo sia sufficiente a restringere il campo dei possibili referenti. Gli autori hanno condotto due esperimenti con gli stessi stimoli, uno con un *task* esplicito (ovvero, veniva chiesto ai partecipanti di valutare se la frase udita fosse compatibile con la

scena rappresentata) e uno senza *task* (il cosiddetto protocollo *look and listen*): nel primo esperimento, la prima saccade verso il referente inizia 127 ms dopo l'attacco del nome nella condizione generica con verbi del tipo *move*, mentre con i verbi restrittivi del tipo *eat* la saccade verso la torta viene lanciata 85 ms prima dell'attacco del nome. Questi movimenti anticipatori vengono registrati anche nel secondo esperimento, portando gli autori a sostenere che siano quindi guidati dall'estrazione di informazioni dal verbo. Negli anni seguenti, l'anticipazione è stata ampiamente indagata, mostrandosi efficace non solo in relazione alla semantica del verbo, ma anche nello studio delle inferenze nella conoscenza degli eventi (tra gli altri, Chambers, San Juan, 2008) e dei ruoli tematici (Kamide et al. 2003; Kukona et al. 2011; Sauppe, 2016). La scelta del referente può essere mediata anche dalla prosodia, come mostrano diversi studi (Dahan et al., 2002; Snedeker e Trueswell, 2003; Dahan e Tenenhaus, 2004; Weber et al., 2006; Watson et al., 2008; Hirose e Mazuka, 2015, 2017; Hirose 2020). Weber e colleghi (2006) esplorano gli effetti del focus contrastivo sulla scelta del referente. L'inquadramento teorico è quello della teoria autosegmentale e metrica nella fonologia dell'intonazione (Goldsmith, 1990; Pierrehumbert, Hirschberg, 1990; Ladd, 1996; Jun, 2005), secondo cui ogni lingua è caratterizzata da un particolare profilo prosodico organizzato in toni. La disposizione di questi toni (che si collocano a livello frasale) distingue tipi di frasi (affermative, interrogative parziali o totali...) e tipi di prominenze (focus ampio e focus contrastivo...). Gli autori analizzano i movimenti oculari in relazione al focus restrittivo prosodicamente marcato nello stimolo uditivo.

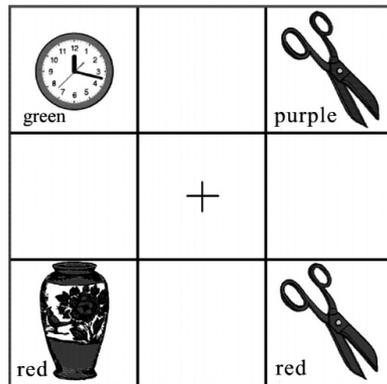


Fig. 5.2. La figura è tratta da Weber et al. (2006) e mostra uno stimolo esemplificativo del loro esperimento.

Sullo schermo (Fig. 5.2) sono posizionati quattro oggetti, di cui un distrattore (qui l'orologio), il target (le forbici viola), il *competitor* (lo stesso oggetto del target ma di un altro colore) e un oggetto dello stesso colore del target. I partecipanti ascoltavano due istruzioni: la prima individuava il target (le forbici viola) e la seconda uno dei due *competitor* (il vaso rosso o le forbici rosse). Ricordiamo che in tedesco la testa del sintagma è a destra; quindi, le coppie di stimoli critici sono del tipo *Klicke jetzt die ROTE Schere an* (con accento frasale contrastivo sull'aggettivo) e *Klicke jetzt die rote VASE an* (con accento frasale sul nome). I risultati mostrano che l'accento sull'aggettivo conduce l'occhio maggiormente sulla forbice rossa che sul vaso rosso (qui ci si riferisce alle fissazioni in proporzione dopo gli 800ms dopo l'onset dell'aggettivo) proprio perché i partecipanti integrano l'informazione prosodica nell'individuazione del referente come parte di un contrasto, ovvero un'opposizione tra due oggetti. Il *visual world paradigm* è stato utilizzato anche per indagare il ruolo della prosodia nella risoluzione dell'ambiguità sintattica, similmente a quanto visto in precedenza. Snedeker e Trueswell (2003), in un complesso studio formato da tre esperimenti collegati l'un l'altro, raccolgono una serie di stimoli del tipo *Tap the frog with the flower* ('tocca la rana con il fiore'), dove 'il fiore' può essere interpretato sia come strumento che come modificatore dell'oggetto. Gli stimoli sono registrati all'interno di un'interazione in laboratorio tra un parlante e un ascoltatore, che segue le indicazioni del primo nel manipolare gli oggetti nella scena davanti alla quale sono posti entrambi. Anche in questo caso, la misura presa in considerazione è la proporzione di fissazioni in una finestra di tempo compresa tra i 200 e i 500 ms a partire dall'onset del complemento oggetto (a sua volta divisa in tre sottosezioni di 100 ms l'una). Nella condizione del sintagma preposizionale interpretabile come strumento, le fissazioni sono dirette verso entrambi i referenti plausibili (la rana con il fiore e la rana senza), mentre nella condizione del modificatore, le fissazioni si dirigono verso l'oggetto target: in entrambi i casi, i movimenti oculari indicano che la frase viene disambiguata tra i 400 e i 500 ms. Considerando il tempo di programmazione di una saccade (circa 200ms), gli autori sostengono, quindi, che la prosodia viene integrata nell'interpretazione del messaggio anche molto presto, ovvero circa 250 ms dopo l'inizio del nome che identifica il referente. Infine, uno dei rami più sviluppati negli studi di comprensione linguistica tramite *visual world paradigm* concerne il riconoscimento di parola. Lo studio di Al-

lopenna et al. (1998), volto a dimostrare la validità di un determinato tipo di modello di comprensione, è certamente uno degli studi di riferimento del visual world paradigm. Indagando l'attivazione di parole fonologicamente vicine, gli autori mostrano come l'accesso alla parola sia in realtà continuo e sembra estendersi lungo un arco di tempo che va oltre il riconoscimento iniziale. I partecipanti dovevano ascoltare un'istruzione contenente la parola target (nel caso di questo esempio, tratto dagli autori, *beaker*). Posti davanti a quattro immagini (*speaker*, 'cassa'; *beaker*, 'becher'; *beetle*, 'coleottero'; *carriage*, 'carriola'), all'onset della parola target, i soggetti tendevano a fissare i due competitor con lo stesso attacco (*beaker* e *beetle*), per poi concentrarsi maggiormente sul *becher* e, sul finire della parola, rivolgere le fissazioni non solo al target ma anche al competitor con cui rima (*speaker*). Questi risultati supportano un modello di riconoscimento continuo proprio perché dimostrano che l'attivazione nel lessico mentale segue lo sviluppo temporale dell'input acustico: anche per questo ultimo aspetto, lo studio è stato preso come base di una prima *linking hypothesis* (crf. supra) che, seppur in maniera non dettagliata, proponeva la sincronia tra i movimenti oculari e la progressione in fasi dell'accesso lessicale. Negli anni seguenti, diversi studi hanno approfondito il riconoscimento di parola sotto vari aspetti, come ad esempio la coarticolazione (Brouwer, 2010) o l'accento di parola (Reinisich et al., 2010). Inoltre, questo filone si è molto sviluppato in riferimento al lessico bilingue (crf. capitolo 6).

Esattamente come per la comprensione, è stato notato come i movimenti oculari siano strettamente legati all'elaborazione del linguaggio anche in compiti di **produzione** (Meyer et al., 1998, Griffin e Bock, 2000; Bock et al., 2003; Meyer et al., 2004; Brown-Schmidt e Tenenhaus, 2006; Gleitman et al., 2007). Indubbiamente, questo accade perché gli occhi riflettono direttamente gli spostamenti dell'attenzione visiva e, spostandosi sull'oggetto, aiutano l'identificazione e il recupero della parola ad esso associato (tra gli altri, crf. Griffin, 2004). In uno dei primi studi, sia cronologicamente che per importanza, Meyer et al. (1998) analizzano il rapporto tra movimenti oculari e pianificazione del parlato tramite un compito di denominazioni di immagini (*picture naming*). Già da questo primo studio, si nota che il parlante fissa gli oggetti nell'ordine in cui li nomina, e che gli oggetti corrispondenti a nomi ad alta frequenza vengono riconosciuti prima e fissati meno a lungo rispetto ai corrispettivi dei nomi a bassa frequenza. Questi risul-

tati vengono confermati anche da Griffin e Block (2000), il cui studio mostra che l'ordine di menzione seguito dagli occhi persiste alla rotazione delle immagini o al cambio di struttura frasale ed è, perciò, vincolato al livello di elaborazione degli eventi piuttosto che alla salienza visiva.

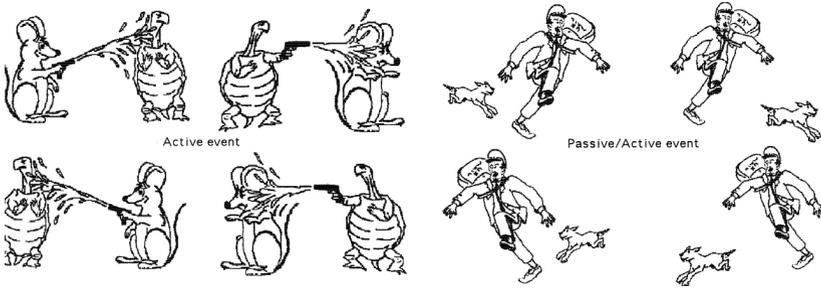


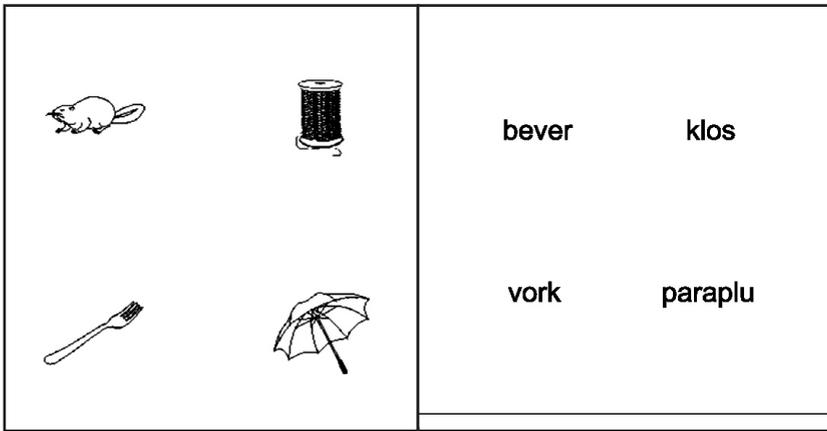
Fig. 5.3. La figura è tratta da Griffin, Bock (2000). Sulla sinistra, le immagini corrispondono sempre a frasi attive, mentre sulla destra le frasi attive corrispondono alle azioni in cui l'agente è l'uomo (viceversa, nelle frasi passive l'uomo è il paziente).

Quattro gruppi di partecipanti vengono posti di fronte a scene (riportate nella figura sopra, tratta da Griffin e Block, 2000) che rappresentano un'azione transitiva, espressa in forma attiva (un topolino che spruzza dell'acqua a una tartaruga, o viceversa) o potenzialmente passiva (il postino rincorre il cane e poi il postino che viene a sua volta rincorso dal cane). Nel gruppo di partecipanti a cui viene chiesto di descrivere la scena in maniera estemporanea, ovvero contemporaneamente alla prima visione dell'immagine, si osserva che l'occhio non segue l'ordine della concatenazione causale degli eventi, ma fissa l'elemento prima di nominarlo (anche qui, in ordine di menzione) indipendentemente da se questo sia agente o paziente, ma partendo dal soggetto della frase. Ciò non vuol dire che il primo elemento fissato diventa il soggetto della frase (infatti, i partecipanti fissano come primo elemento indifferentemente l'agente, il paziente o l'azione senza nessuna tendenza specifica), ma che inizialmente si notano delle fissazioni più corte verso l'oggetto e l'azione rappresentata, per poi concentrarsi più lungamente sul primo elemento della frase che si sta per pronunciare. Tra l'ultima fissazione su un elemento e la denominazione del suo referente lessicale vi è uno scarto di quasi un secondo, periodo che gli autori definiscono come *gaze-voice span*, ovvero il tempo che intercorre tra lo sguardo e la produzione del parlato.

### 5.3. Elementi di base dell'esperimento in visual world

Come abbiamo visto da questa breve rassegna, il *visual world paradigm* prevede l'allestimento di un set di oggetti e un partecipante che ascolta delle istruzioni preregistrate. Generalmente, ai partecipanti viene chiesto o di guardare la scena mentre ascolta la registrazione ("*look and listen*" *paradigm*, come in Altmann e Kamide, 1999) o di manipolare gli oggetti eseguendo l'azione descritta nelle istruzioni (come in Tenenhaus et al., 1995). Negli esperimenti a schermo, gli oggetti sono rappresentazioni più o meno stilizzate (dalle foto, agli sketch, ai disegni) disposti in una griglia. A differenza della scena (cfr. § 4.1), la griglia di oggetti isolati non permette di ricostruire un *gist*, perché non c'è un'idea di fondo che li lega: il partecipante, quindi, è portato a fare ricorso ad associazioni di tipo semantico, fonologico e visivo piuttosto che a interpretazioni sorte dalla conoscenza di un ambiente coerente. A parte il target, gli altri oggetti possono essere distrattori (senza alcuna connessione con il target) o *competitor*, letteralmente oggi che 'comettono' con il target per l'attenzione del soggetto in virtù di una caratteristica comune al target, come una somiglianza fonologica, visiva o una vicinanza semantica (vedi fig. 5.2). Dal punto di vista visivo, si tende a preparare gli stimoli di modo che ci siano meno interferenze possibili: quindi, se lo studio riguarda il linguaggio e non altri piani, l'impatto visivo dovrebbe essere ugualmente distribuito tra gli oggetti. È meglio evitare, per esempio, delle foto in contrapposizione con dei disegni, e in generale porre figure tutte di uno stesso stile e con lo stesso sfondo, avendo cura di ruotarle tra di loro per evitare bias spaziali (Godfroid, 2019). Inoltre, è opportuno bilanciare o controllare le caratteristiche oggettuali che potrebbero guidare l'attenzione al di là dell'input linguistico (cfr. salienza e *guiding features* nella percezione della scena e nella visual search, nel capitolo 4). Trattandosi di un'indagine linguistica, poi, è fondamentale che ci sia consenso sui referenti degli oggetti, per cui immagini ambigue o condizionate culturalmente sarebbero da evitare. Per questo, preliminarmente all'esperimento vero e proprio, le immagini dovrebbero essere sottoposte a dei partecipanti esterni per dei compiti di denominazione, in modo che ci sia accordo sul referente di ciascuna immagine. Uno strumento di riferimento per l'inglese americano è sicuramente il database di Snodgrass e Vanderwart (1980), che è stato poi esteso e adattato in altre sei lingue (tedesco, spagnolo messicano, italiano, bulgaro, ungherese e mandarino) nell'Internatio-

nal Picture Naming Project (IPNP). Infine, una risorsa per l'italiano è anche il database di Lotto, Dell'Aqua e Job (2001), comprendente 266 immagini controllate per diverse variabili (grado di tipicità categoriale del concetto; familiarità del concetto; tempo di produzione del nome; accordo sul nome; accordo sul concetto; lunghezza in lettere del nome della figura; lunghezza in sillabe del nome della figura; frequenza del nome scritto della figura; età di acquisizione del concetto). Oltre alle immagini, Huettig e McQueen (2007) hanno utilizzato una variante della presentazione classica, mostrando ai partecipanti le parole corrispondenti ai referenti degli oggetti (vedi fig. 5.4).



**Fig. 5.4.** La figura è tratta da Huettig, McQueen (2007). Facendo seguito allo stimolo uditivo contenente la parola *beker* "bicchiere", la griglia dell'esperimento 1 (sulla sinistra) mostra un competitor fonologico (*bever*, "castoro"), un competitor semantico (*vork*, "forchetta"), un oggetto con una forma simile (*klos*, "rocchetto") e, infine, un oggetto non collegato al target (*paraplu*, "ombrello"). Nell'esperimento 3, gli stessi stimoli sono rappresentati come parole grafiche (sulla destra).

In base alle domande di ricerca che ci si pone, un'ulteriore possibilità è quella del *blank-screen paradigm*. Altmann (2004) propone un esperimento *look-and-listen* molto simile ad Altman e Kamide (1999), ovvero con frasi contenenti verbi semanticamente restrittivi (e.g. *the man will read the cake*, l'uomo mangerà la torta; *the woman will read the newspaper*, la donna leggerà il giornale). Ai partecipanti viene mostrata una pseudo-scena (chiamata *visual scene*, ovvero una scena molto stilizzata in cui i quadranti con gli item sono immersi in uno sfondo comune, come già in Fig.5.1 di Altmann e Kamide, 1999) con quattro oggetti in quattro quadranti differenti (una donna, un uomo, una torta,

un giornale). L'immagine appare sullo schermo per 5 secondi e poi viene rimpiazzata da uno schermo completamente bianco, con l'input linguistico che parte dopo un secondo. Con questo metodo, lo studio non solo mostra che l'occhio si posa dove prima era presente l'oggetto nominato, ma conferma anche gli effetti anticipatori già osservati nello studio precedente, poiché i partecipanti tendono a guardare il paziente già durante l'ascolto del predicato. Questo paradigma suggerisce, quindi, che l'anticipazione non è dipendente dalla presenza di una scena, quanto piuttosto dall'acquisizione della sua rappresentazione mentale. Come abbiamo visto, un altro elemento essenziale negli esperimenti di *visual world* è l'ordine temporale di presentazione degli stimoli. Negli studi sui movimenti anticipatori, per esempio, mostrare l'immagine prima dell'input acustico (*preview*) può servire al partecipante per familiarizzare con gli oggetti, attivarne l'immagine mentale e localizzarli sullo schermo. Generalmente, comunque, la maggior parte degli studi di *visual world* include una *preview* della scena, in un range variabile tra 1 e 4-5 secondi (Huetting, 2011; Godfroid, 2019). Infine, molti esperimenti prevedono anche una *fixation cross*, ovvero una croce (generalmente posta prima della *preview*) verso la quale sono indirizzati i partecipanti prima dell'inizio del trial. Da un lato, avere un punto iniziale da cui partire uniforma il design sperimentale e permette di avere una base uguale per tutti i partecipanti per misurare direzione e durata delle saccadi iniziali (Godfroid, 2019); da un altro lato, la croce può indurre il partecipante a mantenere la fissazione su quel punto fino all'arrivo dell'istruzione, riducendo così la possibilità di movimenti esplorativi dell'occhio (Salverda, Tanenhaus, 2017).

Anche il file audio deve essere editato in modo che tutti gli item presentino la stessa suddivisione temporale al loro interno. In altre parole, la parte critica (o le parti critiche) dovrebbero avere tutte lo stesso *onset* e lo stesso *offset*. In questo modo, lo sperimentatore può analizzare i dati estratti dalla finestra temporale di interesse per la propria domanda di ricerca (corrispondente, ad esempio, ai millisecondi entro cui si colloca la parola target o alla parola ad essa precedente, come nel caso delle indagini sui movimenti anticipatori).

Per quanto riguarda i partecipanti, se tutti gli studi qui rivisti si basano su trial individuali, non è sempre questo il caso. Keysar et al. (2000), ad esempio, esaminando il ruolo delle conoscenze condivise nella comprensione, costruiscono il loro disegno sperimentale attorno a due compiti di collaborazione (un *referential communication game*, in

cui due persone comunicano con l'obiettivo della realizzazione di un compito comune per cui è necessario che ci sia un accordo sui referenti degli oggetti da manipolare, ossia sui loro nomi). Posti davanti a una griglia con degli oggetti, una persona (*director*) dà istruzioni a un compagno (*addressee*) perché questi sposti gli oggetti seguendo una combinazione precisa data dall'esaminatore. Tuttavia, i due partecipanti hanno davanti due versioni diverse della griglia: nel primo esperimento l'*addressee* non ne è cosciente, mentre nel secondo sa il *director* non vede i suoi stessi oggetti, e le due condizioni sono pensate per fare luce sull'adozione o meno di una prospettiva egocentrica da parte dell'*addressee* (cfr. anche Brown-Schmitt et al., 2008). Anche Snedeker e Trueswell (2003) usano un paradigma simile per analizzare il ruolo della prosodia nella disambiguazione sintattica e, nello specifico, osservare i movimenti oculari dell'ascoltatore mentre elabora in tempo reale gli indizi prosodici del parlante per individuare gli oggetti corretti da spostare.

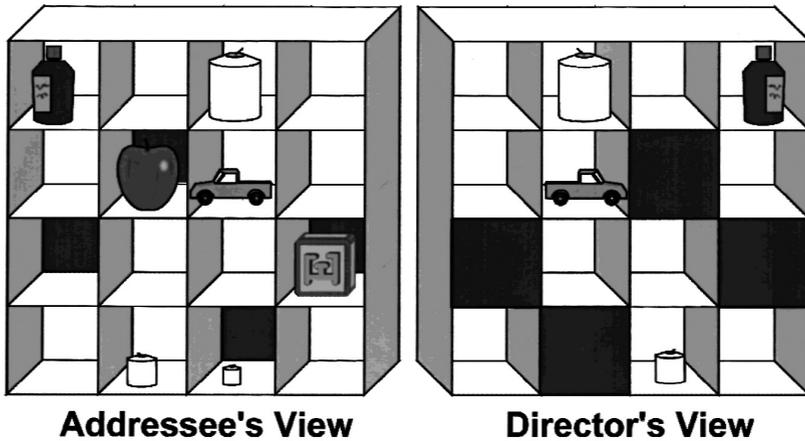


Fig. 5.5. Figura tratta da Keysar et al., (2000). La stessa griglia, vista da due lati diversi, mostra come "candela piccola" la candela visibile da entrambi i lati nell'ultima riga per il *Director*, mentre per l'*Addressee* la candela piccola è quella nascosta alla vista del *Director*.



## 6. L'eye tracking nella Second language research

*Maria Roccaforte*

### 6.1. Eye tracking e meccanismi cognitivi dell'apprendimento linguistico

L'eye tracking è diventato, nel corso degli ultimi venti anni, uno strumento ampiamente utilizzato nella ricerca sulle lingue seconde. Gli studi condotti su apprendenti di una lingua non materna, infatti, sono stati cruciali per sostenere l'ipotesi che un carico cognitivo più pesante comporti fissazioni più lunghe e regressioni più frequenti nei compiti di lettura (Ushioda 2016, Norton 2021). Già a partire dagli anni Duemila infatti ricercatrici e ricercatori di tutto il mondo hanno sfruttato l'eye tracking per misurare empiricamente l'attenzione distribuita in contesto di acquisizione e apprendimento, partendo dal presupposto che esista una stretta relazione tra ciò che viene fissato, come viene fissato e ciò che viene elaborato nella mente (Lai et al., 2013; Rahal & Fiedler 2019; Rayner et al. 2006; Rayner 2009, Eckstein et al. 2017; Godfroid et al. 2013; Issa & Morgan-Short 2019; Pellicer-Sánchez et al. 2020; Wang et al. 2006).

Nel campo della *Second Language Research* (SLR) troviamo poi una serie di studi volti a supportare lo sviluppo delle strategie e dei materiali didattici più efficaci per l'apprendimento della lingua (Remizantseva & Ablyazov 2019; Aldosari & Alrefai 2020) o a esplorare le percezioni di insegnanti di English as Foreign Language (EFL) (Simpson-Vlach & Ellis 2010) e di popolazioni di studenti che presentano una peculiarità dal punto di vista linguistico e talvolta neurotipicità ad esso correlate.

Tuttavia, numerose e di varia natura sono le perplessità che diversi studiosi hanno avanzato rispetto al collegamento tra le metriche dello sguardo raccolte con l'eye tracker e i processi cognitivi in atto nell'i-

stante in cui vengono raccolti. Ad esempio, secondo King et al. (2019, p. 6) le misure di eye tracking sarebbero “semplicemente misure del comportamento visivo capaci di dimostrare solo dove e per quanto tempo l’occhio ha guardato, ma non indicherebbero, se non accompagnate da altri protocolli di indagine, quali sono i processi cognitivi, attentivi, mnestici coinvolti né come avvenga il processo di comprensione o apprendimento” (Holmqvist et al., 2011). In altre parole, secondo questi studiosi, non è possibile utilizzare la registrazione del comportamento dello sguardo per individuare, ad esempio, il momento esatto in cui il significato di una parola diventa accessibile nella comprensione di una frase o di un discorso (per una rassegna si veda Conklin et al. 2018). Durante la lettura, infatti, la durata delle fissazioni sugli elementi di un testo varia in base alla difficoltà linguistica percepita: i lettori tendono a soffermarsi più a lungo su parole sconosciute o difficili e infatti fissazioni prolungate sono diventate nel tempo l’indicatore di un maggiore sforzo cognitivo.

A partire da questo principio, Issa e Morgan-Short (2019) hanno utilizzato l’eye tracking per studiare la distribuzione dell’attenzione durante la lettura di frasi complesse, rilevando che i partecipanti dedicano più tempo alle costruzioni sintattiche che presentano ambiguità grammaticali. In modo simile Rahal & Fiedler (2019) hanno dimostrato che gli apprendenti di una L2 tendono a soffermarsi più a lungo su frasi che richiedono una maggiore elaborazione cognitiva, contenenti parole meno familiari o strutture grammaticali ambigue. Se la durata delle fissazioni può essere utilizzata come indicatore del carico cognitivo, ovvero della quantità di sforzo mentale richiesto per completare un compito, anche un altro comportamento oculare conduce nella stessa direzione: l’aumento del diametro della pupilla. Eckstein et al. (2017) hanno infatti dimostrato che il carico cognitivo può essere misurato anche analizzando il diametro della pupilla e Pellicer-Sánchez et al. (2020) hanno aggiunto che il diametro della pupilla e la durata delle fissazioni aumentano in presenza di testi multimodali letti da apprendenti di una lingua.

Inoltre, uno studio condotto da Keating & Jegerski (2015) ha esaminato come gli apprendenti di L2 risolvono ambiguità grammaticali durante la lettura e dove fissano più a lungo nel tentativo di risolvere strutture ambigue.

L’eye tracking è stato utilizzato per studiare l’apprendimento grammaticale attraverso diverse modalità di input. Uno studio di Montero-

Perez et al. (2015) ha analizzato come l'eye tracking può essere impiegato per monitorare l'apprendimento di regole grammaticali durante compiti di lettura e ascolto simultanei, dimostrando che gli apprendenti prestano maggiore attenzione alle strutture grammaticali complesse quando sono supportate da input visivi.

Per fornire un quadro quanto più possibile aggiornato e completo sul panorama degli avanzamenti della ricerca in ambito SLR, riportiamo pertanto in questo ultimo capitolo le tendenze e i parametri di ricerca di 76 studi pubblicati sulle più importanti riviste scientifiche internazionali di settore, al fine di esplorare le aree di applicazione dell'eye tracking nella ricerca glottodidattica sperimentale e dare conto di quali siano le linee di indagine ricorrenti nel contesto della ricerca che integra la SLR con la tecnica del tracciamento oculare.

## 6.2. Aree di ricerca e stimoli nella L2 con eye tracker

Dal vaglio delle ricerche condotte con eye tracker nell'ambito di studi sulla L2 comparsi negli ultimi dieci anni sulle più importanti riviste scientifiche del settore, è stato possibile individuare le caratteristiche ricorrenti dei disegni sperimentali condotti con l'ausilio dell'eye tracker. Il presente paragrafo, pertanto, rintraccia le tre principali aree di ricerca e le modalità di presentazione dell'input individuate nel corso dello spoglio dei suddetti articoli scientifici.

1. Il primo, e più importante in termini quantitativi, filone di ricerca indaga la comprensione di parole e frasi da parte di lettori non nativi. In questo contesto, la dimensione lessicale è senz'altro l'area di ricerca più studiata. Questo andamento può essere attribuito a due fattori principali. In primo luogo, i disegni sperimentali nella ricerca sull'elaborazione delle parole e delle frasi sono vasti ed esistono già numerosi paradigmi sperimentali e parametri di riferimento, pertanto molti studi si configurano come *replication studies*. In questo genere di studi di norma vengono presentate ai partecipanti singole parole o brevi frasi e viene chiesto loro di eseguire compiti per valutare la comprensione delle parole, come l'associazione con immagini o definizioni, questionari somministrati pre e post test o risposte a voce alta durante il test. Il tracciamento oculare in questo caso viene utilizzato per esaminare i pattern di fissazione durante la lettura delle parole e delle frasi e per confrontare quei pattern alla luce delle risposte corrette o errate fornite o rispetto ai risultati contenuti da un gruppo di controllo.

2. La seconda corrente include lavori che indagano i pattern di movimento oculare nell'individuazione/risoluzione di anomalie o ambiguità a vari livelli linguistici. In questa condizione, i partecipanti sono esposti a frasi con parole mancanti, sostituzioni, errori o disposizioni di testo in modo non canonico. Il tracciamento oculare viene utilizzato per esaminare il processo di ricostruzione e il grado di sensibilità dei partecipanti non nativi rispetto a queste deviazioni dalla norma.

Nel caso in cui il focus sia sulla prosodia, per esempio, vengono presentati discorsi o frasi con variazioni nell'intonazione, nel ritmo e nella velocità per esaminare come lo sguardo dei partecipanti reagisca a tale condizione. Il tracciamento oculare in questo contesto permette di osservare quali strategie di lettura vengano adottate per affrontare queste irregolarità. Nel caso in cui sia sulla sintassi i partecipanti vengono esposti a frasi che presentano strutture grammaticali insolite, omissioni o violazioni delle regole sintattiche. Ad esempio, si possono utilizzare frasi con inversioni soggetto-verbo, dislocazioni a sinistra o a destra, oppure concordanze errate tra soggetto e predicato. Inoltre, alcuni studi sulla sintassi prevedono l'inserimento di frasi ambigue, dove la struttura grammaticale consente più interpretazioni.

L'obiettivo è capire come i lettori non nativi risolvano tali ambiguità e se il loro comportamento e le loro interpretazioni differiscano da quelle dei parlanti nativi. Ad esempio, vengono presentate frasi che lasciano spazio a due interpretazioni, monitorando i movimenti oculari per individuare eventuali difficoltà o ritardi nella comprensione. Nel caso in cui il focus sia sul vocabolario, si possono inserire parole polisemiche o frasi che contengono termini rari o di uso specialistico, per verificare come i partecipanti processino e ricostruiscono il significato basandosi sul contesto. Anche in questo caso, il tracciamento oculare offre informazioni sul tempo di fissazione e sulle regressioni verso i termini problematici, indicando eventuali difficoltà di comprensione o integrazione semantica.

3. Ricerche che indagano l'intenzione comunicativa o il significato implicito di un enunciato. Il tracciamento oculare, in questi casi, viene utilizzato per esaminare dove e come i partecipanti dirigono il loro sguardo durante la comprensione del messaggio e come integrano le informazioni contestuali per costruire un'interpretazione coerente. Si chiede di decifrare frasi con un doppio significato o con intenzioni comunicative non esplicite, come nei casi di sarcasmo, ironia o richie-

ste indirette (es. "Fa freddo qui dentro", che implica l'intenzione di chiudere una finestra). I movimenti oculari possono infatti rivelare le aree di interesse su cui i lettori non nativi si concentrano per elaborare il significato, come il volto di un interlocutore, dettagli rilevanti nell'ambiente o parti specifiche del testo. L'obiettivo finale di questo tipo di ricerca è comprendere meglio come le competenze pragmatiche si sviluppano nei parlanti non nativi e quali strategie possono essere utili per migliorare la loro sensibilità alle sfumature comunicative, contribuendo così a un'interpretazione più fluida e naturale nei contesti reali.

### **6.3. L'uso degli eye tracker nello studio dei manuali di lingua: stato dell'arte e prospettive future**

L'introduzione della tecnologia dell'eye tracking nello studio dell'apprendimento linguistico ha aperto nuove possibilità per analizzare il comportamento visivo e cognitivo degli studenti durante l'interazione con i materiali didattici, in particolare con i manuali di lingua.

Winke et al. (2013) hanno analizzato come gli studenti di lingue seconde interagiscono con esercizi di ascolto e lettura, dimostrando che gli eye tracker possono rilevare aree di confusione o difficoltà. Un altro studio di Schmidt-Weigand et al. (2010) ha esaminato l'efficacia delle immagini nei manuali, evidenziando che immagini ben progettate riducono il carico cognitivo e migliorano la comprensione.

Tuttavia, la maggior parte delle ricerche si concentra su aspetti isolati dei manuali, come la disposizione del testo e delle immagini (Chun, 2012). L'eye tracking potrebbe essere utilizzato per adattare i contenuti dei manuali alle esigenze specifiche degli studenti, fornendo feedback personalizzati in tempo reale. Inoltre, con l'aumento dei manuali digitali, l'eye tracking potrebbe essere integrato con piattaforme interattive per monitorare l'apprendimento e ottimizzare i contenuti e combinando l'eye tracking con altre tecnologie biometriche, come il riconoscimento facciale, si potrebbero analizzare le reazioni emotive degli studenti ai materiali. Infine, l'eye tracking potrebbe essere utilizzato per valutare come gli studenti di diverse culture interagiscono con i manuali, adattandoli alle differenze culturali.

## 6.4. Comprensione multimodale e linguaggio bilingue

L'eye tracking è stato utilizzato anche per studiare la comprensione multimodale e il linguaggio bilingue. Simpson-Vlach & Ellis (2010) hanno condotto uno studio in cui hanno confrontato i movimenti oculari di individui monolingui e bilingui durante la lettura di testi in più lingue. I risultati hanno rivelato che i bilingui tendono a navigare tra le lingue in modo più fluido, dimostrando una maggiore capacità di adattamento alle variazioni linguistiche.

Un altro studio significativo è quello di Pellicer-Sánchez et al. (2020) che ha esplorato come gli apprendenti di L2 elaborano materiali multimodali, come video e testi scritti. L'eye tracking ha rivelato che la comprensione multimodale richiede un maggiore carico cognitivo ma ha effetti positivi sulla memorizzazione e sui livelli attentivi e garantisce un apprendimento più duraturo.

Nonostante i numerosi vantaggi, l'eye tracking presenta vari limiti. Ad esempio, le misure del movimento oculare non possono essere direttamente collegate a specifici processi cognitivi, il che significa che i dati devono essere interpretati con cautela. Inoltre, fattori esterni, come la variazione della luce, la distanza dall'input, problemi di vista o emotività del partecipante o possono influenzare i comportamenti dell'occhio, le fissazioni e perfino la misura del diametro della pupilla, rendendo difficile isolare tali fenomeni dagli effetti del carico cognitivo. Questo richiede protocolli sperimentali rigorosi per garantire la validità dei risultati.

Il tracciamento oculare non è stato adottato in modo diffuso nello studio della scrittura in L2. Questo può essere dovuto al fatto che i ricercatori che si occupano di scrittura stanno ancora esplorando l'applicabilità dell'eye tracking nella ricerca e solo recentemente hanno iniziato a illustrare l'utilità dell'eye tracking in questo ambito (si veda, ad esempio, Ranalli et al., 2019).

## 6.5. Conclusioni e direzioni future

L'eye tracking ha permesso di ottenere importanti progressi nella comprensione dei processi cognitivi che sottendono l'apprendimento delle lingue seconde. Tuttavia, vi sono ancora molte aree inesplorate, come l'uso nel campo della valutazione e nella personalizzazione di strumenti didattici rispetto ai bisogni degli apprendenti. Inoltre,

l'integrazione dell'eye tracking con altre tecnologie educative, come la realtà aumentata e la realtà virtuale, offre opportunità uniche per l'apprendimento delle lingue. Queste tecnologie possono creare ambienti di apprendimento immersivi che stimolano l'interazione e la partecipazione attiva degli studenti, migliorando ulteriormente l'efficacia dell'insegnamento. L'integrazione delle metriche di eye tracking nella valutazione linguistica può portare allo sviluppo di strumenti di valutazione più completi e accurati, in grado di cogliere meglio la natura multiforme della competenza linguistica.

Dal quadro generale è emerso che l'eye tracking ha avuto otto principali aree di applicazione nella ricerca: grammatica, vocabolario, lettura, ascolto, scrittura, validità, conversazione e fonologia e che nelle ricerche condotte sulla L2 sono state usate 12 tipi di misure, temporali e di conteggio delle fissazioni e inerenti alla fissazione.

In secondo luogo, i risultati hanno mostrato che l'eye tracking è stato utilizzato dai ricercatori di L2 per trarre inferenze sull'attivazione e sul mantenimento dei processi attentivi, sull'elaborazione cognitiva e sul carico cognitivo: tra queste l'attenzione è la dimensione più indagata. Gli studi sull'elaborazione cognitiva hanno privilegiato le misure di conteggio della fissazione, la scelta delle metriche usate per indagare il carico cognitivo sembrano invece dipendere principalmente dal tipo di compito assegnato.

In generale, rimandando per ulteriori approfondimenti allo studio di Hu & Aryadoust 2024 ad oggi gli studi condotti con eye tracking in L2 nel campo della SLR non sono ancora in grado di garantire livelli di trasparenza e reduplicabilità, mostrando un alto livello di variabilità e di negli aspetti metodologici riguardanti l'apparecchiatura, nella scelta dei software di analisi, e nelle procedure.



## 7. La metodologia eye tracking nella ricerca linguistica in italiano

*Veronica D'Alesio*

L'eye tracking è stato ampiamente utilizzato, soprattutto negli ultimi venti anni, anche nello studio dell'elaborazione e dell'acquisizione dell'italiano. Nell'ambito della L1, sono stati condotti studi sui diversi livelli linguistici sia su popolazioni adulte che su bambini in ottica evolutiva. Allo stesso modo, gli studi sull'italiano come lingua seconda riguardano sia bambini bilingui che apprendenti adulti a vari livelli di proficiency. Presentiamo in questo capitolo una rassegna narrativa di studi scelti per rappresentare i vari ambiti di applicazione di questa metodologia nello studio dell'italiano come lingua materna e come lingua seconda. Come si noterà, gli argomenti trattati coprono varie branche della linguistica, e le domande di ricerca indagate spaziano notevolmente, per cui la divisione in ambiti qui proposta risponde a una mera esigenza di ordine narrativo e non sempre riesce a rappresentare esaustivamente quei fenomeni che spesso si trovano all'interfaccia di più livelli linguistici. Questo capitolo vuole essere un'introduzione al panorama degli studi condotti in italiano su diverse popolazioni di parlanti, per cui si rimanda ai singoli articoli citati per approfondimenti specifici che esulano dagli scopi di questa rassegna.

### 7.1. Studi sul lessico

Tra i temi affrontati attraverso la presentazione di stimoli in italiano, troviamo l'**accesso lessicale** e la scomposizione morfologica. Marelli e Luzzati (2012) indagano l'accesso dei composti in un paradigma di lettura di frasi, analizzando la trasparenza semantica e la frequenza dei costituenti (e del composto stesso). La trasparenza semantica vie-

ne misurata, tramite giudizi di parlanti nativi, sia in riferimento alla trasparenza dei costituenti (ovvero quanto i costituenti siano vicini al significato del composto) che alla trasparenza del composto (ovvero quanto il significato del composto risulti dalla somma dei significati dei suoi costituenti). Gli autori riportano un effetto di trasparenza semantica e un'interazione tra trasparenza semantica e frequenza, sia del primo costituente che del composto, visibili nella durata della prima fissazione, dimostrando come il primissimo accesso lessicale sia già modulato sia dalle proprietà della parola che dalla sua analisi semantica. Risultati analoghi sono riportati da Amenta e colleghi (2015) sui derivati, in un compito di lettura di frasi. Comparando derivati trasparenti ("copertina", ovvero una piccola coperta) e opachi ("copertina" di un libro), gli autori osservano parimenti un'influenza precoce della componente semantica, riscontrabile un effetto di frequenza della radice lessicale in termini sia opachi che trasparenti. In particolare, nei derivati trasparenti (dove il significato della radice è contiguo al significato del derivato) la frequenza della radice ha un effetto facilitatorio sulla durata della prima fissazione, mentre nei derivati opachi (dove il significato della pseudo-radice non è legato al significato della parola) accade il contrario, ovvero minore la frequenza, maggiore sarà la durata della prima fissazione. A questo si aggiunge anche un effetto della frequenza dell'intera parola sia sulla prima fissazione che sulla *gaze duration*. Espandendo i risultati sui composti nella lettura di frasi, Arcara e colleghi (2014) osservano anche una maggiore durata dei tempi di lettura totali sugli pseudo-composti in cui la prima porzione coincide con una parola vera dell'italiano (ad esempio, "coccodrillo"). Gli autori spiegano questo risultato come una rianalisi dello stimolo che inizialmente viene interpretato come un composto trovandosi come primo segmento una potenziale parola. Anche i composti a testa finale risultano più difficili da elaborare rispetto a quelli a testa iniziale (un risultato apparentemente in contrasto con lo studio di Luzzati e Marelli visto sopra), ulteriore prova della scomposizione che avviene nell'analisi di parole complesse. Anche qui l'impatto sul *total reading time* viene motivato con una possibile rianalisi nel momento in cui la composizione del nome non risulta essere quella canonica, ovvero a testa iniziale come l'ordine sintattico degli elementi in italiano. Infine, ulteriore supporto alla scomposizione morfologica delle parole in lettura è dato dallo studio di Solaja e Crepaldi (2024), che osservano effetti facilitatori per la lettura e l'apprendimento di non-parole con

suffissi reali dell'italiano. L'effetto, visibile sia nella *gaze duration* che nel *total reading time*, è accompagnato da tempi di lettura totali maggiori su parole in cui il significato del suffisso è incongruente con quanto espresso dal senso della frase, evidenziando un ruolo fondamentale del contesto nelle successive rianalisi e integrazioni semantiche denotate dalle *late measures*.

## 7.2. Studi di stampo pragmatico

Un altro consistente filone di ricerca è quello riguardante gli **aspetti semantici e pragmatici** del linguaggio, come lo studio di implicature o significati impliciti. In due esperimenti utilizzando il *visual world paradigm*, Foppolo e colleghi (2015) esaminano l'ambiguità semantica in frasi del tipo "Il grande X è l'unico che sta facendo P", in cui "l'unico" apre due possibili interpretazioni: la prima, detta anaforica, legge "unico" come appartenente alla stessa categoria alla quale appartiene X; la seconda, detta esoforica, allarga l'interpretazione di "unico" a tutte le categorie di entità che in quel momento stanno facendo l'azione P. In termini di informatività, una frase affermativa così fatta è vera in un numero più ampio di situazioni se l'interpretazione è anaforica, mentre per essere vera secondo una lettura esoforica è necessario verificarla in tutti i partecipanti nella situazione, non solo in quelli appartenenti alla stessa categoria dell'elemento X. Quest'ultima è quindi un'interpretazione forte, mentre quella anaforica è debole perché, appunto, necessita del coinvolgimento di un gruppo minore di elementi. Partendo da questo modello, gli stimoli vengono progettati per comprendere quattro diverse categorie di oggetti (in Fig. 7.1., pentagoni, triangoli, cerchi e quadrati) e ulteriori elementi Y (in Fig. 7.1. le emoticon) che fanno l'azione P: in questo modo, la frase affermativa di sopra è vera secondo l'interpretazione anaforica, ma falsa secondo l'interpretazione esoforica. Da questo primo esperimento emerge che l'interpretazione forte, ovvero quella esoforica, è quella preferita dai partecipanti (adulti), in base sia ai pattern di movimenti oculari nella ricerca di X sullo stimolo visivo, sia ai giudizi di verità su ogni frase. Risultati paragonabili si osservano anche in un secondo esperimento, dove le stesse frasi sono presentate con polarità inversa (ovvero negative), di modo che l'interpretazione esoforica sia stavolta quella debole in termini di informatività. Gli autori dimostrano, quindi, che

i loro partecipanti adottano sempre una strategia esoforica. I parlanti nativi adulti tendono a sfruttare massimamente il contesto nell'interpretazione dello stimolo: una volta raggiunta la parte critica "unico", e aver fissato il quadrante target, l'esplorazione si estende anche agli altri quadranti, probabilmente in cerca di altri possibili referenti per risolvere l'ambiguità semantica.

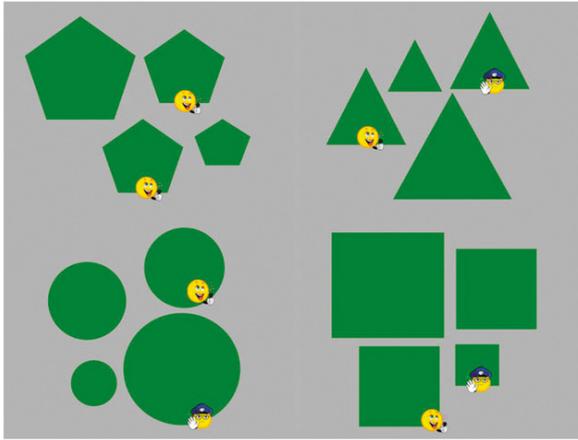


Fig. 7.1. Figura tratta da Foppolo et al. (2015).

Usando stimoli simili, Foppolo e Marelli (2017) indagano il processing delle implicature scalari, precisamente di "alcuni", il cui uso implica la negazione del più forte "tutti" (leggendo, cioè, "alcuni ma non tutti").

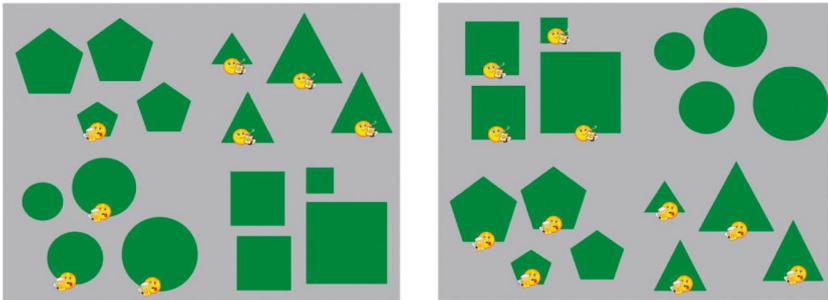


Fig. 7.2. Figura tratta da Foppolo et al. (2017).

Ai partecipanti viene fatta ascoltare una frase (15) in concomitanza dell'immagine di sinistra (in Fig. 7.2) e poi, dopo un intervallo con una

schermata grigia, una seconda frase in concomitanza con l'immagine di destra (7.2). La seconda è la frase critica, che presenta tre condizioni: una con "alcuni" (16) e due con "tutti", in cui solo l'ultima parola (denotante la classe di oggetti) ne rende una vera (17) e una falsa (18). La struttura a due frasi permette di mettere i quantificatori in focus nella fase critica con una struttura VOS, elicitando perciò la computazione dell'implicatura tramite l'attivazione di alternative. Un ulteriore controllo viene dato anche dai giudizi di verità chiesti ai partecipanti, per cui l'analisi viene poi ristretta ai soli partecipanti che interpretano "alcuni ma non tutti" nella frase (16). I risultati mostrano come parlanti madrelingua adulti dirigano lo sguardo al quadrante pertinente all'ascolto del quantificatore, dimostrando che, quindi, le implicature scalari attivate da "alcuni" non richiedono ulteriore tempo per essere processate rispetto a un quantificatore forte come "tutti".

1. Qui il pentagono piccolo sta bevendo una birra.
2. Adesso invece bevono una birra alcuni dei triangoli.
3. Adesso invece bevono una birra tutti i triangoli.
4. Adesso invece bevono una birra tutti i pentagoni.

Anche l'aspettualità del verbo è stata studiata in italiano con il paradigma *visual world* (Foppolo et al., 2021), dimostrando come movimenti anticipatori rivelino l'elaborazione della telicità in maniera incrementale. Comparando frasi con verbi puntuali e durativi (quindi verbi in cui la presenza della conclusione dell'evento è già lessicalmente definita, come "aprire", contro verbi in cui questo non accade, come "colorare") al passato prossimo, ai partecipanti vengono mostrate un'immagine di un oggetto rappresentato ad azione completata (ad esempio, una stella interamente colorata o un giornale aperto) e un'altra in cui lo stesso evento non ha ancora raggiunto il suo termine (quindi, rispettivamente, una stella colorata in parte e un giornale chiuso). In entrambi i casi, lo sguardo si dirige verso l'immagine target già durante l'ascolto del verbo, indicando che la telicità inizia a essere processata sfruttando la perfettività del verbo.

Con riferimento ad account teorici basati su conoscenza del mondo e interpretazioni pragmatiche, in italiano è stata indagata anche la risoluzione dell'anafora. Ronca e Moscati (2019), ad esempio, analizzano il rapporto tra genere grammaticale e stereotipi di genere in un *visual world paradigm* sulle inferenze associate ad alcuni nomi invariabili in

italiano (come “ingegnere”). I loro risultati dimostrano che, se da un lato sussistono delle inferenze basate su stereotipi di genere, queste vengono rapidamente corrette dall’intervento di informazione linguistica come, nel loro caso, le marche morfosintattiche di genere sul participio passato che segue il nome. Vender e colleghi (2025), invece, prendono in esame l’uso di ripresa pronominale o pronomi zero (ellissi) nella subordinata causale in relazione al ruolo dell’antecedente (oggetto o soggetto) nella principale. Dal loro studio sui tempi di lettura frasale, emerge un maggiore sforzo cognitivo in casi in cui il soggetto viene ripreso da un pronome, come evidenziato sia dall’aumento dei tempi di prima lettura (*first-pass*) che dalle misure associate a stadi di elaborazione successiva (*total reading time* e *regression path duration*). Il loro studio, quindi, mettendo in relazione il tipo di ripresa anaforica con la relazione (causale) che ne determina l’antecedente, sottolinea il ruolo dei fattori pragmatici e discorsivi nella risoluzione anaforica in italiano.

### 7.3. Studi in ambito sintattico

Come abbiamo visto nei capitoli precedenti, un’ulteriore linea di indagine con metodologia eye tracking è rappresentato da studi centrati sulla **sintassi**. In italiano, un interessante studio di lettura di lettura di frasi (Concetti e Moscati, 2023) osserva i tempi di lettura sul verbo della principale in frasi come:

5. L’allievo | che | critica | i pittori | durante l’esame | dipinge | un paesaggio.
6. L’allievo | che | i pittori | criticano | durante l’esame | dipinge | un paesaggio.

Dai risultati emerge un effetto facilitatorio sull’area di interesse “dipinge” nelle frasi contenenti una subordinata relativa oggetto (20). L’effetto si riscontra nelle *late measures* (*total reading time* e *regression path duration*), che sono significativamente più corte rispetto alla condizione con subordinata relativa soggetto. Come osservano gli autori, l’asimmetria soggetto-oggetto nelle frasi relative è stata riscontrata in diversi studi di processing ma l’analisi dell’elaborazione del verbo nella principale permette di considerare i vari meccanismi che interven-

gono nella lettura e comprensione di frasi. In questo senso, i tempi di lettura minori nella condizione relativa-oggetto potrebbero riflettere il diverso posizionamento della traccia lasciata nella frase relativa nelle due condizioni, rispettivamente prima (19) e dopo (20) il verbo della subordinata (“critica” / “criticano”). La differenza tra le due costruzioni, infatti, risiede nel fatto che nella condizione (20) la traccia è più vicina al verbo della principale, da cui si deduce un minore sforzo della memoria di lavoro nella riattivazione dei tratti necessari all’operazione di accordo tra il soggetto e il verbo della principale.

#### 7.4. Studi con popolazioni non tipiche

L’italiano, inoltre, è una lingua a ortografia trasparente, per cui generalmente si ha una corrispondenza tra grafemi e fonemi. Per questa sua caratteristica, diversi studi sulla lettura hanno indagato popolazioni con **dislessia**, in cui sono stati riscontrati un maggior numero di fissazioni e di saccadi, una maggiore durata delle fissazioni (Zoccolotti et al., 1999; Bonifacci et al., 2023), un effetto lunghezza prominente e più forte rispetto alla frequenza nei tempi di lettura della parola e una minore probabilità di saltare le parole funzionali (De Luca et al., 1999).

#### 7.5. Processing bilingue in età adulta

In studi sul processing bilingue e sull’apprendimento di una L2 in età adulta, infine, l’eye tracking è stato usato per indagare i processi di lettura di partecipanti sordi e udenti (Tomasuolo et al., 2019), la connessione tra sguardo e richiamo mnemonico nell’apprendimento con sottotitoli (Ragni, 2020) e l’elaborazione di collocazioni in apprendenti di italiano L2 a vari livelli di competenza (Fioravanti et al., 2023). Questo ultimo studio, in particolare, esamina collocazioni tipiche (i.e., non manipolate) e atipiche (i.e., con violazioni grammaticali o sostituzioni lessicali) in un compito di lettura di frasi. I tempi di lettura totali sull’intera collocazione considerata come unica area di interesse (e.g., *passare l’esame/attraversare l’esame*; *aprire gli occhi/aprire lo occhi*) rivelano come sia gli apprendenti che i parlanti nativi di italiano siano sensibili alle manipolazioni di unità multiparola. Inoltre, i pattern di lettura mostrano una differenza significativa tra apprendenti intermedi e avanzati, con questi ultimi più vicini ai parlanti nativi in termini di elaborazione delle collocazioni.

## 7.6. Studi sul processing in età evolutiva

In ottica **acquisizionale**, i movimenti oculari sono stati esaminati in uno studio di caso lungo gli anni della scuola elementare (De Luca et al., 2010), evidenziando una progressiva diminuzione del numero di fissazioni, rifissazioni e regressioni, un aumento di *skipping rate* e un avvicinamento negli anni tra la *gaze duration* e il *total reading time* sulla parola, a confermare il minor numero di riletture necessarie all'elaborazione. Traficante e colleghi (2018) confermano anche nei bambini una differenza tra l'elaborazione di nomi e verbi, già riscontrata negli adulti. In particolare, in un compito di lettura di frasi, la frequenza della base e la frequenza dell'intera parola hanno un effetto diverso in sostantivi derivati da nomi ("caminetto") o da verbi ("punizione"), mostrando come i bambini facciano uso delle componenti morfologiche durante l'elaborazione di parole complesse. In parole derivate da nomi, la durata della prima fissazione è minore quanto più alta è la frequenza della base, ma si allunga all'aumentare della frequenza dell'intera parola. Questo effetto in senso contrario viene spiegato dagli autori da una competizione tra l'attivazione del sostantivo di base e l'attivazione dell'intera parola composta, essendo entrambi candidati plausibili per occorrere in quella data posizione all'interno del contesto frase (sono entrambi nomi e sono anche semanticamente vicini). Lo stesso effetto, invece, non risulta nei sostantivi derivati da verbi, dove l'alta frequenza della base porta a una durata maggiore della prima fissazione, suggerendo un contrasto tra la parola attesa (un sostantivo) e la lettura della prima parte (ovvero, una base verbale). Un altro risultato interessante dello studio è che la competizione nell'attivazione di due nomi altamente frequenti viene risolta nella fase finale di accesso lessicale (*gaze duration*) solo nei lettori più esperti e indipendentemente dal grado di istruzione (che nel loro esperimento andava dalla terza alla quinta elementare), a denotare una maggiore influenza dell'apprendimento implicito rispetto a quello esplicito nella decodifica lessicale. In età prescolare, invece, due *visual world studies* hanno indagato l'elaborazione dei quantificatori negativi (Moscati 2023) e degli aggettivi (Redolfi e Melloni, 2025). Quest'ultimo studio, in particolare, conferma i differenti pattern di processing associati a diverse parti del discorso (in questo caso vengono analizzati nomi e aggettivi) ed estende le osservazioni di pattern differenziati anche tra classi di aggettivi

differenti, con quelli relativi (grande/piccolo) che risultano essere più difficoltosi rispetto a quelli assoluti (aperto/chiuso).

Infine, nel processing bilingue in età evolutiva sono stati esaminati i meccanismi anticipatori dati da indizi morfosintattici di genere e numero (Bosch et al., 2022; Bosch e Foppolo, 2022, 2024) in bambini multilingui e apprendenti di italiano in età scolare.

In conclusione, gli studi sul processing linguistico in età evolutiva evidenziano la complessità e la dinamicità dei meccanismi coinvolti nella lettura e comprensione del linguaggio. L'utilizzo dell'eye-tracking ha permesso di osservare con precisione lo sviluppo delle abilità di decodifica, mostrando come il passaggio dall'infanzia all'età scolare sia caratterizzato da una progressiva automazione dei processi lessicali. Le differenze riscontrate nell'elaborazione di nomi, verbi e aggettivi, nonché l'influenza della struttura morfologica delle parole, sottolineano il ruolo cruciale della competenza grammaticale implicita. Anche nei bambini bilingui emerge una precoce sensibilità agli indizi morfosintattici, suggerendo strategie di anticipazione linguistica già raffinate. Questi risultati contribuiscono non solo alla comprensione dei meccanismi cognitivi alla base del linguaggio, ma anche alla definizione di interventi educativi mirati, capaci di supportare l'acquisizione linguistica in modo efficace e individualizzato.



## Appendice 1. Principali hardware e software impiegati nella ricerca linguistica

PRODUTTORE & MODELLO	TECNOLOGIA	FORMA HARDWARE	SAMPLING RATE	PRECISIONE ACCURATEZZA	SOFTWARE ABBINATO
SR Research – EyeLink II	Infra-red	Indossabile (casco)	500 Hz (binoculare)	Accuratezza 0.5°, risoluzione 0.01° RMS, ritardo 3 ms	Experiment Builder, Data Viewer, (compatibile con Psychopy, Psycho-toolbox, E-prime, Opensesame, Matlab/Python e altri)
SR Research – EyeLink 1000 Plus	Infra-red	Screen-based (Desktop/Tower/Arm/a lungo raggio), compatibile con EEG, fMRI, MEG	fino a 2000 Hz (binoculare)	accuratezza fino a 0.15°–0.5°; risoluzione 0.01° RMS Ritardo < 3ms	
SR Research – EyeLink Portable Duo	Infra-red	Mobile (Screen-based), compatibile con EEG	fino a 2000 Hz stabilizzato, altrimenti fino a 1000 Hz	Accuratezza fino a 0.15°; risoluzione 0.01°	
Tobii – Pro Spark	Near Infra-red	Mobile (Screen-based), compatibile con EEG	60 Hz	Precisione 0.26° RMS, accuratezza 0.45°	Tobii Pro Lab
Tobii – Pro Fusion	Near Infra-red	Mobile (Screen-based), compatibile con EEG	Fino a 250 Hz	Precisione 0.04° RMS, accuratezza 0.3°	
Tobii – Pro Spectrum	Near Infra-red	Screen-based	Fino a 1200 Hz	Precisione 0.01° RMS, accuratezza 0.3°	
Tobii – Pro Glasses 3	Near Infra-red	Indossabile (occhiali)	50 o 100 Hz	Accuratezza 0.6°	Tobii Pro Lab, Controller (App)



## Appendice 2. Definizioni delle metriche di tracciamento rintracciate nel corso dello spoglio delle ricerche sulla Second language research

METRIC	SCALE	DEFINITION
<b>Fixation</b>		
Time to first fixation	Time	The time period from entering the AOI until the first fixation is made
First fixation duration	Time	The duration of the first fixation made in an AOI
Gaze duration	Time	The sum of all fixations on a word prior to an eye movement to another word
First pass reading time	Time	The sum of all fixations recorded for a multi-word interest area up to the point when the eyes leave the interest area
Second pass time	Time	The summed duration of all fixations that are made within an interest area when the eyes visit the area for the second time; This includes cases where the interest area was originally skipped
Rereading time	Time	The summed duration of all fixations in an interest area except for those fixations made during first pass
Average fixation duration	Time	Mean of fixation duration on each AOI
Total fixation duration	Time	The summed duration of all fixations in an AOI
Number of fixations	Count	The number of fixations made in an AOI
Proportion of fixations	Count	The proportion of total fixations that are directed to an AOI, or the number of fixations between AOIs and between experimental groups
Fixation position	Space	Location of a fixation

<b>Saccade</b>		
Saccade duration	Time	The amount of time that it takes to actually move the eyes
Saccade count	Count	The number of saccades counted within an AOI
Saccade length	Space	The distance between two consecutive fixations
<b>Dwell</b>		
First pass time	Time	Time spent for the first entering of an AOI until leaving
Rereading time	Time	Summed duration of revisited time spent within an AOI
Total reading time	Time	Total time spent within an AOI or spent for a reading task
Total visit duration	Time	The summed duration of all visits to a particular interest area
Total number of visits	Count	The total number of visits made to an AOI
Dwell rate	Count	The number of entries into a specific area of interest per minute
<b>Regression</b>		
Regression path duration/go-past time	Time	The time spent on the word itself and any prior parts of the sentence before the reader moves past the critical word to the right
Regression rate	Count	The number of regressions per second, per line, or paragraph, etc.
Regression in	Count	A regressive eye movement that lands in a predefined interest area
Regression out	Count	A regressive eye movement that is launched from a given interest area
<b>Skip</b>		
First-pass skipping rate	Count	The proportion of participants who skipped an AOI when first encountering it
Skip count	Count	The total number of times an interest area was skipped

<b>Pupil</b>		
Pupil diameter	Space	The pupil size for the current position of the eye
Pupil dilation latency	Time	The time elapsing between the onset of increased luminance (or other stimulus) and the beginning of pupil dilation
<b>Blink</b>		
Blink rate	Count	The number of blinks per unit of time
Blink duration	Time	The complete time from when the eyelid starts moving down until it is fully up again
<b>Gaze pattern</b>		
Heatmap	None	The visualization of the screen display, overlaid with a smooth landscape of fixation data represented in different colours
Scanpath	Space	An ordered set of fixations points (depicted by circles) connected by saccades (depicted by lines)



# Bibliografia

- ABU-DAYYEH, N. (2006), *Persisting vision: plans for a modern Arab capital, Amman, 1955-2002*, in "Planning Perspectives", 1, pp. 79-110.
- ABBOTT, M. J., & STAUB, A. (2015), *The effect of plausibility on eye movements in reading: Testing EZ Reader's null predictions*, in "Journal of Memory and Language", 85, pp. 76-87.
- ALLOPENNA, P. D., MAGNUSON, J. S., & TANENHAUS, M. K. (1998), *Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models*, in "Journal of Memory and Language", 38(4), pp. 419-439.
- ALTMANN, G. T. M. (2004), *Language-mediated eye movements in the absence of a visual world: The 'blank screen paradigm'*, in "Cognition", 93, pp. 79-87.
- ALTMANN, G. T. M., & KAMIDE, Y. (1999), *Incremental interpretation at verbs: Restricting the domain of subsequent reference*, in "Cognition", 73(3), pp. 247-264.
- ALTMANN, G. T. M., & KAMIDE, Y. (2007), *The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing*, in "Journal of Memory and Language", 57(4), pp. 502-518.
- AMENTA, S., MARELLI, M., & CREPALDI, D. (2015), *The fruitless effort of growing a fruitless tree: Early morpho-orthographic and morpho-semantic effects in sentence reading*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 41(5), pp. 1587.
- ANTES, J. R. (1974), *The time course of picture viewing*, in "Journal of Experimental Psychology", 103(1), pp. 62-70.
- ARCARE, G., SEMENZA, C., & BAMBINI, V. (2014), *Word structure and decomposition effects in reading*, in "Cognitive Neuropsychology", 31(1-2), pp. 184-218.
- BALOTA, D. A., POLLATSEK, A., & RAYNER, K. (1985), *The interaction of contextual constraints and parafoveal visual information in reading*, in "Cognitive Psychology", 17, pp. 364-390.

- BÉLANGER, N. N., & RAYNER, K. (2015), *What Eye Movements Reveal About Deaf Readers*, in "Current Directions in Psychological Science", 24(3), pp. 220-226.
- BÉLANGER, N. N., SLATTERY, T. J., MAYBERRY, R. I., & RAYNER, K. (2012), *Skilled deaf readers have an enhanced perceptual span in reading*, in "Psychological Science", 23(7), pp. 816-823.
- BERTERA, J. H., & RAYNER, K. (2000), *Eye movements and the span of effective vision in visual search*, in "Perception & Psychophysics", 62, pp. 576-585.
- BICKNELL, K., & LEVY, R. (2011), *Why readers regress to previous words: A statistical analysis*, in *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 33, pp. 931-936.
- BIRCH, S., & RAYNER, K. (2010), *Effects of syntactic prominence on eye movements during reading*, in "Memory & Cognition", 38(6), pp. 740-752.
- BOCK, K., IRWIN, D. E., DAVIDSON, D. J., & LEVELT, W. J. M. (2003), *Minding the clock*, in "Journal of Memory and Language", 48, pp. 653-658.
- BONIFACCI, P., TOBIA, V., SANSAVINI, A., & GUARINI, A. (2023), *Eye-movements in a text reading task: a comparison of preterm children, children with dyslexia and typical readers*, in "Brain Sciences", 13(3), pp. 425.
- BOSCH, J. E., CHAILLEUX, M., YEE, J. E., GUASTI, M. T., & AROSIO, F. (2022), *Prediction on the basis of gender and number in Mandarin-Italian bilingual children*, in "The Acquisition of Gender", pp. 243-272.
- BOSCH, J. E., & FOPPOLO, F. (2022), *Predictive processing of grammatical gender in bilingual children: The effect of cross-linguistic incongruency and language dominance*, in "Lingue e Linguaggio", 21(1), pp. 5-27.
- BOSCH, J. E., & FOPPOLO, F. (2024), *Prediction during spoken language processing in monolingual and multilingual children: Investigating the role of literacy*, in "Linguistic Approaches to Bilingualism", 14(4), pp. 512-543.
- BROUWER, S. (2010), *Processing strongly reduced forms in casual speech*, in "MPI Series in Psycholinguistics", 57, Wageningen: Ponsen & Looijen.
- BROWN-SCHMIDT, S., & TANENHAUS, M. K. (2006), *Watching the eyes when talking about size: An investigation of message formulation and utterance planning*, in "Journal of Memory and Language", 54, pp. 592-609.
- BROWN-SCHMIDT, S., GUNLOGSON, C., & TANENHAUS, M. K. (2008), *Addressees distinguish shared from private information when interpreting questions during interactive conversation*, in "Cognition", 107, pp. 1122-1134.
- BRYSSBAERT, M., & STEVENS, M. (2018), *Power analysis and effect size in mixed effects models: A tutorial*, in "Journal of Cognition", 1(1) 9, pp 1-20.
- BUSWELL, G. T. (1935), *How people look at pictures: A study of the psychology of perception in art*, University of Chicago Press, Chicago.
- CARROL, G., & CONKLIN, K. (2014), *Eye-tracking multi-word units: some methodological questions*, in "Journal of Eye Movement Research", 7(5).

- CARROL, G., & CONKLIN, K. (2020), *Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences*, in "Language and Speech", 63(1), pp. 95-122.
- CHAFFIN, R. (1997), *Associations to unfamiliar words: Learning the meanings of new words*, in "Memory & Cognition", 25, pp. 203-226.
- CHAFFIN, R., MORRIS, R. K., & SEELY, R. E. (2001), *Learning new word meanings from context: a study of eye movements*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 27(1), pp. 225-235.
- CHAMBERS, C. G., TANENHAUS, M. K., EBERHARD, K. M., FILIP, H., & CARLSON, G. N. (2002), *Circumscribing referential domains during real-time language comprehension*, in "Journal of Memory and Language", 47, pp. 30-49.
- CHAMBERS, C. G., TANENHAUS, M. K., EBERHARD, K. M., CARLSON, G. N., & FILIP, H. (1998), *Words and worlds: The construction of context for definite reference*, in *Proceedings of the twentieth annual conference of the Cognitive Science Society*, pp. 220-225.
- CHAMBERS, C. G., TANENHAUS, M. K., & MAGNUSON, J. S. (2004), *Actions and affordances in syntactic ambiguity resolution*, in "Journal of Experimental Psychology: Learning, Memory and Cognition", 30, pp. 687-696.
- CHAMBERS, C. G., & SAN JUAN, V. (2008), *Perception and presupposition in real-time language comprehension: Insights from anticipatory processing*, in "Cognition", 108(1), pp. 26-50.
- CHUN, D. M. (2012), *The role of technology in SLA research*, in "Language Learning & Technology", 16(1), pp. 69-82.
- CLIFTON, C., FERREIRA, F., HENDERSON, J. M., INHOFF, A. W., LIVERSHEDGE, S. P., REICHLER, E. D., & SCHOTTER, E. R. (2016), *Eye movements in reading and information processing: Keith Rayner's 40-year legacy*, in "Journal of Memory and Language", 86, pp. 1-19.
- CLIFTON JR, C., STAUB, A., & RAYNER, K. (2007), *Eye movements in reading words and sentences*, in "Eye Movements", pp. 341-371.
- ÇOKAL, D., STURT, P., & FERREIRA, F. (2007), *Deixis: This and that in written narrative discourse*, in "Discourse Processes", 51(3), pp. 201-22.
- CONCETTI, L., & MOSCATI, V. (2023), *The Unexpected Lightness of the Main Verb: An Eye-Tracking Study on Relative Clauses and Trace Reactivation*, in "Quaderni di Linguistica e Studi Orientali", 9, pp. 45-58.
- CONKLIN, K., & SCHMITT, N. (2012), *The processing of formulaic language*, in "Annual Review of Applied Linguistics", 32, pp. 45-61.
- CONKLIN, K., PELLICER-SANCHEZ, A., CARROL, G. (2018), *Eye tracking - A guide for applied linguistics research*, Cambridge University Press, Cambridge, p. 2.
- COOPER, R. (1974), *The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing*, in "Cognitive Psychology", 6(1), pp. 84-107.

- CUTTER, M. G., DRIGHE, D., & LIVERSEDGE, S. P. (2014), *Preview benefit in English spaced compounds*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 40(6), pp. 1778.
- DAHAN, D., TANENHAUS, M. K., & CHAMBERS, C. G. (2002), *Accent and reference resolution in spoken-language comprehension*, in "Journal of Memory and Language", 47, pp. 292-314.
- DAHAN, D., & TANENHAUS, M. K. (2004), *Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 30, pp. 498-513.
- DE GRAEF, P., CHRISTIAENS, D., D'YDEWALLE, G. (1990), *Perceptual effects of scene context on object identification*, in "Psychological Research", 52, pp. 317-329.
- DE LUCA, M., DI PACE, E., JUDICA, A., SPINELLI, D., & ZOCCOLOTTI, P. (1999), *Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia*, in "Neuropsychologia", 37(12), pp. 1407-1420.
- DE LUCA, M., ZERI, F., SPINELLI, D., & ZOCCOLOTTI, P. (2010), *The acquisition of reading fluency in an orthographically transparent language (Italian): an eye movement longitudinal study*, in "Medical Science Monitor: International Medical Journal of Experimental and Clinical Research", 16(3), pp. SC1-7.
- DUFFY, S. A., MORRIS, R. K., & RAYNER, K. (1988), *Lexical ambiguity and fixation times in reading*, in "Journal of Memory and Language", 27, pp. 429-446.
- DUNCAN, J., & HUMPHREYS, G. W. (1989), *Visual search and stimulus similarity*, in "Psychological Review", 96(3), pp. 433.
- EHRlich, S. F., & RAYNER, K. (1981), *Contextual effects on word recognition and eye movements during reading*, in "Journal of Verbal Learning and Verbal Behavior", 20, pp. 641-655.
- ENGBERT, R., NUTHMANN, A., RICHTER, E. M., & KLIEGL, R. (2005), *SWIFT: a dynamical model of saccade generation during reading*, in "Psychological Review", 112(4).
- FINDLAY, J. M. (1997), *Saccade target selection during visual search*, in "Vision Research", 37, pp. 617-631.
- FINDLAY, J. M. (1997), *Eye Scanning and Visual Search*, in J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 134-159), Psychology Press, 2004.
- FIORAVANTI, I., FORTI, L., D'ALEsIO, V., ROCCAForTE, M., SPINA, S., & KOESTERS GENSINI, S. (2023), *Insights into Phraseological Processing through Stimuli Modification: An Exploratory Eye-Tracking Study on Native Speakers and Learners of Italian*, in "Languages", 9(1), p. 14.
- FOPPOLO, F., MARELLI, M., MERONI, L., & GUALMINI, A. (2015), *Hey little sister, who's the only one? Modulating informativeness in the resolution of pronominal ambiguity*, in "Cognitive Science", 39(7), pp. 1646-1674.
- FOPPOLO, F., & MARELLI, M. (2017), *No delay for some inferences*, in "Journal of Semantics", 34(4), pp. 659-681.

- FOPPOLO, F., BOSCH, J. E., GRECO, C., CARMINATI, M. N., & PANZERI, F. (2021), *Draw a star and make it perfect: Incremental processing of telicity*, in "Cognitive Science", 45(10), e13052.
- FORAKER, S., & MURPHY, G. L. (2012), *Polysemy in sentence comprehension: Effects of meaning dominance*, in "Journal of Memory and Language", 67(4), pp. 407-425.
- FRANCIS, W. N., & KUCERA, H. (1982), *Frequency analysis of English usage: Lexicon and grammar*, Boston: Houghton Mifflin.
- FRISSON, S., & PICKERING, M. J. (1999), *The processing of metonymy: Evidence from eye-movements*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 25, pp. 1366-1383.
- FRISSON, S., RAYNER, K., & PICKERING, M. J. (2005), *Effects of contextual predictability and transitional probability on eye movements during reading*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 31, pp. 862-877.
- GIBBS JR, R. W., & NAYAK, N. P. (1989), *Psycholinguistic studies on the syntactic behavior of idioms*, in "Cognitive Psychology", 21(1), pp. 100-138.
- GIBSON, J. J. (1977), *The theory of affordances*, in *Perceiving, acting and knowing: Towards an ecological psychology*, Erlbaum.
- GLEITMAN, L., JANUARY, D., NAPPA, R., & TRUESWELL, J. C. (2007), *On the give and take between event apprehension and utterance formulation*, in "Journal of Memory and Language", 57(4), pp. 544-569.
- GODFROID, A. (2019), *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*, Routledge.
- GOLDSMITH, J. (1990), *Autosegmental and Metrical Phonology*, Oxford and Cambridge, MA: Basil Blackwell.
- GRIFFIN, Z. M. (2004), *Why look? Reasons for eye movements related to language production*, in J. Henderson, F. Ferreira (Eds.), *The interface of language, vision, and action*, Psychology Press.
- GRIFFIN, Z., & BOCK, K. (2000), *What the eyes say about speaking*, in "Psychological Science", 11, pp. 274-279.
- HAYHOE, M. M., SHRIVASTAVA, A., MRUCZEK, R., & PELZ, J. B. (2003), *Visual memory and motor planning in a natural task*, in "Journal of Vision", 3, pp. 49-63.
- HENDERSON, J. M. (1993), *Eye movement control during visual object processing: Effects of initial fixation position and semantic constraint*, in "Canadian Journal of Experimental Psychology", 47, pp. 79-98.
- HENDERSON, J. M., WEEKS, P. A. JR., & HOLLINGWORTH, A. (1999), *Effects of semantic consistency on eye movements during scene viewing*, in "Journal of Experimental Psychology: Human Perception and Performance", 25, pp. 210-228.
- HENDERSON, J. M., & FERREIRA, F. (2004), *Scene perception for psycholinguists*, in J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action* (pp. 1-58). Hove: Psychology Press.

- HENDERSON, J. M., & HAYES, T. R. (2017), *Meaning-based guidance of attention in scenes as revealed by meaning maps*, in "Nature Human Behaviour", 1(10), pp. 743-747.
- HIROSE, Y. (2020), *Sequential interpretation of pitch prominence as contrastive and syntactic information: contrast comes first, but syntax takes over*, in "Language and Speech", 63(3), pp. 455-478.
- HIROSE, Y., & MAZUKA, R. (2015), *Predictive processing of novel compounds: Evidence from Japanese*, in "Cognition", 136, pp. 350-358.
- HIROSE, Y., & MAZUKA, R. (2017), *Exploiting pitch accent information in compound processing: A comparison between adults and 6-to 7-year-old children*, in "Language Learning and Development", 13(4), pp. 375-394.
- HIROTANI, M., FRAZIER, L., & RAYNER, K. (2006), *Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements*, in "Journal of Memory and Language", 54(3), pp. 425-443.
- HU, X., & ARYADOUST, V. (2024), *A Systematic Review of Eye-Tracking Technology in Second Language Research*, in "Languages", 9(4), p. 141.
- HUESTEGGE, F., REICHLER, E. D., RAYNER, K., & POLLATSEK, A. (2003), *The EZ Reader model of eye-movement control in reading: Comparisons to other models*, in "Behavioral and Brain Sciences", 26(4), pp. 445-476.
- HUETTIG, F., ROMMERS, J., & MEYER, A. S. (2011), *Using the visual world paradigm to study language processing: A review and critical evaluation*, in "Acta Psychologica", 137(2), pp. 151-171.
- HUETTIG, F., & MCQUEEN, J. M. (2007), *The tug of war between phonological, semantic, and shape information in language-mediated visual search*, in "Journal of Memory and Language", 54, pp. 460-482.
- INHOFF, A.W. (1989), *Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations?*, in "Journal of Memory and Language", 28, pp. 444-461.
- INHOFF, A.W. (1989b), *Parafoveal processing of words and saccade computation during eye fixations in reading*, in "Journal of Experimental Psychology: Human Perception and Performance", 15, pp. 544-555.
- INHOFF, A.W. (1990), *Integrating information across eye fixations in reading*, in "Acta Psychologica", 73, pp. 281-297.
- INHOFF, A. W. & RAYNER, K. (1986), *Parafoveal word processing during eye fixations in reading: Effects of word frequency*, in "Perception and Psychophysics", 40(6), pp. 431-439.
- INDRARATHNE, B., & KORMOS, J. (2017), *Attentional processing of input in explicit and implicit learning conditions: an eye-tracking study*, in "Studies in Second Language Acquisition", 39(3), pp. 401-430. Cambridge University Press & Assessment.
- ISSA, B. I., & MORGAN-SHORT, K. (2019), *Effects of external and internal attentional allocation on second language grammar development: An eye-tracking study*, in "Studies in Second Language Acquisition", 41(2), pp. 389-417.

- ITTI, L., & KOCH, C. (2000), *A saliency-based search mechanism for overt and covert shifts of visual attention*, in "Vision Research", 40, pp. 1489-1506.
- JUN, S. A. (Ed.) (2005), *Prosodic typology: The phonology of intonation and phrasing* (Vol. 1). Oxford University Press.
- JUST, M. A., & CARPENTER, P. A. (1980), *A theory of reading: From eye fixations to comprehension*, in "Psychological Review", 87(4), pp. 329-354.
- KAMIDE, Y., ALTMANN, G. T. M., & HAYWOOD, S. L. (2003), *The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements*, in "Journal of Memory and Language", 49(1), pp. 133-156.
- KEATING, G. & JEGERSKI, J. (2015), *Experimental designs in sentence processing research*, in "Studies in Second Language Acquisition", 37(1), pp. 1-32.
- KEYSAR, B., BARR, D. J., BALIN, J. A., & BRAUNER, J. S. (2000), *Taking perspective in conversation: The role of mutual knowledge in comprehension*, in "Psychological Science", 11, pp. 32-38.
- KLIEGL, R., GRABNER, E., ROLFS, M., & ENGBERT, R. (2004), *Length, frequency, and predictability effects of words on eye movements in reading*, in "European Journal of Cognitive Psychology", 16(1-2), pp. 262-284.
- KLIEGL, R., NUTHMANN, A., ENGBERT, R. (2006), *Tracking the mind during reading: The influence of past, present, and future words on fixation durations*, in "Journal of Experimental Psychology: General", 135(1), pp. 12-35.
- KUKONA, A., FANG, S. Y., AICHER, K. A., CHEN, H., & MAGNUSON, J. S. (2011), *The time course of anticipatory constraint integration*, in "Cognition", 119(1), pp. 23-42.
- LADD, D. R. (1996), *Intonational phonology*, Cambridge University Press, Cambridge.
- LAND, M. F., MENNIE, N., & RUSTED, J. (1999), *Eye movements and the roles of vision in activities of daily living: Making a cup of tea*, in "Perception", 28, pp. 1311-1328.
- LOFTUS, G. R., & MACKWORTH, N. H. (1978), *Cognitive determinants of fixation location during picture viewing*, in "Journal of Experimental Psychology: Human Perception and Performance", 4, pp. 565-572.
- LOTTO, L., DELL'ACQUA, R., & JOB, R. (2001), *Le figure PD/DPSS. Misura di accordo sul nome, tipicità, familiarità, età di acquisizione e tempi di denominazione per 266 figure*, in "Giornale Italiano di Psicologia", 28(1), pp. 193-210.
- MACKWORTH, N. H., & MORANDI, A. J. (1967), *The gaze selects informative details within pictures*, in "Perception & Psychophysics", 2(11), pp. 547-552.
- MARELLI, M., & LUZZATTI, C. (2012), *Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency*, in "Journal of Memory and Language", 66(4), pp. 644-664.
- MATSUKI, K., CHOW, T., HARE, M., ELMAN, J. L., SCHEEPERS, C., & MCRAE, K. (2011), *Event-based plausibility immediately influences on-line language comprehension*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 37, pp. 913-934.

- MCCONKIE, G. W., KERR, P. W., REDDIX, M. D., ZOLA, D., & JACOBS, A. M. (1989), *Eye movement control in reading: II. Frequency of refixating a word*, in "Perception & Psychophysics", 46, pp. 245-253.
- MCDONALD, S. A., & SHILLCOCK, R. C. (2003a), *Eye movements reveal the on-line computation of lexical probabilities during reading*, in "Psychological Science", 14, pp. 648-652.
- MCDONALD, S. A., & SHILLCOCK, R. C. (2003b), *Low-level predictive inference in reading: The influence of transitional probabilities on eye movements*, in "Vision Research", 43, pp. 1735-1751.
- MAGNUSON, J. S. (2019), *Fixations in the visual world paradigm: Where, when, why?*, in "Journal of Cultural Cognitive Science", 3(2), pp. 113-139.
- MEYER, A. S., SLEIDERINK, A. M., & LEVELT, W. J. M. (1998), *Viewing and naming objects: Eye movements during noun phrase production*, in "Cognition", 66(2), pp. B25-B33.
- MEYER, A. S., VAN DER MEULEN, F., & BROOKS, A. (2004), *Eye movements during speech planning: Speaking about present and remembered objects*, in "Visual Cognition", 11, pp. 553-576.
- MILLER, P. (2010a), *Similarities and differences in the processing of written text by skilled and less skilled readers with prelingual deafness*, in "The Journal of Special Education", 46(4), pp. 233-244.
- MILLER, P. (2010b), *Phonological, Orthographic, and Syntactic Awareness and their Relation to Reading Comprehension in Prelingually Deaf Individuals: What Can We Learn from Skilled Readers?*, in "Journal of Developmental and Physical Disabilities", 22, pp. 549-580.
- MORAN, R., ZEHETLEITNER, M., LIESEFELD, H. R., MÜLLER, H. J., & USHER, M. (2016), *Serial vs. parallel models of attention in visual search: accounting for benchmark RT-distributions*, in "Psychonomic Bulletin & Review", 23, pp. 1300-1315.
- NUZZO, E., & RASTELLI, S. (2011), *Glottodidattica sperimentale. Nozioni, rappresentazioni e processing nell'apprendimento della seconda lingua*, Carocci, Roma.
- OLIVA, A., & TORRALBA, A. (2006), *Building the gist of a scene: The role of global image features in recognition*, in "Progress in Brain Research", 155, pp. 23-36.
- O'REGAN, J. K. (1990), *Eye movements and reading*, in E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes*, Elsevier, Amsterdam, pp. 395-453.
- O'REGAN, J. K., & JACOBS, A. M. (1992), *The optimal viewing position effect in word recognition: A challenge to current theory*, in "Journal of Experimental Psychology: Human Perception and Performance", 18, pp. 185-197.
- PALMER, J. (1995), *Attention in visual search: Distinguishing four causes of a set-size effect*, in "Current Directions in Psychological Science" (Wiley-Blackwell), 4(4), pp. 118-123.
- RAYNER, K. (1998), *Eye movements in reading and information processing: 20 years of research*, in "Psychological Bulletin", 124(3), pp. 372-422.

- RAYNER, K. (1998), *Eye movements and attention in reading, scene perception, and visual search*, in "Quarterly Journal of Experimental Psychology", 62(8), pp. 1457-1506.
- RAYNER, K. (2009), *The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search*, in "Quarterly Journal of Experimental Psychology", 62(8), pp. 1457-1506.
- RAYNER, K., SLATTERY, T. J., & BÉLANGER, N. N. (2010), *Eye movements, the perceptual span, and reading speed*, in "Psychonomic Bulletin & Review", 17(6), pp. 834-839.
- RAYNER, K., & DUFFY, S. (1986), *Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity*, in "Memory & Cognition", 14, pp. 191-201.
- RAYNER, K., & DUFFY, S. A. (1987), *Eye movements and lexical ambiguity*, in J. K. O'Regan & L. A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition*, North-Holland, Amsterdam, pp. 521-529.
- RAYNER, K., & FISHER, D. L. (1987a), *Eye movements and the perceptual span during visual search*, in J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition*, Elsevier Science Publishers, New York, pp.293-302.
- RAYNER, K., & FISHER, D. L. (1987b), *Letter processing during eye fixations in visual search*, in "Perception & Psychophysics", 42, pp. 87-100.
- RAYNER, K., PACHT, J. M., & DUFFY, S. A. (1994), *Effects of prior encounter and discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations*, in "Journal of Memory & Language", 33, pp. 527-544.
- RAYNER, K., & WELL, A. D. (1996), *Effects of contextual constraint on eye movements in reading: A further examination*, in "Psychonomic Bulletin & Review", 3, pp. 504-509.
- RAYNER, K., KAMBE, G., & DUFFY, S. A. (2000), *The effect of clause wrap-up on eye movements during reading*, in "The Quarterly Journal of Experimental Psychology: Section A", 53(4), pp. 1061-1080.
- RAYNER, K., WARREN, T., JUHASZ, B. J., & LIVERSEDGE, S. P. (2004), *The effect of plausibility on eye movements in reading*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 30, pp. 1290-1301.
- REDOLFI, M., & MELLONI, C. (2025), *Processing adjectives in development: Evidence from eye-tracking*, in "Journal of Child Language", 52(2), pp. 270-293.
- REICHLER, E. D., RAYNER, K., & POLLATSEK, A. (2003), *The EZ Reader model of eye-movement control in reading: Comparisons to other models*, in "Behavioral and Brain Sciences", 26(4), pp. 445-476.
- REICHLER, E. D., POLLATSEK, A., & RAYNER, K. (2006), *E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading*, in "Cognitive Systems Research", 7(1), pp. 4-22.
- REILLY, R. G., & RADACH, R. (2003), *Foundations of an interactive activation model of eye movement control in reading*, in J. Hyönä, R. Radach, and H. Deubel

- (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, North-Holland, Amsterdam, pp. 429-455.
- REINISCH, E., JESSE, A., & MCQUEEN, J. M. (2010), *Early use of phonetic information in spoken word recognition: Lexical stress drives eye-movements immediately*, in "Quarterly Journal of Experimental Psychology", 63(4), pp. 772-783.
- RONCA, D., & MOSCATI, V. (2019), *The interaction of morphological gender with stereotypical information: An eye tracking study on gender inferences*, in "International Journal of Linguistics", 11(4), pp. 111-125.
- SALVERDA, A. P., & TANENHAUS, M. K. (2017), *The visual world paradigm*, in A. de Groot, P. Hagoort (Eds.) *Research methods in psycholinguistics and the neurobiology of language: A practical guide*, pp. 89-110.
- SAUPPE, S. (2016), *Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences*, in "Frontiers in Psychology", 7.
- SCHMIDT-WEIGAND, F., KOHNERT, A., & GLOWALLA, U. (2010), *A closer look at split visual attention in system- and self-paced instruction in multimedia learning*, in "Learning and Instruction", 20(2), pp. 100-110.
- SCHOTTER, E. R., ANGELE, B., & RAYNER, K. (2012), *Parafoveal processing in reading*, in "Attention, Perception, & Psychophysics", 74(1).
- SCHOTTER, E. R., TRAN, R., & RAYNER, K. (2014), *Don't Believe What You Read (Only Once) Comprehension Is Supported by Regressions During Reading*, in "Psychological Science", 25(6).
- SERENO, S. C. (1995), *Resolution of lexical ambiguity: evidence from an eye movement priming paradigm*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 21(3), pp. 582.
- SERENO, S. C., PACTH, J. M., & RAYNER, K. (1992), *The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations*, in "Psychological Science", 3(5), pp. 296-301.
- SIYANOVA-CHANTURIA, A. (2013), *Eye-tracking and ERPs in multi-word expression research: A state-of-the-art review of the method and findings*, in "The Mental Lexicon", 8(2), pp. 245-268.
- SIYANOVA-CHANTURIA, A., CONKLIN, K., & SCHMITT, N. (2011), *Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers*, in "Second Language Research", 27(2), pp. 251-272.
- SIYANOVA-CHANTURIA, A., CONKLIN, K., & VAN HEUVEN, W. J. (2011), *Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences*, in "Journal of Experimental Psychology: Learning, Memory, and Cognition", 37(3).
- SNEDEKER, J., & TRUESWELL, J. C. (2003), *Using prosody to avoid ambiguity: Effects of speaker awareness and referential context*, in "Journal of Memory and Language", 48, pp. 103-130.
- SNEDEKER, J., & TRUESWELL, J. C. (2004), *The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing*, in "Cognitive Psychology", 49(3), pp. 238-299.

- SNODGRASS, J. G., & VANDERWART, M. (1980), *A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity*, in "Journal of Experimental Psychology: Human Learning and Memory", 6(2), pp. 174-215.
- SOLAJA, O., & CREPALDI, D. (2024), *The role of morphology in novel word learning: a registered report*, in "Royal Society Open Science", 11(6), pp. 230094.
- SONBUL, S. (2015), *Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing*, in "Bilingualism: Language and Cognition", 18, pp. 419-437.
- TANENHAUS, M. K., SPIVEY-KNOWLTON, M. J., EBERHARD, K. M., & SEDIVY, J. C. (1995), *Integration of visual and linguistic information in spoken language comprehension*, in "Science", 268(5217), pp. 1632-1634.
- TANENHAUS, M. K., MAGNUSON, J. S., DAHAN, D., & CHAMBERS, C. (2000), *Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing*, in "Journal of Psycholinguistic Research", 29(6), pp. 557-580.
- TATLER, B. W., HAYHOE, M. M., LAND, M. F., & BALLARD, D. H. (2011), *Eye guidance in natural vision: Reinterpreting salience*, in "Journal of Vision", 11(5).
- THORNTON, T. L., & GILDEN, D. L. (2007), *Parallel and serial processes in visual search*, in "Psychological Review", 114(1).
- TOMASUOLO, E., ROCCAFORTE, M., & DI FABIO, A. (2019), *Reading and deafness: Eye tracking in deaf readers with different linguistic background*, in "Applied Linguistics", 40(6), pp. 992-1008.
- TOWNSEND, J. T. (1972), *Some results concerning the identifiability of parallel and serial processes*, in "British Journal of Mathematical and Statistical Psychology", 25, pp. 168-199.
- TOWNSEND, J. T. (1990), *Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished*, in "Psychological Science", 1, pp. 46-54.
- TRAFICANTE, D., MARELLI, M., & LUZZATTI, C. (2018), *Effects of reading proficiency and of base and whole-word frequency on reading noun-and verb-derived words: An eye-tracking study in Italian primary school children*, in "Frontiers in Psychology", 9, 2335.
- TREISMAN, A., & GELADE, G. (1980), *A feature integration theory of attention*, in "Cognitive Psychology", 12, pp. 97-136.
- TRUKENBROD, H. A., & ENGBERT, R. (2007), *Oculomotor control in a sequential search task*, in "Vision Research", 47, pp. 2426-2443.
- UNDERWOOD, G., SCHMITT, N., & GALPIN, A. (2004), *The eyes have it: An eye-movement study into the processing of formulaic sequences*, in N. Schmitt (Ed.), *Formulaic sequences*, John Benjamins, Amsterdam, pp. 153-172.
- VAN ZOEST, W., DONK, M., & THEEUWES, J. (2004), *The role of bottom-up control in saccadic eye movements*, in "Journal of Experimental Psychology: Human Perception and Performance", 30, pp. 746-759.

- VENDER, M., NARDON, A., & VENAGLI, I. (2025), *Topicality, accessibility, and causality in anaphora resolution: An eye-tracking study of null and overt pronouns in Italian*, in "Intercultural Pragmatics", 22(1), pp. 147-172.
- VLASKAMP, B. N. S., & HOOGE, I. T. C. (2006), *Crowding degrades visual search*, in "Vision Research", 46, pp. 417-425.
- VITU, F. (1991), *The existence of a center of gravity effect during reading*, in "Vision Research", 31(7-8), pp. 1289-1313.
- YARBUS, A. (1967), *Eye movements and vision*, New York: Plenum Press.
- WRIGHT, R.D., WARD, L.M. (2008), *Orienting of Attention*, Oxford University Press.
- WARREN, T., & MCCONNELL, K. (2007), *Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading*, in "Psychonomic Bulletin & Review", 14, pp. 770-775.
- WATSON, D. G., TANENHAUS, M., & GUNLOGSON, C. (2008), *Interpreting pitch accents in online comprehension: H vs LH\**, in "Cognitive Science", 32, pp. 1232-1244.
- WEBER, A., BRAUN, B., & CROCKER, M. W. (2006), *Finding referents in time: Eye-tracking evidence for the role of contrastive accents*, in "Language and Speech", 49, pp. 367-392.
- WHITE, S. J. (2008), *Eye movement control during reading: Effects of word frequency and orthographic familiarity*, in "Journal of Experimental Psychology: Human Perception and Performance", 34(1), pp. 205-223.
- WILLIAMS, R., & MORRIS, R. (2004), *Eye movements, word familiarity, and vocabulary acquisition*, in "European Journal of Cognitive Psychology", 16(1-2), pp. 312-339.
- WILLIAMS, C. C., & POLLATSEK, A. (2007), *Searching for an O in an array of Cs: Eye movements track moment-to-moment processing in visual search*, in "Perception & Psychophysics", 69, pp. 372-381.
- WINKE, P., GASS, S., & SYDORENKO, T. (2013), *Factors influencing the use of captions by foreign language learners: An eye-tracking study*, in "Modern Language Journal", 97(1), pp. 254-275.
- WOLFE, J. M. (2007), *Guided Search 4.0: Current progress with a model of visual search*, in W. D. Gray (Ed.), *Integrated models of cognitive systems*, pp. 99-119, Oxford University Press.
- WOLFE, J. M., & HOROWITZ, T. S. (2004), *What attributes guide the deployment of visual attention and how do they do it?*, in "Nature Reviews Neuroscience", 5(6), pp. 495-501.
- WOLFE, J. M., & HOROWITZ, T. S. (2017), *Five factors that guide attention in visual search*, in "Nature Human Behaviour", 1(3).
- ZOCOLOTI, P., DE LUCA, M., DI PACE, E., JUDICA, A., ORLANDI, M., & SPINELLI, D. (1999), *Markers of developmental surface dyslexia in a language (Italian) with high grapheme-phoneme correspondence*, in "Applied Psycholinguistics", 20(2), pp. 191-216.



CONSIGLIO SCIENTIFICO-EDITORIALE  
SAPIENZA UNIVERSITÀ EDITRICE

*Presidente*

AUGUSTO ROCA DE AMICIS

*Membri*

MARCELLO ARCA

ORAZIO CARPENZANO

MARIANNA FERRARA

CRISTINA LIMATOLA

ENRICO ROGORA

FRANCESCO SAITTO

Opera sottoposta a peer review. Il Consiglio scientifico-editoriale, anche attraverso i comitati scientifici di serie, assicura una valutazione trasparente e indipendente delle opere sottoponendole in forma anonima a due valutatori ignoti agli autori e ai curatori. Per ulteriori dettagli si rinvia al sito: [www.editricesapienza.it](http://www.editricesapienza.it)

*This work has been subjected to a peer review. The Scientific-editorial Board, also through the scientific committees of series, ensures a transparent and independent evaluation of the works by subjecting them anonymously to two reviewers, unknown to the authors and editors. For further details please visit the website: [www.editricesapienza.it](http://www.editricesapienza.it)*

COLLANA STUDI E RICERCHE

Per informazioni sui volumi precedenti della collana, consultare il sito:  
[www.editricesapienza.it](http://www.editricesapienza.it) | *For information on the previous volumes included  
in the series, please visit the following website: [www.editricesapienza.it](http://www.editricesapienza.it)*

164. One protein many functions: the non-canonical interactions of SHMT1  
The structural and functional characterization of SHMT1 interactions  
with RNA and in the *de novo* thymidylate synthesis complex  
*Sharon Spizzichino*
165. "One Step Beyond the Hero"  
Disrupting War and Violence in American Literature and Culture  
Collected Essays  
*Giorgio Mariani*
166. For a sociology of local innovation ecosystems  
A work in progress on NRRP and the Rome Technopole  
*Tommaso Fasciani, posthumous work edited by Ernesto d'Albergo*
167. Juegos de miradas: relaciones transatlánticas entre Italia,  
España y América Latina  
*Chiara Bolognese y Beatriz Ferrús Antón (eds.)*
168. Nuovi scenari per ambienti di inclusione sociale  
Ricerche psico-pedagogiche  
*a cura di Guido Benvenuto e Fiorenzo Laghi*
169. L'Africa degli eurocomunisti  
Comunisti italiani e francesi tra crisi dei socialismi africani  
e rapporto Nord-Sud (1969-1984)  
*Gabriele Siracusanò*
170. Conservation Process Model  
An ontology for Conservation in Architecture  
*Marta Acierno and Donatella Fiorani*
171. Percorsi in Civiltà dell'Asia e dell'Africa IV  
Quaderni di studi dottorali alla Sapienza  
*a cura di Mario Casari*
172. Tecniche Sapienti  
Storie di architetture e ingegnere in Sapienza 1910-1968  
*a cura di Claudia Mattoño e Chiara Belingardi*
173. Tracciare la comprensione  
L'eye tracking nella ricerca linguistica  
*Maria Roccaforte e Veronica D'Alesio*







Il volume si propone come una guida introduttiva all'uso dell'eye tracking nel campo della linguistica, per offrire al lettore una panoramica delle principali tecnologie associate a questo strumento e dei paradigmi sperimentali più utilizzati nelle scienze del linguaggio. Particolare attenzione è dedicata all'acquisizione delle lingue seconde e all'elaborazione dell'italiano in bambini e adulti, con esempi tratti da contesti sperimentali. Il testo si rivolge a studenti, studiosi e ricercatori interessati ad approfondire gli aspetti metodologici ed empirici legati all'analisi del comportamento oculare in ambito linguistico, utile per la progettazione e l'interpretazione di studi sperimentali condotti con questo strumento di indagine.

**Maria Roccaforte** insegna Linguistica applicata presso Sapienza Università di Roma. Nelle sue ricerche indaga, anche attraverso l'eye tracking, i pattern di percezione visiva delle lingue vocali da parte di apprendenti sordi e i processi di comprensione delle lingue dei segni da parte di udenti; si occupa di didattica dell'Italiano come L2 e di descrizione linguistica della Lingua dei segni italiana.

**Veronica D'Alesio**, ha conseguito il Dottorato di ricerca in Linguistica presso Sapienza Università di Roma e attualmente svolge attività di ricerca nell'ambito dell'elaborazione del linguaggio in lettura. I suoi interessi scientifici includono l'acquisizione di lingue seconde e l'utilizzo di metodologie online, particolarmente l'eye tracking, nell'indagine linguistica.

ISBN 978-88-9377-387-4



9 788893 773874

