Tampere University

Alina Ida

USABILITY AND ACCESSIBILITY EVALUATION OF FOOD DELIVERY APPLICATIONS IN FINLAND: A CASE STUDY OF WOLT AND FOODORA

Master's thesis
Faculty of Information Technology
and Communication Sciences
Examiners: Päivi Majaranta
Jari Varsaluoma
December 2025

ABSTRACT

Alina Ida: Usability and accessibility evaluation of food delivery applications in Finland: A case study of Wolt and Foodora
Master's thesis
Tampere University
Master's Degree Programme in Computing and Electrical Engineering: Human-Technology Interaction
December 2025

This thesis examines the usability and accessibility of mobile food delivery applications in the Finnish context, focusing on the two dominant platforms, Wolt and Foodora. As food delivery services have become a routine part of everyday consumption, well-designed interfaces and accessible interaction are essential for ensuring efficient, inclusive, and trustworthy use. Despite their widespread adoption, systematic academic evaluation of these applications in Finland remains limited.

The study adopts a mixed-method approach that combines remote usability testing with end users, heuristic evaluation based on Nielsen's ten usability heuristics, and both automated and manual accessibility assessment aligned with the Web Content Accessibility Guidelines 2.2. Usability evaluation focused on common food-ordering tasks, while the accessibility evaluation examined screen reader compatibility, layout behavior under increased accessibility settings, language handling, and visual clarity.

The results show that both applications support successful completion of core ordering tasks, but notable differences exist in interaction quality and accessibility support. Wolt was generally perceived as more stable and predictable, whereas Foodora was more frequently affected by distractions, unclear feedback, and technical interruptions. Recurring issues were identified in both applications, including inconsistent language presentation, limited price transparency prior to checkout, insufficient accessibility labeling, and layout instability when text and display settings were increased. These findings highlight the need for improved language consistency, clearer cost communication, stronger accessibility support, and more predictable interaction design in mobile food delivery services.

This thesis provides practical design recommendations and demonstrates how usability testing, heuristic evaluation, and accessibility assessment can be systematically combined to evaluate mainstream mobile applications.

Keywords: usability, accessibility, mobile applications, food delivery, user experience

USE OF AI IN THESIS

I have utilized AI tools in my thesis:

☐ No
☑ Yes

The AI tools utilized in my thesis and their purposes are described below:

Names and versions of AI tools:

- ChatGPT (OpenAI; GPT-5.1, GPT-5.2)

Purpose of using AI tools

AI tools were used solely for language-related support, including improving grammar, clarity, and academic tone of text written by the author, as well as minor assistance in refining literature search queries.
AI tools were not used to generate research content, analyze data, or draw conclusions.

Sections where AI tools were used

- General writing and editing across multiple sections of the thesis
- Literature search query formulation

I acknowledge that I am fully responsible for the entire content of my thesis, including the parts generated by AI, and accept accountability for any violations of ethical standards in publications.

PREFACE

This thesis focuses on the usability and accessibility of mobile food delivery applications in Finland.

I would like to express my sincere gratitude to my supervisor, Päivi Majaranta, whose course Usability Evaluation Methods provided a strong foundation for the usability testing and evaluation techniques applied in this thesis. Her guidance, feedback, and academic support were essential throughout the research process.

I would also like to thank my other Human-Technology Interaction professors at Tampere University for providing broader theoretical and methodological perspectives that contributed to this work.

I also thank all participants who took part in the usability testing sessions for their time and valuable input, which formed the empirical basis of this study.

Tampere, 15 December 2025

Alina Ida

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

ADHD	Attention-Deficit/Hyperactivity Disorder
dp	density-independent pixels
HCI	Human-Computer Interaction
UX	User Experience
W3C	World Wide Web Consortium
WCAG	Web Content Accessibility Guidelines

1 INTRODUCTION

1.1 Background and motivation

Food delivery applications have become an integral part of everyday life in Finland, particularly among students and working-age adults in urban areas. As the adoption of digital services continues to grow, the usability and accessibility of these applications play a crucial role in ensuring that users can complete essential tasks efficiently and without frustration.

Research on mobile user engagement shows that user-friendly and intuitive mobile features increase perceived value and satisfaction, which in turn strengthens users' intention to continue engaging with mobile services [16]. From this perspective, poorly designed interfaces are likely to reduce user satisfaction and weaken long-term engagement. Mobile accessibility research further highlights that users with different disabilities require specific interaction mechanisms, such as audio feedback, magnification, or screen readers, and that complex multi-level interfaces are particularly problematic for people with visual and motor impairments [15]. When mobile commerce applications fail to incorporate these accessibility principles, they risk excluding these user groups from using their services independently.

Recent research indicates that the popularity of food delivery applications continues to grow for reasons unrelated to the Coronavirus Disease (COVID-19) pandemic. Rungruangjit and Charoenpornpanichkul [14] found that users continued use of food delivery applications is driven by lasting factors such as functional value (e.g., convenience, variety, and value for money), social value (positive social perception), conditional value (promotions, free delivery, and access to preferred restaurants), and epistemic value (curiosity and novelty). These motivations remain strong across generations, suggesting that food delivery services have become a stable part of everyday consumption rather than temporary, crisis-driven behavior. This further reinforces the relevance of examining the usability and accessibility of major food delivery platforms in the current context.

Despite the widespread use of food delivery services, there is still limited academic research focusing specifically on the usability and accessibility of food delivery applications in Finland. This gap is notable given their frequent use in daily life.

Wolt and Foodora are the two dominant platforms operating in Finland's restaurant food ordering and delivery market. According to the Finnish Competition and Consumer Authority, "In Finland, there are only two companies operating in the restaurant food ordering and delivery platform market: Wolt and Foodora," and Wolt holds an estimated 70–80% market share [4]. Market data further confirms this duopoly: Statista's 2025 consumer survey shows that 73% of Finnish respondents used Wolt for online food delivery bookings, while 65% used Foodora during July 2024 - June 2025 [5]. These data demonstrate that Wolt and Foodora are the most widely used food delivery services in Finland, making them the most relevant platforms for investigating usability, user experience (UX), and accessibility within the Finnish food delivery ecosystem.

Consumer behavior data further highlights the importance of restaurant delivery services among Finnish users. According to Statistics Finland, individuals aged 25-34 are the most active online grocery shoppers, with over a quarter reporting that they purchase groceries online, followed by roughly one-fifth of consumers aged 16-24. However, a significantly larger proportion of Finnish consumers order ready-made meals directly from restaurants or via delivery platforms than purchase groceries online [6]. This indicates that food delivery platforms represent a more commonly used digital service category than online grocery shopping, particularly among younger adults, who also form a major user group for Wolt and Foodora.

Moreover, these applications frequently undergo design changes and feature updates, making continuous usability and accessibility evaluation necessary. Understanding the challenges that users, especially novice and users with accessibility needs, encounter can help identify broader design implications for mobile services intended for diverse populations.

1.2 Research gap

Existing research on food delivery applications has focused mainly on general usability, customer satisfaction, or market-specific design issues in countries outside Finland. Only a few studies address accessibility, and these typically examine visually impaired users in web-based restaurant menus or non-European delivery platforms (see Chapter 2 for a detailed review of previous research).

In the Finnish context, the limited available work consists of two bachelor's theses that provide only preliminary insights and do not offer systematic evaluation of usability, accessibility, or multilingual use in food delivery apps (see Chapter 2 for a discussion of these studies). To date, no academic study appears to have systematically evaluated

both Wolt and Foodora in the Finnish context using structured usability testing and accessibility evaluation frameworks. This thesis contributes a comparative, mixed-method evaluation (usability testing, heuristic evaluation, automated and manual accessibility checks) of the two dominant Finnish platforms using the same task set and consistent evaluation framing.

1.3 Aim and research questions

The aim of this thesis is to examine and compare the usability and accessibility of Wolt and Foodora. The study investigates how users perform key tasks, what problems they encounter, and how accessible interfaces are based on recognized evaluation frameworks.

To address this aim, the thesis explores the following research questions:

1. What usability and accessibility issues do users experience when using Wolt and Foodora in Finland?
2. How do these apps compare in terms of UX and accessibility?
3. What recommendations can be made to improve usability and inclusiveness in food delivery applications?

1.4 Scope and limitations

This thesis focuses exclusively on the customer-facing mobile applications of Wolt and Foodora used in Finland. The scope includes usability testing, heuristic evaluation based on Nielsen's heuristics, and accessibility evaluation using Web Content Accessibility Guidelines (WCAG) principles. The empirical component is based on a small sample of 5 participants, which is appropriate for identifying qualitative usability issues but does not allow for statistical generalization.

The study does not address the courier or restaurant-side interfaces, backend systems, or business operations of the companies. In addition, the study is conducted independently and is **not affiliated with, supported by, or endorsed by Wolt or Foodora**.

1.5 Structure of the thesis

This thesis is structured as follows:

- **Chapter 2** reviews the theoretical background related to usability, accessibility, mobile application design, and prior research on food delivery services.
- **Chapter 3** describes the methodology, including usability testing procedures, heuristic and accessibility evaluation methods, participant recruitment, and data analysis.
- **Chapter 4** presents the results of the usability tests and evaluation outcomes.
- **Chapter 5** discusses the findings, methodological considerations, and implications for UX design, followed by recommendations for improving food delivery applications.
- **Chapter 6** concludes the thesis by summarizing the main insights and outlining directions for future work.

2 THEORETICAL BACKGROUND AND RELATED STUDIES

2.1 Mobile application use and constraints

Mobile applications are used in highly variable real-world environments, which introduces specific usability challenges. Zhang and Adipat [9] emphasize that users may be standing, walking, or interacting with an application in bright or dim lighting conditions, all of which affect attention, visibility, and cognitive load. They also highlight enduring constraints such as small screen size and differing display resolutions, limitations that continue to shape mobile interaction design despite advances in device capabilities. As a result, mobile interfaces must prioritize clarity, minimalism, and efficient navigation to support effective use across contexts.

Similarly, Punchoojit and Hongwarittorrn [17] note that limited screen space, reduced precision in touch interactions, and the lack of tactile feedback impose fundamental design restrictions that can increase error rates and user frustration, particularly for individuals with accessibility needs. Together, these studies demonstrate that mobile usability is constrained not only by device characteristics but also by the environment in which applications are used.

2.2 Cognitive load and visual balance in mobile interfaces

Cognitive load theory highlights that the visual structure of an interface strongly influences how easily users can process information. Visual balance, symmetry, and alignment help reduce the mental effort required to interpret elements on a screen. According to Lauer and Pentak [13], symmetry can be achieved through balanced alignment of text, images, and color values, enabling smoother cognitive processing. Conversely, when interfaces lack symmetry, such as through inconsistent alignment or cluttered visual arrangements, users may experience a sense of disorder that increases mental effort and reduces attention.

Bhandari et al. [12] similarly emphasize that visual imbalance in mobile user interfaces elevates cognitive load, particularly on small screens where users must rapidly scan and interpret information. Their findings suggest that consistent alignment, predictable

structure, and clear visual hierarchy are essential for supporting attentional focus and reducing cognitive effort in mobile interaction.

2.3 Accessibility in mobile applications

Accessibility is a fundamental principle in Human-Computer Interaction (HCI) and aims to ensure that digital services can be used by individuals with diverse abilities, contexts, and language backgrounds. Mobile applications require special attention because users interact with them on small screens, in varied lighting conditions, while moving, and often relying on quick visual cues. These conditions can create significant barriers for users with visual, cognitive, motor, or linguistic challenges.

In mobile contexts, accessibility also involves ensuring compatibility with assistive technologies and system-level accessibility settings, such as screen readers, text scaling, and display adjustments. If interface elements are not properly labeled or layouts do not adapt to accessibility settings, users may be unable to access essential information or complete tasks independently.

2.4 WCAG 2.2 as a foundation for accessible mobile design

Although originally developed for web content, WCAG 2.2 are widely applied to mobile applications due to their focus on perceivability, operability, understandability, and robustness [26]. Several WCAG success criteria are particularly relevant for mobile food delivery applications, including:

• Text readability (minimum contrast and scalable text)

• Language of parts (content should not unexpectedly switch languages)

• Clear labels and icons (avoiding reliance on ambiguous symbols)

• Consistent navigation (predictable behavior across screens)

• Avoidance of unnecessary time-based content

2.5 Cognitive accessibility: dyslexia, attention-deficit/hyperactivity disorder (ADHD), and reading difficulties

Text readability is a critical accessibility requirement, particularly for users with dyslexia, low vision, or reading-related difficulties. According to the World Wide Web Consortium (W3C) Web Accessibility Initiative, some users need to adjust text size, spacing, font, and color to read content effectively [11]. Importantly, accessible design

requires that interfaces reflow correctly when such adjustments are made, without loss of information or forced horizontal scrolling.

Users with ADHD may be more affected by interruptions, motion, or unexpected interface changes, which can increase perceived effort and disrupt task flow. As a result, minimizing unnecessary animations and ensuring stable interface behavior can help support sustained attention and more accessible interaction for users with attentional difficulties.

2.6 Visual accessibility and contrast requirements

Visual accessibility plays a key role in mobile interfaces, where users rely heavily on rapid visual scanning. WCAG guidelines recommend sufficient color contrast, clear typography, and adequately sized touch targets [26].

In food delivery applications, those issues may appear in:

• small icons indicating fees

• low-contrast restaurant details

• subtle service-fee labels

• thin or light gray text

Such design choices can be particularly problematic for users with visual impairments or when applications are used in bright sunlight or low-light conditions, as highlighted by Zhang and Adipat [9].

2.7 Multilingual accessibility in Finnish digital services

Finland has a linguistically diverse population. Although Finnish is the first language for most residents, a substantial number of people speak Swedish, Sami, or other languages as their first language, with over 610,000 residents speaking languages other than Finnish, Swedish, or Sami in 2024 [25].

In digital services, this diversity increases the likelihood that users interact with applications in a non-native language. Prior research on multilingual interface design shows that usability is influenced not only by translation accuracy but also by how well interface structure and conventions align with the selected language and cultural context [24].

2.8 Users of food delivery applications

According to the Statista Consumer Insights report Target Audience: Restaurant & Food Delivery Users in Finland, food delivery application users represent a broad demographic spanning all income brackets. The report also indicates that 34% of users fall into the “laggard” category in innovation adoption [8].

This finding highlights the importance of designing food delivery applications that support not only experienced digital users but also individuals with lower technological confidence. For users in the laggard category, unclear navigation, inconsistent feedback, or complex interaction flows may significantly hinder task completion and reduce trust in the service. Consequently, usability and accessibility considerations are essential for ensuring that food delivery platforms remain inclusive and usable for a diverse user base with varying levels of digital literacy.

2.9 Usability principles and heuristic evaluation

ISO 9241-11:2018 defines usability as the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [27]. This definition is useful for interpreting usability testing outcomes in a structured way. Effectiveness reflects whether users can complete tasks successfully, efficiency reflects the effort and smoothness of task completion, and satisfaction reflects users’ subjective comfort, trust, and overall experience during use. In this thesis, the usability testing findings are discussed using this three-part framing to connect observed task performance and user feedback to established usability standards.

Usability evaluation in HCI is often grounded in heuristic principles, general guidelines used to assess interface quality. One of the most widely adopted frameworks is Nielsen’s Ten Usability Heuristics [10], which include principles such as visibility of system status, consistency and standards, recognition rather than recall, error prevention, user control and freedom, and help and documentation.

These heuristics remain highly relevant for mobile applications, where interaction relies heavily on icons, gestures, and condensed layouts. Because of their strong alignment with mobile interaction patterns, Nielsen’s heuristics form the basis for the heuristic evaluation conducted in this study and support the interpretation of usability findings obtained through user testing.

2.10 Previous studies

Previous research on food delivery applications has typically focused on specific populations or regional markets. Juliá-Nehme and Rosell [1] examined older adults' interactions with food delivery applications and identified multiple usability barriers, including unclear feedback, information overload, and difficulty distinguishing interactive elements from static text. While most participants were able to complete tasks, efficiency was low and usability scores fell below established benchmarks. However, the study does not address the general user population in Finland nor compare the country's leading platforms.

Kundu et al. [2] evaluated four food delivery applications in Bangladesh using heuristic evaluation, semiotic analysis, and user surveys. Across methods, they identified recurring issues such as unclear navigation, inconsistent language presentation, ambiguous interactive elements, and information overload. Although these findings suggest that many usability challenges are structural rather than regional, the study does not examine the Finnish context.

Sungboonlue et al. [18] presented a usability-focused redesign of a Thai food delivery application, demonstrating that respecting users' mental models and avoiding unnecessary feature complexity significantly improves perceived usability. While the study does not explicitly address accessibility or European platforms, it provides evidence that many usability problems in food delivery apps are systematic and transferable across markets.

Clariscella and Rahmi [19] analyzed how food delivery application quality influences student satisfaction and reuse intention, finding that content quality and visual presentation were key drivers. However, the study does not address usability or accessibility concerns and does not examine European platforms such as Wolt or Foodora.

In terms of accessibility, Nair et al. [20] investigated the experiences of blind and visually impaired users in food delivery services in the United States. They identified major barriers including screen-reader incompatibility, unlabeled buttons, missing alternative text, and fluctuating accessibility after application updates. This work highlights that accessibility challenges extend beyond interface design to delivery and support processes, but it does not address multilingual mainstream users or the Finnish context.

Similarly, Venkatraman et al. [21] examined accessibility barriers in online restaurant menus for screen reader users and demonstrated how missing structure, unlabeled elements, and poor semantic organization significantly hinder independent use.

Only a small number of studies in Finland have examined food delivery applications, and these are limited to bachelor's theses. Chaudhary [22] conducted a small UX test of Wolt and Foodora with two participants, while Hoang [23] analyzed Wolt's visual design without usability testing or accessibility evaluation. To date, no academic study has systematically evaluated the usability and accessibility of both Wolt and Foodora in the Finnish context, leaving a clear research gap that this thesis addresses.

Together, the literature reviewed in this chapter provides the theoretical foundation for evaluating the usability and accessibility of mobile food delivery applications. The discussed concepts, including mobile interaction constraints, cognitive load, accessibility principles, multilingual usability, and established usability heuristics, highlight design factors that are particularly relevant for food delivery apps, which are frequently used in time-pressured, real-world contexts by a diverse user base. These perspectives inform the selection of evaluation methods, the design of usability testing tasks, and the interpretation of findings in this study. The next chapter describes the methodology used to apply these theoretical concepts to the empirical evaluation of Wolt and Foodora.

3 METHODOLOGY

3.1 Accessibility evaluation

Accessibility evaluation was conducted using a combination of automated and manual methods to assess compliance-related issues and real-world accessibility barriers in the Wolt and Foodora mobile applications. This mixed approach was chosen because automated tools alone cannot fully capture interaction issues experienced by users of assistive technologies, while manual testing benefits from tool-based support to identify systematic technical issues.

Automated accessibility evaluation was performed using Google Accessibility Scanner on Android devices. The tool was used to identify potential accessibility issues related to text alternatives, labeling, color contrast, text resizing, and touch target size. While the scanner does not provide a formal WCAG compliance audit, the detected issues correspond to multiple WCAG 2.2 success criteria, particularly those related to perceivability and operability. The alignment with WCAG 2.2 is especially evident in the detection of touch target size issues, which are defined as Level AA requirements in WCAG 2.2. The automated evaluation was used to systematically screen multiple views of each application and identify recurring accessibility patterns.

Manual accessibility evaluation was conducted using the Android TalkBack screen reader to assess screen reader compatibility and interaction behavior. The evaluation focused on element labeling, reading order, announcement of icons and status information, language handling, and feedback consistency. Manual testing also examined layout behavior under increased system text size and display scaling. This approach allowed the identification of accessibility issues that are not reliably detectable through automated tools alone, such as incorrect screen reader announcements and inconsistent language pronunciation.

3.2 Heuristic evaluation

A heuristic evaluation was conducted to systematically identify usability and accessibility issues in Wolt and Foodora mobile applications. The evaluation was based on Nielsen’s Ten Usability Heuristics [10], which were used as guiding principles throughout the evaluation process, rather than as a post-hoc classification framework.

The evaluation was carried out by two evaluators with backgrounds in UX design and usability. Both evaluators independently inspected the applications while actively considering each of the ten heuristics during interaction, ensuring that observed issues were identified and assessed in relation to specific heuristic principles at the time they occurred.

The evaluators focused on core food-ordering tasks, including browsing restaurants, viewing menus, selecting and customizing items, managing the cart, completing checkout, and accessing order-related support. During the inspection, interface elements, navigation flows, feedback mechanisms, consistency, error prevention, and accessibility-related factors were examined through the lens of the ten heuristics.

Whenever a potential usability issue was identified, it was immediately linked to one or more relevant heuristics and documented with a description and supporting screenshots. This approach ensured that findings were grounded in established usability principles rather than based on ad hoc observations. Accessibility considerations, such as multilingual consistency, cognitive load, and visual clarity, were also taken into account when mapping issues to heuristics.

After completing the individual evaluations, the evaluators compared and discussed their findings to consolidate overlapping issues and agree on final interpretations. Each identified issue was assigned a severity rating from 0 (no usability problem) to 4 (usability catastrophe), following Nielsen’s severity rating guidelines. Severity assessments considered the impact on task completion, frequency of occurrence, and persistence of the issue across the application.

While heuristic evaluation is an effective method for identifying usability and accessibility issues, it has inherent limitations. The evaluation was conducted by two evaluators, which may limit the diversity of perspectives compared to larger expert panels. In addition, heuristic evaluation relies on expert judgment and may not capture all issues experienced by real users in everyday contexts.

Furthermore, the evaluation focused on core food-ordering flows and did not cover all available features of the applications. Finally, although accessibility considerations were taken into account, the evaluation did not include users with disabilities and therefore does not replace user-based accessibility testing.

3.3 Usability testing

Usability testing was conducted to observe how users interact with the Wolt and Foodora mobile applications while performing common food-ordering tasks. Usability

testing is a widely used method in HCI for identifying usability issues by observing users' behavior, task performance, and verbalized thoughts during interaction [28].

The usability tests were conducted remotely. Participants used their own devices in familiar environments, reflecting realistic everyday use of food delivery applications. During the sessions, participants were instructed to think aloud while completing a predefined set of tasks, allowing the researcher to capture users' reasoning, expectations, and points of confusion.

Each session followed the same structure: an introduction and consent confirmation, completion of usability testing tasks, and a short semi-structured interview at the end. The interview was used to collect subjective impressions, perceived difficulties, and trust-related factors that may not be directly observable during task performance.

3.4 Study setup and consent process

The usability testing sessions were conducted using video conferencing software that enabled screen sharing and audio recording. Prior to each session, participants were provided with an information and consent form explaining the purpose of the study, the testing procedure, and their rights, including the right to withdraw at any time without consequences.

Informed consent was obtained before the session began. The study followed ethical user research practices and complied with General Data Protection Regulation (GDPR) principles by minimizing data collection and anonymizing all participant data. Recordings were used solely for research and analysis purposes within this thesis.

3.5 Participants

Five participants took part in the usability testing sessions. Participants were recruited using convenience sampling. The recruitment focused on individuals known to the researcher, as this approach allowed quick access to participants who were willing to participate and represented typical users of food delivery applications. In addition, the recruitment intentionally included participants with minor accessibility-related needs, such as dyslexia, ADHD, or sensitivity to low contrast, to ensure the evaluation reflected a broader range of user experiences. Including participants with these characteristics aligns with inclusive design principles and supports identifying potential barriers that may not be visible when testing only with fully neurotypical users.

The participant group mainly represented the primary target age segment for food delivery services (25-34). All participants had lived in Finland for at least six months.

Most participants reported high confidence in mobile technology use. Prior experience with the evaluated applications varied: the sample included both participants who had never used Wolt or Foodora before and participants who had used one or both applications with different frequencies. This variation allowed the study to capture both novice and experienced-user perspectives.

Gender was not collected in this study because it is not expected to meaningfully affect the use of restaurant delivery platforms in Finland. According to Statistics Finland, gender-based differences in online food ordering are minimal: approximately 23% of men and 24% of women reported ordering meals online in 2023 [7]. Since the behavior of male and female consumers is nearly identical in this context, collecting gender data would not provide additional analytical value for this usability study.

3.6 Questionnaire

Before the usability session, participants filled out a short background questionnaire and consent form. The form collected non-identifying demographic information (age group, experience with food delivery apps, and self-assessed technical skill level). An optional section on accessibility needs (e.g., dyslexia, ADHD, low contrast sensitivity) was included to support an inclusive usability evaluation approach.

No personal identifiers such as name, email, or phone number were collected; each participant was assigned a participant ID to ensure anonymity. The full background and consent form is provided in Appendix A.

3.7 Usability testing tasks

Tasks were selected to reflect common user behaviors and known problem areas identified in previous research, including price transparency, filtering, customization, and users' ability to access help when issues arise. The selection was also informed by Consumer Ombudsman reports highlighting concerns about misleading pricing information and lack of clarity regarding delayed orders [3,4]. Most of the tasks were given in a random order despite codependent ones. The order of tested applications was also switched from one participant to another. The full list of tasks is provided in Appendix B

After completing all tasks, participants took part in a short semi-structured interview. The aim was to capture subjective impressions, perceived difficulties, and trust-related factors. The interview covered themes such as overall experience, navigation,

accessibility, and confidence in using each app. A full list of interview questions is provided in Appendix C.

3.8 Data analysis

The usability testing sessions were audio-recorded and transcribed. The collected qualitative data was analyzed using thematic analysis. An inductive coding approach was applied, allowing themes to emerge from the data rather than being predefined.

The analysis process consisted of the following steps:

- Familiarization with the data through repeated reading of the transcripts.
- Identification of recurring issues, usability problems, and user perceptions.
- Assignment of codes to meaningful data segments related to usability, accessibility, and UX.
- Grouping similar codes into higher-level themes representing common usability and accessibility issues.

The resulting themes were reviewed and refined to ensure consistency and clarity. Representative participant quotes were selected to illustrate each theme. The full thematic coding tables are provided in Appendix C and Appendix D, while the Results section presents a summary of the findings rather than the raw coded data.

3.9 Apps and tools

Applications tested

Both apps were tested using the same researcher account to maintain consistency:

- Wolt (mobile application, version 25.49.9)
- Foodora (mobile application, version 25.48.1)

The same account ensured consistent restaurant listings, saved addresses, and payment settings across sessions.

Applications for the accessibility evaluation

- Android TalkBack was used for manual accessibility evaluation to assess screen reader support, including element labeling, reading order, language handling, and feedback consistency.

- Google Accessibility Scanner was used as an automated accessibility evaluation tool (https://support.google.com/accessibility/android/answer/6376570).

Video conferencing & recording tools

These were used for running remote usability sessions, sharing screens, and capturing data.

- Microsoft Teams - used for remote testing sessions, screen-sharing, recording, and automatic transcription.
- Zoom - used in some sessions for recording and screen-sharing functionality, as one of the participants could not use Teams due to technical difficulties.
- Built-in automatic transcription - From Teams and Zoom to produce initial transcripts for thematic coding. All recordings were stored securely and used only for anonymized analysis.

Devices and operating systems

- Smartphone (participant’s own device) running iOS / Android
- iPad (participant’s own device) - used to test the Foodora and Wolt apps in a real-life mobile context.
- Researcher’s laptop - used to observe, take notes, and store recordings.

Data collection tools

- Google Forms - used for background questionnaires, accessibility questions, and consent collection. No personal identifying data was collected.
- OneDrive - secure storage of recordings, notes, and transcripts.
- Researcher's notes - used during each session to document observations, usability issues, errors, and user behavior.

Analysis tools

- Excel - used for organizing codes, grouping themes, and tracking usability issues.
- Transcripts from Teams/Zoom - cleaned and used as qualitative data for thematic coding.

4 RESULTS

4.1 Automated accessibility evaluation of Foodora

Google Accessibility Scanner was used to conduct an automated accessibility evaluation of the Foodora mobile application on Android. The evaluation covered multiple screens, including the home screen, restaurant listings, promotional banners, category filters, and product cards. The tool reports potential accessibility issues related to screen reader support, touch target size, text exposure, and visual contrast.

One of the most frequently identified issues was unexposed text, where visible information was not included in accessibility labels. This issue affected several types of content, including restaurant ratings (e.g., “3.7 (22 ratings)”), discount information (e.g., “25% off”, “PRO Free delivery”), delivery details, and promotional labels. In these cases, the Scanner indicated that the visible text might not be exposed to accessibility services, meaning that screen readers may not announce this information to users.

The tool also identified duplicate or non-unique accessibility labels across multiple elements. Labels such as “Required”, “Special offers for you”, and “Increase item by 1 unit” were used repeatedly for different interface elements without additional contextual information. The Scanner flagged these cases as multiple items sharing the same description, which may reduce clarity when navigating the interface using a screen reader.

In addition, several interactive elements were reported as potentially missing readable labels. This included large clickable containers such as restaurant cards, rating blocks, and image-based components. In these cases, the Scanner indicated that the item might not have a label readable by screen readers, suggesting that users relying on assistive technologies may not receive sufficient information about the purpose of the element before interacting with it.

WCAG 2.2 Success Criterion 2.5.8 (Target Size, Level AA) defines a minimum target size of 24 × 24 CSS pixels. In Android applications, this requirement is commonly implemented as a minimum touch target size of approximately 48 density-independent pixels (dp), which is reflected in the recommendations provided by Google Accessibility Scanner

The evaluation further revealed touch target size issues affecting small interactive controls, such as information icons, navigation arrows, and quantity adjustment buttons. Some of these elements were reported to be smaller than the recommended minimum size of 48 dp, which may make them more difficult to activate accurately, particularly for users with reduced motor control.

In one instance, the Scanner identified a low text contrast issue, where the contrast ratio between text and background did not meet the recommended minimum. Although this issue was observed less frequently than others, it indicates that certain textual elements may be difficult to perceive for users with low vision or in challenging viewing conditions.

Overall, the automated accessibility evaluation identified multiple types of accessibility issues related to text exposure, labeling, touch target size, and visual contrast. These issues were present across various screens and interface components, demonstrating that while the application is visually functional, several accessibility barriers remain detectable through automated analysis.

4.2 Automated accessibility evaluation of Wolt

Google Accessibility Scanner was also used to conduct an automated accessibility evaluation of the Wolt mobile application on Android. Multiple screens were examined, including the home view, category navigation, promotional banners, restaurant listings, and toolbar elements. During the evaluation, the Scanner reported several accessibility issues related to text exposure, labeling, touch target size, text scaling, and color contrast.

Like the Foodora evaluation, the Wolt application exhibited unexposed text issues, where visible text was not consistently included in accessibility labels. Examples included promotional labels such as “Wolt Rewards BETA,” section headers, pricing information, and category labels. In these cases, the Scanner indicated that visible text might not be available to screen readers, potentially preventing users relying on assistive technologies from accessing essential information.

The evaluation also revealed missing or insufficient accessibility labels on interactive elements. Several icon-based controls, such as filter buttons, toolbar icons, and navigation elements, were flagged as potentially lacking readable labels for screen readers. As a result, the purpose of these elements may not be clear to users navigating the interface using assistive technologies.

In addition, duplicate accessibility descriptions were identified across multiple elements. For example, category items such as “Holiday” were reported as having identical speakable text across more than one clickable element. The Scanner flagged these cases as multiple items sharing the same description, which may reduce clarity and increase ambiguity for screen reader users when navigating between elements.

Touch target size issues were also detected in the Wolt application. Several interactive elements, including toolbar icons and navigation controls, were reported to be smaller than the recommended minimum size of 48 dp. The Scanner indicated that some touch targets were approximately 40 dp by 40 dp, which may make precise interaction more difficult for users with motor impairments or reduced dexterity.

The Scanner further identified text scaling issues, where fixed-size layout containers contained scalable text elements. In such cases, increasing the system font size could cause text to become partially obscured or clipped, particularly within scrollable containers. This suggests that some interface components may not fully support dynamic text resizing.

Unlike the Foodora evaluation, the Wolt application exhibited multiple text contrast issues, observed in both dark mode and light mode. The Scanner reported contrast ratios below the recommended minimum of 4.5:1 for normal text in several instances, including price labels and secondary text elements. These issues may affect the readability of content for users with low vision or in challenging viewing conditions.

During the evaluation, technical limitations were encountered when using the Accessibility Scanner with the Wolt application. The app exhibited noticeable lag, visual glitches, and the appearance of temporary graphical blocks that did not correspond to actual interface elements. These performance issues made it more difficult to reliably interact with the Scanner and may have affected the consistency of the automated results. As a result, the reported findings should be interpreted with caution, as some accessibility issues may not have been detected due to tool–application interaction limitations.

Overall, the automated accessibility evaluation of the Wolt application revealed accessibility issues similar in nature to those identified in Foodora, including unexposed text, insufficient labeling, small touch targets, and text scaling limitations. In addition, Wolt showed more frequent color contrast issues across different display modes. Despite the use of an automated evaluation tool, technical challenges during testing highlight the limitations of automated accessibility assessment for complex, dynamic mobile applications.

4.3 Manual accessibility evaluation of Wolt and Foodora

Manual accessibility testing was conducted using the Android TalkBack screen reader to complement the automated accessibility evaluation and to assess selected WCAG 2.2 success criteria that cannot be fully verified through automated tools. The manual evaluation focused on screen reader output, accessibility labeling, language handling, text scaling behavior, and visual layout under increased accessibility settings. Key screens in both applications were examined, including the home view, restaurant listings, product pages, and the cart view.

Screen reader output and labeling (Foodora)

During TalkBack navigation in Foodora, several issues related to missing or insufficient screen reader output were observed. Delivery-related icons were not announced; instead, TalkBack read only the word “free,” without providing contextual information about delivery conditions. This limits the amount of information available to screen reader users and may lead to ambiguity regarding delivery details.

Category titles were sometimes announced as “heading out of list,” which did not provide a meaningful description of the content or structure of the section. In addition, several icon-based interface elements, such as the “recent” icon, were announced as unlabeled, preventing users from understanding their function.

Product images were not announced at all by TalkBack. Unlike cases where images are identified as “unlabeled,” these elements were entirely skipped during screen reader navigation, meaning that users relying on assistive technologies may not be aware of the presence of product imagery or related content.

These findings relate primarily to WCAG 2.2 Success Criteria 1.3.1 (Info and Relationships) and 4.1.2 (Name, Role, Value), as visible interface information was not consistently exposed to accessibility services.

Screen reader output and labeling (Wolt)

In the Wolt application, TalkBack similarly failed to announce delivery-related icons, resulting in missing contextual information about delivery conditions. Additionally, the progress bar associated with the “Wolt Rewards” feature was not announced, preventing users from perceiving progress or status information through the screen reader.

Several product images were announced as “unlabeled,” indicating that accessibility labels or alternative text were missing. While these elements were detectable by

TalkBack, the lack of descriptive labels reduced their usefulness for users relying on assistive technologies.

These issues also correspond to WCAG 2.2 Success Criteria 1.3.1 (Info and Relationships) and 4.1.2 (Name, Role, Value).

Language handling and screen reader pronunciation (Both applications)

In both Foodora and Wolt, TalkBack attempted to pronounce Finnish-language text using English pronunciation rules when the system language was set to English. This resulted in unclear or incorrect pronunciation of restaurant names, menu items, and category labels.

In the Wolt cart view, the screen reader unexpectedly switched to correct Finnish pronunciation during one interaction, but when the same screen was tested again, it reverted to English-style pronunciation. This inconsistent behavior indicates that language changes within the interface were not reliably communicated to accessibility services.

These observations relate to WCAG 2.2 Success Criterion 3.1.2 (Language of Parts), which requires that changes in language within content are programmatically indicated so that assistive technologies can apply appropriate pronunciation rules.

Text scaling, layout, and visual presentation

Manual testing was conducted by increasing the system text size, both with and without modifying the overall display size.

When only the system text size was increased, Foodora generally handled text resizing well, and no major overlapping or clipping issues were observed. In contrast, Wolt exhibited overlapping of text and promotional advertisements on the main page even when only the system text size was increased, reducing readability and visual clarity.

When both system text size and display size were increased simultaneously, layout issues became more pronounced in both applications. In Wolt, overlapping of promotional text and interface elements was observed on the home screen, as well as overlapping between text and action buttons in the item modification view before adding a product to the cart. In Foodora, text-related issues included recommended items in the cart view being displayed with each letter appearing on a separate line, significantly reducing readability (Figure 1).

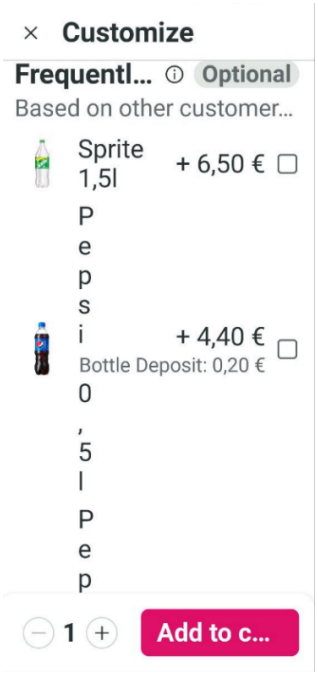

Figure 1. *Layout instability in the Foodora mobile application when system text size and display scaling are increased, resulting in vertically stacked text and reduced readability in the item customization view.*

Additionally, the total price section and confirmation button occupied approximately half of the screen height, limiting the visibility of cart contents. Minor overlapping between text and buttons was also observed in this view.

These findings correspond to WCAG 2.2 Success Criteria 1.4.4 (Resize Text) and 1.4.10 (Reflow), as content did not consistently remain readable or maintain proper layout under increased accessibility settings.

Visual contrast and promotional content (Foodora and Wolt)

In Foodora, promotional labels were displayed using very small text with relatively low contrast, making them difficult to perceive visually. In Wolt, contrast-related issues were also observed, especially in promotional and secondary text elements.

These observations relate to WCAG 2.2 Success Criterion 1.4.3 (Contrast (Minimum)), as some textual elements did not consistently maintain sufficient contrast against their background.

Overall, manual accessibility testing confirmed and extended the issues identified through automated evaluation. Both applications exhibited accessibility barriers related to screen reader labeling, language handling, and layout flexibility under increased accessibility settings.

4.4 Heuristic evaluation results

This section presents the results of a heuristic evaluation of the Wolt and Foodora mobile applications based on Nielsen's ten usability heuristics. The evaluation focused on feedback mechanisms, navigation, consistency, error handling, accessibility-related aspects, and efficiency of use. Identified issues were assigned severity ratings on a scale from 0 (no usability issue) to 4 (usability catastrophe), following Nielsen's severity rating guidelines.

The results are reported under each of Nielsen's ten usability heuristics, which are used as subsection headings. For each heuristic, findings for Wolt and Foodora are presented separately to allow direct comparison between the two applications. Issues identified during the evaluation are accompanied by a severity rating for each app, following Nielsen's severity scale from 0 (no usability issue) to 4 (usability catastrophe). Throughout this section, Wolt is reported first, followed by Foodora, to maintain consistency across heuristics.

Visibility of system status

Wolt provides very strong system status visibility. A spinner is displayed for every loading action, and image loading is handled progressively: food images appear blurred before transitioning into a clear image, clearly indicating the progress. The application relies on animated feedback rather than skeleton screens.

Numerous micro-interactions support feedback clarity, including:

- Animated number changes when adjusting item quantities.
- Unique heart animations when modifying tip amounts.
- Animated indicators showing how bonus points change based on order value.
- Dedicated confirmation screens with both textual and visual confirmation (e.g., tick icon) after payment.
- Refresh animations when navigating to checkout.
- Real-time delivery updates are communicated via notifications. Due to generally fast loading times, slow-loading fallback states could not be evaluated.

Severity: 0 (No usability issue)

Foodora relies primarily on skeleton screens for content loading and uses loading indicators during cart actions. However, loading feedback is inconsistent: in some

places a standard spinner is used, while in others a visually unclear spinning icon or branded Foodora icon appears.

Compared to Wolt, Foodora uses fewer animations for action feedback. Haptic feedback (vibration) is used inconsistently:

- Strong vibration when attempting to open unavailable restaurants or adding restaurants to favorites.
- Medium vibration when adding items

These variations are not explained and may confuse users. Although the app is slightly slower than Wolt, no full freezes were observed.

Severity: 2 (Minor usability issue)

Match between system and the real world

Wolt uses familiar terminology and real-world units consistently (prices in euros, distances in meters/kilometers, time in minutes). Icons align well with their meanings, and food images are generally representative, though occasionally missing.

Restaurant menus vary significantly depending on the restaurant: some are clearly structured, while others contain very large category lists, which may increase cognitive load but reflects real-world variation rather than a system issue.

Severity: 1 (Cosmetic issue)

Foodora also uses real-world terminology and units correctly. However, the app does not clearly indicate when restaurants are closing soon, unlike Wolt, which displays “closing soon” warnings. This may lead users to attempt orders that cannot be completed.

Additionally, restaurants that are closed are still shown, requiring users to first select a scheduled delivery time via a pop-up before proceeding, which disrupts the natural ordering flow.

Severity: 2 (Minor usability issue)

User control and freedom

Wolt supports intuitive navigation, including swipe-back gestures that visually display a left-arrow icon mid-screen during the action. Users can edit carts, change quantities, and return to previous screens easily.

However, entire carts can be deleted without confirmation, which poses a risk of accidental data loss.

Severity: 2 (Minor usability issue)

Foodora similarly allows cart deletion without confirmation. Navigation gestures are supported and visually reinforced with icons. However, user control is weakened by situations where actions silently fail (e.g., items not being added to cart when a restaurant is closing), with feedback only appearing later on a different screen.

Severity: 3 (Major usability issue)

Consistency and standards

Although Wolt demonstrates strong visual and interaction consistency, language consistency is a notable weakness. The application interface is set to English; however, a significant number of restaurants present content partially in Finnish. This inconsistency appears in multiple forms, including:

- Food item names
- Category titles
- Item option labels
- Restaurant-provided descriptions

While Wolt offers a translation feature, its discoverability is poor. The translation option is located below recently ordered items and promotional content, making it difficult to notice, especially for first-time users. Furthermore, the translation function applies only to food descriptions, leaving category names, option labels, and other restaurant-provided text untranslated. As a result, users are frequently required to switch between languages within a single screen.

This behavior violates established interface consistency standards and increases cognitive load, particularly for non-Finnish speakers. It also conflicts with user expectations that the selected application language would be applied consistently throughout the interface.

Severity: 2 (Minor usability issue)

Foodora exhibits similar multilingual consistency issues to Wolt, with the application interface set to English while restaurant-provided content is frequently displayed partially in Finnish. Mixed-language presentation occurs in food item names, category titles, option labels, and additional restaurant descriptions, resulting in screens where users must interpret multiple languages simultaneously.

Although a translation option is available, it is slow to apply and limited in scope. Similar to Wolt, the translation function primarily affects food descriptions, while category names, option labels, and other restaurant-provided text often remain untranslated.

As a result, users may encounter incomplete or delayed translations and are required to manually interpret untranslated elements. This reduces interface predictability and increases cognitive load, particularly for non-Finnish-speaking users. While the underlying issue is comparable to that observed in Wolt, the lack of language selection and slower translation response further weaken consistency and user control.

Severity: 2 (Minor usability issue)

Error prevention

Wolt actively prevents errors by:

- Disabling impossible actions.
- Automatically scrolling users to missing required options.
- Warning users when orders are below the minimum value.
- Blocking orders from closed restaurants.

However, service fees, delivery fees (shown as ranges), and grocery bag fees are only fully revealed at checkout, reducing transparency earlier in the process.

Severity: 2 (Minor usability issue)

Foodora technically marks required item options but does not guide users to them automatically, requiring manual discovery. In one observed case, the app allowed full item configuration but silently refused to add the item to the cart due to restaurant closing, without immediate feedback. Foodora provides improved cost transparency by displaying the service fee already on the restaurant page. However, delivery fees are still presented as a range rather than a fixed value, and additional grocery bag fees are only revealed at checkout, which partially limits early cost predictability.

Severity: 3 (Major usability issue)

Recognition rather than recall

Wolt performs strongly in this area:

- Recently purchased items are displayed when opening a restaurant.
- Orders can be repeated from order history.

- Cart content and order summary are always visible.

Severity: 0 (No usability issue)

Foodora supports order repetition and displays recent items. However, missing guidance for required options forces users to remember information rather than recognize it.

Severity: 2 (Minor usability issue)

Flexibility and efficiency of use

Wolt supports efficient repeat usage through saved orders, favorites, and quick reordering from order history. However, access to the cart is not consistently available across the interface. Unlike Foodora, Wolt does not provide a persistent cart shortcut in the bottom navigation, requiring users to return to the home view or rely on context-dependent entry points to access the cart. This design increases navigation effort, particularly for experienced users who expect quick access to key functions.

Severity: 1 (Cosmetic issue)

Foodora provides a persistent cart tab in the bottom navigation, enabling quick and direct access to the cart from most views, which supports efficient interaction for repeat users. However, overall efficiency is reduced by confusing filter logic and a slower navigation flow compared to Wolt. In addition, Foodora does not provide a dedicated restaurant navigation tab, while groceries are accessible through a persistent tab, which may reduce navigational clarity for some users.

Severity: 2 (Minor usability issue)

Aesthetic and minimalist design

Wolt maintains a clean, uncluttered layout with restrained animations. Touch targets are generally sufficient, though icons could be slightly larger for improved accessibility.

One unclear design element is a play/pause icon beneath home-screen advertisements, which controls promotional rotation but is not self-explanatory and lacks contextual explanation.

Severity: 2 (Minor usability issue)

Foodora’s interface appears more cluttered, particularly due to animated promotional content. Visual noise and moving elements increase cognitive load, especially on smaller screens. Touch targets and icons could be larger for a better accessibility experience.

Severity: 3 (Major usability issue)

Help users recognize, diagnose, and recover from errors

Wolt provides clear, descriptive error messages and allows users to contact support directly about specific past orders. Help is most easily accessed via the Profile section, while alternative paths are less discoverable.

Severity: 1 (Cosmetic issue)

Foodora provides order-related help and support options, but delayed or missing error feedback (e.g., unavailable restaurants) limits users' ability to recover from errors efficiently.

Severity: 3 (Major usability issue)

Help and documentation

In Wolt help and support are well integrated, with contextual options linked to specific orders. Discoverability could be improved outside the Profile section.

Severity: 1 (Cosmetic issue)

Foodora support options are almost the same as in Wolt and can be improved in the identical way.

Severity: 1 (Cosmetic issue)

4.5 Usability testing results

The usability testing results are reported thematically, based on patterns observed across task performance, user behavior, and post-task interviews. Each theme may include findings from both Wolt and Foodora. When an issue is not explicitly attributed to a specific application, it applies to both apps. This reporting approach emphasizes shared and contrasting usability experiences rather than task-by-task outcomes. Identified usability issues were analyzed using thematic coding; the full coding tables are provided in Appendix C and Appendix D.

Overall task performance and user confidence

Overall, participants were able to complete most of the assigned tasks in both Wolt and Foodora without major difficulty. All participants described both applications as generally intuitive and familiar, and several noted that the task flow was easier than expected.

Despite this overall confidence, several tasks revealed recurring usability and accessibility issues, particularly in Foodora. While participants reported feeling capable of using both applications, all participants stated that they either preferred Wolt or trusted it more. This preference was attributed to clearer structure, fewer interruptions, and more predictable behavior.

Task completion issues

Some tasks could not be completed successfully by all participants:

- Changing payment methods (Foodora):

Two participants were unable to add or modify payment methods. They attempted to complete the task via the payment section in account but found only empty page, with no clear option to add new ones. One participant stated:

“It just shows me that it’s empty here in payment method, so I cannot even add anything.”

- Finding information about delayed orders:

One participant was unable to complete this task entirely and searched unsuccessfully within the order confirmation page in both applications. Other participants partially completed the task by stating they would contact customer support, although neither app provided explicit information about handling delayed orders.

- Accessibility settings:

One participant did not find accessibility-related settings in either app, as they did not open the general settings menu and did not expect accessibility options to be located there.

- Filtering restaurants (Foodora):

Two participants were unable to complete the filtering task due to technical or functional issues. All other tasks were completed successfully by most participants.

Language inconsistency

Almost all participants observed that both Wolt and Foodora displayed restaurant information partially in Finnish despite the interface language being set to English. These inconsistencies appeared in menu descriptions, ingredient lists, categories, and promotional banners.In some cases, content was also displayed in Chinese, as illustrated in Figure 2.

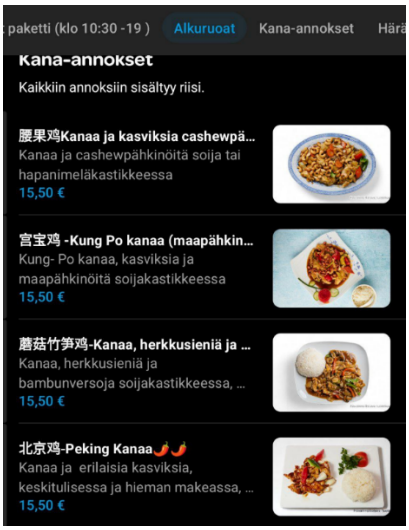

Figure 2. *Example of multilingual content in the Wolt mobile application, where restaurant menu items are shown in Finnish and Chinese.*

Participants reported that unexpected language switching slowed down decision-making and required additional effort to interpret information. In contrast, the participant with dyslexia did not notice the language inconsistency. They explained that they relied primarily on visual cues rather than text: “If there are pictures, I am never reading the text.”

Language and trust

One participant expressed concern about whether customer support would be accessible in English, stating uncertainty about the language used in support interactions. Although both applications offer language selection in support chats, this was not noticed during testing.

This concern appeared to be linked to the broader pattern of language inconsistency across the interfaces, which reduced participants’ confidence and trust in the applications’ predictability.

Intrusive promotional pop-ups and visual noise

Participants encountered frequent promotional pop-ups in Foodora, including animated discount offers and countdown timers that reappeared after dismissal, as illustrated in Figure 3.

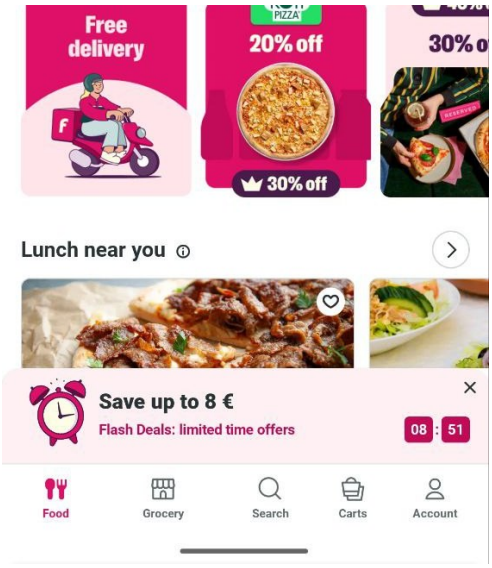

Figure 3. *Promotional pop-up in the Foodora mobile application displaying a time-limited discount offer with a countdown timer.*

These pop-ups disrupted task flow and drew attention away from the main interface. One participant described the experience as follows: “Advertisement everywhere and is popping; advertisement pop-ups and moving images. They distract me. I want to order food.”

Additionally, Foodora home page pop-ups appeared in Finnish even when the app language was set to English, further contributing to confusion and unpredictability, as illustrated in Figure 4.

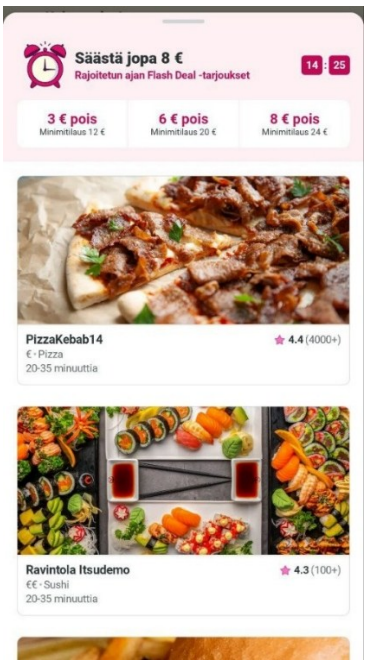

Figure 4. *Promotional banner in the Foodora mobile application displayed in Finnish despite the interface language being set to English.*

Wolt also displayed pop-up elements, but these were smaller, informational, appeared only once, and did not use animation. None of the participants described Wolt's pop-ups as disruptive.

Information findability

None of the participants were able to locate clear information in either application on what to do if an order was delayed. Participants searched within order details and help sections but did not find guidance related to delays or cancellations. As a result, they assumed they would need to contact customer support directly.

Price transparency and price formation

Several participants experienced confusion regarding how the final price of an order was calculated. Some assumed that the price shown in the restaurant menu was final and expected only a delivery fee to be added. The presence of service fees was often overlooked.

In Wolt, the service fee was only noticed at checkout, while in Foodora it was displayed under the restaurant rating but described as visually subtle. One participant commented: "I don't understand: the delivery fee was zero euro and here I can see it's one euro ninety-nine." Participants reported surprise at the final price and uncertainty about what each fee represented.

Technical issues

Technical problems were observed exclusively in the Foodora application. One participant using an iPad experienced repeated freezing and unresponsive screens: "Ok, it is not moving. I think it froze… I have to reload the app again."

Another participant reported that the app became slow when modifying cart items, displaying loading screens after each interaction. These delays disrupted task flow and reduced perceived responsiveness.

No technical issues were observed in Wolt across any device.

Overall user satisfaction

Despite the identified issues, participants reported overall satisfaction with both applications. Several described the apps as intuitive and easier to use than expected. However, aggressive pop-ups, visual clutter, filtering problems caused Foodora to be perceived as less reliable.

In contrast, Wolt was consistently described as calmer, clearer, and more trustworthy, leading all participants to express a preference for using it in everyday situations.

5 DISCUSSION

5.1 Combining automated and manual accessibility evaluation

The findings of this study demonstrate that automated and manual accessibility evaluation methods play complementary roles when assessing mobile applications. Automated evaluation efficiently identified recurring technical accessibility issues, including unexposed text, insufficient accessibility labels, small touch targets, and contrast problems. The repetition of these issue types across multiple screens suggests that accessibility barriers in both applications are systemic rather than isolated implementation errors.

However, manual testing revealed accessibility and usability issues that were not consistently detectable through automated tools. Screen reader testing exposed missing or misleading announcements, inconsistent language pronunciation, and layout breakdowns when accessibility settings were increased. These findings support prior research emphasizing that mobile application evaluation must account for real-world usage constraints, such as small screen size, varied lighting conditions, and divided attention [9], which automated tools alone cannot fully capture.

5.2 Accessibility, language handling, and information perception

Language inconsistency emerged as one of the most significant accessibility challenges in both applications. Restaurant content frequently appeared in Finnish or Chinese despite the interface language being set to English, resulting in incorrect screen reader pronunciation and reduced predictability. This behavior violates WCAG 2.2 Language of Parts requirements and aligns with research showing that usability in multilingual interfaces depends not only on translation availability but also on consistent language signaling and predictable structure [24].

In the Finnish context, where a substantial proportion of users interact with digital services in a non-native language [25], such inconsistencies can significantly affect comprehension, trust, and perceived reliability. Manual accessibility testing further showed that essential information, such as delivery conditions and progress indicators,

was not consistently exposed to assistive technologies, limiting users' ability to form an accurate understanding of system state.

5.3 Cognitive load, visual design, and interaction stability

The usability issues observed in this study reflect cognitive load mechanisms described in prior research. Foodora's intrusive promotional pop-ups, animated elements, and visually subtle fee indicators increased mental effort and disrupted task flow. These design characteristics correspond to the types of visual imbalance and attentional disruption identified by Lauer and Pentak [13] and Bhandari et al. [12] as increasing cognitive load on mobile devices. This study extends these findings by demonstrating how cognitive load is amplified in commercial mobile applications when visual imbalance is combined with time pressure, pricing uncertainty, and promotional interruptions.

In contrast, Wolt's more restrained visual design, consistent feedback mechanisms, and limited use of animation appeared to support smoother cognitive processing. Participants described Wolt as calmer and more predictable, suggesting that visual balance and interaction consistency reduced the effort required to interpret information and complete tasks. These findings reinforce the importance of minimalism, predictability, and visual hierarchy in mobile interface design, particularly under cognitively demanding conditions.

5.4 Usability testing in terms of effectiveness, efficiency, and user satisfaction

From the perspective of established usability standards, particularly ISO 9241-11:2018, the usability testing results can be interpreted in terms of effectiveness, efficiency, and user satisfaction. In terms of effectiveness, most participants were able to complete the assigned tasks in both applications, indicating that the core functionality supports successful task completion. However, certain tasks, such as changing payment methods, finding information about delayed orders, and filtering restaurants in Foodora, could not be completed by all participants, revealing gaps in task support.

Efficiency differed more clearly between the two applications. Wolt generally enabled smoother task execution through clearer feedback, predictable navigation, and fewer interruptions. In contrast, Foodora's intrusive promotional content, inconsistent feedback, and occasional technical delays disrupted task flow and increased the effort

required to complete tasks, reducing overall efficiency even when tasks were ultimately completed.

5.5 Overall usability and accessibility trade-offs between Wolt and Foodora

Rather than identifying a clearly superior application, the findings reveal distinct usability and accessibility trade-offs. Foodora offers advantages in cart discoverability and earlier visibility of service fees but is undermined by intrusive promotional content, layout instability under accessibility settings, and greater cognitive disruption. Wolt provides clearer system feedback and a calmer interaction experience but struggles with multilingual consistency, incomplete translation coverage, and limited early cost transparency.

These trade-offs illustrate that usability and accessibility are shaped not by individual features alone but by the cumulative interaction of visual design, language handling, feedback mechanisms, and accessibility support. Given the broad and diverse user base of food delivery applications in Finland, including less tech-confident users [8], addressing accessibility, language consistency, and promotional design is essential for improving inclusivity, trust, and overall UX.

5.6 Recommendations for improving usability, accessibility, and consumer clarity

Based on the findings, the following key recommendations are identified for improving usability, accessibility, and user trust in food delivery applications:

- *Improve early price transparency by clearly presenting all mandatory fees before checkout.*
- *Provide clearer and more accessible guidance for problem situations such as delayed or cancelled orders.*
- *Ensure consistent and user-controlled language handling across all interface elements and restaurant content.*
- *Strengthen accessibility support by improving screen reader labeling, touch target sizes, and announcement of system status.*
- *Improve layout stability when text size and display scaling are increased.*
- *Address application-specific usability issues, including filtering and payment flows in Foodora and cart accessibility in Wolt.*

First, price transparency should be improved earlier in the ordering flow. Mandatory costs such as service fees and additional charges should be clearly presented before checkout rather than appearing only at the final stage. Participants frequently misunderstood how the final price was formed, and this aligns with concerns raised by the Finnish Consumer Ombudsman regarding insufficiently clear pricing practices in food delivery services [3]. Presenting all mandatory fees in a prominent and consistent manner would support informed decision-making and reduce user frustration.

Second, support information for problem situations should be made more visible and comprehensive. Participants were unable to find clear guidance on what to do if an order is delayed, how cancellations work in practice, or when contacting support is necessary. As a result, users assumed they would need to message customer support for most issues. Expanding FAQ sections and making them easier to access could reduce unnecessary support contact and increase users' sense of control. This recommendation is consistent with consumer complaints noted by the Consumer Ombudsman regarding delayed deliveries and uncertainty around cancellation procedures [4].

Third, language handling should be made consistent and user-controlled across the entire interface. All interface elements and restaurant-provided content should either follow the selected application language or be reliably and fully translated. Although both Wolt and Foodora offer an automatic translation feature, none of the participants were able to locate it during usability testing, and it was only discovered during heuristic evaluation. In addition, the translation function was inconsistent and applied only to certain content types, leaving categories, options, and labels untranslated. Given the large variation between restaurants, ranging from very simple menus to highly complex ones with many categories and options, incomplete translation significantly increases cognitive load and the risk of misunderstanding. Making translation options easier to find, applying them consistently, or clearly indicating untranslated content would improve predictability and accessibility in a multilingual context.

Related to this, allergen and ingredient information requires particular attention. Wolt explicitly warns users that automatic translations may be incorrect and advises contacting the restaurant directly regarding allergens. However, restaurant comments are not always reviewed, and users may still rely on translated menu text when making decisions. If allergen information is mistranslated or missing, users may unintentionally order food they cannot safely consume. Ensuring that allergen-related information is

clearly labelled, reliably translated, or highlighted as requiring confirmation would reduce health-related risks.

Fourth, accessibility support should be strengthened across both applications. Essential information must be consistently exposed to assistive technologies by providing meaningful accessibility labels for icons, images, and interactive elements, avoiding non-unique labels, and ensuring that progress indicators and delivery information are announced correctly by screen readers. In addition, touch targets and clickable elements should be increased in size to meet WCAG 2.2 target size recommendations, improving usability for users with motor impairments and reducing accidental interactions for all users.

Fifth, layout stability under increased accessibility settings should be improved. Both applications exhibited overlapping, clipping, or reduced readability when text size and display scaling were increased. Ensuring proper reflow and preventing layout breakdowns would support users with dyslexia, low vision, or other reading-related difficulties and reduce reliance on visual guessing.

Finally, several application-specific usability improvements are recommended. In Foodora, filtering functionality and payment-related flows should be fixed to prevent task failure, and adding a persistent restaurant navigation tab, similar to the existing grocery tab, could improve discoverability and navigation efficiency. In Wolt, access to the cart could be made more consistent and visible across views to reduce navigation effort and improve user control.

Together, these recommendations highlight that improving usability and accessibility in food delivery applications requires addressing not only individual interface elements but also broader issues of language consistency, information transparency, accessibility support, and predictable interaction design.

6 CONCLUSIONS

The objective of this thesis was to evaluate and compare the usability and accessibility of the Wolt and Foodora mobile applications in the Finnish context. The study employed a mixed-method approach combining usability testing with end users, heuristic evaluation based on Nielsen's ten usability heuristics, and automated and manual accessibility assessment aligned with WCAG 2.2. By applying multiple complementary evaluation methods, the study aimed to provide a comprehensive understanding of how usability and accessibility challenges manifest in widely used food delivery applications.

6.1 Summary of findings

Research question 1: What usability and accessibility issues do users experience when using Wolt and Foodora in Finland?

The study identified several recurring usability and accessibility issues affecting both applications. Common usability problems included limited price transparency before checkout, difficulty locating information related to delayed orders, and inconsistent error feedback. From an accessibility perspective, both applications exhibited insufficient screen-reader support, including missing or unclear accessibility labels, unannounced icons, and incomplete exposure of visible information to assistive technologies. In addition, both applications demonstrated layout instability when text size and display scaling were increased, leading to overlapping elements, clipped content, or reduced readability. Language inconsistency was a particularly significant issue, as restaurant content frequently appeared in languages different from the selected interface language, affecting comprehension and screen-reader pronunciation.

Research question 2: How do these apps compare in terms of UX and accessibility?

The comparative analysis showed that Wolt and Foodora differ in how usability and accessibility issues affect the overall UX. Wolt generally provided clearer system feedback, smoother navigation, and higher perceived stability, which contributed to greater user trust and satisfaction. However, Wolt showed notable weaknesses in multilingual consistency, incomplete translation coverage, and limited early visibility of

mandatory fees. Foodora, while offering earlier visibility of some pricing information, was more frequently associated with increased cognitive load due to intrusive promotional content, inconsistent feedback in failure situations, and occasional technical performance issues. From an accessibility standpoint, both applications exhibited similar categories of problems, but Wolt showed more frequent contrast issues, while Foodora displayed more pronounced layout and interaction disruptions under increased accessibility settings.

Research question 3: What recommendations can be made to improve usability and inclusiveness in food delivery applications?

Based on the findings, several recommendations can be made to improve usability and accessibility in food delivery applications. First, mandatory costs such as service fees and additional charges should be clearly presented earlier in the ordering process to improve price transparency and support informed decision-making. Second, guidance for problem situations, including delayed orders and cancellations, should be more visible and explicit to reduce user uncertainty and reliance on customer support. Third, language handling should be consistent and user-controlled across the entire interface, ensuring that all restaurant-provided content follows the selected application language or is reliably translated. Fourth, accessibility support should be strengthened by improving accessibility labeling, ensuring that essential information is exposed to screen readers, increasing touch-target sizes, and maintaining sufficient text contrast. Finally, interface layouts should be designed to remain stable and readable when accessibility settings such as increased text size or display scaling are applied.

6.2 Contributions, limitations, and future work

This thesis contributes to the limited body of research on food delivery applications in Finland by providing a systematic, comparative evaluation of the two dominant platforms using established usability and accessibility frameworks. By combining usability testing with heuristic and accessibility evaluation, the study offers practical insights into how design decisions affect both general UX and inclusive access in real-world mobile applications. The findings demonstrate that accessibility challenges in mainstream consumer applications often arise from cumulative interaction design decisions rather than from isolated technical failures.

Beyond the empirical findings, this thesis contributes methodologically by demonstrating how automated accessibility tools, manual screen reader testing, heuristic evaluation, and usability testing can be systematically combined to evaluate

mobile applications in a real-world context. The study also highlights the importance of treating multilingual consistency as both a usability and accessibility concern, rather than as a purely localization-related issue. These insights extend existing usability research by showing how accessibility barriers emerge from the interaction between language handling, visual design, and dynamic interface behavior in mainstream consumer applications.

This study has several limitations. First, participants were recruited via convenience sampling, and the sample size was small (5 participants), which supports qualitative issue discovery but limits representativeness and statistical generalization. The participant group also consisted mostly of tech-confident users, which may underestimate difficulties faced by less experienced users. Second, usability sessions were conducted remotely on participants' own devices and networks, and differences in device type, OS version, and performance may have influenced observed issues (e.g., delays or freezing). Third, accessibility evaluation combined automated and manual methods, but automated scanner results are indicative rather than a formal WCAG compliance audit, and the manual screen reader testing was limited in scope and did not include users with significant disabilities. Finally, the evaluated applications evolve rapidly through updates, and some findings may change as features and user interface elements are revised.

Future research should involve larger and more diverse participant groups, include users with disabilities directly, and conduct cross-platform and longitudinal evaluations to assess how usability and accessibility evolve across application updates.

REFERENCES

[1] Juliá-Nehme, B., & Rosell, J. (2025). Interaction and Design Barriers for Older Adults in Food Delivery Apps: A Usability Study. *International Journal of Human–Computer Interaction*, 41(9), 5761–5778. https://doi.org/10.1080/10447318.2024.2365484

[2] Kundu, Dipanjali, et al. Evaluating the Usability of Food Delivery Mobile Application in Bangladesh with Heuristics: A Multimethod Study. *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)*, IEEE, pp. 1–6, https://doi.org/10.1109/QPAIN66474.2025.11171641

[3] Finnish Competition and Consumer Authority. (2025). Consumer Ombudsman intervened in Foodora's practices – misleading information was provided about prices and the duration of offers. https://www.kkv.fi/en/current/press-releases/consumer-ombudsman-intervened-in-foodoras-practices-misleading-information-was-provided-about-prices-and-the-duration-of-offers/

[4] Finnish Competition and Consumer Authority. (2025). Platform company Wolt ceased competition-restricting terms during FCCA investigation – upcoming report will reveal the impact of changes. https://www.kkv.fi/en/current/press-releases/platform-company-wolt-ceased-competition-restricting-terms-during-fcca-investigation-upcoming-report-will-reveal-the-impact-of-changes/

[5] Statista. (2025). Online food delivery bookings by brand in Finland as of June 2025 [Graph]. https://www-statista-com.libproxy.tuni.fi/forecasts/1188039/online-food-delivery-bookings-by-brand-in-finland

[6] Statista. (2023). Share of shoppers buying groceries and meal deliveries online in Finland in 2023, by age group [Graph]. https://www-statista-com.libproxy.tuni.fi/statistics/1399008/online-grocery-shoppers-by-generation-finland/

[7] Statista. (2023). Share of shoppers buying groceries and meal deliveries online in Finland in 2023, by gender [Graph]. https://www-statista-com.libproxy.tuni.fi/statistics/1399029/grocery-shoppers-by-gender-finland/

[8] Statista. (2024). Target audience: Restaurant & food delivery users in Finland. Statista Consumer Insights Global. https://www-statista-com.libproxy.tuni.fi/study/119386/target-audience-restaurant-and-food-delivery-users-in-finland/

[9] Zhang, Dongsong, and Boonlit Adipat. Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction*, vol. 18, no. 3, 2005, pp. 293–308, https://doi.org/10.1207/s15327590ijhc1803_3

[10] Nielsen, J. (1994). Ten usability heuristics for user interface design. Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/

[11] W3C Web Accessibility Initiative. (n.d.). Customizable text. W3C. https://www.w3.org/WAI/perspective-videos/customizable/

[12] Bhandari, Upasna, et al. Cognitive Load and Attention for Mobile Applications: A Design Perspective. *Human-Computer Interaction. Interaction Platforms and Techniques*, edited by Masaaki Kurosu, Springer International Publishing, 2016, pp. 278–84, https://doi.org/10.1007/978-3-319-39516-6_26

[13] Lauer, D. A., & Pentak, S. (2011). *Design basics* (8th ed.). Cengage Learning.

[14] Rungruangjit, Warinrampai, and Kitti Charoenpornpanichkul. What Motivates Consumers' Continued Usage Intentions of Food Delivery Applications in Post-COVID-19 Outbreak? Comparing Generations X, Y and Z. *Journal of Asia Business Studies*, vol. 18, no. 1, 2024, pp. 224–51, https://doi.org/10.1108/JABS-06-2023-0234

[15] Eskicioglu, Omer Can, et al. Safe and Sound Mobile Application: A Solution for Aid People with Visual Disabilities' Mobility. *Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion*, ACM, 2020, pp. 22–28, https://doi.org/10.1145/3439231.3440616

[16] Kim, Young Hoon, et al. A Study of Mobile User Engagement (MoEN): Engagement Motivations, Perceived Value, Satisfaction, and Continued Engagement Intention. *Decision Support Systems*, vol. 56, 2013, pp. 361–70, https://doi.org/10.1016/j.dss.2013.07.002

[17] Punchoojit, Lumpapun, and Nuttanont Hongwarittorrn. Usability Studies on Mobile User Interface Design Patterns: A Systematic Literature Review. *Advances in Human-Computer Interaction*, edited by Thomas Mandl, vol. 2017, no. 2017, pp. 1–22, https://doi.org/10.1155/2017/6787504

[18] Sungboonlue, Pongsatorn, et al. A Study of Redesigning Food Delivery Application in Thailand. *Telkomnika*, vol. 20, no. 5, 2022, pp. 1073–82, https://doi.org/10.12928/telkomnika.v20i5.24094

[19] Clarisella, K., & Rahmi, R. (2025). Online Food Delivery App Quality and Its Impact on Preferences and Satisfaction. *Journal of Distribution Science*, *23*(7), 99–116. https://doi.org/10.15722/JDS.23.07.202507.99

[20] Nair, Adit, et al. Investigating the User Experience and Challenges of Food Delivery Applications for the Blind and Visually Impaired. *Proceedings of the 25th International Conference on Mobile Human-Computer Interaction*, edited by Luis A. Leiva et al., ACM, 2023, pp. 1–6, https://doi.org/10.1145/3565066.3608688

[21] Venkatraman, Nithiya, et al. AccessMenu: Enhancing Usability of Online Restaurant Menus for Screen Reader Users. *Proceedings of the 22nd International Web for All Conference*, ACM, 2025, pp. 36–47, https://doi.org/10.1145/3744257.3744275

[22] Chaudhary, R. (2023). User experience testing in food delivery app: Case study of Foodora and Wolt [Bachelor's thesis, Laurea University of Applied Sciences]. Theseus. https://www.theseus.fi/handle/10024/803893

[23] Hoang, Q. (2024). The impact of visual and interaction design on user experience: A case study of Wolt's application [Bachelor's thesis, Metropolia University of Applied Sciences]. Theseus. https://www.theseus.fi/handle/10024/877355

[24] Hillier, Mathew. The Role of Cultural Context in Multilingual Website Usability. *Electronic Commerce Research and Applications*, vol. 2, no. 1, 2003, pp. 2–14, https://doi.org/10.1016/S1567-4223(03)00005-X

[25] Clark, D. (2025). Population of Finland 2014–2024, by language [Graph]. https://www.statista.com/statistics/529490/population-of-finland-by-language/

[26] World Wide Web Consortium (W3C). (2024). Web Content Accessibility Guidelines (WCAG) *2.2* (W3C Recommendation). https://www.w3.org/TR/WCAG22/

[27] International Organization for Standardization. (2018). ISO 9241-11:2018 Ergonomics of human-system interaction, Part 11: Usability: Definitions and concepts. ISO.

[28] Rubin, Jeffrey, and Dana Chisnell. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*, *Second Edition*. 2nd ed., Wiley, 2008.

APPENDIX A: BACKGROUND QUESTIONNAIRE AND CONSENT FORM

Participant information & consent – Wolt and Foodora usability research

This form collects basic background information and your consent to participate in a usability study on food delivery applications in Finland.

No personal identifying data (such as your name or email) is collected.

You will be assigned a participant number for anonymity.

The study involves evaluating the usability and accessibility of two food delivery apps (Wolt and Foodora).

Your answers in this form will help the researcher understand your previous experience with similar apps and any accessibility needs relevant to mobile app use.

Participation is voluntary, and you may withdraw at any time.

Researcher: Alina Ida, Master's student, Tampere University

Supervisor: Päivi Majaranta

If you have questions about the study, you may contact the researcher (alina.ida@tuni.fi)

Participant ID (given by the researcher) *

Мой ответ

Age group *

- ◯ 18-24
- ◯ 25-34
- ◯ 35-44
- ◯ 45-54
- ◯ 55-64
- ◯ 65+

Have you lived in Finland for at least 6 months? *

- ◯ Yes
- ◯ No

What food delivery apps have you used before? *

- ☐ Wolt
- ☐ Foodora
- ☐ Uber Eats
- ☐ DoorDash
- ☐ Bolt Food
- ☐ Other
- ☐ I have never used any food delivery apps

How often do you use food delivery apps *

- ◯ Never
- ◯ Less than once a month
- ◯ 1-3 times per month
- ◯ 1-2 times per week
- ◯ 3+ times per week
- ◯ Prefer not to say

If you ever used Wolt, how often do you use it

- ◯ Less than once a month
- ◯ 1-3 times per month
- ◯ 1-2 times per week
- ◯ 3+ times per week

If you ever used Foodora, how often do you use it

- ◯ Less than once a month
- ◯ 1-3 times per month
- ◯ 1-2 times per week
- ◯ 3+ times per week

Your technical skill level (in mobile apps/technology) *

- ◯ Very low (I struggle with mobile apps)
- ◯ Low
- ◯ Average
- ◯ High
- ◯ Very high (I am confident with technology)

Accessibility (optional & anonymous)

You may skip this section. This information helps evaluate accessibility, but it is entirely optional

Do you have any accessibility needs that affect your mobile app usage?

- [ ] Difficulty reading small text
- [ ] Difficulty seeing low-contrast text or icons
- [ ] Color blindness
- [ ] Dyslexia or other reading difficulties
- [ ] ADHD or attention-related difficulties
- [ ] Motor/dexterity difficulties
- [ ] Hearing difficulties
- [ ] Other
- [ ] I prefer not to answer
- [ ] No accessibility needs

Do you normally use any accessibility settings on your phone?

- [ ] Larger text size
- [ ] High contrast mode
- [ ] Screen reader
- [ ] Voice control
- [ ] Reduced animations
- [ ] None
- [ ] Prefer not to say

Consent

By participating in this study, I confirm that: *

- [ ] I am 18 years old or older
- [ ] I voluntarily agree to participate in this usability study
- [ ] I understand that the session will be recorded (audio/video) and automatically transcribed
- [ ] I understand that the data will be anonymized and used only for research purposes in a Master's thesis at Tampere University
- [ ] I understand that I may withdraw at any time without giving a reason
- [ ] I agree to the recording and analysis of my interaction during the test session

APPENDIX B: USABILITY TESTING TASKS

Task 0. Home page exploration

Scenario:

You just opened the app as if you were planning to order food later today. Take a moment to look around the home page and explore what you see.

Goal:

Evaluate first impressions, layout understanding, and overall navigation clarity.

Success criteria:

Participant can describe what they see and navigate without confusion.

Task 1. Finding something to eat

Scenario:

It’s close to lunchtime. Explore the app and find any restaurant or dish you might want to order.

Goal:

Assess natural browsing behavior and discoverability of content.

Success criteria:

Participant reaches any restaurant or dish in a reasonable time.

Task 2. Finding sushi

Scenario:

You suddenly feel like sushi. Find sushi options available near you.

Goal:

Evaluate search efficiency, input usability, and relevance of results.

Success criteria:

Participants find sushi restaurants or dishes.

Task 3. Finding fast delivery options

Scenario:

You are very hungry and want food as soon as possible. Find restaurants that can deliver in about 20 minutes.

Goal:

Examine filter discoverability and understanding of delivery-time filters.

Success criteria:

Participants apply the correct filter to narrow results.

Task 4. Checking restaurant details

Scenario:

Before ordering, you want to see important details. Open any restaurant and find:

– Delivery fee

– Minimum order

– Estimated delivery time

Goal:

Evaluate clarity and accessibility of restaurant information.

Success criteria:

Participant correctly identifies all three required details.

Task 5. Ordering for yourself and friends

Scenario:

You’re having a movie night. Choose a dish you’d like, customize it, and add it to the cart. Then add three drinks for your friends.

Goal:

Assess customization flow, option visibility, and multi-item ordering.

Success criteria:

Participant successfully adds a customized item and three drinks.

Task 6. Updating the cart

Scenario:

Your friends tell you they have already bought drinks, and they want you to order two portions of your dish instead of one. Update your cart accordingly.

Goal:

Test clarity of edit options, deletion, duplication, and price transparency.

Success criteria:

Participants modify the cart without assistance.

Task 7. Reviewing order before checkout

Scenario:

You are ready to order but want to double-check everything. Go to check out and review all important details. Stop before paying.

Goal:

Evaluate checkout layout, trust indicators, and clarity of fees.

Success criteria:

Participant finds total cost, delivery fees, estimated delivery time, and address.

Task 8. Changing the delivery address

Scenario:

You notice the app is using the wrong delivery address. Update it to the correct one.

Goal:

Assess discoverability and usability of address management.

Success criteria:

Participants update their address successfully.

Task 9. Exploring accessibility options

Scenario:

You want to make the app more comfortable to use. Perhaps larger text, language changes, or dark mode. Look for any accessibility or display-related settings.

Goal:

Determine whether accessibility settings are present and discoverable.

Success criteria:

Participant finds accessibility-related settings or confidently reports none.

Task 10. What to do if order is delayed

Scenario:

Imagine your order might arrive late. Find information about what the app recommends in that situation.

Goal:

Evaluate help center usability and information architecture.

Success criteria:

Participant reaches the correct help article or support instructions.

Task 11. Adding a payment method

Scenario:

You want to prepare the app for future orders. Try to add a new payment method but stop before confirming anything.

Goal:

Assess ease of finding payment settings and user trust with payment flow.

Success criteria:

Participant reaches payment method screen without finalizing.

Task 12. Checking past orders

Scenario:

You want to see what you ordered previously. Find the order history and open it.

Goal:

Evaluate recognizability and clarity of order history.

Success criteria:

Participant opens order history.

APPENDIX C: INTERVIEW QUESTIONS

1. Overall experience

How would you describe your overall experience using each app (Wolt and Foodora)?

2. Task usability

Which tasks felt the most difficult or confusing?

What made them difficult?

3. Navigation & interface

How easy was it to find the features you needed (e.g., filters, restaurant information, cart, help)?

Did the layout or structure of either app feel more intuitive? Why or why not?

4. Accessibility & language

Did you encounter any elements (text, icons, colors, contrast, labels) that were hard to understand or read?

Did either app feel more user-friendly for someone with accessibility needs (e.g., dyslexia, ADHD, vision sensitivity)?

5. Pop-ups, distractions & interruptions

How did you feel about any pop-ups or promotional screens that appeared while using the apps?

Did they help, distract, or confuse you?

6. User confidence & trust

Did you feel confident performing tasks in each app?

If not, what made you unsure?

Which app would you trust more to place a real order? Why?

7. Final comparison & improvement

If you could change one thing in Wolt and one thing in Foodora to improve usability or accessibility, what would it be?

APPENDIX D: THEMATIC CODING PART 1

	Wolt							
Participant ID	ease of use and understanding of the service	lack of accessibility settings	FAQ issues	unclear information about price formation	filtering issues	not showing all the restaurants	language issues	Pop-ups issue
P1	"It was very straightforward, I would argue more straightforward than Uber Eats (modifying the cart). Wolt app felt kind of intuitive and I think that it is easier to navigate then Uber Eats, so I was pleasantly surprised (interview)."	"So i can adjust the language...and that's that, yeah , i think it is just the language."	"For Wolt adding more answers to FAQ, so you don't have to contact the support."					
	Foodora							
	"Overall, I think that both apps are pretty straightforward."	"Oh yeah, it has language and that's about it."	"I would go to help center, and it will not give me anything helpful, well, in theory I would be able to get help in the help center, but I guess not."	"I am not sure that the service fee is, does it go on the top of the order? I don't think other restaurant had a service fee of 7 percent."	"I thought I sorted them by fast delivery, but looking at it, it is not sorted at all. I think in theory I would be able to click delivery and then fast delivery, but apparently not really."			
P2		"Wolt had more accessibility. Foodora I couldn't find any accessibility options other than language (interview)."				"I have to scroll down past a bunch of junk. It looks like popular restaurants, ads, late nights, recommendations, more discounts before I get to a full list of all the restaurants."	"Oh, this one's customizable. I can pick a sauce. Wait, no, a side dish. I don't know. It's in Finnish, even though I've selected the language of the app to be in English."	"The one on Foodora took up my entire screen. It was in Finnish. I couldn't translate it. Yeah, it was. It was overall very invasive and very aggressive. (interview)"
	Wolt							
							"OK, just pick a pizza. Looks like customizable. It's also in Finnish, even though I have English selected."	
P3			"Yeah, I'm not able to find what I should do if my order is late. I don't think it mentions anywhere here."	"I don't understand the delivery fee was 0 euro and here I can see it's one euro 99."			"It's again in Finnish. I thought it was in even categories are in Finnish, but here it is in English. I'm not sure it's up in English or Finnish. It's it is all in Finnish...OK, great. Now it's in Chinese. So it's either in Chinese or in Finnish. I see."	
	Foodora							
		"Yeah, but it's it's actually so small and so like so hard to see since some light grey color...I don't really want to use this application at night where it will just burn my eyes because it's so bright. It only has the light mode"		"So when I order, I expect my items to be 7 euro and plus delivery, but there's also service fee. I'm very confused about service fee."		"What? Can I just apply filter without mentioning the name of restaurant? I don't understand. If I want to see the list of all of the restaurants and sort them by time, why can't I do that? Maybe I'm doing something wrong. Can I have a list of like all of the restaurants?"	"OK, it's for some reason it's again in Finnish. Yeah, even though we choose English, it is still in Finnish."	"Advertisement like everywhere and like popping, advertisement pop-ups and moving, literally moving. I don't like that moving images. They distract me. I want to order food." (interview)
P4			"There was no information in FAQ about delayed orders."		"I don't think it is sorting them correctly." (applying the filter)			"What is that? Do I have to press it? How do I close it?" (home page pop-up appeared)
	Wolt							
				"I think I would trust both of the applications, but I also didn't realize what they don't show you the service fees included because I have to see the service fee first then I have to place it to to my items fee. (interview)"			"Why is it in Finnish for some reason?"	
P5				"Oh, why is it more expensive? Oh, OK. It's service fee, I think."			"The only option I can understand is mozarella. What does it mean?"(going through item's options)	
	Foodora							
		"I don't like that Foodora icons are so small because in Wolt icons are quite big for rectangles, like big rectangles, so it is easy to press them." (talking about categories icons/images)						

APPENDIX E: THEMATIC CODING PART 2

Participant ID::	unclear feedback	unclear information on payment methods	delivery details	freezing	unclear availability of restaurants	no option to add payment methods
P1		"Weird thing is that i think it is automatically connected to my apple pay, maybe not, maybe it just show an option of apple pay but is not going to charge it. but if it is connected, it is kind of weird, because it didnt ask me for confirmation at any point."**(WOLT)**	"In Wolt there was an option how I want my order delivered, like to the door. In foodora I didn't see, and I think it is very important, so we are on the same page about how order is being delivered." **(FOODORA)**	"Ok, it is not moving, I think it froze. Yeah it is froze. Maybe it is because I am using my iPad and not a phone, but it is completely frozen, let me try to open it again... it froze again, it froze, so I have to reload the app once again... it froze again, it is not working." **(FOODORA)**	"Maybe it has like a button open now? No, I dont see any buttons. Here maybe? No." (paticipant was struggling to see if restaurants are open) **(WOLT)**	
P2	"Oh, it looks like there's more that they had hidden. And then add to cart. (not clear required items)."**(FOODORA)**					"OK, also here on the accounts page, it looks like there's an option for payments. It's empty here. Wait, why is there no option to add? What? There's no option to add payments. What? That seems. That seems wrong. Hold on. ... It should be on this page where your payment methods are and there's nothing here." **(FOODORA)**
P3				"Why was it lagging? Why? Why is it doing that? I'm just like. Doing the amount bigger and. It has this like loading loading screens after I try to add more items. " **(FOODORA)**		"It just shows me that it's empty here in payment method, so I cannot even add anything. Why, I want to add a new payment method and I I can't do it from the payment section." **(FOODORA)**
P5					"Is it open? I dont think so. If it is saying that it will deliver soon then it is open?" **(FOODORA)**	