



CHALMERS
UNIVERSITY OF TECHNOLOGY



UNIVERSITY OF GOTHENBURG

Korpi

Developing a User-Friendly and Useful Frontend for a Corpus Search Engine

Bachelor's thesis in Computer science and engineering

Ludwig Alexandersson

Erik Dreifeldt

Pomesh Kumar

Lisa Wadenbrant

Ida Åberg

BACHELOR'S THESIS 2025

Korpi
**Developing a User-Friendly and Useful Frontend
for a Corpus Search Engine**

Ludwig Alexandersson

Erik Dreifeldt

Pomesh Kumar

Lisa Wadenbrant

Ida Åberg



UNIVERSITY OF
GOTHENBURG



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2025

Korpi

Developing a User-Friendly and Useful Frontend for a Corpus Search Engine

Ludwig Alexandersson Erik Dreifeldt Pomesk Kumar Lisa Wadenbrant Ida Åberg

© Ludwig Alexandersson, Erik Dreifeldt, Pomesk Kumar, Lisa Wadenbrant, Ida Åberg
2025.

Supervisor: Peter Ljunglöf, Department of Computer Science and Engineering,
Språkbanken Text

Examiners: Patrik Jansson, Department of Computer Science and Engineering

Graded by teacher: Magnys Myreen, Department of Computer Science and Engineering

Bachelor's Thesis 2025

Department of Computer Science and Engineering

Chalmers University of Technology and University of Gothenburg

SE-412 96 Gothenburg

Telephone +46 31 772 1000

Typeset in L^AT_EX

Gothenburg, Sweden 2025

Korpi

Developing a User-Friendly and Useful Frontend for a Corpus Search Engine

Ludwig Alexandersson, Erik Dreifeldt, Pomesk Kumar, Lisa Wadenbrant, Ida Åberg

Department of Computer Science and Engineering

Chalmers University of Technology and University of Gothenburg

Abstract

Corpora are vast collections of texts that allow linguists to study how languages are used and how they change. The Language Bank of Sweden (Språkbanken) houses and provides access to corpora via *Korp*, a research platform for analysing texts. This thesis discusses the development and evaluation of a more user-friendly frontend for Korp named *Korpi*. It aims to address three issues with Korp: its slow speed when searching across large corpora, an interface which is not suited for mobile devices, and complex functionality that may not be needed for the general public.

Korpi was developed in an iterative manner with features being implemented and evaluated throughout the project. Key elements of the user experience were identified and strategies to ensure a responsive design were utilized. A mobile-first approach was adopted ensuring that Korpi was suited to mobile devices.

Korpi was evaluated by two different focus groups divided by age. Various aspects of Korpi, as well as Korp, were evaluated. These related to the user-friendliness and usefulness of the two applications. Questionnaires were used to quantify the results on a scale from 1 to 5 for each of the evaluation aspects.

Findings from the focus groups show that Korpi is a user-friendly application that can be considered useful. The design and streamlined implementation of search engine features were appreciated. Korpi was seen as a better option for first-time users wanting an introduction to linguistics as opposed to Korp.

Keywords: corpus, linguistics, user-friendly design, search engines, web development

Sammandrag

Korpusar är stora textsamlingar som används inom språkforskning för att studera språkbruk och språklig förändring. Språkbanken vid Göteborgs universitet tillhandahåller via plattformen Korp tillgång till korpusar. Trots sin funktionella bredd är Korp främst utformad för forskare, vilket gör den svåränvänd för den breda allmänheten.

Denna uppsats behandlar utvecklingen av ett mer användarvänligt gränssnitt till Korp, kallat Korpi. Projektet syftar till att åtgärda tre huvudsakliga problemområden i Korp: låg prestanda vid sökning i stora korpusar, bristande mobilanpassning samt komplex funktionalitet. Korpi utvecklades iterativt med fokus på användarupplevelse och responsiv design, enligt ett "mobile-first"-angreppssätt. En komponentbaserad arkitektur möjliggjorde flexibel utveckling och testning.

Korpi har stöd för både det ursprungliga Korp-API:et och en ny, snabbare backend kallad Korpsearch. Utvärderingen genomfördes med två fokusgrupper, indelade efter ålder. Både Korpi och Korp bedömdes utifrån aspekter som relaterade till användarvänlighet och användbarhet. Resultaten kvantifierades med hjälp av enkäter med en femgradig skala.

Resultat från fokusgrupperna visade att Korpi är en användarvänlig applikation som också kan anses användbar. Designen och den optimerade implementeringen av sökmotorn uppskattades. Korpi sågs som ett bättre alternativ för förstagångs användare som ville ha en introduktion till lingvistik, i motsats till Korp.

Keywords: korpus, lingvistik, användarvänlig design, sökmotor, webbutveckling

Acknowledgements

We would like to thank our supervisor Peter Ljunglöf for his guidance and support, and for providing a backend Korpsearch. His input and advice have helped shape Korpi. We would also like to thank Språkbanken for allow us to attend a helpful meeting with their engineers and for helping us when confusion arose with Korp. We thank the project group DATX11-VT24-05 for giving inspiration, helping us with the report structure. Finally, we give our utmost appreciation and thanks to the seven participants in the focus groups for contributing to the evaluation and development of Korpi by providing a fresh and fruitful perspective.

Ludwig Alexandersson, Erik Dreifeldt, Pomesk Kumar, Lisa Wadenbrant, Ida Åberg,
Gothenburg, June 2025

Contents

List of Figures	xv
List of Tables	xvii
1 Introduction	1
1.1 Background	1
1.2 Purpose	2
1.3 Problem Areas	2
1.4 Method	2
1.5 Scope and Limitation	3
2 Theoretical Framework	4
2.1 Corpora and Corpus Search Engines	4
2.2 User Experience	5
2.3 UX elements in Website Development	5
2.4 Accessibility	6
2.5 Design Patterns	6
2.5.1 Centre Stage	6
2.5.2 Escape Hatches	7
2.5.3 Accordion	7
2.6 Mobile-First	7
2.7 Technical Resources	8
2.7.1 Korp API and Korpsearch	8
2.7.2 React	9
2.7.3 Figma	9
2.7.4 Bootstrap	9
2.7.5 GitLab	9
2.7.6 Netlify	9
3 Methodology	10
3.1 Development Phases in a UX Project	10
3.1.1 Concept Phase	10
3.1.2 Development Phase	10
3.1.3 Detailing Phase	10
3.2 User Experience Evaluation	11
3.2.1 Qualitative and Quantitative Research	11
3.2.2 Focus Groups	11

3.2.3	Evaluation Questions	12
3.3	Component-Based Architecture	13
4	Development Process and Implementation	14
4.1	Use of Focus Groups	14
4.2	Evaluation Elements	14
4.3	Concept Phase	15
4.3.1	Functionality in Concept Phase	15
4.3.2	Design in Concept Phase	17
4.3.3	Evaluation in Concept Phase: First Focus Group Meeting	18
4.4	Development Phase	21
4.4.1	Design in Development Phase	21
4.4.2	Implemented Functionality in Development Phase	22
4.4.3	Evaluation in Development Phase: Second Focus Group Meeting	22
4.4.4	Functionality Added After the Second Meeting	27
4.5	Detailing Phase	28
4.5.1	Design in Detailing Phase	28
4.5.2	Implemented Functionality in Detailing Phase	29
4.5.3	Evaluation in Detailing Phase: Third Focus Group Meeting	29
4.6	Coding the Web Application	30
4.6.1	Frontend	30
4.6.2	Backend	31
5	Results	32
5.1	Final Design of Korpi	32
5.1.1	Landing Page	32
5.1.2	Choose Corpus	32
5.1.3	Extended Search	33
5.1.4	Settings Modal	33
5.1.5	Modify Search Modal	34
5.1.6	History Panel	35
5.1.7	Tour	35
5.1.8	Results Page	36
5.1.9	Dictionary	37
5.1.10	Statistics Page	38
5.2	Final Functionality of Korpi	38
5.3	Final Evaluation of Korpi	39
6	Discussion	43
6.1	Fulfilment of Purpose	43
6.1.1	Design for User-Friendliness	43
6.1.2	Functionality for Usefulness	44
6.2	Possible Sources of Error	45
6.3	Ethical Considerations	46
6.3.1	Regarding Information Access and Copyright in Corpora	46
6.3.2	Ethics in Web Development and Security	46
6.3.3	Social Aspects	47

6.3.4	Focus Groups	47
6.4	Future Improvements	47
7	Conclusion	49
	References	51
A	Focus Group Tasks, Questions, and Responses	I
A.1	Tasks and Questions for Focus Group Meeting 1	I
A.1.1	Tasks for testing Korp	I
A.1.2	Questions related to Korp - younger focus group	I
A.1.3	Questions related to Korp - older focus group	I
A.1.4	Questions related to initial design prototype of Korpi - younger focus group	I
A.1.5	Questions related to initial design prototype of Korpi - older focus group	II
A.1.6	Questions related to navigation and information presentation for Korp and Korpi	II
A.2	Responses from Focus Group Meeting 1	II
A.3	Tasks and Questions for Focus Group Meeting 2	III
A.3.1	Tasks for testing the first version of Korpi	III
A.3.2	Questions related to the first version of Korpi	III
A.3.3	Questions related to the new design prototype of Korpi - only asked to younger focus group	IV
A.3.4	Figma Design Proposal For Extended Search	IV
A.3.5	Questions related to the statistics demo - only asked to younger focus group	IV
A.4	Responses From Focus Group Meeting 2	V
A.5	Tasks and Questions for Focus Group Meeting 3	V
A.5.1	Tasks for testing final Korpi	V
A.5.2	Questions related to final Korpi	V
B	Focus Group Survey	VI
C	Focus Group Material for Discussion Regarding Functionality	X
D	Design	XII
D.1	Concept Design	XII
D.2	Detailing Phase Design	XIV
D.3	Future Improvements	XIV

List of Figures

2.1	Språkbanken Korp on mobile (left) and desktop (right) screens. . . .	5
4.1	First prototype of Korpi's landing page	17
4.2	First prototype of Korpi's results page	17
4.3	New Figma prototype of Korpi in development phase: landing page .	21
4.4	New Figma prototype of Korpi in development phase: results page. .	21
4.5	Figma design of the extended search in Korpi	28
4.6	Korpi's landing page after second focus group meeting	28
5.1	Landing Page	32
5.2	Choose Corpus Modal	33
5.3	Extended Search Panel	33
5.4	Settings Modal	34
5.5	Result Panel in Dark-mode	34
5.6	Modify Search Modal	35
5.7	History Panel	35
5.8	Landing Page Tour	36
5.9	Result Page	36
5.10	Result Card Information	37
5.11	Results Panel with all tabs closed.	37
5.12	Definition Panel	38
5.13	Statistics Panel	38
5.14	Survey responses for final version of Korpi	40
A.1	Figma design proposal for extended search in Korpi	IV
B.1	Focus group survey page 1	VII
B.2	Focus group survey page 2	VIII
B.3	Focus group survey page 3	IX
C.1	Functionality that was discussed with the older focus group.	X
C.2	Functionality that was discussed with the younger focus group.	XI
D.1	First prototype of Korpi's side menu for mobile and desktop	XII
D.2	First prototype of Korpi's settings page for mobile and desktop	XII
D.3	First prototype of Korpi's calendar result page	XIII
D.4	First prototype of Korpi's choose corpus page for mobile and desktop	XIII
D.5	First design of Korpi's result page when search result is expanded . .	XIV

D.6 Extended Search when used in Result Page XIV

List of Tables

2.1	UX elements in Website Development	6
4.1	Translation between terms in section 2.3 and terms used with focus groups.	15
4.2	First focus group meeting: Younger focus group	19
4.3	First focus group meeting: Older focus group.	20
4.4	Second focus group meeting: Younger focus group evaluating Korpi.	23
4.5	Second focus group meeting: Older focus group evaluating Korpi.	25
4.6	Second focus group meeting: Younger focus group	27
5.1	Main implemented and potential future functionality of Korpi.	39
5.2	Summary of written answers from Younger and Older groups to all feedback questions. Answers are translated from Swedish to English.	41
A.1	First focus group meeting: Younger focus group evaluating Korp.	II
A.2	First focus group meeting: Older focus group evaluating Korp.	III
A.3	Second focus group meeting: Younger focus group	V

1

Introduction

Linguistics is the science of language and a linguist is someone who in a structured way studies and describes a language. One methodology within linguistics is that of corpus linguistics, which is used for studying the use of a language [1]. In corpus linguistics, the use of a language is studied in bodies of text called corpora, from the Latin word *corpus* meaning “body” [1]. Today, a corpora is almost always an electronic corpora and the definition is “a collection of texts which is stored on some kind of digital medium and used by linguists to retrieve linguistic items for research or by lexicographers for dictionary-making” [1, p. 3]. The necessary size of a corpora depends on the purpose of a linguistic investigation [1]. For studying frequent aspects of a language a million words is often enough, for some studies 500 million words or more is suitable.

1.1 Background

*The Language Bank of Sweden (Språkbanken)*¹ is a Swedish organization where partner organizations² are collaborating to develop a research e-infrastructure based on language data [2]. Språkbanken has developed several digital research platforms and made them available for free - in order to support research where language data is important [3]. One of these research platforms is *Korp*³. In Korp, a user can search for words in corpora that are annotated with different kinds of language attributes such as word class (see figure 2.1). The search results consist of the sentences in the corpus that contain the queried word. It is also possible to find statistics in Korp, e.g. the number of occurrences of a word and its distribution over time.

Being a research platform, Korp is primarily targeted towards professionals in the field. Although it offers extensive functionality, Korp is slow when performing searches across large corpora, affecting the overall user experience. The user interface is difficult to understand and use without thorough explanation of the various functionalities. Furthermore, it is not adapted to smaller screens such as mobile phones, hampering its availability to the general public.

However, according to the supervisor of this project, the core functionality of Korp

¹<https://sprakbanken.se>

²University of Gothenburg, The Institute for Language and Folklore (Isolf), National Library of Sweden, KTH Royal Institute of Technology, Linköping University, Lund University, The Swedish National Archives, Stockholm University, Umeå university, and Uppsala University.

³<https://sprakbanken.gu.se/korp/>

could potentially be interesting for the public. Indeed, an assumption made throughout the project is that many people are interested in linguistics. Yet the public may have less patience with a non-responsive web application that is not user-friendly and overwhelming with features.

This thesis therefore addresses the development and evaluation of a new version of Korp, called *Korpi*, one that is targeted towards the general public.

1.2 Purpose

The purpose of the project is to develop a frontend for an existing corpus search engine based on Korp; from being functionally-oriented towards professionals, to being user-friendly for the general public while remaining useful. The application should be easy to use and understand on both mobile and desktop devices.

1.3 Problem Areas

In order to fulfil the purpose, these are problem areas that need to be addressed:

- **Problem 1 - Design for user-friendliness**

The first problem area pertains to making the web application user-friendly. This includes adapting the web application for both mobile and desktop devices, as well as presenting large amounts of lexical information in a digestible manner, without overwhelming the user.

- **Problem 2 - Functionality for usefulness**

The second problem area is how to ensure that the web application is useful for users. A difficulty in this project is that the core functionalities of the app are already provided, a search engine for large corpora. Users might not know what they would like to use these functionalities for, as they are not necessarily applicable to the general public.

1.4 Method

Both the overall set-up of the project and the more detailed function and design choices of the final product have been based on a theoretical framework put together in a process of literature review.

The theoretical framework is described in chapter 2, a description of the methods used in chapter 3, and the more detailed function and design choices along with the implementation of the project in chapter 4.

Briefly, the set-up of the project can be described as a process of parallel coding and designing of different stages of prototypes that have been tested by and discussed with two different focus groups. Evaluation of the product was also carried out by the project group.

1.5 Scope and Limitation

While Korp is designed for linguists, Korpi will be developed with the general public in mind. The application aims to be accessible, regardless of prior knowledge in linguistics. Therefore, it is essential that the web interface is intuitive and easy to use.

This project will build upon existing code for processing text searches in corpora. Although modifying the original code is not part of the project scope, any requests made to change the backend will be forwarded to the supervisor. Thus, Korpi will serve solely as a frontend to the pre-existing backend. As such, the project assumes no responsibility for the corpora themselves, the accuracy or ordering of search results, or the types of corpora employed. Korpi merely presents the results returned from the user's query, and it is the user's responsibility to interpret this data. The ethical aspects of the project are further discussed in section 6.3.

2

Theoretical Framework

This chapter gives the theoretical background regarding the project. Describing how linguists use corpora, and different theories on how to achieve user-friendliness and usefulness. The chapter also introduces the technical resources used.

2.1 Corpora and Corpus Search Engines

As introduced in Chapter 1, corpora are collections of lexical data, often in the form of written texts. For example, the entire works of Shakespeare form a corpus whereas all articles written by a publication in a year, may form another corpus. Corpora can therefore be comprised of hundreds of texts and millions of words. Using corpora, linguists can study changes in language over time [1]. This could encompass the utilization of various words or phrases over the years and how their meanings have changed. This can include “differences between genres, geographical varieties, spoken and written language, text from different time periods and so on” [1, p. 8]

In the digital age, corpora are composed of electronic documents which can be searched through the use of a corpus search engine [1]. The differentiating factor between texts in corpora and texts in a book are annotations of individual words. Words in a corpus are annotated with certain attributes: lemma, word class, definition and so on. These annotations allow for contextual information of a word to be taken into account, for example, the word *run* can either be a verb or a noun, depending on the context.

A corpus search engine enables linguists to explore various corpora. Språkbanken provides Korp as noted earlier, targeted for researchers. When a query is made to a corpus search engine the results returned contain the searched word surrounded the context the word has been found in. This is illustrated in figure 2.1 where the word *katt* (eng. cat) has been searched. The additional context gives researchers the ability to understand how a particular word is used, what types of words frequently precede and succeed it, and so on.

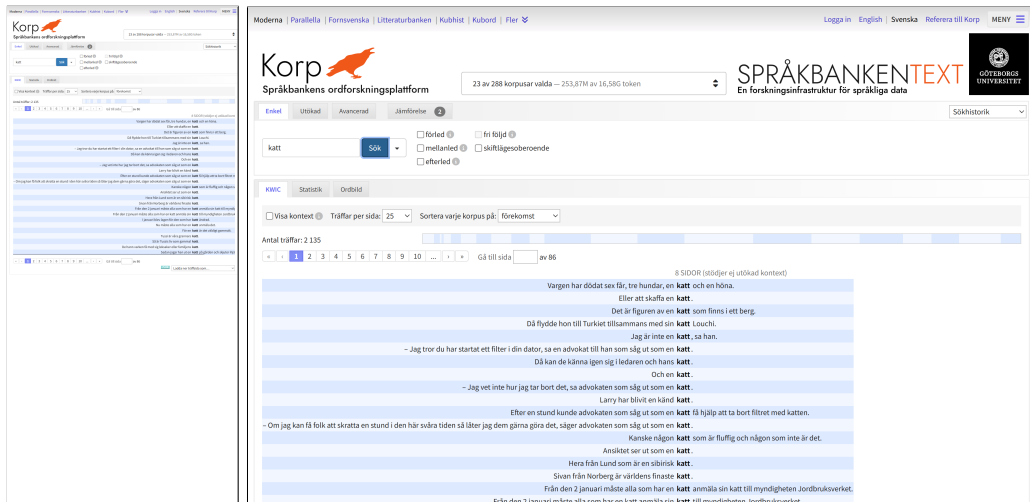


Figure 2.1: Språkbanken Korp on mobile (left) and desktop (right) screens. The website is not adapted to smaller screens. Results show the query *katt* centred in context lines. Each of the lines is from the same corpus (8 Sidor) and show words used before and after the query.

2.2 User Experience

User Experience (UX) refers to interaction design that goes beyond just the user interface. The term was introduced by Don Norman in the 1990s during his time at Apple to describe a broader view of how users experience products [4]. Since then, UX has become an important part of both digital and physical product development.

Earlier approaches to design mainly focused on functionality, system performance, and visual appeal, aiming to make sure tasks could be completed efficiently. UX instead centres the entire user experience including emotional reactions, perceptions, and context of use [5]. While designers may have a sense of what is aesthetically pleasing, UX emphasizes that the user’s perspective must come first. UX is therefore a human-centred and iterative process depending on real user feedback to improve final designs.

2.3 UX elements in Website Development

Effective website development requires addressing multiple user experience elements organized within a five-plane conceptual framework of abstraction levels [6]. These levels are interdependent, with lower-level decisions depending on those at the higher-level. The *strategy* level establishes website objectives, encompassing both product objectives and user needs, which subsequently inform content and functionality requirements at the *scope* level. The *structure* and *skeleton* levels on the other hand, address website organization through both functional and informational dimensions, with functionality enabling task completion and information facilitating user comprehension. Navigation, defined as users’ ability to discover and traverse the digital environment [7, p. 563], represents a critical component spanning these intermediate levels. Finally, the *surface* level encompasses users’ sensory experience,

primarily focusing on visual design elements including typography and colour.

Table 2.1: UX elements in Website Development

Abstraction levels	Elements
Strategy (abstract level)	- Product objectives - User needs
Scope	- Functional specifications - Content requirements
Structure	- Interaction design - Information architecture
Skeleton	- Information design - Interface design - Navigation design
Surface (concrete level)	- Sensory design

2.4 Accessibility

An important consideration in designing interactive systems is how to ensure accessibility for people with disabilities. There are two approaches for how to deal with this issue, namely inclusive design and universal design [7]. Inclusive design is about designing specifically for certain disabilities. It can relate to specific technologies for people with disabilities, such as voice input and screen readers. Universal design, on the other hand, builds on the philosophy that the design should be usable by a wide range of people. The design should be flexible so that it suits different abilities and preferences. It should be simple, intuitive with information being easily perceptible, no matter the users' sensory abilities.

2.5 Design Patterns

Design patterns are reusable solutions to reoccurring problems to improve the user experience. The usage of design patterns are standard practice in web-development. There are numerous different patterns, but the patterns described in this section are the most vital for the project.

2.5.1 Centre Stage

The Centre Stage layout prioritizes the primary content by placing it at the centre of the interface [8]. This could involve editing a document, filling out a form, or browsing through search results. The design ensures that the most important information receives the appropriate space and focus. By drawing attention to the centre, the layout naturally guides the user's eye to where they should begin. This establishes a strong focal point, reducing the cognitive effort and minimizing the need for excessive scanning or guesswork. Such layouts are especially effective for content-heavy pages, where clarity and direction are essential for a smooth and focused user experience.

Supporting content, such as navigation, filters, or other secondary information, are placed in smaller panels around the main area. This creates a clear hierarchy between primary and supporting elements, helping users grasp the structure of the interface more quickly.

2.5.2 Escape Hatches

Escape hatches are clear and simple navigation elements, such as buttons or links [8]. They help users return from a screen to a previously visited or familiar page. Escape hatches are useful in interfaces with limited navigation options, like modals or error pages. If users find themselves on an unfamiliar or confusing page, whether they followed a link or ended up on an error screen, an escape hatch provides a direct way out without causing frustration. Escape hatches reduces the risk of users abandoning the site by making it easy to leave pages or dead-end screens. Users feel more safe to explore the site if they know that they can always come back to a more familiar page.

2.5.3 Accordion

An accordion divides similar or related content into vertical sections that can be expanded or collapsed individually [8]. This can be useful when a page contains a large amount of information that needs to fit within a limited space. Instead of showing all the content at once and risking overwhelming the user, the accordion allows them to reveal only the information they choose to see. Other solutions, such as tabs, might not allow multiple sections to be open at the same time. Therefore, accordions offer more flexibility in how users interact with and view the content.

A well-known example of accordion use can be found on Reddit, particularly in the comment sections. Replies to top-level comments are initially hidden under a collapsed thread. Users can expand these threads individually to view deeper discussions without being overwhelmed by all replies at once.

The items contained in an accordion are often similar in width but can vary in height. There are, however, some limitations to using accordions. For example, if a user opens multiple large sections at once, the lower items may be pushed out of view. Still, accordions are a well-known and established design pattern, especially in web development. By allowing users to decide which sections remain visible, accordions help them customize their own environment and manage complexity more effectively.

2.6 Mobile-First

Mobile-first design gives priority to mobile devices as opposed to the more classic approach in which desktops are the primary focus. This ensures accessibility, engagement, and usability for users on mobile devices.

Some illustrative principles for mobile-first design are:

- **Responsive and optimized content**

Mobile-first design focuses on responsive content that smoothly can adjust to different sized screens. Making sure the design is consistent and user-friendly, including adapting layout, fonts and visuals for readability and usability on smaller screens [9].

- **Adapting content for smaller screens and touch-based navigation**

To fit smaller screens, content should be concise, engaging visually and divided into manageable parts. Developers should make sure that all interactive elements are touch-friendly and easily accessible for mobile users [9].

- **Intuitive user interface and navigation**

A mobile interface that is intuitive should be easy to use, visually appealing, and consistent over all parts, helping users navigate without creating confusion or frustration. This enhances the overall using experience [9].

- **Implementing clear and consistent navigation structures**

Clear and consistent navigation is important to intuitive mobile usage. Menus and icons need to be organized and familiar, having structured pathways that helps users progress through content efficiently [9].

2.7 Technical Resources

This section gives a description of technical resources that will be used for developing the web application.

2.7.1 Korp API and Korpsearch

There are two APIs (Application Programming Interfaces) that are used in the development of Korpi, namely *Korp API* and *Korpsearch*.

As noted earlier, Korp is the research platform used by Språkbanken for text analysis in large corpora [10]. It provides both a frontend and backend, with the latter referred to as Korp API [11]. Queries made by the user through the frontend are sent to central servers using various end-points. A complete list of commands and commands is outlined in the Korp API's documentation [11]. The responses can then be parsed to display results in Korpi.

Korpsearch provides an alternative API which speeds up response times for queries [12]. It is one of the prerequisites for a more responsive Korpi. It utilizes inverted indices to greatly speed up queries to corpora databases. An inverted index is a data structure that maps individual words to their locations in different documents. Instead of searching each document individually for a particular word, the query can look up the word in a table and find the documents it maps to, in constant time. The trade-off is that the inverted indices must be built beforehand, a potentially time-consuming process. However, for large databases where the content does not change frequently, such as corpora, this can be done server-side as needed. If a new document is added to a corpus, the indices can be rebuilt without the client being

aware.

2.7.2 React

React is a mature and tested Javascript framework released more than a decade ago by Facebook and is amongst the most popular web technologies [13], [14]. It allows for the creation of user interfaces in a modular and organized manner. A webpage is composed of many different components, for example, it may contain a navigation bar at the top, menu on the side, gallery, main content and so forth. With React these components can be built separately, tested individually and composed together at a later stage to create the final page.

2.7.3 Figma

Figma is a web-based prototyping tool that allows for the creation designs and interfaces [15]. No coding is required and focus is purely on the look-and-feel as well as general layout of a webpage. The designs can be used by the coding team whilst developing the final application.

2.7.4 Bootstrap

Bootstrap is a popular framework used for styling frontends [16]. It offers multiple pre-designed components, supporting mobile-first, responsive interfaces. It enables developers to create consistent layouts and apply styling to various elements, such as navigation bars, buttons, and modals, without needing to write custom styles. Bootstrap can be used alongside libraries like React, where Bootstrap handles the styling and layout of the page, while React manages the logic and interactive elements.

2.7.5 GitLab

GitLab is a web-based platform to manage code and enable collaborative software projects used by developers [17]. It is based on Git, a version control system that can track changes in code over time [18]. With Gitlab one can store the codebase in a cloud-based repository so that multiple group members can work together, allowing for parallel development.

2.7.6 Netlify

Netlify is a popular cloud-based service that allows for the hosting of web content, ensuring a layer of security and reliability [19]. Code written locally can be built and packaged by Netlify to produce a website that can be accessed worldwide and on multiple devices.

3

Methodology

This chapter describes the processes and methods used when developing a web application. It discusses how to divide a UX project based on user evaluation. Furthermore, iterative and component-based methods for developing a web application are discussed.

3.1 Development Phases in a UX Project

The process of creating a UX-centred product can be simplified by dividing the project into three phases: the *concept*, *development*, and the *detailing* phase [5].

3.1.1 Concept Phase

The first step in a UX-driven project is understanding the client's demands as well as the user needs [5]. This is done through research, involving users, and defining intentions and objectives. Concept sketches are done and outcomes are evaluated against initial goals. It is essential to distinguish between what users want in comparison to the client requests. A balance between these two perspectives must be achieved [5]. It is the responsibility of the design team to look beyond the client's and users immediate wishes and refine the concept accordingly. Time, scope and other resources must be addressed in this phase.

3.1.2 Development Phase

In the development phase, the initial concepts are transformed into a functional prototype. Product requirements are defined for data, design, interfaces and functionalities. After this phase the product should be close to functional [5]. The scope and limitations must be taken into account and revised as the changes to the product emerge. The design and coding teams collaborate to build and refine the product iteratively. Feedback from user tests as well as the client is the primary guidance.

3.1.3 Detailing Phase

Lastly, design and functionalities are integrated into a complete product and the detailing phase is concerned with the refinement of the product. Finalizing the layout based on earlier feedback and testing is the priority, since the core functionalities should be finished at this stage. The product in this phase is also reviewed, ensuring that it is both functional and visually appealing [5].

3.2 User Experience Evaluation

UX evaluation refers to a collection of methods used to assess how people experience a system whether it is a product, service, or interface. UX evaluation involves choosing a relevant set of constructs, identifying appropriate methods, and interpreting user interaction within the context of use [20]. In the case of web applications, it can be how users navigate, understand, and use the products' features.

UX evaluation is shaped by a few key methodological choices. One is whether the evaluation is formative, conducted during development to guide design decisions, or summative, applied to a finished product [20]. Another is the type of data collected: quantitative (e.g. task completion time, error rate) or qualitative (e.g. user feedback, observations). The research needs to specify the scale of analysis focusing on momentary impressions, episodic experiences (e.g. using a feature), or overall evaluations after extended use.

When evaluating a website, certain experience constructs are particularly relevant. Usability concerns whether users can complete tasks efficiently and without confusion. Utility refers to whether the functions offered are actually useful to the user. Aesthetics considers the visual design and whether the interface is perceived as attractive or coherent. Constructs like stimulation (whether the interface engages the user), or value (whether the system matters to the user) can also be assessed. As can be noted, these evaluation aspects are closely related to the abstraction levels related to website development described in chapter 2.

UX evaluation methods are typically divided into explicit, implicit, and creative categories. In a web application project, explicit methods are most applicable, such as usability testing, interviews, and surveys – as they directly capture the user's conscious reflections on their interaction. Implicit methods, such as observation or reaction – time measurement, may supplement this by capturing user behaviour that is not articulated in speech. Creative methods (e.g. paper prototyping or co-design) are often used earlier in design but can still support evaluation if user involvement is ongoing.

3.2.1 Qualitative and Quantitative Research

Qualitative research is used to study the complexity of human experience, focusing on words, actions, and interpretations rather than numerical data – which instead is the sole focus in quantitative research. While quantitative data is often treated as more conclusive – particularly in technical disciplines – qualitative methods provide the explanatory depth needed to interpret user behaviour [21]. Unlike quantitative methods which abstract behaviour into metrics, qualitative approaches aim to capture meaning, motivation, and context.

3.2.2 Focus Groups

Focus groups are a qualitative method that involves a small, clearly defined group of participants (often 7–10). The representation of a group depends on the end goals of

a project but should be either homogenous or heterogenous, but not randomized [22]. This group is engaged in moderated discussion about a specific product, concept, or system [23]. Focus groups are characterised by their structured yet interactive nature. While widely adopted today in fields such as marketing and healthcare, the method was first formalised in the 1940s as a response to the limitations of interviewer-dominated techniques [22]. In UX evaluation, focus groups can be used for formative evaluation. Because of their qualitative nature, focus groups allow researchers to explore not just what users think directly, but the steps leading to why they think it. This yields a lot of data for interpretation of overall user experience.

A focus group has a facilitator which guides the group through a set of open-ended questions, allowing the conversation to flow naturally while keeping it centred on the study's objectives [23]. Participants are encouraged to react to each other's input, leading to a collaborative exploration of views and experiences. The group setting often facilitates *piggybacking* where participants build on one another's contributions, helping to surface shared needs, expectations, and reactions [22].

3.2.3 Evaluation Questions

The following guidelines serve as support through the development of user evaluation questions. They are based on various UX concepts that discuss the design of evaluation meetings and how the facilitators of said meetings should structure questions [5].

Firstly, questions should be about past behaviour instead of future intentions. Asking about the future is ineffective, as it is difficult for people to predict their own behaviour [5]. If they use other products with similarities in a certain way, that might give a clue as to how the product in development will be utilised.

Secondly, using observation as a method of evaluation is important for achieving a broad overview of user experience. Below are some key points to assess when observing:

- **Sequence and order:** How do users perform tasks, and in what order? This can provide insight into the usability of the system.
- **Valuable tools:** What features or aids do they use? This can highlight which parts of the interface contribute to perceived utility.
- **Comments and body language:** What do they say? How do they react? When noting this, it is important to not be disruptive of the session by engaging or helping users when they become frustrated or ask questions [21]. To truly observe one must not intervene.

Lastly, task-based evaluation involves assigning users specific tasks to complete, rather than allowing them to explore the product freely. This approach helps evaluate usability by addressing key questions: Can users successfully complete the tasks? If not, what obstacles are they facing [5]?

3.3 Component-Based Architecture

A component-based architecture simplifies modularity, reusability, and scalability in a project. This methodology involves dividing the interface into different components [24]. These components handle a specific functionality or visual element. By defining functionality or logic in interfaces, singular components can be developed and modified independently without affecting other parts of the system. This type of architecture enables the structure of a web application to be consistent and reusable.

4

Development Process and Implementation

This chapter describes how the implementation project was conducted as well as important decisions and partial results. The realisation of the project was divided into the three phases as discussed earlier in section 3.1, namely the concept phase, development phase and detailing phase.

4.1 Use of Focus Groups

Formative evaluation of the product was done using focus groups with users, since UX was central to the product. Seven participants were chosen to represent the general public, meaning they were not professionals in the linguistic field. The participants were selected based on their gender and age differences. Many factors can contribute to a group representing the general public, but age differences and having a 50-50 group of men and women were dominant factors in deciding participants. As a consequence, other personal factors were less diverse since people that voluntarily participated were mostly engineers. This meant that the work-field and education levels were somewhat similar for all participants.

The participants were then divided into two groups, one of four and one of three. One group represented the younger population and the other represented the older population. The younger group consisted of two men in their twenties and two women who were in their late forties. The older focus group consisted of two men and one woman who were all senior citizens. The reason for having both younger and older participants was to ensure accessibility for different age groups. The approach to accessibility in the project was *universal design*, see description in section 2.4. The focus groups met three times each during the course of the project, with identical meeting structures to each other.

The evaluation was done with a qualitative approach. The exception to this was a summative survey of the product, which was also qualitative. The questions and tasks the focus groups were developed according to section 3.2.3.

4.2 Evaluation Elements

In order to obtain a thorough evaluation of Korpi, the different abstraction levels as described in section 2.3, were addressed in the meetings with the focus groups.

However, to avoid making the evaluation overly complex, the structure and skeleton levels were combined into one level. When these levels were in focus, a separation was made between the concepts of functionality and information instead (see section 2.3). The reason for this was the importance of understanding both ease of movement through the web application (while performing tasks), and ease of understanding presented information. Another simplification, regarding communication, was using a different terminology than in chapter 2. A translation can be found in table 4.1 below. *Information presentation* in table 4.1 simply refers to how all kinds of information is presented, for example information could be presented using the accordion design pattern as described in section 2.5.3.

Table 4.1: Translation between terms in section 2.3 and terms used with focus groups.

Terms in section 2.3	Terms used with focus groups
User needs (part of Strategy)	- Usefulness
Scope	- Functionality (including different types of content when applicable)
Structure and skeleton	- Navigation - Information presentation
Sensory design (the only part of Surface)	- Design (in practice referring only to visual design)

4.3 Concept Phase

The first prototype was based on the functionality of Korp as well as on information about what linguists normally use corpora for, the latter described in chapter 2. Though the public will not have the same needs as a linguist, the intention was to start out from what is relevant in Korp.

4.3.1 Functionality in Concept Phase

The Minimum Viable Product (MVP), the most basic functionality necessary, was agreed upon by the project team together with the supervisor. It specified the essential functionality that a frontend for a corpus search engine would require.

- **Mobile-First**
The application should mainly be for mobile use, therefore the design should not be hindered by smaller screen sizes and mobile use.
- **Search Function**
The application should allow a simple search where the user can type in a word in a search bar and retrieve all results containing a specified word in the selected corpora.
- **Result Presentation**
The search results should be displayed in a clear and accessible format, prefer-

ably in a table structure with the searched word being in the centre of the sentences (i.e. concordances, see chapter 2). Each result provides contextual information, and the interface should allow for users to quickly interpret the output.

- **Corpus Selection**

The user should be able to select any number of available corpora. The corpora should be broken down into categories, such as newspapers, journals, fiction, and other types of text collections. The categories should be easy to understand and navigate through and it should be possible to select an entire category at once.

- **Individual Result Information**

Search results should be interactive. When they are clicked (or tapped on mobile), additional information should be shown in a compact card. This includes the source, date and author of the result.

- **Extended Search Options**

The application should allow for more fine-grained, extended searches where words can be searched by additional attributes. These attributes should include word class and the ability to search for the dictionary form of word. Users should also be able to search based on part-of-speech, for example by searching only for nouns, adjectives, verbs, or adverbs. In addition to this, it should also be possible to search for several words at the same time.

- **Menu**

A dedicated menu and navigation bar should provide access to background information about the platform, such as its relation to Korp and Språkbanken. A help button and a home button need to be present in the navigation bar.

4.3.2 Design in Concept Phase

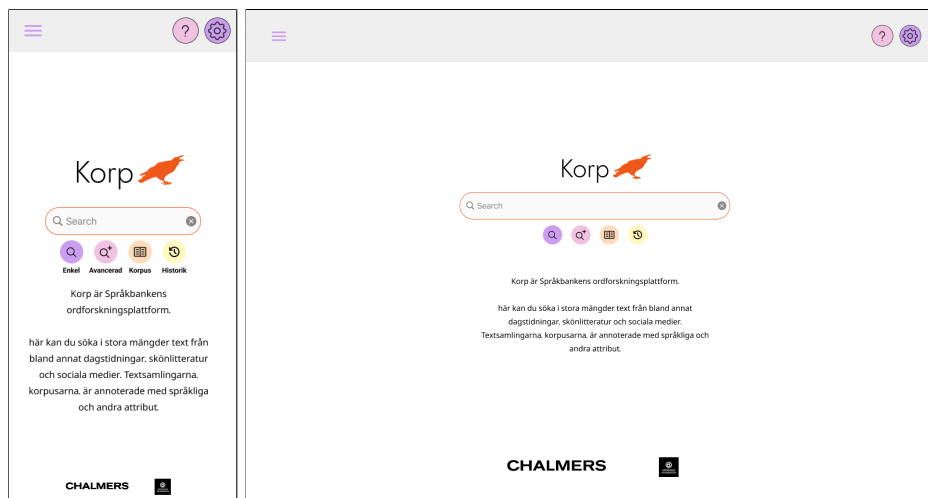


Figure 4.1: First prototype of Korpi's landing page for mobile and desktop, showing the centred search bar and associated icons, see text for more information.

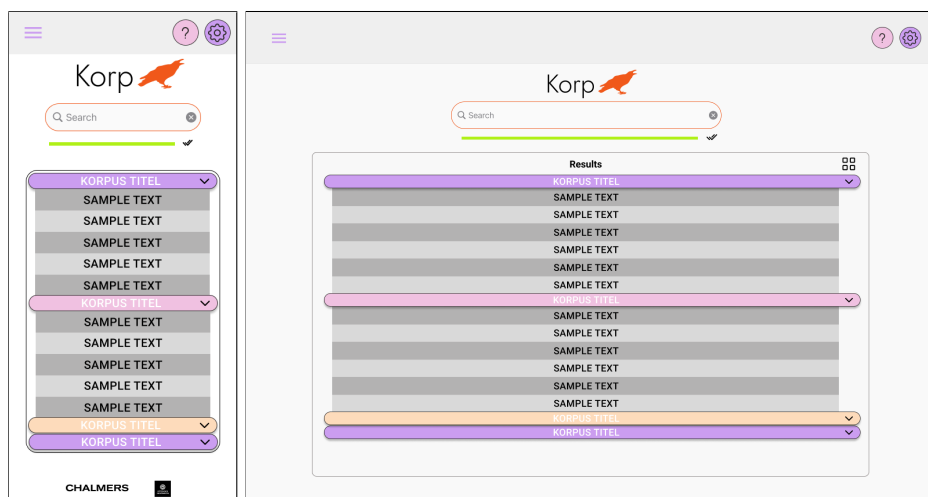


Figure 4.2: First prototype of Korpi's results page for mobile and desktop, showing the repositioning of the search bar, giving more space to show results.

When designing the web application, standard design patterns were applied alongside a mobile-first philosophy. The mobile-first approach meant that all designs were initially created with smaller screens in mind. The most commonly used patterns are listed in Section 2.5.

The most important design choices at this stage were to have a search bar as the main focus, adopting a centre-stage design, and to not make the user feel overwhelmed. Since the search functionality is the main feature of the application, it should be the first thing the user notices (figure 4.1). Inspiration was drawn from Google, as its simple and intuitive search-first design has proven highly effective [25].

For the results page, accordions were conceptualized to allow users to easily filter and explore different corpora (figure 4.2). Maintaining the centre-stage philosophy, the search bar was repositioned to the top to give more space to the results, which become the primary focus after a query has been made. The composition of search results were inspired by Korp's way of centring the searched word(s) by using concordance, as described earlier in section 2.1.

Additional figures of the design in the concept phase can be found in Appendix D.1.

4.3.3 Evaluation in Concept Phase: First Focus Group Meeting

The purpose of the first focus group meeting was to evaluate Korp and to evaluate the concept prototype of Korpi.

The meeting had the following agenda:

1. Testing Korp (and being observed).
2. Answering questions about Korp.
3. Testing the Figma concept prototype of Korpi (and being observed).
4. Answering questions about the Figma concept prototype of Korpi.
5. Answering questions about navigation and information presentation related to both Korp and the Figma concept prototype of Korpi.

The questions that were asked can be found in Appendix A.1.

The results from the first focus group meeting can be found in tables 4.2 and 4.3 below, and in tables A.1 and A.2 in Appendix A.2. A general observation from the first focus group meeting was that the participants in the older focus group tended to more passively observe what could be done in the prototype and in Korp before exploring, to a greater extent than the participants in the younger focus group.

Table 4.2: First focus group meeting: Younger focus group evaluating Figma concept prototype of Korpi.

First focus group meeting: Younger focus group evaluating Figma concept prototype of Korpi		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Design for user-friendliness	General user-friendliness	- Simple and easy to understand, but partly due to limited functionality.
	Navigation	- Go back button is requested. - One integrated search function, including the extended search options, is requested. This would simplify navigating the application.
	Information presentation	- Easier to understand what to do than in Korp, due to the central search function. - More space for the search results and search history is requested, e.g. by moving logos and search bar. - The alternative search result presentation in matrix form is difficult to understand. - Not clear that the progress bar is loading.
	Visual design	- Better and more modern visual design than Korp. - Positive that Korpi is more colourful than Korp. - The specific choice of colours needs more consideration. White text on light background is difficult to read. - A neutral font when presenting larger amounts of text is requested, e.g. arial.
Functionality for usefulness	Interesting functionality	- More sorting and/or filtering options are requested. - Statistics is requested. - Not interesting with a too simple web application with limited functionality.
	Usefulness	- No relevant feedback related to usefulness in this part of the evaluation.

Table 4.3: First focus group meeting: Older focus group evaluating Figma concept prototype of Korpi.

First focus group meeting: Older focus group evaluating Figma concept prototype of Korpi		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Design for user-friendliness	General user-friendliness	- All feedback related to user-friendliness is categorized under the evaluation aspects navigation, information presentation, and visual design.
	Navigation	- No relevant feedback related to navigation in this part of the evaluation. According to the focus group, there is still little functionality to navigate.
	Information presentation	- Positive that the user is not overwhelmed by information. - Good presentation of search results.
	Visual design	- Good visual layout in general.
Functionality for usefulness	Interesting functionality	- More sorting and/or filtering options are requested. - To be able to click a search result and get to the original source is requested. - A user guide, preferably with pictures, is requested. - Statistics is not considered interesting.
	Usefulness	- No relevant feedback related to usefulness in this part of the evaluation.

4.4 Development Phase

This phase covered the development of the first version of Korpi as a web application and its evaluation in the second focus group meeting.

4.4.1 Design in Development Phase

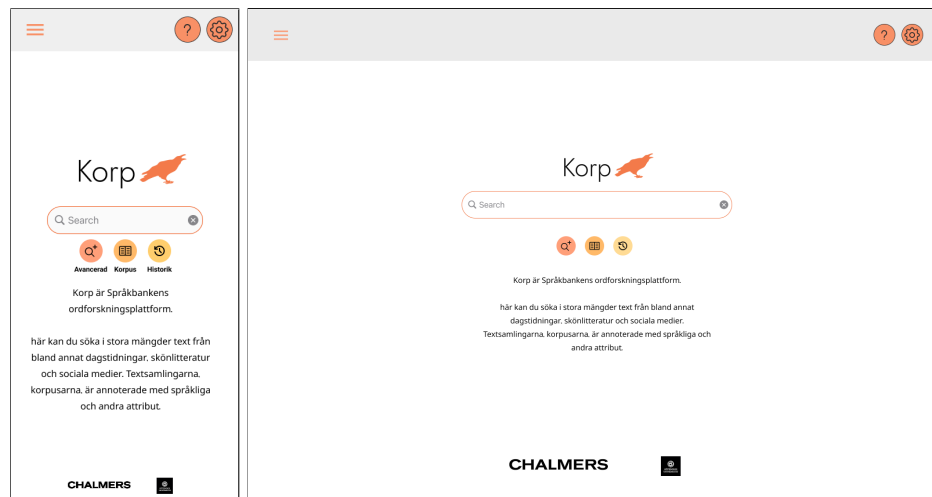


Figure 4.3: New Figma prototype of Korpi in development phase: landing page. The colour scheme has been updated to be more coherent by using similar colours.

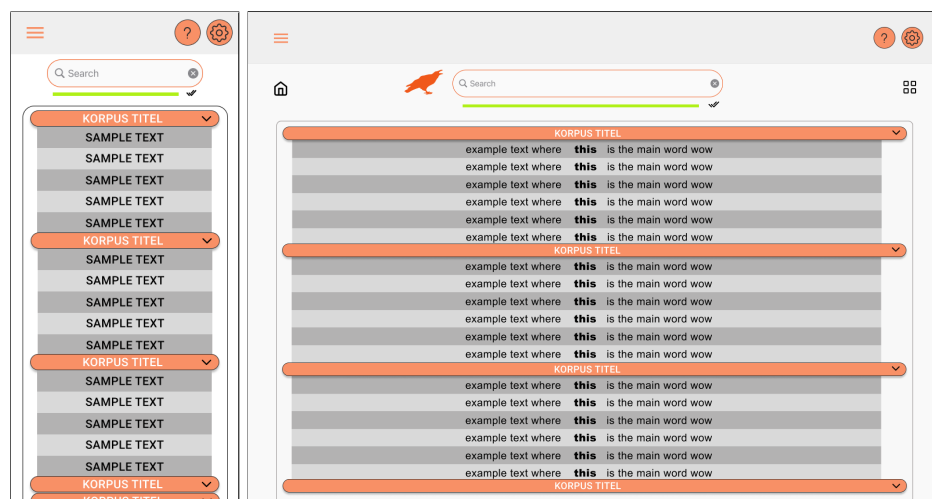


Figure 4.4: New Figma prototype of Korpi in development phase: results page. The result panel has been enlarged to reduce unnecessary white space on the sides.

The feedback from the focus group was acknowledged and carefully considered, and efforts were made to implement their suggestions as accurately as possible. The first focus group provided several key insights about the design. Most importantly, they felt that the Result Panel was too small and suggested it should occupy as much space as possible. Some users reported feeling stuck on certain pages due to a lack of clear navigation options. They also noted that the colour scheme appeared random

and lacked clear meaning, and that white text on light-coloured backgrounds was sometimes difficult to read. On a positive note, participants appreciated the centred search bar and praised the clean design of the landing page. As shown in figures 4.3 and 4.4, the colour scheme was adjusted to offer higher contrast and follow a more intentional design. In figure 4.4, the Result Panel was enlarged, and escape hatches, such as a home icon, was added to improve navigation and help users avoid feeling stuck.

4.4.2 Implemented Functionality in Development Phase

In the development phase, the search functionality and corpus selection as defined in the MVP were developed. This version included the ability to search in one corpus and view the results. Search history was also implemented, an initiative from the project group.

4.4.3 Evaluation in Development Phase: Second Focus Group Meeting

The main purpose of the second focus group meeting was to evaluate the first version of Korpi and to discuss what specific functionality was requested. However, with the younger focus group there was also time to evaluate a new Figma prototype of Korpi and to test a statistics demo.

The meetings had the following agenda:

1. Testing Korpi (and being observed).
2. Answering questions about Korpi.
3. For younger focus group: Testing a new Figma prototype of Korpi with an extended search function (and being observed).
4. For younger focus group: Answering questions about the new Figma prototype of Korpi.
5. For younger focus group: Testing a statistics demo (and being observed)
6. For younger focus group: Answering questions about the statistics demo
7. Discussing what specific functionality is requested.

The questions that were asked can be found in Appendix A.3 and the new Figma prototype for extended search is seen in Appendix A.3.4.

The results from the second focus group meeting can be found in tables 4.4 and 4.5 below, and in tables A.3 and A.2 in Appendix A.4. The relevant slides with the suggested functionality that was discussed with the focus groups can be found in Appendix C. The results from this discussion are presented in table 5.1 in section 5.2.

Table 4.4: Second focus group meeting: Younger focus group evaluating Korpi.

Second focus group meeting: Younger focus group evaluating Korpi		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Design for user-friendliness	General user-friendliness	<ul style="list-style-type: none"> - Korpi is difficult to understand. - Korpi is considered faster than Korp.
	Navigation	<ul style="list-style-type: none"> - A larger and more central corpus selection button is requested. It should stand out from the other buttons on the landing page, since it represents important functionality. - Not self-evident what the corpus selection button with a book icon represents. Information when hovering over is requested. - Difficult to know what corpora a user has searched in when viewing search results. - Difficult to know what button on the landing page is currently selected. - It is requested that selecting how search results are presented should be done on landing page/results page and not under settings. E.g. number of search results that should be displayed.
	Information presentation	<ul style="list-style-type: none"> - Unclear organization of the corpora that can be selected. - The search results should be structured according to corpus. Or this should be an option. - Unclear order of search results, why some come before others. - There should be more space for displaying the search results. E.g. the search bar could be moved when displaying the results. - Parts of words are sometimes displayed in the text context of the search terms in the list of search results. - Some information is missing, e.g. how large the corpora are and when they are updated. - It is difficult to know what <i>kontextstorlek</i>, i.e. context size, means.

Second focus group meeting: Younger focus group evaluating Korpi cont.		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Design for user-friendliness	Visual design	<ul style="list-style-type: none"> - The landing page looks good. It looks like a home page, not a tool. - The modal showing corpora to choose between could have a better visual design. - Good with colours and the choice of colour is good. - Dark mode looks good, not too dark. - Good that the individual search results have alternating colours. - The height of the bars for the individual search results could be reduced. - The layout of borders could be more consistent. Shadowed borders are not appreciated.
Functionality for usefulness	Interesting functionality	<ul style="list-style-type: none"> - A filtering function where you could filter according to newest and oldest search results is requested. - An extended search function related to different inflections is requested. - To be able to search for several words at the same time is requested. - Statistics including graphs is requested.
	Usefulness	<ul style="list-style-type: none"> - No relevant feedback related to usefulness in this part of the evaluation.

Table 4.5: Second focus group meeting: Older focus group evaluating Korpi.

Second focus group meeting: Older focus group evaluating Korpi		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Design for user-friendliness	General user-friendliness	- Difficult to understand what to do. Request to be able to understand better without tour/user guide.
	Navigation	- Navigation is considered ok. - Difficult to know what corpuses are selected, both in corpus selection modal and on search results page. - Not very visible that it is possible to scroll in the corpus modal. - The search bar for searching FOR corpora is mistaken for the main search bar for searching IN corpora. - It looks like there are no search results when the search is not completed.
	Information presentation	- Positive that “cleaner” look than Korp, i.e. not so much information. - Not clear in what order search results are presented. - Not clear what corpora are default. - Better if corpora more relevant to the public are placed higher up in the corpus modal. - Not clear what type of texts all corpora contain, e.g. “L2-Korp”. - Spelling mistakes are not highlighted in the search function. - Words in settings are not easy to understand, e.g. API, context size and the Swedish word <i>provstorlek</i> . - Suggestion of left-aligned text in search results.
	Visual design	- No relevant feedback related to visual design in this part of the evaluation.

Second focus group meeting: Older focus group evaluating Korpi cont.		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Functionality for usefulness	Interesting functionality	<ul style="list-style-type: none"> - Positive that not too little functionality, which could be boring. - Positive that there is a tour/user guide. - To be able to click a search result and get to the original source is requested. - Information about the search term itself (e.g. meaning and origin) is requested. - To be able to search for several words at the same time is requested. - More corpora to search in are requested. - Personal preferences as default corpora is requested.
	Usefulness	<ul style="list-style-type: none"> - The search bar for searching FOR corpora is not considered useful. - Not clear what the web application should be used for, especially not since there are alternatives on the web. - Request that it should be suggested in the web application what it could be used for.

Table 4.6: Second focus group meeting: Younger focus group evaluating new Figma prototype of Korpi.

Second focus group meeting: Younger focus group evaluating new Figma prototype of Korpi		
Problem areas	Evaluation aspects	Summary of feedback related to evaluation aspects
Design for user-friendliness	General user-friendliness	- The extended search option functionality was difficult to understand.
	Navigation	- More information when hovering over different parts of the prototype is requested. - No need for plus sign when adding new search terms to the search field.
	Information presentation	- The linguistic settings are confusing. Both what the linguistic terms mean, what search term(s) they refer to, and the fact that one could search for e.g. the word “blue” as a verb.
	Visual design	- Minor comments. Removal of borders and reduced size of headings in search results are suggested. Word class boxes should also have better readability.
Functionality for usefulness	Interesting functionality	- Information/fun facts to know about the search terms is requested. - Synonyms that could inspire new searches is requested.
	Usefulness	- No relevant feedback related to usefulness in this part of the evaluation.

4.4.4 Functionality Added After the Second Meeting

After the second meeting, there was still functionality that needed to be implemented before the project was ready to enter the detailing phase. Hence, both extended search and statistics were added to Korpi. Additionally, each search result now included more detailed information and allowed users to access the original source articles. To help users understand what to do on the web application, a tour was added that showcased the most important components of the current page. Following requests from the older focus group, a definition of the searched word was added, explained in more detail in section 4.6.

Outside of the responses from the focus groups, there were other functionalities implemented to further improve Korpi. A settings modal were added with multiple parameters: number of corpora shown per page, number of results per corpus, the sentence length around the searched word, a way to toggle between Dark-mode and Light-mode and a button to toggle between the two API’s.

When all of the functionality from the MVP, requests from the focus groups and

improvements from the project group were implemented, Korpi was ready to step into the detailing phase.

4.5 Detailing Phase

In the detailing phase, the web application was finalized and a final evaluation in the form of a questionnaire was conducted.

4.5.1 Design in Detailing Phase

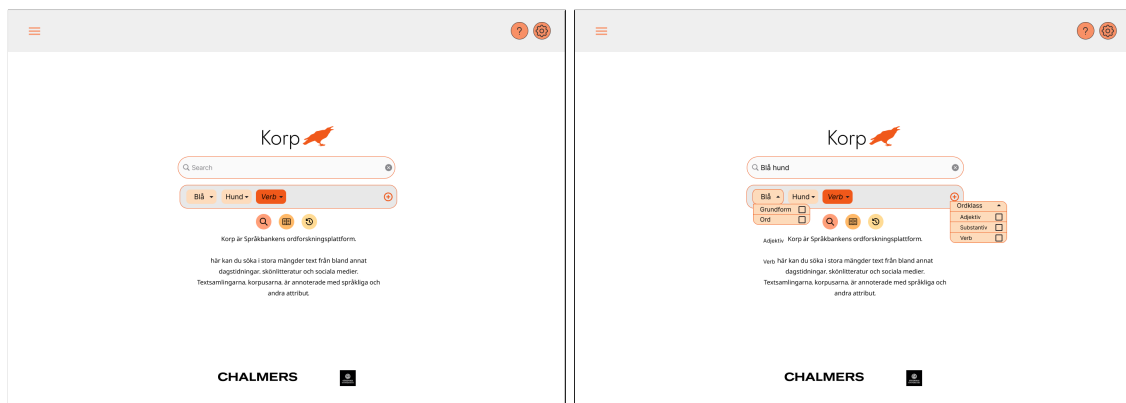


Figure 4.5: Figma design of the extended search in Korpi after the second focus group meeting, featuring improved clarity in the presentation of word class boxes and enhanced options for user selection.

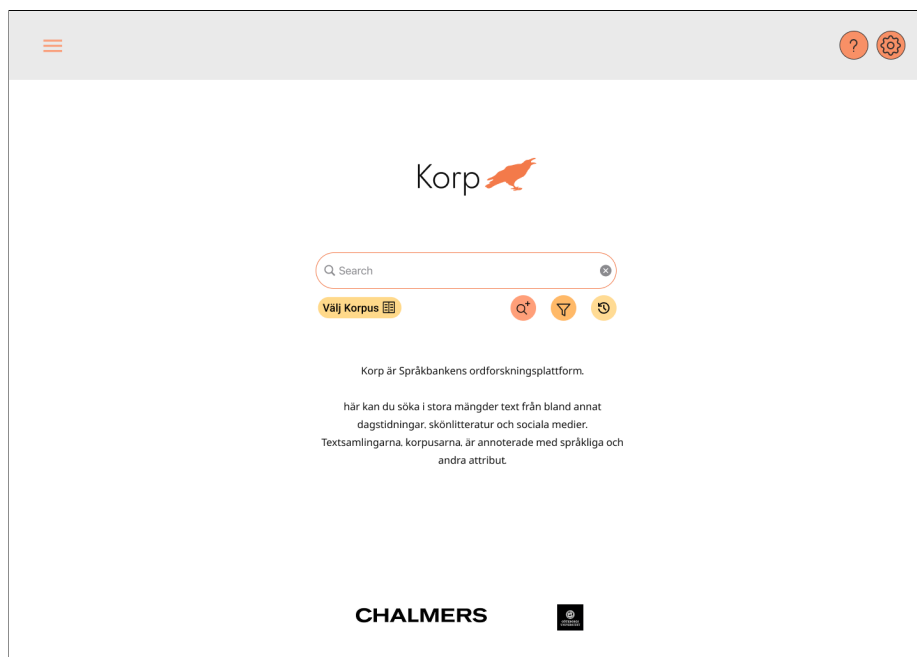


Figure 4.6: Korpi's landing page after second focus group meeting with improved select corpus button, now with the text *Välj Korpus*.

At the second focus group meeting the proposed design for the extended search was discussed, see figure A.1 in Appendix A.3.4. The feedback indicated that the design was difficult to understand, with a key suggestion to improve the way word classes, such as nouns and verbs, are displayed. Specifically, it was recommended that the word class boxes be, italicized or styled differently for better readability.

Figure 4.5 shows the new design for the extended search, featuring improved clarity in the presentation of word class boxes and enhanced options for user selection. The container beneath the search bar was introduced to ensure compatibility with the technical requirements of the web application.

Further design concerns included the results page not displaying which corpora were selected. Participants also suggested that the “Select corpus” button should be more prominent as it plays a central role in the web application’s functionality. It was also recommended to replace the book icon with the text “Select corpus”, as an icon alone was difficult to decipher.

To emphasise the “Select corpus” button it was enlarged and placed directly under the search bar. The text “Select corpus” was also added to make it clearer, see figure 4.6.

Further refinements including how the design for additional information about each search result was created can be seen in figure D.5 in Appendix D.2.

4.5.2 Implemented Functionality in Detailing Phase

The core functionality was already implemented when starting the detailing phase, although some minor changes were made. The settings parameters that were directly connected to the search result, i.e, number of corpora per page, number of results per corpus and sentence length around the search word, were moved to another modal called Modify Search (*sw. Anpassa Sökning*). This was in turn connected to a new button in the button group under the search bar on the landing page and, around whilst on the results page.

4.5.3 Evaluation in Detailing Phase: Third Focus Group Meeting

The main purpose of the final focus group meeting was to perform an evaluation of the final version of Korpi.

The meeting had the following agenda:

1. Testing Korpi on a computer (and being observed).
2. Testing Korpi on a mobile phone (and being observed)
3. Answering a survey.
4. Answering questions about/discussing Korpi.

The tasks that were observed during testing and the questions that were discussed

can be found in Appendix A whilst the survey is in Appendix B.

The results from the final focus group meetings are presented in chapter 5.

4.6 Coding the Web Application

Coding began shortly after the first Figma design was created, illustrated in figure 4.1 and 4.2. Much of the early work was primarily directed towards the frontend, with coding being done in Javascript. The initial codebase was designed to be robust and allow for expansion as more features were added. The designs were divided into separate components and composed into different pages, as dictated by component-based architecture (section 3.3). Early work was primarily focused on the landing and results pages as well as their constituents.

Feedback from the focus groups was taken into account regularly, and suggested improvements to the design or function of the application were implemented. Depending on the suggestion, a new component or a new library were added to the code. These components were then incrementally tested and added to the application.

As an example, the focus groups proposed that it would be helpful to have the definition of a word be displayed in the results. The Definition component was then designed and coded. It was tested separately to ensure that queries to the dictionary API returned the correct results. Finally, the component was placed into the application with styling, ensuring that it fit with the look-and-feel of Korpi. Other features were similarly implemented in an iterative manner, going from Figma design to code prototype.

4.6.1 Frontend

React was used for the frontend. Its emphasis on component-based development allowed for fast translation of designs to modular components in code. For example, the navigation bar was a single component that could be reused and placed on different pages. In turn, the bar was composed of sub-components (buttons) that could be tested and coded separately. Dividing the application in this manner made collaboration easier, as different parts could be developed independently by different developers. Furthermore, errors and debugging were more manageable due to the modular design of the application.

A prominent part of the frontend was to make the interface feel quick and responsive. The application's interactions with the provided API endpoints had to be considered, together with the response sizes. A query to multiple corpus results in large amounts of information that must be parsed and displayed.

Initially the response for each corpus were retrieved by first making a call to obtain start and end position for each corpus. Subsequently a separate call for each corpus were made, with start and end position for corresponding corpus along with all corpora as arguments. This caused the backend to go fetch the results an unnecessary

amount of times. This was not thought to be a problem since both Korpsearch's and Korp's API:s cache their results. It was first recognized as a problem when searching for words in a lot of corpora against Korp's API.

The first steps were to investigate what influenced the response time of the queries. It was found that the initial call retrieving the start and end position was the main problem, and second problem was the handling of arguments when using multiple corpora. The start and end positions did not seem to apply specifically to results that had not been cached yet. When the first search of a non-cached word occurred, the whole result was returned, regardless of start and end values chosen. This was fixed by skipping the initial large query. When requesting each corpus, we only used that specific corpus and a set end value as arguments. Hence, instead of querying all corpora at the same time, multiple batches of requests for single corpora were sent. Through this, we no longer needed to retrieve the start and end positions and avoid having the backend to go through redundant results multiple times.

4.6.2 Backend

The backend for the application was not developed by the project team as Korpsearch¹ was provided to the project team. Korpsearch was designed with the same request and response structure as Korp API. This meant that Korpi could act as a frontend for both services with little changes to the code. During development, this proved beneficial as features that were not yet implemented by Korpsearch could be tested against the larger Korp API without changing the frontend.

¹Korpsearch, <https://github.com/heatherleaf/korpsearch>

5

Results

This chapter presents the final version of Korpi that was developed, as well as the focus groups' final evaluation of the application.

5.1 Final Design of Korpi

This section outlines the finalised design of the web application, including the landing and results pages, highlighting the important components. The final version is hosted on *korpi.netlify.app* with the code repository available on Gitlab¹.

5.1.1 Landing Page

The landing page was designed to put the search bar in the centre as it is the most crucial component. Above the bar is the application's logo and below it are the main buttons; Choose Corpus (*sw. Välj Korpus*) followed by icons for Extended Search (*Utökad Sökning*), Modify Search (*sw. Anpassa sökning*), and History (*sw. Historik*) (see figure 5.1).

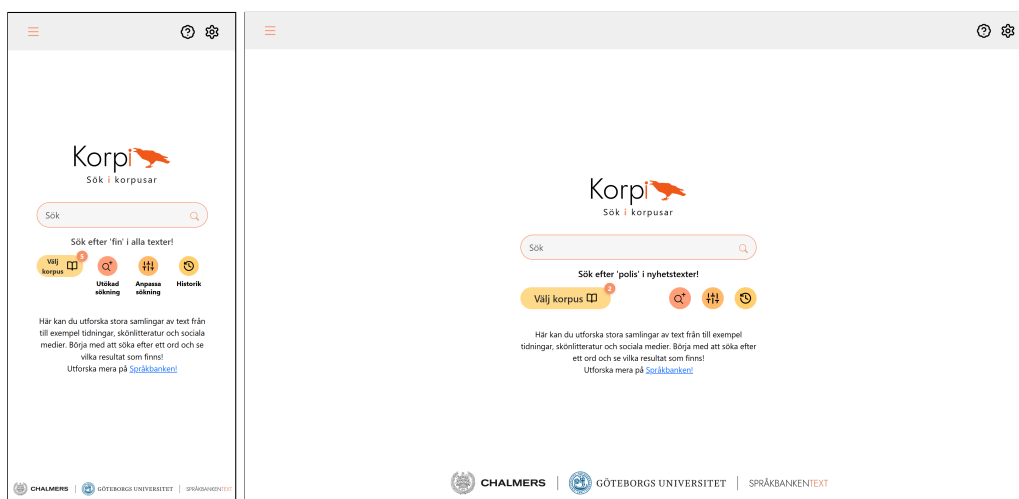


Figure 5.1: Korpi's final landing page. The emphasis is on the search bar, with it taking centre stage. Left mobile view, right desktop view.

5.1.2 Choose Corpus

Choose Corpus (*sw. Välj Korpus*) was designed to be a modal which shows all available corpora, subdivided into categories, see figure 5.2. Multiple corpora in a

¹<https://git.chalmers.se/pomesh/korpi>

category can be selected at once, and unwanted corpora can be removed. To select all corpora from one category, the checkbox to the left of the category name can be checked.

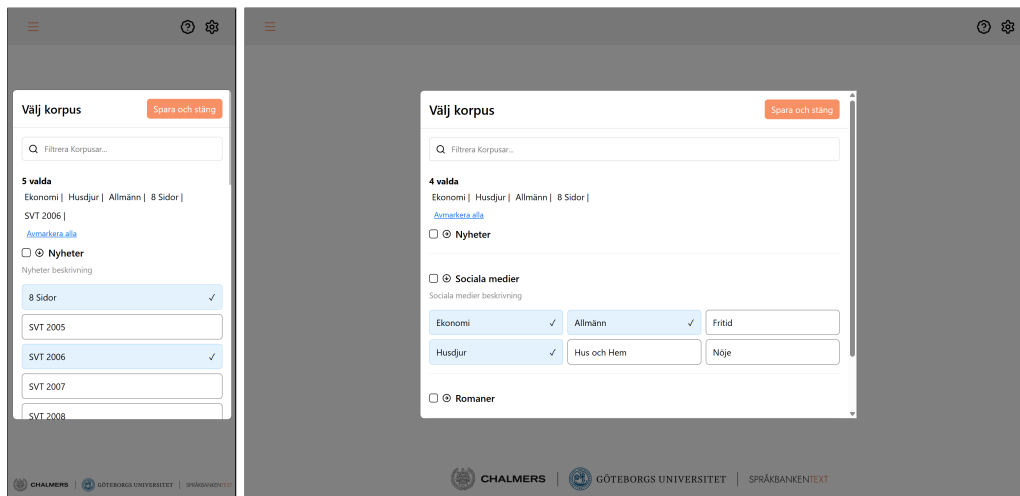


Figure 5.2: Choose Corpus (sw. Välj Korpus) modal.

5.1.3 Extended Search

Extended Search is illustrated in figure 5.3. When enabled, queries with multiple words can be searched. These words can then be reordered by clicking the arrows on the sides of each word component. Further, it is possible to add wildcard entries for various word classes. For example, with extended search the user can query a corpus for all adjectives followed by their searched word, e.g. “Adjective *car*” returns *red car*, *broken car* and so on.

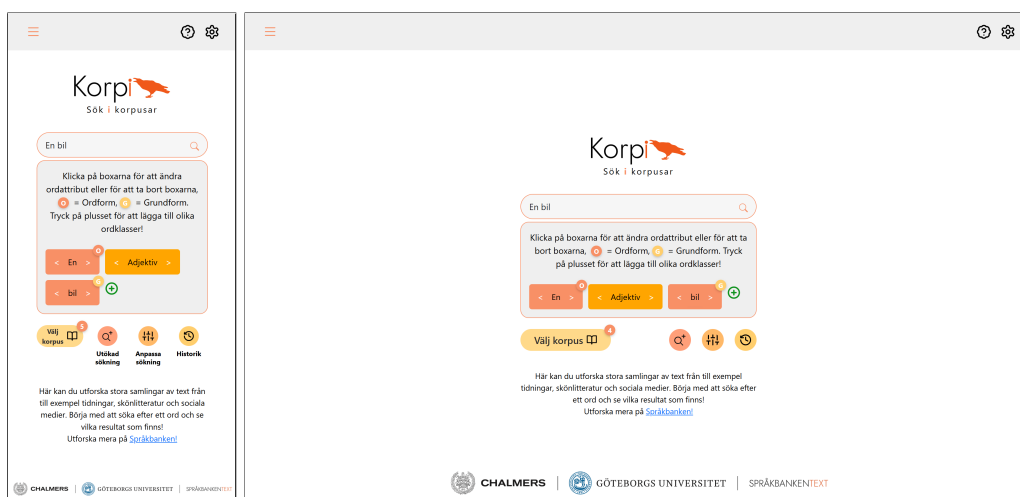


Figure 5.3: Extended Search Panel with three words chosen.

5.1.4 Settings Modal

The settings modal, figure 5.4 allows the user to switch between light and dark modes for the application as well as choose the API to query: *Språkbanken API*

5. Results

(Korp API) and *Ny API* (Korpsearch). In figure 5.5, dark-mode version of the page is displayed.

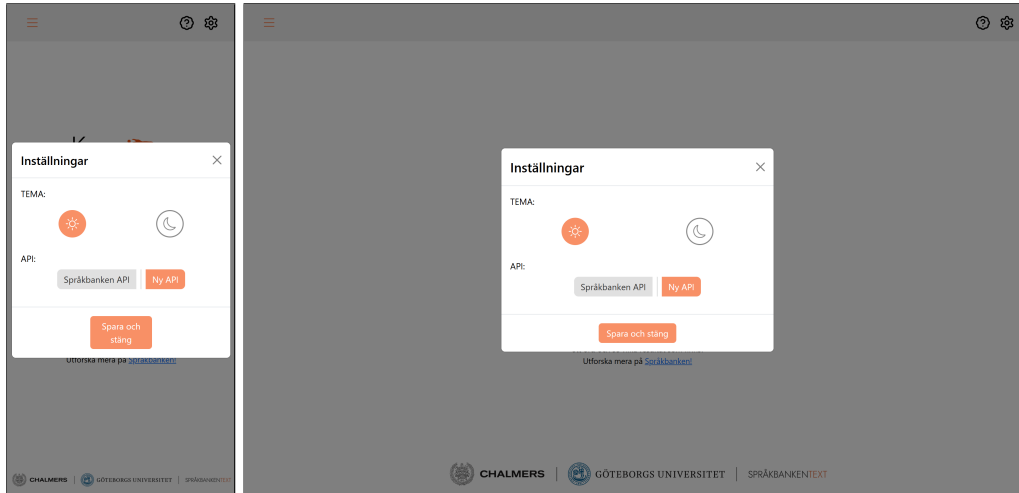


Figure 5.4: Korpi's final Settings Modal.

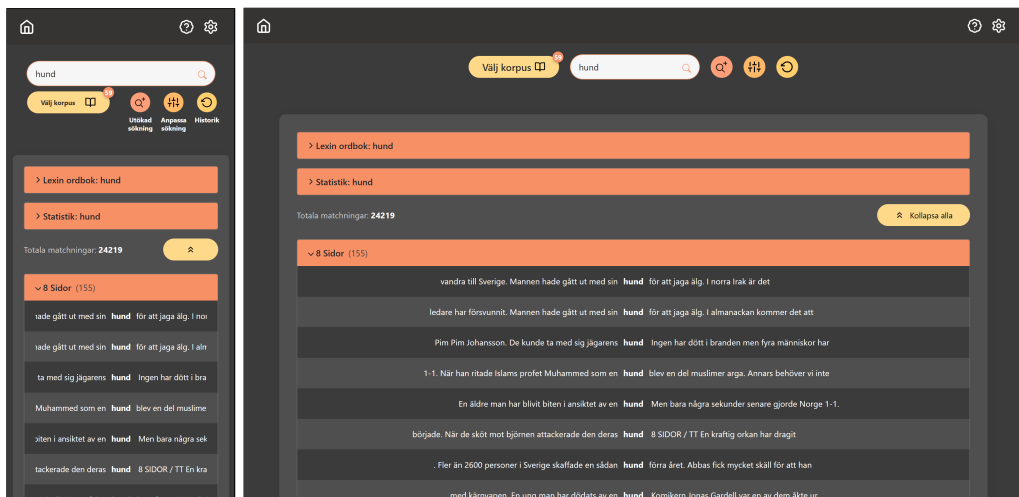


Figure 5.5: Result Panel in Dark-mode

5.1.5 Modify Search Modal

The modify search modal, figure 5.6 allows the user to change how results are viewed in the results panel. This includes the number of corpora shown per page, number of results per corpus and the sentence length (number of words in each result).

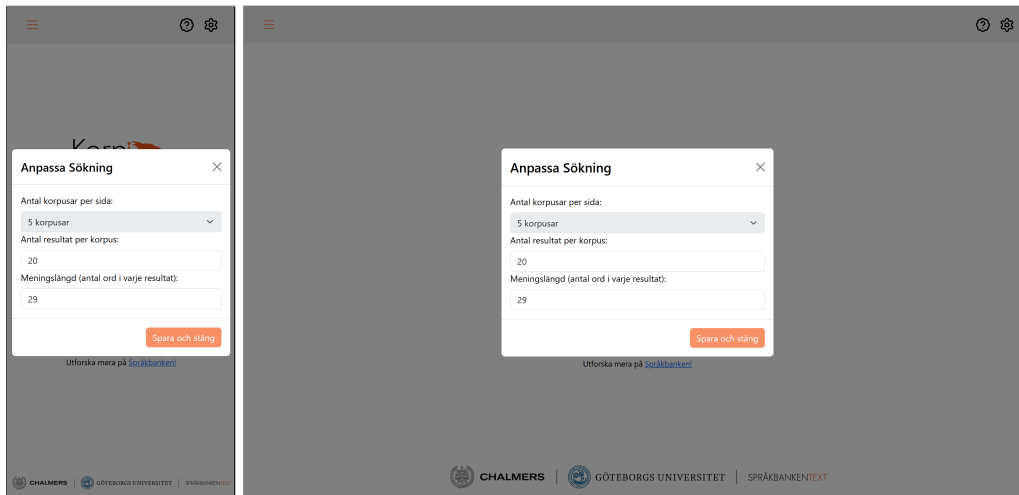


Figure 5.6: Modify Search Modal

5.1.6 History Panel

The history panel shows the user's latest search history, figure 5.7. Each search can be clicked and will take the user to that specific search. Furthermore, the date and time the search was made is displayed. An option to remove the item from the search history is provided.

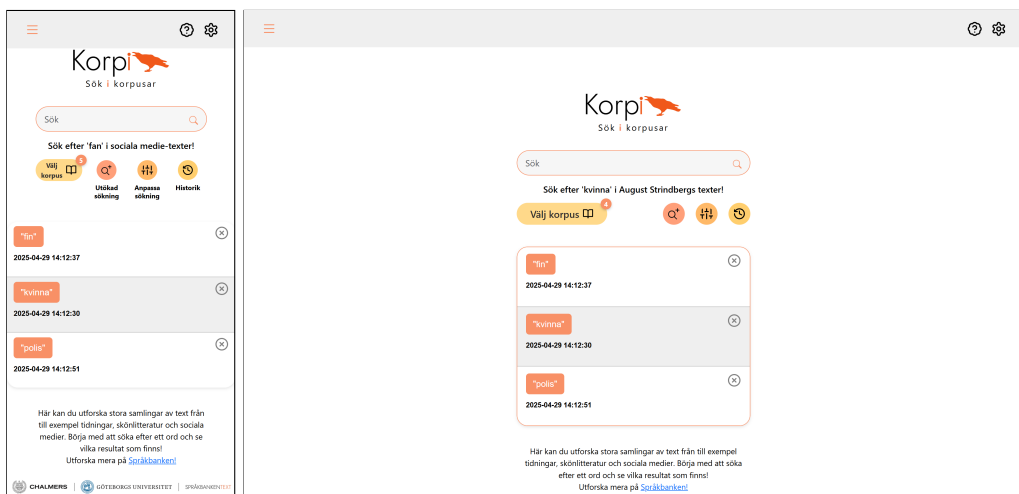


Figure 5.7: History Panel

5.1.7 Tour

The question mark icon on the top right, in the navigation bar, can be clicked to start the tour, see figure 5.8. There are different tours for the landing and results page. The tours take the user through the steps needed to make a search, change corpora, view results and so on.

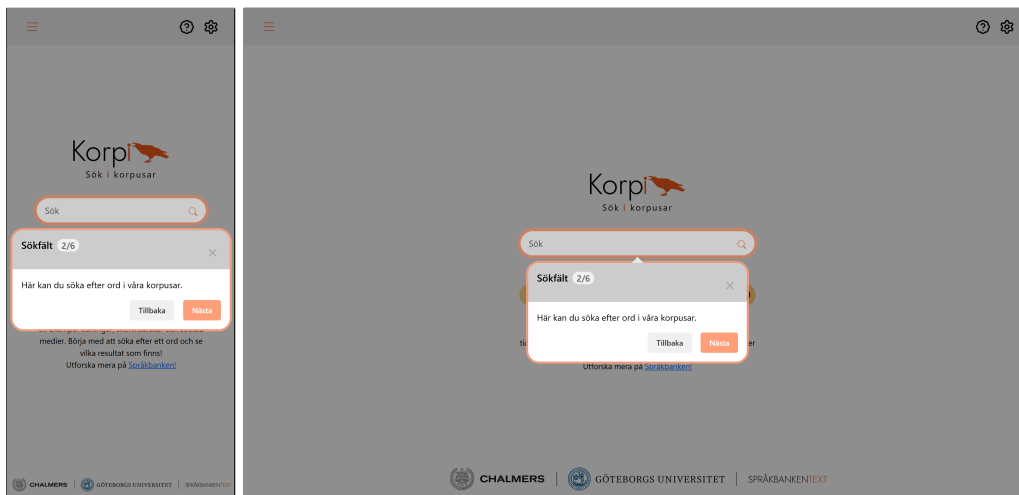


Figure 5.8: Landing Page Tour

5.1.8 Results Page

The final version of the results page layout is shown in figure 5.9. Some elements such as the Navigation Bar and its icons are the same as those on the landing page. However, there are changes to the search bar and its associated icons. These are repositioned to the top of the screen, around the search bar, giving more space to display the results. The main components for the results page are the result panels and cards, the statistics panel and the definition panel.

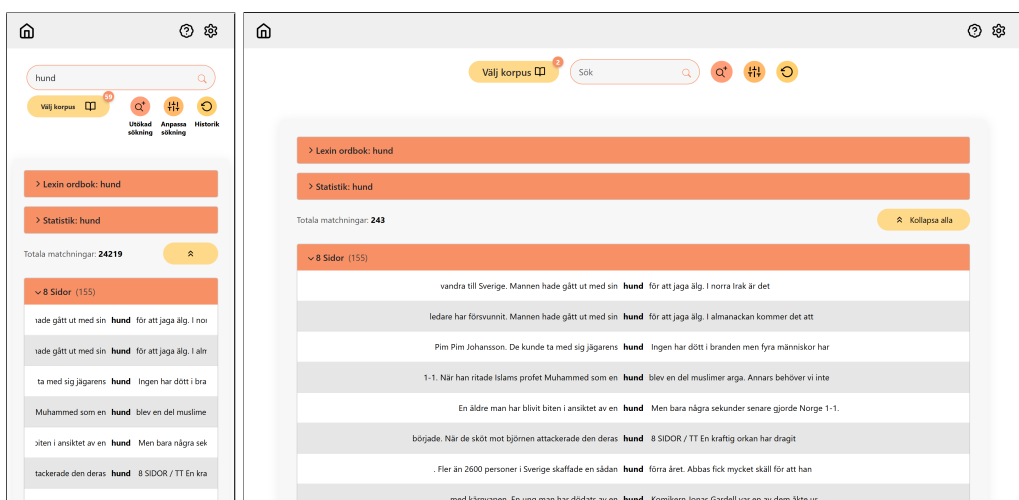


Figure 5.9: Result Page

For more information about the results, each card can be clicked, see figure 5.10. This expands the card to show extra details such as title, date, author, and the link to the source of the sentence.

The Collapse All (sw. Kollapsa alla) button can be used to collapse all of the results from each corpora (figure 5.11). The user can also control which corpora

they want to see by manually clicking on their accordion items, which are collapsible (figure 5.10). When the user has collapsed all the results, the Collapse All turns into Expand All (*sw. Expandera alla*) which in turn expands all the result.

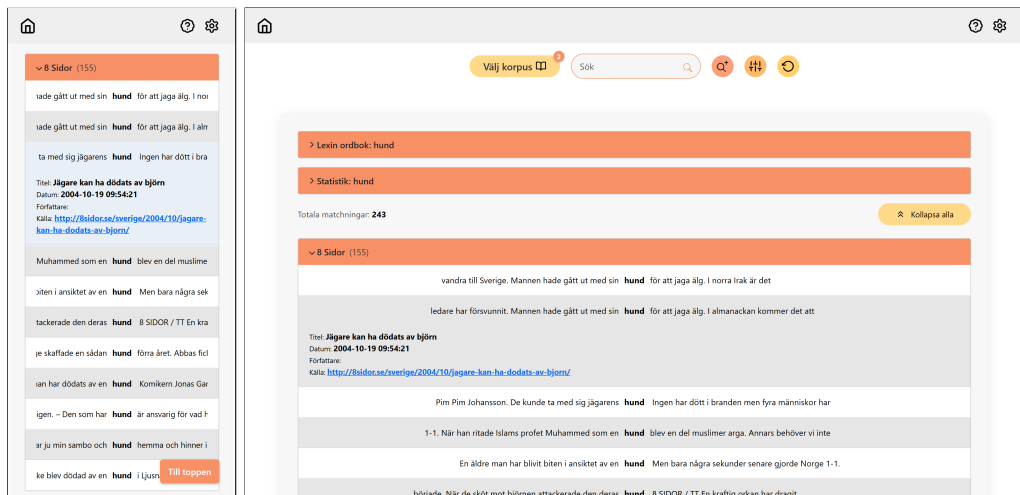


Figure 5.10: Result Card Information

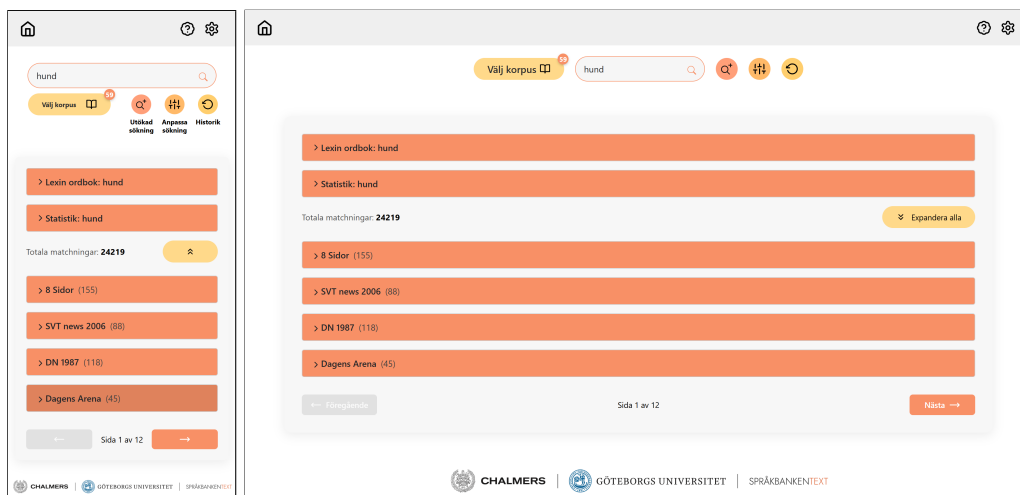


Figure 5.11: Results Panel with all tabs closed.

5.1.9 Dictionary

The dictionary panel (*Lexin ordbok*) is similarly an accordion item, populated by definitions of every searched word, figure 5.12. The panel shows three definitions, sometimes together with example sentences. A link to a publicly available dictionary (Lexin) is provided for the complete set of definitions and sample sentences.

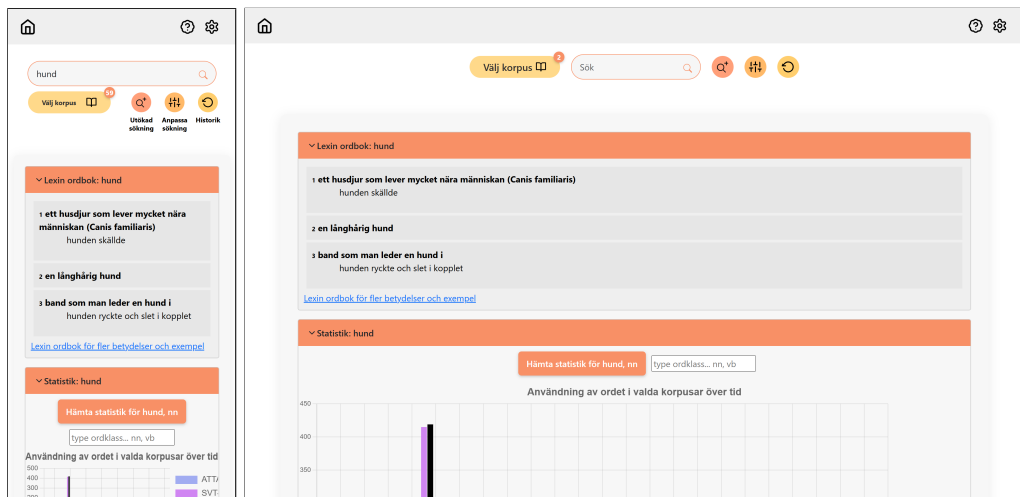


Figure 5.12: Definition Panel. A separate accordion item is created for each searched word. The top three definitions are displayed, sometimes together with example sentences. A link to the definitions in the dictionary Lexin is also provided.

5.1.10 Statistics Page

The statistics panel is also a part of the accordion. The graph is populated by number of occurrences by year of the searched word, in the selected corpora figure 5.13. If a query consists of multiple words, as in extended search, statistics is retrieved for each word and displayed in separate panels.

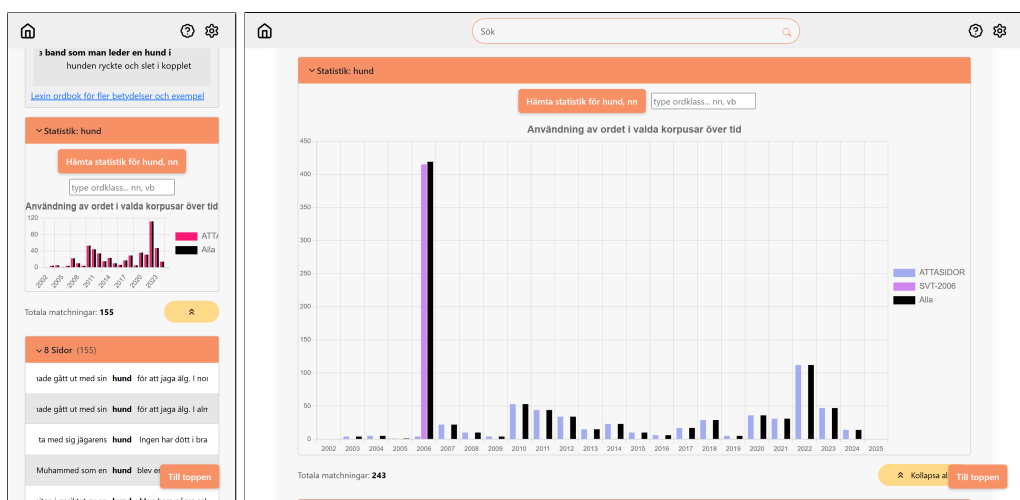


Figure 5.13: Statistics Panel

5.2 Final Functionality of Korpi

The main final functionality of Korpi, and potential future functionality as suggested by the focus groups, is listed in table 5.1. The implemented functionality is divided into functionality in the MVP, additional functionality that was requested by the focus groups and additional functionality initialized by the project team. It can be

noted that the focus groups were not interested in more functionality of a purely linguistic/academic kind.

Table 5.1: Main implemented and potential future functionality of Korpi.

Implemented and potential future functionality of Korpi
<p>Functionality requested by supervisor, Minimal viable product (MVP):</p> <ul style="list-style-type: none"> - Search Function - Result Presentation - Corpus Selection - Individual Result Information - Extended Search Options - Menu
<p>Additional functionality requested by focus groups:</p> <ul style="list-style-type: none"> - Statistics (Frequency, Frequency change over time – both tables and graphs) - Tour - Dictionary
<p>Additional functionality initialized by project team:</p> <ul style="list-style-type: none"> - Organization of search results according to corpus - Modify Search (number of results per corpus, number of corpora per page, size of text surrounding individual search result) - Search history - Settings
<p>Potential future functionality (suggested by the focus groups):</p> <ul style="list-style-type: none"> - Filtering of search results (Alphabetic order, New/old search results) - Statistics (Frequency comparison of words) - Tips for how to use Korpi - Suggestions of synonyms - Lists and statistics of e.g. new/outdated words or words that have changed meaning

5.3 Final Evaluation of Korpi

This section outlines the final evaluation of Korpi by the focus groups. How the final focus group meetings were conducted was described in section 4.5.3.

The survey that was given to the focus groups consisted of one quantitative part with twelve categories, where the participants were asked to grade different aspects of Korpi from one (poor) to five (very good), see figure 5.14. The categories were as follows: useful, responsiveness, speed (searches), navigation, design, interesting, ease of understanding, experience when using, presentation of information, presentation of statistics, presentation of search results, and Korpi in its entirety. The focus group participants were also presented with 13 written-answer questions. The questions and a summary of the answers can be read in table 5.2. A discussion of the survey also took place during the focus group meetings, and important clarifications from the focus group members are also included in table 5.2.

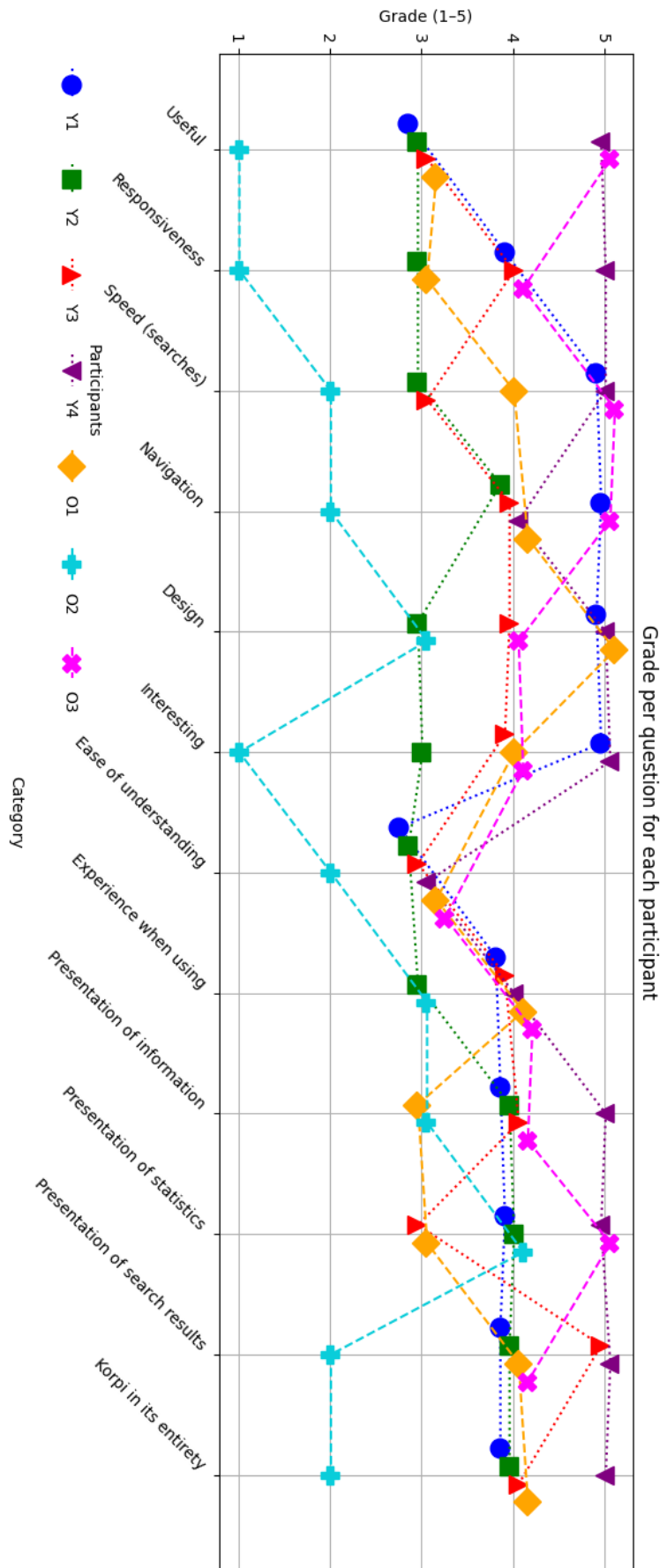


Figure 5.14: This graph shows the grade (1-5, where 1 is poor and 5 very good), each member gave the final version of Kori with respect to the evaluation aspects. Y1-4 are members of the younger focus group, and O1-3 are from the older group, they are presented as a unique combination of colour and shape. Member O3 missed the last question and therefore did not answer.

Table 5.2: Summary of written answers from Younger and Older groups to all feedback questions. Answers are translated from Swedish to English.

Question	Group	Summary of Comments
Is the site interesting? If not, what is missing for it to become interesting?	Younger	Korpi is seen as interesting; more clarity and examples desired. Colours help. More statistics requested.
Is the site interesting? If not, what is missing for it to become interesting?	Older	Not seen as very interesting; site lacks clarity. Users want editable results.
How was your experience using the site?	Younger	Experience generally positive.
How was your experience using the site?	Older	Not entirely positive; frustrations with limited search functionality.
How was navigation and modal use?	Younger	Navigation okay but not as praised as design. No issues with modals.
How was navigation and modal use?	Older	Some issues with modals (trial and error). Navigation considered good.
Which information on the website is easy or difficult to understand?	Younger Difficult	Advanced/extended search hard to understand (e.g., needing a space).
Which information on the website is easy or difficult to understand?	Younger Easy	Most things were easy to understand, excluding advanced search.
Which information on the website is easy or difficult to understand?	Older Difficult	Mixed responses; detailed elements were hard to grasp.
Which information on the website is easy or difficult to understand?	Older Easy	One noted search result presentation was easy to understand.
What do you think about how the search results are presented?	Younger	Positive; grouping and colour choices praised.
What do you think about how the search results are presented?	Older	Also positive, though possibly less enthusiastic than younger group.

Continued on next page

5. Results

Question	Group	Summary of Comments
What do you think about the dictionary and word statistics?	Younger	Generally good, but some dictionary entries seemed odd.
What do you think about the dictionary and word statistics?	Older	Perceived as good.
What do you think about the design of Korpi?	Younger	Positive overall; colour choices specifically praised.
What do you think about the design of Korpi?	Older	Only positive remarks.
Is the mobile version comparable to the desktop version?	Younger	Generally comparable; corpus module and statistics harder in dark mode.
Is the mobile version comparable to the desktop version?	Older	Comparable or slightly better than desktop; one noted mobile cuts off full sentences.
Overall thoughts: Good	Younger	Interesting, user-friendly, visually appealing. Unique to non-linguists.
Overall thoughts: Bad	Younger	Doubts about usefulness to public. Too few corpora leads to empty results.
Overall thoughts: Additional thoughts	Younger	Could suit academics. Requests for more features, simpler corpora selection, more info.
Overall thoughts: Good	Older	Unique tool combining texts and dictionaries. Useful for specific groups.
Overall thoughts: Bad	Older	Usefulness questioned. Needs more features. Advanced search difficult. Bugs.
Overall thoughts: Additional thoughts	Older	Target audience unclear. Suggests educational videos. Interesting and has potential.

6

Discussion

The purpose of this project was to “*develop a frontend for an existing corpus search engine based on Korpi; from being functionally-oriented towards professionals, to being user-friendly for the general public while remaining useful. The application should be easy to use and understand on both mobile and desktop devices*”.

In this chapter, the fulfilment of this purpose as well possible sources of error and ethical considerations are outlined. Finally future improvements are discussed.

6.1 Fulfilment of Purpose

The final version of Korpi and the evaluation by the focus groups is presented in chapter 5, Results. The majority of the focus group participants gave Korpi an above average grade in its entirety (see figure 5.14, *Korpi in its entirety*). According to the focus groups, Korpi worked well, with the same interface, on both mobile devices and laptop. In the subsections below, the findings of the project with respect to the two problem areas are discussed.

6.1.1 Design for User-Friendliness

The majority of the participants in the focus groups gave Korpi an above average grade (see figure 5.14) regarding the *Experience when using Korpi*. This could be due to Korpi’s focus on navigation, presentation of information and visual design. A decision that has been important for the general understanding and that relates to navigation, information presentation and visual design, is the use of the centre-stage design pattern. Centre-stage in this case meant the search field was large and placed in the centre, similar to other popular search engines. This made the web application easier to understand and provided a visual anchor for starting the search.

Navigation is an aspect that determines user-friendliness. The focus groups did not have any negative comments about the general navigation of Korpi 5.2. The simplicity of the navigation is due to a number of factors. Firstly, the use of modals made the different parts of the web application easily accessible without leaving the current page. These features were made possible through good utilization of screen space. Secondly, the size and colour of buttons used in Korpi made their function clearly visible. Finally, the application had clear escape hatches to simplify finding the way back to landing page, e.g. home button.

The information presentation has been another evaluation aspect in the project.

One feature that was central to this aspect was the accordion design pattern. This was used since accordions are a clear way to present varying information in a coherent manner, in one single component. In Korpi, both statistics and search results have been presented in the same component, separated by an accordion. The accordion generally made it possible to present a lot of information in a digestible manner, without overwhelming the user.

The visual design of Korpi, especially the colours, was also appreciated by the focus groups. The choice of colours was consequent in the web application. The orange colour was chosen in order to fit with with colour of the logo of Korp. Thus, Korpi could visually serve as part of the existing applications provided by Språkbanken.

In the final evaluation questionnaire, a question that scored low was the general ease of understanding of the web application, see figure 5.14. The main reason for this is the extended search function, which was generally considered unintuitive. For example, it was difficult to understand that all search terms had to be separated by space and what it actually meant to search for e.g. adverbs or adjectives in general. Especially on the results page, since two search bars are on the screen at once, the users did not realize which bar was active, see figure D.3 in the appendix. Further, the query in the top bar was still present in the search box, even when users were typing in the bar below. This made it difficult to understand which word or phrase was being searched. This problem is understandable, since the extended search is still a work in progress and in need of refinements.

6.1.2 Functionality for Usefulness

The question of whether something is useful or not is complex. Is something only useful if it lives up to its intended purpose? When evaluating usefulness of a product, the purpose of use needs to be clear. The purpose of use was difficult to define in this project. Instead the approach was to ask the focus groups if they found Korpi useful. Generally, no one knew if they themselves would use Korpi, but imagined it could be helpful for people learning languages. There is also perceived usefulness, which refers to the belief that something is useful, even if it may not be personally beneficial. Thus, the question “is this useful” which was asked regarding Korpi, has been deemed flawed by the project team in hindsight. It cannot be known which of the two interpretations of usefulness, for themselves or others, participants thought of when answering the questions regarding usefulness. The conclusion was that asking about the usefulness of something without providing context, as was done in this project, was a poor choice. This was evident in the survey responses, where the ratings for usefulness varied significantly.

A clearer question is if the focus groups thought of the web application as interesting. This could fall under the category of perceived usefulness. The grade they gave was above average on how interesting Korpi is (figure 5.14, label *Interesting*). In general, the younger focus group was more interested in statistics and the older focus group more interested in understanding the meaning of words and seeing examples of their use.

The same figure illustrates that Korpi is less useful than interesting, however the difference is not that large. However, taking the discussions during the focus groups into consideration, it is doubtful if the web application is really thought of as useful.

The focus groups, especially the older focus group, were requesting suggestions regarding how to use Korpi. It was difficult for them to find use-cases for a corpus search engine. This request refers to something other than the general help function that is already implemented. The current help function explains the purpose of the different parts of the web application rather than provide concrete use-cases for Korpi.

Since the participants of the focus groups found it difficult to know what was useful for them, it made sense to instead address what was interesting for them. In order to get a picture of what type of functionality could be worth building on in the development process. This could potentially be a way to go about interacting with focus groups in other projects where the usefulness question is undefined. Even though it would be important to still remember that the web application in the end also has to be useful.

6.2 Possible Sources of Error

There are a number of factors that potentially weakens the findings of the report. Firstly, the use of focus groups must be scrutinized as they were central to both the development and evaluation of Korpi. This includes the size and composition of the groups. The number of participants was small, totalling seven people across two groups. This was smaller than the average number of participants of focus groups. This allowed for more in-depth understanding of the user experience. However, a smaller sample size also entails less reliable data. It is difficult to perform statistically sound analyses with such numbers and in turn difficult to form quantitative conclusions. Further, it is not possible to infer that the findings from these small groups are applicable for the general public. However, in the project there was not sufficient resources to perform both an in-depth investigation and a broader survey targeting many users. Understanding a few users in depth was considered to be more important in the beginning of a development process, when no close to finished product yet exists for the broader public to form an opinion about. Regarding the composition of the groups, it turned out that all of the focus group participants but one had an engineering background.

The second possible error source concerns the survey evaluation of Korpi. In the quantitative part of the survey, the application was evaluated against a scale of 1-5, corresponding to “Not Good” to “Good” respectively. Retrospectively, this was too coarse of a gradation. There were no specific guidelines to what each of the grades pertained and the respondents may have had difficulty in choosing between a four or a five for example. It could have been better to make the scale range from e.g. 1 signifying “Bad”, 3 signifying “Good”, and 5 signifying “Very good”, in order to better catch the differences between different degrees of “Good”. Additionally, the

categories that were quantitatively evaluated were presented as singular words (e.g. *usefulness* or *user-friendliness*), which may have allowed individuals to interpret these words too freely.

The third possible error source regards the focus groups' evaluation of the final product. It is likely that the results from the final evaluation are skewed since the focus group were the ones suggesting changes to Korpi. This could have inclined them to give a higher grade than if they had interacted with the same improvement for the first time.

Finally, it is important note that the focus groups have only evaluated Korpi in isolation and not in direct comparison to other corpus search engines except for Korp. This implies that even Korpi is considered useful by the current respondents, there could still be a competing tool that is chosen ahead of Korpi but that has not been evaluated.

6.3 Ethical Considerations

Whilst working with corpora and search engines, there are a number of ethical considerations that need to be addressed. The relevant issues for Korpi are discussed below.

6.3.1 Regarding Information Access and Copyright in Corpora

As corpora are collections of large bodies of text, composed from a multitude of sources, their contents may contain sensitive information. This can take the form of explicit or hateful language which may provoke users. Similarly, personal details published in a source article can show up in the results of a query, including the addresses or other personal information. Further, if an article is redacted in the future, perhaps due to false information, it may still show up in the search results unless it is also removed from the database. Finally, another major issue with corpus search engines is that of copyright. Theoretically it is possible to query words such that the entire source article can be recreated from searches alone.

The issues outlined above are all crucial to corpus search engines and it is important to acknowledge them. However, due to the limited scope of the project, they were not addressed in totality with the current application. They are primarily backend considerations, which was not developed by the group. With Korpi being a front-end application, the above issues are the concern of the Korp API and Korpsearch. Korpi is merely a portal to the corpora, the results it displays are those that the servers provide. There is no censorship of content.

6.3.2 Ethics in Web Development and Security

Since the application is hosted and served on the web, considerations about security need to be discussed. Navigating to Korpi requires the download of assets and code to the user's web browser. Malicious code is always a concern in these situations.

For Korpi, the entire code base is available in a public repository. Care has been taken to not include potentially unsafe libraries during development and deployment. This does not necessarily preclude potential exploits that may be discovered at later dates. However, continuous development of the application to remove outdated libraries is outside the project's scope.

The current prototype is hosted on Netlify. An overview of their security assurances is given by Netlify's official documentation [26]. Nevertheless it is important to outline some network issues that can arise by using Korpi. Even though Korpi has no backend code, it still acts as a portal to the servers housing various corpora and other data. The search bar is a potential entry point for sending malicious code to the servers. The backend can also be flooded by sending multiple queries as is the case in denial-of-service attacks. The measures needed to combat these issues are the responsibility of the backend however, and not of the frontend.

6.3.3 Social Aspects

Korpi could be used in combination with other tools to help learning the Swedish language. Some Swedish words might not be perfectly translatable to other languages and simply looking up a dictionary definition may not fully capture their true meaning. Instead one could search in Korpi and find out how the word is actually used in natural language as well as grammar and example usages. One example of this is the Swedish word *lagom*. When translating *lagom* to English, the translation is often *just right*, which is not quite the real meaning. If this was searched on Korpi, the result would be multiple sentences where *lagom* is used and might give the learner a more complete picture of the actual meaning.

6.3.4 Focus Groups

The use of focus groups involved the handling of certain personal data to enable repeated contact with participants. This included information such as names and ages as well as other, potentially relevant details. Furthermore, data were stored regarding how the participants interacted with the web application, as well as their opinions and feedback. Legal frameworks such as the General Data Protection Regulation (GDPR) governed the processing of personal data [27]. However, from an ethical standpoint, it was equally important that any information provided to the project in confidence was not disclosed to third parties.

6.4 Future Improvements

The scope of the project imposed certain limitations on the development of Korpi. Firstly, being a frontend application, Korpi could only query and display results from corpora that were predefined in the backend servers of Korp and Korpsearch. Adding additional corpora cannot be done through Korpi alone. Secondly, some functionality that focus groups found interesting could only be implemented by the backend. This includes support for additional statistics which is not present in Korpsearch, i.e. frequency of words in individual corpora. The current statistics module queries the original Korp API instead of Korpsearch for this reason. Finally

the source, author and date of the individual results is only available from Korp API and not Korpsearch.

Another limitation to Korpi's development was time. The minimal viable product was implemented and evaluated in the given time frame but it was not free from bugs. On some page refreshes the application does not query the server leading to no results being displayed. The statistics panel needs two button presses to display the graph of frequencies from Korp API. The history panel can sometimes delete all history or even resurrect deleted items in some cases. Extended search could have been made clearer and implemented draggable buttons for the individual buttons. Further time and bug-testing is needed to eliminate these issues.

7

Conclusion

Korpi is a user-friendly web application for the public that works well on both mobile and desktop devices. It provides an intuitive interface, clean design, and is streamlined to emphasize the essential functionality of a corpus search engine. Through Korpi, it is possible to search for words in corpora and the foundation of extended search is in place. Other functionalities such as search history, statistics and lexicon have also been implemented on a basic level.

The application was evaluated by two focus groups of differing ages using a survey including a questionnaire. The results show that Korpi is user-friendly because of the design choices implemented. Korpi can already be considered useful in its current form and can be viewed as a better option for first-time users that want an introduction to linguistics than Korp. Yet, there is still room to make Korpi even more useful; the statistics and lexicon functionalities can be improved, and more attributes can be added into the extended search function.

Bibliography

- [1] H. Lindquist and M. Levin, *Corpus Linguistics and the Description of English*, 2nd ed., Edinburgh, UK: Edinburgh University Press, 2018.
- [2] Språkbanken. “Om oss.” (Apr. 2025), [Online]. Available: <https://www.sprakbanken.se/om-oss> visited 1-4-2025.
- [3] Språkbanken. “Språkbanken text.” (Apr. 2025), [Online]. Available: <https://sprakbanken.se/om-oss/organisation-och-verksamhet/sprakbanken-text> visited 1-4-2025.
- [4] J. Nielsen. “A 100-year view of user experience.” (Dec. 2017), [Online]. Available: <https://www.nngroup.com/articles/100-years-ux/> visited 10-4-2025.
- [5] M. Arvola, *Interaktionsdesign och UX: Om att skapa en god användarupplevelse*. Studentlitteratur, 2020.
- [6] G. J. James, *The Elements of User Experience: User-Centered Design for the Web and Beyond*, 2nd ed., Berkeley, CA USA: New Riders, 2011.
- [7] D. Benyon, *Designing Interactive Systems*, 3rd ed., Edinburgh, UK: Pearson Education, 2014.
- [8] J. Tidwell, C. Brewer, and A. Valencia, *Designing Interfaces: Patterns for Effective Interaction Design*. O’Reilly Media, 2020.
- [9] J. Evanick, “Implementing mobile-first strategies in online education,” in *Creative Approaches to Technology-Enhanced Learning for the Workplace and Higher Education*, ser. Lecture Notes in Networks and Systems, vol. 1150, Springer, 2024, pp. 157–182. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-72430-5_14.
- [10] L. Borin, M. Forsberg, and J. Roxendal, “Korp – the corpus infrastructure of språkbanken,” in *Proceedings of LREC 2012. Istanbul: ELRA*, vol. Accepted, 2012, pp. 474–478.
- [11] Språkbanken. “Korp API (8.1.0).” (2025), [Online]. Available: <https://ws.spraakbanken.gu.se/docs/korp> visited 11-2-2025.
- [12] P. Ljunglöf, N. Smallbone, M. Thoreson, and V. Salomonsson, “Binary indexes for optimising corpus queries,” in *Proceedings of the 20th Conference on Natural Language Processing (KONVENS 2024)*, P. H. Luz de Araujo, A. Baumann, D. Gromann, B. Krenn, B. Roth, and M. Wiegand, Eds., Vienna, Austria: Association for Computational Linguistics, Sep. 2024, pp. 149–158. [Online]. Available: <https://aclanthology.org/2024.konvens-main.17/>.
- [13] StackOverflow. “Stack overflow developer survey 2024.” (Mar. 30, 2025), [Online]. Available: <https://survey.stackoverflow.co/2024>.

- [14] Meta Platforms. “React – A JavaScript library for building user interfaces.” (2024), [Online]. Available: <https://react.dev/> visited 8-5-2025.
- [15] Figma. “Figma.” (Sep. 2024), [Online]. Available: <https://www.figma.com/> visited 20-1-2025.
- [16] The Bootstrap Authors. “Bootstrap – The most popular HTML, CSS, and JS library in the world.” (2024), [Online]. Available: <https://getbootstrap.com/> visited 8-5-2025.
- [17] GitLab. “GitLab – The DevSecOps Platform.” (2024), [Online]. Available: <https://about.gitlab.com/> visited 8-5-2025.
- [18] GitLab. “Git Basics | GitLab Documentation.” (2024), [Online]. Available: <https://docs.gitlab.com/topics/git/> visited 8-5-2025.
- [19] Netlify, Inc. “Netlify – The Modern Web Development Platform.” (2024), [Online]. Available: <https://www.netlify.com/> visited 8-5-2025.
- [20] E. L.-C. Law, V. Roto, M. Hassenzahl, A. P. Vermeeren, and J. Kort, “Understanding, scoping and defining user experience: A survey approach,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '09, Boston, MA, USA: Association for Computing Machinery, 2009, pp. 719–728. DOI: 10.1145/1518701.1518813. [Online]. Available: <https://doi.org/10.1145/1518701.1518813>.
- [21] C. B. Seaman, “Qualitative methods in empirical studies of software engineering,” *IEEE Transactions on Software Engineering*, vol. 25, no. 4, pp. 557–572, 1999. DOI: 10.1109/32.799955.
- [22] F.-H. Leung and R. Savithiri, “Spotlight on focus groups,” *Canadian Family Physician*, vol. 55, no. 2, pp. 218–219, 2009, PMID: PMC2642503. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2642503/>.
- [23] R. A. Krueger, *Focus Groups: A Practical Guide for Applied Research*. Sage Publications, 2014.
- [24] GeeksForGeeks. “Component-based architecture – system design.” (Jul. 2024), [Online]. Available: <https://www.geeksforgeeks.org/component-based-architecture-system-design/> visited 29-4-2025.
- [25] StatCounter. “Market share of leading desktop search engines worldwide from january 2015 to march 2025 [graph].” (Apr. 2025), [Online]. Available: <https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/> visited 29-4-2025.
- [26] Netlify. “Netlify security checklist.” (Feb. 2025), [Online]. Available: <https://docs.netlify.com/platform/security-checklist/> visited 28-4-2025.
- [27] European Parliament and Council of the European Union. “Regulation (EU) 2016/679 of the European Parliament and of the Council,” of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). (Apr. 27, 2016), [Online]. Available: <https://data.europa.eu/eli/reg/2016/679/oj> visited 25-3-2025.

A

Focus Group Tasks, Questions, and Responses

A.1 Tasks and Questions for Focus Group Meeting 1

A.1.1 Tasks for testing Korp

Testa Korp i 5 min. Exempelsökningar: sök på “Gud” i Göteborgs-Posten, sök på substantivet “får”

A.1.2 Questions related to Korp - younger focus group

1. Vad uppfattar du att språkverktyget kan användas till?
2. Påminner språkverktyget dig om något du använt tidigare?
3. Blir du lockad av att använda språkverktyget?
4. Vad skulle du använda språkverktyget till?
5. Tycker du att något är otydligt i språkverktyget?
6. Vilka funktioner tycker du är relevanta för dig? Saknas något?

A.1.3 Questions related to Korp - older focus group

1. Vad uppfattar du att språkverktyget kan användas till?
2. Påminner språkverktyget dig om något du använt tidigare?
3. Blir du lockad av att använda språkverktyget?
4. Vad skulle du använda språkverktyget till?
5. Vilka funktioner är relevanta för dig?
6. Tycker du att något är otydligt i språkverktyget?

A.1.4 Questions related to initial design prototype of Korpi - younger focus group

1. Påminner språkverktyget dig om något du använt tidigare?
2. Blir du lockad av att använda språkverktyget?
3. Vad skulle du använda språkverktyget till?
4. Tycker du att något är otydligt i språkverktyget?
5. Saknas något som finns i Korp?

A.1.5 Questions related to initial design prototype of Korpi - older focus group

1. Påminner språkverktyget dig om något du använt tidigare?
2. Blir du lockad av att använda språkverktyget?
3. Tycker du att något är otydligt i språkverktyget?
4. Saknas något som finns i Korp?

A.1.6 Questions related to navigation and information presentation for Korp and Korpi

1. Hur känns det att navigera sidorna och funktionerna (flow)?
2. Vad tycker du om hur resultaten presenteras?
3. Övriga tankar

A.2 Responses from Focus Group Meeting 1

Table A.1: First focus group meeting: Younger focus group evaluating Korp.

First focus group meeting: Younger focus group evaluating Korp		
Problem areas	Evaluation aspects	High-level results
Design for user-friendliness	General user-friendliness	- Korp was not considered user-friendly. - Korp was considered slow.
	Navigation	-
	Information presentation	- It was considered unclear how search results are prioritized.
	Visual design	- Korp was considered dull.
Functionality for usefulness	Interesting	- It was considered interesting to see how words are used. - It was considered interesting to be able to search for different forms of a word. - It was considered interesting to see statistics. - It was considered interesting to be able to search among many, large corpora. - It was not considered interesting to conduct investigations of a more linguistic/grammatical kind.
	Useful	- Korp was not considered useful for the focus group participants. - Tips for how to use Korp were requested.

Table A.2: First focus group meeting: Older focus group evaluating Korp.

First focus group meeting: Older focus group evaluating Korp		
Problem areas	Evaluation aspects	High-level results
Design for user-friendliness	General user-friendliness	- Korp was not considered user-friendly. - Korp was considered difficult to understand. - Korp was considered slow when searching in many corpora.
	Navigation	-
	Information presentation	- Korp was considered having too much information. - The search results were not considered to be presented in a good way. Too much “just text” and need for horizontal scrolling.
	Visual design	- Korp was not considered to have a visual design that attracts users.
Functionality for usefulness	Interesting	- It was considered interesting to see the meaning of words. - It was considered interesting to see examples how difficult words are used. - It was especially considered interesting to be able to see meaning and use of new words.
	Useful	- Korp was not considered useful for the focus group participants. - Google is considered just as useful as Korp, e.g. for looking for meaning of words and examples.

A.3 Tasks and Questions for Focus Group Meeting 2

A.3.1 Tasks for testing the first version of Korpi

Testa prototypen och jämför med Korp i 15 min. Exempelsökningar: sök på “Ful”, sök först på “insulin” i medicinska texter, sedan i SVT-Nyheter.

A.3.2 Questions related to the first version of Korpi

1. Blir du mer lockad av att använda språkverktyget nu än senast?
2. Är det lätt att navigera webbsidan?
3. Hur tycker du korpus valet ska presenteras? (inte design)
4. Vad tycker du om hur resultat är presenterade?
5. Är det viktigt med korpus-titlar i resultatpanelen?
6. Vad tycker du om inställningar?

7. Vilka funktioner saknas?
8. Vad är du nöjd med i nya Korpi / missnöjd med?

A.3.3 Questions related to the new design prototype of Korpi - only asked to younger focus group

1. Utifrån första mötet, känner du att vi har implementerat det som du önskade?
2. Vad saknas?

A.3.4 Figma Design Proposal For Extended Search

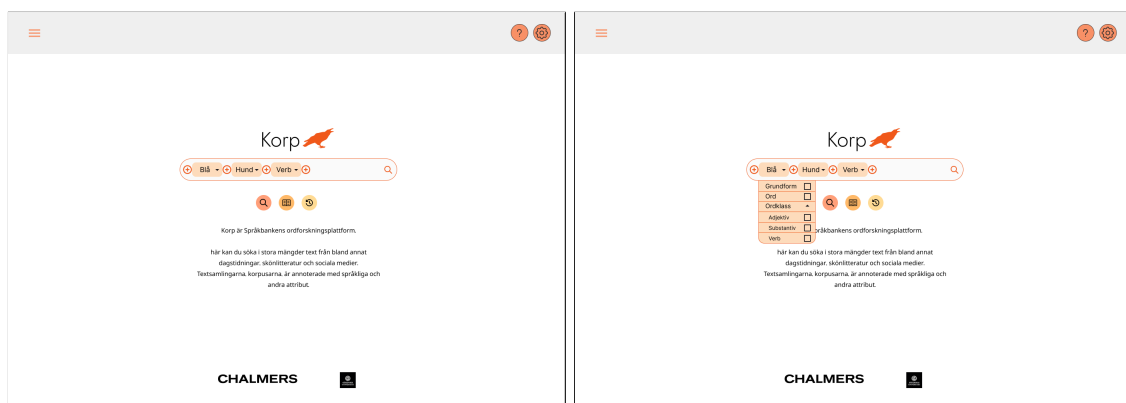


Figure A.1: Figma design proposal for extended search in Korpi

A.3.5 Questions related to the statistics demo - only asked to younger focus group

1. Är detta det ni menade med att se mer statistik?
2. Hur tycker ni det presenteras?
3. Var vill ni se statistiken? (ex: som en egen flik, vid sidan av resultatpanelen, om man klickar på knappar och filter (att den är mer gömd)?

A.4 Responses From Focus Group Meeting 2

Table A.3: Second focus group meeting: Younger focus group evaluating statistics demo.

Second focus group meeting: Younger focus group evaluating statistics demo		
Problem areas	Evaluation aspects	High-level results
Design for user-friendliness	General user-friendliness	- The statistics was difficult to understand and interpret.
	Navigation	-
	Information presentation	- More graphs are requested. - Unclear wording. Not clear what a lemma is.
Functionality for usefulness	Visual design	-
	Interesting	- More statistics is requested (number of hits, change over time, comparison between different types of words).
	Useful	-

A.5 Tasks and Questions for Focus Group Meeting 3

A.5.1 Tasks for testing final Korpi

Uppgifter att utföra under 15 minuter:

1. Gå igenom hjälpfunktionen.
2. Tryck på en exempelsökning.
3. Sök på ordet "katt" i korpusarna SVT nyheter 2007 och SVT nyheter 2010. Välj att visa fem resultat per korpus.
4. Sök på ordet "och" och alla former av ordet "katt" (samma korpusar och antal resultat per korpus som ovan). Sök på alla adverb (samma korpusar och antal resultat per korpus som ovan).
5. Gör om första sökningen "katt" igen, genom att använda historikfunktionen.
6. Ändra till dark mode. Säg till när ni kommit hit!
7. Sök på "trääd" och ta fram statistik för detta ord.
8. Gå till ursprungskällan för ett av dina sökresultat då du sökt på "trääd".

A.5.2 Questions related to final Korpi

1. Vad är det viktigaste ni vill framföra?
 - (a) Positivt
 - (b) Negativt

B

Focus Group Survey

Slutvärderingsformulär - Korpi

Namn: _____

(1 = Inte bra, 5 = Bra)
Markera en bubbla per rad.

Skala:	1	2	3	4	5
Användbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Responsiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Snabb (sökningar)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Navigation (att hitta på sidan)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Design	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Intressant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tydlig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Upplevelse under användning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Informationspresentation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statistikpresentation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sökresultatpresentation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Öppna frågor

Går det att förstå hur man öppnar och stänger olika fönster?

Vilken information är svårare att förstå på webbsidan?

Vilken information är lättare att förstå på webbsidan?

Vad tycker du om hur sökresultaten presenteras?

Vad tycks om designen?

Vad tycks om navigationen av sidan?

Vad tycks om lexikonbeskrivningen av ord och statistik av ord i resultat?

Vad upplever du när du använder sidan?

Figure B.2: Focus group survey page 2

Är sidan intressant? Om inte, vad saknas för att den ska kunna bli det?

Mobil app:

Är mobilversionen lika bra som datorversionen? Stack något ut som bra eller dåligt?

Korpi i sin helhet:

Skriv dina slutgiltiga åsikter och tankar om Korpi:

Vad är bra?

Vad är dåligt?

Övriga tankar:

Korpi i sin helhet

○ ○ ○ ○ ○

Figure B.3: Focus group survey page 3

C

Focus Group Material for Discussion Regarding Functionality

10

Intressant funktionalitet

Övergripande funktionalitet:

1. Söka på ord i ett brett sortiment av korpusar och genrer (för allmänheten, för personer som lär sig svenska och för professionella)
2. Se statistik, både tabeller och grafer
 - a. Hur ofta ord används och i olika korpusar
 - b. Hur användning har förändrats över tid och i olika korpusar (ex: nyord eller användning av t.ex. "vore" eller "våran")
 - c. Jämförelse av ord (t.ex. om "de" är vanligare än "dom")
3. Få tips om hur språkverktyget kan användas
4. Ha tillgång till en användarguide
5. Fler lingvistikfunktioner?

Detaljerad funktionalitet:

1. Kunna söka på olika ordformer (ordklasser och böjningar)
2. Kunna söka på flera ord
3. Kunna välja att söka bland många olika (stora) korpusar
4. Kunna välja att söka bland grupperingar av korpusar
5. Kunna förstå hur sökresultat prioriteras
6. Kunna filtrera sökresultat
7. Kunna se länk till fulltext för ett sökresultat
8. Kunna få information om sökordet (t.ex. betydelse eller synonymer)
9. Röststyrning

**Har vi förstått er rätt?
Är det tillräcklig funktionalitet?
Vilka tycker ni är viktigast?**

Figure C.1: Functionality that was discussed with the older focus group.

10

Intressant funktionalitet

Övergripande funktionalitet:

1. Söka på ord i ett brett sortiment av korpusar och genrer (för allmänheten, för personer som lär sig svenska och för professionella)
2. Se statistik
 - a. Hur ofta ord används
 - b. Hur användning har förändrats över tid (ex: nyord)
 - c. Jämförelse av ord
3. Få tips om hur språkverktyget kan användas
4. Ha tillgång till en användarguide
5. Synonymer?
6. Fler lingvistikfunktioner?

Detaljerad funktionalitet:

1. Kunna söka på olika ordformer
2. Kunna söka på flera ord
3. Kunna välja att söka bland många olika (stora) korpusar
4. Kunna välja att söka bland grupperingar av korpusar
5. Kunna förstå hur sökresultat prioriteras
6. Kunna filtrera sökresultat
7. Kunna nå källan som en resultatmening kommer från
8. Röststyrning

**Har vi förstått er rätt?
Är det tillräcklig funktionalitet?
Vilka tycker ni är viktigast?**



Figure C.2: Functionality that was discussed with the younger focus group.

D

Design

D.1 Concept Design

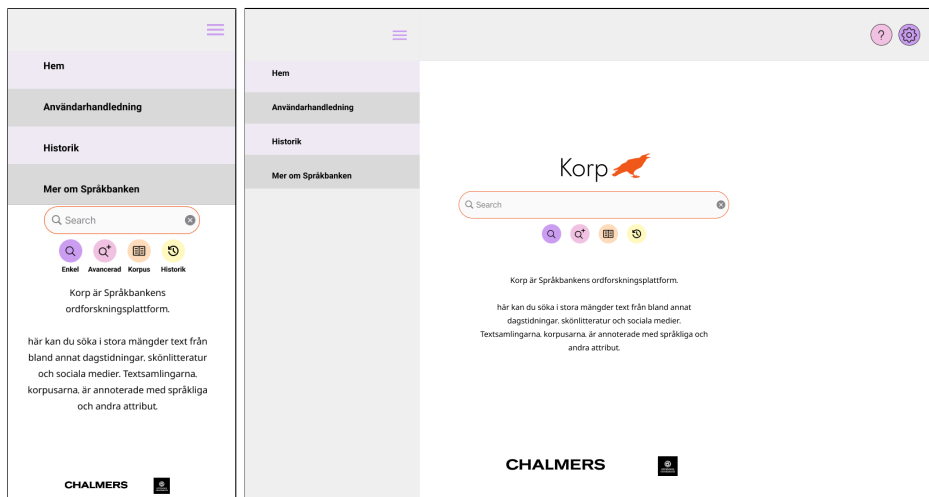


Figure D.1: First prototype of Korpi's side menu for mobile and desktop

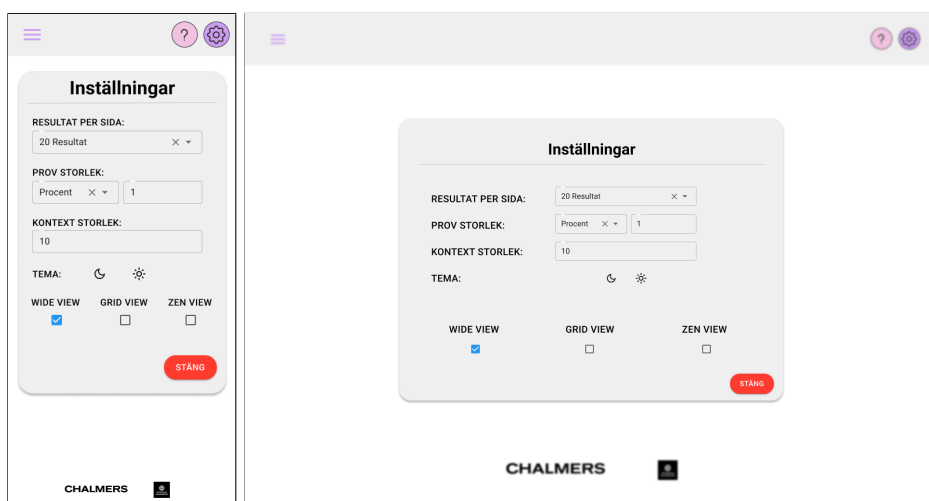


Figure D.2: First prototype of Korpi's settings page for mobile and desktop

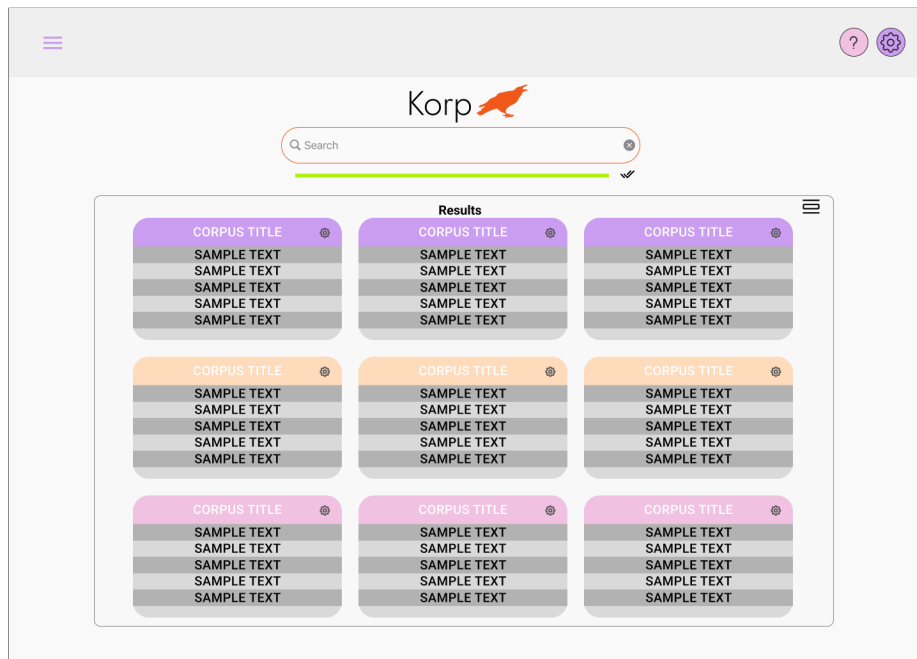


Figure D.3: First prototype of Korpi's calendar result page

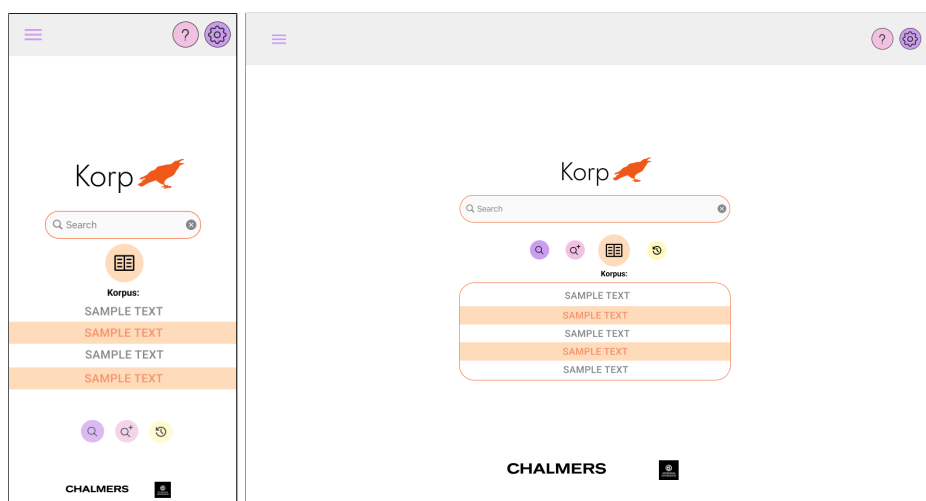


Figure D.4: First prototype of Korpi's choose corpus page for mobile and desktop

D.2 Detailing Phase Design

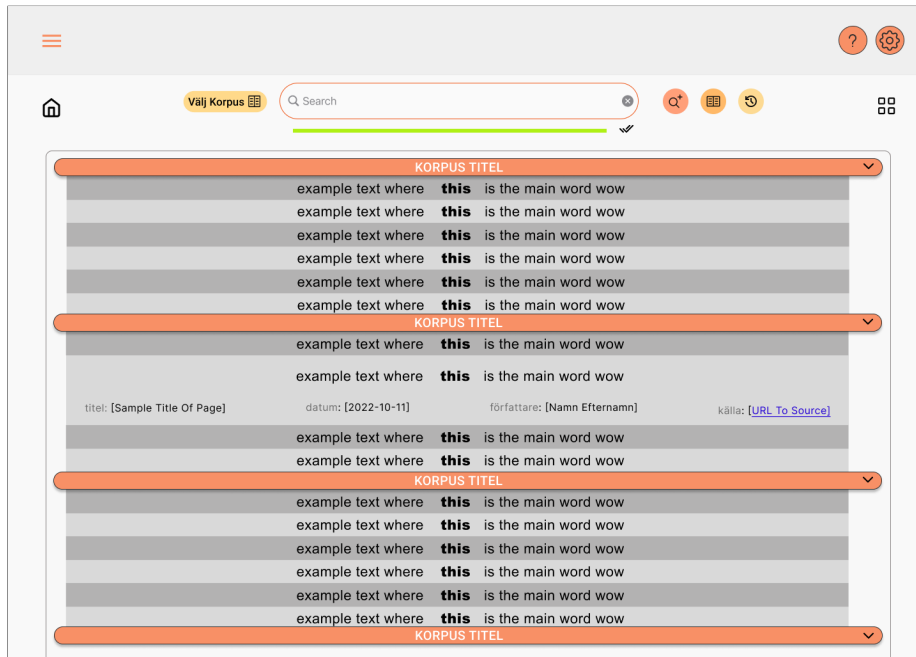


Figure D.5: First design of Korpi's result page when search result is expanded

D.3 Future Improvements

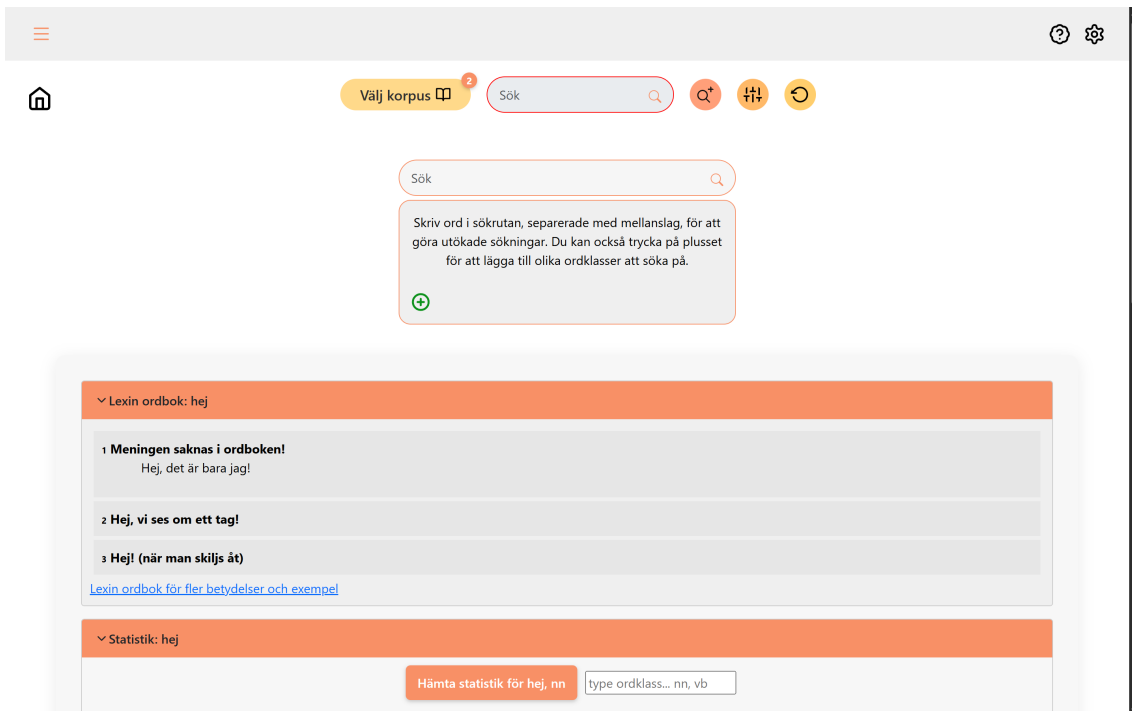


Figure D.6: Extended Search when used in Result Page