

Progettazione di un'Interazione Multimodale per l'Elaborazione del Testo su Android



Candidato: **NUZZO LUCA**

Relatore: **SEGNAN MARINO**

Programmazione Mobile: Android

Introduzione

II

Problema

Applications that require extensive text entry or typing are standout examples of mobile-problematic experiences.

C. Banga, J. Weinhold, Essential Mobile Interaction Design (Usability), Pearson Education, 2014

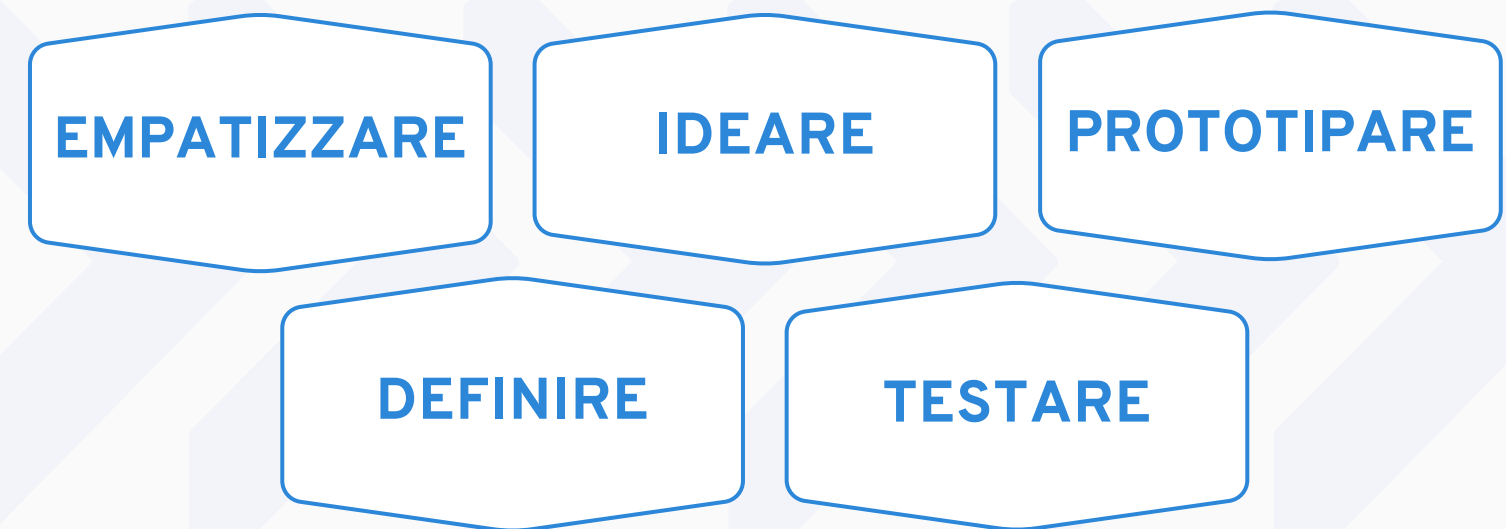
Obiettivi

Per agevolare la manipolazione di contenuti testuali su dispositivi mobili:

- progettare un'interazione multimodale basata su voce e tocco
- sviluppare un'applicazione Android per valutare la soluzione proposta.

Nozioni e Metodi

- *conversation design*
- *human-centered design*
- *design thinking*





Casi d'uso

■ ■ A sinistra

Comando Multimodale

"Taglia questa frase"

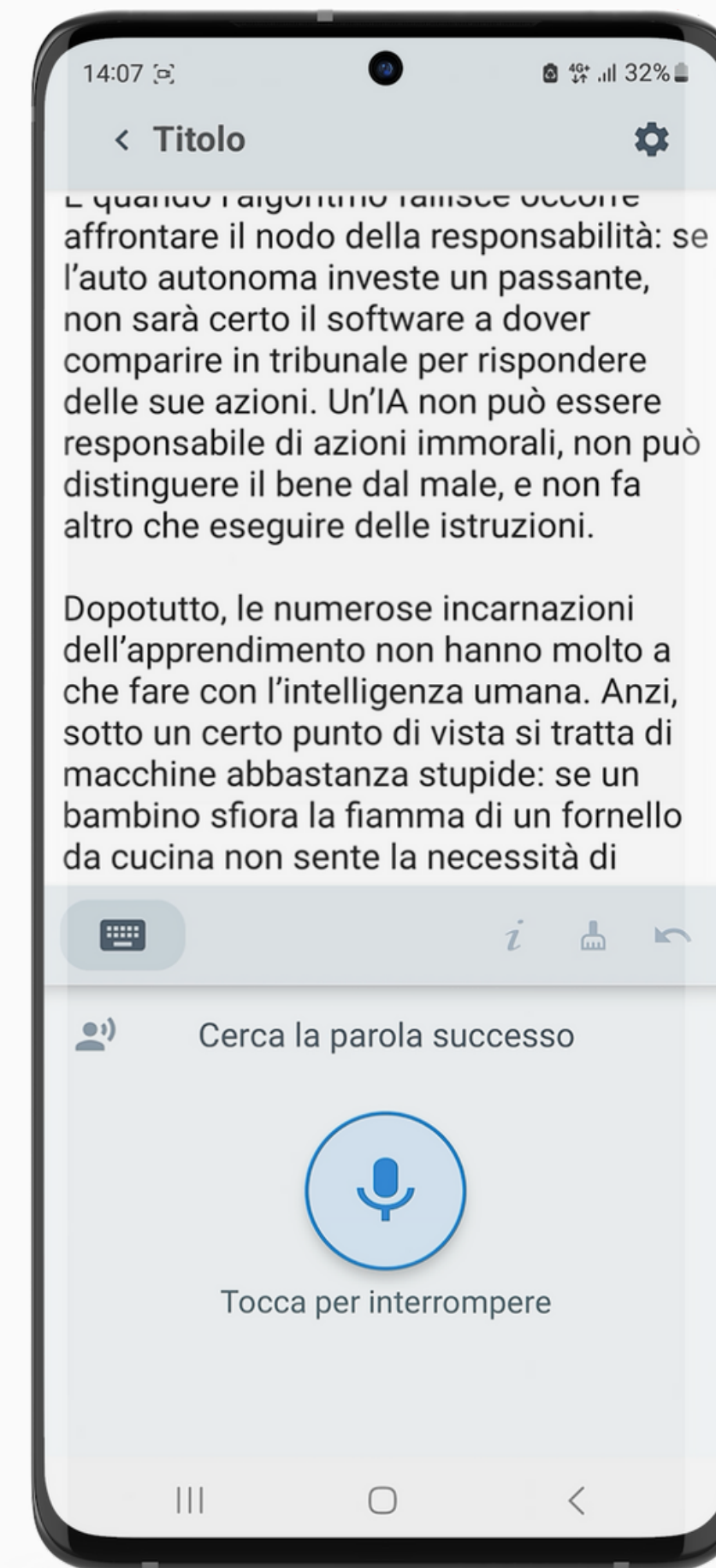
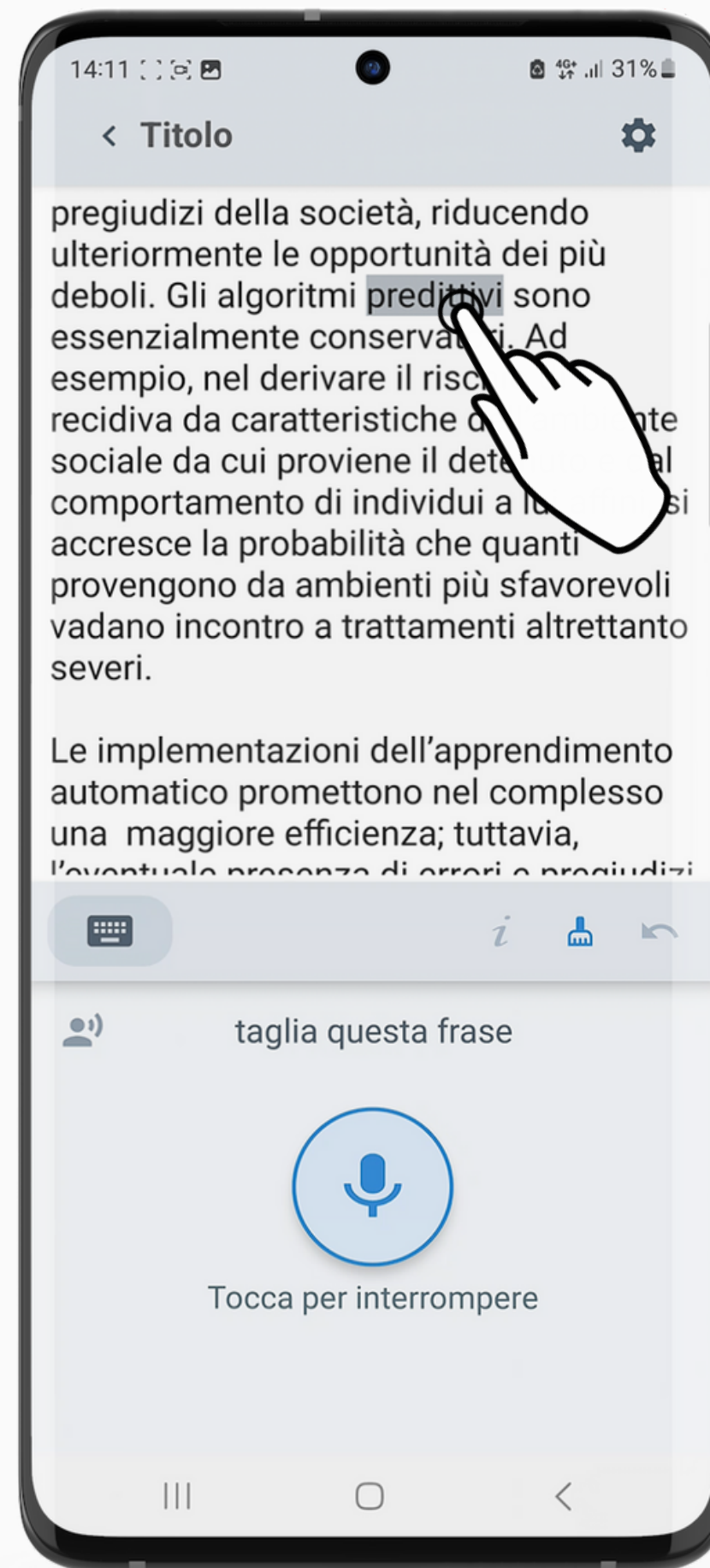
⊕ *Tap* sulla frase

□ ■ A destra

Comando Vocale Completo

"Cerca la parola *successo*"

⊗ Nessun input tattile



Empatizzare ●○○○○

IV

Acquisire una conoscenza profonda di bisogni, comportamenti e motivazioni degli utenti

👤 Stato dell'arte

Soluzioni dominanti sul mercato:

- *iOS Voice Control*
- *Android Voice Access*

Proposte accademiche correlate:

- M. Zhao et al., *Voice and Touch Based Error-tolerant Multimodal Text Editing and Correction for Smartphones*
- J. Yang et al., *DoThisHere: Multimodal Interaction to Improve Cross-Application Tasks on Mobile Devices*

👤 Guerrilla User Research

Ricerca con un numero ristretto di partecipanti (3x2) e in ambienti non controllati.

- Domande iniziali per conoscere l'utente
- Task di scrittura e manipolazione del testo
- Domande finali per conoscere o approfondire delle problematiche

👤 Analisi della letteratura

Aree tematiche:

- Digitazione ed elaborazione del testo su *mobile*
- Analisi e design della conversazione
- Interazione vocale e interfacce multimodali



Definire



Sintetizzare i risultati della ricerca con definizioni formali del problema e dei bisogni degli utenti

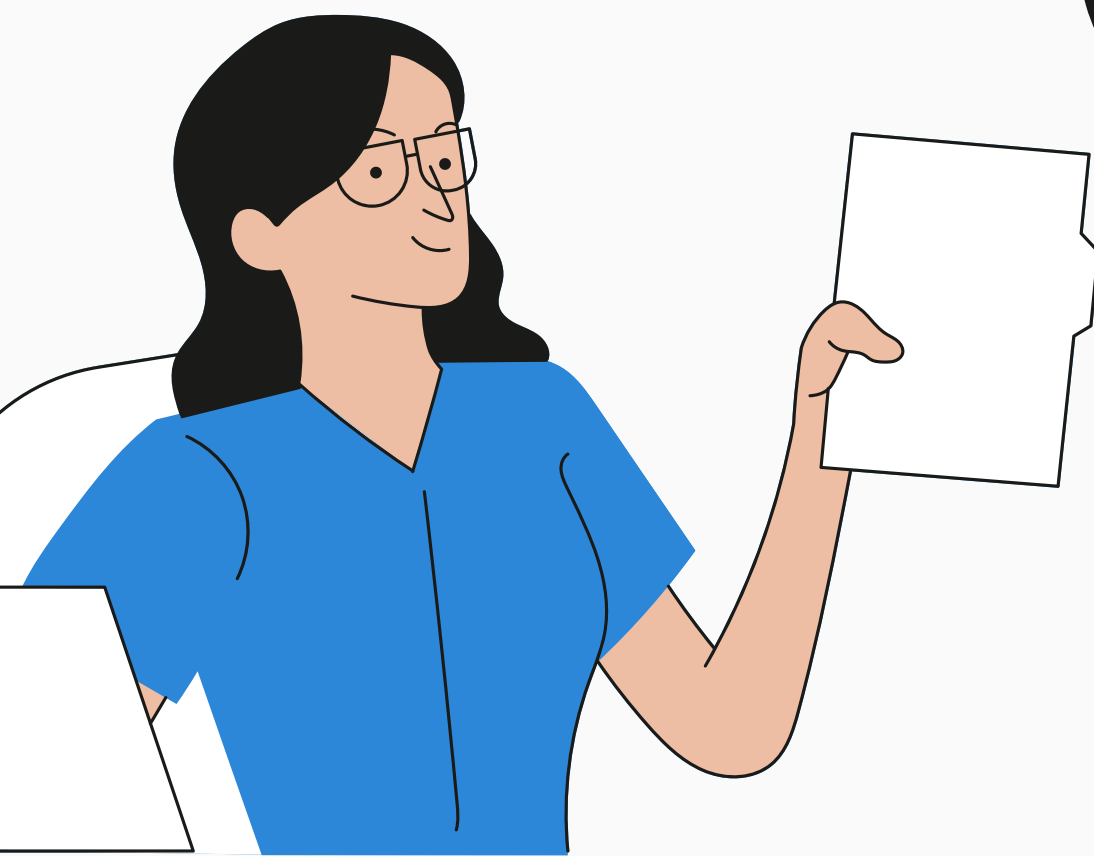
OBIETTIVO
PERCHÈ?
MUST-HAVE

Progettare un'interazione multimodale basata su comandi vocali per agevolare e snellire le operazioni di elaborazione del testo su *mobile*, a partire dalle operazioni di copia, taglia, incolla e rimuovi.

**Minimum
Viable Product**

Agevolare la
selezione del
testo

Consentire di
scegliere la
modalità di input



«È la fase della generazione di idee, di divergenza del pensiero in cui l'obiettivo non è farsi venire l'idea giusta, ma quante più idee possibili»

D. Bottà, User eXperience Design, Hoepli, 2018, p. 22

Intento	Comando vocale completo	Comando multimodale
#copy {Copia} {Copiare}	# tutto [il testo] # [[il] testo] da ^parola/e [fino] a parola/e # [[il] testo] da ^parola/e [fino alla fine] # [[il] testo] [dall'inizio] fino a parola/e	# [[{questa} {la}] parola] # [{questa} {la}] frase # [{questo} {il}] {paragrafo} {capoverso} {verso} # [[il] testo] da ^parola/e # [[il] testo] fino a ^parola/e
#cut {Taglia} {Tagliare}	# tutto [il testo] # [[il] testo] da ^parola/e [fino] a ^parola/e # [[il] testo] da ^parola/e [fino alla fine] # [[il] testo] [dall'inizio] fino a ^parola/e	# [[{questa} {la}] parola] # [{questa} {la}] frase # [{questo} {il}] {paragrafo} {capoverso} {verso} # [[il] testo] da ^parola/e # [[il] testo] fino a ^parola/e

Estratto della lista di utterance riconoscibili per gli intenti #copy e #cut

Spazio della conversazione

Sono state individuate tutte le operazioni (o intenti) possibili, ordinandole in base a priorità e facilità d'implementazione.

Utterance-Intent Matching

Si è generata una lista di espressioni vocali (o *utterance*) riconoscibili per ciascun intento, distinguendo tra comandi vocali completi e comandi multimodali.



Iterazione

La lista è stata ciclicamente riformulata per rispondere alle evidenze emerse in fase di prototipazione e testing.

VII

Prototipare

◆ Pseudocodice

Notazione semplificata di un linguaggio di programmazione.

Strategia ad alto livello per l'analisi dell'input multimodale e l'elaborazione di parametri, in funzione dei requisiti di design.

◆ Wireframes

...

Estratto di pseudocodice nella prima iterazione del design

procedure <u>GET OPERATION</u>	
1	input:
2	$T \leftarrow$ touch input
3	$V \leftarrow$ voice input
4	definizioni:
5	w : location parameters of the text to edit
6	i : intent specified by speech (V)
7	$i \leftarrow$ search intent from V
8	processo:
9	if intent i is not found then
10	return
11	if intent i is <i>#paste</i> then
12	if touch input T then
13	$w \leftarrow$ get paste location params
14	else
15	$w \leftarrow$ get position of cursor
16	else if intent i is <i>#search</i> then
17	if touch input T then
18	$w \leftarrow$ get string from T
19	else $w \leftarrow$ get string from V
20	else
21	if touch input T then
22	$w \leftarrow$ GET LOCATION PARAMS BY COMBINING T AND V
23	else $w \leftarrow$ GET LOCATION PARAMS FROM V
24	return $o^* = \langle i, w \rangle$

VII

Prototipare

◆ Pseudocodice

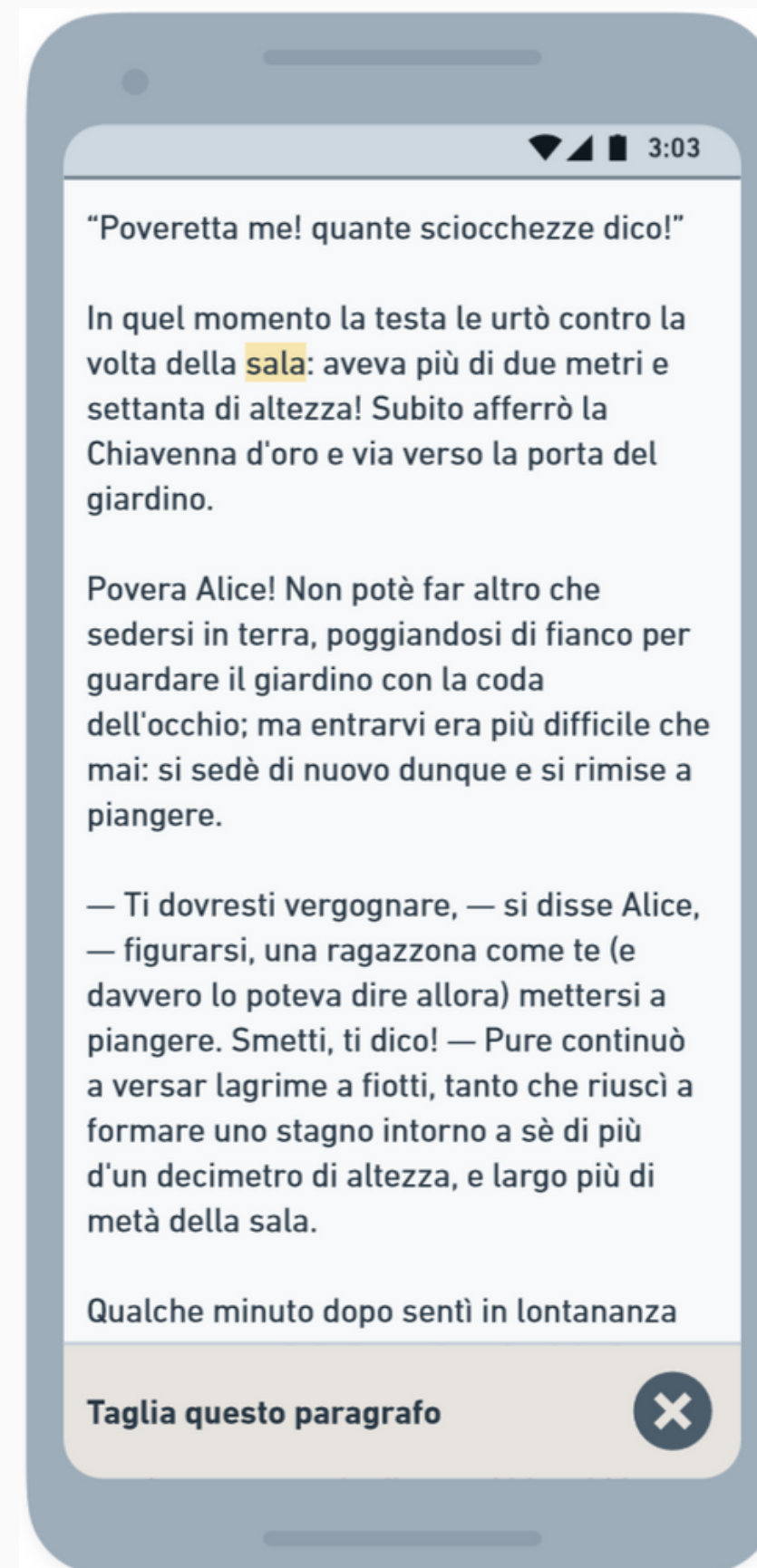
...

◆ Wireframes

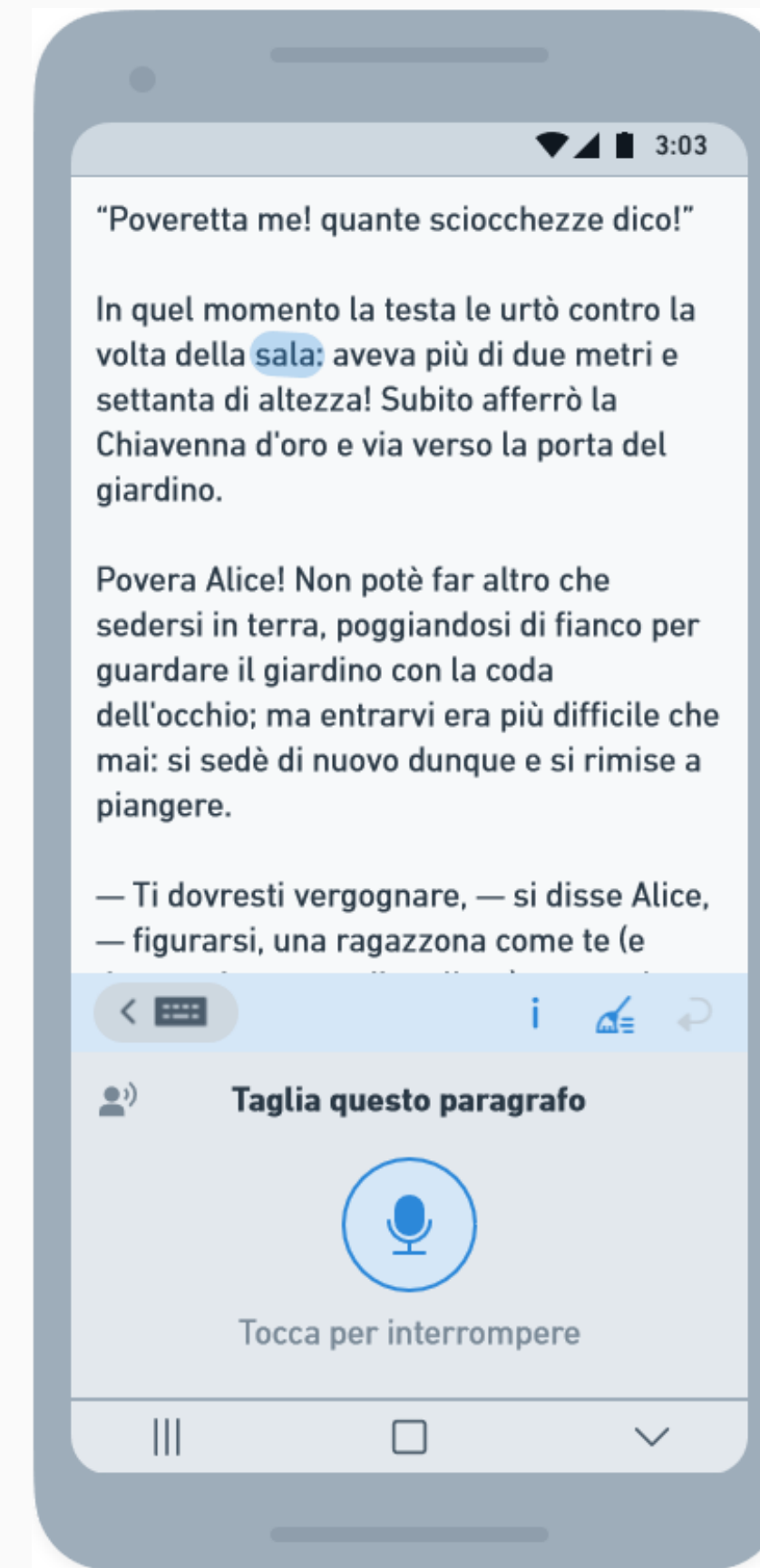
Rappresentativi del layout, della navigazione e dei principali elementi di un'interfaccia.

Creati per migliorare UI e UX, lo pseudocodice e l'*utterance-intent matching*, aumentando progressivamente il livello di fedeltà.

Prototipo Low-Fidelity



Prototipo High-Fidelity





Wizard of Oz

Ogni partecipante ha interagito con il prototipo. Si è simulato il comportamento di quest'ultimo scegliendo quella risposta che più si addice al comando dell'utente.

«Discount Usability» Jakob Nielsen

Tecniche rapide, poco costose e non troppo sistemiche per individuare rapidamente i più gravi problemi di usabilità.

- 3-4 partecipanti per sessione (x3)
- ambiente controllato
- copione di riferimento predefinito
- registrazione audio/video

Testare

VIII

Raccogliere i feedback degli utenti sul prototipo

Report

Impressioni degli utenti, pattern d'interazione, utterance più comuni, aspetti più problematici e possibili soluzioni.

(Allegato) [Lista delle evidenze riscontrate](#), classificate come:

- Nota positiva o funzionalità supportata
- Segnalazione *in-scope*
- Segnalazione *out-of-scope* o non prioritaria

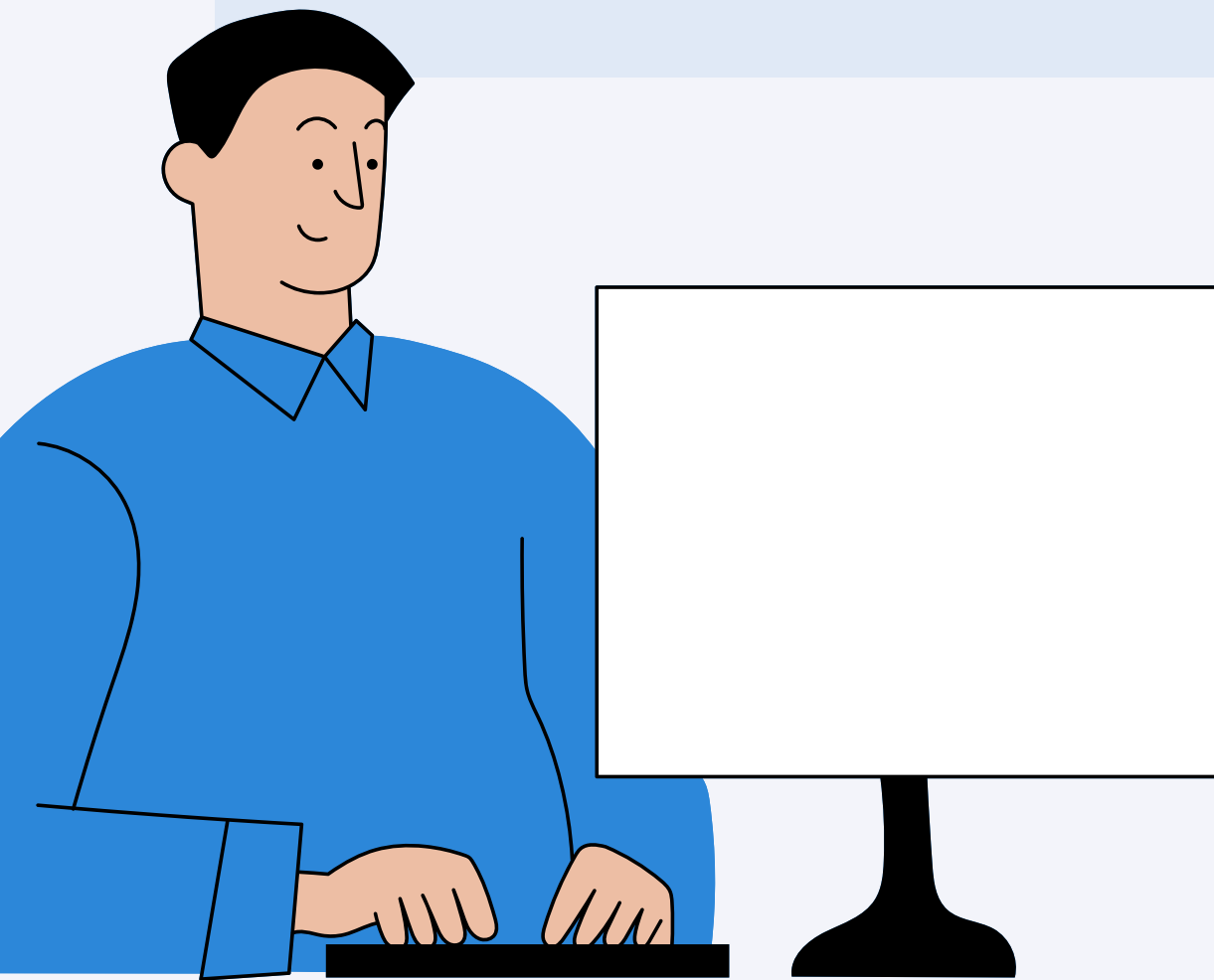
Iterazione: Requisiti di sviluppo

La lista ha orientato le successive iterazioni del design ed è stata aggiornata ad ogni ciclo.

Sviluppo di un prototipo di app Android per valutare la soluzione proposta

DESIGN
ARTIFACTS

Utterance-Intent Matching
Wireframe High-Fidelity
Pseudocodice
Requisiti di sviluppo



◆ **Single-Activity Architecture**

Una singola *activity* ospita due *fragment* che si alternano a schermo, rispettivamente dedicati all'interazione multimodale e al dialogo fra utente e sistema.

◆ **Android Speech API**

```
public class SpeechRecognizer
```

Per accedere al servizio di riconoscimento vocale

```
public interface RecognitionListener
```

Metodi per intercettare e gestire gli eventi correlati



Traguardi raggiunti

✓ Più di un *Minimum Viable Product*

La strategia elaborata con lo pseudocodice si è rivelata adeguata al **soddisfacimento degli obiettivi**. Sono stati soddisfatti quasi tutti i requisiti prioritari e alcuni requisiti secondari.

✓ Intenti e *Utterance*

Sono stati implementati tutti gli intenti prioritari e le annesse utterance, nonché la maggior parte di quelli non prioritari. **Espressioni vocali** inedite sono state aggiunte in fase di sviluppo.



✓ Maggiore Usabilità

Rispetto alle soluzioni dominanti, l'app rende possibili delle interazioni multimodali più **flessibili** e più tolleranti nei confronti degli errori umani.

✓ UI & UX

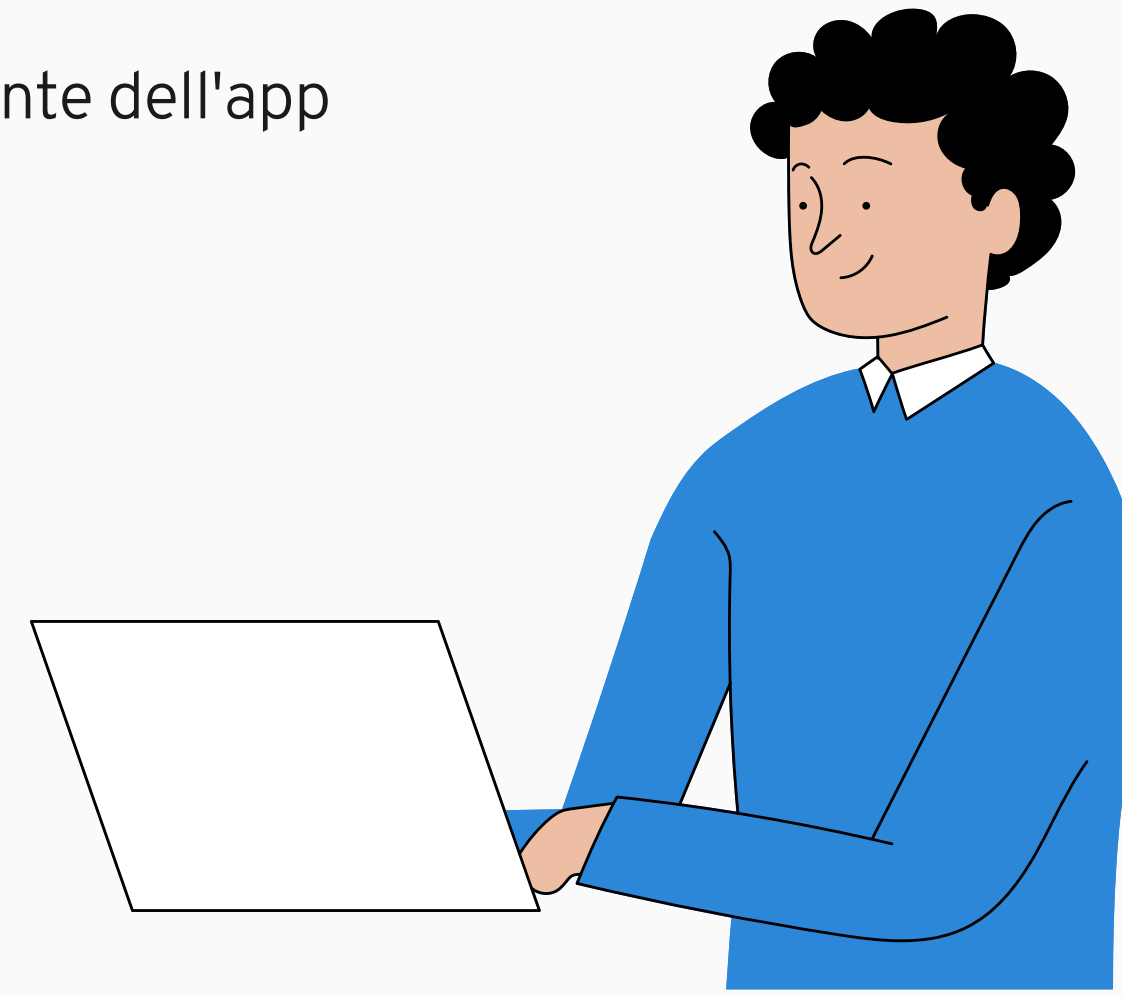
L'interfaccia e l'**interazione** proposte nel wireframe ad alta fedeltà sono state adeguatamente traslate all'app Android.



L'applicazione soddisfa gli obiettivi del progetto, offrendo un'interazione multimodale potenzialmente capace di agevolare e snellire le operazioni di manipolazione del testo su *mobile*.

Sviluppi Futuri

- Svolgere nuovi test di usabilità incentrati sul prototipo di app Android
- Migliorare la gestione e ripartizione dei *task* tra ciascun componente dell'app
- Potenziare il riconoscimento dei comandi vocali
- Affinare l'intercettazione di eventi legati all'input tattile
- Accrescere le *gesture* possibili per l'utente finale



Grazie per l'attenzione!

Progettazione di un'Interazione Multimodale per l'Elaborazione del Testo su Android

Candidato: **NUZZO LUCA**

Relatore: **SEGNAN MARINO**

Programmazione Mobile: Android