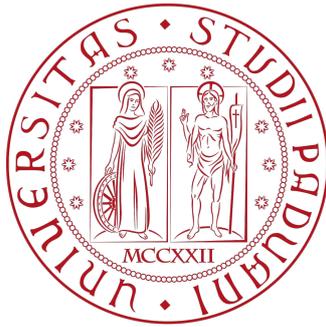


**UNIVERSITÀ DEGLI STUDI DI PADOVA**

**DIPARTIMENTO DI SCIENZE STATISTICHE**

**CORSO DI LAUREA MAGISTRALE IN SCIENZE STATISTICHE**



**TESI MAGISTRALE**

**Analisi statistica dell'inclusività linguistica  
tramite modelli a effetti misti incrociati**

**Relatore:** DOTT. PIETRO BELLONI

DIPARTIMENTO DI SCIENZE STATISTICHE - UNIVERSITÀ DEGLI STUDI DI  
PADOVA

**Correlatrice:** PROF.SSA LAURA NADAL SANCHIS

DIPARTIMENTO DI STUDI LINGUISTICI E CULTURALI COMPARATI -  
UNIVERSITÀ CA' FOSCARI VENEZIA

**Laureando:** MATTEO ALLEGRI

**Matricola N°:** 2062087

**ANNO ACCADEMICO 2023/2024**



# Indice

|   |           |
|---|-----------|
| <b>Introduzione</b>   | <b>2</b>  |
| <b>1 Dati e presentazione degli studi</b>                                       | <b>6</b>  |
| 1.1 Indagine di percezione Studio Psychopy . . . . .                            | 6         |
| 1.2 Items sperimentali . . . . .  | 8         |
| 1.3 Partecipanti . . . . .  | 9         |
| 1.4 Tecnica e procedura . . . . .   | 9         |
| 1.5 Dataset finale . . . . .  | 10        |
| 1.6 Indagine di percezione Studio Eyetracking . . . . .                         | 11        |
| 1.7 L'Eyetracking . . . . .   | 12        |
| 1.8 Items sperimentali . . . . .  | 13        |
| 1.9 Tecnica e procedura . . . . .   | 14        |
| 1.10 Partecipanti . . . . .   | 15        |
| 1.11 Dataset finale . . . . .   | 15        |
| <b>2 Modelli</b>  | <b>17</b> |
| 2.1 Modelli a effetti misti incrociati . . . . .                                | 18        |
| 2.1.1 Modello lineare a effetti misti . . . . .                                 | 20        |
| 2.1.2 Procedura di stima tramite massima verosimiglianza . . . . .              | 22        |
| 2.1.3 Procedura di stima tramite massima verosimiglianza<br>ristretta . . . . . | 23        |
| 2.1.4 Modello logistico a effetti misti . . . . .                               | 24        |
| 2.1.5 Modello logistico multinomiale a effetti misti . . . . .                  | 26        |
| 2.1.6 Modello di Cox a effetti misti . . . . .                                  | 27        |
| 2.2 Modelli per dati ordinali . . . . .   | 29        |
| 2.2.1 Modelli CUB . . . . .   | 30        |

|          |  |            |
|----------|--|------------|
| 2.2.2    | Modelli CUBE e IHG CUB . . . . .                               | 31         |
| <b>3</b> | <b>Analisi dello studio Psychopy</b>                           | <b>33</b>  |
| 3.1      | Analisi preliminare . . . . .                                  | 33         |
| 3.2      | Analisi esplorativa . . . . .                                  | 34         |
| 3.3      | Modelli per il tempo di lettura . . . . .                      | 37         |
| 3.4      | Modelli per l'esito della risposta . . . . .                   | 42         |
| 3.5      | Analisi di sopravvivenza . . . . .                             | 47         |
| 3.6      | Modello logit multinomiale con effetti misti . . . . .         | 51         |
| 3.7      | Modelli CUB e CUBE . . . . .                                   | 56         |
| 3.8      | Conclusioni sullo studio . . . . .                             | 62         |
| <b>4</b> | <b>Studio Eyetracking</b>                                      | <b>64</b>  |
| 4.1      | Valori anomali . . . . .                                       | 65         |
| 4.2      | Modelli per la frase completa . . . . .                        | 67         |
| 4.2.1    | Modelli per tempo totale di lettura . . . . .                  | 68         |
| 4.2.2    | Modelli per la prima fase di lettura . . . . .                 | 73         |
| 4.2.3    | Modelli per la rilettura . . . . .                             | 77         |
| 4.3      | Modelli per i tempi di lettura delle parole iniziali . . . . . | 82         |
| 4.3.1    | Modelli per tempo totale di lettura . . . . .                  | 82         |
| 4.3.2    | Modelli per la prima fase di lettura . . . . .                 | 87         |
| 4.3.3    | Modelli per la rilettura . . . . .                             | 92         |
| 4.4      | Modelli per la quinta parola dell'item . . . . .               | 96         |
| 4.4.1    | Modelli per tempo totale di lettura . . . . .                  | 96         |
| 4.4.2    | Modelli per la prima fase di lettura della quinta parola       | 101        |
| 4.4.3    | Modelli per la rilettura . . . . .                             | 105        |
| 4.5      | Conclusioni sullo studio . . . . .                             | 109        |
| <b>5</b> | <b>Limiti dello studio</b>                                     | <b>111</b> |
| <b>6</b> | <b>Conclusioni</b>   | <b>112</b> |



# Introduzione

Il dibattito sull'utilizzo del linguaggio inclusivo in Italia ha preso avvio fin dalla fine degli anni '80 del Novecento, principalmente grazie agli importanti contributi di Alma Sabatini, i cui lavori sono stati promossi dal Consiglio dei Ministri. Le ricerche di Sabatini (1986, 1987), infatti, hanno rappresentato un punto di riferimento per la sensibilizzazione e la riflessione sulla disparità di genere nel linguaggio nella cultura italiana. Questi lavori non solo hanno ispirato un ampio dibattito pubblico, ma hanno anche stimolato ulteriori ricerche e discussioni su come il linguaggio possa influenzare e riflettere dinamiche di genere nella società. I lavori di Sabatini si sono sviluppati a partire dall'osservazione di tendenze simili in studi condotti in altre lingue, in particolare basandosi sulle ricerche effettuate negli Stati Uniti durante gli anni '70. Questi studi americani hanno esplorato come le strutture linguistiche potessero perpetuare distinzioni di genere e disparità, fornendo così un modello e una base teorica per le indagini di Sabatini sulla lingua italiana.

La crescente attenzione verso un linguaggio più inclusivo emerge come risposta alle iniziative di varie organizzazioni che sostengono l'importanza di rappresentare equamente entrambi i generi nella comunicazione. In particolare, si è osservata una tendenza a sottovalutare la presenza femminile attraverso l'uso di un linguaggio che, per abitudine, favoriva forme maschili, contribuendo così a perpetuare una certa invisibilità delle donne nei media e nella letteratura. Questa pratica ha anche facilitato la diffusione di stereotipi che riguardano esclusivamente un genere, con terminologie che tendevano a oscurare l'equivalente femminile, alimentando così una certa parzialità nella percezione dei ruoli di genere. Questi fenomeni sono stati classificati sotto il concetto di "genere sociale", come definito da Bazzanella (2010). Questo termine aggrega e descrive le varie manifestazioni di stereotipi e le assegnazioni

di ruoli sociali e culturali che sono tradizionalmente etichettati come femminili o maschili, basandosi su caratteristiche distinte. Tale classificazione è stata ulteriormente esplorata da Sartori (2009), e successivamente inclusa nella sintesi di Galeandro (2021), evidenziando come le distinzioni basate su genere possano influenzare profondamente sia il linguaggio che la comunicazione sociale. In questo contesto, la promozione di un linguaggio inclusivo non mira solamente a un maggiore bilanciamento nella rappresentazione dei generi, ma anche a una più accurata riflessione sulla natura e l'impatto delle parole usate quotidianamente. L'obiettivo è quello di sensibilizzare come il linguaggio non solo rifletta, ma anche modelli la realtà sociale, influenzando la percezione delle identità di genere e la loro rappresentazione.

Con l'emergere e la proliferazione dei media moderni, si è assistito a una crescente diffusione di immagini e messaggi che utilizzano un linguaggio che tende a privilegiare l'uso del maschile generico, tradizionalmente considerato neutro e capace di rappresentare indistintamente entrambi i generi. Questa pratica linguistica, che sceglie il genere maschile per esprimere concetti ritenuti universali, si radica in un'antica convenzione che associa maggiore autorevolezza e prestigio a tali termini. I sostantivi derivati, che includono termini legati a titoli, cariche e professioni di alto profilo come *ministro* o *presidente*, riflettono questa tendenza: nonostante la lingua italiana permetta la formazione di equivalenti femminili, la persistenza del maschile in questi contesti solleva questioni di rappresentazione.

La discussione si estende alla gerarchia implicita nei sostantivi che designano persone, dove le differenze di genere possono implicare differenze di percezione del prestigio. Al contrario, nei sostantivi comuni e/o astratti che non possiedono caratteristiche animate, come per esempio i termini *poltrona* e *sgabello*, il genere non incide sulla percezione del prestigio, che può essere influenzato da altri fattori (Giusti, 2021).

Il problema si è fatto più acuto quando le donne hanno iniziato a ricoprire ruoli tradizionalmente attribuiti agli uomini, portando alla luce la necessità

di un adeguamento del linguaggio che non era stato precedentemente considerato urgente. Il sistema linguistico e sociale, trovandosi di fronte a questi cambiamenti, spesso non ha dimostrato di essere completamente preparato ad integrare pienamente il genere femminile in tutte le sue espressioni professionali e di potere. Questa evoluzione dei ruoli ha sollevato la necessità di rivedere le modalità espressive per rendere il linguaggio più inclusivo e rappresentativo di una realtà sociale in continuo cambiamento.

Fino ad ora si è trattato il tema della femminilizzazione, ma negli ultimi anni si è visto l'emergere di una nuova questione, ossia la differenza tra sesso e identità di genere fino ad ora considerati termini abbastanza interscambiabili. Mentre il sesso di una persona può essere identificato alla nascita come maschile o femminile basandosi su caratteristiche biologiche (Ghigi, 2019), l'identità di genere di quella stessa persona potrebbe essere percepita internamente come maschile, femminile, una combinazione di entrambi o nessuno dei due, riflettendo una realtà più complessa e personale (Sartori, 2009). A questo punto, si è passati dalla necessità di femminilizzare la lingua italiana, al bisogno di neutralizzarla (neutralizzazione della lingua) al fine di includere ogni identità di genere (Galeandro, 2021). Le strategie si possono riassumere qui di seguito (Kenda, 2022):

- L'asterisco [\*] *Marco non è andat\* al cinema perché è stanc\**
- La chiocciola [@] *Marco non è andat@ al cinema perché è stanc@*
- La lettera [x] *Marco non è andatx al cinema perché è stancx*
- La lettera [u] *Marco non è andatu al cinema perché è stancu*
- Il trattino basso [\_] *Marco non è andat\_ al cinema perché è stanc\_*
- La schwa larga [ɜ] per il plurale *Marco e Lucia non sono andatɜ al cinema perché sono stanchɜ*
- La schwa corta [ə] per il singolare *Marco non è andatə al cinema perché è stancə.*

Il presente lavoro è stato svolto in collaborazione con la prof.ssa Laura Nadal Sanchis, che ha condotto lo studio e raccolto i dati e il cui contributo è stato fondamentale per l'analisi. In questa tesi si propongono diversi metodi per l'analisi dei dati relativi all'utilizzo di forme sperimentali volte a promuovere l'inclusività linguistica. L'obiettivo principale è comprendere come queste forme sperimentali influenzino i tempi di lettura e l'accuratezza delle risposte a un questionario. Attraverso l'applicazione di modelli statistici basati su approcci differenti, si intende esplorare le dinamiche che sottendono le variazioni nei tempi di risposta e nell'accuratezza delle risposte, fornendo una comprensione approfondita dell'impatto delle forme inclusive.

L'elaborato seguirà la seguente struttura: nel Capitolo 1 sono descritti due studi condotti per rispondere a questa domanda di ricerca e la struttura dei dati raccolti. Nel capitolo 2 è stata riportata la metodologia con cui si sono svolti gli studi, dando particolare attenzione sui modelli a effetti misti incrociati oltre a un approccio basato sulla modellazione tramite modelli CUB, ovvero metodi per l'analisi di dati ordinali. Nel Capitolo 3 verranno presentati i risultati nell'analisi dello studio Psychopy mentre nel Capitolo 4 quelli per lo studio dell'Eye-tracking. Nel Capitolo 5 verranno discussi alcuni limiti incontrati nelle analisi e infine, nel Capitolo 6, verrà fornita un'interpretazione dei risultati, cercando di dare una risposta alle domande formulate in precedenza.

# Capitolo 1

## Dati e presentazione degli studi

### 1.1 Indagine di percezione Studio Psychopy

Come tecnica applicata nel campo della pragmatica sperimentale, le indagini di percezione consentono di superare gli studi teorici basati sull'intuizione e l'introspezione del linguista e di verificare ipotesi sull'accettazione che determinate strutture discorsive generano tra parlanti di diversi background linguistici. Questo tipo di esperimento consiste nella realizzazione di ricerche di percezione, che, oltre a quantificare lo sforzo cognitivo durante la lettura di stimoli testuali, mirano ad analizzare l'effettiva comprensione di tali stimoli da parte dei parlanti. In particolare, gli studi condotti a complemento degli esperimenti sull'elaborazione informativa (tracciamento dei movimenti oculari) consentono di verificare se, indipendentemente dai costi di elaborazione, i parlanti danno effettivamente un'interpretazione inclusiva agli elementi sperimentali contenenti elementi contrassegnati con informazioni di genere o se, al contrario, attribuiscono loro un'interpretazione esclusivamente maschile (soprattutto nel caso del maschile generico) o se, in generale, hanno difficoltà ad assegnare qualsiasi tipo di interpretazione.

Come ogni forma di sperimentazione, si tratta di un metodo di osservazione che integra gli studi sui corpus, ovvero una collezione di testi selezionati e organizzati per facilitare le analisi linguistiche: mentre questi sistematizzano gli usi reali della lingua e descrivono come un sistema linguistico si plasma sul piano individuale della parola, gli studi sperimentali osservano le reazioni cognitive che suscitano gli stimoli linguistici (Loureda et al., 2019).

L'indagine di percezione è, inoltre, un compito che richiede al parlante l'applicazione di processi coscienti basati sulla sua competenza linguistica, comunicativa o culturale, poiché gli viene richiesto di determinare consapevolmente il tipo di rappresentazione mentale che una particolare forma linguistica suscita.

Nelle indagini di percezione sono stati utilizzati stimoli testuali corrispondenti a quattro condizioni sperimentali per analizzare la variabile marcatura di genere. Dunque, sono stati creati quattro tipi di item critici nei quali veniva modificata solo la desinenza o marchio di genere del primo sintagma nominale con funzione di soggetto e dell'aggettivo con funzione di complemento predicativo, come nel seguente esempio

- a. Gli alunni rimasero scioccati dai brutti voti. (maschile generico)
- b. L3 alun<sup>n</sup>3 rimasero scioccat<sup>3</sup> dai brutti voti. (schwa larga)
- c. L \* alun<sup>n</sup>\* rimasero scioccat\* dai brutti voti. (asterisco)
- d. Gli/Le alunni/e rimasero scioccati/e dai brutti voti. (sdoppiamento)

La presentazione di ciascuno degli elementi critici è seguita da una domanda che ha lo scopo di verificare l'interpretazione dei morfemi o grafie di genere da parte dei parlanti:

A chi può far riferimento “l3 alun<sup>n</sup>3” in questa frase?

Per rispondere alla domanda, i partecipanti ricevevano sempre quattro possibili opzioni tra cui scegliere:

- 1. A Mario, Giovanni ed altri uomini.
- 2. A Marco, Adriana ed altre persone.
- 3. A Lucia, Luisa ed altre donne.
- 4. Non lo so, nessuna delle precedenti.

La risposta (1) indica un'interpretazione escludente, ovvero esclusivamente maschile, del sintagma nominale che apre la frase; una risposta del tipo (2) evidenzia un'interpretazione inclusiva del marchio di genere; la risposta (3) indica un'inclinazione a un'interpretazione esclusivamente femminile; infine, la risposta (4) dà la possibilità di non optare per nessuna delle opzioni, al fine di non forzare la scelta di una determinata interpretazione.

## 1.2 Items sperimentali

Sono stati creati in totale 32 item sperimentali per implementare ciascuna delle quattro condizioni. Tutti gli items si adattano alla stessa struttura sintattica e contengono sempre parole ad alta o altissima frequenza al fine di ridurre l'effetto di possibili variabili di confusione (Keating e Jegerski, 2015). Inoltre, ogni elemento critico non si presenta isolato, ma seguito da un terzo segmento che fornisce maggior contesto ed evita che l'attenzione del partecipante si focalizzasse unicamente sull'item oggetto di analisi:

*I bambini rimasero impressionati dall'affettuoso pagliaccio. All'inizio pensavano fosse pauroso.*

Oltre agli item sperimentali, vengono introdotti filler items o frasi di riempimento in un rapporto 2:1, al fine di evitare l'effetto apprendimento e che i partecipanti arrivassero a riconoscere l'oggetto di studio. Le frasi di riempimento sono tematicamente legate alle frasi sperimentali, ogni elemento critico era preceduto e seguito da un filler, in modo che la sequenza di tre frasi raccontasse una breve storia, come nel seguente esempio:

FILLER

Il Ministero dell'Istruzione ha annunciato importanti cambiamenti nel sistema scolastico.

ITEM

I/Le maestri/e diventarono dubbiosi/e per la nuova normativa. Sembrava troppo rigida.

FILLER

Iniziare le lezioni alle 7:45 era troppo presto per la scuola elementare.

Gli stimoli vengono distribuiti in quattro liste sperimentali in modo contro-bilanciato secondo un sistema di quadrato latino. Questo sistema permette a tutti gli informatori di leggere più volte ogni condizione (8 volte nel presente esperimento), ma sempre replicate su argomenti diversi. In altre parole, chi legge la versione del maschile generico *gli alunni* non legge anche la condizione inclusiva *ls alunns*, ma riceve la grafia della schwa lunga, ad esempio nell'enunciato *ls maestrz*.

### 1.3 Partecipanti

I dati sono raccolti da 88 madrelingua italiani. Tutti i partecipanti hanno terminato il ciclo di studi obbligatori. La maggior parte sono studenti dell'Università Ca' Foscari di Venezia e rientrano in una fascia di età compresa tra i 18 e i 23 anni. Si tratta di un gruppo misto di donne e uomini.

### 1.4 Tecnica e procedura

Le indagini sono state programmate utilizzando come software Psychopy, che consente di misurare i tempi di reazione e registrare le risposte ottenute da uno stimolo, e sono state distribuite online attraverso la piattaforma Pavlovia. L'acquisizione dei dati è stata effettuata in diverse sessioni presso i la-

boratori informatici dell'Università Ca' Foscari di Venezia. I partecipanti leggevano le istruzioni nelle prime due schermate programmate per l'esperimento. Successivamente, ricevevano due elementi di prova che consentivano loro di familiarizzare con la procedura e, infine, accedevano a una delle quattro liste sperimentali composte da elementi critici e distrattori che erano stati assegnati casualmente. Le frasi sperimentali erano presentate in un ordine pseudo-randomizzato, cioè ogni elemento critico appariva preceduto e seguito dall'enunciato di riempimento corrispondente e si evitava che la stessa condizione sperimentale venisse visualizzata due volte di seguito, altrimenti gli elementi venivano riprodotti ogni volta in un ordine diverso.

## 1.5 Dataset finale

Il dataset si compone di un totale di 2816 osservazioni, il che corrisponde a 32 osservazioni per ogni partecipante coinvolto nello studio. Per ciascuna di queste osservazioni, come descritto precedentemente, sono stati registrati diversi elementi: l'item specifico che è stato analizzato, accompagnato dalle due frasi definite filler. Inoltre, oltre a queste informazioni, il dataset dispone delle seguenti variabili:

- il tema dell'item, variabile categoriale a 32 livelli;
- la condizione sperimentale analizzata, variabile categoriale con 4 livelli;
- la risposta scelta, variabile categoriale rappresentante la risposta scelta dal soggetto;
- la risposta corretta, variabile categoriale rappresentante la posizione della risposta inclusiva;
- i tempi di lettura dell'item sperimentale oltre che i tempi di lettura dei filler, tutte variabili continue rappresentanti il tempo di lettura in millisecondi.

## 1.6 Indagine di percezione Studio Eyetracking

Noveck (2018) definisce la pragmatica come "lo studio di come le proprietà linguistiche e i fattori contestuali interagiscono nell'interpretazione degli enunciati". La pragmatica non tratta il linguaggio da un punto di vista grammaticale o di spiegazione di regole che tutti i parlanti devono rispettare, ma considera il rapporto che quest'ultima ha con gli utenti, il contesto e le circostanze comunicative. Quando si parla di pragmatica sperimentale ci sono diversi metodi, tecniche e test sperimentali di carattere psicologico che vengono utilizzati per studiare un atto di comunicazione linguistica, in particolare per studiare i prodotti e i processi di comprensione (Irrazábal e Molinari Marotto, 2004). L'eyetracking è uno di questi metodi. Non è una tecnica offline o ritardata, cioè che fornisce dati sul "risultato di un processo di comunicazione" come lo studio precedente di Psychopy. L'eyetracking appartiene al gruppo dei metodi online, ovvero quelli che "forniscono un accesso ai processi cognitivi immediato o con un ritardo temporale minimo" (Loureda et al, 2016). I metodi online consentono di studiare in tempo reale processi neurali e mentali altamente automatizzati. I dati ottenuti possono fornire un'interpretazione dell'attività e delle competenze coinvolte nella costruzione di una rappresentazione mentale che è stata comunicata (Keating e Jegerski, 2015). Questi metodi e tecniche sperimentali si basano sul cosiddetto occhio-mente (Just e Carpenter, 1980); cioè, esiste una stretta relazione tra il sistema percettivo, cioè l'occhio e lo sguardo, e il sistema cognitivo, cioè la mente. La mente elabora l'area di interesse in cui l'occhio si fissa in quel momento, generalmente senza indugio tra lo sguardo e il processo di comprensione (Irrazábal e Molinari Marotto, 2004). In altre parole, registrare il movimento oculare consente di riflettere l'elaborazione testuale e semantica. A questo assunto Just e Carpenter ne aggiungono un altro, quello dell'immediatezza, cioè, man mano che si affronta una parola del testo proposto, il lettore la interpreta tenendo conto che a volte i tentativi possono rivelarsi infruttuosi perché ci sono diversi

passaggi che la nostra mente mette in atto: "decodificare la parola, scegliere un solo significato, attribuirlo al suo referente e determinarne il ruolo nella frase e nel discorso" (Daneman e Carpenter, 1980).

## 1.7 L'Eye-tracking

L'Eye-tracking esiste da cento anni e rappresenta uno strumento prezioso per monitorare e registrare il movimento oculare e lo sguardo con forme di varie misure. Come affermano diversi autori (Rayner, 1998; Richardson e Dale, 2005; Holmqvist et al., 2011), esistono circa 120 tipi di misurazioni dei movimenti oculari da cui ne emergono quattro principali che variano a seconda della domanda di ricerca: a) misure di movimento, b) misure di posizione, c) misure numeriche e d) misure di latenza. Per questo esperimento vengono considerate le misure numeriche, soprattutto due: i movimenti saccadici e le fissazioni. Le prime si riferiscono ai movimenti degli occhi tra le parole da un punto all'altro del testo, cioè i "salti" (che possono comprendere da sette a nove lettere), che si alternano a "pause", le fissazioni, cioè i punti in cui l'occhio non si muove e si fissa su una parola. Può esserci molta variabilità durante un movimento saccadico ed un fissaggio, ma di solito durano circa 200-250 millisecondi (ms) (Holmqvist et al., 2011). Questa variabilità può portare a tempi di elaborazione maggiori per fissaggi più lunghi o a un numero maggiore di questi (movimenti di progressione e regressione) a seconda della difficoltà del testo (Rayner, Campbell and Bovee, 2014). Inoltre, va sottolineato che durante questi salti non si interrompe il sistema generale di comprensione e allo stesso tempo non vengono elaborate nuove informazioni (Duchowski, 2007). Si rivela quindi un sistema molto valido per elaborare dati su tutto il testo, ogni frase, ogni parola o sintagma.

## 1.8 Items sperimentali

Per l'esperimento relativo all'eyetracking, si è dapprima partiti con la creazione degli item che contenessero nomi senza alcune incongruenze come l'esclusione di nomi tradizionalmente maschili, evitare parole con ambiguità di genere o l'uso di termini neutri. Sono stati quindi presi nomi che non presentavano queste incongruenze. Per controllare il disegno sperimentale sono state utilizzate parole ad alta frequenza, senza polisemia o ambiguità sintattiche, parole invariabili e sono state evitate quelle che avevano un cambiamento ortografico per arrivare ad un totale di 32 elementi. In secondo luogo, è stata analizzata la struttura sintattica tenendo conto delle regole della lingua italiana per non rendere gli elementi troppo artificiosi. Ogni item presentava il seguente ordine sintattico: Articolo + nome / Verbo / Aggettivo predicativo / Complemento indiretto.

Al termine delle osservazioni preliminari, sono stati creati 32 gruppi di 4 frasi ciascuno con le 4 condizioni sperimentali definite secondo Arunachalam (2013). Le condizioni sono:

- a per il maschile generico,
- b per la schwa lunga,
- c per l'asterisco,
- d per lo sdoppiamento.

Viene di seguito riportato un esempio:

- a. Gli ebrei divennero gioiosi per l'attesa celebrazione.
- b. L<sub>3</sub> ebrez divennero gioios<sub>3</sub> per l'attesa celebrazione.
- c L\* ebre\* divennero gioios\* per l'attesa celebrazione.
- Gli/Le ebrei/e divennero gioiosi/e per l'attesa celebrazione.

Per evitare che i partecipanti si rendessero conto dell'obiettivo e delle finalità

della ricerca e che si presentasse l'effetto di ripetizione nel leggere gli elementi critici uno dopo l'altro, è stato indispensabile aggiungere agli stimoli critici (gli item sperimentali) degli stimoli di riempimento, ovvero i filler (Keating e Jegerski, 2015), che per non creare bias nella lettura, hanno più o meno la stessa lunghezza degli elementi sperimentali e non si pongono con la forma interrogativa.

In aggiunta a quanto detto finora, bisogna considerare la posizione delle parole chiave o più importanti, ovvero quelle che hanno il linguaggio inclusivo all'interno del testo. Queste non devono essere alla fine per evitare il wrap-up effect che si manifesta con la tendenza a leggere più velocemente una porzione di testo per poi rallentare nella frase finale (Raney et al. 2016). Inoltre, sia le parole che le frasi chiave devono apparire nelle stesse posizioni e bisogna bilanciare la frequenza e la lunghezza delle parole poiché sono fattori che influenzano il numero di fissazioni e regressioni e i tempi di lettura.

## 1.9 Tecnica e procedura

Lo studio è stato condotto utilizzando come software l'eyetracker EyeLink 1000 plus, Experiment Builder di SR Research EyeLink. È un software di programmazione grafica drag-and-drop per creare esperimenti di psicologia e neuroscienza da implementare su computer. Dopo alcune istruzioni viene inserito un esempio per abituare l'occhio del lettore a ciò che apparirà di seguito e familiarizzare con la metodologia di ricerca. Viene lasciata l'impostazione che permette al partecipante di decidere quando far iniziare l'esperimento in quanto si tratta di una lettura auto controllata senza tempo definito per la sua esecuzione. Viene quindi rilasciata una versione di prova del trial per verificare che il testo sia visualizzato correttamente sullo schermo, assicurando così la validità e l'affidabilità del trial stesso.

Una volta organizzate queste prime sezioni si può proseguire con la fase più importante, quella di implementazione delle liste sperimentali nel trial.

Successivamente, si procede con la creazione di quattro diverse versioni dello stesso esperimento, che vengono chiamate "deploy". In questo passaggio, gli stimoli critici sono distribuiti in quattro liste distinte. Per ogni lista, viene scelto un identificativo che riassume il tema principale di quella lista. Nella colonna "text" di ogni lista, vengono inseriti gli stimoli critici insieme ai loro elementi di riempimento, i "filler". Gli stimoli nelle liste sono poi disposti in ordine casuale per evitare che un ordine fisso possa influenzare i risultati del trial e per bilanciare l'attenzione dei partecipanti che tende a essere più alta all'inizio e diminuire verso la fine a causa dell'effetto della stanchezza.

## 1.10 Partecipanti

I dati sono stati presi da 88 soggetti, molti dei quali studenti di Ca' Foscari. Tutti i soggetti attraversano la stessa situazione sperimentale, ma sono sottoposti a condizioni di stimolazione diverse. I partecipanti devono rispettare alcuni requisiti: avere l'italiano come lingua madre, entrare nella fascia di età 18-40, aver completato gli studi secondari, non avere lesioni o disturbi oculari e non essere dislessici.

## 1.11 Dataset finale

Il dataset si compone di un totale di 8448 osservazioni, il che corrisponde a 96 osservazioni per ogni partecipante coinvolto nello studio. Queste osservazioni sono state suddivise in base ai 32 item analizzati, con 3 osservazioni specifiche per ciascun tema. La prima osservazione analizza la frase intera, la seconda si concentra esclusivamente sulle prime due parole della frase, mentre la terza osservazione considera solamente la quinta parola (queste posizioni indicano le parole definite con un linguaggio inclusivo). Per ciascuna di queste osservazioni sono stati registrati diversi elementi: l'item specifico che è stato analizzato, accompagnato dalle due frasi definite filler. Inoltre, oltre a queste informazioni, il dataset dispone delle seguenti variabili:

- il tema dell'item, variabile categoriale con 32 livelli;
- l'area d'interesse in ogni item, variabile categoriale a livelli:
  - Enunciado (frase completa);
  - Género 1 (le due prime parole della frase);
  - Género 2 (la quinta parola della frase);
- la condizione sperimentale analizzata, variabile categoriale con 4 livelli;
- il tempo di lettura totale, variabile continua in cui viene definito il tempo di lettura dell'intero item in millisecondi;
- il tempo della prima fase di lettura, variabile continua rappresentante il tempo della prima fase di lettura in millisecondi;
- il tempo di rilettura, variabile continua rappresentante l'eventuale tempo di rilettura dell'item in millisecondi;
- il numero di caratteri, variabile numerica in cui viene indicato il numero di caratteri (spazi esclusi) dell'item analizzato;
- il numero di parole dell'item;
- variabili demografiche quali sesso ed età.

# Capitolo 2

## Modelli

In questo capitolo verranno presentati i modelli che verranno utilizzati per l'analisi. Tra questi, un'attenzione particolare sarà dedicata ai modelli a effetti misti, i quali incorporano effetti casuali che sono fondamentali per gestire la variabilità intra-gruppo. Questa caratteristica permette di modellare efficacemente la correlazione tra le osservazioni all'interno degli stessi gruppi o cluster, assicurando un'analisi più accurata e rappresentativa delle dinamiche di gruppo. Verranno trattati inizialmente i modelli lineari a effetti misti, esaminando come questi modelli generalizzino i modelli lineari semplici per includere sia effetti fissi che casuali e per cui verrà esposta la procedura di stima. Successivamente, verranno approfonditi i modelli logistici a effetti misti, utilizzati per dati con risposta binaria. Verranno discussi in seguito i modelli logit multinomiali per risposte con molteplici categorie. Il capitolo proseguirà trattando il modello di Cox a effetti misti, una estensione del modello di sopravvivenza di Cox che include effetti casuali per gestire la presenza di rischi proporzionali non osservati che possono variare tra i soggetti.

È importante sottolineare che, data la natura dei dati trattati negli studi esaminati, i modelli presentati nel capitolo, almeno inizialmente, adotteranno una struttura di effetti casuali incrociati. Questa scelta metodologica risponde alla necessità di modellare accuratamente la complessità e la correlazione intrinseca tra i dati, che sono tipicamente presenti in studi che coinvolgono più fattori o livelli di raggruppamento. L'impiego degli effetti casuali incrociati consente di tenere conto contemporaneamente di diverse fonti di variabilità. Infine, verrà introdotta una classe di modelli per risposte ordinali, ovvero i

modelli CUB, e una loro estensione, i modelli CUBE, non tenendo però conto di effetti casuali.

## 2.1 Modelli a effetti misti incrociati

In questo capitolo verranno presentati i modelli a effetti misti incrociati, che rappresentano un'estensione significativa dei modelli multilivello tradizionali. I modelli a effetti misti incrociati offrono una flessibilità notevole in quanto permettono la generalizzazione su più tipi di unità, come diversi partecipanti e elementi di test. Questo approccio non solo gestisce le variazioni intra-gruppo ma modella anche la correlazione tra le osservazioni all'interno degli stessi gruppi o cluster. I vantaggi di adottare questi modelli non si limitano alla flessibilità nel trattamento degli effetti casuali incrociati, ma includono anche una maggiore robustezza contro i dati mancanti e la capacità di integrare variabili predittive sia discrete che continue.

Proseguendo, si esaminerà come la modellazione a effetti misti permette di incorporare direttamente nel modello statistico gli effetti che si manifestano durante il corso di un esperimento e di considerare altre covariate potenzialmente rilevanti. Questo approccio è particolarmente prezioso in contesti come gli studi cronometrici, dove i fattori temporali come l'apprendimento o la fatica possono influenzare significativamente i risultati. Negli studi cronometrici, è comune osservare variazioni nelle prestazioni dei soggetti dovute all'apprendimento o alla fatica. I modelli a effetti misti consentono di modellare questi cambiamenti longitudinali, riducendo la contaminazione del rumore sperimentale e migliorando l'accuratezza nell'identificazione dei predittori critici.

La risposta a una prova target può essere fortemente influenzata dalle prove precedenti. In contesti come la decisione lessicale, la reazione a parole precedenti può prevedere significativamente la latenza di risposta successiva. Attraverso l'uso di effetti misti, è possibile controllare statisticamente queste dipendenze, isolando meglio gli effetti di interesse.

Anche le caratteristiche qualitative dei trial precedenti, come la correttezza delle risposte o la categoria grammaticale delle parole, possono influenzare le performance nei trial successivi. I modelli a effetti misti permettono di incorporare questi dettagli nel modello statistico, fornendo una comprensione più profonda delle dinamiche sperimentali. I vantaggi generali dei modelli a effetti misti si applicano anche ai modelli con effetti casuali incrociati, tra cui l'assenza di ipotesi di omoschedasticità e la robustezza contro i dati mancanti. In questo quadro, l'uso di modelli a effetti misti incrociati assume un ruolo cruciale, consentendo di valutare l'interazione tra i fattori fissi e gli effetti casuali incrociati in modo più dettagliato, distinguendosi per la loro capacità di catturare la variabilità intra-individuale e inter-individuale in risposta agli stessi stimoli. Questa analisi dettagliata è resa possibile poiché, nei modelli a effetti misti incrociati, gli effetti casuali per soggetti ed elementi sono trattati come provenienti da distribuzioni normali indipendenti, consentendo di modellare le correlazioni intra-classe senza presupporre una struttura gerarchica annidata. Specificamente, se consideriamo un disegno sperimentale in cui ogni soggetto è esposto allo stesso set di elementi, i modelli a effetti misti incrociati utilizzano matrici di varianza-covarianza per gli effetti casuali per quantificare e gestire la correlazione tra le risposte.

Negli esperimenti controllati, come definito in Quené e Van den Bergh (2008), le osservazioni ripetute sono generalmente condotte utilizzando lo stesso campione di elementi di test per tutti i partecipanti. Questo approccio sperimentale può limitare la generalizzazione dei risultati solo agli elementi di test selezionati, poiché non è possibile estendere con sicurezza i risultati ad altri possibili elementi. Inoltre, i residui risultano correlati tra loro, indicando che questi errori non sono indipendenti, il che evidenzia la presenza di due effetti casuali incrociati e non annidati nel disegno sperimentale.

Data la complessità derivante da questi limiti, è essenziale adottare un approccio statistico più avanzato. I modelli multilivello tradizionali possono essere estesi per includere più effetti casuali incrociati e non annidati, come

suggerito da Goldstein (1999). Tali modelli con effetti casuali incrociati sono ideali per analizzare i risultati di esperimenti controllati, perché consentono la generalizzazione simultanea e congiunta ad altri partecipanti e ad altri elementi del test, derivante dall'inclusione simultanea di entrambi i fattori casuali in un'unica analisi. Di seguito verranno introdotti i modelli a effetti misti utilizzati per l'analisi.

### 2.1.1 Modello lineare a effetti misti

In questa sezione verranno analizzati i modelli lineari a effetti misti, come definiti in Diggle (2002). Questi modelli presuppongono l'esistenza di caratteristiche latenti, comuni a tutte le osservazioni di una stessa unità statistica. Tali caratteristiche sono modellate come realizzazioni di variabili casuali, note come effetti casuali. La presenza di questi effetti nel modello implica una correlazione tra le osservazioni all'interno della stessa unità, permettendo così di rappresentare la dipendenza tra le diverse misurazioni.

Viene indicato con  $\mathbf{Y}$  il vettore  $N$ -dimensionale di tutte le misurazioni di tutte le unità statistiche, il modello può essere espresso in forma matriciale come

$$Y = X\beta + \sum_{j=1}^r Z_j\gamma_j + \varepsilon$$

dove:

- $X$  è la matrice di disegno degli effetti fissi, con dimensione  $Np$ ;
- $\beta=(\beta_1, \dots, \beta_p)$  rappresenta il vettore  $p$ -dimensionale degli effetti fissi, il quale definisce la relazione media fra la risposta e l'insieme di variabili esplicative a livello di popolazione;
- $Z_j$  sono le matrici di disegno degli effetti casuali;
- $\gamma_j=(\gamma_{j1}, \dots, \gamma_{jN})$  sono i vettori aleatori  $n$ -dimensionali contenenti le variabili casuali per ogni unità statistica relative al  $j$ -esimo effetto casuale,
- $\varepsilon=(\varepsilon_1, \dots, \varepsilon_n)$  è il vettore casuale per il termine d'errore.

L'eterogeneità osservata tra le diverse unità statistiche presenti nel campione è spiegata attraverso l'inclusione degli effetti casuali nel modello. Questi effetti casuali sono fondamentali per rappresentare la variazione individuale che si manifesta tra le unità, consentendo di distinguere le caratteristiche specifiche di ciascuna unità rispetto agli effetti fissi, che invece si applicano uniformemente a tutto il campione. In questo modo, il modello può adeguatamente riflettere sia la diversità intrinseca presente nel gruppo di dati sia le tendenze comuni a tutte le osservazioni. Risulta quindi ragionevole supporre che gli effetti casuali si configurino come vettori aleatori con una media che si assume essere nulla.

Viene definita la matrice  $\mathbf{Z}=(Z_1, \dots, Z_r)$ , matrice di disegno per tutti gli effetti casuali ottenuta affiancando le matrici  $Z_j$  e  $\gamma=(\gamma_1, \dots, \gamma_r)$  il vettore ottenuto unendo ogni vettore degli effetti casuali. Definite queste matrici, è possibile ridefinire il modello nella seguente forma:

$$Y = X\beta + Z\gamma + \varepsilon \quad (2.1.1)$$

È evidente che i modelli con effetti casuali ampliano i tradizionali modelli di regressione lineare aggiungendo una componente,  $Z\gamma$ , al predittore lineare. Di conseguenza, la variabile risposta è composta dalla somma di effetti fissi e di effetti casuali, arricchendo così il modello con una maggiore flessibilità per catturare le variazioni sia sistematiche che aleatorie tra le osservazioni.

La specificazione più semplice del modello prevede che  $\gamma_j \sim N(0, \sigma^2 D)$  e  $\varepsilon \sim N(0, \sigma^2 I)$ . La formulazione presuppone l'indipendenza tra gli effetti casuali e gli errori casuali, nonché l'indipendenza tra gli effetti casuali associati a diverse unità e quelli relativi alla stessa unità, e tra gli errori casuali associati a diverse unità e quelli relativi alla stessa unità. Definendo il vettore  $Y_i$  contenente le osservazione dell'individuo  $i$ , nel caso più semplice è possibile supporre che  $Y_i \sim N(X_i\beta, \Sigma_i)$  indipendenti, dove  $X_i$  rappresenta la matrice di disegno dell' $i$ -esimo individuo mentre  $\Sigma_i$ , matrice di varianze e covarianze

e definita pari a

$$\Sigma_i = \sigma_\varepsilon^2 I + \sum_{j=1}^r \sigma_j^2 [Z_j Z_j^T]_{ii},$$

dove l'elemento  $[Z_j Z_j^T]_{ii}$  indica il blocco diagonale  $i$ -esimo della matrice  $Z_j Z_j^T$ .

A questo punto è possibile definire la matrice di varianze e covarianze del vettore  $Y$  definito in precedenza

$$V = \sigma_\varepsilon^2 I_N + \sum_{j=1}^r \sigma_j^2 Z_j Z_j^T,$$

la quale è rappresentata da una matrice diagonale a blocchi, visto l'assunto di indipendenza, avente sulla diagonale le matrici  $\Sigma_i$ .

## 2.1.2 Procedura di stima tramite massima verosimiglianza

Il vettore ignoto dei parametri viene definito come  $\theta = (\beta, \alpha)^T$ , in cui  $\beta$  è il vettore degli effetti fissi mentre  $\alpha = (\sigma_\varepsilon^2, \sigma_1^2, \dots, \sigma_r^2)$  il vettore relativo alla varianza dell'errore casuale oltre che alle varianze degli effetti casuali.

La procedura di stima classica si basa sulla massimizzazione della verosimiglianza, che per  $\theta$  è pari a

$$L(\theta, y) \propto \prod_{i=1}^n |\Sigma_i|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)^T \Sigma_i^{-1} (y_i - X_i \beta) \right\}.$$

La stima di massima verosimiglianza  $\hat{\theta}$  si ottiene massimizzando la precedente equazione, oppure risolvendo in  $\beta$  l'espressione

$$\sum_{i=1}^n X_i^T \Sigma_i^{-1} (y_i - X_i \beta) = 0$$

per il vettore degli effetti fissi e in  $\sigma_j^2$  l'espressione

$$\sum_{i=1}^n (y_i - X_i \beta)^T \Sigma_i^{-1} [Z_j Z_j^T]_{ii} \Sigma_i^{-1} (y_i - X_i \beta) - \text{tr}(\Sigma_i^{-1} [Z_j Z_j^T]_{(ii)}) = 0 \quad (2.1.2)$$

per i parametri del vettore  $\alpha$ .

Per risolvere il sistema di queste due equazioni, assumendo che le varianze

degli effetti casuali siano fissate, si ottiene la soluzione per  $\beta$  in funzione di  $\alpha$  pari a

$$\hat{\beta}_\alpha = \left( \sum_{i=1}^n X_i^T \Sigma_i^{-1} X_i \right)^{-1} \sum_{i=1}^n X_i^T \Sigma_i^{-1} y_i = (X^T V^{-1} X)^{-1} X^T V^{-1} y \quad (2.1.3)$$

Dall'ultima equazione è possibile notare come, se la matrice  $V$  fosse nota, si otterrebbe lo stimatore di massima verosimiglianza per la componente degli effetti fissi che risulta pari allo stimatore dei minimi quadrati generalizzati. Nel caso in cui la matrice  $V$  non sia nota, è possibile ricavare la stima del vettore  $\alpha$  (2.1.2) sostituendo a  $\beta$  l'espressione 2.1.3. Il risultato del sistema di equazione che si ottiene restituisce la stima di massima verosimiglianza  $\hat{\alpha}$ .

Una volta ottenuta questa, grazie alla proprietà di equivarianza dello stimatore di massima verosimiglianza, si procede con la stima delle matrici  $\Sigma_i$  e di conseguenza della matrice  $V$ , indicate rispettivamente come  $\hat{\Sigma}(\hat{\alpha})$  e  $\hat{V}(\hat{\alpha})$ . Lo stimatore  $\hat{\beta}$  risulta avere distribuzione normale oltre che a essere non distorto e, per il teorema di Gauss-Markov, è lo stimatore a varianza minima. La matrice che rappresenta le varianze e le covarianze dello stimatore di massima verosimiglianza per gli effetti fissi è uguale a:

$$Var(\hat{\beta}) = \left( \sum_{i=1}^n X_i^T \Sigma_i^{-1} X_i \right)^{-1} = (X^T V^{-1} X)^{-1},$$

la quale può essere stimata sostituendo le stime alle rispettive matrici.

### 2.1.3 Procedura di stima tramite massima verosimiglianza ristretta

Nonostante lo stimatore di massima verosimiglianza sia consistente ed efficiente sotto l'ipotesi di normalità, gli stimatori delle varianze  $\sigma_j^2$  sono solamente asintoticamente non distorti. Si predilige infatti l'uso di uno stimatore corretto basato sulla verosimiglianza marginale rispetto al semplice stimatore di massima verosimiglianza. Si tratta della verosimiglianza ottenuta a partire dal modello per una trasformazione  $AY$  di  $Y$ , con densità non dipendente da  $\beta$ . Questa trasformazione fornisce i residui linearmente indipendenti della

regressione lineare di  $Y$  su  $X$ . Per qualsiasi matrice  $A$ , la log-verosimiglianza marginale per  $\alpha$  è

$$l_r(\alpha; y) \propto -\frac{1}{2}(Ay)^T V^{-1}(Ay) + \log|V| + \log|X^T V^{-1} X|.$$

La stima REML (Restricted maximum likelihood) del vettore per  $\alpha$  è calcolabile massimando la funzione di log-verosimiglianza oppure ricercando la soluzione che annulla il vettore delle derivate parziali prime.

Gli stimatori REML per gli effetti fissi, denotati come  $\hat{\beta}_R$  derivano dalla 2.1.3 calcolando  $\hat{\beta}_{\alpha R}$ . Gli stimatori  $\hat{\beta}_R$  e  $\hat{\alpha}_R$  sono asintoticamente normali e incorrelati. La varianza invece dello stimatore  $\hat{\beta}_R$  segue la stessa forma di quello dello stimatore di massima verosimiglianza, in cui la matrice di varianze e covarianze  $\Sigma_i$  viene stimata con  $\hat{\Sigma}_i(\hat{\alpha}_R)$  invece che  $\hat{\Sigma}_i(\hat{\alpha})$ .

## 2.1.4 Modello logistico a effetti misti

I modelli logistici a effetti misti costituiscono un'estensione dei modelli logistici standard, essendo particolarmente efficaci nell'analisi di dati con strutture gerarchiche. Questi modelli si distinguono per l'integrazione di effetti fissi e casuali, permettendo una modellazione accurata delle variazioni intra e inter-gruppo. Questa configurazione conferisce ai modelli a effetti misti una robustezza particolare nell'assegnare correttamente la varianza relativa a ogni livello di aggregazione dei dati.

Un modello logistico a effetti misti stima il logaritmo delle probabilità del risultato su vari predittori per stimare la probabilità che  $Y = 1$  si verifichi. La forma generale di un modello logistico a effetti misti ha la seguente struttura:

$$\text{logit}\left(E(Y_{ij} = \pi_{ij} = 1|X_j, u_j)\right) = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = X\beta + Z_j u_j$$

dove  $\beta$  è il vettore degli effetti fissi,  $Z_j$  sono le matrici di disegno degli effetti casuali,  $u_j$  sono gli effetti casuali specifici per ogni soggetto e si assume che seguano distribuzioni normali tali per cui  $u_j \sim N(0, \sigma_u^2)$ .

Nei modelli logistici a effetti misti per dati binari, la probabilità dell'outcome  $y_j$ , condizionata dagli effetti specifici del cluster  $u_j$ , può essere formulata come

$$P(y_j = 1|u_j) = \frac{\exp(X_j\beta + Z_ju_j)}{1 + \exp(X_j\beta + Z_ju_j)}$$

oppure come

$$P(y_j = 0|u_j) = \frac{1}{1 + \exp(X_j\beta + Z_ju_j)}.$$

Sotto l'assunzione comune che le risposte in ogni cluster siano indipendenti dopo aver condizionato sugli effetti casuali, la probabilità condizionata di  $y_j$  può essere scritta come:

$$f(y_j|u_j; \beta) = \prod_{i=1}^{n_j} P(y_{ij} = 1|u_j)^{y_{ij}} P(y_{ij} = 0|u_j)^{1-y_{ij}}.$$

Poiché gli effetti casuali sono non osservati, le inferenze per gli effetti fissi  $\beta$  e la matrice di covarianza degli effetti casuali  $G$  sono stimate integrando sugli effetti casuali,  $u_j$ . Si ottiene quindi la funzione di verosimiglianza marginale, definita come:

$$L(\beta, G) = \prod_{j=1}^J \int f(y_j|u_j; \beta) r(u_j; G) du_j,$$

dove  $r(u_j; G)$  è una distribuzione di probabilità di  $u_j$ , che si presume abbia una distribuzione normale multivariata con vettore di media 0 e matrice di covarianza  $G$ . Quando più effetti casuali sono inclusi nel modello, l'integrale in precedenza definito diventerà multidimensionale, il che può aumentare la complessità computazionale del modello.

Tuttavia, a differenza dei modelli lineari a effetti misti, la verosimiglianza non ha una soluzione esplicita. Pertanto, sono richieste approssimazioni della verosimiglianza (o della log-verosimiglianza) del modello per ottenere le stime dei parametri. Tre tecniche comuni per i modelli logistici a effetti misti sono la quadratura gaussiana (adattativa), l'approssimazione di Laplace e la pseudo-verosimiglianza penalizzata. Una descrizione dettagliata di queste tecniche è disponibile nell'articolo di McNeish (2016).

## 2.1.5 Modello logistico multinomiale a effetti misti

I modelli logit multinomiali con effetti misti, definiti in Hartzel, Agresti e Caffo (2001) sono un caso particolare di un modello lineare generalizzato misto multivariato, come descritto da Tutz e Hennevogl (1996). Dapprima viene definito il modello lineare generalizzato misto multivariato, dopodiché si procede ad integrare il modello logit multinomiale con effetti casuali nel suo contesto.

Sia  $Y_{ij}$  un vettore di risposta per la  $j$ -esima osservazione dell'individuo  $i$ , con  $i = 1, \dots, n$ . Includendo gli effetti casuali  $u_i$  la distribuzione condizionata  $f(Y_{ij}|u_i)$  appartiene alla famiglia esponenziale multivariata con

$$\mu_{ij} = E(Y_{ij}|u_i) = h(n_{ij}), \text{ con } n_{ij} = Z_{ij}\beta + W_{ij}u_i,$$

dove  $\beta$  è il vettore degli effetti fissi,  $Z_{ij}$  e  $W_{ij}$  sono le matrici del modello per gli effetti fissi e casuali mentre la funzione  $h(n_{ij})$  è un vettore di funzioni di collegamento inverse. Si assume inoltre che gli  $u_i$  siano indipendenti dalla funzione di densità  $g(u)$ , come la normale multivariata.

Sia  $Y_{ij}^{(s)}$ ,  $s=1, \dots, n_{ij}$ , dove  $n_{ij}$  rappresentano gli esiti categorici per l'osservazione  $j$ -esima dell'individuo  $i$ . Per esprimere il modello multinomiale come un modello lineare generalizzato misto multivariato, occorre riformulare ogni risposta categorica come un vettore di risposta con dummy,  $Y_{ij}^{(s)} = Y_{ij1}^{(s)}, \dots, Y_{ij,(R-1)}^{(s)}$  dove  $Y_{ijr}^{(s)}$  vale

$$\begin{cases} 1 & \text{se } Y_{ij}^{(s)} = r, r = 1, \dots, R, \\ 0 & \text{altrimenti.} \end{cases}$$

Quindi per  $n_{ij}$  ripetizioni indipendenti,  $y_{ij} = \sum_{s=1}^{n_{ij}} y_{ij}^{(s)}$  si distribuisce come multinomiale con indici  $n_{ij}$  e parametri  $\pi'_{ij} = (\pi_{ij1}, \dots, \pi_{ij,(R-1)})$ .

Il modello multinomiale generale è definito dall'equazione precedente in termini del vettore di risposta  $y_{ij}$ . I casi specifici risultano specificando la funzione di collegamento inverso  $h(n_{ij})$  e le matrici del modello. In seguito si specifica la funzione per la definizione del modello logit multinomiale a effetti

misti, la quale è pari a

$$h_r(n_{ij}) = \frac{\exp(n_{ijr})}{1 + \sum_{l=1}^{R-1} \exp(n_{ijl})}, r = 1, \dots, R - 1.$$

Si assume, per la definizione della funzione di verosimiglianza, una normale multivariata con media 0 e matrice di varianza e covarianza  $\Sigma$  per  $u$ . La funzione di verosimiglianza risulta pari a

$$L(\beta, \Sigma) = \prod_{i=1}^n \int_R \cdots \int_R \left\{ \prod_{j=1}^{T_i} f(y_{ij}|u_i; \beta) \right\} g(u_i; \Sigma) du_i.$$

Per risolvere questi integrali, vengono utilizzati diversi approcci quali la quadratura adattiva di Gauss-Hermite. Una descrizione dettagliata di queste tecniche è disponibile nell'articolo di Hartzel, Agresti e Caffo (2001).

## 2.1.6 Modello di Cox a effetti misti

Il Modello di Cox a effetti misti rappresenta un'estensione del modello di Cox a rischi proporzionali, come definito in Therneau (2007). Il modello originale di Cox assume che il rischio di evento per un individuo sia una funzione esponenziale delle variabili esplicative fisse. Tuttavia, il modello standard di Cox non tiene conto dell'eterogeneità non osservata tra i soggetti, che può derivare da differenze genetiche, ambientali o da altre variabili non misurate. L'omissione di questi effetti casuali può portare a stime distorte dei coefficienti delle covariate e a conclusioni errate riguardo l'effetto delle variabili di interesse. Per superare questi limiti, il modello di Cox a effetti misti introduce componenti casuali nel modello, permettendo di controllare l'eterogeneità non osservata tra i soggetti. In questo modello esteso, si assume che il logaritmo del rischio di ciascun individuo non sia solo una funzione delle covariate fisse, ma includa anche termini casuali specifici per i soggetti o gruppi di soggetti, che catturano l'influenza di fattori non misurati. Questi effetti casuali sono generalmente modellati come variabili casuali normalmente distribuite, il che permette l'integrazione flessibile di varie strutture di correlazione tra i soggetti.

Per definire il modello, si introduce inizialmente la funzione di rischio,

$$\lambda_i(t) = \lambda_0(t)e^{X\beta + Zb},$$

dove  $\lambda_0$  è un tasso di rischio di base non specificato, la matrice  $X$  e il vettore dei parametri  $\beta$  rappresentano gli effetti fissi del modello mentre  $Z$  e  $b$  gli effetti casuali.

Per la definizione del modello si assume che gli effetti casuali seguano una distribuzione normale multivariata,  $b \sim N(0, \Sigma)$ , come proposto in Ripatti e Palmgren (2000). Da questo è possibile definire la verosimiglianza parziale del modello a effetti casuali, la quale viene definita come

$$L = \int_R PL(\beta, b) \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-b'\Sigma b/2] db. \quad (2.1.4)$$

L'Equazione 2.1.4 descrive un integrale multidimensionale complesso: la verosimiglianza parziale di Cox (PL) si presenta come un prodotto di rapporti e la variabile  $b$  ha una dimensione che cresce con l'ordine di  $n$ . Per facilitare i calcoli legati a questa verosimiglianza, si applica l'approssimazione di Laplace all'integrale, seguendo il metodo delineato da Ripatti e Palmgren (2000). Quest'approssimazione porta alla seguente equazione:

$$\exp[PPL(\beta, b)] = \exp[\log(PL(\beta, b) - b'\Sigma b/2)]$$

e sostituisce la verosimiglianza parziale penalizzata (PPL) con una serie di Taylor del secondo ordine centrata su  $(\hat{\beta}, \hat{b})$ . Questa verosimiglianza parziale penalizzata deriva dalla log-verosimiglianza parziale di un modello di Cox a rischi proporzionali standard, considerando sia  $b$  che  $\beta$  come covariate comuni, a cui si sottrae una penalizzazione quadratica su  $b$ . Ulteriori riorganizzazioni e l'identificazione dell'integrale residuo come una densità gaussiana multivariata conducono alla stima approssimata della log-verosimiglianza integrata

$$L = PPL(\hat{\beta}, \hat{b}) = -\frac{1}{2}(\log|\Sigma| + \log|H_{\hat{b}\hat{b}}|),$$

con  $H$  matrice delle derivate seconde della PPL e  $H_{\hat{b}\hat{b}}$  parte della matrice corrispondente ai coefficienti degli effetti casuali  $b$ . La massimizzazione della

log-verosimiglianza integrata rispetto a  $(\beta, b)$  porta a stime approssimate di massima verosimiglianza dei coefficienti di regressione, così come stime delle varianze degli effetti casuali.

Nell'ambito della regressione a rischi proporzionali di Cox, l'utilizzo di approssimazioni permette la stima efficace di modelli a effetti misti. Questo approccio è supportato dall'applicazione di varie tecniche numeriche, che facilitano il trattamento dei calcoli necessari per modelli che includono effetti casuali. Questo metodo consente di eseguire analisi di sopravvivenza più complesse, integrando gli effetti casuali per ottenere risultati più accurati e dettagliati.

## 2.2 Modelli per dati ordinali

Nell'ambito dello studio, l'impiego di modelli a effetti misti si è rivelato cruciale per analizzare le variazioni intra-individuali e inter-individuali nei dati. Questi modelli, grazie alla loro capacità di incorporare effetti casuali, consentono di trattare efficacemente la correlazione tra osservazioni all'interno di cluster. Nonostante l'utilità dei modelli a effetti misti, si è reso necessario estendere l'analisi attraverso l'utilizzo di modelli CUB. Questa scelta deriva dall'esigenza di indagare con maggiore precisione le componenti di preferenza e incertezza manifestate nelle risposte dei soggetti. I modelli CUB, specializzati nell'isolare e quantificare tali componenti, si configurano come strumenti analitici particolarmente idonei per decomporre le dinamiche sottostanti alle scelte dei rispondenti.

L'integrazione dei modelli CUB, affiancati ai modelli a effetti misti, costituisce un avanzamento significativo dell'analisi, poiché permette non soltanto di esaminare l'influenza di variabili contestuali e caratteristiche individuali sulle risposte, ma anche di valutare come le predisposizioni personali e l'incertezza intrinseca influenzino le decisioni dei partecipanti di fronte a stimoli determinati.

### 2.2.1 Modelli CUB

I modelli CUB sono stati introdotti nella letteratura per analizzare dati ordinali e rientrano nel quadro delle variabili latenti, poiché le scelte discrete risultano da variabili non osservabili. I CUB appartengono alla classe dei modelli di mistura, poiché le scelte discrete osservate, tradotte in dati di valutazione o classificazione, sono spiegate dalla combinazione di due principali componenti latenti, ovvero il sentimento e l'incertezza. Questi si riferiscono al sentimento personale verso l'oggetto e all'incertezza nella scelta della risposta. Formalmente, la valutazione osservata è l'esito di una variabile casuale discreta, che è definita come una mistura di una variabile casuale binomiale traslata e una uniforme. L'uso della binomiale traslata come componente per modellare il processo di conteggio in una scelta binaria tra categorie ordinate è stato proposto in precedenza da D'Elia nel 2000.

Formalmente avremo:

$$P(R = r|\theta) = \pi b_r(\zeta) + (1 - \pi)P(U = r), r = 1, \dots, m$$

con  $\theta = (\pi, \zeta)'$ ,  $\zeta \in [0, 1]$ ,  $\pi \in (0, 1]$ ,  $b_r(\zeta) = P(V = r) = \binom{m-1}{r-1} \zeta^{m-r} (1-\zeta)^{r-1}$ , dove  $V$  rappresenta la densità di una variabile casuale binomiale traslata e  $P(U = r) = \frac{1}{m}$  è la distribuzione di probabilità di una variabile casuale uniforme su supporto  $1, \dots, m$ . La variabile  $V$  viene introdotta per modellare la componente emotiva ( $1 - \zeta$  viene definito come parametro di sentimento), mentre  $U$  ha lo scopo di misurare la componente di incertezza ( $1 - \pi$  è definito come parametro di incertezza). Iannario (2010) ha confermato che i modelli CUB sono identificabili per  $m > 3$ .

Le questioni inferenziali legate ai modelli CUB sono state trattate utilizzando metodi di massima verosimiglianza. In particolare viene utilizzato l'algoritmo EM per massimizzare la funzione di log-verosimiglianza specifica per questi modelli, la quale assume la seguente forma:

$$L(\theta) = \sum_{r=1}^m n_r \log P(R = r|\theta).$$

Un limite dei modelli CUB è che questi non sono adeguati per affrontare una possibile variabilità tra le risposte, dato che la posizione e la variabilità di una variabile casuale binomiale sono fortemente e strettamente correlate.

## 2.2.2 Modelli CUBE e IHG CUB

Per ovviare il problema dell'elevata variabilità osservata nelle risposte quando si utilizzano i modelli CUB, è stata avanzata un'estensione parsimoniosa denominata modelli CUBE (Combination of a Uniform and a (shifted) BEta-binomial), introdotta da Iannario nel 2014. Questa proposta prevede la sostituzione della Binomiale traslata con una distribuzione Beta-Binomiale. Tale adattamento si rivela essenziale per affrontare la presenza di sovradisersione nei dati ordinali, fenomeno che può essere ricondotto alla variabilità intrinseca nei sentimenti personali dei rispondenti, come osservato da Jansen nel 1990.

La densità per la distribuzione Beta-Binomiale  $g_r(\zeta, \phi)$ , definita su un supporto da 1 a  $m$ , illustra come questa estensione modelli la probabilità per ogni categoria di risposta  $r$ :

$$g_r(\zeta, \phi) = \binom{m-1}{r-1} \frac{\prod_{k=1}^r (1 - \zeta + \phi(k-1)) \prod_{k=1}^{m-r+1} (\zeta + \phi(r-1))}{(1 - \zeta + \phi(r-1))(\zeta + \phi(m-r)) \prod_{k=1}^{m-1} (1 + \phi(k-1))},$$

con  $r = 1, 2, \dots, m$ . Quindi è possibile ricavare la distribuzione di probabilità per un modello CUBE, definita come

$$P(R = r|\theta) = \pi g_r(\zeta, \phi) + (1 - \pi) \frac{1}{m}$$

dove il parametro aggiuntivo  $\phi$ , che caratterizza la distribuzione, è una misura diretta della sovradisersione. La sovradisersione è legata al concetto di variabilità reciproca e aumenta quando i rispondenti danno risposte per lo più diverse tra loro.

Per quanto riguarda il processo di stima, questo si avvale dell'algoritmo EM, il quale è caratterizzato da un processo di convergenza quasi certo ma

prolungato (McLachlan e Krishnan, 1997). Di conseguenza, si rivela essenziale avvalersi di valori iniziali precisi e congruenti. Tali valori possono essere ottenuti attraverso l'impiego di stimatori coerenti dei parametri del modello Beta-Binomiale, come quelli proposto da Tripathi et al. (1994). La funzione di log-verosimiglianza di un modello CUBE, per  $r=(r_1, \dots, r_n)$  e parametri  $\theta = (\pi, \zeta, \phi)$  risulta pari a:

$$l(\theta) = \sum_{i=1}^n \log \left\{ \pi \left[ \beta_{r_i}(\zeta, \phi) - \frac{1}{m} \right] + \frac{1}{m} \right\} = \sum_{i=1}^m n_r \log \left\{ \pi \left[ \beta_r(\zeta, \phi) - \frac{1}{m} \right] + \frac{1}{m} \right\}$$

poiché le informazioni contenute in  $r$  sono equivalenti a quelle contenute nel vettore delle frequenze assolute  $n=(n_1, \dots, n_m)$ . Per dati valori dei parametri dei modelli CUBE, ovvero per:

- $\pi = 1$ ,
- $\zeta = \frac{(m-1)\theta}{1+m(\theta-2)}$ ,
- $\phi = \frac{1-\theta}{1+\theta(m-2)}$ ,

come dimostrato in Iannario (2012), il modello risulta equivalente a un modello IHG definito in D'Elia (2003). I modelli IHG derivano da distribuzioni unimodali di sentimenti estremi senza incertezza e in cui IHG indica la distribuzione Inversa Ipergeometrica. In questo caso, per il  $i$ -esimo soggetto, un parametro  $\theta_i$  caratterizza il modello e fornisce la probabilità di osservare una valutazione corrispondente alla prima o ultima categoria, pertanto il parametro è valutabile come una misura diretta di preferenza. Di seguito viene riportata la funzione di log-verosimiglianza del modello IHG:

$$\sum_{i=1}^n \sum_{k=1}^m d_{ik} \log P(R = r_k), d_{ik} \begin{cases} 1 & \text{se } R_i = r_k \\ 0 & \text{altrimenti.} \end{cases}$$

# Capitolo 3

## Analisi dello studio Psychopy

### 3.1 Analisi preliminare

In questa sezione verrà condotta un'analisi del primo dataset, ovvero quello legato allo studio Psychopy, utilizzando i modelli esposti in precedenza. Come definito nella nella sezione 1.5, i dati raccolti comprendono diverse variabili: i tempi di lettura degli item e dei filler, il numero di caratteri degli item, i temi degli item, le condizioni sperimentali analizzate, la risposta scelta e quella corretta, nonché un identificativo univoco per ciascun soggetto.

L'obiettivo principale di questa analisi è esaminare la relazione tra il tempo di lettura degli item e l'esito della risposta rispetto alle quattro condizioni sperimentali considerate. Tale analisi si propone di fornire una comprensione dettagliata di come queste variabili interagiscano tra loro e di come influenzino il comportamento dei partecipanti.

La scelta di impiegare modelli a effetti misti è stata dettata dalla necessità di rappresentare accuratamente sia la natura intrinseca delle variabili di interesse sia la struttura dei dati raccolti. In particolare, l'adozione di modelli a effetti misti incrociati risulta adeguata per gestire la variabilità tra i soggetti e tra i temi degli item, consentendo così una stima più precisa degli effetti delle condizioni sperimentali. I modelli adattati in questa sezione sono quelli definiti nel capitolo 2, dove sono stati definiti in dettaglio le strutture e le metodologie di stima utilizzate.

## 3.2 Analisi esplorativa

È stata condotta un'analisi esplorativa preliminare per esaminare la relazione tra le condizioni sperimentali e la prima variabile dipendente, ovvero il tempo di lettura. La Figura 3.1 illustra come il tempo di lettura varia per ciascuna delle condizioni sperimentali analizzate A, B, C e D.

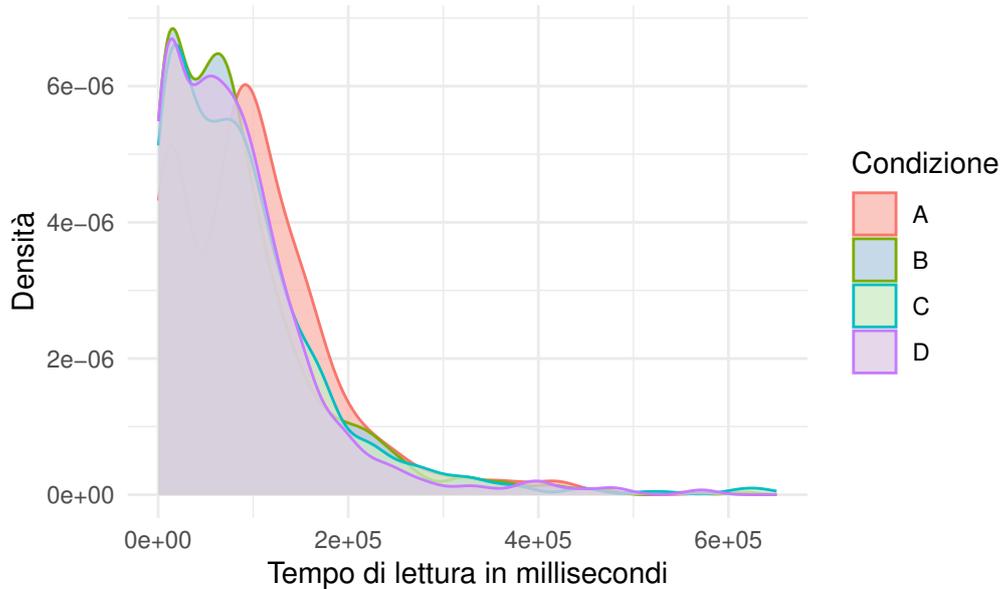


Figura 3.1: Distribuzione del tempo di lettura per condizione

La Figura mostra che la maggior parte dei tempi di lettura sono concentrati verso valori bassi, con una coda destra molto lunga che si estende verso valori più alti. Questo indica che, in generale, i partecipanti tendono a completare la lettura in tempi relativamente brevi, con alcune eccezioni che impiegano tempi significativamente più lunghi. Le curve di densità per le quattro condizioni sono sovrapposte, suggerendo che non ci sono differenze marcate tra le condizioni in termini di distribuzione del tempo di lettura.

Considerata la natura delle distribuzioni osservate, caratterizzate come detto da una concentrazione di valori bassi e una coda destra molto lunga, risulta appropriato applicare una trasformata logaritmica ai tempi di lettura. Nella Figura 3.2 viene riportata la distribuzione del logaritmo del tempo di lettura per le 4 condizioni analizzate:

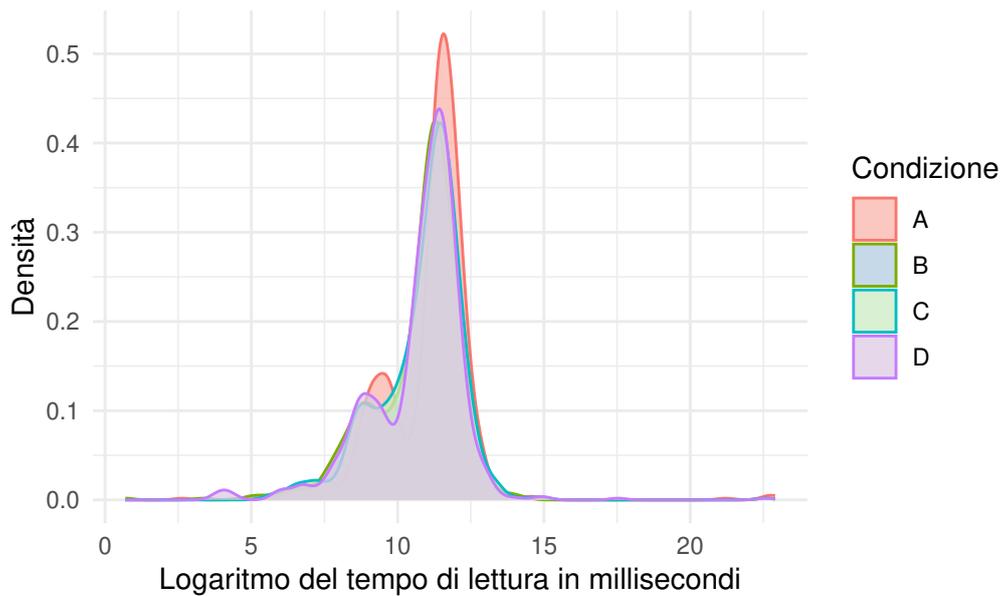


Figura 3.2: Distribuzione del tempo di lettura in scala logaritmica per condizione

Si pone in seguito l'attenzione sulla variabile risposta inclusiva/non inclusiva. Nell'analisi, una risposta è considerata inclusiva se la risposta scelta dal partecipante coincide con la risposta corretta, che in questo studio è sempre l'opzione inclusiva. Al contrario, una risposta è classificata come non inclusiva se la risposta scelta dal partecipante non coincide con la risposta corretta. La Figura 3.3 rappresenta la distribuzione delle risposte inclusive e non inclusive per ciascuna delle quattro condizioni sperimentali A, B, C e D.

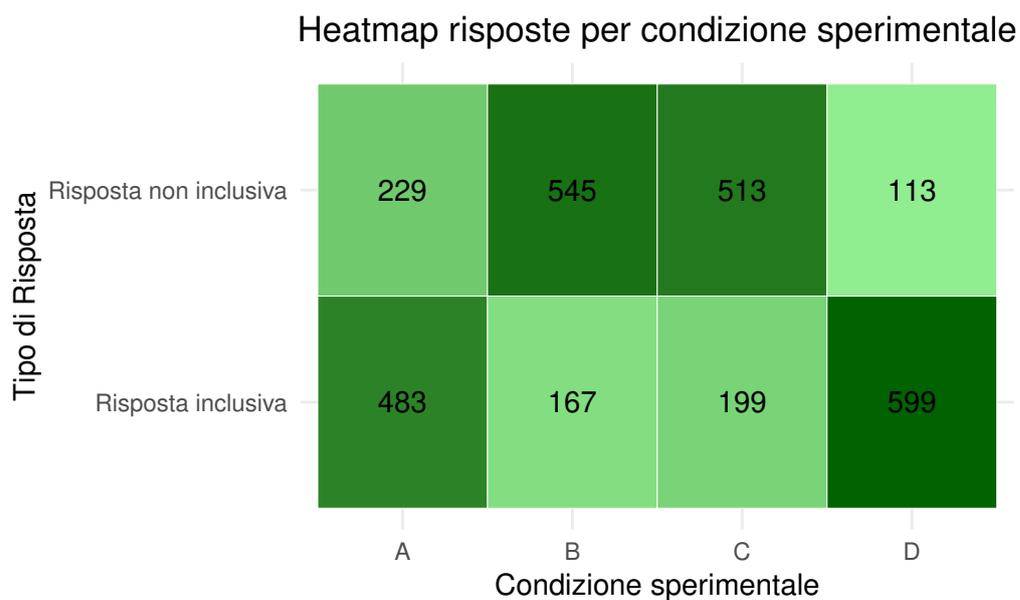


Figura 3.3: Distribuzione delle risposte per condizione

Dalla Figura si osserva che la distribuzione delle risposte varia significativamente tra le condizioni. Nella condizione A, il numero di risposte inclusive (483) supera di gran lunga il numero di risposte non inclusive (229), suggerendo una marcata propensione dei partecipanti a fornire risposte inclusive in questa condizione. Al contrario, nella condizione B, la situazione è invertita: le risposte non inclusive (545) sono notevolmente più numerose rispetto alle risposte inclusive (167). Le condizioni C e D mostrano comportamenti differenti. Nella condizione C, le risposte non inclusive (513) sono più frequenti delle risposte inclusive (199). La condizione D presenta un'alta frequenza di risposte inclusive (599) rispetto alle risposte non inclusive (113), suggerendo un comportamento analogo alla condizione A.

Nella sezione successiva, verranno proposti modelli statistici multivariati per analizzare in modo più dettagliato i tempi di lettura e la correttezza delle risposte, al fine di comprendere meglio le dinamiche sottostanti e le influenze delle condizioni sperimentali.

### 3.3 Modelli per il tempo di lettura

In questa sezione vengono definiti i modelli per l'analisi del tempo di lettura. Seguendo l'approccio adottato negli studi di Stetie e Zunino (2022) e alla luce delle considerazioni dell'analisi esplorativa riguardo la distribuzione del tempo di lettura, l'analisi è stata condotta utilizzando la trasformata logaritmica del tempo di risposta.

I modelli proposti sono modelli lineari a effetti misti, metodi adatti per risposte continue, a partire da modelli a effetti misti incrociati per partecipanti e temi. Nei modelli sono incluse le intercette casuali per partecipanti e temi, al fine di considerare e gestire la variabilità intrinseca associata a questi due livelli di raggruppamento. L'inclusione delle intercette casuali per i partecipanti permette di tenere conto delle differenze individuali nella capacità di risposta, mentre l'inclusione delle intercette casuali per i temi considera la variabilità dovuta alla diversa difficoltà o alla natura dei temi stessi.

Vengono inoltre introdotti come effetti fissi la condizione sperimentale e il numero di caratteri dell'item. L'inclusione della condizione sperimentale come effetto fisso consente di valutare l'influenza diretta di questa variabile sui tempi di lettura, permettendo di isolare e quantificare l'effetto specifico della condizione sperimentale. Vengono considerati il numero di caratteri dell'item come un ulteriore effetto fisso. È infatti ragionevole ipotizzare che una variazione sostanziale nel numero di caratteri possa esercitare un'influenza significativa sui tempi di lettura. Un maggiore numero di caratteri implica un aumento della quantità di informazione che il partecipante deve processare, il che potrebbe tradursi in un tempo di lettura più lungo. Dall'analisi del modello appena proposto però il numero di caratteri non risulta significativamente diverso da zero, con un p-value pari a 0.353. La motivazione sottostante a questo risultato potrebbe essere attribuita alla distribuzione relativamente uniforme del numero di caratteri tra i diversi item (la distribuzione è attorno agli 80 caratteri). Questa uniformità nella distribuzione potrebbe spiegare la

stima non significativa del parametro considerato.

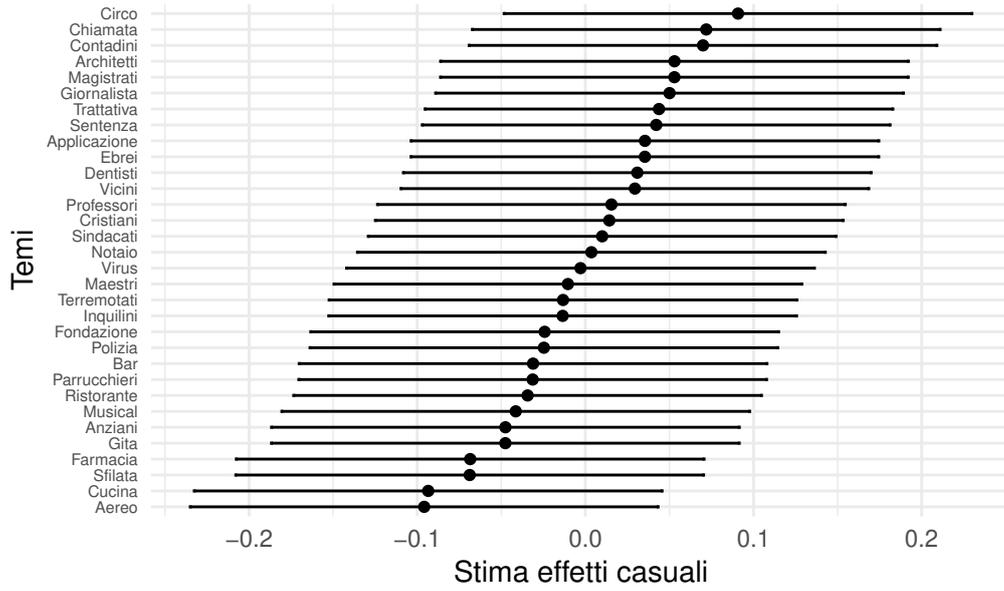
Si procede quindi alla stima di un modello per il tempo di lettura logaritmico togliendo il parametro relativo al numero di caratteri, preferendo un modello più parsimonioso. Nella Tabella 3.1 viene riportato il modello stimato, in cui viene considerata la condizione sperimentale come unico effetto fisso oltre che all'inserimento di effetti incrociati per tema e partecipante.

| <b>Effetto</b>                 | <b>Stima</b> | <b>Errore Std.</b> | <b>t value</b> | <b>P-value</b> |
|--------------------------------|--------------|--------------------|----------------|----------------|
| <b>Effetti fissi</b>           |              |                    |                |                |
| Intercetta                     | 11.06020     | 0.12321            | 89.764         | < 0.001        |
| Condizione B                   | -0.35025     | 0.06292            | -5.566         | < 0.001        |
| Condizione C                   | -0.25101     | 0.06294            | -3.988         | < 0.001        |
| Condizione D                   | -0.36695     | 0.06292            | -5.832         | < 0.001        |
| <b>Effetti casuali</b>         |              |                    |                |                |
| <b>Partecipante</b>            |              |                    |                |                |
| Intercetta                     | 1.1272       | 1.0617             |                |                |
| <b>Tema</b>                    |              |                    |                |                |
| Intercetta                     | 0.0078       | 0.0885             |                |                |
| <b>Varianza residua</b>        | 1.3768       | 1.1734             |                |                |
| <b>Criteri di Informazione</b> |              |                    |                |                |
| AIC                            | 9307.36      |                    |                |                |
| BIC                            | 9349.04      |                    |                |                |

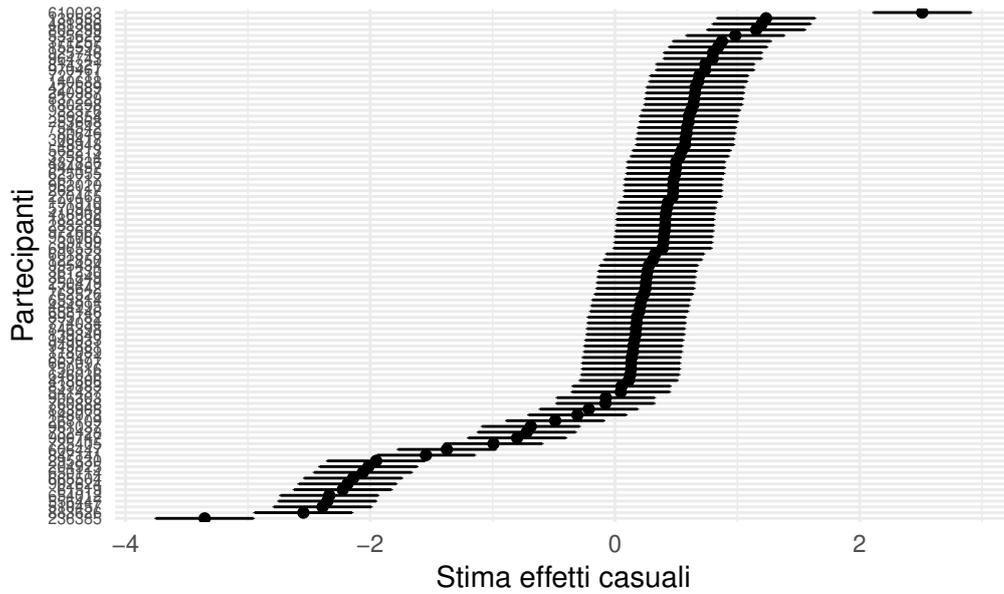
Tabella 3.1: Risultati del Modello Lineare Misto

L'analisi della varianza dell'effetto casuale associato ai temi ha evidenziato un valore piuttosto ridotto, pari a 0.0078. Questo risultato potrebbe indicare che non vi sono differenze sostanziali tra i temi in termini di impatto sulla variabile di risposta. Tale osservazione potrebbe suggerire che le variazioni tra i temi non incidano significativamente sulla variabilità complessiva della risposta. Per rappresentare graficamente la situazione degli effetti casuali, in

Figura 3.4, si propongono due grafici con le relative stime.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 3.4: Effetti casuali stimati con il modello lineare a effetti misti incrociati

Dall'analisi di questi grafici emerge chiaramente come la varianza tra i partecipanti sia significativamente maggiore rispetto a quella tra i temi. Questo risultato rispecchia le stime dei parametri di varianza ottenute dal modello, confermando che la variabilità tra i partecipanti contribuisce in misura più rilevante alla variabilità complessiva dei dati. Considerate queste premesse, nella successiva fase dell'analisi verrà adattato un modello con un singolo ef-

fetto casuale per partecipante, adottando così un approccio più parsimonioso, riportando l'output nella Tabella 3.2.

| <b>Effetto</b>                 | <b>Stima</b> | <b>Errore Std.</b> | <b>t value</b> | <b>P-value</b> |
|--------------------------------|--------------|--------------------|----------------|----------------|
| <b>Effetti fissi</b>           |              |                    |                |                |
| Intercetta                     | 11.0548      | 0.1197             | 92.392         | < 0.001        |
| Condizione B                   | -0.3341      | 0.0622             | -5.371         | < 0.001        |
| Condizione C                   | -0.2373      | 0.0622             | -3.816         | < 0.001        |
| Condizione D                   | -0.3572      | 0.0622             | -5.743         | < 0.001        |
| <b>Effetti casuali</b>         |              |                    |                |                |
| <b>Partecipante</b>            |              |                    |                |                |
| Intercetta                     | 1.102        | 1.050              |                |                |
| <b>Varianza residua</b>        | 1.377        | 1.174              |                |                |
| <b>Criteri di Informazione</b> |              |                    |                |                |
| AIC                            | 9308.01      |                    |                |                |
| BIC                            | 9343.74      |                    |                |                |

Tabella 3.2: Risultati del Modello Lineare misto con effetto casuale per partecipante

L'intercetta del modello presenta una stima di 11.0548, con un p-value inferiore a 0.001, indicando che il tempo di lettura logaritmico medio è significativamente diverso da zero per la condizione di base, identificata dal maschile generico. Il parametro associato alla condizione B ha una stima di -0.3341 e un p-value inferiore a 0.001, suggerendo un effetto negativo significativo sul tempo di lettura logaritmico. Questo implica che, rispetto alla condizione di riferimento, la condizione B riduce il tempo di lettura logaritmico medio di 0.3341. Analogamente, il parametro associato alla condizione C presenta una stima di -0.2373 con un p-value inferiore a 0.001. Anche in questo caso, l'effetto negativo è significativo, riducendo il tempo di lettura logaritmico medio di 0.2373 rispetto alla condizione di riferimento. Il parametro associato alla condizione D, con una stima di -0.3572 e un p-value inferiore a 0.001, mostra

anch'essa un effetto negativo significativo sul tempo di lettura logaritmico. In particolare, questa condizione riduce il tempo di lettura logaritmico medio di 0.3572 rispetto alla condizione di riferimento.

L'effetto casuale per i partecipanti ha una stima di 1.102. Questo valore riflette la variabilità interindividuale nel tempo di lettura logaritmico, indicando che ci sono differenze significative tra i partecipanti che il modello è in grado di catturare.

In sintesi, i risultati indicano che le condizioni sperimentali hanno un effetto negativo significativo sul tempo di lettura logaritmico rispetto alla condizione del maschile generico, riducendolo rispetto alla condizione di riferimento. Inoltre l'inclusione dell'effetto casuale per i partecipanti nel modello evidenzia una considerevole variabilità tra individui, fornendo una comprensione più completa delle dinamiche che influenzano il tempo di lettura logaritmico.

Dopo aver adattato entrambi i modelli, e considerando la situazione tra modelli annidati, si propone di eseguire un test del rapporto di verosimiglianza per verificare la significatività dell'effetto casuale associato ai temi. Questo test consentirà di determinare se l'inclusione dell'effetto casuale relativo ai temi apporta un miglioramento significativo alla bontà di adattamento del modello rispetto a un modello più parsimonioso che considera solo l'effetto casuale per partecipante. I modelli proposti, come definito in 2.1.1, vengono stimati utilizzando la funzione di verosimiglianza ristretta. Di conseguenza, non è possibile utilizzare un normale test del rapporto di verosimiglianza per valutare la significatività degli effetti casuali. Pertanto, si rende necessario adottare il test del rapporto di verosimiglianza ristretto, come descritto da Scheipl, Greven e Küchenhoff (2008). Questo approccio permette di considerare adeguatamente le specifiche caratteristiche dei modelli a effetti misti, garantendo una valutazione più accurata dell'effetto casuale associato ai temi.

Nonostante il parametro di varianza dell'effetto casuale associato ai temi sia pari a 0.0078, il test del rapporto di verosimiglianza ristretto ha prodotto

un p-value di 0.047, portando a preferire il modello completo. Dal risultato ottenuto, è fondamentale considerare che anche una varianza apparentemente piccola può avere un impatto significativo sulla variabilità complessiva dei dati.

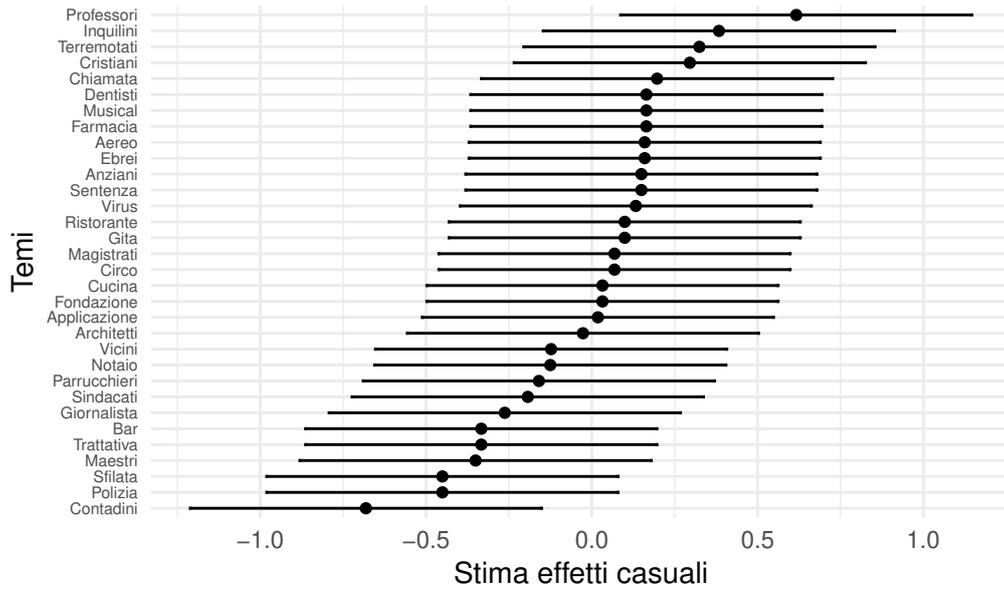
### **3.4 Modelli per l'esito della risposta**

L'analisi si è successivamente focalizzata sull'accuratezza delle risposte, valutando la probabilità che queste fossero corrette o errate mediante l'impiego di modelli logistici a effetti misti. Questa fase dell'analisi si differenzia dalla precedente per il suo obiettivo specifico, che non è più incentrato sul tempo di risposta dei soggetti, bensì sulla precisione delle loro risposte. Come nei modelli precedenti, anche in questa fase sono state incluse nei modelli le intercette casuali per partecipanti e temi, al fine di considerare e gestire la variabilità intrinseca associata a questi due livelli di raggruppamento. L'inclusione delle intercette casuali per partecipanti permette di tenere conto delle differenze individuali nella capacità di risposta, mentre l'inclusione delle intercette casuali per temi considera la variabilità dovuta alla diversa difficoltà o alla natura dei temi stessi. Questa metodologia consente di ottenere stime più accurate delle probabilità di risposte corrette, migliorando così la comprensione delle dinamiche sottostanti l'accuratezza delle risposte nei diversi contesti tematici e tra i diversi partecipanti. Viene quindi adattato un modello logistico a effetti misti incrociati per tema e partecipante, con effetti fissi per condizione sperimentale e numero di caratteri dell'item. Quest'ultimo, però, come per l'analisi precedente, risulta non significativo e per questo viene rimossa questa variabile. L'output del modello è riportato nella Tabella 3.3.

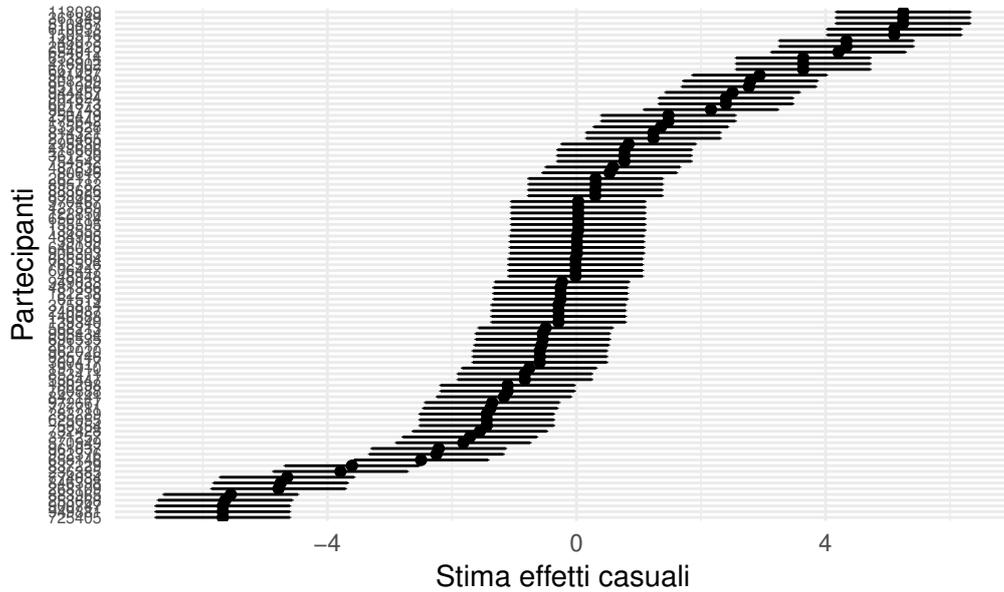
| <i>Risposta corretta</i> |             |                |         |
|--------------------------|-------------|----------------|---------|
| Predittori               | Odds Ratios | Standard Error | P-value |
| Intercetta               | 3.67        | 0.3284         | < 0.001 |
| Condizione B             | 0.03        | 0.1905         | < 0.001 |
| Condizione C             | 0.05        | 0.1764         | < 0.001 |
| Condizione D             | 5.68        | 0.1941         | < 0.001 |
| <b>Effetti casuali</b>   |             |                |         |
| Predittore               | Stima       | Standard Error |         |
| $\sigma^2$               | 3.29        |                |         |
| $\tau_{00}$ partecipante | 8.0008      | 2.8286         |         |
| $\tau_{00}$ tema         | 0.1371      | 0.3703         |         |
| AIC                      | 1988.2      |                |         |
| BIC                      | 2023.8      |                |         |

Tabella 3.3: Risultati del modello logistico misto con effetto casuale per partecipante e tema

L'analisi della varianza dell'effetto casuale associato ai temi ha un valore pari a 0.1371. Questo risultato potrebbe indicare che non vi sono differenze sostanziali tra i temi in termini di impatto sulla variabile di risposta. Per rappresentare graficamente la situazione degli effetti casuali, si propongono in Figura 3.5 due grafici con le relative stime.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 3.5: Effetti casuali stimati con il modello logistico a effetti misti incrociati

Anche in questo caso, emerge chiaramente come la varianza tra i partecipanti sia significativamente maggiore rispetto a quella tra i temi. Questo risultato rispecchia le stime dei parametri di varianza ottenute dal modello, confermando che la variabilità tra i partecipanti contribuisce in misura più rilevante alla variabilità complessiva dei dati.

Alla luce di questi risultati, è stato proposto un modello più parsimonioso. Nello specifico, si suggerisce l'adozione di un modello ridotto che incorpori

un effetto casuale unicamente per i partecipanti, eliminando l'effetto casuale associato ai temi. Tale approccio consente di focalizzare l'attenzione sulle variazioni interindividuali, che sembrano avere un ruolo più rilevante nella spiegazione della variabilità osservata.

L'output del modello è riportato nella Tabella 3.4.

| <i>Risposta corretta</i> |             |                |         |
|--------------------------|-------------|----------------|---------|
| Predittori               | Odds Ratios | Standard Error | P-value |
| Intercetta               | 3.61        | 0.3120         | < 0.001 |
| Condizione B             | 0.03        | 0.1839         | < 0.001 |
| Condizione C             | 0.05        | 0.1716         | < 0.001 |
| Condizione D             | 5.27        | 0.1902         | < 0.001 |
| <b>Effetti casuali</b>   |             |                |         |
| Predittore               | Stima       | Standard Error |         |
| $\sigma^2$               | 3.29        |                |         |
| $\tau_{00}$ partecipante | 7.607       | 2.758          |         |
| AIC                      | 1996.3      |                |         |
| BIC                      | 2026.0      |                |         |

Tabella 3.4: Risultati del modello misto con effetti fissi e casuali

La tabella mostra i risultati dell'analisi del modello logistico a effetti misti, comprendente sia effetti fissi che casuali. L'odds ratio dell'intercetta è di 3.61 indicando che, nella condizione di riferimento (maschile generico), la probabilità di fornire una risposta corretta è 3.61 volte maggiore rispetto alla probabilità di fornire una risposta errata. La condizione B ha un odds ratio di 0.03, anch'esso altamente significativo ( $p < 0.001$ ), indicando una riduzione drastica dell'odds di una risposta corretta del 97% rispetto alla condizione di riferimento. Analogamente, la condizione C mostra un odds ratio di 0.05, suggerendo una riduzione del 95% nell'odds di una risposta corretta ( $p <$

0.001). Diversamente, la condizione D presenta un odds ratio di 5.27, indicando un aumento significativo dell'odds di una risposta corretta di circa 5.27 volte rispetto alla condizione di riferimento ( $p < 0.001$ ). Gli effetti casuali del modello sono rappresentati dalle stime della varianza degli effetti casuali per i partecipanti. Questa è stimata a 7.607 con un errore standard di 2.758, indicando una sostanziale variabilità tra i partecipanti rispetto alla risposta corretta.

In conclusione, i risultati indicano che le condizioni B e C hanno un effetto negativo significativo sull'odds di una risposta corretta, mentre la condizione D ha un effetto positivo significativo. L'inclusione degli effetti casuali per i partecipanti è supportata dalla significativa variabilità tra i partecipanti stessi, come evidenziato dalla stima di  $\tau_{00}$ . Questi risultati forniscono una comprensione dettagliata delle variazioni nelle risposte corrette tra le diverse condizioni e tra i partecipanti, evidenziando l'importanza degli effetti casuali nel modello.

I modelli proposti, come definito nella sezione 2.2, sono stati stimati utilizzando la funzione di verosimiglianza. Di conseguenza, per valutare la significatività degli effetti casuali e determinare il modello migliore, viene utilizzato il test del log-rapporto di verosimiglianza. Questo approccio consente di confrontare adeguatamente i modelli annidati e identificare quello che offre la miglior qualità di adattamento ai dati.

Nonostante il parametro di varianza dell'effetto casuale associato ai termini sia pari a 0.1371, il test del rapporto di verosimiglianza ha prodotto un p-value  $< 0.001$ , portando a preferire il modello completo con un livello di significatività del 5%. Dato il risultato ottenuto, è fondamentale riconoscere che anche una varianza relativamente piccola può avere un impatto significativo sulla variabilità complessiva dei dati nel contesto di un modello statistico complesso.

## 3.5 Analisi di sopravvivenza

In questa sezione viene utilizzato un modello di sopravvivenza per analizzare il tempo di lettura come tempo all'evento, adottando un approccio che si integra con i modelli a effetti misti per soggetto e tema. Questo metodo consente di studiare come il rischio di dare una risposta corretta vari nel tempo in relazione alla condizione sperimentale, esaminando l'effetto dinamico di quest'ultima sul tempo di risposta, offrendo una visione dettagliata dell'impatto delle variazioni sperimentali. Il modello proposto è un modello di Cox a effetti misti, definito nella sezione 2.1.6. Il modello di Cox a effetti misti introduce componenti casuali nel modello, permettendo di controllare l'eterogeneità non osservata tra i soggetti e temi. Il modello a effetti misti proposto valuta come variabile relativa all'evento la correttezza della risposta. Tale variabile è codificata in modo che l'evento di interesse, ovvero una risposta corretta, sia indicato con il valore 1, mentre una risposta errata è trattata come censura e indicata con il valore 0. La scelta di trattare una risposta errata come una censura è motivata dal desiderio di focalizzarsi sull'analisi del tempo fino al verificarsi di una risposta corretta, mantenendo al contempo l'integrità dei dati di soggetti che non hanno fornito risposte corrette entro il periodo di osservazione. L'output del modello è riportato nella Tabella 3.5, in cui vengono inclusi gli effetti casuali per partecipanti allo studio e tema.

| <i>Effetti fissi</i> |             |                  |                 |                |
|----------------------|-------------|------------------|-----------------|----------------|
| <b>Effect</b>        | <b>Coef</b> | <b>Exp(Coef)</b> | <b>SE(Coef)</b> | <b>P-value</b> |
| Condizione B         | -0.87626    | 0.4163           | 0.09545         | < 0.001        |
| Condizione C         | -0.78204    | 0.4575           | 0.08898         | < 0.001        |
| Condizione D         | 0.62159     | 1.8619           | 0.06457         | < 0.001        |

| <i>Effetti casuali</i> |           |                 |
|------------------------|-----------|-----------------|
| <b>Gruppo</b>          | <b>SD</b> | <b>Varianza</b> |
| Partecipante           | 2.5837194 | 6.6756060       |
| Tema                   | 0.1750785 | 0.0306525       |

Tabella 3.5: Risultati del modello di Cox a effetti misti

L'analisi della varianza dell'effetto casuale associato ai temi ha rivelato un valore relativamente ridotto, pari a 0.0306525. Questo risultato suggerisce l'assenza di differenze sostanziali tra i temi in termini dell'effetto sulla variabile di risposta. Viene quindi proposto un modello più parsimonioso. In particolare, si propone un modello ridotto che incorpori un effetto casuale unicamente per i partecipanti, eliminando l'effetto casuale associato ai temi. Tale modifica metodologica non solo è giustificata dalla non significativa variazione tra i temi, ma potrebbe anche migliorare l'efficienza del modello riducendo la complessità senza compromettere la capacità di catturare le variazioni pertinenti ai partecipanti. L'output del modello è riportato nella Tabella 3.6.

| <i>Effetti fissi</i> |             |                  |                 |                |
|----------------------|-------------|------------------|-----------------|----------------|
| <b>Effect</b>        | <b>Coef</b> | <b>Exp(Coef)</b> | <b>SE(Coef)</b> | <b>P-value</b> |
| Condizione B         | -0.85559    | 0.42503          | 0.09469         | < 0.001        |
| Condizione C         | -0.74681    | 0.47388          | 0.08794         | < 0.001        |
| Condizione D         | 0.59820     | 1.81885          | 0.06395         | < 0.001        |

| <i>Effetti casuali</i> |                 |           |                 |
|------------------------|-----------------|-----------|-----------------|
| <b>Gruppo</b>          | <b>Variable</b> | <b>SD</b> | <b>Varianza</b> |
| Partecipante           | Intercetta      | 1.022084  | 1.044656        |

Tabella 3.6: Risultati del modello di Cox a effetti misti

Nell'ambito dei modelli a effetti misti, si tende a preferire la stima REML, poiché questa ha rivelato una maggiore affidabilità; al contrario, le stime ottenute mediante il metodo ML si rivelano frequentemente sovrastimate. Analogamente, viene proposta l'adozione della stima REML anche per il modello di Cox a effetti misti. Ciononostante, le evidenze sia teoriche che empiriche a sostegno della sua superiorità risultano limitate. In particolare, Cortinas (2004) evidenzia che la stima REML manifesta un bias superiore rispetto a quella ML. Le stime dei modelli di Cox proposti vengono stimati tramite il metodo ML. Per questa ragione, trattandosi di un contesto di modelli annidati, è possibile definire il miglior modello tramite test del log-rapporto di verosimiglianza. L'analisi dei modelli suggerisce un'incrementata adattabilità dei dati per il modello con effetti casuali incrociati, come evidenziato dal valore della statistica Chi-quadrato e dal corrispondente p-value estremamente basso (32.689, p-value < 0.001).

Infine, viene proposta una breve interpretazione del modello a effetti misti incrociati. Nel contesto dell'analisi dei modelli di Cox, il coefficiente esponenziale, ovvero l'hazard ratio, rappresenta una misura per quantificare l'impatto delle covariate sui tempi di sopravvivenza. Per la condizione B, l'hazard ratio pari a 0.4163 suggerisce che il rischio di evento per i soggetti sotto questa

condizione è ridotto al 41.63% rispetto ai soggetti nella condizione di riferimento, indicando una notevole diminuzione del rischio. Analogamente, per la condizione C, l'hazard ratio di 0.4575 conferma che il rischio associato è pari al 45.75% di quello nella condizione di base, sottolineando anch'esso una riduzione del rischio. Differentemente, per la condizione D, l'hazard ratio evidenzia un incremento significativo del rischio; con un valore di 1.8619, il rischio è quasi raddoppiato, aumentando dell'86.19% rispetto alla condizione di riferimento. Inoltre i p-value mostrano che le stime sono statisticamente significative. Questi risultati dimostrano come le condizioni sperimentali analizzate influenzino significativamente la stima del rischio.

### 3.6 Modello logit multinomiale con effetti misti

Fin'ora sono stati utilizzati dei modelli per analizzare due tipologie di risposte: una risposta continua, rappresentante il tempo (in scala logaritmica) di lettura e la risposta binaria, categorizzata come inclusiva o non inclusiva. In questa sezione, si estende l'analisi includendo una variabile risposta avente 3 categorie: inclusiva, non inclusiva e indefinita. Quest'ultima categoria è utilizzata per catturare le risposte dei partecipanti che hanno scelto l'opzione "Non lo so, nessuna delle precedenti". Tale estensione mira a fornire una comprensione più approfondita dell'impatto della condizione sperimentale, nonché delle caratteristiche individuali dei soggetti e delle peculiarità dei temi, sulle risposte fornite. L'obiettivo è quindi quello di esplorare in modo più dettagliato come queste variabili influenzino la scelta della categoria di risposta da parte dei partecipanti. Viene di seguito riportata la distribuzione della risposta considerata per ogni condizione analizzata in Figura 3.6.

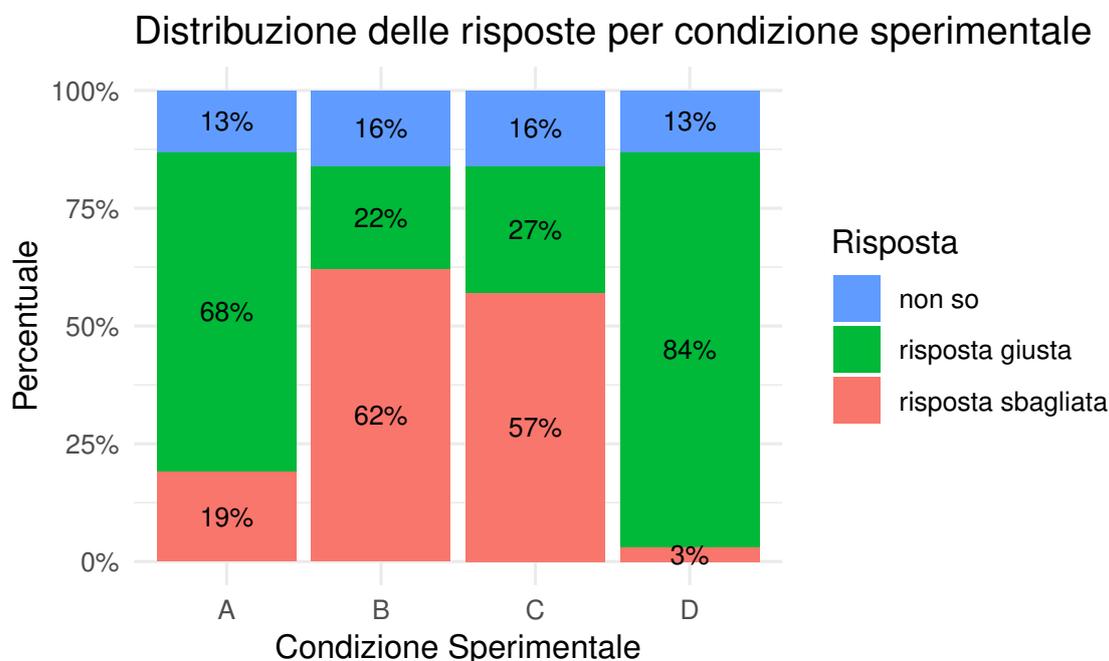


Figura 3.6: Distribuzione delle risposte per condizione analizzata

Per l'analisi di questa tipologia di risposta si propone quindi un modello

logit multinomiale, definito nella sezione 2.1.5. Viene adattato questo modello includendo effetti casuali incrociati per soggetto e partecipante, riportando nella Tabella 3.7 le stime ottenute.

| <b>Effetto</b>                        | <b>Stima</b> | <b>Errore Std.</b> | <b>Pvalue</b> |
|---------------------------------------|--------------|--------------------|---------------|
| <b>Risposta corretta vs sbagliata</b> |              |                    |               |
| Intercetta                            | 1.8006       | 0.2542             | < 0.001       |
| Condizione B                          | -3.4102      | 0.1841             | < 0.001       |
| Condizione C                          | -2.9069      | 0.1717             | < 0.001       |
| Condizione D                          | 2.4828       | 0.2718             | < 0.001       |
| <b>Risposta "Non so" vs sbagliata</b> |              |                    |               |
| Intercetta                            | -1.8061      | 0.3966             | < 0.001       |
| Condizione B                          | -0.7231      | 0.3025             | 0.016         |
| Condizione C                          | -0.5238      | 0.2998             | 0.080         |
| Condizione D                          | 2.3455       | 0.3889             | < 0.001       |
| <b>Livello di Raggruppamento</b>      | <b>Stima</b> | <b>Errore Std.</b> |               |
| <b>Partecipante</b>                   |              |                    |               |
| Categoria 1 vs 0                      | 3.9861       | 2.790              |               |
| Categoria 3 vs 0                      | 0.4866       | 2.583              |               |
| <b>Tema</b>                           |              |                    |               |
| Categoria 1 vs 0                      | 0.09659      | 0.0003488          |               |
| Categoria 3 vs 0                      | 0.06391      | 0.0002691          |               |

Tabella 3.7: Modello Logit Multinomiale a Effetti Misti

Il modello logit multinomiale a effetti misti stimato confronta le categorie di risposta a coppie, includendo sia effetti fissi per la condizione sperimentale sia effetti casuali associati agli individui e ai temi. Le stime degli effetti fissi indicano che le condizioni sperimentali esercitano un impatto significativo sulle risposte dei partecipanti. In particolare, le condizioni B e C riducono significativamente la probabilità di fornire una risposta inclusiva (categoria

1) o una risposta "non so" (categoria 3) rispetto a una risposta non inclusiva (categoria 0), mentre la condizione D aumenta tale probabilità. Gli effetti casuali rivelano una variabilità significativa tra gli individui e, in misura minore, tra i temi. La varianza associata agli individui suggerisce l'esistenza di differenze sostanziali tra i partecipanti che non sono spiegate dalle condizioni sperimentali, evidenziando l'importanza delle caratteristiche individuali non misurate. La varianza associata ai temi, sebbene presente, risulta essere molto più ridotta, indicando che i temi contribuiscono in misura minore alla variabilità delle risposte rispetto agli individui. Questo suggerisce che i temi hanno un impatto meno variabile sulle risposte, probabilmente perché i partecipanti rispondono in modo più simile quando esposti a specifici temi, rispetto a come variano le loro risposte individualmente. Seguendo l'approccio utilizzato precedentemente, viene adattato un modello più parsimonioso con singola intercetta casuale per partecipante. Viene riportato l'output del modello presentato nella Tabella 3.8.

| <b>Effetto</b>                        | <b>Stima</b> | <b>Errore Std.</b> | <b>Pvalue</b> |
|---------------------------------------|--------------|--------------------|---------------|
| <b>Risposta corretta vs sbagliata</b> |              |                    |               |
| Intercetta                            | 1.7904       | 0.2448             | < 0.001       |
| Condizione B                          | -3.3809      | 0.1830             | < 0.001       |
| Condizione C                          | -2.8853      | 0.1705             | < 0.001       |
| Condizione D                          | 2.4712       | 0.2714             | < 0.001       |
| <b>Risposta "Non so" vs sbagliata</b> |              |                    |               |
| Intercetta                            | -1.8047      | 0.3912             | < 0.001       |
| Condizione B                          | -0.7268      | 0.3010             | 0.016         |
| Condizione C                          | -0.5111      | 0.2979             | 0.086         |
| Condizione D                          | 2.3332       | 0.3876             | < 0.001       |
| <b>Livello di Raggruppamento</b>      | <b>Stima</b> | <b>Errore Std.</b> |               |
| <b>Partecipante</b>                   |              |                    |               |
| Categoria 1 vs 0                      | 3.8560       | 3.850              |               |
| Categoria 3 vs 0                      | 0.5065       | 3.762              |               |

Tabella 3.8: Modello Logit Multinomiale con effetto casuale per partecipante

Come discusso da He et al. (2021), l'inferenza per le componenti di varianza degli effetti casuali presenta diverse problematiche di natura metodologica. La statistica del rapporto di verosimiglianza per testare l'ipotesi che una singola componente di varianza sia pari a zero segue una distribuzione nulla che è approssimativamente una mistura tra una distribuzione degenera a 0 e una distribuzione chi-quadrato con 1 grado di libertà (Self e Liang, 1987). Inoltre, in uno studio recente, è stato osservato che il test basato sugli effetti casuali per modelli logit multinomiali non è invariante rispetto alla scelta del livello di riferimento (Liu et al., 2021). Questo risulta particolarmente interessante, poiché il modello di regressione logistica, spesso considerato come un caso speciale della regressione logistica multinomiale, è noto per possedere la proprietà di invarianza.

La mancanza di tale proprietà nel modello multinomiale può creare problematiche pratiche significative, poiché si potrebbe giungere a conclusioni potenzialmente contraddittorie a seconda della scelta del livello di riferimento. Viste le problematiche metodologiche discusse, si riconosce che la valutazione delle stime delle componenti di varianza per gli effetti casuali è complessa. Tuttavia, nonostante la varianza del tema sia vicina a zero, e considerata l'assenza di test robusti che possano fornire una valutazione definitiva, si è deciso di preferire un modello completo con doppio effetto casuale. Questa decisione è supportata dai risultati ottenuti dai modelli precedenti, i quali hanno evidenziato l'importanza di includere gli effetti casuali per catturare adeguatamente la variabilità intrinseca nei dati.

### 3.7 Modelli CUB e CUBE

Nella Sezione 2.2 sono stati introdotti i modelli per l'analisi di dati ordinali. La variabile risposta per questa tipologia di modelli viene riclassificata in quattro livelli ordinali: risposta completamente errata con connotazione femminile (1), risposta completamente errata con connotazione maschile (2), risposta indefinita (3), e risposta inclusiva (4). Gli esperti di linguistica hanno ipotizzato che una risposta totalmente connotata come femminile sia priva di validità contestuale, pertanto essa è considerata l'errore più grave. A sostegno di tale assunzione, si presenta il seguente esempio: nella condizione A, si esamina l'item "Gli alunni rimasero scioccati dai brutti voti. Era stato un semestre difficile." Non vi sono presupposti logici che giustifichino una risposta esclusivamente femminile in tale contesto, rendendo tale opzione inadeguata e illogica. Analogamente, lo stesso ragionamento si applica alla condizione D. Al contrario, per le condizioni B e C potrebbero sorgere dubbi riguardo l'ordinamento, ma per garantire la coerenza dell'analisi, si assume che tale ordinamento sia applicabile uniformemente a tutte le condizioni esaminate. Dati questi presupposti, si procede adattando i modelli CUB considerando la risposta con 4 livelli.

Nella Tabella 3.9 viene rappresentata la distribuzione della risposta ordinale rispetto alla condizione sperimentale.

|   | Condizione A | Condizione B | Condizione C | Condizione D | Totale |
|---|--------------|--------------|--------------|--------------|--------|
| 1 | 7            | 426          | 392          | 9            | 834    |
| 2 | 127          | 7            | 5            | 9            | 148    |
| 3 | 90           | 110          | 113          | 95           | 408    |
| 4 | 472          | 153          | 186          | 583          | 1394   |

Tabella 3.9: Distribuzione della risposta per ogni condizione sperimentale

Dalla tabella si evince come le risposte per le categorie estreme siano le più frequenti. La variabilità nelle risposte suggerisce che la percezione o

l'esperienza dei partecipanti varia notevolmente a seconda delle condizioni. Data questa condizione di distribuzione dei dati, infatti, vengono riscontrati dei problemi di convergenza provando ad adattare un qualsiasi modello CUB poichè, come definito nella sezione 2.2, questi modelli non sono adeguati per affrontare una possibile alta variabilità tra le risposte, dato che la posizione e la variabilità di una variabile casuale binomiale sono fortemente e strettamente correlate.

L'approccio utilizzato in questa sezione consiste nell'adattare inizialmente un modello IHG passando poi alla stima di un modello CUBE in cui è possibile valutare le componenti di feeling, incertezza e dispersione.

Il modello IHG, in cui è possibile modellare solamente la componente di feeling, viene stimato prevedendo come unica covariata la condizione sperimentale analizzata. Il parametro di feeling del modello rappresenta l'inclinazione o la predisposizione dei partecipanti verso specifiche categorie di risposta. Nel contesto del linguaggio inclusivo, la condizione sperimentale potrebbe influenzare significativamente questa inclinazione. Viene riportato l'output del modello appena descritto nella Tabella 3.10.

| Coefficienti   | Stime     | Std. Error | Wald   | P-value |
|----------------|-----------|------------|--------|---------|
| Costante       | 0.43362   | 0.06553    | 6.62   | 0.020   |
| Condizione B   | -2.56316  | 0.09405    | -27.25 | < 0.001 |
| Condizione C   | -2.34584  | 0.09215    | -25.46 | < 0.001 |
| Condizione D   | 1.09004   | 0.11368    | 9.59   | 0.004   |
| Log-likelihood | -2754.747 |            |        |         |
| AIC            | 5517.494  |            |        |         |
| BIC            | 5541.22   |            |        |         |

Tabella 3.10: IHG Model

Nel modello IHG proposto, si osserva un'interazione tra le condizioni sperimentali e la predisposizione dei partecipanti alle diverse categorie di risposta. Il coefficiente relativo alla costante risulta positivo, indicando una tendenza

dei partecipanti a orientarsi verso la condizione A, dove il linguaggio non viene modificato. Per la condizione B si registra un significativo decremento nella predisposizione alle risposte, evidenziato da un coefficiente negativo. Questo fenomeno potrebbe riflettere una minor familiarità o una maggiore difficoltà nel riconoscere o accettare questa forma inclusiva, che distorce l'uso linguistico consueto. In modo analogo, la condizione C manifesta una tendenza simile con un impatto negativo sulla predisposizione dei rispondenti, suggerendo che anche questa forma di inclusione possa risultare meno intuitiva o più complessa da elaborare per i partecipanti. Contrastando con i risultati delle condizioni B e C, la condizione D, dove si adotta lo sdoppiamento, mostra un impatto positivo sulla predisposizione dei rispondenti. Questo coefficiente positivo riflette una maggiore accettazione o forse una più semplice integrazione di questa modalità linguistica nel contesto comunicativo, indicando che lo sdoppiamento potrebbe essere percepito come una forma di inclusione più naturale rispetto alle altre condizioni analizzate.

Per poter includere nel modello anche le componenti di incertezza e dispersione a favore di un'analisi più dettagliata, viene quindi successivamente stimato un modello CUBE, modello più complesso definito nella sezione 2.2.2 e che permette di tenere conto di diverse componenti rispetto alle scelte dei soggetti. Viste le limitate covariate a disposizione, ma anche la volontà dello studio di esaminare in profondità l'effetto della condizione sperimentale, viene utilizzata quest'ultima come variabile dipendente per tutte e 3 le componenti da analizzare. Questa scelta risulta anche logica rispetto allo studio. Nei modelli CUBE, il parametro di feeling rappresenta l'inclinazione o predisposizione dei partecipanti verso specifiche categorie di risposta e, in questo contesto, è evidente come la condizione sperimentale possa influenzare notevolmente queste inclinazioni. Parallelamente, l'incertezza nei modelli CUBE misura il grado di certezza percepito dai partecipanti nel fornire le loro risposte. Variazioni nelle condizioni sperimentali possono compromettere la chiarezza dell'informazione disponibile, aumentando l'incertezza nelle rispo-

ste. Infine, la dispersione valuta l'eterogeneità delle risposte tra i soggetti, riflettendo come le diverse interpretazioni delle condizioni testate possano influenzare la coerenza delle risposte. Data inoltre la distribuzione delle risposte tra le diverse condizioni riportate nella tabella 3.9, può essere ragionevole considerare la condizione sperimentale come parametro relativo alla dispersione. Quindi, fatte queste considerazioni, viene adattato un modello CUBE in cui si la condizione sperimentale viene utilizzata per specificare ognuna delle 3 diverse componenti. Nella Tabella 3.11 viene riportato l'output del modello stimato.

| Parametri              | Stime     | Std.Err. | Wald       | p-value |
|------------------------|-----------|----------|------------|---------|
| Uncertainty            |           |          |            |         |
| Costante               | -1.938135 | 1.217989 | -1.591258  | 0.112   |
| Condizione B           | 4.189452  | 1.931940 | 2.168521   | 0.030   |
| Condizione C           | 4.177006  | 1.938900 | 2.154317   | 0.031   |
| Condizione D           | -1.823498 | 0.908682 | -2.006750  | 0.045   |
| Feeling                |           |          |            |         |
| Costante               | 1.599908  | 0.080418 | 19.894987  | <0.001  |
| Condizione B           | -2.917877 | 0.203122 | -14.365137 | <0.001  |
| Condizione C           | -2.529832 | 0.173897 | -14.547916 | <0.001  |
| Condizione D           | 1.406414  | 0.144222 | 9.751727   | <0.001  |
| Overdispersion         |           |          |            |         |
| Costante               | -0.789058 | 0.146710 | -5.378345  | <0.001  |
| Condizione B           | 4.588525  | 2.049609 | 2.238732   | 0.025   |
| Condizione C           | 4.479093  | 2.905283 | 1.541706   | 0.023   |
| Condizione D           | -2.929022 | 0.928244 | -2.078141  | 0.038   |
| <b>Log-likelihood:</b> | -2522.203 |          |            |         |
| <b>AIC:</b>            | 5068.406  |          |            |         |
| <b>BIC:</b>            | 5139.586  |          |            |         |

Tabella 3.11: Modello CUBE

Nel modello CUBE esaminato, i risultati evidenziano la variazione delle percezioni e delle reazioni dei partecipanti in relazione alle diverse condizioni sperimentali. Nella sezione relativa all'incertezza, la costante mostra una stima negativa che non risulta statisticamente significativa, indicando che la condizione di base non modifica in modo sostanziale il livello di incertezza percepito dai partecipanti. Al contrario, le condizioni B e C registrano un incremento significativo dell'incertezza, come dimostrato dalle stime positive e significative. Questo suggerisce che in queste condizioni si è generato un aumento dell'incertezza tra i partecipanti. D'altra parte, la condizione

D mostra una riduzione dell'incertezza, sottolineando il potenziale di questa configurazione nel fornire maggiore chiarezza e sicurezza nelle risposte.

Passando alla sezione "Feeling", la costante in questo caso rivela una notevole predisposizione positiva verso le risposte corrette nella condizione del maschile generico. Le condizioni B e C, tuttavia, presentano stime negative, riflettendo potenzialmente la confusione o la complessità introdotta dal linguaggio inclusivo usato in queste configurazioni. In netto contrasto, la condizione D mostra un effetto positivo, suggerendo che questa particolare modalità di presentazione del linguaggio potrebbe facilitare una migliore comprensione e accettazione da parte dei partecipanti.

Infine, nella sezione riguardante la dispersione, osserviamo che la condizione A presenta un coefficiente negativo, con un p-value  $<0.001$ , indicando una maggiore omogeneità nelle risposte per questa condizione. Le condizioni B e C mostrano dei coefficienti di overdispersion positivi e significativi, indicando un aumento significativo della dispersione delle risposte rispetto alla Condizione A. Infine, la Condizione D presenta un coefficiente negativo e significativo, indicando una diminuzione significativa della dispersione delle risposte rispetto alla Condizione A. Questo è coerente con la distribuzione delle frequenze, dove si osserva una forte concentrazione di risposte nella quarta categoria.

In conclusione, l'analisi del modello CUBE illustra come le modifiche nelle condizioni sperimentali influenzino in modi complessi e significativi l'incertezza, il sentimento e la dispersione delle risposte. Nello specifico, le condizioni B e C, associate a una maggiore incertezza e una minore predisposizione, suggeriscono che le modalità di linguaggio inclusivo adottate in queste condizioni possano creare confusione o difficoltà di comprensione tra i partecipanti. Al contrario, la condizione D, che mostra una minore incertezza e una maggiore predisposizione, indica che lo sdoppiamento del linguaggio è percepito come più chiaro e accettabile. L'aumento della dispersione delle risposte nelle condizioni B e C rispetto alla Condizione A evidenzia l'impatto delle varianti

linguistiche introdotte, suggerendo che il linguaggio inclusivo in queste configurazioni potrebbe essere più controverso o difficile da integrare. La Condizione D, mostrando una dispersione ridotta, è associata a una maggiore chiarezza percepita, indicando che questa modalità di linguaggio inclusivo è più facilmente accettabile per i partecipanti, poiché le risposte tendono ad essere più coerenti e meno variabili.

### **3.8 Conclusioni sullo studio**

L'analisi dei modelli adattati sui dati dello studio Psychopy ha fornito una comprensione approfondita degli effetti delle diverse condizioni sperimentali sulle risposte dei partecipanti. Dai risultati ottenuti dai modelli stimati, emerge chiaramente che le condizioni hanno influenzato significativamente le performance dei partecipanti.

La condizione A, utilizzata come riferimento, ha permesso di valutare gli effetti delle altre condizioni. Le condizioni B e C, rispettivamente schwa lunga e asterisco, hanno mostrato un impatto negativo sulle prestazioni, che seppur caratterizzate da tempi di risposta più brevi, comportano una minore probabilità di fornire risposte corrette rispetto alla condizione A. Questo suggerisce che tali condizioni aumentano la difficoltà o l'incertezza percepita dai partecipanti, determinando un peggioramento complessivo delle loro performance. Al contrario, la condizione D, lo sdoppiamento, ha manifestato un effetto positivo, facilitando risposte più rapide e accurate rispetto alla condizione A. Questa condizione sembra ridurre la complessità o migliorare la comprensione del compito, portando a una performance migliorata.

Il modello multilogit ha ulteriormente chiarito come le condizioni influenzino le probabilità di risposte corrette, errate o di incertezza. Le condizioni B e C hanno ridotto significativamente la probabilità di risposte corrette rispetto alla condizione A, mentre la condizione D ha aumentato questa probabilità, confermando la tendenza osservata nei modelli precedenti.

Il modello CUBE ha fornito ulteriori informazioni sulle dinamiche di incertezza, sentimento e dispersione nelle risposte. Le condizioni B e C hanno aumentato l'incertezza e ridotto il sentimento positivo rispetto alla condizione A, mentre la condizione D ha avuto l'effetto opposto, migliorando il sentimento positivo e riducendo l'incertezza. Questo modello ha anche evidenziato che le condizioni B e C aumentano significativamente la dispersione delle risposte, indicando una maggiore variabilità individuale nella percezione della difficoltà del compito.

Gli effetti casuali nei modelli hanno rivelato una significativa variabilità tra i partecipanti e, in misura minore, tra i temi. La varianza tra i partecipanti indica che esistono differenze individuali sostanziali nelle risposte, indipendentemente dalla condizione sperimentale. Questo potrebbe essere attribuibile a fattori intrinseci dei partecipanti, quali le differenze cognitive, esperienziali o motivazionali. La variabilità tra i temi, sebbene presente, è meno pronunciata, suggerendo che le caratteristiche dei temi influenzano le risposte in misura minore rispetto alle caratteristiche individuali dei partecipanti.

In sintesi, i risultati di questa ricerca sottolineano l'importanza delle condizioni sperimentali nel modellare le risposte dei partecipanti e la necessità di considerare le differenze individuali nelle analisi. La condizione A, utilizzata come benchmark, evidenzia come le condizioni B e C siano più difficili da leggere e interpretare, mentre la condizione D agevola la lettura.

# Capitolo 4

## Studio Eyetracking

Questo capitolo è dedicato all'analisi dei dati raccolti tramite eyetracking, utilizzando modelli lineari a effetti misti, con effetti casuali incrociati. Questi modelli rappresentano un approccio flessibile per l'analisi di dati che includono misurazioni ripetute. In particolare, i modelli implementati in questo studio sono stati configurati per gestire le dipendenze intra-soggetto e intra-tema che emergono dalla natura ripetuta delle osservazioni raccolte. Tale configurazione prevede la presenza di effetti casuali sia per il soggetto che per il tema, permettendo di controllare l'autocorrelazione interna e di catturare l'eterogeneità sia tra i soggetti che tra i temi studiati.

In aggiunta agli effetti casuali, il modello incorpora vari effetti fissi, che includono il numero di caratteri della frase, la condizione sperimentale, essenziale per testare le ipotesi specifiche dello studio, nonché variabili demografiche quali il sesso e l'età dei partecipanti. La condizione sperimentale è di particolare interesse, in quanto rappresenta la variabile indipendente principale attraverso cui si intendono esplorare le dinamiche e le risposte specifiche dei soggetti. Saranno sviluppati e valutati diversi modelli, differenziati in base alle tre aree di interesse e alle variabili dipendenti analizzate nel presente studio. Per ciascuna area di interesse (frase completa, prime due parole e quinta parola), verranno stimati modelli per il tempo totale di lettura, il tempo della prima fase di lettura e il tempo della fase di riletture, con tutti i valori normalizzati rispetto al numero di parole analizzate.

## 4.1 Valori anomali

In questa sezione si discutono le strategie adottate per la gestione dei valori anomali nel dataset, essenziali per garantire l'affidabilità e la validità delle analisi condotte. È fondamentale precisare che le definizioni di valori anomali adottate nel presente studio sono state elaborate dai linguisti. La presenza di valori anomali può infatti introdurre distorsioni significative nelle stime e influenzare l'interpretazione dei risultati. Pertanto è fondamentale identificare e trattare adeguatamente tali valori prima di procedere con ulteriori analisi. Le tre condizioni specifiche per identificare le osservazioni anomale valutando l'area d'interesse della frase completa sono:

- un'osservazione viene considerata anomala se il tempo di fissazione durante la prima fase di lettura è registrato pari a zero. Tale condizione può indicare un errore nella registrazione dei dati o una mancata attenzione da parte del partecipante, rendendo il dato non rappresentativo del comportamento di lettura tipico;
- un'osservazione è classificata come anomala se il tempo medio di lettura per parola, calcolato sia durante la prima lettura che durante le letture successive, è inferiore a 80 millisecondi. Valori così bassi possono suggerire una lettura superficiale o inattenta, o potenziali errori nella raccolta dei dati.
- si considera anomala un'osservazione se il valore medio di lettura per parola, considerando il tempo totale di fissazione sull'area d'interesse, supera gli 800 millisecondi. Questo può indicare una lettura eccessivamente lenta, che potrebbe essere attribuita a difficoltà di comprensione, a strategie di lettura particolarmente meticolose o a errori di registrazione.

Nel contesto del presente studio, la metodologia adottata per identificare e gestire le osservazioni anomale ha portato alla selezione di 79 osservazioni considerate come outliers. Di queste, una significativa maggioranza, preci-

samente 52 osservazioni, è stata identificata in base alla seconda condizione definita precedentemente. Questa condizione riguarda i tempi di lettura medi per parola anormalmente bassi, sia durante la prima lettura che durante le letture successive, che risultano inferiori a 80 millisecondi. Questo dato evidenzia un pattern specifico nel comportamento di lettura nell'area d'interesse della frase completa, dove una grande proporzione degli outliers è attribuibile a velocità di lettura eccezionalmente rapide, suggerendo che tali osservazioni possono rappresentare errori di registrazione o comportamenti di lettura non tipici che necessitano di essere esclusi per evitare distorsioni nelle analisi.

La rimozione di queste 79 osservazioni assicura dei risultati che riflettano comportamenti di lettura più rappresentativi dei partecipanti. Il totale di osservazioni per l'area d'interesse "Frase completa" sono 2814. Nella Tabella 4.1 viene rappresentata la frequenza degli outliers per le 3 condizioni.

| Condizione             | Numero di Outlier | Percentuale sul Totale (%) |
|------------------------|-------------------|----------------------------|
| 1                      | 13                | 0.46                       |
| 2                      | 52                | 1.85                       |
| 3                      | 14                | 0.50                       |
| <b>Totale Outliers</b> | <b>79</b>         | <b>2.81</b>                |

Tabella 4.1: Riepilogo degli Outlier per Condizione

## 4.2 Modelli per la frase completa

In questa sezione viene proposta un'analisi per la prima area di interesse analizzata, ovvero l'intera frase letta dai partecipanti. È stata condotta un'analisi esplorativa preliminare per esaminare la relazione tra le condizioni sperimentali e la prima variabile dipendente, ovvero il tempo di lettura normalizzato per numero di parole dell'intero item. La Figura 4.1 rappresenta la densità delle osservazioni per differenti condizioni sperimentali A, B, C e D rispetto al tempo di lettura pro parola, espresso in millisecondi.

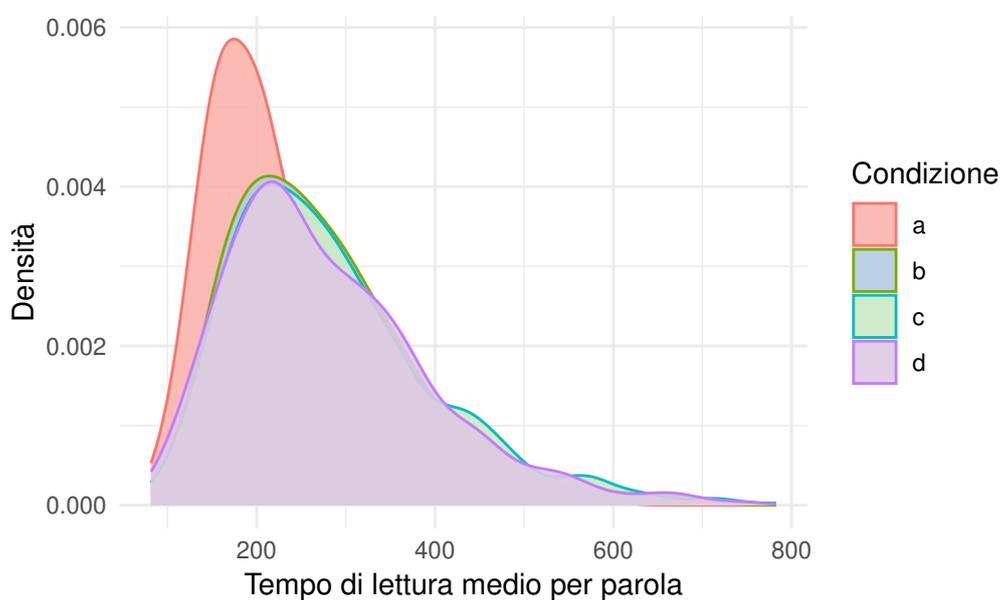


Figura 4.1: Distribuzione del tempo di lettura pro parola per condizione

Dalla figura è possibile notare come la condizione A mostri una distribuzione con una moda verso i tempi di lettura pro parola più brevi, il che indica che molti partecipanti hanno completato la lettura in tempi relativamente ridotti. Le condizioni B, C e D mostrano distribuzioni caratterizzate da code più lunghe, indicando una variazione maggiore nei tempi di lettura pro parola tra i partecipanti. Nelle sezioni successive verranno delineati e approfonditi i diversi modelli sviluppati per l'area di interesse presa in esame.

### 4.2.1 Modelli per tempo totale di lettura

L'analisi iniziale si concentra sulla valutazione del tempo totale di lettura dell'item, considerato come variabile dipendente e normalizzato rispetto al numero di parole. I modelli adattati in questa sezione sono modelli lineari a effetti misti, appropriati per la modellazione di variabili continue.

Il primo modello stimato include, oltre agli effetti casuali per partecipante e tema, degli effetti fissi per la variabile condizione, sesso, età, numero di caratteri oltre che alle interazioni tra numero di caratteri e condizione. I coefficienti stimati relativi a sesso ed età non risultano però significativi. Ciò significa che i maschi hanno un tempo di lettura medio che non differisce significativamente rispetto alle femmine. Una conclusione analoga è valida per l'età: infatti un aumento unitario dell'età del soggetto non comporta una variazione del tempo di lettura statisticamente significativa, per questo motivo i coefficienti legati a sesso ed età sono stati tolti dal modello. Viene quindi stimato un nuovo modello con effetti casuali per soggetto e tema, oltre che a un effetto fisso per la condizione, numero di caratteri e le interazioni tra le due. Essendo però tutte le interazioni non significative, si procede togliendo questi parametri a favore di un modello più parsimonioso, mostrato in Tabella 4.2.

| Predittori       | Stime | Errore standard | P-value |
|------------------|-------|-----------------|---------|
| Intercetta       | 76.45 | 37.4953         | 0.042   |
| Condizione B     | 54.07 | 4.3121          | <0.001  |
| Condizione C     | 62.94 | 4.3115          | <0.001  |
| Condizione D     | 39.72 | 6.8183          | <0.001  |
| Numero caratteri | 2.99  | 0.7576          | <0.001  |

| Effetti casuali          |         |
|--------------------------|---------|
| $\sigma^2$               | 6292.79 |
| $\tau_{00}$ Partecipante | 4661.05 |
| $\tau_{00}$ Tema         | 322.22  |

Tabella 4.2: Modello lineare a effetti misti incrociati

Si può notare come la varianza a livello di tema ( $\tau_{00}$  tema) sia molto inferiore rispetto a quella per soggetto, indicando una minore variabilità nel tempo di lettura tra i diversi temi. Per rappresentare graficamente la situazione degli effetti casuali, si propongono di seguito due Figure, 4.2 e 4.3, con le relative stime.

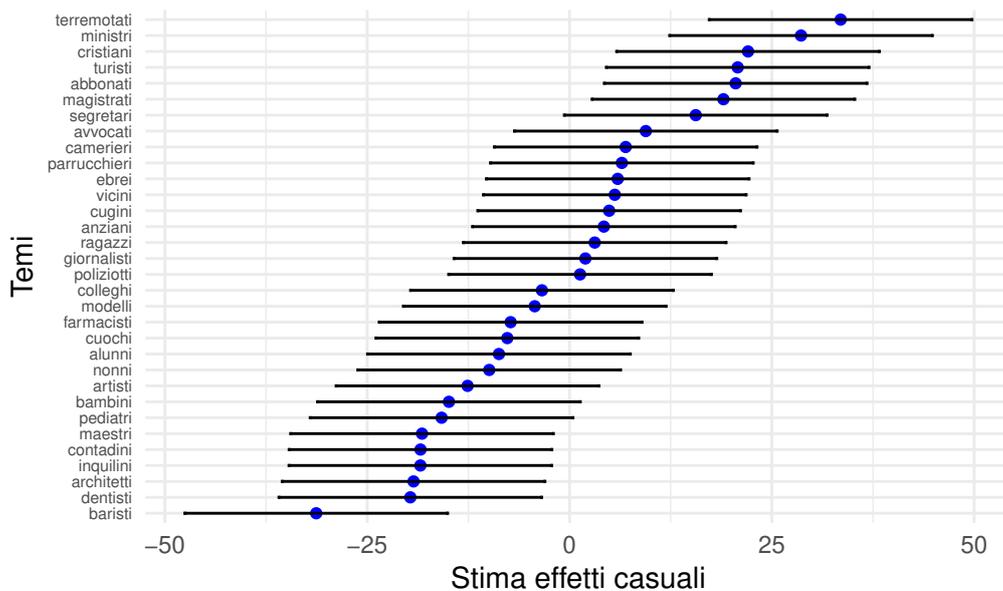


Figura 4.2: Effetti casuali per tema

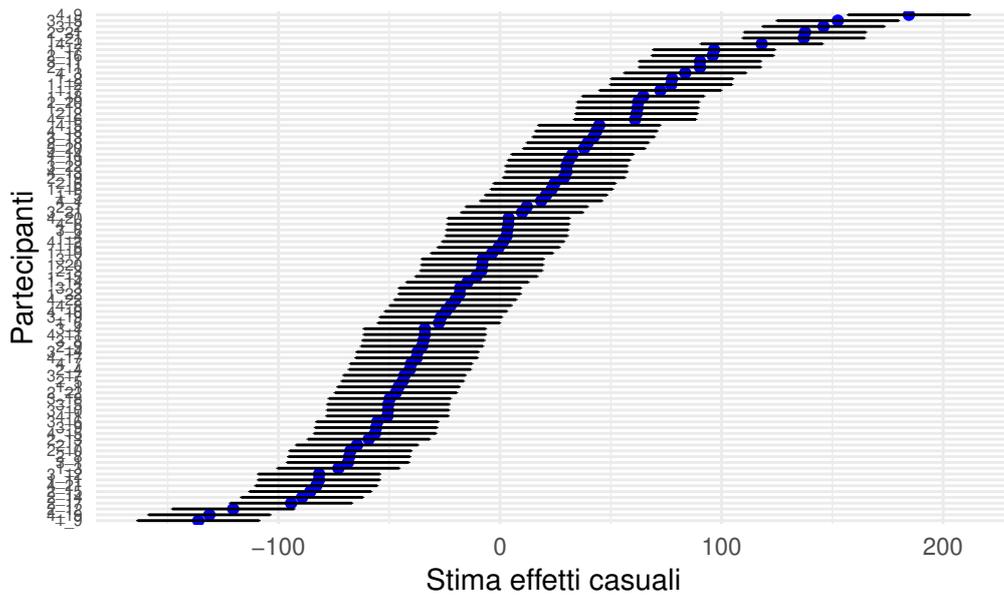


Figura 4.3: Effetti casuali per partecipanti

Dall'analisi di queste figure, emerge chiaramente come la varianza tra i partecipanti sia significativamente maggiore rispetto a quella tra i temi. Questo risultato rispecchia le stime dei parametri di varianza ottenute dal modello, confermando che la variabilità tra i partecipanti contribuisce in misura più rilevante alla variabilità complessiva dei dati. In considerazione di questi risultati, è stato proposto un modello più parsimonioso. Nello specifico, si suggerisce l'adozione di un modello ridotto che incorpori un effetto casuale unicamente per i partecipanti, eliminando l'effetto casuale associato ai temi, come riportato nella Tabella 4.3.

| Predittore               | Stima   | Standard Error | $p$    |
|--------------------------|---------|----------------|--------|
| Intercetta               | 75.41   | 18.4189        | <0.001 |
| Condizione B             | 54.26   | 4.4065         | <0.001 |
| Condizione C             | 63.12   | 4.4056         | <0.001 |
| Condizione D             | 39.57   | 5.0309         | <0.001 |
| Numero caratteri         | 3.01    | 0.3452         | <0.001 |
| Effetti casuali          |         |                |        |
| $\sigma^2$               | 6606.32 |                |        |
| $\tau_{00}$ Partecipante | 4648.43 |                |        |

Tabella 4.3: Modello a effetti misti, con effetto casuale per partecipante

Nel modello analizzato, l'intercetta ha una stima di 75.41 millisecondi, rappresentando il tempo medio di lettura per parola atteso nella condizione A. Questo valore è significativo, indicando un chiaro effetto nel tempo di lettura. La condizione B mostra un incremento medio del tempo di lettura per parola di 54.26 millisecondi rispetto alla condizione di base. La condizione C evidenzia un aumento più marcato, con una differenza media di 63.12 millisecondi rispetto alla condizione di base. Infine, la condizione D porta a un aumento del tempo di lettura per parola di 39.57 millisecondi rispetto alla condizione di base, confermando anch'essa un effetto significativo sul tempo di lettura. Il coefficiente associato al numero di caratteri è pari a 3.01, suggerendo che, per ogni carattere aggiuntivo in una parola, il tempo medio di lettura aumenta in media di 3.01 millisecondi. Questo effetto è significativo e contribuisce in modo rilevante alla variazione del tempo di lettura.

Questi risultati indicano che le diverse condizioni influenzano in modo significativo il tempo di lettura per parola, e che anche il numero di caratteri di una parola gioca un ruolo cruciale nella determinazione del tempo di lettura. Il parametro associato alla varianza tra individui è risultato particolarmente elevato, il che sottolinea l'esistenza di significative differenze interindividua-

li nei tempi di lettura. Questo parametro fa parte degli effetti casuali del modello e rappresenta la variabilità nei tempi di lettura che non può essere spiegata dagli effetti fissi, ma che è attribuibile alle caratteristiche uniche di ciascun soggetto.

Dopo aver stimato entrambi i modelli e considerando la natura annidata dei modelli stessi, si propone di eseguire un test del rapporto di verosimiglianza per valutare la significatività dell'effetto casuale associato ai temi. Questo test permetterà di determinare se l'inclusione dell'effetto casuale relativo ai temi migliora significativamente l'adattamento del modello rispetto a un modello più parsimonioso che considera solo l'effetto casuale per partecipante. I modelli, come descritto nella sezione 2.1.1, sono stimati utilizzando la funzione di verosimiglianza ristretta (REML). Dunque, non è possibile utilizzare il test del rapporto di verosimiglianza standard per valutare la significatività degli effetti casuali. Pertanto, si adotta il test del rapporto di verosimiglianza ristretto, come delineato da Scheipl, Greven e Küchenhoff (2008), per valutare l'aggiunta del parametro. Questo metodo considera le specifiche caratteristiche dei modelli a effetti misti, garantendo una valutazione accurata dell'effetto casuale associato ai temi. L'approccio proposto assicura che l'analisi tenga conto delle peculiarità dei modelli a effetti misti, fornendo una valutazione robusta e affidabile della significatività dell'effetto casuale aggiuntivo.

Nonostante la varianza dell'effetto casuale associato ai temi non sia elevata, il test del rapporto di verosimiglianza ristretto ha indicato un p-value inferiore a 0.001, suggerendo una preferenza per il modello completo. Questo risultato evidenzia che anche una varianza apparentemente ridotta può influenzare in modo significativo la variabilità complessiva dei dati. È quindi essenziale riconoscere l'importanza di considerare tali effetti nel contesto dell'analisi. Di conseguenza, si adotterà questo approccio per tutti i modelli successivi, valutando singolarmente e complessivamente ogni effetto casuale. Questa metodologia garantirà un'analisi rigorosa e dettagliata, permettendo di identificare e includere gli effetti casuali che contribuiscono significativamente alla

variabilità complessiva del modello.

## 4.2.2 Modelli per la prima fase di lettura

Viene quindi analizzata la seconda variabile dipendente, ovvero quella che rappresenta il tempo normalizzato per parola per la prima fase di lettura dell'item. È stata eseguita un'analisi esplorativa preliminare per indagare la relazione tra le condizioni sperimentali e il tempo per la prima fase di lettura. La Figura 4.4 mostra la distribuzione delle osservazioni per le diverse condizioni sperimentali (A, B, C e D) in relazione alla variabile dipendente considerata.

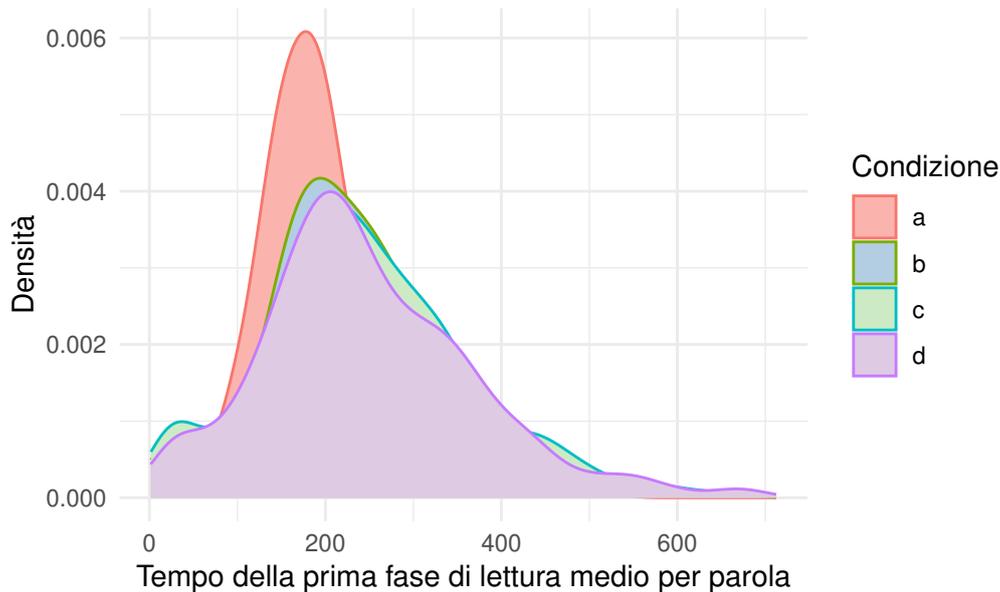


Figura 4.4: Distribuzione del tempo per prima fase di lettura per condizione

Dalla figura si osserva che la distribuzione per la condizione A presenta una moda verso tempi per la prima fase di lettura per parola più brevi, suggerendo che un numero significativo di partecipanti ha completato la lettura in tempi relativamente ridotti. Al contrario, le distribuzioni per le condizioni B, C e D mostrano code più lunghe, indicanti una maggiore dispersione dei tempi di lettura per parola tra i partecipanti. Questo evidenzia una variabilità più ampia nei tempi di lettura associata a queste condizioni.

Per la stima dei modelli seguenti è stata adottata una logica sequenziale

come quella impiegata nelle analisi precedenti. In particolare, ogni variabile è stata valutata singolarmente per determinare il suo effetto specifico e il suo contributo alla spiegazione della variabilità dei dati, permettendo di identificare gli effetti significativi delle variabili considerate. La valutazione sistematica delle variabili ha contribuito a costruire modelli che assicurano al contempo la parsimonia e l'efficacia del modello finale.

Inizialmente, vengono inclusi nel modello, oltre agli effetti casuali per partecipante e tema, il numero di caratteri, la condizione sperimentale, le interazioni tra condizione e numero di caratteri e alcune variabili demografiche come il sesso e l'età. Tuttavia, queste sono state successivamente rimosse a seguito della loro mancanza di significatività statistica. Questa decisione ha permesso di semplificare il modello senza compromettere l'integrità dei risultati. In seguito, vengono eliminate anche le interazioni tra la variabile relativa alla condizione sperimentale e il numero dei caratteri, parametri tutti non significativamente diversi da 0. Viene quindi riportato in Tabella 4.4 l'output del modello stimato con effetti fissi la condizione sperimentale e il numero di caratteri, oltre al doppio effetto casuale.

| Predittori               | Stime   | Standard Error | <i>p</i> – <i>value</i> |
|--------------------------|---------|----------------|-------------------------|
| Intercetta               | 75.24   | 37.2950        | 0.044                   |
| Condizione B             | 42.16   | 4.8673         | <0.001                  |
| Condizione C             | 50.08   | 4.8665         | <0.001                  |
| Condizione D             | 35.86   | 7.1744         | <0.001                  |
| Numero caratteri         | 2.47    | 0.7556         | 0.001                   |
| Effetti casuali          |         |                |                         |
| $\sigma^2$               | 8028.71 |                |                         |
| $\tau_{00}$ Partecipante | 3813.65 |                |                         |
| $\tau_{00}$ Tema         | 293.75  |                |                         |

Tabella 4.4: Stima del modello lineare a effetti misti, con effetti casuali incrociati

Dall'analisi della varianza degli effetti casuali emerge chiaramente come la varianza tra i partecipanti sia significativamente maggiore rispetto a quella tra i temi, come confermato come confermato dalle Figure 4.5 e 4.6.

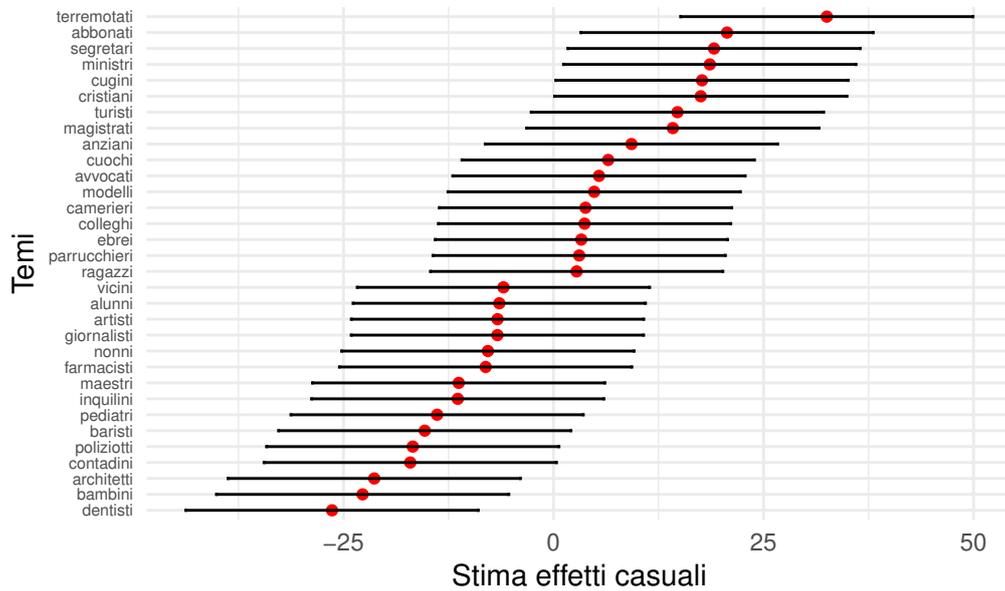


Figura 4.5: Effetti casuali per tema

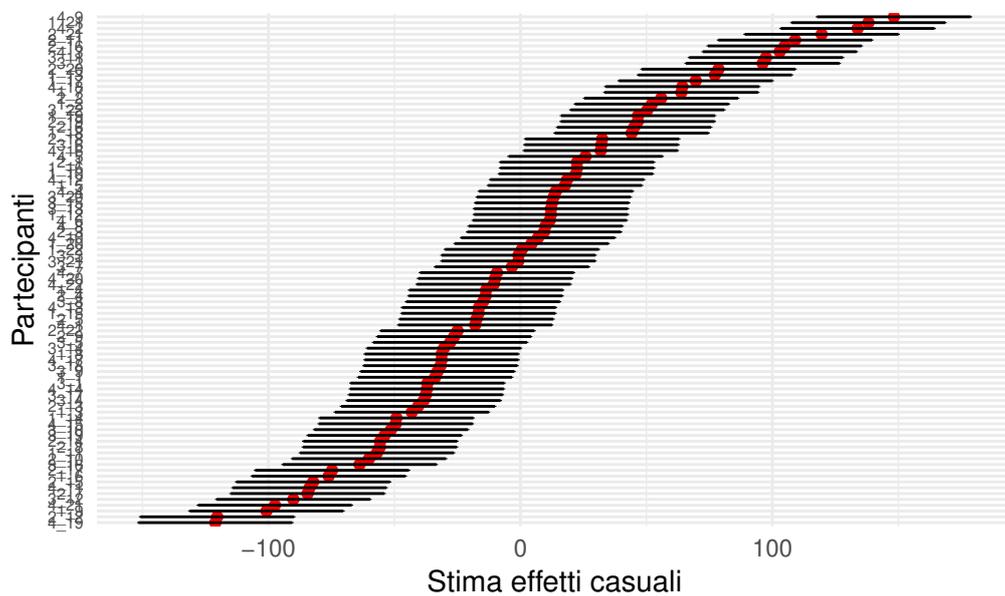


Figura 4.6: Effetti casuali per partecipanti

Seguendo la logica metodologica adottata nei modelli precedenti, è stato esaminato l'effetto della rimozione dell'effetto casuale associato ai temi dal modello. I risultati di questa analisi sono riportati nella Tabella 4.5.

| Predittori               | Stime   | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|---------|----------------|-------------------------|
| Intercetta               | 86.02   | 20.0904        | <0.001                  |
| Condizione B             | 42.35   | 4.9430         | <0.001                  |
| Condizione C             | 50.23   | 4.9420         | <0.001                  |
| Condizione D             | 37.31   | 5.6434         | <0.001                  |
| Numero caratteri         | 2.24    | 0.3872         | <0.001                  |
| Effetti casuali          |         |                |                         |
| $\sigma^2$               | 8313.10 |                |                         |
| $\tau_{00}$ Partecipante | 3802.12 |                |                         |

Tabella 4.5: Modello lineare a effetti misti, con effetto casuale per partecipante

Nel modello analizzato, l’intercetta ha una stima di 86.02 millisecondi, rappresentando il tempo medio della prima fase di lettura per parola atteso nella condizione A, risultando significativamente diverso da zero. I coefficienti associati alle altre condizioni indicano differenze medie rispetto alla condizione di base. La condizione B mostra un incremento medio del tempo di lettura per parola di 42.35 millisecondi rispetto alla condizione di base. Analogamente, la condizione C evidenzia un aumento medio di 50.23 millisecondi. Infine, la condizione D porta a un incremento del tempo di lettura di 37.31 millisecondi rispetto alla condizione di base.

Questi risultati confermano l’effetto significativo delle condizioni sperimentali sul tempo della prima fase di lettura, evidenziando variazioni rilevanti rispetto alla condizione di base. Il coefficiente stimato per il numero di caratteri risulta pari a 2.24, indicando che un aumento unitario del numero di caratteri comporta un aumento del tempo della prima fase di lettura pro parola pari a 2.24 millisecondi. Il parametro associato alla varianza tra individui è risultato particolarmente elevato, il che sottolinea l’esistenza di significative differenze interindividuali nelle prime fasi di lettura, rappresentando la variabilità nei

tempi delle prime fasi di lettura che non possono essere spiegate dalle variabili fisse considerate nel modello, ma che sono attribuibili alle caratteristiche uniche di ciascun soggetto. Viene inoltre valutato il miglior modello tramite test del rapporto di verosimiglianza ristretta, che porta a favorire il modello con effetti casuali incrociati, con un p-value  $< 0.001$ , dimostrando che anche una varianza apparentemente bassa ha un impatto significativo sulla variabilità complessiva dei dati.

### 4.2.3 Modelli per la rilettura

In questa sezione viene analizzata la terza variabile dipendente, ovvero quella che riguarda il tempo di rilettura per parola dell'item. È stata eseguita un'analisi esplorativa preliminare per indagare la relazione tra le condizioni sperimentali e il tempo della rilettura. La Figura 4.7 mostra la distribuzione delle osservazioni per le diverse condizioni sperimentali in relazione alla variabile dipendente considerata.

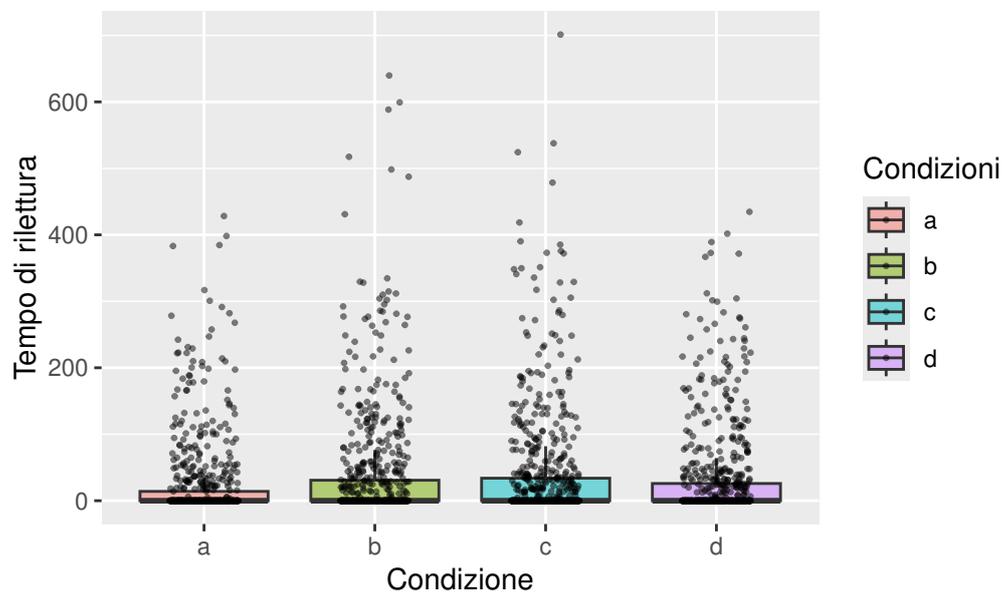


Figura 4.7: Distribuzione del tempo di rilettura per condizione

Il box plot rappresenta la distribuzione del tempo di rilettura per parola suddiviso per quattro condizioni sperimentali. La distribuzione dei punti offre un'idea visiva della densità dei dati: una maggiore concentrazione di punti

indica una maggiore frequenza dei valori in quell'intervallo. La condizione A presenta una concentrazione di tempi di rilettura più brevi, evidenziata da un'alta densità vicino alla base del box. Le condizioni B, C e D, invece, mostrano mediane pari alla condizione A ma una maggiore dispersione nei tempi di rilettura.

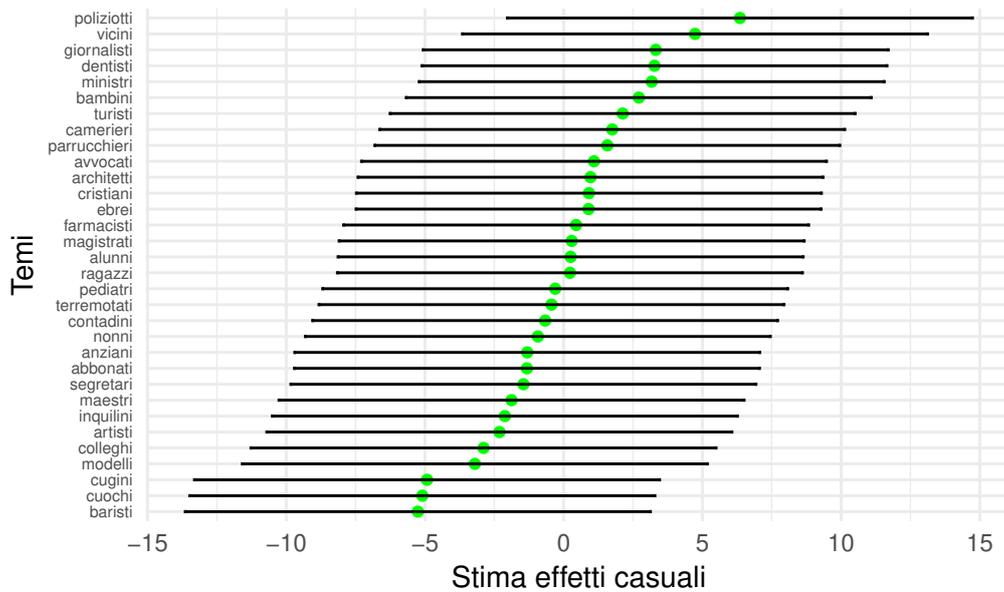
Per la stima dei modelli, viene seguita la metodologia precedentemente adottata. Inizialmente, vengono incluse nel modello le variabili relative al sesso e l'età. Tuttavia, queste variabili sono state rimosse a causa della loro mancanza di significatività. Questa decisione porta alla semplificazione del modello senza compromettere l'integrità dei risultati.

Successivamente, sono state eliminate anche le interazioni tra la variabile relativa alla condizione sperimentale e il numero dei caratteri, poiché tutti i parametri associati non erano significativamente diversi da zero ( $p\text{-value} > 0.5$ ). Questo processo porta alla costruzione di un modello più parsimonioso, mantenendo solo i predittori che contribuiscono significativamente alla spiegazione della variabilità dei dati. Di seguito, nella Tabella 4.6, viene riportato l'output del modello con effetti casuali incrociati.

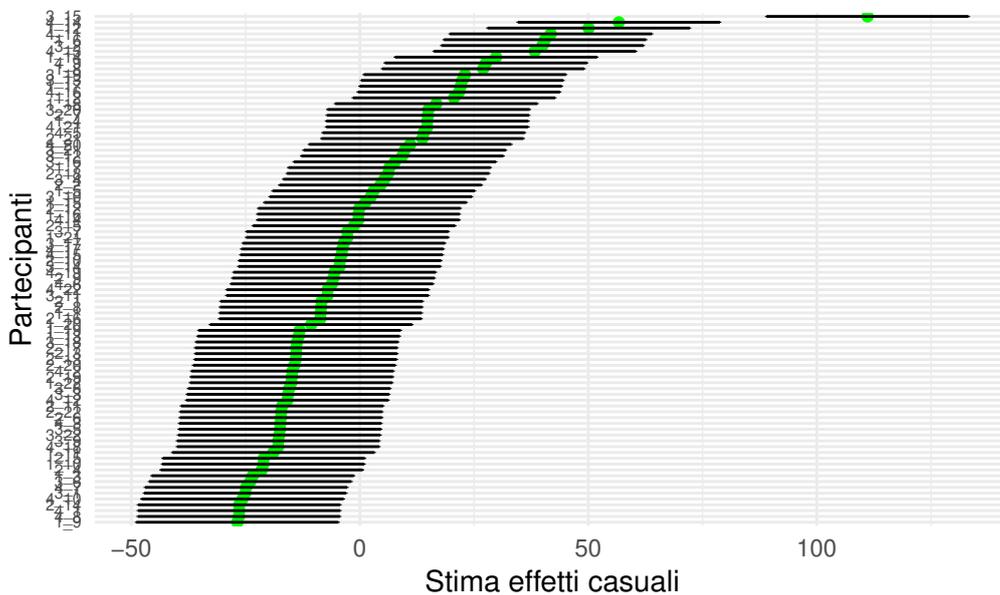
| Predittori               | Stime   | Standard Error | $P - value$ |
|--------------------------|---------|----------------|-------------|
| Intercetta               | -10.15  | 17.7936        | 0.568       |
| Condizione B             | 11.95   | 3.8545         | 0.002       |
| Condizione C             | 12.98   | 3.8537         | 0.001       |
| Condizione D             | 2.12    | 4.6069         | 0.645       |
| Numero caratteri         | 0.74    | 0.3600         | 0.041       |
| Effetti casuali          |         |                |             |
| $\sigma^2$               | 5052.26 |                |             |
| $\tau_{00}$ Partecipante | 645.02  |                |             |
| $\tau_{00}$ Tema         | 25.67   |                |             |

Tabella 4.6: Modello a effetti misti incrociati per la rilettura

L'analisi della varianza degli effetti casuali rivela che la varianza attribuibile alle differenze tra i partecipanti è significativamente maggiore rispetto a quella associata ai temi. Questo risultato è ulteriormente supportato dall'evidenza fornita dalla Figura 4.8.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 4.8: Effetti casuali per temi e partecipanti

Dalla figura emerge chiaramente che la varianza tra i partecipanti è significativamente superiore a quella tra i temi. Dati questi risultati, viene proposto

un modello più parsimonioso. In particolare, viene stimato un modello ridotto che includa solo un effetto casuale per i partecipanti, eliminando l'effetto casuale associato ai temi. In Tabella 4.7 viene riportato l'output del modello adattato.

| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | -11.5370 | 15.0794        | 0.444                   |
| Condizione B             | 11.98    | 3.8627         | 0.002                   |
| Condizione C             | 13.01    | 3.8619         | 0.001                   |
| Condizione D             | 1.97     | 4.4098         | 0.656                   |
| Numero caratteri         | 0.76     | 0.3026         | 0.012                   |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 5077.16  |                |                         |
| $\tau_{00}$ Partecipante | 644.00   |                |                         |

Tabella 4.7: Modello lineare misto con intercetta casuale per partecipante

L'analisi del modello relativo al tempo di rilettura rivela che non emergono differenze significative tra le condizioni sperimentali A e D. Le stime dei coefficienti per le condizioni B e C invece risultano statisticamente diverse da zero rispetto alla condizione di base. Il numero di caratteri inoltre risulta statisticamente significativo a un livello di confidenza del 5%, indicando che un aumento unitario del numero di caratteri comporta un aumento significativo del tempo di rilettura. Per quanto riguarda gli effetti casuali, la varianza tra i partecipanti è notevolmente superiore rispetto a quella osservata tra i temi, indicando che la componente di variabilità attribuibile ai partecipanti gioca un ruolo significativamente più rilevante nella determinazione della variabilità complessiva rispetto alla variabilità imputabile ai temi. Viene inoltre valutato il miglior modello tramite test del rapporto di verosimiglianza ristretta, che, differentemente dai casi precedenti, porta a preferire un modello con singolo effetto casuale per partecipante, fornendo un pvalue di 0.1747, accettando

l'ipotesi che la componente di varianza per tema sia pari a zero.

## 4.3 Modelli per i tempi di lettura delle parole iniziali

Finora, l'analisi si è concentrata sull'intero item. Ogni item è strutturato seguendo regole costanti, in cui le prime due parole sono determinate dalle condizioni linguistiche sperimentali considerate. Per ottenere una comprensione più dettagliata di come i partecipanti elaborano queste informazioni visivamente, questa sezione esamina i tempi di lettura normalizzati per le prime due parole. L'obiettivo è capire come i tempi di lettura variano specificamente per le parole scritte con un linguaggio inclusivo.

L'attenzione alle prime due parole è giustificata dal fatto che esse rappresentano una parte critica dell'input linguistico che i partecipanti devono processare. Analizzando i tempi di lettura per queste parole, è possibile ottenere informazioni più precise su come il linguaggio inclusivo influisce sulla velocità della lettura. Inoltre, questo approccio permette di isolare gli effetti delle condizioni linguistiche sperimentali e di comprendere meglio le dinamiche cognitive coinvolte nel processo linguistico.

Come per l'area d'interesse precedente, anche in questa verranno modellate le tre diverse variabili dipendenti, ovvero totale del tempo di lettura, tempo per la prima fase di lettura e tempo per la fase di rilettura.

### 4.3.1 Modelli per tempo totale di lettura

La prima variabile dipendente esaminata per l'area relativa alle prime due parole dell'item è il tempo totale di lettura, inteso come la somma della fase iniziale di lettura e delle eventuali riletture. Viene proposta un'analisi esplorativa volta a relazionare questa variabile dipendente con le diverse condizioni sperimentali adottate, sulle quali si intende focalizzare l'attenzione. La Figura 4.9 riporta la distribuzione delle risposte per ciascuna delle condizioni sperimentali analizzate.

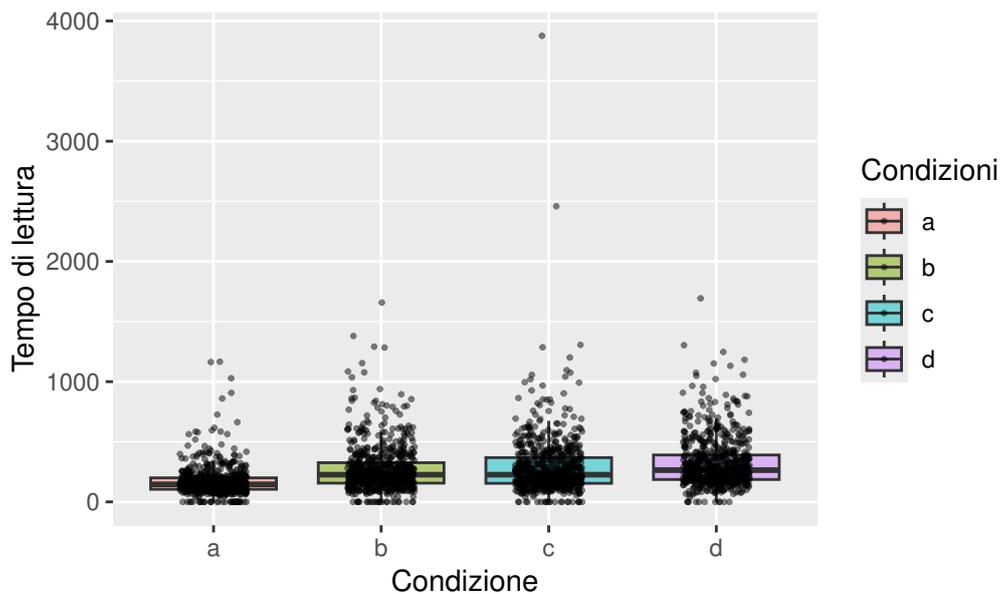


Figura 4.9: Distribuzione del tempo di lettura per condizione

I dati sono rappresentati attraverso boxplot con punti sovrapposti, permettendo di visualizzare sia la dispersione dei dati sia alcune statistiche di base. Le mediane dei tempi di lettura per le diverse condizioni sono relativamente simili, indicando una consistenza nei tempi medi di lettura attraverso le condizioni. Tuttavia, si osserva una variabilità notevole all'interno di ciascuna condizione. Questo suggerisce che, sebbene la maggior parte dei tempi di lettura si concentri intorno a valori simili, ci sono casi isolati che richiedono un tempo significativamente maggiore, potenzialmente indicando difficoltà o particolarità nel processamento delle parole in alcune specifiche condizioni, come la B e la C. La distribuzione complessiva dei dati evidenzia quindi una certa omogeneità tra le condizioni, pur con variazioni individuali notevoli.

Successivamente, vengono stimati modelli lineari a effetti misti per una comprensione più approfondita dei dati. Il primo modello stimato tiene conto, oltre che agli effetti casuali per partecipante e tema, degli effetti fissi per la variabile condizione, sesso, età, numero di caratteri e le interazioni tra condizione sperimentale e numero di caratteri. Dall'analisi emerge che i coefficienti relativi alle stime di sesso ed età non risultano significativi, con p-value rispettivamente pari a 0.204 e 0.733. Questo implica che il tempo medio di lettura

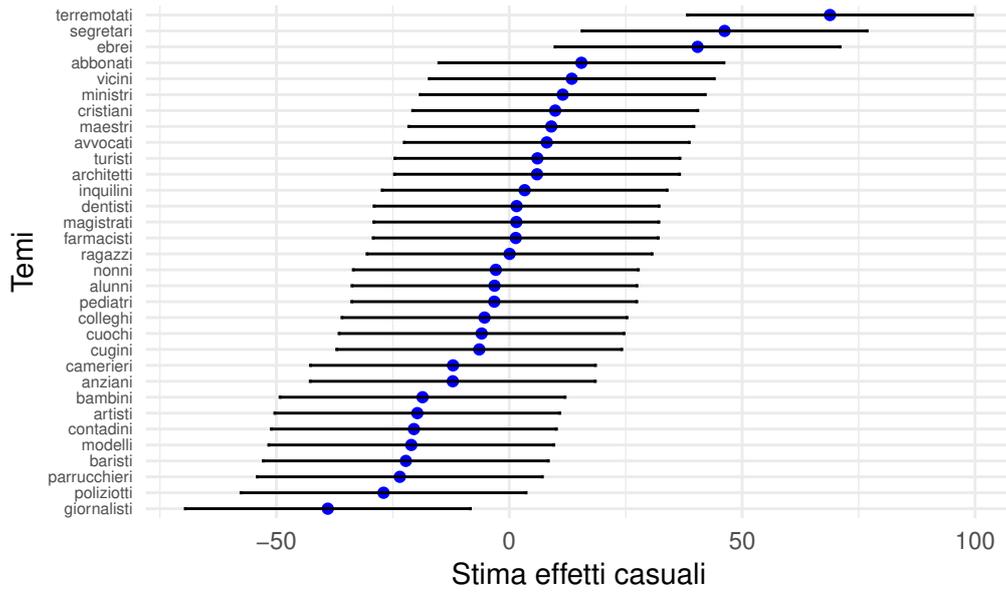
per parola dei partecipanti di sesso maschile non differisce significativamente da quello delle partecipanti di sesso femminile. Analogamente, l'età non mostra un effetto significativo sul tempo di lettura: un incremento unitario dell'età dei partecipanti non determina una variazione statisticamente significativa del tempo di lettura. Pertanto, queste variabili sono state escluse dal modello.

Successivamente, è stato stimato un nuovo modello con effetti casuali per soggetto e tema, oltre a un effetto fisso per la condizione sperimentale e le interazioni tra condizione e numero di caratteri. Tuttavia, anche queste interazioni non risultano significativamente diverse da zero a un livello di confidenza del 5%, riportando l'output in Tabella 4.8.

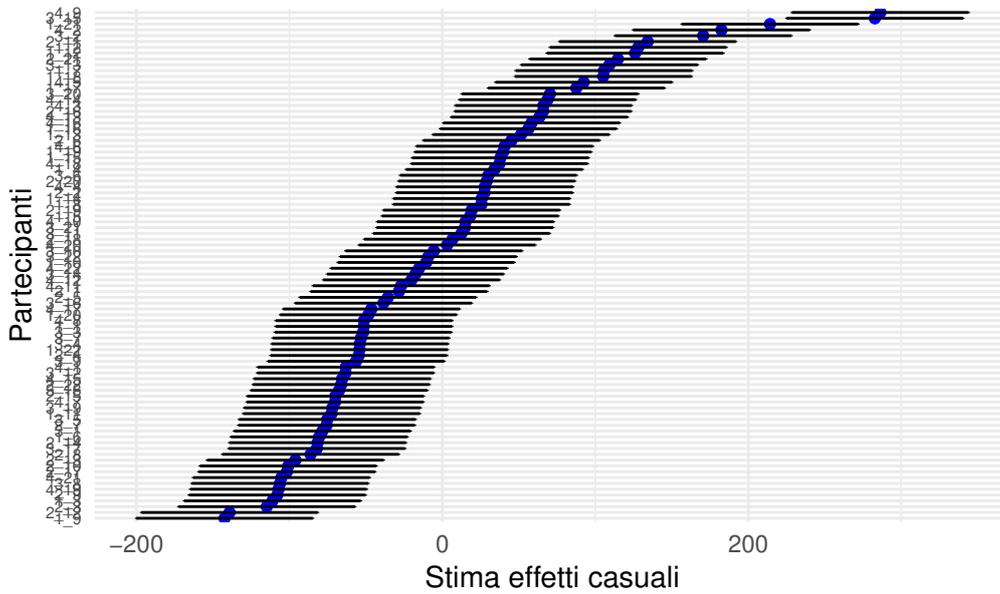
| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | 36.47    | 31.816         | 0.252                   |
| Condizione B             | 97.11    | 9.386          | <0.001                  |
| Condizione C             | 113.64   | 9.399          | <0.001                  |
| Condizione D             | 81.22    | 17.971         | <0.001                  |
| Numero caratteri         | 13.98    | 3.053          | <0.001                  |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 30183.41 |                |                         |
| $\tau_{00}$ Partecipante | 8785.31  |                |                         |
| $\tau_{00}$ Tema         | 719.95   |                |                         |

Tabella 4.8: Modello lineare a effetti misti, con effetti casuali incrociati

Si può notare come la varianza a livello di tema ( $\tau_{00}$  tema) sia molto inferiore rispetto a quella per soggetto, indicando una minore variabilità nel tempo di lettura tra i diversi temi, come si può verificare dall'analisi grafica riportata in Figura 4.10.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 4.10: Rappresentazione grafica degli effetti casuali

Viste le considerazioni fatte riguardo l'effetto casuale per tema, seguendo anche l'approccio utilizzato fin'ora, viene adattato un modello con singolo effetto casuale per partecipante, come riportato in Tabella 4.9.

| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | 35.78    | 21.441         | 0.095                   |
| Condizione B             | 97.17    | 9.414          | <0.001                  |
| Condizione C             | 113.52   | 9.428          | <0.001                  |
| Condizione D             | 80.86    | 13.278         | <0.001                  |
| Numero caratteri         | 14.06    | 1.865          | <0.001                  |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 30887.29 |                |                         |
| $\tau_{00}$ Partecipante | 8764.67  |                |                         |

Tabella 4.9: Modello lineare a effetti misti con effetto casuale per partecipante

Nel modello analizzato, l'intercetta ha una stima di 35.78 millisecondi, indicando il tempo medio di lettura per parola atteso nella condizione A, risultando significativo solo a un livello di confidenza del 10%. Gli altri coefficienti mostrano differenze medie rispetto alla condizione di base: la condizione B presenta un incremento medio del tempo di lettura di 97.17 millisecondi, mentre la condizione C evidenzia un aumento medio di 113.52 millisecondi e la condizione D un aumento medio di 80.86 millisecondi. Tutte queste differenze sono statisticamente significative, indicando un effetto rilevante delle condizioni sperimentali sul tempo di lettura rispetto alla condizione di base.

Il coefficiente stimato per il numero di caratteri risulta significativo, con un valore di 14.06, suggerendo che un aumento unitario del numero di caratteri porta a un incremento del tempo di lettura per parola di 14.06 millisecondi. Il parametro associato alla varianza sottolinea l'esistenza di significative differenze interindividuali nei tempi di lettura. Questo parametro, appartenente agli effetti casuali del modello, rappresenta la variabilità nei tempi di lettura non spiegata dalle variabili fisse considerate nel modello, ma attribuibile alle caratteristiche uniche di ciascun soggetto.

Dato che si tratta di modelli annidati, è stato eseguito un test di verosi-

miglianza ristretta per determinare se l'effetto casuale associato ai temi degli item possa essere considerato nullo o se sia appropriato includerlo nel modello per spiegare parte della variabilità dei dati. Il test confronta i modelli suggerendo una preferenza per il modello completo con effetti casuali incrociati ( $p\text{-value} < 0.001$ ), sottolineando l'importanza dell'inclusione dell'effetto casuale per tema.

### 4.3.2 Modelli per la prima fase di lettura

In questa sezione viene analizzata la seconda variabile dipendente nell'area d'interesse riguardante le prime due parole dell'item, ovvero il tempo della prima fase di lettura normalizzato per il numero di parole. Come valutato precedentemente, anche per questa variabile si svolge una breve analisi esplorativa, relazionando quest'ultima con le condizioni sperimentali analizzate, come riportato nella Figura 4.11.

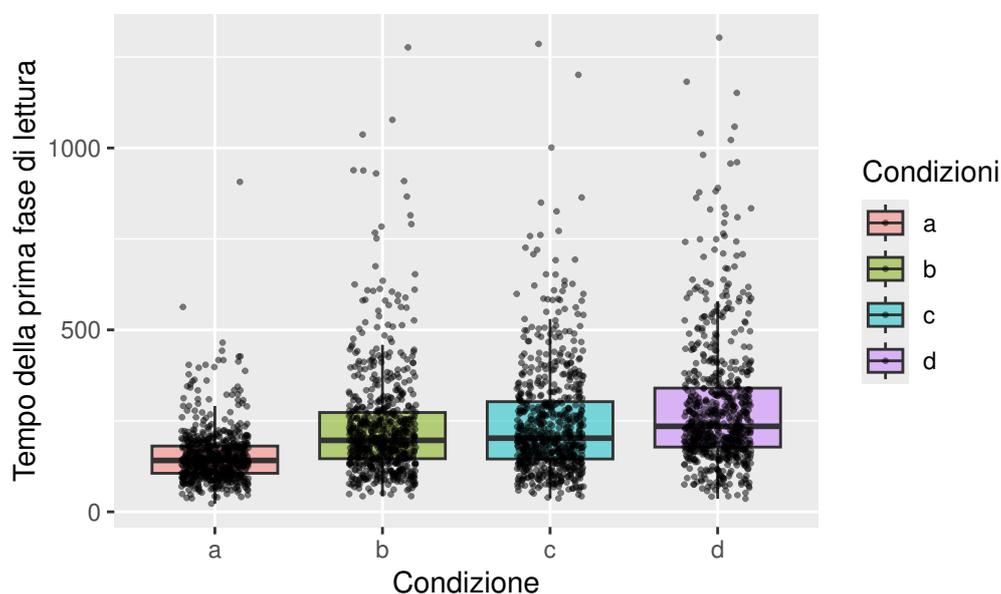


Figura 4.11: Distribuzione della prima fase di lettura per condizione

Dalla figura di evince che le mediane dei tempi di lettura nella prima fase mostrano variazioni minime tra le condizioni, con la condizione A che presenta una mediana inferiore rispetto alle altre. La maggiore dispersione dei tempi di lettura nelle condizioni B, C e D, rispetto alla condizione A, è indi-

cativa di una maggiore variabilità interna, suggerendo che queste condizioni influenzano maggiormente i tempi di lettura iniziali. La presenza di numerosi valori estremi nelle condizioni B e C sottolinea ulteriormente la complessità o difficoltà aggiuntiva nel processamento delle prime due parole in queste condizioni, mentre i valori elevati nella condizione D potrebbero essere attribuiti alla maggiore lunghezza dei caratteri inclusivi.

Per la stima dei modelli nel presente studio, viene adottata una strategia sequenziale, seguendo la logica metodologica precedentemente delineata. Inizialmente, il modello ha incluso variabili come condizione sperimentale, sesso, età, numero di caratteri, le interazioni tra condizione sperimentale e numero di caratteri, oltre a due intercette casuali per soggetto e tema. I coefficienti relativi a sesso ed età sono risultati non significativi ( $p\text{-value} = 0.231$  e  $0.757$ ), portando alla loro esclusione dal modello per semplificarne la struttura senza comprometterne l'integrità. Anche i coefficienti delle interazioni non sono risultati significativamente diversi da zero. Il primo modello proposto per il tempo della prima fase di lettura assume quindi come effetti fissi la condizione sperimentale analizzata e il numero di caratteri delle parole, oltre che gli effetti casuali incrociati per partecipante e tema. Nella Tabella 4.10 viene riportato l'output del modello appena definito.

| Predittore       | Stima | Standard Error | <i>P</i> – <i>value</i> |
|------------------|-------|----------------|-------------------------|
| Intercetta       | 47.14 | 24.451         | 0.050                   |
| Condizione B     | 76.48 | 7.027          | <0.001                  |
| Condizione C     | 86.61 | 7.030          | <0.001                  |
| Condizione D     | 80.65 | 13.785         | <0.001                  |
| Numero caratteri | 11.01 | 2.367          | <0.001                  |

| Effetti casuali          |          |
|--------------------------|----------|
| $\sigma^2$               | 16145.46 |
| $\tau_{00}$ Partecipante | 4378.35  |
| $\tau_{00}$ Tema         | 462.45   |

Tabella 4.10: Modello lineare a effetti misti, con effetti casuali incrociati

La varianza associata ai temi risulta significativamente inferiore rispetto a quella tra soggetti, indicando una minore variabilità nei tempi di lettura tra i diversi temi. Questo fenomeno è ulteriormente confermato dalle Figure 4.12 e 4.13.

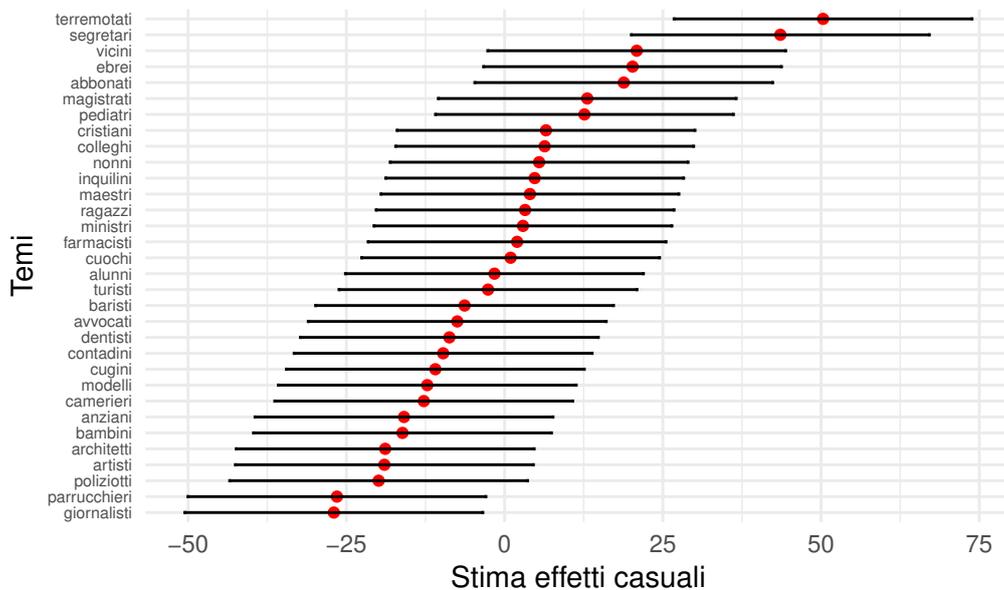


Figura 4.12: Effetti casuali per tema

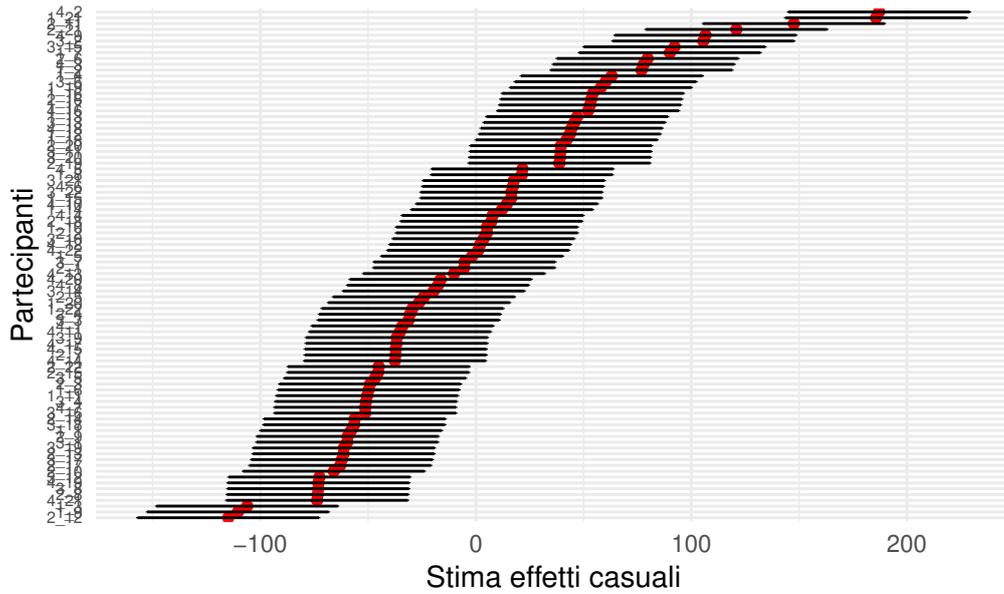


Figura 4.13: Effetti casuali per partecipanti

Per le considerazioni di cui sopra, viene quindi stimato un modello con effetto casuale esclusivamente per i partecipanti. I risultati vengono riportati nella Tabella 4.11.

| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | 55.16    | 15.803         | <0.001                  |
| Condizione B             | 76.99    | 7.057          | <0.001                  |
| Condizione C             | 87.06    | 7.061          | <0.001                  |
| Condizione D             | 84.69    | 9.910          | <0.001                  |
| Numero caratteri         | 10.19    | 1.389          | <0.001                  |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 16599.78 |                |                         |
| $\tau_{00}$ Partecipante | 4350.88  |                |                         |

Tabella 4.11: Modello lineare a effetti misti, con effetto casuale per partecipante

Nel modello analizzato, l'intercetta ha una stima di 55,16 millisecondi, indicando il tempo medio della prima fase di lettura atteso nella condizione A, risultando statisticamente significativo. Gli altri coefficienti mostrano dif-

ferenze medie rispetto alla condizione di base: la condizione B presenta un incremento medio del tempo di lettura di 76,99 millisecondi, mentre la condizione C evidenzia un aumento medio di 87,06 millisecondi e la condizione D un incremento medio di 84,69 millisecondi. Questi risultati confermano che tutte le condizioni sperimentali considerate hanno un effetto significativo sul tempo della prima fase di lettura pro parola rispetto alla condizione di base. Inoltre, il parametro relativo al numero di caratteri è significativo, con una stima di 10,19, indicando che un aumento unitario del numero di caratteri comporta un incremento del tempo della prima fase di lettura. Il parametro associato alla varianza tra individui risulta elevato, il che sottolinea l'esistenza di significative differenze interindividuali nelle prime fasi di lettura.

Confrontando infine i due modelli stimati, essendo i modelli annidati, è stato condotto un test di verosimiglianza ristretta per valutare se l'effetto casuale associato ai temi degli item fosse nullo o se fosse opportuno includerlo nel modello per spiegare parte della variabilità dei dati. I risultati del test hanno indicato una preferenza per il modello completo con effetti casuali incrociati, evidenziando l'importanza di includere l'effetto casuale per tema per una descrizione accurata della variabilità dei dati.

### 4.3.3 Modelli per la riletture

Viene valutata la terza variabile dipendente, quella riguardante l'eventuale tempo di riletture delle prime due parole. La distribuzione della risposta viene di seguito relazionata con le condizioni sperimentali analizzate nello studio, come riportato nella Figura 4.14.

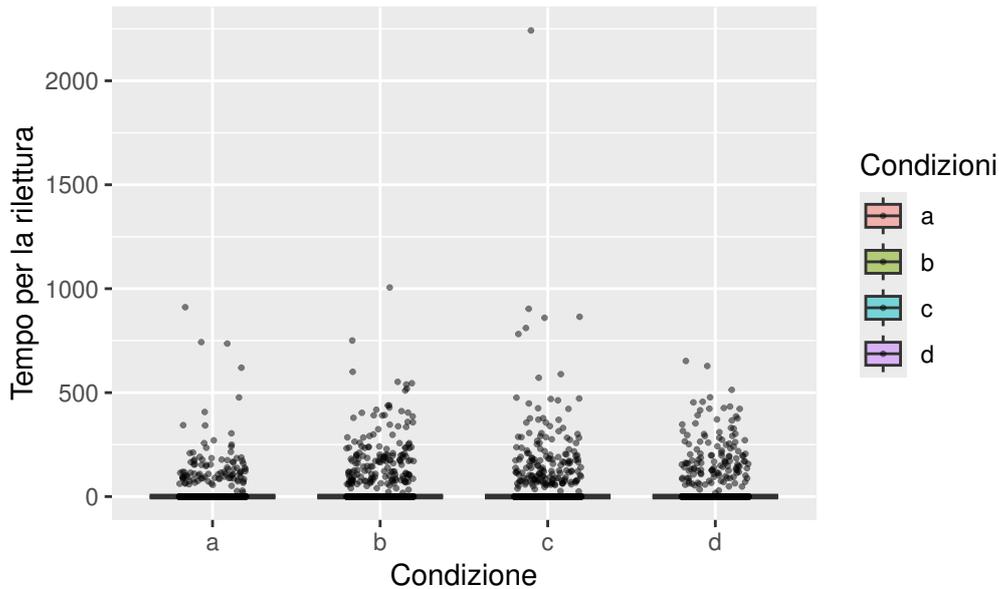


Figura 4.14: Distribuzione dei tempi di riletture per condizione

È importante notare la presenza di numerosi valori pari a zero, indicando che molti soggetti non hanno effettuato riletture. Questa caratteristica è visibile in tutte le condizioni, suggerendo che la riletture non è un comportamento comune per la maggior parte dei partecipanti. Le mediane dei tempi di riletture sono molto basse per tutte le condizioni, riflettendo la tendenza generale a non rileggere o a farlo per brevi periodi. Tuttavia, si osservano alcuni valori estremi, in particolare nelle condizioni B e C, dove alcuni soggetti hanno impiegato tempi di riletture significativamente più lunghi. Questo suggerisce che, sebbene la maggior parte dei partecipanti non rilegga, esistono individui che necessitano di tempi aggiuntivi per processare le informazioni linguistiche in queste specifiche condizioni.

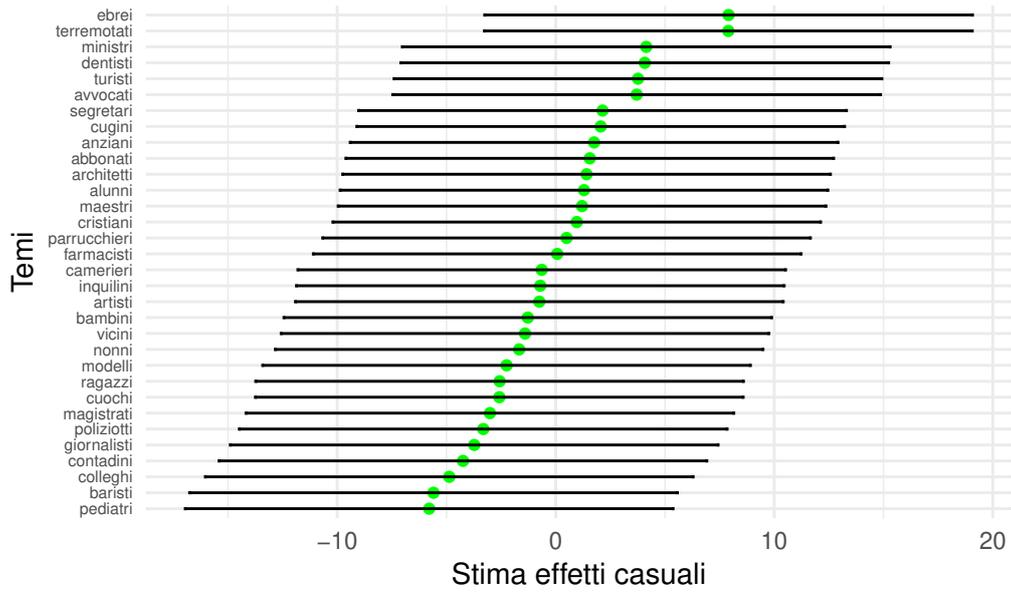
Per la stima dei modelli è stata seguita la logica metodologica precedentemente adottata. Inizialmente, il modello includeva variabili come sesso ed

età. Tuttavia, queste variabili sono state rimosse in seguito alla loro mancanza di significatività statistica (p-value = 0.732 e 0.467). Questa decisione ha permesso di semplificare il modello senza comprometterne l'integrità e la robustezza dei risultati. Inoltre, anche i coefficienti delle interazioni tra condizione sperimentale e numero di caratteri sono risultati non significativamente diversi da zero, suggerendo che le interazioni non apportavano un contributo rilevante al modello complessivo. Viene quindi stimato un modello con effetti fissi la condizione sperimentale e il numero di caratteri, oltre che gli effetti casuali incrociati per tema e partecipanti, riportando l'output nella Tabella 4.12.

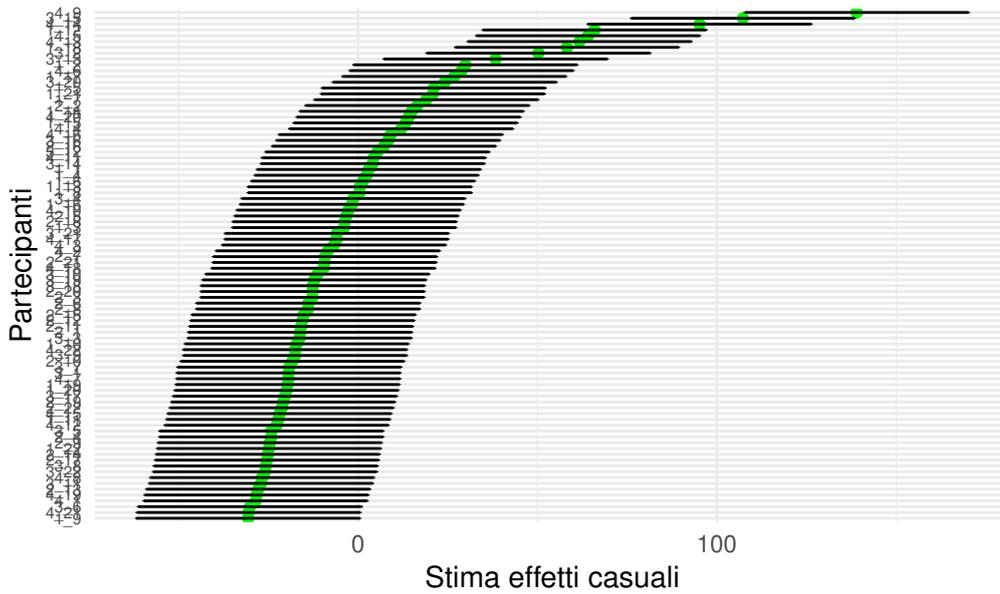
| Predittori               | Stime    | Standard Error | P-value |
|--------------------------|----------|----------------|---------|
| Intercetta               | -2.34    | 13.052         | 0.858   |
| Condizione B             | 21.17    | 5.402          | <0.001  |
| Condizione C             | 23.74    | 5.410          | <0.001  |
| Condizione D             | 0.83     | 8.255          | 0.920   |
| Numero caratteri         | 2.64     | 1.243          | 0.034   |
| Effetti casuali          |          |                |         |
| $\sigma^2$               | 10137.09 |                |         |
| $\tau_{00}$ Partecipante | 1235.99  |                |         |
| $\tau_{00}$ Tema         | 44.58    |                |         |

Tabella 4.12: Modello lineari a effetti misti per la rilettura, con effetti casuali incrociati

La varianza associata ai temi è risultata notevolmente inferiore rispetto a quella tra soggetti, indicando che i tempi di lettura variano meno tra i diversi temi rispetto a quanto variano tra i singoli soggetti. Questo risultato è ulteriormente supportato dai grafici in Figura 4.15.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 4.15: Rappresentazione grafica degli effetti casuali

Per le considerazioni di cui sopra, viene quindi stimato un modello più parsimonioso con effetto fisso la condizione sperimentale, numero di caratteri e effetto casuale per soggetto, riportato nella Tabella 4.13.

| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | -3.01    | 11.5189        | 0.794                   |
| Condizione B             | 21.13    | 5.4045         | <0.001                  |
| Condizione C             | 23.71    | 5.4125         | <0.001                  |
| Condizione D             | 0.49     | 7.6225         | 0.949                   |
| Numero caratteri         | 2.71     | 1.0705         | 0.012                   |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 10180.46 |                |                         |
| $\tau_{00}$ Partecipanti | 1234.50  |                |                         |

Tabella 4.13: Modello lineare a effetti misti, con effetto casuale per partecipante

L'analisi del modello relativo al tempo di rilettura indica l'assenza di differenze significative tra le condizioni sperimentali A e D. Al contrario, le stime dei coefficienti per le condizioni B e C risultano significativamente diverse rispetto alla condizione A. Il numero di caratteri emerge come una variabile statisticamente significativa, suggerendo che un aumento unitario nel numero di caratteri comporta un incremento significativo del tempo di rilettura per parola. Tuttavia, l'intercetta del modello non significativa complica ulteriori interpretazioni dei risultati. Infine, il parametro associato alla varianza tra individui risulta elevato, evidenziando l'esistenza di differenze interindividuali rilevanti nei tempi di rilettura.

Inoltre, è stato valutato il miglior modello tramite il test del rapporto di verosimiglianza ristretta. A differenza dei casi precedenti, questo test ha indicato una preferenza per un modello con un singolo effetto casuale per partecipante, con un p-value di 0.2296, supportando l'ipotesi che la componente di varianza per tema possa essere rimossa in quanto irrilevante per spiegare la variabilità riguardante i tempi di rilettura.

## 4.4 Modelli per la quinta parola dell'item

Dopo aver condotto un'analisi sull'intero item e sulle prime due parole di questi, l'attenzione si focalizza ora sulla quinta parola, anch'essa scritta utilizzando un linguaggio inclusivo. Ogni item è strutturato secondo regole costanti, con la quinta parola che assume un ruolo cruciale in questa fase dell'analisi.

L'obiettivo di questa sezione è esaminare come i partecipanti elaborano questa specifica parola, valutando i tempi di lettura. L'analisi approfondita della quinta parola, definita in modo inclusivo, consente di acquisire una comprensione più dettagliata dell'influenza del linguaggio inclusivo sulla velocità della lettura. Questo approccio permette di isolare gli effetti delle condizioni linguistiche sperimentali, migliorando la comprensione delle dinamiche cognitive coinvolte nel processo linguistico.

Analogamente alle analisi delle aree di interesse precedenti, anche in questa sezione verranno modellate tre diverse variabili dipendenti: il tempo della prima fase di lettura, il tempo della fase di rilettura e il tempo totale di lettura, dato dalla somma dei tempi precedenti. Questi modelli forniranno una comprensione più approfondita delle modalità con cui i partecipanti leggono la quinta parola dell'item, scritta con linguaggio inclusivo.

Sebbene ci sia la consapevolezza che l'analisi potrebbe risultare simile a quella precedente, l'obiettivo è osservare come variano questi tempi quando l'attenzione si sposta dall'inizio della frase a una parola situata nel mezzo della stessa.

### 4.4.1 Modelli per tempo totale di lettura

La prima variabile dipendente esaminata per l'area relativa alla quinta parola dell'item è il tempo totale di lettura, definito come la somma della fase iniziale di lettura e delle eventuali riletture. Viene presentata un'analisi esplorativa finalizzata a relazionare questa variabile dipendente con le diverse condizioni sperimentali presentate, sulle quali si desidera focalizzare l'attenzione. Di se-

guito, in Figura 4.16, viene rappresentata la distribuzione delle risposte per ciascuna delle condizioni sperimentali analizzate.

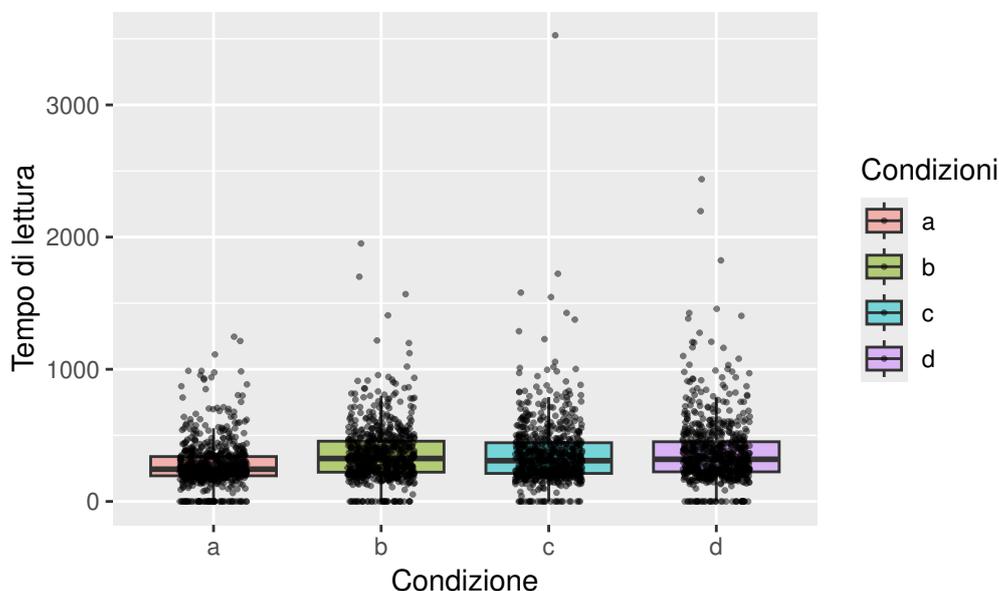


Figura 4.16: Effetti casuali per tema

Il grafico suggerisce che, sebbene i tempi di lettura medi siano relativamente costanti tra le condizioni, esiste una notevole variabilità individuale che potrebbe essere influenzata dalle specifiche caratteristiche delle condizioni sperimentali. In particolare, si osservano alcuni tempi di lettura molto elevati, soprattutto nelle condizioni B, C e D, suggerendo che alcuni partecipanti potrebbero aver incontrato difficoltà maggiori nel processare la quinta parola in queste condizioni.

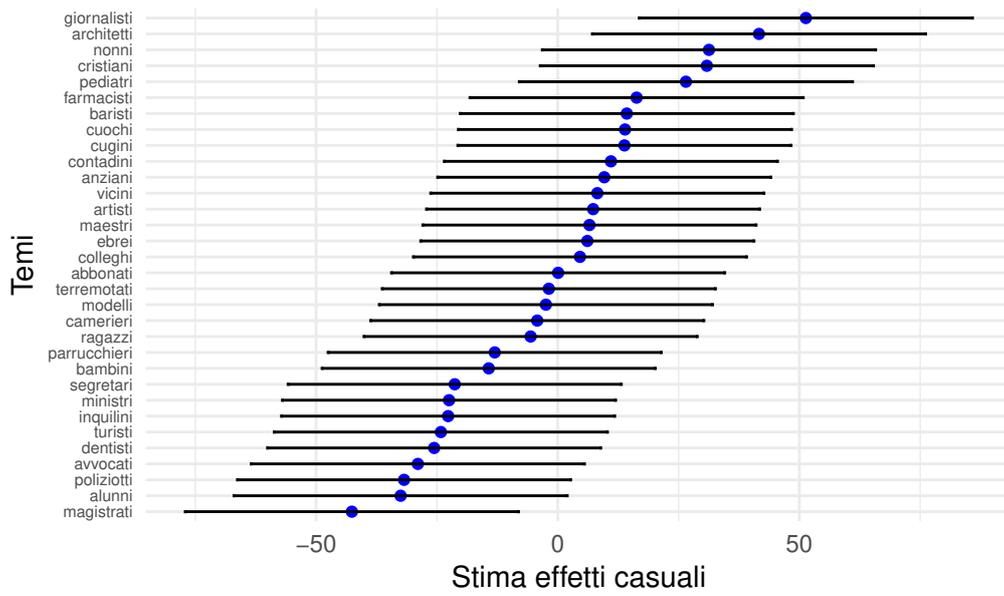
Successivamente, vengono stimati modelli lineari a effetti misti per una comprensione più approfondita dei dati. Il primo modello stimato include effetti casuali per partecipante e tema, e considera anche effetti fissi per le variabili condizione, sesso, età, numero di caratteri e le interazioni tra condizione sperimentale e numero di caratteri. I coefficienti relativi a sesso ed età non risultano significativi (p-value rispettivamente di 0.376 e 0.548), indicando che il tempo di lettura medio per parola non differisce significativamente tra maschi e femmine. Allo stesso modo, un aumento unitario della variabile età non produce una variazione statisticamente significativa nel tempo di lettura.

Pertanto, le variabili sesso ed età sono state rimosse dal modello. Successivamente, viene stimato un nuovo modello che mantiene gli effetti casuali per soggetto e tema, e include effetti fissi per la condizione e il numero di caratteri, senza considerare le interazioni tra queste ultime variabili. Le interazioni tra condizione sperimentale e numero di caratteri sono infatti risultate non significative per ogni parametro stimato, portando alla proposta di un modello più parsimonioso che esclude tali interazioni. Questo approccio permette di ottenere un modello più semplice e interpretabile, mantenendo l'integrità dei risultati e concentrandosi sulle variabili che hanno un impatto significativo sui tempi di lettura. Viene riportato nella Tabella 4.14 l'output del modello lineare definito, con effetti fissi le condizioni sperimentali e il numero di caratteri, oltre che agli effetti casuali incrociati per tema e partecipanti.

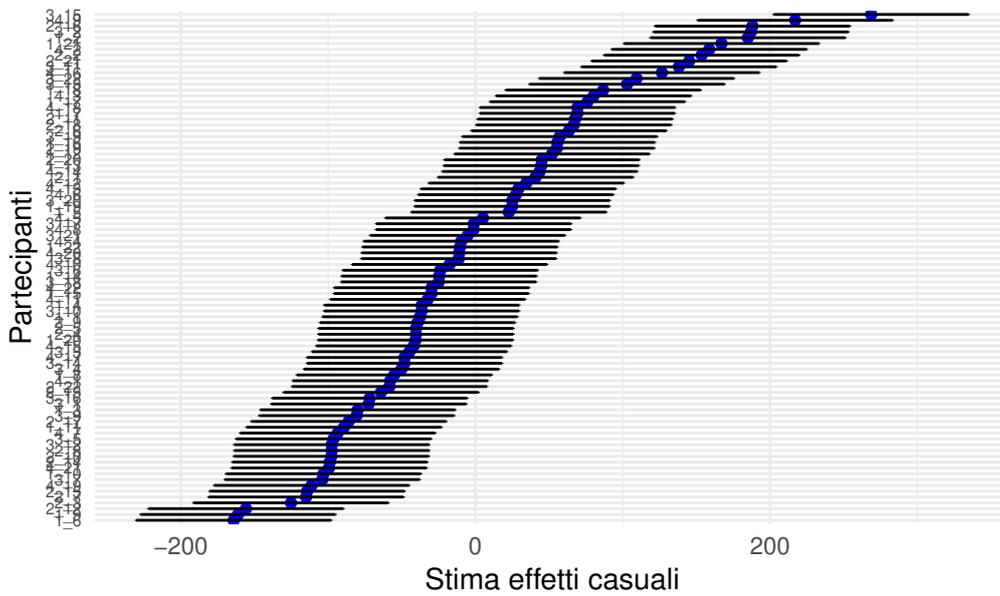
| Predittori               | Stime    | Standard Error | P-value |
|--------------------------|----------|----------------|---------|
| Intercetta               | 162.15   | 30.971         | <0.001  |
| Condizione B             | 82.38    | 10.732         | <0.001  |
| Condizione C             | 82.01    | 10.736         | <0.001  |
| Condizione D             | 69.72    | 12.319         | <0.001  |
| Numero caratteri         | 12.65    | 3.043          | <0.001  |
| Effetti casuali          |          |                |         |
| $\sigma^2$               | 40591.56 |                |         |
| $\tau_{00}$ Partecipanti | 9679.05  |                |         |
| $\tau_{00}$ Tema         | 835.29   |                |         |

Tabella 4.14: Modello lineare a effetti misti per il tempo di lettura della quinta parola

Si può notare come la varianza a livello di tema ( $\tau_{00}$  Tema) sia molto inferiore rispetto a quella per soggetto, indicando una minore variabilità nel tempo di lettura tra i diversi temi, come si può verificare dalla Figura 4.17.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 4.17: Rappresentazione degli effetti casuali

Viene quindi proposto nella Tabella 4.15, seguendo la logica usata per le aree d'interesse precedenti, un modello più parsimonioso includendo un singolo effetto casuale per partecipante.

| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | 167.86   | 21.31          | <0.001                  |
| Condizione B             | 82.57    | 10.84          | <0.001                  |
| Condizione C             | 81.81    | 10.84          | <0.001                  |
| Condizione D             | 70.97    | 11.45          | <0.001                  |
| Numero caratteri         | 12.02    | 1.86           | <0.001                  |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 41400.13 |                |                         |
| $\tau_{00}$ Partecipanti | 9654.88  |                |                         |

Tabella 4.15: Modello lineare a effetti misti per il tempo di lettura della quinta parola, con singolo effetto casuale per partecipante

Nel modello analizzato, l’intercetta ha una stima di 167.86 millisecondi, indicando il tempo medio di lettura per la quinta atteso nella condizione A. Le altre condizioni sperimentali mostrano variazioni medie rispetto alla condizione di base: la condizione B presenta un incremento medio di 82.57 millisecondi, la condizione C un incremento medio di 81.81 millisecondi e la condizione D un aumento medio di 70.97 millisecondi. Questi incrementi confermano un effetto significativo delle condizioni sperimentali sul tempo di lettura rispetto alla condizione di base. Inoltre, anche il coefficiente per il numero di caratteri è significativo, indicando che un aumento unitario nel numero di caratteri comporta un incremento del tempo di lettura per parola di 12.02 millisecondi. Il parametro associato alla varianza tra individui è risultato particolarmente elevato, il che sottolinea l’esistenza di significative differenze interindividuali nei tempi di lettura.

Dati i due modelli proposti, risulta interessante valutare se l’aggiunta dell’effetto casuale per tema risulta adeguata. Pertanto, considerando che si tratta di modelli annidati, viene eseguito un test del log-rapporto di verosimiglianza ristretta. I risultati di questo test indicano una preferenza per il modello con effetti casuali incrociati (p-value <0.001), suggerendo l’inclusione

dell'effetto casuale per tema.

#### 4.4.2 Modelli per la prima fase di lettura della quinta parola

Viene considerata la seconda variabile dipendente, rappresentante il tempo per la prima fase di lettura della parola. Viene presentato, in Figura 4.18, un grafico per un'analisi preliminare, valutando come questi tempi variano in base alla condizione sperimentale utilizzata.

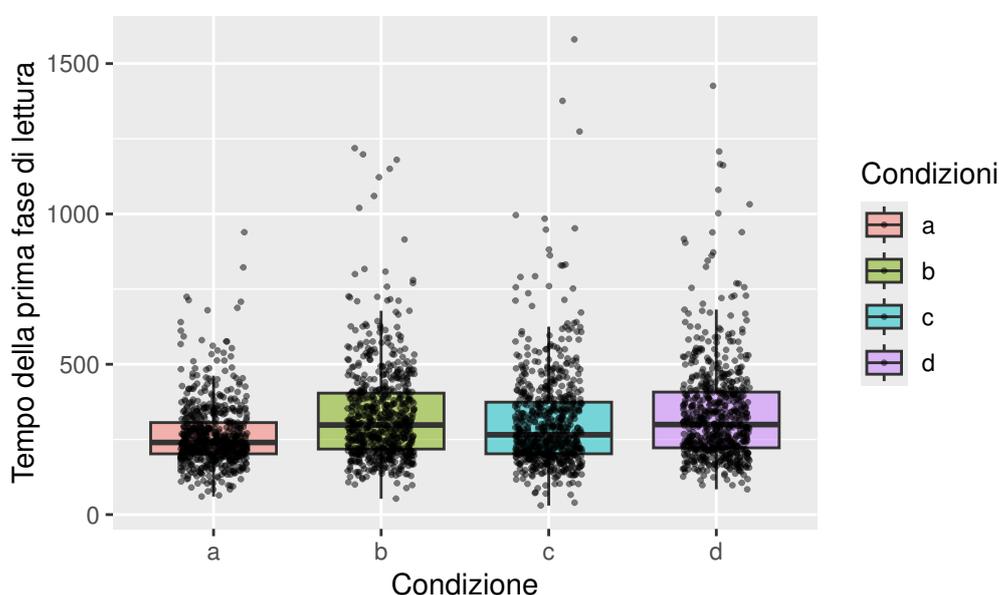


Figura 4.18: Distribuzione dei tempi della prima fase di lettura per distribuzione

Le mediane dei tempi di lettura sono comparabili tra le diverse condizioni, suggerendo una consistenza nei tempi medi di lettura iniziali. Tuttavia, si nota una variabilità considerevole all'interno di ciascuna condizione, come evidenziato dall'ampiezza delle scatole e dalla dispersione dei punti. Le condizioni B, C e D, in particolare, mostrano una maggiore dispersione, indicando che alcuni partecipanti hanno sperimentato tempi di lettura significativamente più lunghi.

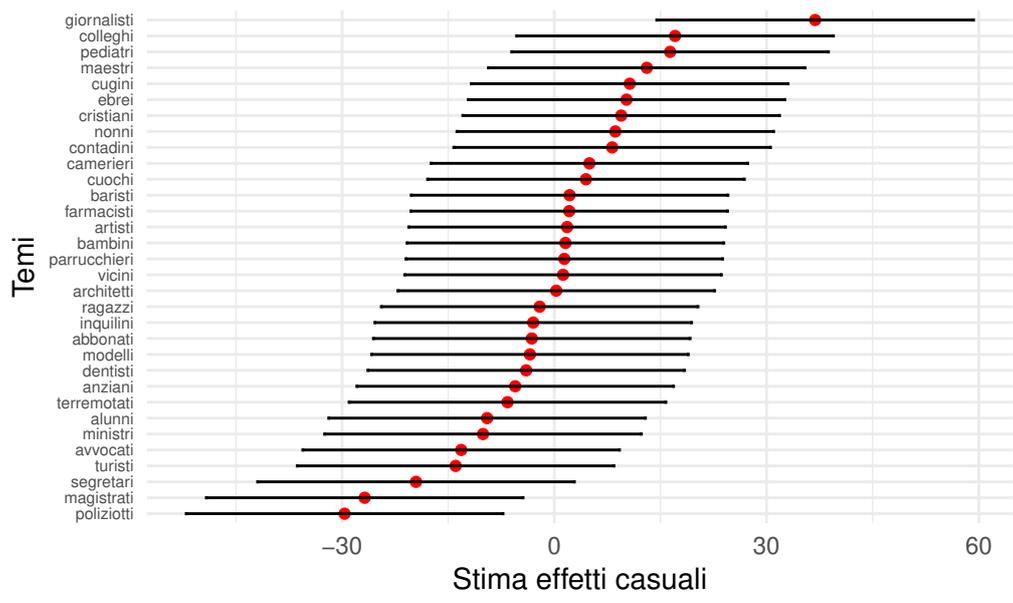
Per la stima dei modelli nel presente studio è stata adottata una strategia sequenziale, seguendo la logica metodologica precedentemente impiegata. Inizialmente, il modello ha incluso variabili quali la condizione sperimentale,

il sesso, l'età, il numero di caratteri, le interazioni tra condizione sperimentale e numero di caratteri, oltre alle due intercette casuali per soggetto e tema. Tuttavia, i coefficienti relativi a sesso ed età non sono risultati significativi (p-value = 0.200 e 0.773), portando alla loro esclusione dal modello a causa della mancanza di significatività statistica. Questa decisione ha consentito di semplificare il modello senza comprometterne l'integrità dei risultati. Inoltre, anche i coefficienti relativi alle interazioni tra condizione sperimentale e numero di caratteri non sono risultati significativamente diversi da zero. La Tabella 4.16 riporta i risultati del modello stimato.

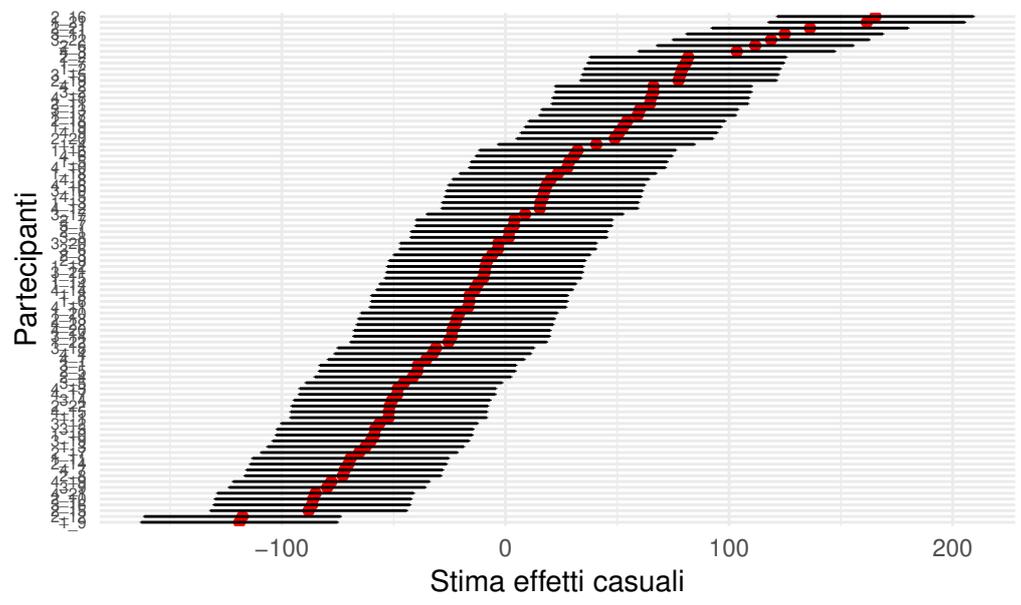
| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | 162.09   | 19.936         | <0.001                  |
| Condizione B             | 68.59    | 7.377          | <0.001                  |
| Condizione C             | 48.76    | 7.357          | <0.001                  |
| Condizione D             | 51.91    | 8.295          | <0.001                  |
| Numero caratteri         | 10.92    | 1.928          | <0.001                  |
| Effetti casuali          |          |                |                         |
| $\sigma^2$               | 17815.11 |                |                         |
| $\tau_{00}$ Partecipante | 4454.68  |                |                         |
| $\tau_{00}$ Tema         | 300.79   |                |                         |

Tabella 4.16: Modello lineare a effetti misti, con effetti casuali incrociati

Dall'analisi della varianza degli effetti casuali emerge chiaramente come la varianza tra i partecipanti sia significativamente maggiore rispetto a quella tra i temi, come confermato dall'analisi grafica:



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 4.19: Rappresentazione grafica degli effetti casuali

Per le considerazioni di cui sopra, viene quindi stimato un modello con effetto casuale esclusivamente per i partecipanti. I risultati vengono riportati nella Tabella 4.17.

| Predittori               | Stime    | Standar Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|---------------|-------------------------|
| Intercetta               | 163.77   | 14.650        | <0.001                  |
| Condizione B             | 68.80    | 7.434         | <0.001                  |
| Condizione C             | 48.70    | 7.415         | <0.001                  |
| Condizione D             | 52.32    | 7.832         | <0.001                  |
| Numero caratteri         | 10.75    | 1.271         | <0.001                  |
| Effetti casuali          |          |               |                         |
| $\sigma^2$               | 18105.95 |               |                         |
| $\tau_{00}$ Partecipante | 4446.83  |               |                         |

Tabella 4.17: Modello lineare a effetti misti per la prima fase di lettura, con effetto casuale per partecipante

Nel modello analizzato, l’intercetta è stimata a 163.77 millisecondi, rappresentando il tempo medio della prima fase di lettura atteso nella condizione A, risultando significativo. Le altre condizioni sperimentali mostrano incrementi medi nel tempo della prima fase di lettura rispetto alla condizione di base: 68.80 millisecondi per la condizione B, 48.70 millisecondi per la condizione C e 52.32 millisecondi per la condizione D, confermando l’effetto significativo delle condizioni sperimentali. Il coefficiente relativo al numero di caratteri, pari a 10.75, indica che un aumento unitario nel numero di caratteri comporta un incremento del tempo della prima fase di lettura. Infine, il parametro associato alla varianza tra individui risulta elevato, sottolineando significative differenze interindividuali nelle prime fasi di lettura.

Per valutare l’importanza del parametro relativo all’effetto casuale per tema viene definito il test del log rapporto di verosimiglianza ridotta vista la presenza di modelli annidati. L’analisi suggerisce un’incrementata adattabilità dei dati per il modello con effetti casuali incrociati, restituendo un p-value per il test proposto pari a 0.0003, preferendo quindi un modello con doppio effetto casuale.

### 4.4.3 Modelli per la rilettura

Viene esaminata la terza variabile dipendente, relativa al tempo di rilettura della quinta parola. La distribuzione di questa variabile viene analizzata in relazione alle diverse condizioni sperimentali considerate nello studio, come riportato nella Figura 4.20.

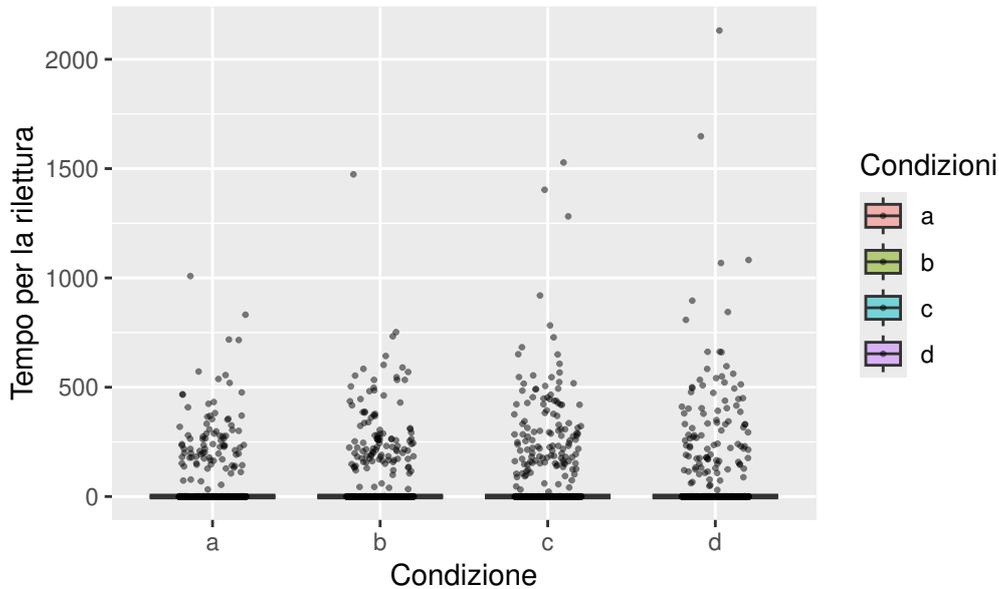


Figura 4.20: Distribuzione dei tempi di rilettura per condizione

Dalla figura si evince che molti valori sono pari a zero, indicando che numerosi soggetti non hanno effettuato fasi di rilettura della quinta parola. Questo fenomeno è evidente in tutte le condizioni, suggerendo che la rilettura della quinta parola non è un comportamento comune tra i partecipanti. Le mediane dei tempi di rilettura sono basse per tutte le condizioni, riflettendo la tendenza generale a non rileggere. Tuttavia, si osserva una variabilità significativa, in particolare le condizioni B, C e D mostrano una maggiore dispersione e la presenza di valori estremi con tempi di rilettura più elevati, suggerendo che alcuni partecipanti hanno impiegato tempi significativamente più lunghi nella rilettura in queste condizioni.

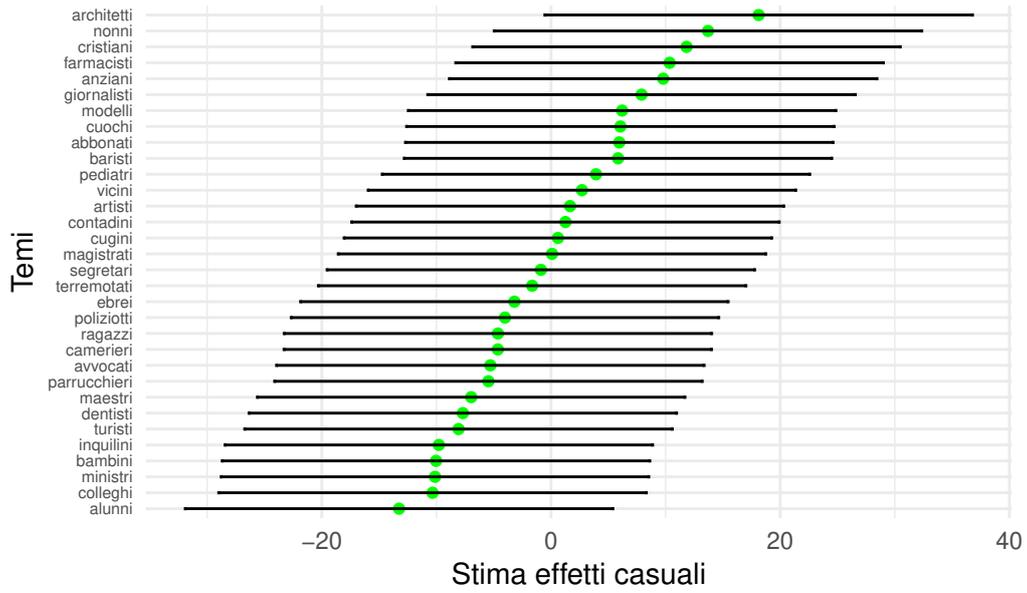
Per la stima dei modelli si è seguita la logica precedentemente impiegata. Inizialmente, vengono incluse nell'analisi variabili come sesso ed età. Tuttavia, queste variabili sono state successivamente rimosse a causa della loro

manca di significatività (p-value = 0.803 e 0.264). Questa decisione consente di semplificare il modello senza comprometterne l'integrità dei risultati. A differenza delle altre variabili dipendenti, in questo caso il numero di caratteri non risulta significativamente diverso da zero. Viene quindi adattato un modello lineare a effetti misti con effetti fissi la condizione sperimentale, oltre che agli effetti casuali incrociati per tema e partecipante, riportato nella Tabella 4.18.

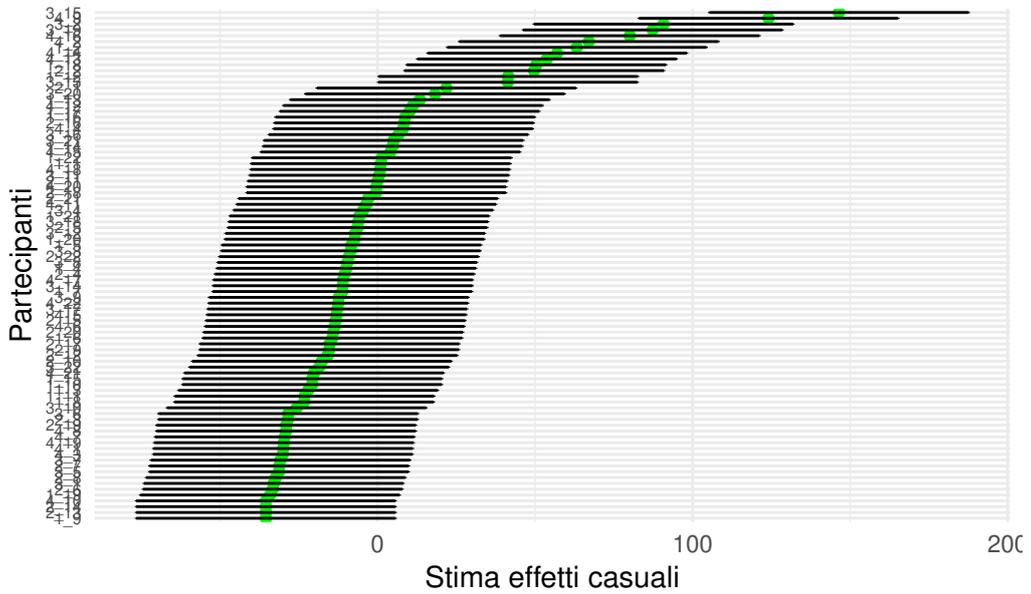
| Predittori               | Stime    | Standard Error | <i>P – value</i> |
|--------------------------|----------|----------------|------------------|
| Intercetta               | 35.02    | 7.132          | <0.001           |
| Condizione B             | 10.71    | 7.317          | 0.143            |
| Condizione C             | 23.55    | 7.319          | 0.001            |
| Condizione D             | 16.24    | 7.314          | 0.026            |
| Effetti casuali          |          |                |                  |
| $\sigma^2$               | 18866.71 |                |                  |
| $\tau_{00}$ Partecipante | 1706.21  |                |                  |
| $\tau_{00}$ Tema         | 153.17   |                |                  |

Tabella 4.18: Modello lineare a effetti misti per la rilettura, con effetti casuali per tema e partecipante.

L'analisi della varianza degli effetti casuali rivela che la varianza attribuibile alle differenze tra i partecipanti è significativamente maggiore rispetto a quella associata ai temi. Questo risultato è ulteriormente supportato dall'evidenza fornita dall'analisi grafica riportata in Figura 4.21.



((a)) Effetti casuali per tema



((b)) Effetti casuali per partecipanti

Figura 4.21: Rappresentazione grafica degli effetti casuali

Pertanto viene adattato un modello più parsimonioso, togliendo l'effetto casuale relativo ai temi e riportando l'output in Tabella 4.19.

| Predittori               | Stime    | Standard Error | <i>P</i> – <i>value</i> |
|--------------------------|----------|----------------|-------------------------|
| Intercetta               | 35.02    | 6.800          | <0.001                  |
| Condizione B             | 10.76    | 7.346          | 0.143                   |
| Condizione C             | 23.50    | 7.349          | 0.001                   |
| Condizione D             | 16.24    | 7.343          | 0.027                   |
| Effetti casuale          |          |                |                         |
| $\sigma^2$               | 19020.05 |                |                         |
| $\tau_{00}$ Partecipante | 1701.30  |                |                         |

Tabella 4.19: Modello lineare a effetti misti, con effetto casuale per partecipante

L'intercetta ha un valore di 35.02 millisecondi, indicando il tempo medio di rilettura per parola atteso nella condizione di A. La condizione C mostra un incremento significativo del tempo medio di rilettura per parola, con una stima di 23.50 millisecondi. Analogamente, la condizione D presenta un effetto positivo significativo, con un aumento stimato di 16.24 millisecondi. Al contrario, la condizione B, nonostante mostri un aumento stimato di 10.76 millisecondi, non presenta una differenza significativa rispetto alla condizione di base. Viene inoltre confermato il parametro associato alla varianza tra individui, il quale risulta elevato, sottolineando significative differenze interindividuali nelle fasi di rilettura della parola analizzata.

Infine, è stato condotto un test del log-rapporto di verosimiglianza ristretta tra i due modelli annidati per valutare l'adeguatezza dell'inclusione dell'effetto casuale per tema. I risultati di questo test hanno indicato che, per il tempo di rilettura, il modello senza l'inclusione dell'effetto casuale per tema è preferibile. Pertanto, si è deciso di accettare il modello più parsimonioso, che esclude l'effetto casuale per tema, poiché offre una rappresentazione adeguata dei dati senza aggiungere complessità non necessaria.

## 4.5 Conclusioni sullo studio

L'analisi dei modelli sviluppati in questo studio ha permesso di approfondire la comprensione degli effetti dell'utilizzo di differenti forme linguistiche sui tempi di lettura di diversi item. I risultati dei modelli stimati indicano chiaramente che le condizioni sperimentali hanno avuto un'influenza significativa sulle performance dei partecipanti.

I risultati ottenuti dimostrano che il maschile generico risulta l'opzione che genera minori costi di elaborazione rispetto alle tre alternative di linguaggio inclusivo, ovvero schwa, asterisco e sdoppiamento, in tutte le aree di interesse e nelle tre variabili dipendenti analizzate. Da questo primo risultato si può dedurre che l'uso del maschile generico conferma la sua prevalenza come opzione abituale. I partecipanti sono infatti più abituati a leggere frasi in cui il sintagma nominale e l'aggettivo presentano la marcatura grammaticale maschile. L'asterisco e la schwa hanno registrato i tempi di lettura medi per parola più alti, soprattutto durante la fase di prima lettura e di lettura totale. Questo risultato era prevedibile, in quanto le forme linguistiche inclusive sono ancora relativamente nuove e non sono ampiamente conosciute. Inoltre, nonostante siano opzioni che possono superare la dicotomia con entrambi i generi, la loro unica applicazione nel campo scritto le limita molto.

In particolare, per l'asterisco, questi tempi più lunghi sono probabilmente dovuti alle incomprensioni che può causare nei lettori. Lo stesso si può dire per la schwa perché per alcuni potrebbe essere percepita più maschile (Gheno, 2022) ma allo stesso tempo per alcuni potrebbe essere interpretata come [e] dalle somiglianze fonetiche. Per quanto riguarda lo sdoppiamento, i risultati confermano l'ipotesi che questa modalità comporti costi di elaborazione maggiori rispetto al maschile generico, ma inferiori rispetto all'asterisco e allo schwa, sia nella fase di prima lettura che nel tempo di lettura totale. Questa differenza può essere attribuita al fatto che la forma dello sdoppiamento è relativamente più conosciuta rispetto alle altre due condizioni.

Tutti i marchi di linguaggio inclusivo, se confrontati con il maschile generico, quindi, comportano un aumento dei tempi di elaborazione e dei costi cognitivi, soprattutto durante la prima lettura. È in questa fase che il lettore incontra per la prima volta tali marchi e deve effettuare le prime operazioni di decodifica.

I tempi impiegati per la riletture risultano generalmente bassi. Tuttavia, anche durante questa fase, le tre opzioni di linguaggio inclusivo si sono rivelate difficili da elaborare. Nonostante i tempi di riletture per parola siano ridotti, esse mostrano un sovraccosto considerevole e statisticamente significativo rispetto alla condizione di base. La presenza di valori negativi nella fase di riletture del sintagma nominale e dell'enunciato indica che tale fase non è obbligatoria; il lettore effettua una verifica solo se ha incontrato difficoltà nelle fasi precedenti che non è riuscito a risolvere.

In conclusione, il maschile generico si conferma come l'opzione di genere inclusivo che comporta i minori costi di elaborazione, se confrontato con le nuove forme di linguaggio inclusivo quali lo schwa, l'asterisco e lo sdoppiamento. Questo risultato suggerisce che i lettori sono maggiormente abituati al maschile generico, rendendone l'elaborazione cognitiva più agevole. Le nuove forme di linguaggio inclusivo, sebbene potenzialmente efficaci, richiedono un periodo di adattamento più lungo per essere integrate nei processi di lettura e comprensione.

# Capitolo 5

## Limiti dello studio

Nonostante i risultati ottenuti forniscono importanti indicazioni sulle dinamiche di lettura e comprensione del linguaggio inclusivo, lo studio presenta alcuni limiti che meritano di essere discussi.

Un primo limite riguarda il numero limitato di variabili incluse nell'analisi. Sebbene lo studio abbia considerato alcune variabili fondamentali, l'inclusione di ulteriori variabili demografiche e psicometriche avrebbero potuto arricchire l'analisi. Ad esempio, la raccolta di dati attraverso un questionario mirato avrebbe permesso di ottenere informazioni dettagliate sulle caratteristiche dei soggetti, come il loro pensiero riguardo al linguaggio inclusivo o la loro familiarità con tali forme linguistiche. Questo approccio avrebbe potuto offrire una comprensione più approfondita delle differenze individuali nelle risposte.

Un secondo limite significativo, soprattutto per il secondo studio, è l'assenza di metodi adeguati per l'analisi delle variabili di risposta continue. Attualmente, esistono pochi modelli in grado di gestire efficacemente gli effetti misti, con la maggior parte dei modelli disponibili progettati per includere esclusivamente effetti fissi, limitazione che ha impedito un'analisi più dettagliata e robusta delle risposte continue.

In sintesi, mentre lo studio ha contribuito in modo significativo alla comprensione del linguaggio inclusivo, l'inclusione di ulteriori variabili demografiche e psicometriche oltre l'adozione di modelli metodologici più avanzati per l'analisi delle variabili continue potrebbero migliorare ulteriormente la qualità della ricerca.

# Capitolo 6

## Conclusioni

Lo studio ha esaminato l'influenza delle diverse condizioni sperimentali sulle performance dei partecipanti, analizzando due dataset distinti. I risultati hanno evidenziato un impatto significativo delle condizioni sperimentali, sia sui tempi di lettura che sulla qualità delle risposte. Il maschile generico è emerso come l'opzione che comporta i minori costi di elaborazione, risultando più familiare e agevole per i partecipanti. Al contrario, le nuove forme di linguaggio inclusivo hanno mostrato un incremento nei tempi di lettura e nei costi cognitivi a causa della loro relativa novità e complessità. Le analisi hanno inoltre rivelato una significativa variabilità individuale nelle risposte, suggerendo che le differenze cognitive, esperienziali o motivazionali tra i partecipanti giocano un ruolo cruciale nelle loro performance. La variabilità tra i temi, sebbene presente, è risultata meno pronunciata, indicando che le caratteristiche individuali dei partecipanti hanno un'influenza maggiore rispetto alle caratteristiche dei temi stessi.

In sintesi, lo studio evidenzia la necessità di considerare le differenze individuali e la familiarità con le forme linguistiche nella progettazione degli esperimenti e nell'interpretazione dei risultati. Le nuove forme di linguaggio inclusivo, pur richiedendo un periodo di adattamento più lungo, necessitano di ulteriori indagini per valutarne l'impatto sulla comunicazione. Le future ricerche dovrebbero includere ulteriori variabili demografiche e psicometriche per approfondire la comprensione delle dinamiche cognitive associate all'uso del linguaggio inclusivo, al fine di migliorare la qualità delle analisi.

## Bibliografia

- Arfini, Elia, Rossella Ghigi e Sveva Magaraggia (2019). «Can feminism be right? A content analysis of discourses about women by female Italian right-wing politicians». In: *Rassegna Italiana di Sociologia* 60.4, pp. 693–719.
- Bazzanella, Carla et al. (2010). «Genere e lingua». In: *Enciclopedia dell'italiano*. Vol. 1. Enciclopedia Italiana Treccani, pp. 556–557.
- CORTINAS ABRAHANTES, Jose (2004). «Estimation procedures for mixed-effects models with applications to normally distributed and survival data». In.
- D'Elia, Angela et al. (2000). «Il meccanismo dei confronti appaiati nella modellistica per graduatorie: sviluppi statistici ed aspetti critici». In: *Quaderni di Statistica* 2, pp. 173–203.
- D'Elia, Angela (2003). «Modelling ranks using the inverse hypergeometric distribution». In: *Statistical Modelling* 3.1, pp. 65–78.
- Daneman, Meredyth e Patricia A Carpenter (1980). «Individual differences in working memory and reading». In: *Journal of verbal learning and verbal behavior* 19.4, pp. 450–466.
- Diggle, Peter (2002). *Analysis of longitudinal data*. Oxford university press.
- donna, Commissione nazionale per la realizzazione della parità tra uomo e e Alma Sabatini (1986). *Raccomandazioni per un uso non sessista della lingua italiana: per la scuola e per l'editoria scolastica*. Presidenza del Consiglio dei ministri, Direzione generale delle informazioni.
- Duchowski, Andrew e Andrew Duchowski (2007). «Eye tracking techniques». In: *Eye tracking methodology: Theory and practice*, pp. 51–59.
- Galeandro, Simona (2021). «Femminilizzazione versus neutralizzazione». In: *Testo e Senso* 23, pp. 65–73.
- Gheno, Vera (2022). *Femminili singolari+: Il femminismo è nelle parole*. effequ.

- Giusti, G (2021). «In che modo la lingua riflette la cultura di chi parla e non la realtà dei fatti». In: *Agcult/Lettere lente* 1.
- Goldstein, Harvey (2011). *Multilevel statistical models*. John Wiley & Sons.
- Hartzel, Jonathan, Alan Agresti e Brian Caffo (2001). «Multinomial logit random effects models». In: *Statistical Modelling* 1.2, pp. 81–102.
- He, Qianchuan et al. (2021). «Random effect based tests for multinomial logistic regression in genetic association studies». In: *Genetic epidemiology* 45.7, p. 736.
- Holmqvist, Kenneth et al. (2011). *Eye tracking: A comprehensive guide to methods and measures*. oup Oxford.
- Iannario, Maria (2010). «On the identifiability of a mixture model for ordinal data». In: *Metron* 68, pp. 87–94.
- Iannario, Maria et al. (2012). «CUBE models for interpreting ordered categorical data with overdispersion». In: *Quaderni di statistica* 14, pp. 137–140.
- Iannario, Maria (2014). «Modelling uncertainty and overdispersion in ordinal data». In: *Communications in Statistics-Theory and Methods* 43.4, pp. 771–786.
- Irrazábal, Natalia e Carlos Molinari Marotto (2004). «Técnicas experimentales en la investigación de la comprensión del lenguaje». In: *Revista Latinoamericana de Psicología* 37.3, pp. 581–594.
- Jansen, J (1990). «On the statistical analysis of ordinal data when extravariation is present». In: *Journal of the Royal Statistical Society Series C: Applied Statistics* 39.1, pp. 75–84.
- Just, Marcel A e Patricia A Carpenter (1980). «A theory of reading: from eye fixations to comprehension.» In: *Psychological review* 87.4, p. 329.
- Keating, Gregory D e Jill Jegerski (2015). «Experimental designs in sentence processing research: A methodological review and user’s guide». In: *Studies in Second Language Acquisition* 37.1, pp. 1–32.

- Kenda, Jana (2022). «Grammatica inclusiva in italiano: Le alternative linguistiche offerte e il riscontro dell'opinione pubblica». In: *Linguistica* 62.1-2, pp. 205–222.
- Li, Hongmin e Woonghee Tim Huh (2011). «Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications». In: *Manufacturing & Service Operations Management* 13.4, pp. 549–563.
- Loureda, Óscar, Laura Nadal Sanchis, Inés Recio et al. (2016). «Partículas discursivas y cognición: sin embargo y la conexión argumentativa». In: *Geométrica explosión. Estudios de lengua y literatura en homenaje a René Lenarduzzi*. Ca'Foscari.
- Loureda, Óscar et al. (2019). «Empirical studies of the construction of discourse». In: *Empirical Studies of the Construction of Discourse*, pp. 1–327.
- McLachlan, Geoffrey J e Thriyambakam Krishnan (2007). *The EM algorithm and extensions*. John Wiley & Sons.
- McNeish, Daniel (2016). «Estimation methods for mixed logistic models with few clusters». In: *Multivariate Behavioral Research* 51.6, pp. 790–804.
- Meernik, Clare et al. (2016). «Eye tracking outcomes in tobacco control regulation and communication: a systematic review». In: *Tobacco regulatory science* 2.4, p. 377.
- Noveck, Ira (2018). *Experimental pragmatics: The making of a cognitive science*. Cambridge University Press.
- Quené, Hugo e Huub Van den Bergh (2008). «Examples of mixed-effects modeling with crossed random effects and with binomial data». In: *Journal of Memory and Language* 59.4, pp. 413–425.
- Raney, Gary E, Spencer J Campbell e Joanna C Bovee (2014). «Using eye movements to evaluate the cognitive processes involved in text comprehension». In: *JoVE (Journal of Visualized Experiments)* 83, e50780.

- Rayner, Keith (1998). «Eye movements in reading and information processing: 20 years of research.» In: *Psychological bulletin* 124.3, p. 372.
- Richardson, Daniel C e Rick Dale (2005). «Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension». In: *Cognitive science* 29.6, pp. 1045–1060.
- Ripatti, Samuli e Juni Palmgren (2000). «Estimation of multivariate frailty models using penalized partial likelihood». In: *Biometrics* 56.4, pp. 1016–1022.
- Sabatini, Alma e Marcella Mariani (1987). *Il sessismo nella lingua italiana*. Presidenza del Consiglio dei ministri, Direzione generale delle informazioni.
- Sartori, Francesca et al. (2009). *Differenze e disuguaglianze di genere*. Il mulino.
- Scheipl, Fabian, Sonja Greven e Helmut Küchenhoff (2008). «Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models». In: *Computational statistics & data analysis* 52.7, pp. 3283–3299.
- Self, Steven G e Kung-Yee Liang (1987). «Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions». In: *Journal of the American Statistical Association* 82.398, pp. 605–610.
- Stetie, Noelia Ayelén e Gabriela Mariel Zunino (2022). «Non-binary language in Spanish? Comprehension of non-binary morphological forms: a psycholinguistic study». In.
- Therneau, Terry (2003). «On mixed-effect Cox models, sparse matrices, and modeling data from large pedigrees». In: URL [http://www. mayo. edu/biostatistics](http://www.mayo.edu/biostatistics).
- Tripathi, Ram C, Ramesh C Gupta e John Gurland (1994). «Estimation of parameters in the beta binomial model». In: *Annals of the Institute of Statistical Mathematics* 46, pp. 317–331.

- Tutz, Gerhard e Wolfgang Hennevogl (1996). «Random effects in ordinal regression models». In: *Computational Statistics & Data Analysis* 22.5, pp. 537–557.
- Zunino, Gabriela Mariel e Noelia Ayelén Stetie (2022). «Binary or non-binary? Gender morphology in Spanish: differences dependent on the task». In: *Alfa: Revista de Linguística (São José do Rio Preto)* 66, e14546.