

Jonna Jantunen

# **“WE INVEST IN THE INTELLIGIBILITY OF LANGUAGE”**

Linguistic Accessibility of the English Kanta.fi Website

# ABSTRACT

Jonna Jantunen: "We Invest in the Intelligibility of Language": Linguistic Accessibility of the English Kanta.fi Website  
Master's Thesis  
Tampere University  
Master's Programme in Multilingual Communication and Translation Studies, English translation and interpreting  
December 2025

---

Accessibility is about providing products and services that are available, usable, and understandable to as many users as possible. At its core, it is about ensuring equality and, ultimately, the realization of human rights, particularly for persons with disabilities. Lately, the accessibility of digital services has gained additional attention, particularly since the EU published the Web Accessibility Directive (2016/2102) in 2016. However, this attention on accessibility has focused primarily on the technical aspects, despite the fact that access to information is a fundamental human right that cannot be achieved with technical solutions alone; language plays a central role in ensuring the accessibility of information.

The goal of this thesis is to understand how linguistically accessible the English kanta.fi website is. Linguistic accessibility is defined as the extent to which the linguistic material on a website can be understood by everyone, regardless of their abilities or personal traits, and can thus be used to achieve the desired goals. The kanta.fi website provides information about Kanta Services, a set of nationwide health care and social welfare services used to store and exchange data about patients. Due to their importance in the Finnish health care and social welfare systems, it is crucial that information about the services is available to everyone – all users have the right to know what data is stored about them and how it is used. Additionally, since over 10% of the Finnish population has a foreign first language, and English is the only foreign language the website is offered in, the accessibility of the English translation is particularly important. The right to health care does not depend on one's language skills, and neither should the ability to get information about the available services.

The research data consists of ten randomly selected texts from the kanta.fi website. The research question was approached with the method of heuristic evaluation: each text was compared to a set of heuristics – the guidelines for linguistic accessibility – created for this thesis. The results of the evaluation were then analyzed to form an understanding of the types of issues present and the overall state of linguistic accessibility.

The study shows that despite their claim to "invest in the intelligibility of language", the English kanta.fi website is not linguistically accessible to all users. While the research data comprised of only ten texts, the findings across the texts were consistent, indicating the results are likely applicable to the entire website. The most prevalent accessibility issues were related to consistency, ambiguity, and the amount of information provided in the texts. Some of these findings might be explained by the translation process and could be improved by modifications to the process.

The results of this thesis provide insight into the linguistic accessibility issues present on the kanta.fi site and could be used to improve accessibility of the English texts on the site. Further research could be conducted to extend the analysis to linguistic elements outside the main texts, such as user interface texts. Additionally, involving real users could provide fruitful information to further improve accessibility.

Keywords: linguistic accessibility, accessibility, accessible language, Kanta Services, administrative communication, heuristic evaluation

The originality of this thesis has been checked using the Turnitin Originality Check service.

# TIIVISTELMÄ

Jonna Jantunen: "We Invest in the Intelligibility of Language": Englanninkielisen kanta.fi-sivuston kielellinen saavutettavuus  
Pro gradu -tutkielma  
Tampereen yliopisto  
Monikielisen viestinnän ja käännöstieteen maisteriohjelma, englannin kääntämisen ja tulkkauksen opintosuunta  
Joulukuu 2025

---

Saavutettavuuden tavoitteena on varmistaa tuotteiden ja palveluiden saatavuus, käytettävyyys ja ymmärrettävyys mahdollisimman monelle käyttäjälle. Pohjimmiltaan kyse on erityisesti vammaisten henkilöiden yhdenvertaisuuden ja ihmisoikeuksien toteutumisen turvaamisesta. Viime vuosina digitaalisten palveluiden saavutettavuus on ollut pinnalla, erityisesti vuonna 2016 julkaistun EU:n saavutettavuusdirektiivin (2016/2102) ansiosta. Tämä huomio on kuitenkin keskittynyt pääasiassa teknisiin ratkaisuihin huolimatta siitä, että oikeus tietoon on perustavanlaatuinen ihmisoikeus, jonka toteutumista ei voida turvata pelkästään teknisillä ratkaisuilla; kielellä on merkittävä rooli tiedon saavutettavuuden varmistamisessa.

Tämän tutkielman tavoitteena on muodostaa käsitys englanninkielisen kanta.fi-verkkosivuston kielellisestä saavutettavuudesta, eli siitä, kuinka hyvin kuka tahansa verkkosivuston käyttäjä, riippumatta hänen henkilökohtaisista ominaisuuksistaan tai kyvyistään, voi ymmärtää sivuston kielellisen materiaalin ja siten hyödyntää sitä tavoitteidensa saavuttamisessa. Kanta.fi tarjoaa tietoa Kanta-palveluista, jotka ovat kokoelma potilaan tietojen säilömiseen ja jakamiseen tarkoitettuja valtakunnallisia sosiaali- ja terveyspalveluita. Kanta-palvelut ovat keskeinen osa suomalaista sosiaali- ja terveydenhuoltojärjestelmää, minkä vuoksi olisikin tärkeää varmistaa, että tieto palveluista on kaikkien saatavilla – kaikilla käyttäjillä on oikeus tietää, mitä tietoja heistä tallennetaan ja miten sitä käytetään. Englanninkielisen sivuston kielellinen saavutettavuus on lisäksi erityisen tärkeää, sillä yli 10 % Suomen väestöstä on vieraskielisiä, ja englanti on ainoa vieras kieli, jolla sivusto on saatavilla. Oikeus terveydenhuoltoon ei riipu yksilön kielitaidosta, joten jokaisella pitäisi myös olla mahdollisuus saada tietoa saatavilla olevista palveluista.

Tutkimusaineisto koostuu kymmenestä kanta.fi-sivustolta satunnaisesti valitusta tekstistä. Lähestyin tutkimuskysymystä heuristisen arvioinnin menetelmällä: vertailin jokaista tekstiä tätä tutkielmaa varten luotuihin heuristiikkoihin, eli kielellisen saavutettavuuden ohjeisiin. Lopuksi analysoin löydöksiä tunnistaakseni tyypillisimmät ongelmat ja tehdäkseni päätelmiä aineiston kielellisestä saavutettavuudesta.

Kanta.fi:n väitteestä "we invest in the intelligibility of language" huolimatta tutkimuksen tulokset osoittavat, että englanninkielinen sivusto ei ole kielellisesti saavutettava kaikille käyttäjille. Vaikka tutkimusaineisto koostui vain kymmenestä tekstistä, löydökset olivat johdonmukaisia läpi tekstien, joten tulosten voi katsoa kuvaavan koko sivuston kielellistä saavutettavuutta. Yleisimmät ongelmat liittyivät johdonmukaisuuteen, monitulkintaisuuteen sekä annetun tiedon määrään. Osa löydöksistä saattaa selittyä käännösprosessilla, jonka muokkaaminen voisikin parantaa kielellistä saavutettavuutta.

Tutkimuksen johtopäätökset antavat käsityksen kanta.fi-sivuston kielellisen saavutettavuuden ongelmista, ja niitä voisikin käyttää englanninkielisen käännöksen saavutettavuuden parantamisen tukena. Tutkimusasetelmaa voisi jatkossa laajentaa kattamaan myös varsinaisen sisällön ulkopuoliset kielelliset elementit kuten käyttöliittymätekstit. Myös loppukäyttäjien mukaan ottaminen voisi tarjota hedelmällistä tietoa kielellisen saavutettavuuden parantamiseksi.

Avainsanat: kielellinen saavutettavuus, saavutettavuus, saavutettava kieli, Kanta-palvelut, viranomaisviestintä, heuristinen arviointi

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

# USE OF AI IN THESIS

I have utilised AI tools in my thesis:

- No
- Yes

# TABLE OF CONTENTS

- 1 INTRODUCTION ..... 1
- 2 LEGISLATION AND GUIDANCE..... 7
  - 2.1 Equality and Access to Information..... 7
  - 2.2 Web Accessibility Legislation ..... 8
  - 2.3 Web Content Accessibility Guidelines (WCAG) ..... 10
- 3 WEB ACCESSIBILITY ..... 13
  - 3.1 Accessibility Is for Everyone ..... 14
  - 3.2 Usability ..... 15
- 4 LINGUISTIC ACCESSIBILITY ..... 17
  - 4.1 Understandability ..... 18
  - 4.2 Language Variants on a Spectrum of Accessibility ..... 19
  - 4.3 Linguistic Accessibility Ensures Access to Information ..... 21
  - 4.4 Linguistic Accessibility Benefits Everyone ..... 22
  - 4.5 Linguistic Accessibility for Language Minorities ..... 24
- 5 RESEARCH DATA ..... 28
  - 5.1 Administrative Texts ..... 28
  - 5.2 Administrative Texts in Health Care ..... 29
  - 5.3 Kanta.fi ..... 31
  - 5.4 Research Material ..... 34
- 6 METHOD ..... 36
  - 6.1 Heuristic Evaluation ..... 36
  - 6.2 Guidelines for Linguistically Accessible Texts ..... 38

6.2.1 Background .....	38
6.2.2 Modifications.....	39
6.2.3 Linguistic Accessibility Guidelines.....	41
6.3 Research Method.....	43
7 ANALYSIS .....	44
7.1 Grammar and Spelling .....	44
7.1.1 Consistent Grammar.....	45
7.1.2 Ambiguity.....	46
7.1.3 Grammatical and Sentence Structures.....	48
7.1.4 Spelling.....	50
7.2 Vocabulary .....	50
7.2.1 Consistent Vocabulary .....	51
7.2.2 Abstractness.....	53
7.2.3 Word Choice.....	54
7.3 Structure .....	55
7.3.1 High-Level Structure .....	56
7.3.2 Headings .....	58
7.3.3 Order of Information .....	60
7.3.4 Cohesion.....	61
7.4 Information Density and Clarity .....	62
7.4.1 Amount of Information .....	63
7.4.2 Links.....	66
7.4.3 Non-Text Content.....	68
7.5 Summary of Findings and the Overall Linguistic Accessibility ...	71
8 CONCLUSIONS.....	75
WORKS CITED .....	80
Research Data.....	80
References .....	80
SUOMENKIELINEN LYHENNELMÄ.....	I

# 1 INTRODUCTION

Accessibility, particularly that of digital services, has received recurring attention in public discussions for a while now, and the topic maintains prominence as new and renewed legislation and guidelines are published. In Finland, web accessibility first gained significant traction in 2016 with the publication of the EU Web Accessibility Directive (2016/2102) that mandates all public sector bodies' websites and mobile applications to adhere to accessibility requirements based on the Web Content Accessibility Guidelines (WCAG) 2.1. In 2019, the European Accessibility Act (2019/882) extended the requirements to apply to more products and services, this time involving private organizations as well. These directives, alongside the national legislation enforcing the requirements set forth by them, provide concrete steps for organizations to take to ensure accessibility in the digital realm.

Accessibility is about creating products and services that are usable and understandable to all, regardless of their abilities and possible barriers to function (ISO 2018). In the web, this means ensuring websites are technically accessible to a variety of users and assistive technologies and creating content that can be understood by users with varying abilities. In short, accessibility is about ensuring equality, particularly for persons with disabilities. Although the web accessibility legislation is relatively recent, this fundamental idea of equality has been prevalent in both national and global legislation for decades, as is evident from the first article of the Universal Declaration of Human Rights from 1948: “All human beings are born free and equal in dignity and rights” (UN, n.d.).

One of the key factors in ensuring equality is the fundamental human right of access to information; the right to have access to and understand information provides the foundation necessary to live and function in a society (Hirvonen et al. 2020, 19–20; Koponen and Nurminen 2020, 306). The fulfilment of this right is often dependent on language, yet, despite its apparent importance, linguistic accessibility has been left to little attention in legislation. In Finland, the Administrative Procedure Act (434/2003) requires public authorities to use appropriate, clear, and understandable language, but provides little information on how that should be achieved. Similarly, the current version of the WCAG does not provide any guidelines for language use, despite understandability being one of its four main principles (W3C 2024d; Leskelä and Uotila 2020, 233). Nevertheless, linguistic accessibility has been

addressed more lately: the European Accessibility Act (2019/882, 31) contains some requirements for language use, as does the draft version of the WCAG 3, which is projected to develop into a standard in a few years (W3C 2023; 2025b). This increased attention on language use and its effects on accessibility seems to indicate that linguistic accessibility might continue to grow into a more prominent topic and a clearer legal requirement, simultaneously increasing the need for information and research – a need I aim to in part answer with this thesis.

The purpose of this thesis is to understand how linguistically accessible the English kanta.fi website for citizens is. I consider linguistic accessibility of the written material on the website and evaluate how easily users with varying language and cognitive skills can understand it and use it to reach their desired goals. This idea of using texts for a specific purpose is central in my definition of linguistic accessibility – texts are viewed as tools, and their accessibility depends on whether they can be successfully used for their intended purpose (Suominen 2019, 61; Suojanen et al. 2015, 7). This viewpoint works particularly well for the kind of informative and instructional texts that comprise my research data, as those are generally read to understand or do something, that is, to reach a specific goal. This is also the reason I have decided to use the term *user* rather than *reader* – to emphasize the reasons for reading these texts and thus the role of language in ensuring the users' goals are met.

The kanta.fi website provides information about and guidance on the Kanta Services, a nationwide collection of digital services for managing and storing social welfare and health care data (Kanta Services 2024d). The choice to focus on kanta.fi in this thesis stems from the widespread use of the Kanta Services and their undeniable importance in the Finnish health care system, due to which the accessibility of information about these services is also essential. Whereas most of these services are aimed at professionals, the data in the services belongs to the citizens, and it is crucial that they have the information necessary to understand and affect why and how their data are stored and used (Kanta Services 2025a; 2025d). Additionally, the context of health care adds to the relevance of linguistic accessibility: health care and particularly sickness can be stressful, and stress has been shown to increase the need for accessible language even for those who do not normally need it (Leskelä 2019, 109; WHO 2024).

In this thesis, I focus on the Citizens section of kanta.fi. The information it provides is aimed at the general public, resulting in significant variance in the intended users and their abilities, further emphasizing the need for linguistic accessibility (Pasquier 2018, 453). This variance is

particularly significant amongst the users of the English website: since English is the first language of only about 6% of the over 600,000 Finnish residents with a foreign first language (Statistics Finland 2025), the majority of those using kanta.fi in English are not native, or necessarily even fluent, speakers of the language (Kela 2025, 8). In addition, particularly newer migrants might be unfamiliar with the Finnish health care system and available services, and navigating that system can be easier when the information about it is available in an accessible format (WHO 2024). It is also worth noting that some of the texts on kanta.fi provide information and instructions related to the MyKanta service that citizens can use to view and manage their data saved in the Kanta Services. It is an important part of the Finnish health care system and allows citizens to independently handle some of their health-care-related matters (Kanta Services 2025d). Yet, despite its relevance, the only part of MyKanta available in English is its name – the tool itself has not been translated. Due to this, the accessibility of the English texts covering the use of MyKanta is of particular importance, as those can be the only way for some users to use the service.

To understand how linguistically accessible the kanta.fi website is, it is important to investigate what kind of linguistic accessibility issues are present on the site and whether any issues are more prominent than others. To answer these questions, I analyze ten randomly selected texts using the methods of heuristic evaluation: comparing the texts to a list of linguistic accessibility guidelines and looking for occasions where those guidelines are violated. Heuristic evaluation is an easy, quick, and cost-efficient, yet effective, method for finding problems in the evaluated product (Nielsen 1994, 160), which is why I chose to use it in this thesis. The method stems from the field of usability research, but today it is a common tool used in various fields and for various products, including texts. Heuristic evaluation results in a list of issues with information on which guideline they violate, and the method is thus often used to identify and fix issues in a product. However, it also suits my purposes well: I use the list to understand the type and frequency of accessibility issues present in the texts and to draw conclusions about the state of linguistic accessibility on the website. I also consider whether there are significant differences between the texts, as the amount and type of variance is important in trying to understand the applicability of the results to the entire website.

The guidelines used for the heuristic evaluation are a revised and extended version of the criteria for evaluating the linguistic accessibility of a text that I created in my bachelor's thesis (Jantunen 2020, 11–13). While it is typical in heuristic evaluation to customize the heuristics

for each use case, these guidelines are not tailored specifically for this thesis. Instead, they are designed to work as both instructions and heuristics outside this thesis as well – ideally, these guidelines could be used more broadly to support writers in creating accessible texts as well as to evaluate the linguistic accessibility of different texts. However, the guidelines are not the main focus of this thesis, and I do not evaluate or analyze their suitability for anything besides the research data, though that could be an interesting topic for future research.

Linguistic accessibility as such has not received significant attention in research, which is likely at least partially due to the technical focus of the related legislation and guidelines. Similarly, the development of online communication in general, particularly in the public sector, has been focusing on technical usability and user interface design (Suominen 2019, 15). However, extensive research has been done on topics closely related to linguistic accessibility, such as readability, understandability, and the various textual factors affecting comprehension, and the linguistic accessibility guidelines in this thesis are largely based on these previous studies. Understandability, in particular, is often used interchangeably with linguistic accessibility, and the difference between the two is primarily in the viewpoint; with this thesis, I want to focus the attention to the ideas of access and equality that come with the term *accessibility*, rather than focusing purely on the understandability of the texts.

One form of linguistic accessibility that has been studied more is simplified language: an artificially created language variant that is aimed specifically at those with trouble understanding standard language and that uses only the clearest and easiest features of language to ensure understandability (Leskelä 2019, 93; Leskelä and Uotila 2020, 227). For example, in a paper related closely to the topic of this thesis, Sarah Ahrens (2020) studied administrative texts written in simplified German, focusing specifically on second language learners as the target group. While simplified language is undeniably important and should be used alongside accessible standard language, I wanted to focus on the accessibility of standard language, the variant used most often on public authorities' websites.

Another similarity between my thesis and Ahrens' (2020) paper is the focus on administrative communication, that is, communication from public authorities to the general public. While it has not been studied much from the perspective of linguistic accessibility, other aspects of administrative communication have been researched from various standpoints. For instance, in her doctoral dissertation, Suominen (2019) focused on improving the usability of administrative texts in municipalities and, based on interviews with the writers of these texts, created heuristics

for writing online municipal communication. Suominen focused only on Finnish texts, which is fairly typical: since public authorities are legally only required to provide information in the national languages of Finnish and Swedish, administrative communication research conducted in Finland tends to focus on those. While understandable, this choice leaves out the important audience of foreign language speakers. As the number of foreign language speakers in Finland is growing yearly (Statistics Finland 2025), it is important to bring attention to the understandability, usability, and accessibility of texts in other languages as well.

Some research has also been done on linguistic accessibility, though that is mostly limited to theses, such as Kervinen's master's thesis (2020) on digital service content creators' views on linguistic accessibility and its improvement, though her perspective on the topic is considerably different from mine. From a more similar perspective, master's theses related to linguistic accessibility have been written by Invenius (2024), who evaluated the cognitive and particularly linguistic accessibility of the city of Vantaa's websites, and Laakso (2021), who studied the usability and equality of the English patient instructions on Finnish public hospitals' websites. Despite not directly mentioning linguistic accessibility, Laakso's thesis is particularly similar to mine – besides the slightly different perspective on the topic, the most significant difference is her comparative analysis of the language versions. This is a common research method in translation studies, and, for example, the majority of administrative communication research that does cover translation to foreign languages tends to focus on this as well. Although this type of information is valuable, I have intentionally decided to not compare the translations to their source texts, as I want to explore the accessibility of the English texts as they are – those using the texts are likely not able to compare them either. As far as I know, this kind of research on the linguistic accessibility of a specific service or website is limited.

Finally, the kanta.fi website itself has not been researched much, nor has the linguistic accessibility of the Kanta Services. Some studies have been done on the Kanta Services – see, for example, Palojoki and Vuokko (2025) for an assessment of the services and the development of digitalization in health care – but the website has not been the focus in any of these studies. Additionally, the linguistic aspect has received little attention even in studies focused on the MyKanta service aimed at the general public; see, for example, Eriksson-Backa et al. (2021) for a study on enablers for and barriers to using MyKanta.

In the following chapter, I introduce the legislation governing equality and accessibility in Finland. Chapters 3 and 4 cover the concepts of web accessibility and linguistic accessibility,

providing more information on why they matter and who they benefit. In chapter 5, I introduce the research data, as well as explore the relevance of linguistic accessibility in health-care-related administrative communication. Chapter 6 is devoted to the research method, first introducing heuristic evaluation and then delving into details about the guidelines and how I performed the analysis – the results of which are presented in chapter 7. Finally, I present the conclusions of this study in chapter 8.

## **2 LEGISLATION AND GUIDANCE**

### **2.1 Equality and Access to Information**

At its core, accessibility is about equality, diversity, and, ultimately, the realization of human rights. These ideas are not new to the Finnish legislation, as equality is demanded in various laws, starting with the Constitution of Finland (731/1999), which defines equality as one of the fundamental human rights – everyone is equal before the law, and no one should be treated differently or put in an unequal position based on their personal traits such as sex, age, ethnicity, opinion, language, or disability. To further promote equality, the Non-Discrimination Act (1325/2014) was enacted in 2014, strengthening the protection of the law for those who have experienced discrimination and demanding reasonable adjustments to ensure equality. Furthermore, Finland has ratified the UN Convention on the Rights of Persons with Disabilities (CRPD), which aims to “promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities, and to promote respect for their inherent dignity” (UN 2006, 4).

In a way, these demands for equality already include the demand for accessibility, as the lack of accessibility does place some people in an unequal position. For instance, a building entrance with stairs is inherently inaccessible to wheelchair users, and not providing them with alternative, accessible solutions (such as a ramp) prevents them from entering the building and using the services within it. Although the inaccessible entrance is likely not a deliberate decision to exclude specific groups of people, they are nevertheless being denied entrance based on a personal trait and thus discriminated against; indirect discrimination is still a violation of the Non-Discrimination Act (1325/2014). It is also important to note that measures taken to ensure de facto equality for persons with disabilities are not considered discriminatory towards those without disabilities (UN 2006, 7; Non-Discrimination Act 1325/2014). After all, equality does not mean that everyone is treated the same.

One important factor in ensuring equality is access to information. The CRPD, for example, states that in order to “enable persons with disabilities to live independently and participate in all aspects of life”, they need equal access to all aspects of life, including information and communication, as well as information and communication technologies and systems (UN 2006, 10). In addition to enabling equality, access to information in itself is a fundamental

human right, and it is defined as such in various conventions and laws, such as the UN Universal Declaration of Human Rights from 1948 (UN, n.d.). Similarly, the Constitution of Finland (731/1999) states that everyone has the right to express, publish, and receive information. Additionally, everyone has the right to participate in societal actions and decision-making in matters concerning themselves, and it is one of the main tasks of public authorities in Finland to increase individuals' possibilities to do so, regardless of their personal traits or abilities (Horowitz et al. 2019, 128–29). For an individual to be able to effectively participate in societal actions and decision-making, they must have access to the relevant information (149). Access to information, therefore, plays a significant role in the fulfillment of other constitutional rights.

However, access to information alone is not enough. For the information to be useful, it must also be understandable, that is, linguistically accessible (the relationship between understandability and linguistic accessibility is discussed further in chapter 4.1). Using accessible language to deliver information ensures everyone is able to receive the information and be treated equally, and thus have their constitutional rights fulfilled. Although linguistic accessibility itself is not directly mentioned in legislation, the Administrative Procedure Act (434/2003) does require public authorities to use appropriate, clear, and understandable language. However, the requirement does not clearly define what understandable language means, which can lead to varying interpretations and levels of understandability. The preamble for the law does provide a little clarification, as it mentions unambiguity, consistency, and expressing matters in such a way that the customer gets enough information and can be expected to unequivocally understand the content (HE 72/2002). Additionally, the Administrative Procedure Act (434/2003) requires authorities to ensure translation and interpretation services in official matters to anyone who cannot otherwise be understood, including both speakers of foreign languages and those who cannot be understood due to an illness or a disability.

## **2.2 Web Accessibility Legislation**

So, the implicit idea of accessibility has existed in national and international legislation for a long time. Besides the examples introduced in the previous chapter, different aspects of accessibility are, more or less indirectly, considered in various Finnish laws regarding, for example, postal services and audiovisual services (Åkermarck 2020, 35). The idea of accessibility is thus not a new one; it is just more explicit today, particularly in the digital realm.

Perhaps the most prominent piece of accessibility legislation today, at least in Finland, is the Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies – the Web Accessibility Directive. It aims to improve the quality of digital services and increase equality in today’s digital society by creating uniform minimum requirements for the accessibility of public sector websites and mobile applications in the EU (Ministry of Finance 2017). Since all EU member states have ratified the CRPD and therefore committed to ensuring equality for persons with disabilities, the directive does not technically introduce any new requirements (Web Accessibility Directive 2016/2102, 2). Instead, it brings more attention to the topics of equality and accessibility in the digital world and provides clear guidelines for member states to create accessible online services. The directive came into effect on December 12, 2016, and required member states to create national legislation to comply with it (Ministry of Finance 2017). In Finland, this is the Act on the Provision of Digital Services (306/2019), which came into effect on April 1, 2019.

In the Web Accessibility Directive (2016/2102, 1), ‘accessibility’ refers to the principles and techniques that must be followed when planning, developing, maintaining, and updating websites and mobile applications, and the content within them, in order to make them more usable for everyone, especially for persons with disabilities. For a digital service to be considered accessible, it should fulfill certain accessibility requirements, include an accessibility statement describing the accessibility of the service, and provide a way for users to give feedback on the service’s accessibility (Web Accessibility Directive 2016/2102, 6–7). The accessibility requirements are defined in the European Standard (ETSI) EN 301 549 Accessibility Requirements Suitable for Public Procurement of ICT Products and Services in Europe, which is based on the Web Content Accessibility Guidelines (WCAG) 2.1 (Web Accessibility Directive 2016/2102, 7–8; Leskelä 2019, 51). It is worth noting that some of the standard’s requirements are not relevant to the directive, and though the standard draws heavily from the WCAG 2.1, the two are not identical (European Commission 2025). However, as none of these differences have to do with language use, they are not relevant to the scope of this thesis and will therefore not be covered here – the WCAG is discussed in the next chapter though.

In 2019, the accessibility legislation and requirements in the EU extended, as the Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility

requirements for products and services – the European Accessibility Act – came into effect. It aims to help member states implement CRPD requirements, further promote common accessibility regulations within the EU, and remove barriers to help the internal markets function better (European Accessibility Act 2019/882, 1, 3). The European Accessibility Act applies to a variety of products and services placed on the market after June 28, 2025, including, for example, computers and operating systems, self-service terminals, banking services, products and services related to passenger transport, e-commerce, e-books, audiovisual media services, and answering emergency communications to 112 (European Accessibility Act 2019/882, 14–15). A significant difference between the scope of the European Accessibility Act and the Web Accessibility Directive is that the latter only applies to public sector organizations, while the new directive applies equally to both public and private sector operators (8–9). Additionally, the requirements of the European Accessibility Act, outlined in Annex I of the directive, explicitly mention language-related topics, such as using words consistently and following a logical structure to ensure the understandability of information (31, 33, 38).

As these developments show, accessibility legislation is still changing, improving, and expanding, and the role of accessible language is becoming more prominent. This might indicate that linguistic accessibility could become a clearer legal requirement in the near future, and I believe organizations will start to pay more attention to the understandability of their communication, not just its technical accessibility.

### **2.3 Web Content Accessibility Guidelines (WCAG)**

The Web Content Accessibility Guidelines (WCAG) refers to a set of criteria, guidelines, and recommendations created by the World Wide Web Consortium (W3C), together with multiple international individuals and organizations, in order to establish a single, international standard to support the creation of accessible web content (W3C 2024d; 2023). It focuses on improving web content accessibility for persons with disabilities, but notes that the guidelines will also improve overall usability for all users (W3C 2023). As noted earlier, the WCAG and the Web Accessibility Directive are intrinsically linked, as the directive relies on the European standard ETSI EN 301 549, which is based on the WCAG (Leskelä 2019, 51).

The WCAG, and subsequently the Web Accessibility Directive, is divided into four main principles: perceivable, operable, understandable, and robust (W3C 2023; Web Accessibility Directive 2016/2102, 5). These principles create the foundation for accessible web content: the navigation, user interface components, and information on a website must be perceivable with different senses, usable and navigable with different methods, understandable to all users, and robust enough to be used by different users, with or without the help of assistive technologies, and to remain accessible as technologies develop (W3C 2024b). The principles are divided into a total of 13 guidelines, which in turn consist of testable success criteria that are complemented with techniques that help fulfill the criteria (W3C 2023). Each success criteria is assigned a level of conformance: A, AA, or AAA (W3C 2023). The Web Accessibility Directive mandates level AA conformance, meaning all level A and AA success criteria should be fulfilled – level AAA is considered the best, and those criteria are currently not required by law (European Standard (ETSI) EN 301 549 2015; W3C 2023).

The Web Accessibility Directive is based on the WCAG version 2.1, which was published in June, 2018 (W3C 2023). The guidelines are continuously reviewed and updated, and the newest version, WCAG 2.2, was published in October, 2023 (W3C 2023). In addition to updating and maintaining the WCAG 2 versions, the W3C is currently working on a new version, WCAG 3. It is currently in an exploratory phase, with the latest working draft published in September 2025, and is intended to develop into a standard in a few years (W3C 2025b). The fundamental idea and goals behind the new guidelines remain unchanged, but WCAG 3 aims to be more flexible for different types of content, be easier to understand, and cover a wider range of user needs, focusing especially on users with cognitive disabilities (W3C 2025b). From the point of view of linguistic accessibility, perhaps the most significant change is the inclusion of rules related to language use; the current draft includes a guideline called “Text and Wording” that contains topics such as unambiguous text, double negatives, verb tenses, and uncommon words (W3C 2025a).

One of the main principles of WCAG 2, as mentioned earlier, is understandability, which is explained as follows: “users must be able to understand the information as well as the operation of the user interface (the content or operation cannot be beyond their understanding)” (W3C 2024b). As is typical, what understandability means, especially from the point of view of language use, remains vague despite the explanation (Leskelä and Uotila 2020, 233). WCAG 2.2 does include a few criteria related to language use, such as helping the user identify

unfamiliar words and abbreviations, but all language-related criteria are level AAA and thus not required by legislation (W3C 2023). Additionally, they focus primarily on technical solutions: the techniques are more about providing methods to identify unknown abbreviations than about aspects of language use that would improve understandability, such as avoiding unfamiliar abbreviations or explaining them in text (W3C 2023).

The limited attention on language use, and other cognitive functions in legislation and guidelines is in part due to the varying causes for the need for linguistic and cognitive accessibility. The challenges faced by individuals with these disabilities, and thus the accommodations they would benefit from, can vary significantly – what is accessible to one user might hinder the accessibility for another (Selovuuo 2019, 32). Language use in particular is difficult to address in international guidelines, as any rules introduced in them would have to be applicable to a wide range of languages. Additionally, assessing cognitive accessibility, unlike that of technical accessibility, cannot be automated or performed without end users, making it more difficult, time-consuming, and expensive (Leskelä and Uotila 2020, 233). The W3C has openly admitted that they have faced challenges in defining guidelines to support those with cognitive disabilities, in part due to the varying nature of their needs (W3C 2023). However, they are continuously improving their guidelines in this regard, for example, by providing supplementary guidance for cognitive accessibility (W3C 2023; 2024a). It seems clear that despite the challenges, addressing cognitive accessibility is important to the W3C, and it is evident also in the plans for the upcoming WCAG 3.

### 3 WEB ACCESSIBILITY

Accessibility is a very broad concept, covering everything from wheelchair-accessible buildings to websites that can be used without sight – and much more (Leskelä 2019, 48–49). The focus of this thesis is on web accessibility, which refers to building, developing, and delivering online services, products, and content in such a way that they are available to everyone, regardless of their abilities or possible barriers to function, in an understandable and usable form (Hirvonen et al. 2020, 13; Selovuuo 2019, 11; ISO 2018). It is about much more than just the legal requirements and guidelines discussed in the previous chapter; it is about justice, fairness, and equality in digital services, and it benefits all users (Selovuuo 2019, 13, 18). For example, not everyone sees colors the same way, uses the same browsers and browser settings, or navigates on websites in a similar manner – web accessibility is about acknowledging and accepting these differences and taking them into account when planning and implementing digital services.

Accessibility can be viewed and analyzed from various points of view, from the different areas of life and society, such as education or culture, to the barriers to access, focusing on physical or cognitive accessibility, to name a few (Hirvonen et al. 2020, 20). In the context of web accessibility, it is typical to focus on the final product or service, analyzing the different aspects of that (21). For example, it is common to consider the technical accessibility of a website and the accessibility of the content within it separately, in part because the responsibility of those areas typically belongs to different people. However, it is important to remember that the overall accessibility of an online service, such as a website, is dependent on the accessibility of the different parts that make up the service, and accessibility should be considered in them all (Jylhä 2020, 70; Hirvonen et al. 2020, 20).

As mentioned earlier, web accessibility has been prominent in public discussions in the past decade or so, but the idea is not a new one. In fact, already in 1997, less than a decade after inventing the World Wide Web, Tim Berners-Lee started the World Wide Web Accessibility Initiative to remove accessibility barriers for all people with disabilities (Friedman and Bryen 2007, 206). According to him, “The power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect” (W3C 2024c). The idea behind the web has thus always been to remove barriers to communication and interaction. Computers and online services can allow persons with disabilities to communicate and interact with others in ways

that would not be possible for them in the physical world, providing them with independence they might not otherwise have (W3C 2024c). Nevertheless, it is important to keep in mind that despite the accessibility considerations, the web is still inherently inaccessible to some, for example, those with unreliable internet connections and those who lack the necessary skills to use digital services (Horowitz et al. 2019, 137). So, while this thesis focuses on web accessibility, the importance of offering services and information via other modes as well should not be forgotten or overlooked.

### **3.1 Accessibility Is for Everyone**

For a long time, society has tended to define a so-called ‘normal’ person, whose world those who are different are trying to adapt to, and accessibility has been focused on promoting that adaptation. However, recently, the focus has been increasingly shifting from this to viewing the environment, rather than the person and their traits, as the limiting factor creating restrictions and obstacles that prevent access (Hirvonen et al. 2020, 20). The European Accessibility Act (2019/882, 7–8), for example, states that “accessibility should be achieved by the systematic removal and prevention of barriers, preferably through a universal design or ‘design for all’ approach, which contributes to ensuring access for persons with disabilities on an equal basis with others”. With the goal of implementing CRPD requirements, the European Accessibility Act (2019/882, 1) keeps the focus on users with disabilities but also notes that any, permanent or temporary, functional limitations, together with environmental barriers, can result in the need for alternative, accessible solutions. Therefore, the requirements of the directive can benefit groups such as the elderly and those who are pregnant, but also, for example, people with heavy luggage unable to carry it up the stairs (1).

However one defines disability, the audiences of accessibility are diverse: different users have different traits and face different barriers, and thus need different solutions to ensure accessibility (Selovuuo 2019, 30–32). Some examples of disabilities and difficulties that can lead to the need for accessible solutions in the web include difficulties with perception or concentration, vision and hearing impairments, memory disorders, intellectual and learning disabilities, neurological disorders, cognitive disabilities, psychological limitations, and difficulties brought on with age (Selovuuo 2019, 40; Hirvonen et al. 2020, 16). This is a significant portion of the population, as some estimate that accessible content can help 1.5–2.5 million people in Finland alone (Selovuuo 2019, 15). Worldwide, approximately 1.3 billion

people, 16% of the global population, “experience a significant disability”, according to estimates by the WHO (2023). Additionally, the number is projected to increase significantly in the following years due to the aging population and increase in noncommunicable diseases (WHO 2023).

Despite the focus on persons with disabilities, accessibility benefits everyone. After all, anyone can, at any point in their lives, have a temporary functional limitation causing the need for accessible solutions. These functional limitations can even be entirely caused by external, environmental factors such as language or culture – or even a noisy environment in which anyone can benefit from subtitles originally intended for hard of hearing and deaf users (Hirvonen et al. 2020, 16). It is also important to remember that anyone can become disabled at any point in their lives: 97% of disabilities are acquired, 89% of them through an illness (Maaß and Hansen-Schirra 2022, 43). Finally, in addition to benefiting all users, accessibility benefits businesses too: users needing accessible solutions are consumers just as much as anyone else, and businesses can lose potential customers by creating inaccessible services (Selovuo 2019, 15). Additionally, many of the techniques used to increase the accessibility of a website also increase its usability for all users and can help it show up higher in search results, potentially increasing traffic to the site (16).

### **3.2 Usability**

Striving for accessibility increasingly means using principles and techniques that benefit everyone and make services usable to all, rather than creating separate solutions for persons with disabilities (Hirvonen et al. 2020, 16–17). Usability, just like accessibility, is a complex concept and has varying definitions depending on the viewpoint, and, regardless of the definitions, both are essential in online services. In essence, usability is about how well a product or service works in the situation it is used in (Hirvonen et al. 2020, 19). It is always about the user’s subjective experience, so there is no one right answer to making a service or product usable; instead, usability depends on the user and their characteristics, the type of task they are trying to complete, the equipment they are using, and other environmental and intrapersonal factors (Suojanen et al. 2015, 14).

This close connection between usability and accessibility leads to the interchangeable use of the terms, particularly in informal conversations. Perhaps the most significant difference is that

accessibility emphasizes the fair and equal treatment of all users, while usability is more focused on the actual use of the product or service (Suominen 2019, 59). Accessibility can also be seen as a part of usability, and for example, an ISO standard on software accessibility defines accessibility as the “usability of a product, service, environment or facility by people with the widest range of capabilities” (ISO 2008). Alternatively, usability can be seen as part of accessibility, as can be seen, for example, in how I defined web accessibility earlier: the building, development, and delivery of online services and content in such ways that they are available, understandable, and *usable* to all.

While the focus of this thesis is on accessibility, usability cannot be left entirely outside the scope due to the close relation between the two. Additionally, the method I used for analysis, heuristic evaluation (discussed in chapter 6), comes from the field of usability research. Finally, my approach to the textual analysis is inspired by usability – I focus on the texts as tools the users use to achieve some goal and analyze their accessibility from this viewpoint.

## 4 LINGUISTIC ACCESSIBILITY

As I have noted before, accessibility is more than just access, and for information to be truly accessible, it must also be understandable. For example, if one was given a book written in a language they do not understand, they would technically have access to the information in the book, but it would not be accessible to them, as they would be unable to understand, and thus use, the information (Hirvonen et al. 2020, 20). Similarly, information can be inaccessible even if it is written in a familiar language, if the language is too complex for one's level of comprehension or, for instance, meant for an audience of experts in a field unfamiliar to the user. This is what linguistic accessibility is all about – ensuring the language used to deliver information is clear, simple, and therefore understandable enough that everyone within the intended audience can comprehend and use the information.

In the context of web accessibility, linguistic accessibility is concerned with the linguistic material on a website, or other online service, and covers everything from user interface texts and instructions to blog posts and online discussions, as well as audio clips, videos, and images that contain text – any use of language, be that oral or written, can be assessed for linguistic accessibility. The focus of this thesis is solely on written texts on a website, so the accessibility of spoken language will not be covered here. Nevertheless, it is important to keep in mind that written language is not the most accessible solution for everyone. For some, it might be easier to understand spoken language, and others might benefit from images alongside or instead of the text. Though it is not possible to cover it within the confinements of this thesis, accessibility can be significantly improved when information is offered in multiple formats and via multiple channels.

The Finnish Ministry of Justice (2018, 19) defines linguistic accessibility, specifically in the context of health and social services, as the patient or customer being aware of the services and receiving care, being understood, understanding instructions, and therefore being able to take responsibility of their own care. Additionally, in order for information and services to be linguistically accessible, they must be accessible in other aspects too; for example, digital services must be technically accessible, information must be available via different channels, and interpreting services must be available (19). Finally, services must match the users' needs in order to be useful and thus accessible to them (86). Although the Ministry of Justice focuses

on health and social services, their definition of linguistic accessibility can easily be generalized and applied to other contexts as well.

To match the scope and focus of this thesis, I define linguistic accessibility as the extent to which the linguistic material on a website – for example, user interface texts, instructions, and videos – can be understood by everyone, regardless of their abilities or personal traits, and can be used to achieve the desired goals. This definition is largely based on the definition from the Ministry of Justice described above. My focus, as noted earlier, is on written texts, and the discussion here will reflect that, but the definition can be applied to all linguistic material equally. A central part of the definition is the ability to use the information to achieve goals: texts are usually read with some goal in mind, and the information gathered from the texts is used to achieve that goal. For this to happen, the user must comprehend the information well enough to be able to connect it with other, pre-existing and simultaneously processed information (Hirvonen et al. 2020, 24). This is mostly true for informative or instructional texts, such as the ones analyzed in this thesis, but it can be applied to fictional texts read for recreation as well. Then, the goal can be, for example, to enjoy the experience and have fun; if the language is inaccessible, users might have to spend their energy on trying to understand what is being said, which can hinder their enjoyment.

## **4.1 Understandability**

Linguistic accessibility cannot be discussed without mentioning understandability. Understandability, just like accessibility, has various definitions depending on the viewpoint and the researcher. Rello et al. (2013, 203), for example, define it as the ease with which the reader can comprehend the main idea of the text and draw conclusions from it. This is quite similar to my definition of linguistic accessibility, as drawing conclusions from a text can be seen as using the text to reach some goal. Selovuo (2019, 83), on the other hand, defines understandability as providing content to users in different ways so that the information is accessible to them and notes that it covers not only language but also considering how different users use the content or understand different types of messages; for example, deaf people can watch a video but cannot hear the audio, while blind people are dependent on just the audio. This definition is a good example of the overlap between (linguistic) accessibility and understandability. Although the two concepts are very similar and often used interchangeably, the term ‘linguistic accessibility’ takes on a more social standpoint: it has to do with who has

access to information, can process and comprehend it, and can therefore participate in society and impact future information (Heikkinen 2012, 43). This is why I have chosen to use the term ‘linguistic accessibility’ rather than ‘understandability’ – I want to bring attention to the way language affects equality and the fulfillment of basic human rights, especially from the point of view of persons with disabilities.

Another related term, sometimes used interchangeably with understandability, is readability. Rello et al. (2013, 203) define readability as “the legibility of a text, that is, the ease with which text can be read”. Often, readability is seen as consisting of two parts: legibility and ease of reading. Legibility is primarily a visual feature and has to do with how clear and easily distinguished different typographic elements, such as letters and words, are (Suojanen et al. 2015, 50). The ease of reading, then, focuses on the style and complexity of writing and considers how easy it is to read whole paragraphs and texts (51). Compared to understandability, which is usually seen as a more complex concept dependent on the reader’s individual traits, readability is linked to quantifiable things such as the length of words and sentences (54).

## **4.2 Language Variants on a Spectrum of Accessibility**

When discussing linguistic accessibility, it is important to remember that there are different levels or variants of language that are used in different contexts and meant for different audiences and purposes. The type of variance that is relevant in this context, though, is related to the complexity, and thus the accessibility, of language, which can be seen as a spectrum. For instance, specialized language, characterized by the use of specialized vocabulary – jargon – and used primarily in scientific and vocational contexts, such as in research papers and reports, would be placed in the more complex and less accessible end of the spectrum (Leskelä 2019, 93). However, it is also crucial to consider the intended audiences: specialized language is used when communicating with an expert audience about a topic related to their area of expertise, so the language could be seen as accessible to its intended audience. The level of complexity in the language variant should reflect the intended audiences of the text.

The sort of middle ground on a spectrum of language complexity is standard language. It is the variant that is shared among all speakers of the language, clearly distinguishing it from specialized language only shared by experts of a specific field, but also from, for example,

regional dialects (Hiidenmaa 2005). Standard language is factual and informative in style, and it is used when communicating with a general audience; for example, public authorities use standard language when communicating with the general public, as their goal is for as many people as possible to understand the message (Hiidenmaa 2005; Leskelä 2019, 94). Standard language can be modified to make it more accessible, to take those who struggle with understanding standard language into consideration, while still aiming the text at general, large audiences (Leskelä 2019, 94). This usually involves methods such as simplifying long and complex sentences and explaining any jargon that might be used (95). The resulting accessible standard language, sometimes also called plain language, is a form of standard language that pays special attention to wording and structure to make the information as easy to find, clear, and understandable as possible (ISO 2023). In this thesis, I use the term ‘accessible language’ to refer specifically to accessible standard language. However, accessible language can also be used as an umbrella term for all accessible forms of language, such as simplified language.

Simplified language, sometimes referred to as easy-to-read or easy language, is an artificially created language variant that has been modified in content, vocabulary, and structure to include only the aspects of standard language that are the easiest to understand (Kulkki-Nieminen 2010, 20–21; Leskelä 2019, 93). As an artificial language variant, simplified language has clear and strict rules that must be followed. Simplifying a text, for example, involves much more than just shortening and simplifying the structure of the original standard language version – in addition to the removal of complex structures and words, clarifications and examples are often added to facilitate understanding (Kulkki-Nieminen 2010, 120). In comparison, accessible standard language makes some modifications to simplify the text, but those modifications are more dependent on the writer and follow general understandable writing guidelines rather than strict, predetermined rules (Leskelä 2019, 94). Similar to specialized language, simplified language is aimed at a specific audience: those with language disorders or other difficulties understanding standard language (Leskelä and Uotila 2020, 227). These difficulties can be caused by various factors, such as intellectual or learning disabilities, memory disorders, mental health issues, substance use disorders, or limited skills in the specific language (229–30). The Finnish Center for Easy Language estimates that in Finland approximately 11–14% of the population need simplified language, and the percentage keeps growing due to the aging population and increased immigration (230).

### **4.3 Linguistic Accessibility Ensures Access to Information**

As discussed earlier, accessibility plays a crucial role in ensuring equality and thus fulfilling some of the fundamental human rights we all should have – particularly for persons with disabilities. Living in a society requires information about how the society functions and how one can participate in it, as well as about the different services it offers, such as health care or education (Koponen and Nurminen 2020, 304). Not having access to this information, or not understanding it, can cause issues in various aspects of life, and everyday activities such as following the news, buying bus tickets, or following recipes while cooking can become difficult or even impossible (Leskelä 2019, 56). Without understandable language, no service or product can be truly accessible, even if the legal accessibility requirements are met. Language is the key to ensuring access to information.

As the primary channel for transmitting information, language plays an important part in almost all areas of life (Koponen and Nurminen 2020, 306). An individual's abilities to, for example, engage in dialogue and participate in social and public life are all dependent on language, as is the realization of many fundamental rights (UNESCO 2007, 2). Linguistically inaccessible texts can even hinder individuals' abilities to participate in democracy and thus prevent democracy from being carried out (Suominen 2019, 62). Moreover, if information about the rights of individuals is not linguistically accessible, individuals might not be aware of their rights, which can easily lead to them unknowingly giving up those rights (Ministry of Justice 2018, 20). Additionally, linguistic factors have even been shown to play a significant role in the eradication of extreme poverty and hunger, achieving universal primary education, and responding to diseases like AIDS and malaria (UNESCO 2007, 2). Linguistic accessibility, therefore, is essential.

The digital world and advances in technology have made information more easily available than ever and given people more options to communicate (Horowitz et al. 2019, 26). Today, nearly all information is online, and the web is an important resource in almost all aspects of life, from education and employment to commerce and health care, recreation, and much more (Koponen and Nurminen 2020, 306; W3C 2024c). In Finland, public authorities are increasingly prioritizing digital communication channels, for example, with the help of services like Suomi.fi Messages, and legislative changes to support this are planned to come into effect at the beginning of 2026 (Ministry of Finance 2024). In addition to information, people are

online as well – according to the Survey of Adult Skills, the percentage of adults using the internet in OECD countries increased from 76% in 2012 to 93% in 2023 (OECD 2024a, 34). Additionally, Eurostat’s research (2024) showed that 92% of internet users in Finland used the internet to interact with public authorities. This also means that web accessibility is an essential requirement for people to be able to receive information and participate in society – to fulfill their constitutional rights – and proper attention should be given to the accessibility of online services, including the language used in them (Koponen and Nurminen 2020, 306; W3C 2024c).

The ubiquity of the internet and the unprecedented amount of information available at the tip of our fingers also bring new challenges – having easy and near-instant access to vast amounts of information is not solely positive. The internet allows for a rapid, ever-present spread of information, including the spread of misinformation, disinformation, and conspiracy theories, leaving users to face the challenges of identifying the correct, trustworthy information from a range of conflicting sources (OECD 2024, 34–35; Rathore and Farooq 2020, S162). The downsides of this so-called infodemic are evident, for example, when looking at the information overload regarding COVID-19: during previous pandemics, lack of information caused significant issues, but during COVID-19 the problem was the opposite, as information spread rapidly online and users had to try to find out which of it was correct and authenticated (Rathore and Farooq 2020, S162). One of the main ways to find reliable information and avoid spreading misinformation is to look for information on the official websites of authorities (S164). To support this, public authorities’ websites should be easily available, clearly identified, and, most importantly, accessible to everyone. This also means the information on those sites should be linguistically accessible, as inaccessible language can lead to misunderstandings or cause the users to look for information elsewhere, which again increases the risk of finding and potentially spreading incorrect information (S164). Accessibility of information like this can save lives, and inaccessible information can negatively impact not only the health of the person who cannot access the information, but also the people around them.

#### **4.4 Linguistic Accessibility Benefits Everyone**

The largest group needing accessible solutions consists of those with cognitive disabilities such as learning disabilities, intellectual disabilities, neurological disorders, and memory disorders (Selovuo 2019, 15; Leskelä 2019, 49). These disabilities have to do with cognitive functions, such as perception, memory, attention, reasoning, problem solving, decision making,

calculating – and linguistic functions (ISO 2020). Because language is one of the key cognitive functions, linguistic accessibility is sometimes considered to be a part of cognitive accessibility. In this thesis, however, linguistic accessibility is viewed from the perspective of the final product, as a part of the accessibility of content. This is simply a matter of perspective, and linguistic accessibility can be considered through either lens. Regardless of the viewpoint, those with cognitive disabilities can significantly benefit from linguistic accessibility as well.

As mentioned earlier in the context of the WCAG, cognitive and linguistic accessibility are difficult to address in guidelines and legislation, due to their subjective nature and significant variance in the difficulties and needs of the users. Additionally, the need for cognitive and linguistic accessibility has even been doubted altogether, as those needing it are assumed to get linguistic support from their family and friends (Leskelä and Uotila 2020, 232). This clearly goes against the principles of equality and prevents those who need linguistic support from living independent lives, as well as puts an unfair burden on their loved ones (232). Accessible language, just like other accessibility solutions, ensures independence for many (Leskelä 2019, 48).

Similar to other forms of accessibility, linguistic accessibility too benefits everyone while ensuring crucial access for some. Anyone can be in a situation where they would benefit from linguistic accessibility, as common, everyday things such as being tired, being in a noisy environment, or having a headache can temporarily lower a person's cognitive skills (Leskelä 2019, 102). Additionally, accessible language lightens the cognitive load for anyone and allows them to focus on the topic at hand, which can be particularly helpful when trying to learn about something completely new or understand a difficult topic (Leskelä 2019, 57; Selovuo 2019, 48–49). Linguistically accessible information is faster and easier to understand, which also benefits businesses: clear information decreases the likelihood of errors made by users and lowers the number of times users contact customer support for help, thus saving the company resources and increasing customer satisfaction (PLAIN, n.d.).

In addition to persons with disabilities that affect language and understanding, linguistic accessibility can benefit anyone with poor literacy skills for reasons not directly related to a disability. This can be due to, for example, not being used to reading, having a lower level of education, dealing with an unfamiliar language, or aging (Jylhä 2020, 67). Poor literacy skills have been linked to an increased risk of social exclusion, dropping out of education, problems with income and livelihood, and difficulties in life management – and through these, added

costs to the state (Leskelä 2019, 56). Additionally, those with higher proficiency in literacy are more likely to report better health and higher levels of satisfaction in life (OECD 2024b). Overall, the literacy rates in Finland are higher than average, and have actually been rising over the past decade, as the results of the 2023 Survey of Adult Skills show (OECD 2024a, 102–4). However, that does not negate the need for linguistic accessibility – cognitive disabilities are fairly common, and, additionally, things like mental health issues and immigration can cause, often temporary, cognitive difficulties that can be reflected in lower literacy proficiency (Leskelä 2019, 102).

One significant group that scored lower than average on literacy proficiency in the survey was older adults, that is, people between 55 and 65 years of age (OECD 2024b). When viewing these results, it is important to note that the survey was conducted on adults aged 16–65 years, which means that the results do not provide insights for users over 65 years at all. It is a well-established fact that aging affects cognitive skills, particularly memory and attention, and old age is the single biggest factor contributing to the need for accessible language (Maaß and Hansen-Schirra 2022, 43; Leskelä 2019, 102). Additionally, the number of over 65-year-olds in Finland has been growing for a long time, and the growth is predicted to continue – in 2023, there were 1.3 million persons over 65 years of age, and Statistics Finland (2024) estimates that the number will reach 1.5 million in the beginning of the 2040s, and 2 million by 2070. The need for linguistic accessibility will thus keep growing as well. From a business perspective, it is also worth noting that as the number of elderly people is growing, so is their purchasing power; for example, one study from Germany showed that every 3<sup>rd</sup> euro was spent by someone over the age of 60 (Maaß and Hansen-Schirra 2022, 43). Investing in linguistic accessibility can thus have financial benefits for businesses as well.

#### **4.5 Linguistic Accessibility for Language Minorities**

Linguistic accessibility is also important for one major group that is particularly central to the topic of this thesis – language minorities. With that, I refer to those who do not speak the majority language, or languages, of the country natively or fluently. In Finland, the majority language is Finnish, but in the context of this thesis, I consider Swedish to be on the same level as Finnish, as they are both official national languages and all public authorities are required to provide information in both languages on their websites. This is an example of linguistic rights, that is, the right to use specific languages in specific situations, often required by law (Ministry

of Justice 2018, 19). The concept of linguistic rights is an essential part of the conversation about language minorities. Additionally, it is closely related to, and can be seen as a part of linguistic accessibility, especially in multilingual societies, as the languages in which information is available in is a central aspect of linguistic accessibility (Nurminen and Koponen 2020, 152). Some even view the two terms as synonymous and define linguistic accessibility as having to do with accessing the information in a specific language, often focusing on speakers of minority languages (Matamala and Ortiz-Boix 2016, 11; Schuster 2012, 312).

In addition to the official national languages, Finnish legislation covers the use and protection of Sami languages, Romani languages, and Finnish sign languages, as well as the right for interpretation or translation due to illness or disability (Ministry of Justice 2018, 20). Although the right to interpretation and translation, the right to be understood, and other related linguistic rights are also granted to speakers of other languages, at least in certain situations such as health care and other official matters, there is no requirement for public authorities to provide information in languages other than Finnish and Swedish (25–26). However, most, if not all, public authorities do generally provide at least some of the information on their websites in English as well. This is important to keep in mind as this thesis focuses specifically on that English-language information.

The most significant language minority, and the one central to my focus, consists of migrants and those with a migrant background, such as children of immigrants (Leskelä 2019, 104–5). The migrant population in Finland has been growing annually for a long time, and the growth has been increasing lately: the number of foreign language speakers has increased by almost 200,000 after 2019, and at the end of 2024, over 600,000 people in Finland – 10.8% of the population – had a foreign first language (Statistics Finland 2025). However, it is worth noting that because bilingualism on an individual's level is not official recognized, that number likely includes some people who are also native speakers of Finnish or Swedish despite having a foreign language as their official first language in the population register, such as children of immigrants (Määttä et al. 2022, 5). In addition to more permanent immigration, Finland gets an average of 6000–7000 foreign exchange students and interns every year (Vipunen, n.d.) – and that statistic only includes those who have stayed or are planning to stay for at least a year, so the true number is likely higher.

In 2024, the most common foreign first languages were Russian, Estonian, Arabic, Ukrainian, English, Somali, and Farsi, all with over 20,000 speakers (Statistics Finland 2025). However,

many migrants declare the official language of their country of origin or their language of education as their first language, instead of their actual first language and the language they use at home, which in part explains the high number of speakers in languages such as Russian, Arabic, Farsi, English, and French (Määttä et al. 2022, 28). So, while, for example, English is among the top five most-common foreign first languages, the number of those who speak it as a first language might be significantly lower.

Linguistic accessibility is essential for migrants. Firstly, accessible language, and particularly simplified language, have been shown to help with learning a new language, which is a crucial part of integrating into the society (Kulkki-Nieminen 2010, 26). The Survey of Adult Skills also showed that those with a migrant background – including both those who were born abroad and those whose parents were born abroad – had significantly lower literacy proficiency compared to the native population (OECD 2024a, 85). In fact, out of the 31 countries participating in the survey, Finland had the largest difference in literacy proficiency between the native population and foreign-born population (85). Secondly, accessible language is especially helpful in new situations and when trying to understand difficult topics, and migrants, especially new migrants, are in a new place and situation and have to learn about, for example, how the society works and how they can access essential services such as health care – this process can be significantly eased if information is available in an accessible format (Leskelä 2019, 109).

Offering essential information in as many languages as possible would help ensure everyone has access to it, though resources rarely allow that (Koponen and Nurminen 2020, 307). Due to its status as a lingua franca, most public authorities offer information in English, in addition to Finnish and Swedish, but other language versions are rare. Since the information is only available in a limited number of languages, it is important to ensure those languages are accessible. As at least new migrants outside the Nordic countries are probably most familiar with English, they will likely rely heavily on the English-language information. This also means that a significant number of those reading the English version are not native, or necessarily even fluent, speakers of English, further emphasizing the need for linguistic accessibility.

The importance of translation and interpretation services should not be overlooked though – especially since the increased emphasis on English does have some concerns, such as placing different migrant groups in unequal positions based on their level of English skills and potentially decreasing the migrants' interest in learning the local language (Leskelä 2019, 19; Koponen and Nurminen 2020, 305). Additionally, focusing on providing accessible information

in the country's official languages promotes learning those languages and thus helps with integrating into society. So, while it is important to increase the number of languages information is available in, and make the information as linguistically accessible as possible, it is also crucial to ensure the accessibility of the information in the official languages, for example by using simplified language, and encourage migrants to learn those languages. Nevertheless, when it comes to crucial information, such as information about one's health and well-being, it is essential that it is available and accessible to all – migrants, for example, should not be required to learn a certain level of Finnish or Swedish to be able to use health care services. I will not focus on these issues more here, but it is important to acknowledge them.

It is also worth noting that the use of machine translation is extremely common: for example, in just April of 2021, Google Translate had been used to translate 20 billion web pages globally (Diño 2021). Additionally, several studies have shown that migrants, particularly the younger generations, use machine translation to help them understand official texts and, for example, communicate with health care personnel (Pym et al. 2022, 74). These users are generally aware of the possibility of errors, but still rely on machine translation as it provides independence (74). The output of machine translation can be improved by pre-editing the text before it gets translated, and the principles of pre-editing are very similar to the guidelines of accessible language: explaining potentially ambiguous words, using short sentences, and avoiding complex syntax, to name a few (Pym et al. 2022, 76). Simplifying the texts to be more accessible can thus also help prevent some of the potential translation errors caused by machine translators. So, in addition to facilitating understanding, accessible language can help machine translators to provide more correct translations (Koponen and Nurminen 2020, 315–16).

## **5 RESEARCH DATA**

The purpose of this thesis is to evaluate the linguistic accessibility of the English language texts on the kanta.fi website. Kanta is a public sector entity dealing with health and social services, and the kanta.fi website is maintained by public authorities. This means that the website should meet the requirements of the Web Accessibility Directive and other applicable legislation. Additionally, as mentioned earlier, it is particularly important for public authorities' websites to be accessible and understandable to promote the finding and using of correct and authenticated information – especially when it comes to essential topics such as health care.

To place the website and the texts within it in context, I first briefly explain what administrative texts are and discuss their importance in the health care context. After that, I introduce the Kanta Services and the kanta.fi website and explain why I have chosen to focus on them. Finally, I go over the data I have selected for analysis.

### **5.1 Administrative Texts**

Administrative communication, also called public communication, refers to the communication from public authorities to the general public, often done in the form of administrative texts (Kotus, n.d.). It is about sharing information from public organizations and institutions to the general public, usually with the goals of notifying the public about decisions and actions, promoting values, and helping maintain social cohesion (Pasquier 2018, 441). The language used for administrative communication – administrative language – is controlled not only by the norms of standard language, but also by legislation and principles that govern all actions of public authorities (Kotus, n.d.). This includes, for example, the Administrative Procedure Act's (434/2003) requirement to use clear and understandable language discussed in chapter 2.

As administrative texts are aimed at broad, increasingly multicultural, audiences, it is particularly important to pay attention to the linguistic accessibility of those texts (Pasquier 2018, 453). This is also recognized in the Finnish government communications guidelines from 2023 which state that understandability and accessibility are requirements for successful communication and should therefore be central in all communication and communication channels (Prime Minister's Office 2023, 11). According to the guidelines, in addition to the legal requirement for understandable language, public authorities should use simplified

language when necessary as well as communicate the most important topics also in English and the minority languages used in Finland (10–11). After all, the government and public authorities, along with their communication, are for the people and should therefore be accessible and understandable to all (7).

This recommendation is reflected in, for example, Kela's language policy for customer communication, in which they state their goal to increase communication in English and improve the understandability and clarity of the English used in their communication (Kela 2025, 8). As the majority of those who use English to communicate with Kela are not native speakers of the language, the accessibility of it is crucial (8). Kela does also acknowledge the need for more multilingual communication and cooperate with Infonland.fi to provide information in more languages, as they alone lack the resources to provide separate communication materials for all language groups (10). Kela's language policies are relevant to the topic of this thesis, as Kela is responsible for the kanta.fi website and other communication related to the Kanta Services. Since I was not able to find a separate language policy for the Kanta Services, one could assume Kela's policies apply to them as well.

## **5.2 Administrative Texts in Health Care**

Health care communication, in the broad sense, refers to any and all communication related to health, sickness, health care, and the field of medicine (Torkkola et al. 2002, 22). In this thesis, I approach the topic from the perspective of administrative communication and focus on the health-care-related communication provided by public authorities. This is usually general and informative in nature and covers topics such as what health care services are available and how they can be used. As a significant portion of the health care services in Finland are provided by public authorities, much of health care communication is also administrative communication, which means it should follow the legislation and general principles that govern all other administrative communication as well.

Health is one of the most important aspects of life, and access to health care is one of the fundamental human rights defined in the Constitution of Finland (731/1999). Language plays a significant role in health care, as the realization of linguistic rights affects patient safety and ensures the fulfillment of the right of self-determination (Ministry of Justice 2018, 20). The Act on the Status and Rights of Patients (785/1992) states that the patient must be given information

on their health and matters related to it in such a way that they sufficiently understand it, with the help of interpretation services if necessary, and requires the patient's individual needs, mother tongue, and culture to be taken into consideration in their treatment and care as much as possible. An ombudsman has determined that a patient's self-determination rights cannot be ignored on the basis of the patient not understanding the languages used by authorities (Ministry of Justice 2018, 25). While these requirements primarily apply to situations where a patient is seeking health care from a professional, I would argue that they could also be applied to the distribution of more general health care information available online. After all, when the information available is accessible, users are less likely to contact the service providers asking for clarifications and better prepared to take responsibility of their own care (Torkkola et al. 2002, 8).

The accessibility of health care information is crucial, particularly due to the importance of the topic. The primary audience of accessible solutions, persons with disabilities, are already at a disadvantage when it comes to health and health care. This inequality stems from the unfair conditions they face, including discrimination, poverty, and exclusion from education and employment, as well as barriers in the health care system (WHO 2023). Due to the lack of accessible information, persons with disabilities are often left out of public health interventions, increasing the risk factors for non-communicable diseases such as smoking, alcohol consumption, poor diet, and lack of physical activity, which in part leads to a higher risk of developing conditions such as depression, diabetes, obesity, and poor oral health (WHO 2023). Persons with disabilities would thus significantly benefit from accessible health information. Additionally, in the health care context, the need for linguistic accessibility is higher for everyone, as even those with high literacy skills can struggle with health literacy. Especially the use of unfamiliar medical terms can hinder understanding, but simply the context of health care in itself plays a role: health care texts are typically only read when necessary, which generally means that the user or their loved one is sick or otherwise affected by some health issues – these kinds of new and potentially stressful situations are known to negatively impact the abilities to focus and understand information (Torkkola et al. 2002, 18; WHO 2024).

In addition to persons with disabilities, accessible information about health care and health care services is particularly important for migrants as well; studies have shown that their health and well-being are generally poorer than that of the native population, in part due to language barriers and the resulting lack of information (Tirronen et al. 2020, 8; Ministry of Justice 2018,

34). Additionally, despite the need for medical services being approximately the same between the native population and the migrant population in Finland, migrants tend to use health care services less (THL 2024). While migrants come from a variety of backgrounds and have their own, unique experiences and issues, experiences of discrimination stemming from language barriers and inefficient awareness of health care and the available services have been observed across different groups and language minorities (Ministry of Justice 2018, 34). In Finland, immigrants with a residence permit and a permanent municipality of residence are entitled to the same health and social services as Finnish citizens, as are undocumented children (THL 2024). Additionally, undocumented adults are entitled to urgent medical care as well as non-urgent services under the discretion of medical professionals (THL 2024). The right to health care is thus not dependent on one's language skills, and in order for migrants to get information about their rights and the health care services available, that information should be made as accessible as possible.

### **5.3 Kanta.fi**

Kanta.fi is the website and primary communication channel for the Kanta Services, providing up-to-date information and guidance to all users (Kanta Services 2025c). The Kanta Services, often referred to as just Kanta, is a nationwide solution consisting of a “set of digital services that store citizens' social welfare and health care data”, such as patient and client data repositories, a pharmaceutical database, and a prescription service (Kanta Services 2025c). The majority of these are intended to be used by professionals, and the services are used across public and private operators in the health care and social welfare sectors: the staff records data in their own patient or client data systems, from where the data is stored in the Kanta Services and made available for use by other professionals as needed (Kanta Services 2024d; 2024b). Kanta was created to support the providers of health care and social welfare services, as well as to help citizens manage their health and well-being as easily as possible (Kanta Services 2024d). The services give professionals in different organizations across the country access to their patients' data, regardless of where the data was recorded, and thus help avoid unnecessary repetition of information and tests, saving time and resources and allowing the professionals to focus on providing high-quality care (Kanta Services 2024b).

The Kanta Services are developed and managed in co-operation by various Finnish public authorities: Kela (Social Insurance Institution of Finland) and THL (Finnish Institute for Health

and Welfare) are responsible for the development of the Kanta Services under the strategic guidance of the Ministry of Social Affairs and Health (Kanta Services 2024a). They also collaborate with multiple other actors, such as the well-being services counties, private service providers, pharmacies, system providers, and other government agencies (Kanta Services 2024a). Communication related to the Kanta Services – including the kanta.fi website – is handled by Kela (Kanta Services 2024a). The use of the Kanta Services is managed, and mandated, by law: legislation for Kanta was created in 2007 and deployment of the services started in 2010 (Kanta Services 2025a). Today, all pharmacies and public sector health care service providers, as well as the majority of private health care service providers use the Kanta Services (Kanta Services 2025a). Additionally, social welfare services are required to start recording data into Kanta by 2026 (Kanta Services 2024b).

The one Kanta service that most individuals in Finland are familiar with is MyKanta (or OmaKanta in Finnish). It is the most frequently used health care and social welfare service in Finland, with over a million different users each month, and is arguably the single most important online service in the health care and social welfare sector (Kanta Services 2025c). Unlike the majority of the Kanta Services, MyKanta is aimed at the citizens – it allows them to view the information recorded about them in Kanta easily, anywhere and anytime (Kanta Services 2025d). It came into use in 2010, and today it can be used, for example, to view treatment-related records such as vaccinations and appointment records, check the results of laboratory tests, and request prescription renewal (Kanta Services 2025d).

Despite the undeniable importance of MyKanta, it is unfortunately not available in languages other than Finnish and Swedish, rendering the use of it difficult or even impossible to users who do not speak those two languages. This means that up to 10% of the population is unable to use this essential service due to language barriers. Of course, some of those with a foreign first language do also speak Finnish or Swedish, so the 10% is most likely an exaggeration. Still, a significant portion of the population is unable to use MyKanta due to a language barrier. As mentioned before, on average, immigrants have poorer health than the native population and are less likely to use health care services even if needed – having access to a service like MyKanta could lower the threshold to use these services. Additionally, MyKanta provides independence to its users, as they can handle matters related to their care independently via the service, instead of having to contact a service provider.

Kanta.fi, on the other hand, is available in English. The website includes instructions for using MyKanta, and having those instructions available in English can help some users navigate the service. However, these instructions have one significant flaw, which I will discuss more in the analysis chapter: none of the instructions, or any other pages for that matter, mention in any capacity that MyKanta is not available in English. In addition to these instructions, kanta.fi provides a lot of useful information about the Kanta Services and various topics related to health care and social welfare, such as picking up prescriptions from a pharmacy, which can be particularly helpful to those unfamiliar with the Finnish health care and social welfare systems.

I chose to focus on kanta.fi in this thesis because the Kanta Services are central in the Finnish health care system, and kanta.fi provides important information about the services to all citizens. The information on the website can be relevant to anyone, and particularly new migrants can benefit from the guidance for using the health care systems in Finland. Therefore, it would be crucial for that information to be available and accessible to as many users as possible. Additionally, those who want to use MyKanta but do not speak Finnish or Swedish are entirely dependent on the English instructions on kanta.fi, further emphasizing the importance of their accessibility.

As kanta.fi is maintained by public authorities, they must comply with the legal accessibility requirements and provide an accessibility statement. At the time I collected my data (March 2025), the accessibility statement for kanta.fi stated that the accessibility had been assessed by an external accessibility expert and that the statement was created based on that assessment (Kanta Services 2025b). According to the text in the statement, it had last been updated in January 2023, though the “Last updated” timestamp at the bottom of the page was dated in January 2025. It is possible that the actual content was last updated in 2023, and the latter date indicates, for example, some minor linguistic edits, though that is just a guess. According to the statement, the service “meets the critical accessibility requirements and largely the A- and AA-level accessibility criteria required by law” (Kanta Services 2025b). The known inaccessible content identified in the accessibility statement includes file links that lack information about the file format on several pages, some inaccessible attachments, and some stock illustrations that have alt texts even though they should not. Regarding the stock illustrations with alt texts, they name the pages that have them: Home, Instructions and services, and Prescriptions. Interestingly, pages titled “Instructions and services” or “Prescriptions” do not exist, and the pages with similar titles did not have any stock illustrations with alt texts.

Clearly, the accessibility statement is not up to date – which, in a way, is to be expected if it has not been updated for over two years.

The kanta.fi accessibility statement includes, besides the legally mandated sections, a list of measures they take to ensure accessibility. In addition to measures such as providing accessibility training to their employees, understandable language is mentioned as well: “We invest in the intelligibility of language. The language of online content is checked prior to publication” (Kanta Services 2025b) – though, ironically, “we invest in the intelligibility of language” is not a very intelligible way to express the idea. Additionally, in an article from 2024, they announce that they have updated some of the instructions for using MyKanta and specifically state that they have “put a lot of effort into the clarity and correctness of language” (Kanta Services 2024c). Based on these statements, as well as Kela’s goal to improve their English communication (discussed briefly in chapter 5.1), one could assume the language used in kanta.fi to be quite understandable and accessible.

## **5.4 Research Material**

The goal of this thesis was to form an understanding of the overall linguistic accessibility on the English kanta.fi website. In order to do that, I selected, evaluated, and analyzed 10 different texts from the site. The selection of the texts was done mostly at random, meaning I did not read the texts before selection, but I did look through the site and tried to gather a diverse selection of data, from instructions to more informative texts, covering a variety of topics.

The kanta.fi website is divided into three main sections based on the intended user groups: Citizens, Professionals, and System Developers. My analysis only includes texts from the Citizens section, since those are aimed at the general public and should therefore be linguistically accessible and, for example, not contain any specialized jargon that the other two might have. The Citizens section consists of four main menus or categories: Instructions, Data usage and security, Current issues, and About Kanta Services. The texts in the Instructions menu are, as the name suggests, mostly instructional in nature and cover topics most relevant to an average user – topics such as performing various actions in MyKanta and using pharmacy services. That is why the majority of the texts in my dataset, seven out of the ten texts, are from the Instructions menu. Additionally, I analyzed one text from the Data usage and security menu and two from the About Kanta Services menu. The pages in the Current issues category were

primarily very short announcement-style texts with information about, for example, upcoming maintenance breaks and ongoing technical issues. Since these pages were purely about the technical aspects of the Kanta Services and not, for example, the use of those services, I determined them to be less important for the average citizen, who is more likely to need help with topics such as using MyKanta or renewing a prescription. Additionally, short texts like these are generally easier to understand than longer texts with more complex structures, which is another reason I decided to not include any of them in this analysis.

The titles of the pages picked for analysis are listed in table 1 below, along with a code I assigned to each page to use to refer to them in the analysis and examples. The first seven are from the Instructions menu, the eighth is from the Data usage and security menu, and the last two are from the About Kanta Services menu.

**Table 1. Research data**

Code	Page title
T1	MyKanta
T2	Log in securely to MyKanta
T3	Health care visits and diagnoses in MyKanta
T4	Prescriptions and prescription renewal in MyKanta
T5	Using pharmacy services
T6	Take medicines and prescriptions with you when you travel abroad
T7	The data saved in Kanta are shown in MyKanta
T8	How can I have a say in the use of my data?
T9	Statistics
T10	Research and knowledge management

All of the texts are informative in nature, and the majority of them are also at least somewhat instructional; some contain clear, step-by-step instructions for doing something, whereas others are more general and descriptive. For each page, I analyzed everything from the page heading to the end of the page, just before the information about when the page was last updated. Elements outside the main content, such as menus and headers, were excluded from analysis.

The data was collected on March 27, 2025. To prevent any updates to the web pages from affecting the data during my analysis, I copied each text into a text document and used these copies for analysis. While copying the texts, I made sure that the contents and styles matched the web page ensuring, for example, that all links were the same as on the web page and all headings were marked as such in the copy. Since visual aspects, such as fonts and colors, are outside the scope of this thesis, the copies did not have to match the visual look of the originals. I did, however, save screen captures from each page, just in case I needed to check any visual details from those.

## 6 METHOD

To understand how linguistically accessible the data was, I created a set of guidelines and used those to perform a heuristic evaluation on the texts. The guidelines are based on a set of criteria I created in my bachelor's thesis (Jantunen 2020); for this thesis, I expanded them to cover a few new topics and made some minor clarifications to the existing criteria. I also decided to call them 'guidelines' rather than 'criteria', as I feel the latter sounds quite restrictive. My intention was not to create strict rules, especially since many aspects of language use are dependent on context, and that is acknowledged in many of the guidelines too.

Next, I briefly explain heuristic evaluation as a method and discuss my reasoning for choosing it. After that, I introduce the guidelines, providing background information and justifications for the additions. Finally, the method of using the guidelines to evaluate the texts and draw conclusions from them is described at the end of this chapter.

### 6.1 Heuristic Evaluation

Heuristic evaluation is a method of assessing the usability of a product with the goal of identifying usability problems (Nielsen 1994, 155). The method originates from the field of usability research and was first described by Jakob Nielsen and Rolf Molich in 1990 (249) as "looking at an interface and trying to come up with an opinion about what is good and bad about it". Originally, it was used specifically to evaluate the usability of user interfaces, but the method has since spread to various fields and products, including texts.

Heuristic evaluation involves a group of experts analyzing the usability of a product based on a list of heuristics, that is, a list of rules, guidelines, or usability principles (Suojanen et al. 2015, 77). These expert evaluators are typically usability specialists, but they might also have expertise in the subject matter (Nielsen 1992, 373). The general recommendation, based on Nielsen's (1992, 376) research, is to use three to five evaluators, each of whom performs the evaluation independently and preferably goes over the product at least two times to ensure as many of the usability problems as possible are found. The outcome of the evaluation is a list of usability problems with references to the heuristics they break (Suojanen et al. 2015, 80). Since heuristic evaluation focuses on finding usability problems, suggestions for fixing those

problems are typically not included, though indicating the heuristic the problems break often in itself provides at least some hint to a solution (Nielsen 1994, 159).

Compared to other usability methods, heuristic evaluation's main advantage is its effectiveness and cost-efficiency – Nielsen (1994, 160) himself has described heuristic evaluation as a “discount usability method”. Performing the evaluation only requires a list of heuristics and a few usability experts with some time to spare. Unlike, for example, usability testing, heuristic evaluation can be performed anywhere, is generally quite fast to complete, and does not require recruiting real users to participate in the testing (Nielsen 1994, 160). However, this lack of real users is also the biggest disadvantage of heuristic evaluation: without the involvement of real users and real situations of use, the results always have some degree of uncertainty, as finding all usability problems via heuristic evaluation is unlikely (Suojanen et al. 2015, 93; Nielsen and Molich 1990, 251–52). That is why Nielsen (1994, 165) always recommends combining different usability evaluation methods to get the best results, and particularly advocates for usability testing, “the most fundamental usability method”, as it provides valuable insight from real users.

Since different projects have different needs, each heuristic evaluation is usually performed with a list of heuristics created specifically for that purpose, though these are often based on existing similar heuristics or other guidelines, such as company style guides (Suojanen et al. 2015, 80–81). It is also common to directly use different guidelines and checklists as heuristics, even if they were not originally designed for that (Suojanen et al. 2015, 78). In the field of translation studies, heuristics are rarely used; however, most translation projects – and other writing projects – typically include revision, which can be seen as implied heuristic quality control (Suojanen et al. 2015, 130). Some heuristics designed specifically for translations do also exist, such as the user-centered translation (UCT) heuristics created by Suojanen et al. (2015, 89). These heuristics were designed for translators to use as an evaluation tool during revision as well as an interactive tool during the translation process. This concept is similar to the idea behind my guidelines, which are intended to function both as heuristics and as guidelines to follow while writing.

## 6.2 Guidelines for Linguistically Accessible Texts

### 6.2.1 Background

The guidelines used in this thesis are largely based on the linguistic accessibility criteria that I created and tested in my bachelor's thesis (Jantunen 2020, 11–13). In addition to the criteria itself, I created a list of questions to evaluate each criterion, which I used to guide my analysis (16–17). While I did not use the terms 'heuristics' or 'heuristic evaluation' in my bachelor's thesis, the criteria – or the list of questions – could easily function as heuristics, and the testing I performed was done in the same style as heuristic evaluation.

The criteria was based primarily on the WCAG 2.1 technique G153 on making the text easier to read and the Selkomittari, a tool created by the Finnish Center for Easy Language to measure the understandability and readability of a text, as well as research and my previous knowledge on the topic (Jantunen 2020, 11). Similar to the guidelines in this thesis, the criteria were intended to function both in evaluating the accessibility of a text and to support writing accessible texts. Additionally, I aimed to keep them as language independent as possible by avoiding any strict and specific rules, but since I am primarily familiar with Finnish and English, the criteria were bound to be heavily influenced by those languages (10). I decided to create my own criteria because the guidelines for linguistic accessibility that I was able to find were, in my opinion, lacking or vague. For example, the WCAG technique G153 focused on word and sentence level rules, lacking any guidance related to the order and amount of information, even though those have been found to be important factors affecting linguistic accessibility, and it was, somewhat ironically, not very accessible itself as it was repetitive, not written consistently, and fairly difficult to find (Leskelä 2019, 118–20; Jantunen 2020, 10).

As the criteria seemed usable and useful, based on the limited testing in my bachelor's thesis, I decided to use them as the basis for the guidelines in this thesis. As mentioned earlier, it is typical to modify existing heuristics to fit the needs of the project, and I did so here too. However, to support the goal that the guidelines could be used outside this thesis as well, the guidelines are not tailored specifically to fit the needs of this study. They are, for example, not only focused on accessible English, despite the data for this thesis being entirely in English. Instead, I simply aimed to improve the criteria by making small modifications and additions to clarify the guidelines and make them more usable, and to cover a few topics that were left

outside the scope of the bachelor's thesis – these are discussed more in the next chapter. Finally, instead of having all the guidelines in one long list, I sorted them into four categories based on the aspect of the text they were focused on: Grammar and Spelling, Vocabulary, Structure, and Information Density and Clarity.

### **6.2.2 Modifications**

The modifications and additions done to the criteria were based on the questions I used to evaluate the fulfillment of the criteria in my bachelor's thesis and the findings from the analysis, as well as additional research done for this thesis. The additions in particular were all aspects of (linguistic) accessibility often mentioned in related research and similar guidelines, such as the WCAG and the UCT heuristics mentioned earlier. To cover these additional topics, I added a total of five new guidelines, more clearly addressing ambiguity and vagueness as well as expanding the criteria to include some aspects of accessibility that are not purely linguistic but that have to do with language use, such as text descriptions for images. Additionally, to improve the clarity of some of the old criteria, I made minor modifications to their phrasing – I am not covering these in more detail as the contents of the criteria were not changed.

Ambiguity, vagueness, and abstract language all negatively affect the linguistic accessibility of a text by making it harder to read and understand and increasing the possibility of misunderstandings (Suojanen et al. 2015, 55). These aspects were implied in many of the old criteria, but to emphasize their importance, I added two guidelines, one under the Grammar and Spelling category and another under Vocabulary. The first addition, guideline G7, is about avoiding ambiguity caused by grammar and formatting, and is based primarily on the WCAG 2 supplemental guidance for cognitive accessibility, specifically the guideline that states: “Use clear, unambiguous formatting and punctuation” (W3C 2021). A common formatting ambiguity that might confuse anyone is the way dates are marked: the practices vary between different languages and countries, and, for example, 1.2.2025 would be understood as February 1<sup>st</sup> in the UK but as January 2<sup>nd</sup> in the US. Decoding ambiguous structures can be cognitively demanding, which can hinder understanding especially for users with cognitive disabilities (W3C 2021). Additionally, screen readers and other assistive technologies can have trouble with ambiguous formatting, for example, Roman numerals might be read as letters (W3C 2021).

The second addition, guideline V5, deals with avoiding abstract and figurative language. The lack of concretization that leads to abstract language is a common problem, and one not typically addressed in similar guidelines despite having a significant impact on the readability and understandability of the text (Kulkki-Nieminen 2010, 44–45; Suominen 2019, 217, 307; Suojanen et al. 2015, 55). However, it is often implied in guidelines such as “avoid jargon”, as was the case with the criteria in my bachelor’s thesis.

The three final additions have to do with the aspects of accessibility that are not purely linguistic but that are related to – and can often be improved with – language: link texts, text alternatives, and referencing elements outside the text the user needs to be able to locate. All of these belong to the final category of Information Density and Clarity – and are, in fact, included in WCAG 2.2 as well. The first of these, guideline I4, is about giving links short, unique, and descriptive names. This is a fairly common guideline, and is in fact already a legal requirement as the WCAG 2.2 level A success criterion 2.4.4 Link Purpose (In Context) states that the purpose of a link must be clear from the link text alone, or from the link text with its programmatically determined link context, unless it would be ambiguous to all users (W3C 2023). The WCAG, as discussed earlier in chapter 2.3, is primarily focused on the technical aspects of accessibility, but this criterion can also be fulfilled with purely linguistic resources by simply writing a clear and descriptive link text, which is why I added the guideline. Clear link texts are especially important to screen reader users, as screen readers often provide the option to list all the links on a page, and that list will not include any text outside the link text that might provide essential information about where the link will lead (Jylhä 2020, 71).

Guideline I5 is about providing text alternatives to all non-text content and, similar to I4, serves primarily screen reader users – although, for example, images can be very helpful to some users and increase accessibility for some, those users who cannot see the images rely on text alternatives to provide the information the image relays to those who can see it (Jylhä 2020, 71). Additionally, this helps those who are relying on machine translation to get the information in a language they can understand: machine translation tools do not usually pick up text from images, but they do translate the text alternatives (Pym et al. 2022, 81–82). This, too, is a level A requirement in WCAG 2.2, as guideline 1.1 Text Alternatives mandates that all non-text content, unless purely decorative, has a text alternative (W3C 2023). These alternatives can be in various formats, such as subtitles in videos or text descriptions of images. The latter are commonly called alt texts, as the description is typically written in the alt attribute of an image.

The alt text is a short description of the image, and it is left empty to indicate a decorative image that a screen reader can ignore. There are also other strategies to provide text alternatives for images, especially when a longer description is needed, but for the sake of simplicity, I am not distinguishing between the different strategies here – the technical implementation of the alternative is not relevant from a linguistic perspective, so I typically refer to them as alt texts or text alternatives. The text alternatives are, naturally, written language, and should follow the same guidelines of linguistic accessibility as any other text.

The final addition, guideline I6, states that when the text refers to something outside its immediate context that the user needs to be able to locate, the reference should not rely solely on visual elements such as size, shape, color, or location. For instance, instructing a user to “click the big red button” is not very helpful to someone who does not see the user interface, and can be problematic for those with color vision deficiencies. Instead, additional ways to identify the correct element should be provided, for example, by including text on the element or using technical solutions to provide screen reader users some link to the element. Similarly, while using only visual features to describe the element is not accessible, providing additional visual details, such as the shape and location, can help those who cannot rely on the color alone. In WCAG 2.2, this aspect is covered with success criterion 1.3.3 Sensory Characteristics (W3C 2023). As with the previous two, this benefits screen reader users, but also, for example, anyone with trouble distinguishing colors, understanding directions, or remembering instructions, particularly if the element that the user needs to locate is far from the text (Jylhä 2020, 71).

### **6.2.3 Linguistic Accessibility Guidelines**

The final linguistic accessibility guidelines, including the additions and modifications described above, consist of four categories and a total of 24 guidelines. I also gave each guideline a code consisting of a letter and a number – the letter comes from the first letter of the category title, and all guidelines under each category are numbered starting from 1. The purpose of these codes is to allow easy reference to the guidelines without having to repeat the entire guideline at every mention, and the letters provide a clear indication to the aspect of text the guideline concerns. The guidelines are listed in tables 2–5 below, each category in their own table.

**Table 2. Grammar and Spelling**

Code	Guideline
G1	Consistently follow the grammar and spelling rules of the language or variant you are using.
G2	Use the active voice unless there is a specific reason to use the passive voice (for example, in situations where the subject is unknown or irrelevant).
G3	Use simple sentence structures. For example, in English and Finnish, prefer the Subject-Verb-Object structure.
G4	Use verb tenses consistently.
G5	Use verbs to express what is being done. For example, in English, instead of saying “They performed an investigation” say “They investigated”.
G6	Make clear pronoun references. Repeat the subject if it makes the text easier to understand.
G7	Avoid ambiguity. Use clear, unambiguous sentence structures and provide enough information. For example, write out dates instead of using just numbers.

**Table 3. Vocabulary**

Code	Guideline
V1	Use common, everyday words. Avoid professional jargon and other terms with specialized meanings. Replace long or unfamiliar words with shorter, more common words.
V2	Explain words that might be hard to understand if it is necessary to use them.
V3	Avoid using abbreviations. Explain all uncommon abbreviations and acronyms the first time you use them.
V4	Use names, terms, and labels consistently. Use only one name to refer to the same thing throughout the text.
V5	Avoid abstract and figurative language. Provide concrete examples and link abstract ideas with real world things the user might be familiar with.

**Table 4. Structure**

Code	Guideline
S1	Structure the text logically. For example, follow the order in which things happen or are done.
S2	Divide the text into logical sections. Use subheadings to help the reader understand the structure of the text.
S3	Use descriptive headings. Make sure the heading matches the content of the paragraphs below it.
S4	Add cohesion by indicating logical relationships between phrases, sentences, paragraphs, or sections of the text. For example, in English, use cohesive devices such as “additionally”, “but”, or “however”.
S5	Tell the most important thing first.
S6	Use lists instead of sentences that contain long series of words or phrases separated with commas.

**Table 5. Information Density and Clarity**

Code	Guideline
I1	Tell the user everything that they need to know, but nothing more. If necessary, direct the user elsewhere for additional information.
I2	Cover only one idea per paragraph. Start a new paragraph whenever you introduce a new topic or viewpoint.
I3	Cover only one point per sentence.
I4	Give links short, descriptive, and unique names, and use those names consistently. If the link leads to a file or material in a different language, include that information in the link text.
I5	Provide text alternatives for non-text content such as images, unless they are purely decorative. Follow accessible writing guidelines. The text alternatives should serve the same purpose and convey the same information as the non-text content.
I6	Do not rely solely on visual elements like size, shape, color, or location when referring to something outside the immediate context that the user needs to be able to locate.

### 6.3 Research Method

To analyze the data, I performed a heuristic evaluation using the guidelines introduced above as heuristics. While heuristic evaluation is typically used to find, and then fix, problems in the evaluated product, and it is often used as a part of an iterative development process, I used it as a tool to perform a textual analysis on a selection of texts and thus form an overall idea of the state of linguistic accessibility on the kanta.fi website. Since the goal was not to find all usability problems – or in this case, linguistic accessibility problems – it was justified to not follow the general recommendation of having at least three evaluators. Instead, I performed the analysis alone. Of course, I alone might not have been able to find all the problems, and it is possible that I missed some issues. However, it is unlikely to have a significant impact on the results, as even a partial list of problems can reasonably be used to draw conclusions about the overall state of linguistic accessibility on the site.

To perform the evaluation, I first went through each text with the guidelines next to me and marked all the problems I noticed to form a general overview of the texts. After that, I went through each text multiple times again, focusing on one guideline at a time, to find any issues I missed in the first round. During this process, I noticed some problems with consistency not just within individual texts but also between the texts. To get a better understanding of that, as the final step of the evaluation, I looked through all the texts together and paid close attention to the consistency between them. I then analyzed the results of the evaluation to form an understanding of types of problems present in the texts and thus the overall state of linguistic accessibility. Finally, since some of the guidelines deal with visual aspects and benefit primarily screen reader users, I tested a few things using Windows Narrator to get at least some idea on how certain aspects work with a screen reader. This testing, although limited, gave me at least some insight into the way screen reader users experience these pages. Of course, for more reliable results, it would be best to do more extensive testing on these matters as well.

I believe that since I performed the evaluation with such scrutiny, and because I have some expertise in both heuristic evaluation and especially linguistic accessibility, I was able to find the majority of, or at least the most significant of, the problems in these texts. Though it is likely that I did not find every single issue as I was the only one performing the evaluation, I trust that my findings are enough to form an overall idea of the state of linguistic accessibility on the kanta.fi site.

## 7 ANALYSIS

In this chapter, I present the key findings of the evaluation and analysis. Since I used heuristic evaluation as a tool for the analysis, the initial findings were focused on specific guidelines. However, because the different guidelines within a category all relate to the same general textual aspect and therefore inevitably affect one another, the analysis focuses primarily on the categories and the bigger themes evident in the findings, rather than covering each guideline separately. It is also worth noting that the different categories, and the guidelines within them, also affect one another, and similar themes can be seen across all categories. That is to be expected, as language use is rarely simple and straightforward, and the findings reflect that.

Next, I discuss each category separately and introduce the different types of problems found related to the guidelines in that category, as well as provide examples from the data to support my findings. Each subchapter is further divided into sections based on the most prevalent themes or findings across the category. Even though the analysis does not focus on individual guidelines, when discussing the findings and examples, the related guideline is indicated in parentheses. Additionally, I analyze the findings and their severity and impact on the linguistic accessibility of the texts in order to draw final conclusions about the state of linguistic accessibility on the kanta.fi website. Finally, I provide a summary and overview of the findings in the different categories at the end of this chapter.

### 7.1 Grammar and Spelling

The first category of guidelines deals with aspects of grammar, spelling, and formatting. Since grammar varies between languages, and because some flexibility in the grammatical structures used in a text is necessary to ensure a smooth and coherent flow, the guidelines do not include any strict grammar rules. Instead, they focus on consistency, clarity, and simplicity on a more general level. Some of the guidelines are concerned with specific grammatical features though, such as the passive voice (G2) and nominalizations (G5), as those can have a significant impact on linguistic accessibility regardless of language. Still, sometimes, for example, the use of the passive voice is justified and the more accessible solution – it is important to consider each guideline in the context of the text and evaluate how it should be followed rather than assume strict conformance at all times. After all, they are simply guidelines, not strict requirements.

This is particularly true for the grammar-related guidelines, as grammatical aspects tend to naturally vary in almost all texts.

### 7.1.1 Consistent Grammar

While some flexibility is natural and expected in fluent and accessible writing, consistency is also necessary to ensure linguistic accessibility, which is why it is included in almost all guidelines across different categories. Even when not explicitly mentioned, consistency is always implied at least in the sense that the guidelines should be followed consistently. Therefore, it is no surprise that consistency was a prevalent theme in the findings across categories as well – and the Grammar and Spelling category is no exception.

Many of the grammatical inconsistencies (G1) found in the data were minor and unlikely to significantly hinder understandability, as well as easy to correct; for instance, the serial comma, or Oxford comma, was not used consistently. Similarly, while contractions were primarily not used, the negative was contracted on a handful of occasions: *can't* instead of cannot and *don't* instead of do not. While the use of contractions itself can negatively impact linguistic accessibility, it is unlikely that the inconsistency in its use would have a notable effect on the understandability of a text. However, if the variance is significant and frequent, the likelihood of these types of minor inconsistencies affecting the overall linguistic accessibility of a text increases. Additionally, the accumulation of smaller issues like these, especially in combination with other potential problems, does negatively impact linguistic accessibility.

An example of the type of grammatical inconsistency that is more likely to cause issues with accessibility is the inconsistent use of verb tenses, covered by guideline G4. All the evaluated texts were written primarily in the present tense, with other tenses used when necessary – for example, the past tense was used when referring to a previous appointment the user might have had. However, something that stood out in the data was the overuse of the future tense, usually without a justifiable reason. Example 1 is a typical example of this phenomenon: while the sentence is not grammatically incorrect or necessarily even difficult to understand, the use of the future tense is not justified, especially since the surrounding text was in the present tense. This type of inconsistent use of verb tenses also has the potential to confuse the user about when the actions happen. A similar inconsistency, albeit not in verb tenses, is illustrated in example 2. The noun *data*, used throughout all the evaluated texts, can be considered either singular or

plural (OED 2024), but the decision on which to use should be consistent (G1). In these texts, it was more often viewed as plural, but on multiple occasions the singular was used instead, sometimes even mixing the two forms in consecutive sentences, like in example 2.

- (1) The person prescribing your medicine *will save* your prescription electronically in Kanta. (T5, emphasis added)
- (2) it is important that your *data is* visible to the professional treating you and managing your affairs. You can access MyKanta to manage how your *data are* used (T8, emphasis added)

### 7.1.2 Ambiguity

Another overarching theme in the guidelines is the goal to avoid ambiguity. This is directly addressed in guideline G7, but violations of other guidelines can also contribute to ambiguity. Overall, ambiguity caused by grammatical structures, spelling, or formatting was not frequent in the data. However, ambiguous sentences due to word choice or lack of information were more common – those are covered in the next chapter. Ambiguous structures, such as the one in example 3, can be interpreted in multiple ways and thus lead to misunderstandings. This example is the heading of an info box, and it is trying to communicate that you should determine what policies apply in your destination country before travelling there. However, due to the use of *in* and the way the sentence is structured, one could also understand it as an order to determine the policies when you are physically in the destination country. While that does not logically make sense with the rest of the sentence (as you cannot do something in a country before you travel there), it can be ambiguous and hard to understand for users with cognitive difficulties.

- (3) Determine the policies in your destination country before you travel (T6)

Ambiguity can also be caused by, for example, unclearly written dates or times, such as the use of the 12-hour clock in example 4. It is important to keep in mind that there are no universal definitions for what is ambiguous – as with many other aspects of (linguistic) accessibility, individuals' needs vary, and the audiences must always be considered. For instance, if this was written primarily for an American audience, the 12-hour clock would be the more accessible solution, as that is the format commonly used in the US. However, in Finland, the 24-hour clock is used, and it is more common worldwide (Costa 2025), so it would likely be the more accessible choice here, though the use of *a.m.* and *p.m.* does help avoid misunderstandings. Additionally, the spaces in the phone number are not non-breaking, and just like in example 6 below, the phone number spread onto two rows, at least with my browser and display settings. This can be troublesome for users with cognitive difficulties as they might not notice or

understand that the number continues onto the next row. Finally, example 4 also demonstrates incorrect punctuation (G1) with the mix of longer en-dashes and shorter hyphens: in general, the hyphen is used primarily to connect compound words, while the en-dash should be used to indicate a span or a range, such as the days and times in the example.

- (4) For help with using Suomi.fi identification, please contact Public Service Info: tel. +358 295 000 (Mon-Fri 8:00 a.m.–9:00 p.m., Sat 9:00 a.m.–3:00 p.m.) (T2)

Example 5 demonstrates a variety of linguistic accessibility issues, including ambiguity. First, the formatting of the list obfuscates the meaning of the sentence, as it is unclear whether both of the listed items are required or if they are presented as alternatives (G7). This was the only example of a situation like this in these texts though, so while this is very unclear and therefore linguistically inaccessible, it does not seem to be a recurring issue on the kanta.fi website. However, the last sentence in the example demonstrates a more common ambiguous structure: the phrase “will be provided” gives the impression that the information will be given automatically, but typically these have to be specifically requested. Additionally, the list is grammatically wrong (G1), as a single list item should only contain one sentence, particularly when it is meant to complete the introductory sentence as is the case here. In general, the lists used across the evaluated texts did not consistently follow the same punctuation and formatting rules, although, as mentioned earlier, that does not have a major impact on linguistic accessibility. The type of formatting in this example is especially problematic, as it can make it harder to decode the overall meaning of the sentence and understand which parts relate to what. Furthermore, example 5 also violates guideline I1 about providing enough information (discussed more in chapter 7.4), because the items listed are not elaborated on later and they can be hard to understand – it is unclear, for example, what exactly *diagnostic information about your illness* means in this context. Finally, the last sentence in the list violates both guidelines G2 and G3 with its unjustified use of the passive voice and the unnecessarily complicated *be done by someone* structure – both recurrent issues in the research data.

- (5) If you wish to take prescription medicine with you when you travel abroad, take with you
- a signed copy of the prescriptions in English. You can get a copy from a pharmacy or the person issuing the prescription.
  - a summary of your medication and the justification for it, diagnostic information about your illness, or your medical records. These will be provided to you by the person prescribing the medicine. (T6)

The unnecessary use of the passive voice (G2) can significantly impede linguistic accessibility by increasing ambiguity and complicating sentence structures. However, the passive is not always an inaccessible choice, and it should not be avoided in all situations. For instance, when

one wants to keep the attention on the user and express something that happens in the background or is done by an unknown or irrelevant actor, using the passive is justified and can even improve linguistic accessibility. On the other hand, whenever the user is instructed to do something, the active voice should be used, as it helps the user identify what they are responsible for and what they need to do. In the research data, the active voice was always used when directly addressing the user, and the use of the passive was mostly justified – as is illustrated in example 6. Still, the findings include multiple instances where the actor was mentioned yet the passive was still used. Many of these would have been clearer in the active voice; for instance, the last sentence in example 5 could have been rephrased to “You will get these from the person prescribing the medicine”, which both brings the focus to the user, the *you*, and is structurally clearer and simpler.

(6) For example, if you attended an appointment at a health centre, corrections and additions to MyKanta entries related to that visit *will also be made* there. (T3, emphasis added)

### 7.1.3 Grammatical and Sentence Structures

As can be seen, the unnecessary passive can also complicate sentence structures (G3), particularly when the *be done by someone* structure is used to indicate the actor. These types of complicated structures, especially in combination with long sentences, can be harder to read and understand, thus negatively affecting linguistic accessibility. Although the evaluated texts mostly consisted of short and fairly simple sentences, each text also contained some complex structures. For instance, some sentences in the data were complicated by the use of long and complex phrases such as *due to the fact that* (T3) and *the extent to which* (T7, T8). Another instance of a complex structure is illustrated in example 7, which is complicated by the additional information separated by commas in the middle of the sentence. If the information was deemed necessary, the sentence should have been ordered so that the added information was not presented in the middle of it. However, in this specific case, the extra information adds little to no value, as it is not new to the user: managing your own prescriptions was covered before this sentence. While these types of additions can be used to increase cohesion, it should not be done at the expense of clarity and simplicity – besides, the *also* in the sentence is already acting as a cohesive device.

(7) In MyKanta, in addition to your own prescriptions, you can also request a prescription renewal on behalf of your child or an adult loved one. (T4)

Additionally, on some occasions, the sentence structure was complicated (G3) due to the lack of parallelism in the structures, as demonstrated in example 8 below. Typically, elements linked to each other with a connective of some kind should be grammatically similar to facilitate readability and understandability. Example 8 does not follow this idea of parallel structures, as the two examples it gives are structurally different: *the quantity of data* and *on how the services are used*. Additionally, the examples listed should be grammatically coherent with the rest of the sentence (G1), but here, if we remove the first example, the sentence reads *on things like on how the services are used* – the repetition of the preposition is both ungrammatical and confusing. This example also illustrates another grammatical structure that can unnecessarily complicate sentences: the nominalization of verbs (G5). Of course, nominalizations are not always the inaccessible choice, but when they are used instead of a verb to express action, they can be harder to understand. Here, the use of *for publication* is an example of that, and the same idea could have been expressed more clearly with, for example, *and publish those*. Another, more serious, example of a nominalized structure is the use of *carries out the supervision* in T10. However, overall, nominalizations were rarely used in the research data, so the problem was not frequent.

- (8) We compile statistics on these data for publication in Kela's Info Tray, where you can find information on things like the quantity of patient and client data stored in Kanta and on how the services are used in social welfare and health care. (T9)

In addition to the above-mentioned issues, example 8 is also quite long and could have been clearer if split into two separate sentences – this relates to guideline I3, which is discussed more in chapter 7.4. Finally, the sentence starts with the pronoun *we*, though who or what it refers to is not explicitly clarified anywhere in the text (G6). One could assume that it refers to the Kanta Services, since the text is on their website, but a user with cognitive difficulties might not be able to easily make that connection. Unclear pronoun references are yet another issue that can cause ambiguity in sentences. However, this example is the only one with an unclear pronoun reference in the data, so the issue was not frequent. Although, in a few instances, the repetition of a noun instead of a pronoun could have clarified the sentence, overall pronouns were used quite sparingly and repeating the noun was more common – which, rather counterintuitively, decreased linguistic accessibility in some cases. Sometimes, the use of a pronoun instead of a noun could have made the overall sentence a bit clearer and easier to understand. This is illustrated by example 9, where the repetition of *personal data* makes the sentence longer and harder to process, especially since the two mentions are so close to each other. Replacing the second occurrence of *personal data* with *it* could have improved the sentence's linguistic

accessibility. Although, the meaning of the sentence would have remained vague, as it does not provide enough information to clearly explain pseudonymization, nor does it provide links to further information about the topic, thus violating guideline I1 (see chapter 7.4).

- (9) Pseudonymisation refers to the processing of personal data in such a way that personal data can no longer be linked to a specific person without additional information. (T2)

#### 7.1.4 Spelling

The final aspect I want to bring up here is spelling; linguistically accessible texts should use correct and consistent spelling (G1). Issues with spelling were uncommon throughout all the texts – for instance, only one spelling mistake was found: in T10, the word *health* had been misspelled as *heath*. Mistakes such as this one, where the incorrectly spelled word is still a real word, are fairly common, as spell checker software are likely to not identify them as incorrect. Similar to the aspects of grammar, most of the findings were related to consistency rather than correctness, though inconsistent spelling was less common than inconsistent grammar. One example of inconsistent spelling is the word *health care*, which was sometimes spelled as one word, *healthcare* – though inconsistent and thus a violation on guideline G1, its impact on linguistic accessibility is minor.

Additionally, the variant used was consistent (G1) with just one exception: in T9, the words *inquire* and *enquiries* were used in close proximity to one another. One of these is a noun and one a verb, but the spelling should be consistent – *inquire* and *inquiry* are more common in American English, while *enquire* and *enquiry* are more common in British English, the variant primarily used on the kanta.fi website. Both are correct and although using *inquire* and *inquiry* in a text otherwise written in British English would likely not cause significant accessibility issues, it would be best to adhere to one spelling to limit the chances of misunderstanding. Although, in this case, understandability is more likely to be hindered by the uncommon word choice; replacing these with more common alternatives, such as *ask* and *question*, would be more accessible and more in line with guideline V1 discussed in the next chapter.

## 7.2 Vocabulary

The guidelines in the Vocabulary category have to do with the clarity, understandability, and consistency of individual words and phrases. Even though a single word is just a small part of

a sentence, let alone an entire text, just one poorly chosen word can have a significant negative impact on the text's linguistic accessibility. Preferring common and frequently used words and avoiding figurative language helps ensure that users with limited language skills and vocabulary understand the contents of the text. Additionally, just as with aspects of grammar, using words consistently and avoiding ambiguous phrasing can facilitate understanding, particularly for users with cognitive difficulties or limited language skills.

### 7.2.1 Consistent Vocabulary

As noted before, a significant number of the findings across categories had to do with consistency. When it comes to vocabulary, the most obvious way inconsistencies manifest is by using multiple names or labels for the same idea – this is reflected in guideline V4, and the texts did have multiple findings violating this guideline. Some of these findings were less severe, such as the multiple verbs used in T2, T3, and T4 to indicate interacting with buttons in the MyKanta user interface: *click*, *select*, and *press* were all used in similar contexts. These are all common words and using them all in the same text does not cause a major linguistic accessibility issue, though it would be best to adhere to one to ensure clarity and understandability. Another example of a more minor inconsistency is the use of *Kanta* and *Kanta Services*, demonstrated in examples 10 and 12. There did not seem to be any clear logic to using one over the other; for example, *Kanta* was not always used after *Kanta Services* had been mentioned – if it would have been, one could have assumed *Kanta* to be short for *Kanta Services*. It could, of course, still be so, but it was not made clear in the texts. In fact, after reading through all of the research data, it is still somewhat unclear to me whether the two differ somehow and if they do, what the difference is. Regardless, the two are clearly related, so the likelihood of this inconsistency causing major issues with understanding the overall text is not too significant.

- (10) When you interact with health care or social welfare services, data is saved in the Kanta Services by a doctor, nurse, social worker, or other professional. Pharmacies also record information in Kanta. You can browse this data yourself in MyKanta. (T7)

Example 10 also demonstrates another, more significant, inconsistency: the interchangeable use of *data* and *information*. These were used interchangeably in all the evaluated texts, and often both were used even within the same paragraph, as in the example above. *Data* and *information* are commonly used interchangeably, especially in non-specialized contexts, but there is a difference: *data* refers to the raw facts and figures that themselves do not carry meaning, while

*information* refers to data that has been processed to provide meaning and context (Henry et al. 2007, 600). Although kanta.fi is not the type of specialized context that would warrant making a difference between the two words, using them both to mean the same thing in the same text can cause confusion, especially to users who are not so familiar with the words and might, logically, expect them to mean different things. In addition to these two, in T10, the words *data content*, *data material* and *material* were used a total of nine times, all in contexts where only *data* would have sufficed and helped avoid potential confusion.

Finally, I want to bring up two inconsistencies related to the services and service providers mentioned in the data. First, the Kanta Services, as mentioned before, are for both health care and social welfare data and, quite often, both were mentioned in the texts. However, on multiple occasions, only health care was mentioned, even if social welfare was mentioned earlier and the context clearly applied to both, such as in example 11: the second sentence only mentions *health care organisations*, despite *social welfare* being mentioned alongside *health care* in the first sentence. This has the potential to confuse users and lead to misunderstandings, as some might incorrectly interpret the lack of a mention about social welfare to indicate the sentence only pertains to health care. The risk of misunderstandings is exacerbated by the fact that occasionally the sentence does only apply to health care, as, for instance, prescriptions are not used in social welfare services. On the other hand, *social welfare and health care* is a long modifier and can thus complicate sentence structures and make them harder to read, so the need to shorten the phrase is understandable. However, leaving out one part of it is not an ideal solution, as it can lead to more issues than it solves. Instead, using a collective word to cover both contexts or leaving out the modifiers from later mentions could have improved accessibility. In the example below, for instance, the latter mention of *health care* could have been omitted without significantly altering the meaning – when neither are mentioned, the user is more likely to connect *organisations* to both *health care* and *social welfare* mentioned in the previous sentence.

- (11) Health care and social welfare services professionals can access the data through their client or patient information systems. The professional treating you can also use Kanta to view information that has been recorded about you in other health care organisations, such as another wellbeing services county or a private clinic. (T8)

The final inconsistency is related to what the services and service providers were called throughout the data. For the most part, *service* was used to refer to the services on a more general level, but when referring to a service provider, such as a health care center, a variety of words were used in addition to *service provider*: *unit*, *provider*, *organisation*, and *operator*.

Occasionally, *services* was used in this context as well. This issue is demonstrated by example 12, where all three mentions of a provider use a different word. Using all these different words can lead to confusion and hinder linguistic accessibility if the user does not understand the potential difference – or lack of one – between the words. Although some variety in words might occasionally be necessary, for example to differentiate between more general references to service providers and more concrete organizations in examples, the inconsistent use of six words is not justified.

- (12) All public health care *units* use Kanta, but some private *operators* do not yet use the Kanta Services. Check which private health care *organisations* use Kanta. (T3, emphasis added)

### 7.2.2 Abstractness

The vocabulary category also contained several findings causing ambiguity, primarily by the use of abstract, poorly defined words and phrases (V5) – apart from a few examples, such as *around the clock* in T9, the texts were void of figurative language though. A common abstract phrase used in half the evaluated texts was *managing your affairs*. Alone, it does not mean anything specific, as both *manage* and *affairs* can refer to a variety of different things. This ambiguity is emphasized by the fact that *health*, *health care* or *social welfare* were never mentioned with the phrase, despite it being exclusively used to refer to managing affairs related to health and social welfare. Though one could argue that it should be clear based on the general context of the texts (Kanta Services), it might not be obvious to someone with cognitive difficulties. Whereas in some findings its meaning could be relatively easily deduced from the surrounding text, in others it was far too vague, as in example 13: what *managing your affairs* covers is left ambiguous, especially since *treating you* is mentioned separately – what other affairs the professionals manage is unclear. This is particularly problematic since the sentences are about who can see your data and when, and that cannot be unequivocally determined based on the information provided.

- (13) Your data can only be viewed and accessed by the health care or social welfare professionals who *manage your affairs*. The data are only used when it is necessary in order to treat you and *manage your affairs*. (T7, emphasis added)

Another example of abstractness is the use of *within a reasonable time* in T3 – without any additional clarification, this gives no useful information to the user. Similarly, in example 14, the use of *usually* implies that there are situations where the statement is not true, but what those situations might be is not explained, rendering the meaning ambiguous. In addition to the

abstract phrases, examples 13 and 14 could also be seen to be inaccessible due to not providing enough information, thus also violating guideline I1 (covered in chapter 7.4).

- (14) You will usually receive a reimbursement for your medicine directly at the pharmacy, when the pharmacy checks your right to the reimbursement electronically with Kela. (T5)

### 7.2.3 Word Choice

The use of uncommon or specialized words (V1) can also contribute to abstractness and ambiguity, as is illustrated by the emphasized phrase in example 15. While *preparation* can be used to refer to a product, usually specifically a medicinal substance, that has been made ready for use, this meaning is quite specialized and used primarily in the fields of medicine and chemistry (OED 2025). Much more commonly, *preparation* is understood as the act of preparing for something. Unless one is familiar with this specific use of the word, the meaning of the sentence is difficult to decode, despite the examples it contains. Additionally, the structure of the phrase is unnecessarily complex (G3), further hindering its linguistic accessibility.

- (15) the prescribed medication is *a preparation intended to be taken as a specific course of medicine*, such as antibiotics or cough medicine (T4, emphasis added)

In general, when the use of uncommon words was necessary – that is, for example, when more common alternatives do not exist – they were explained (V2) or a link to more information was provided (I1). However, some, primarily medical or health-care-related, terms were not explained even though they might be unfamiliar to an average user and especially to a user with cognitive difficulties. Some examples of these were *living will* in T1, *special permit product* in T4, *narcotics or central nervous system agents* in T6, *declaration of intent* in T8, and *regulatory tasks* in T10 – as well as the specialized use of *preparation* covered above. Besides these terms, the majority of the words used in the data were common and frequent, though all texts also had occasions where more complex or uncommon words were used without a justified reason (V1). The following examples were all used in more than one text and could have easily been replaced with a more common alternative, indicated in parentheses after the examples: *rectify* (correct), *receive* (get), *provide* (give), *purchase* (buy), *obtain* (get), and *quantity* (amount or number).

The final guideline not yet covered in this chapter, V3, has to do with abbreviations and acronyms. Overall, only two common abbreviations were used in the data: *e.g.* and *i.e.* Although common and unlikely to significantly impede the understandability of the text, the abbreviations

were not necessary. The more common one of these, *e.g.*, was used three times in the data, whereas the non-abbreviated forms *for example* and *for instance* were used a total of 15 times, so the abbreviation was not consistent. The other abbreviation, *i.e.*, was used a total four times: twice to explain Findata (as in example 16) and twice to clarify the meaning of some phrase (as in example 17) – none of these were strictly necessary. When explaining Findata, the abbreviation could have been removed without affecting the meaning or grammar of the sentence, and in other cases it could have been replaced with, for example, *meaning*. To ensure linguistic accessibility, it is always best to avoid abbreviations altogether when possible.

- (16) please apply for permission for the secondary use of the data from Findata, i.e. the Finnish Social and Health Data Permit Authority (T9)
- (17) the prescription is subject to dose dispensing, i.e. the pharmacy supplies the medicine to the customer dispensed into single doses. (T4)

Additionally, some acronyms or short names for organizations were used in the texts, and they were not always explained. It seemed like the organizations that could be considered ‘common knowledge’ among the Finnish population, such as Kela, were typically not explained, whereas less commonly known organizations were elaborated on, as was done with Findata in example 16 above. Although this approach might make sense when writing to a Finnish audience who is familiar with the most common Finnish organizations (though users with cognitive difficulties might still benefit from more information), the differences between the audiences of the original and the translation should be considered in the translation process. The most significant audience of the English translation, as discussed earlier, consists of migrants who might be unfamiliar with the Finnish organizations and their tasks, and could therefore benefit from clarifications. Additionally, when some explanation was provided, it was not always clear; in T10, the Finnish Institute for Health and Welfare (THL, Terveystieteiden ja hyvinvoinnin laitos) was referred to twice: first with just the acronym *THL*, and a few paragraphs later with only the full name *National Institute for Health and Welfare*. Both refer to the same organization, but a user unfamiliar with it cannot determine that from this text. Typically, all acronyms should be explained the first time they are used in each text, if it is necessary to use them, and the same could be said about organizations in general.

### 7.3 Structure

The third category of guidelines has to do with the structure of the text: is the order and division of information logical, are headings descriptive and accurate, and is the text cohesive and easy

to follow. Unlike with the previous two categories, topics of consistency and ambiguity are not explicitly mentioned in these guidelines. Nevertheless, the implied requirement to follow the guidelines consistently is present. Additionally, an illogical structure can contribute to ambiguity, though that is not explicitly covered here. As has been mentioned earlier, many of the guidelines affect each other and even overlap, and common themes can be observed in all categories.

### 7.3.1 High-Level Structure

In general, the overall structure of the texts made sense: looking at the different sections and their subheadings reveals that the majority of the texts present information in a logical order, for example, following the order that something is done in or moving from general to specific (S1). Additionally, the division of information within the texts seems logical (S2). For example, T2 “Log in securely to MyKanta” starts with instructions for logging in, then explains how to log out, and finally covers some frequently asked questions related to logging in. Similar structures were used in T3 and T4, which also contained step-by-step instructions for using MyKanta. Another example of a logical structure in a different style of text can be seen in T8 “How can I have a say in the use of my data?”, which contains the following subheadings: *Consent to data sharing, Denials of consent, Consent to the transfer of a Patient Summary to European countries, Restriction of use of data for research purposes, and Further information* – as can be seen, the text starts with general topics and then moves onto more specific issues.

It is worth noting that T8, as well as T1, are top-level pages of sections, and their primary goal is to provide brief introductions and links to the topics covered in more detail in the section’s subpages. Formulating a logical structure in a collective page like this can be difficult, which can explain why in T1, although the structure is not necessarily illogical, the subheadings do not seem to form a coherent whole either. Though the lack of coherence can be problematic from the perspective of linguistic accessibility, it cannot always be avoided in situations like this and is likely not a significant issue considering the context. Interestingly, T7 and T9 are also top-level pages of sections, but they do not have a similar introductory function as T1 and T8. Instead, they provide information that is not available in the subpages and act more as complete standalone pages rather than just collections of links. In fact, none of the links to other kanta.fi pages included in T7 and T9 lead to their subpages. The overall structure of the website is not within the scope of this thesis, so I will not cover this in more detail, but the differing use

of these top-level pages is an interesting finding nonetheless – especially since T1 and T7 are both from the Instructions menu.

T7 is also an example of a text with several structural issues, starting from its title *The data saved in Kanta are shown in MyKanta*, which seems more like a statement than an informative title and leaves the user unsure about what to expect from the text (S3). Additionally, the first heading of this text – *What are the benefits of Kanta?* – does not match the title, nor does it fully fit in with the rest of the text’s contents. Finally, the overall structure of the text was not very logical: the text is about the data that is stored in Kanta, but what kind of data is stored was the last topic covered on the page (S1, S5). Before that, topics such as who can see the data, how long the data is stored, and how to consent to data sharing were covered. From a user’s perspective, it would make more sense to first introduce the type of data that is stored in Kanta, as that is relevant information to know when considering, for example, who can see the data. T10 shared many of these structural issues with T7, which is interesting as the pages are from two completely different menus on the website: T7 is under Instructions while T10 is under About Kanta Services. These examples demonstrate how the different aspects of structure, and the different guidelines, affect one another and how issues in one can lead to issues in another, eventually resulting in an overall confusing structure and an inaccessible text. However, the issues with T7 and T10 are an exception, and the majority of the texts were structured logically and the issues within them were less severe and less frequent. For the most part, titles were descriptive (S3) and texts were divided into logical sections with their own subheadings (S2), making them easy to scroll through and navigate.

Another example of a different kind of problem in the overall structure of a text (S1) is the use of the *Read more* sections in T5 and T6. These sections provide a list of links that lead to pages with more information on topics covered in the texts, which in itself is a useful strategy and can increase accessibility. Similar sections were used in the majority of the evaluated texts, though typically they only contained links to additional information related to the topic of the text, whereas links related directly to the topics at hand were within the suitable sections. In T5 and T6, however, almost all the links were in the *Read more* section, even though some of them could have been more helpful near the paragraphs where the topics were first mentioned. For instance, in T5, the link to an explanation of the direct reimbursement system could have better served the needs of the users if it had been placed directly after the paragraph that covered getting a direct reimbursement at a pharmacy, rather than at the end of the page where it might

be missed. To keep the text focused on its topic (using pharmacy services), not including the details of the reimbursement system in the text is a logical choice, but as that information could be necessary for users not familiar with the system, including a link is essential to ensure linguistic accessibility – and the link is more easily found when it is placed near the context it relates to. This is, again, also related to guideline I1 (discussed in chapter 7.4) which, in addition to providing enough information, covers directing the user elsewhere for more details.

The links within the different sections were most often presented in a list format as the last thing within the section before the next subheading. Presenting links (and other information) in lists is a common and helpful strategy that can increase accessibility (S6). Sometimes, these lists only had one link in them, which is grammatically incorrect, as one item cannot constitute a list (G1). However, these single-item lists could be justified with the consistent identification of links: when links are always introduced in a similar manner, it is easy for the user to identify and find the links based on the visual layout. This justification is only valid if the same logic is followed consistently throughout the texts though, which was not the case here. For the most part, lone links at the end of a section were in a list, but there were a few occasions where the link was its own paragraph instead. To ensure consistency and allow users to easily identify and find links, it could be argued that the most accessible solution would be to present all links to further information in lists, even if some of those lists would have just one item. This, of course, only benefits sighted users – for a screen reader user, these single-item lists could be somewhat bothersome, so they should not be used excessively, or they should be technically implemented so that a screen reader does not announce them as lists.

### **7.3.2 Headings**

As discussed previously, informative texts, such as the ones in my dataset, are usually read to perform some action or reach some goal, and users are generally only interested in the information that helps them do that. Therefore, the texts are rarely read from beginning to end – instead, users glance through the headings to find the information that concerns them. That is why it is important to use descriptive headings (and titles) and ensure they accurately describe the content below them (S3). The headings used in the data did largely follow this guideline, with only a few exceptions. For example, in T2, bookmarking the kanta.fi website was covered under the heading *How to log out of MyKanta* despite the two topics not being related. Instead of modifying the heading to include bookmarking, it would have been logical to separate the

bookmarking paragraph into its own section, perhaps, for example, some kind of a hint box. A different kind of mismatch between the heading and the contents is illustrated in example 18, which shows a heading and the first sentence after it. As can be seen, the point of view changes: the heading implies that the user is acting on behalf of someone else, while the content below is written from the perspective of someone else acting on behalf of the user. While the topic itself remains the same, the change in perspective can disrupt understandability for someone with cognitive difficulties.

(18) Acting on behalf of another person

If someone else collects your medication from a pharmacy, they must present your Kela card or your prescription's printed patient instructions. (T5)

Many of the texts also included collapsible sections that provided additional information, for example, in the form of frequently asked questions and their answers. For the purposes of this thesis, I treated the headings of those sections the same as any other headings, because they do essentially act as such. In fact, since the content in those sections is initially hidden, the users are even more dependent on the headings to determine whether the content interests them, further accentuating the need for descriptiveness. In a few questions-and-answers-type sections, the heading (question) did not fully match the content (answer). For instance, in T4, the answer to the question *What prescriptions can I renew in MyKanta?* included information about why a doctor might not renew the prescription. Since that information does not directly relate to the question, a user looking for information on why their prescription was not renewed might not be able to locate it.

Example 19, on the other hand, demonstrates the opposite: the question is too broad for the answer. There are multiple methods of identification, but only issues with online banking codes are mentioned in the answer. Either the answer should include information about the other methods as well, or the question should specify that it is about online banking codes only. The answer is problematic in other ways as well, primarily due to lack of information, as the *required extra e-identifier* is not explained, nor is it mentioned anywhere else in the text. Additionally, the answer also lacks cohesion between the two sentences, and the word order in the question is grammatically wrong: it should start with *Why can't I* instead of *Why I can't*.

(19) Why I can't identify myself?

If you have online banking codes but are still unable to identify yourself in various public administration e-services, it is possible that you don't have the required extra e-identifier added to the online banking code. Contact your own bank. (T2)

### 7.3.3 Order of Information

So far, I have focused on the macro-level structure of the texts which was mostly coherent and logical. When viewing the texts in more detail though, some more illogicalities in their structure emerge. For example, T6 starts with a list of things to bring with you when taking prescription medication abroad, introduced in example 5 in chapter 7.1. After that, the text contains an info box noting that the requirements on what to bring vary between countries, so you should check the policies of your destination country before you travel – something that would be helpful to know when reading the list of items to bring (S1, S5). Due to the order the information is presented in, some users might not realize that the list does not cover all situations and miss the note about checking the policies themselves. It is also unclear which situations the list covers, as that is not mentioned anywhere. Finally, the list is the first thing on the page after the title and a summarizing lede paragraph, so it could logically be understood to be the most important piece of information (S5). Yet, I at least would argue that the fact that the policies vary would be the most relevant information to a user reading the text.

As guideline S5 notes, texts should always start with the most important information and, similarly, so should sections within the texts and even individual paragraphs – and the majority of the data did follow this principle. However, when viewing individual paragraphs or sentences, there were some cases where the order of information was not optimal for the user. Example 20 illustrates this: for the user, what they should do is more relevant than what they should not do, yet here the information is presented the other way around. Additionally, the use of the passive (G2) obfuscates the instructions, as does the lack of a direct mention of who the user should contact (G7, I1). This means the user must be able to decode the paragraph and understand the implicit meaning within it – something that can be very difficult or even impossible to someone with cognitive difficulties. This kind of ambiguity was discussed more in chapter 7.2 above.

- (20) A request to restrict or object to the processing of data cannot be addressed to Kanta, as Kanta is not the controller of the data. The controllers of the social welfare client data and patient data stored in Kanta are wellbeing services counties or private health care providers, for example. (T10)

Similarly, example 21 demonstrates a sentence-level issue in the order of information with a step from the instructions for logging in to MyKanta. When reading instructions, the most important thing for the user is the action they should perform, so steps should always start with that information (S5). Here, that would be selecting an authentication method rather than being

redirected. Typically, information about what the system does is included at the end of the step that triggers the system's actions, because the focus should be on the user's tasks. In addition, instructions tend to use the imperative form to ensure clarity, so the phrasing *you can choose* in the example is atypical, as selecting the authentication method is not an optional step.

- (21) 3. You will be redirected to Suomi.fi identification where you can choose a strong electronic authentication method (T2)

### 7.3.4 Cohesion

The final aspect of structure I want to discuss is cohesion. As discussed above, the high-level structure of the texts was, for the most part, cohesive and unlikely to cause significant trouble. However, on the sentence level, the lack of cohesion became a more frequent issue, and almost all texts had some occurrences of consecutive sentences not being coherently linked to one another (S4), as is demonstrated in examples 22 and 23. The first example is particularly problematic because, in addition to the lack of a connective word, the two sentences seem to not be connected by meaning either. In example 23, the lack of cohesion is emphasized by the repetitive information – the two sentences are almost identical in meaning, yet they are not connected with a cohesive device. The main difference between the sentences is the change of perspective, in addition to the mention of *social welfare services* in the second sentence, which does not provide enough new information relevant to the user to justify repetition like this. Similar structures were present in a couple of other texts as well.

- (22) You will receive a confirmation whether the renewal request was successfully submitted. When the doctor renews the prescription, they can see prescriptions prescribed by other doctors in Kanta. (T4)
- (23) The Kanta Services are used in both public and private health care services around Finland. Kanta collects information from public and private social welfare and health care services. (T7)

Occasionally, cohesion was disrupted by connecting sentences incorrectly, as in example 24. Using a demonstrative pronoun, such as *this*, is a common strategy to increase cohesion, and it does connect the two sentences in the example as well. However, looking at the meanings of the sentences, it becomes evident that the two cannot logically be linked as they are here: the data covering the entire population does not make it available in a uniform format. So, while the two sentences are linguistically connected and they seem cohesive when read, the connection is logically flawed and can thus impede linguistic accessibility.

- (24) What is special about Kanta's client and patient data is that, as a rule, they cover data from the entire country and for all citizens. This makes it possible for researchers to obtain data for research purposes in a uniform format from one place and with a single application. (T10)

Finally, lists are generally an accessible way to present related information, and they can increase cohesion as well, particularly when used to connect whole sentences or ideas. A paragraph requires connective words to cohesively link sentences to one another, but items on a list are inherently linked to each other and thus cohesive – as long as they are related to each other. Lists were used quite extensively throughout the data, and there were no sentences that contained long series of words or phrases separated by commas (S6). However, on a few occasions, a list could have helped clarify the message, even if the alternative was not one long sentence, as in example 25. The answer it presents is not very cohesive as the sentences and paragraphs are not clearly linked to one another. Besides adding connective words to the text (S4), cohesion could be improved by adding a list to introduce the possible reasons why the information is not shown in MyKanta before providing additional information (S6). This could also help avoid possible ambiguity (G7); currently only one of the possible reasons is explicitly identified as such, so the user must be able to pick out the reasons from the answer, which can be difficult for users with cognitive difficulties.

- (25) Why is there no information about my visit in MyKanta?

The Health Information section in MyKanta shows information that the health care service professional who treated you has recorded in the Kanta Services. MyKanta shows information starting from the date from which the Kanta Services were taken into use by the health care service unit you are attending.

All public health care units use Kanta, but some private operators do not yet use the Kanta Services. Check which private health care organisations use Kanta.

If your visit data are not shown, it may also be due to the fact that the data have not yet been approved or entered by the health care service. The delay in storing data in the Kanta Services varies between organisations. (T3)

## 7.4 Information Density and Clarity

The last set of guidelines deal with what information is given, how clear it is, and how it is presented – the order of information is considered in the Structure category and was discussed in the previous chapter. It is crucial for accessibility that the users get all the information they need and are directed elsewhere for further details if necessary. This also helps avoid ambiguity, a common theme across the guidelines. Naturally, the information provided to the user must also be correct, as even the most linguistically accessible text can be inaccessible and unusable

if the information is incorrect. However, the correctness of information is beyond the scope of linguistic accessibility and thus this thesis. Additionally, to ensure clarity, information should be presented in a logical way, limiting the amount of new information per sentence and per paragraph. This is particularly important for those who struggle with reading long sentences and processing information. Clarity can also be affected by various visual features, which is why this category contains several guidelines that have to do with visual elements. Visual aspects of the text, such as layout, typography, or colors, are not directly considered here, as the focus of this thesis is on written material. However, the accessibility of some visual elements can be improved with textual features – for instance, images become more accessible when they have a useful and accessible text alternative. These considerations are particularly important for screen reader users, though they can benefit anyone.

#### **7.4.1 Amount of Information**

The structure of a text, discussed in the previous chapter, is closely related to the concept of information density, that is, how much information each sentence and paragraph contains. The research data was primarily accessible in information density: all paragraphs in the data were short and none introduced more than one main idea (I2). On the contrary, sometimes one idea or viewpoint was unnecessarily split into two connected paragraphs. This is illustrated in example 26, where the second, one-sentence-long, paragraph starts with a clear reference to the first: *for this reason*. Separating closely related information like this can impede accessibility and understandability, as it can be more difficult for the user to process the information as one whole when it is presented as two separate paragraphs. Short paragraphs are generally more accessible than longer ones, but dividing information just for the sake of shortening the paragraphs is not always an accessible solution, especially if the combined paragraph would not be excessively long, as was the case in these findings. Similar to the paragraphs, most sentences in the data were short and contained only one point (I3). However, there were some exceptions to this, as occasionally several ideas or viewpoints were introduced in one sentence that would have been more accessible if divided into two – this is demonstrated by example 27.

- (26) If you notice any errors or omissions in the information shown in MyKanta, please contact the treatment unit or professional that you have dealt with directly. Social welfare and health care services are responsible for recording and correcting data.

For this reason, incorrect information shown in MyKanta cannot be updated or rectified by Kanta Services. (T7)

(27) Prescription data and patient data will be stored for 12 years after the patient's death or 120 years after the patient's birth if there is no information on the time of death. (T7)

Whereas the data contained only a few findings related to information density, more issues were related to the overall amount of information in the texts: users should be told everything they need to know, but nothing more (I1). As has become clear in the previous chapters, the amount of information given affects a plethora of other aspects of linguistic accessibility, which is why I have already introduced several examples where this guideline was violated – for instance, example 5 in chapter 7.1, examples 13 and 14 in chapter 7.2, and example 19 in chapter 7.3. Despite the multiple examples, for the most part, the texts did provide enough information. In fact, sometimes the information given was perhaps too detailed, such as in example 28: who exactly processes the request is not really relevant information to the user. Alternatively, if the fact that sometimes a nurse is the one processing the request is deemed important enough to mention, it should also be clarified what those situations are to avoid ambiguity.

(28) [the renewal request] will be processed by a doctor or, in some cases, by a nurse (T4)

However, there was one significant piece of information that, in my opinion, should have been included in all texts that mentioned MyKanta even in passing, yet was not brought up anywhere: the fact that MyKanta is not available in English. In fact, not only was that information not included in any of the evaluated texts, but it was also not mentioned anywhere else on the kanta.fi website – or at least I was not able to find a mention of it, which at the very least means the information was not easily findable. As discussed in chapter 5.3, MyKanta is a central tool in the Finnish health care system, and the fact that the tool is not available in English is already a major accessibility issue as it can prevent those with limited Finnish or Swedish skills from independently managing their health-care-related matters. This issue is exacerbated by the fact that the information about the languages the service is available in is not clearly presented on the kanta.fi website. Similarly, a few of the texts mentioned the kanta.fi chatbot, but lacked the information that the chatbot is only available in Finnish and Swedish, though that information at least was available on the customer service page.

Because the lack of an English translation is not mentioned anywhere on the kanta.fi website, and because the name of the service has been translated into English, users could easily and justifiably assume that the service is available in English. That is, until they log in and are directed to a Finnish page without warning. This can be especially problematic for screen reader users, as most users in this situation would likely try to find the language menu to change the language (possibly assuming the redirect led to the wrong language version) and finding that

menu with a screen reader can be difficult, particularly when it is speaking a language the user does not know. This sudden change in language can also be difficult for those with cognitive difficulties, as they might get disoriented and have trouble locating the language menu even if they can see the site. Of course, in this case, finding the language menu does not help, since the site is not available in English. It should not be left to the user to figure out the lack of a translation on their own, especially when it could be rectified by something as simple as adding one sentence to all pages mentioning MyKanta.

The instructions for using MyKanta did use the Finnish user interface texts to identify various elements the user was to interact with and provided translations for them, as in example 29. One could argue that, based on this, the user should be able to deduce that the service is not available in English. However, not all users will read these instructions, and even if they do, users with cognitive difficulties might not be able to make a connection like that. Additionally, the Finnish names were only mentioned if the direct name of the element was mentioned – in example 30, the name of the drop-down list is not mentioned directly, and since the page has two lists, the user cannot know which to use if they do not know what *visit type* is in Finnish. It is also worth noting that even if the name of the list was given, the available options are not translated or provided in the instructions, so the user might still face issues navigating the service despite the attempt to provide some support via the English translations of the instructions.

- (29) Select the ‘**Terveystiedot**’ (Health information) section in the menu. (T3)
- (30) Select a visit type from the drop-down list. (T3)

Finally, the information about the MyKanta service’s language was not included in any of the links to log in to MyKanta (I4). This might be because the links do not directly lead to MyKanta, they lead to the Suomi.fi e-Identification service that is used to identify oneself and thus log in to MyKanta, and that service is available in English. Therefore, saying the link leads to a Finnish page would technically be incorrect, but it should clearly be stated somewhere, preferably also with the links, that the MyKanta service itself is not available in English. Suomi.fi e-Identification is a third-party service used solely for identification, so, naturally, it cannot provide information about the language of the page the user is directed to after identification – that information should be clear before moving onto the identification service.

## 7.4.2 Links

In general, if the language of the page the link leads to differs from the language of the page the link is on, that information should be included in the link text (I4) – this was mostly true in the evaluated texts. A few links did lack that information though, and a couple of link names were in Finnish, which is not accessible nor acceptable, even if the links lead to a Finnish page. Additionally, a few links led to Finnish pages (and indicated that), even though the destination page was also available in English. It is possible that users will not open the links at all if they lead to a Finnish page as they might, logically, assume that the page is not available in English. It would have been beneficial to check each destination page and modify the link to lead to the English translation if one was available, even though that might have required some additional work in the translation process.

Furthermore, an accessible link text is short, descriptive, and unique, meaning the same link text is not used with multiple different links on the same page. Besides the login links, which all had the same text, no links were repeated within a text, so all the link texts in the data were unique. Most of them were also descriptive and fairly short, though some exceptions did exist. Example 31 demonstrates a link text that would generally be considered accessible: it is short, and the link text matches the title of the page that the link leads to. However, since the links to log in to MyKanta were named *Log in to MyKanta*, it is possible that a user could confuse the two and interpret this to also be a login link. The login links were always presented in the form of a button, which can help sighted users distinguish them, but that does not entirely prevent the confusion – and a screen reader might not necessarily even distinguish between a text-form link and a button that acts as a link. To avoid this mix-up, the link text could have included some indicator that it is not a login link, such as “Read more about”.

(31) [Log in securely to MyKanta](#) (T1, underline added to indicate hyperlink)

On the other hand, that type of additional information is only useful if it helps the user distinguish similar links from each other. Typically, it is not necessary, as the existence of a link already indicates that it leads to more information. This is demonstrated by example 32: the link could have started after the *Learn in more detail about* without negatively impacting accessibility. It can become particularly troublesome if multiple links start with a similar phrase, as many screen reader software allow the user to list all links on a page. Instead, using key words at the beginning of the link texts allow screen reader users to quickly navigate between

them. Additionally, multiple consecutive links starting with the same phrase can cause issues for sighted users too: in T8, six consecutive links, presented in two lists with three items each and separated by a heading and one paragraph, started with *Find out more about*. As with any list, if the items start with the same phrase, it can be difficult to tell them apart and find the information that matters. In other texts, similar situations were typically handled by having the introductory phrase before the list, which helps the user find the necessary information and improves linguistic accessibility.

(32) Learn in more detail about what health information is shown in MyKanta (T7, underline added to indicate hyperlink)

Finally, example 33 illustrates a link that in context makes perfect sense and is accessible, but the link text (underlined in the example) alone is vague. This can be a problem for screen reader users utilizing the option to list all links on a page: that list only contains the link text, so users will not get the surrounding context that provides essential information. In a similar vein, a couple of link texts – more specifically, buttons that acted as links – were just *Read more*, which is not descriptive enough without the surrounding context. It is likely that some technical solutions to provide additional context for screen reader users exist, and I cannot for sure say if such were used here, but, based on my very limited testing, I could determine that at least some links lacked the additional context.

(33) Check which private health care organisations use Kanta. (T3, underline added to indicate hyperlink)

Links can also be helpful in controlling the amount of information included in the texts. As mentioned earlier, it is important that the user gets all the information necessary to, for example, perform the action they are reading the text for. However, it is rarely possible, or even logical, include every detail of every aspect in the text – that can lead to an information overload and decrease linguistic accessibility as it becomes harder for the user to find the information that matters to them. Therefore, a useful strategy is to provide the essential information in the text itself and include links to additional information so the user can easily look up related details they might not be familiar with (I1). This strategy was used extensively in the texts, and I discussed it briefly in the previous chapter as well. Still, there were some occasions when a link was not provided even if it would have been useful. For example, consenting to data sharing was briefly mentioned in T7, and it would have been helpful to also include a link to a page where the user can read more about it, as a page covering that does exist on the kanta.fi website.

### 7.4.3 Non-Text Content

In addition to accessible link texts (I4), text alternatives for non-text content (I5) are particularly important for screen reader users. It is of course important to remember that visual presentation of information can be more understandable and accessible to some users, and many might benefit from supporting visuals alongside the text. For instance, user instructions often include images of the product or user interface to help users find the element referenced in the instructions; the data did not contain such visuals, though the instructions for using MyKanta could have benefitted from them. However, that is outside the scope of this thesis. Here, my focus was on checking whether any information would be missed if the user only had access to the textual content – as is often the case for screen reader users. Most of the texts in the data did not contain any non-text information: three texts contained infographics and one an embedded video, in addition to which many of the texts had some decorative images, that is, pictures that contain no information.

The one embedded video in the data was titled *Kanta – a pillar of wellbeing*. The video explained the Kanta Services as a whole, but since it was included on a page focused on MyKanta (T1), it seemed a little out of place, violating some of the structure-related guidelines discussed in the previous chapter. However, linguistically, the video was accessible: it provided subtitles that covered everything that was said, and all visual information was also said out loud, so the information could be received with only sight or hearing. It is worth noting though that if the video was not viewed in full screen mode, the subtitles were initially hidden behind the control bar that contains, for example, the play/pause button – at least on my screen. It was not possible for the user to manually hide the bar, and it popped back up if the cursor was moved over the video. This is not related to linguistic accessibility directly, but it can significantly impede the overall accessibility for those who need the subtitles, so it is worth mentioning.

Images were more common in the data, as many of the texts included short info boxes about various topics, typically placed at the end of the page and containing a decorative image suiting the topic at hand – such as the one in Figure 1. Since the images do not contain information, their alt texts should be empty so screen readers can ignore the images; for the most part, that was true for the images in the data. There were a couple that described the image in the alt text though, as is illustrated by the alt text of the image in Figure 1 presented in example 34. Additionally, one image in T10 had an empty alt text, but it was not implemented correctly: instead of being fully empty, the alt text consisted of quotation marks. Therefore, a screen reader

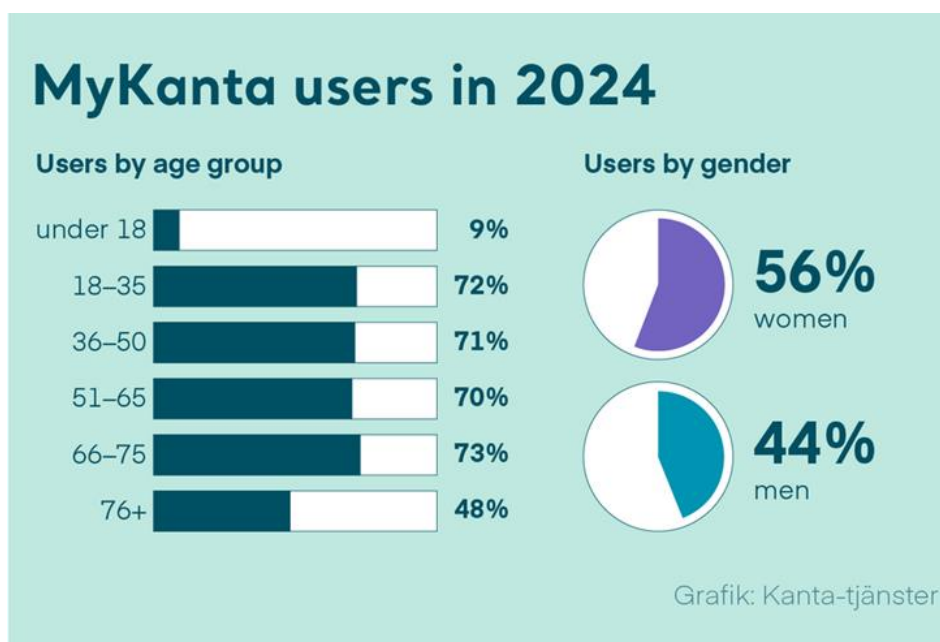
would not ignore the image; instead, it would announce the existence of an image but, because the alt text did not contain any text, would not provide a description of it. This makes the alt text inaccessible, as a user who cannot see the image cannot tell if it is decorative or if it contains information. And in this case, the image was an infographic on the use of Kanta data and contained information that was not presented anywhere else.



**Figure 1. Info box at the end of T6**

(34) Passport and a smartphone. (T6)

The majority of the non-decorative images in my dataset were in T9, “Statistics”, and showed key figures about Kanta usage and users. All of these images had alt texts that described the statistics, but the numbers in the alt texts did not fully match the ones in the images. This is illustrated by Figure 2 and its alt text presented in example 35. Considering the year in the image is 2024 while the alt text mentions the year 2022, it seems likely that the images were updated to contain newer statistics, but the alt texts were not. Interestingly, the information in the figure itself is also ambiguous (G7): clearly, the *users by age group* part shows the percentage of MyKanta users in each age group, but it is unclear whether the same is true for the *users by gender* charts. Because the percentages add up to 100%, the statistic can be interpreted in two ways: either it follows the same logic as the age group statistic and indicates that 56% of all women use MyKanta, or it describes the gender distribution among MyKanta users, stating that 56% of them are women.



**Figure 2. Example of a figure from T9**

- (35) Figure 4: MyKanta was used by 3.5 million people in 2022. There were a total of 37 million logins to MyKanta. MyKanta is used by 11% of under-18s. Over 80% of people aged 18–65 use MyKanta. MyKanta is used by 72% of those aged 66–75 and by 43% of those aged over 76 (T9)

The *users by gender* statistic in Figure 2 is not mentioned in the alt text presented in example 35, but the alt text mentions the number of MyKanta users and logins, which are not shown in Figure 2 (I5). These disparities could also be explained by the images being updated and the alt texts not, or it could be intentional in an attempt to keep the alt texts a reasonable length by including information about one image in the alt text of another – the number of users and logins were included in the previous image in the text. Although this is not likely to cause significant issues for someone who cannot see the images, it can confuse those relying on machine translations of the alt texts to help them understand the images. Additionally, the gender distribution was not mentioned in any alt text, so at least some information was unavailable to those who could not see the image. It is also worth noting that all of the alt texts in T9 started with *Figure* – while guideline I5 does not mention this, it is not necessary to declare the image as such in the alt text, because a screen reader will tell the user that there is an image. The figures were also numbered as 1, 2, and 4 – the missing figure 3 can lead the user to think that they are missing an image.

Finally, the alt texts often contained strings of sentences that did not cohesively link to one another (S4). This was especially the case with the figures in T9 and is also illustrated by

example 35. The lack of cohesion in the example is likely caused by the attempt to fit all the information presented in the figure to an alt text while still keeping it short. This might also explain the use of the atypical phrase *under-18s*. If the same information was presented in a normal paragraph, I would have recommended a list format instead (S6), but a list is not necessarily a suitable solution for an alt text. Nevertheless, the information could have been presented in a more accessible and cohesive way. One option could have been to include only the diagrams as images and describe them in the caption visible and accessible to everyone – that way, no textual information would be hidden in an image and the long and incohesive alt texts could be avoided.

The final guideline, I6, is concerned with the way user interface elements that the user must be able to locate and interact with are described. Referring to these elements only by their visual features such as color or location – for instance, “Select the red button at the bottom of the page” – is problematic particularly to screen reader users who cannot see the user interface, as well as users with color vision deficiencies. Additionally, users with cognitive difficulties might have trouble understanding directional instructions. My research data did not contain any findings related to this guideline, though it is worth noting that only three of the ten texts included instructions on operating a user interface, so the majority of the data did not have anything relevant to this guideline. Nevertheless, the lack of findings does not mean this guideline is irrelevant; it simply means the evaluated texts followed this guideline well. Of course, this is one of the new guidelines I added for this thesis, so it would have been beneficial to also evaluate how well it suits the relevant findings. However, despite the lack of testing, I still believe that this guideline is relevant and thus justified to keep in the collection of guidelines. After all, the way user interface elements are referred to can have a significant impact on the accessibility of a service.

## **7.5 Summary of Findings and the Overall Linguistic Accessibility**

Overall, the results of the analysis were relatively even across the guidelines and throughout the data – no texts stood out in the number or severity of findings, and no category or individual guideline alone had a significant presence in the findings. Naturally, some guidelines had more findings than others, and certain guidelines had more severe findings, but the majority of findings were relatively minor. Still, it is important to keep in mind that these findings accumulate and, when combined, can cause more severe issues. For instance, most of the

findings related to the correctness and consistency of grammar and spelling were minor, but there was a significant number of those findings, increasing their overall impact on linguistic accessibility. Similarly, findings that alone seem significant can, in the end, have only a minor impact on accessibility if those findings are not repeated – one such finding was the ambiguously formatted list introduced with example 5 in chapter 7.1.

The most prevalent findings, that I consider to have the most significant impact on the overall accessibility of the texts, were related to the consistent use of names and labels (V4) and the amount of information provided (I1), as well as the general ambiguity caused by the violation of various guidelines – in addition to the two already mentioned, ambiguity was caused by, for example, complex sentence structures (G3), the passive voice (G2), ambiguous structures (G7), abstract language (V5), and a lack of cohesion both in the overall structure (S1, S2) and within sentences and paragraphs (S4). Just as with most other guidelines, the findings related to these also varied in severity. However, since ambiguity is such an overarching theme affected by a multitude of factors, it is difficult to point to individual findings as being more severe than others. Instead, the reason I consider ambiguity to be such a severe issue in the data is the number of findings that, in one way or another, contribute to the overall ambiguity of the texts.

The issues in consistency, on the other hand, are easier to categorize based on severity. For instance, the inconsistent verbs used to indicate clicking a button in the user interface are much less of an issue than the interchangeable use of *data* and *information* – which is one of the issues I believe to be most likely to cause confusion, at least when it comes to individual words and sentences (V4). Similarly, the lacking explanation of pseudonymization in example 9 is a fairly minor violation compared to the absence of any mention of the fact that MyKanta is not available in English (I1). I covered that extensively in chapter 7.4 because I firmly believe that to be one of the most significant issues in the data. Despite its potential for causing severe accessibility issues, the violation could be rectified by the simple addition of a mention about the language change to the relevant contexts.

The lack of such mention can possibly be explained by the translation process, at least in part. While the mentions of user interface elements in T2, T3, and T4 did take the lack of a translation into account by referring to the elements with their Finnish names and providing English translations for them, it is possible that the translations otherwise followed the source texts quite strictly and did not add any new information. Because this thesis is not about the translation process or the accuracy or suitability of the translations, I did not compare the evaluated texts

with the Finnish ones available on the website. Since the users of these English texts cannot rely on the Finnish ones if they do not understand something, I did not do that either. Therefore, any possible translation errors would have been missed in the analysis, and it is possible that some of the findings could be explained by something in the translation process. For instance, the lack of clarification on the Finnish health care system and organizations could indicate that at least no major changes were made to the contents of the texts during the translation process. Since the audience of the English translation is likely different from the audience of the original, certain changes, such as additional explanations, could have improved the overall linguistic accessibility of the texts. The translation process is outside the scope of this thesis but acknowledging the effect it might have had on the accessibility of the texts is important, particularly when considering the real-life implications of the results – knowing where the issues might stem from is helpful when trying to resolve them.

When it comes to the different texts, there was less variance than with the guidelines. The most frequently violated guidelines varied across the texts, but, overall, the linguistic accessibility of all the evaluated texts was relatively equal. T7 and T10, as discussed in chapter 7.3, had more issues with structure, which in turn negatively affected the overall linguistic accessibility of those texts, making these perhaps the least accessible ones in the dataset, though the differences were not major. Conversely, T2, T3, and T4 – the texts instructing the user to operate in MyKanta – were the ones with the least issues. However, they, along with T1 “MyKanta”, were the ones where the information about the lack of an English version of MyKanta would have been most crucial. Otherwise, the overall number and severity of findings in these texts was a bit less than average. That is likely because these texts were structurally simple and had less content than some other texts; since they were focused on instructing the user to perform actions in MyKanta, a significant portion of the content was step-by-step instructions, which, in general, are simpler in form than informative paragraphs. Nevertheless, they too had their issues and, as noted earlier, the overall level of linguistic accessibility between the texts did not vary significantly.

Due to the nature of heuristic evaluation as a method, the findings and their analysis are focused on negative aspects. Of course, it is clear from the previous chapters that, aside from the last guideline I6, the data contained multiple violations of each guideline. Additionally, most guidelines had at least one finding in each of the ten evaluated texts. However, most findings were not severe or consistent, and some guidelines were followed well. For instance, the

paragraphs in the texts were all short and did not contain multiple ideas (I2), nominalizations (G5) and abbreviations (V3) were rare, and lists were used extensively (S6). In addition, the majority of the words used in the texts were common (V1) and the high-level structure of most texts was logical (S1, S2). Therefore, despite some of the more severe and persistent issues, the overall linguistic accessibility of the texts was adequate enough that one could reasonably assume that a user with sufficient language skills and no or only minor cognitive difficulties would be able to use and understand the texts without significant trouble.

Still, as discussed in chapter 5, the kanta.fi website and texts within its Citizens section are aimed at the general public, that is, everyone in Finland. These types of websites should make lower-literacy users a priority and ensure the texts are understandable and accessible enough that they can be understood by everyone, not just by high-literacy users (Nielsen 2005). Additionally, since kanta.fi is an essential part of the Finnish health care system, it is particularly important for it to be accessible to all. For example, immigrants are already at a disadvantage when it comes to health and health care, as are persons with disabilities, who also typically use these services more than the average citizen. Therefore, it is vital that the information regarding these services is as (linguistically) accessible as possible – and available in as many languages as possible. So, while the texts could be understood by some users, that is not enough, and my analysis does not indicate that the English kanta.fi website is linguistically accessible.

## 8 CONCLUSIONS

In this thesis, I have explored the concept of linguistic accessibility and analyzed how it was realized on the English kanta.fi website for citizens. The goal was to understand how linguistically accessible the website is, that is, how easily the written information on the site can be understood and used by a wide variety of users, including those with limited language skills and those with cognitive difficulties. To do that, I analyzed ten randomly selected texts from the kanta.fi website using the methods of heuristic evaluation: comparing the texts to a list of guidelines – heuristics – and looking for occasions where those guidelines were violated.

To form an overall idea of the linguistic accessibility of the site, I needed to understand what kind of accessibility issues were present on the site and how prominent they were. The analysis outlined in chapter 7 revealed a variety of accessibility issues ranging in severity and frequency. Apart from guideline I6, all guidelines had multiple findings, and no single guideline stood out in the number or severity of findings. Since my goal was not to form a comprehensive list of flaws to fix issues on the site, I did not categorize the individual findings in terms of severity, nor did I collect a comprehensive list of the findings. Instead, I used the findings as the basis for my analysis and attempted to construct an understanding of the overall linguistic accessibility. While a single problem can significantly impede the understandability of a specific passage, it is not likely to have a major effect on the overall linguistic accessibility of a text, let alone an entire website. Similarly, several minor issues that alone do not constitute a significant problem can in large numbers have a substantial impact on the overall linguistic accessibility. Because of this, rather than focusing on individual findings or guidelines, my analysis was focused on the categories and themes present across guidelines and categories.

The most prevalent themes that had findings across the different categories and guidelines were consistency, ambiguity, and amount of information. This is not necessarily surprising – while some of these had their own guidelines as well, they are the types of general, overarching themes that can be seen, on some level, in most of the guidelines. This prevalence in the guidelines explains their frequency in the results as well; more niche topics, such as correct spelling, cannot have findings in multiple guidelines. Still, the frequency of these themes is an important finding, particularly when considering the applicability of these results to the linguistic accessibility of the whole website.

Just as no guideline stood out in frequency or severity, the different texts were relatively similar in terms of accessibility: all texts had multiple findings from all categories of guidelines, and the overall number and severity of findings was similar. As can be expected, the structure and content of the texts affected their overall accessibility: texts with a simpler structure, such as those focused on step-by-step instructions, generally had fewer issues than longer informative texts covering more complex topics. Still, the differences were not significant enough to single out any individual text.

So, are the texts linguistically accessible? Based on the results of my analysis, no. While the severity of the issues varied and the number of very severe issues was not significant, the number of findings in general indicates that the texts cannot be declared linguistically accessible, at least not to all users. Granted, one could reasonably assume that someone with sufficient English skills and no or only minor cognitive difficulties would likely be able to understand the majority of the texts with little trouble. However, that is not enough – due to the importance of the Kanta Services in the Finnish health care and social welfare systems, it would be crucial to ensure information about those services is available and accessible to all. Because of this, it would also be critical to provide information in other commonly spoken foreign languages alongside English. Despite my focus on English, one of the best ways to ensure linguistic accessibility, particularly to foreign language speakers, is to provide translations in as many languages as possible. Additionally, since it is realistically not possible to offer translations in all foreign languages, special attention should be given to ensuring the linguistic accessibility of the translations that are offered.

Users with cognitive difficulties can face issues understanding the texts, especially due to the issues caused by ambiguity, inconsistency, and lack of clarifying information present across the analyzed texts. Similarly, users with limited English skills might struggle with the texts – inconsistent terms and complex structures in particular can be problematic for them. Additionally, the texts can, in part, be inaccessible to those who are not familiar with the Finnish health care and social welfare systems, such as new migrants, as many organizations, systems, and practicalities that are likely to be familiar to a Finnish person were not explained. This is probably a reflection of the translation process; while I did not compare the translations to their source texts, one can reasonably assume that the lack of clarification in the translation is likely because that clarification was not necessary in the source text, and the translation did not provide additional information despite the different target audiences. This potential effect of

the translation process was illustrated particularly well by the lack of a mention about the fact that the MyKanta service is not available in English – something that would be useful information for those who do not speak Finnish or Swedish, and that can cause significant confusion to those with cognitive difficulties as well as screen reader users.

Since the findings across the categories and guidelines as well as the different texts were similar, these results may be used to form an overall idea of the linguistic accessibility of the texts on the English kanta.fi website, despite the relatively small sample size. It is unlikely that the evaluated texts, selected at random across different sections and covering a variety of topics and structures, would be the outliers of the website. This is particularly true for the written material that the majority of the data consisted of – results regarding less prominent material such as images and videos are less reliably applicable to the entire site and would require further investigation. Similarly, any content outside the scope of this thesis, such as user interface texts, could provide more insights into the linguistic accessibility of the entire website and would be interesting topics for future research. It is also important to remember that linguistic accessibility does not exist in a vacuum, and even the most linguistically accessible texts can become inaccessible if the website they are on is inaccessible. Therefore, studying the other aspects that affect the overall accessibility of the texts, such as navigation, layout, and the correctness of information, alongside linguistic accessibility, could provide fruitful insights and help create a more accessible website for all users.

It is of course important to remember that individuals' needs vary and no single research, particularly by one researcher, can make a definitive declaration on the accessibility of an entire website. This lack of input from real users is a known weakness of heuristic evaluation (Nielsen and Molich 1990, 251–52). To get more reliable and useful information about the linguistic accessibility of these texts, real users in the target audiences should be involved. Feedback from users with limited language skills or cognitive difficulties would be especially useful. Additionally, since some of the guidelines dealt with visual features and accessibility solutions that would be particularly helpful for screen reader users, involving those users in testing could provide valuable insights.

It is also worth noting that I performed the evaluation alone, so the results are bound to reflect my subjective views. Following the general recommendation of 3–5 evaluators (Nielsen 1992, 376) might have yielded more findings and provided more insights into the severity of issues: problems identified by multiple evaluators are more likely to be noticed by real users as well

and thus more likely to cause accessibility issues when the texts are used. Still, as my goal was not to find all accessibility issues or to categorize them by severity but rather to form an understanding of the overall state of linguistic accessibility on the site, I believe my findings alone are sufficient for the purpose. In addition, the consistency of the findings indicates that the issues are likely to cause problems for real users as well. However, analyzing these texts with multiple evaluators and involving real users by performing usability testing could be an interesting subject for follow-up research.

The heuristics used in this thesis worked well for this purpose, and I believe they could be usable in other contexts as well, both as heuristics for evaluating linguistic accessibility and as guidelines to support writing accessible texts – at least to some extent. However, the guidelines themselves were not the focus of this thesis and their use in other contexts was not considered in detail. Due to the limitations and effects of my language skills and the scope of this thesis, the guidelines are more likely to work with informative and instructional texts written in Finnish or English but might not be so successful with other types of texts and other languages. Nevertheless, the guidelines could prove helpful in other use cases as well, but further testing and research would be necessary to ensure their usability. Additionally, it could be interesting to explore their suitability for other types of texts and texts in other languages, and to develop and/or expand them based on the results.

Overall, the results of the analysis aligned with what I expected. Issues with consistency, ambiguity, and amount of information are, in general, quite typical in a variety of texts. Still, as the kanta.fi website contains a few statements indicating they are paying special attention to the understandability of the language used on the site, one could assume the texts would be relatively linguistically accessible. However, the fact that they expressed this sentiment in their accessibility statement by noting that they “invest in the intelligibility of language” (Kanta Services 2025b) gave some indication that the language they use might not be the simplest or clearest and thus not the most linguistically accessible.

It is possible that the Finnish (and potentially the Swedish) texts are more linguistically accessible than the English ones analyzed here. I am not familiar with the kanta.fi content creation process or their translation process, so I cannot make any definitive claims about them. However, in general, while the original Finnish texts are more typically written in-house by writers intimately familiar with the service, translations are commonly purchased from third-party translation companies and thus produced by people less familiar with the service

and topic – this disparity can also be reflected in the translations. It is also possible that the translators have not been instructed to, or had the opportunity to, focus on linguistic accessibility the same way the writers of the Finnish texts might have. These would be fruitful topics for further research. Another interesting topic for follow-up research could be to perform a similar comparative study on the translations and the source texts. Additionally, it would be illuminating to find out what the organizations behind kanta.fi do to ensure the understandability of their texts and how that goal is considered in the translation process – interviewing the content creators and translators for kanta.fi could provide insights into this.

Despite the limitations of this thesis, I believe the results are useful and usable. Although certain modifications or additions to the research design could have yielded more reliable results, the analysis provided solid enough evidence to answer the research question with at least some level of certainty: the English kanta.fi website for citizens is not linguistically accessible to all users. The findings of this thesis could be used to concretely support improving the linguistic accessibility of the website. Even though I did not produce the typical outcome of a heuristic evaluation, a list of problems and the heuristics they break, the analysis identified themes and common issues across the texts; this information can be helpful in finding and fixing the linguistic accessibility issues present on the website and thus be used to improve the understandability of the texts.

Since the website provides crucial information about the Kanta Services, which are used extensively in health care and social services throughout the country, it would be critical to ensure its linguistic accessibility for all. It would also be important to understand and consider the different target audiences of the English translation and the Finnish source text: it is likely that the users of the English texts are not as familiar with the Finnish society and systems as the readers of the Finnish texts. They are also likely not native speakers of English, while the Finnish texts are most likely read primarily by native or at least fluent speakers of Finnish. The results of this thesis could be used to help ensure that more users are able to understand the information on the kanta.fi website and help them function in the Finnish health care and social services systems – thus helping them participate in the Finnish society and ensuring their equal treatment and the fulfillment of their fundamental human rights.

## WORKS CITED

### Research Data

- Kanta Services. 2024a. "Take Medicines and Prescriptions with You When You Travel Abroad." Kanta.fi, June 13. <https://www.kanta.fi/en/when-travelling>.
- Kanta Services. 2024b. "Using Pharmacy Services." Kanta.fi, April 23. <https://www.kanta.fi/en/at-the-pharmacy>.
- Kanta Services. 2025a. "The Data Saved in Kanta Are Shown in MyKanta." Kanta.fi, February 18. <https://www.kanta.fi/en/data-in-kanta>.
- Kanta Services. 2025b. "Statistics." Kanta.fi, March 7. <https://www.kanta.fi/en/statistics>.
- Kanta Services. 2025c. "Research and Knowledge Management." Kanta.fi, March 12. <https://www.kanta.fi/en/research-and-knowledge-management>.
- Kanta Services. 2025d. "Prescriptions and Prescription Renewal in MyKanta." Kanta.fi, January 29. <https://www.kanta.fi/en/prescriptions-and-prescription-renewal>.
- Kanta Services. 2025e. "Log in Securely to MyKanta." Kanta.fi, February 18. <https://www.kanta.fi/en/log-in-securely-to-mykanta>.
- Kanta Services. 2025f. "How Can I Have a Say in the Use of My Data?" Kanta.fi, February 18. <https://www.kanta.fi/en/consent-and-denials-of-consent-to-data-sharing>.
- Kanta Services. 2025g. "Health Care Visits and Diagnoses in MyKanta." Kanta.fi, January 29. <https://www.kanta.fi/en/appointments-and-diagnoses>.
- Kanta Services. 2025h. "MyKanta." Kanta.fi, February 18. <https://www.kanta.fi/en/mykanta>.

### References

- Act on the Status and Rights of Patients (785/1992) = Laki potilaan asemasta ja oikeuksista (785/1992). <https://www.finlex.fi/fi/laki/ajantasa/1992/19920785>.
- Administrative Procedure Act (434/2003) = Hallintolaki (434/2003). <https://www.finlex.fi/fi/laki/ajantasa/2003/20030434>.
- Ahrens, Sarah. 2020. "Easy Language and Administrative Texts: Second Language Learners as a Target Group." In *Easy Language Research : Text and User Perspectives*, edited by Silvia Hansen-Schirra and Christiane Maaß. Easy – Plain – Accessible 2. Frank & Timme. <https://doi.org/10.26530/20.500.12657/42088>.
- Åkermarck, Mikael. 2020. "Puheenvuoro: Viestinnän saavutettavuus lainsäädännössä." In *Saavutettava viestintä: yhteiskunnallista yhdenvertaisuutta edistämässä*, edited by Maija Hirvonen and Tuija Kinnunen. Gaudeamus.
- Constitution of Finland (731/1999) = Suomen perustuslaki (731/1999). <https://www.finlex.fi/fi/laki/ajantasa/1999/19990731>.
- Costa, Daniel. 2025. "24-Hour Clock | Definition, History, Uses, & Facts." In *Encyclopedia Britannica*. January 17. <https://www.britannica.com/topic/24-hour-clock>.

- Diño, Gino. 2021. “Sure, Google Translate Volumes Are Staggering, But Here Are Hotter Insights From Google I/O 2021.” Slator, June 22. <https://slator.com/sure-google-translate-volumes-are-staggering-but-here-are-hotter-insights-from-google-i-o-2021/>.
- European Accessibility Act = Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the Accessibility Requirements for Products and Services (Text with EEA Relevance), CONSIL, EP, 151 OJ L (2019). <http://data.europa.eu/eli/dir/2019/882/oj/eng>.
- Eriksson-Backa, Kristina, Noora Hirvonen, Heidi Enwald, and Isto Huvila. 2021. “Enablers for and Barriers to Using My Kanta – A Focus Group Study of Older Adults’ Perceptions of the National Electronic Health Record in Finland.” *Informatics for Health & Social Care* 46 (4): 399–411. 153654692. <https://doi.org/10.1080/17538157.2021.1902331>.
- European Commission. 2025. “Web Accessibility Directive — Standards and Harmonisation.” May 5. <https://digital-strategy.ec.europa.eu/en/policies/web-accessibility-directive-standards-and-harmonisation>.
- European Standard (ETSI) EN 301 549:2015 on the Accessibility Requirements Suitable for Public Procurement of ICT Products and Services in Europe, 2015. 1-134.* 2015. [https://ec.europa.eu/eip/ageing/standards/ict-and-communication/accessibility-and-design-all/en-3015492015\\_en](https://ec.europa.eu/eip/ageing/standards/ict-and-communication/accessibility-and-design-all/en-3015492015_en).
- Eurostat. 2024. “Digital Economy and Society Statistics - Households and Individuals.” Statistics Explained, July 19. [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital\\_economy\\_and\\_society\\_statistics\\_-\\_households\\_and\\_individuals](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals).
- Friedman, Mark, and Diane Bryen. 2007. “Web Accessibility Design Recommendations for People with Cognitive Disabilities.” *Technology and Disability* 19 (4): 205–12. <https://doi.org/10.3233/TAD-2007-19406>.
- HE 72/2002. 2002. “Hallituksen Esitys Eduskunnalle Hallintolaiksi Ja Laiksi Hallintolainkäyttölain Muuttamisesta.” Finlex, May 24. <https://www.finlex.fi/fi/hallituksen-esitykset/2002/72>.
- Heikkinen, Vesa. 2012. “Ymmärrettävyytutkimuksen kysymyksiä.” *Kielikuvia* 2/2012: 40–44.
- Henry, Raymond M., Gordon E. McCray, Russell L. Purvis, and Tom L. Roberts. 2007. “Exploiting Organizational Knowledge in Developing IS Project Cost and Schedule Estimates: An Empirical Study.” *Information & Management* 44 (6): 598–612. <https://doi.org/10.1016/j.im.2007.06.002>.
- Hiidenmaa, Pirjo. 2005. “Näkökulmia yleiskieleen.” *Kielikello*, April. <https://kielikello.fi/nakokulmia-yleiskieleen/>.
- Hirvonen, Maija, Tuija Kinnunen, and Liisa Tiittula. 2020. “Viestinnän saavutettavuuden lähtökohtia.” In *Saavutettava viestintä: yhteiskunnallista yhdenvertaisuutta edistämässä*, edited by Maija Hirvonen and Tuija Kinnunen. Gaudeamus.
- Horowitz, Minna, Hannu Nieminen, Marko Ala-Fossi, et al. 2019. *Viestintä kuuluu kaikille: kansalaisten viestinnälliset oikeudet ja mahdollisuudet Suomessa*. Gaudeamus.
- Invenius, Sonja. 2024. “Kielellinen saavutettavuus kaupunkiviestinnässä: Vantaan kaupungin verkkosivuston kognitiivisen saavutettavuuden piirteitä.” Master’s Thesis, Tampere University. <https://trepo.tuni.fi/handle/10024/156840>.

- ISO. 2008. *Ergonomics of Human-System Interaction — Part 171: Guidance on Software Accessibility*. ISO 9241-171:2008. <https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-171:ed-1:v1:en>.
- ISO. 2018. *Ergonomics of Human-System Interaction — Part 11: Usability: Definitions and Concepts*. ISO 9241-11:2018. <https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-11:ed-2:v1:en>.
- ISO. 2020. *Cognitive Accessibility — Part 1: General Guidelines*. ISO 21801-1:2020. <https://www.iso.org/obp/ui/en/#iso:std:iso:21801:-1:ed-1:v1:en>.
- ISO. 2023. *Plain Language — Part 1: Governing Principles and Guidelines*. ISO 24495-1:2023. <https://www.iso.org/obp/ui/en/#iso:std:iso:24495:-1:ed-1:v1:en:ref:3>.
- Jantunen, Jonna. 2020. “Understanding Understandability: Evaluating Understandability as a Part of the Linguistic Accessibility of a Text and Its Translation.” Bachelor’s Thesis, Tampere University. <https://urn.fi/URN:NBN:fi:tuni-202008036386>.
- Jylhä, Virpi. 2020. “Puheenvuoro: Viestinnän apuvälineet näkövammaisten arjessa.” In *Saavutettava viestintä: yhteiskunnallista yhdenvertaisuutta edistämässä*, edited by Maija Hirvonen and Tuija Kinnunen. Gaudeamus.
- Kanta Services. 2024a. “Kanta Partners.” Kanta.Fi, September 27. <https://www.kanta.fi/en/partners>.
- Kanta Services. 2024b. “The Kanta Services as Part of Social Welfare and Health Care.” Kanta.Fi, August 29. <https://www.kanta.fi/en/kanta-and-the-social-welfare-and-health-care-services>.
- Kanta Services. 2024c. “We Have Updated the Instructions for Managing Affairs to Provide Clearer Guidance on Using MyKanta.” Kanta.Fi, April 24. [https://www.kanta.fi/en/article/-/asset\\_publisher/1QjC602jKPR6/content/uudistimme-asiointiohjeet-ohjaavat-entista-selkeammin-omakannan-kaytossa](https://www.kanta.fi/en/article/-/asset_publisher/1QjC602jKPR6/content/uudistimme-asiointiohjeet-ohjaavat-entista-selkeammin-omakannan-kaytossa).
- Kanta Services. 2024d. “What Are the Kanta Services?” Kanta.Fi, September 24. <https://www.kanta.fi/en/what-are-kanta-services>.
- Kanta Services. 2025a. “A Brief History of the Kanta Services.” Kanta.Fi, January 23. <https://www.kanta.fi/en/a-brief-history-of-the-kanta-services>.
- Kanta Services. 2025b. “Accessibility Statement for the Kanta.Fi Web Service.” Kanta.Fi, January 2. <https://www.kanta.fi/en/accessibility-statement-for-the-kanta.fi-web-service>.
- Kanta Services. 2025c. “Kanta Services Communications.” Kanta.Fi, January 3. <https://www.kanta.fi/en/system-developers/communication-and-social-media>.
- Kanta Services. 2025d. “MyKanta.” Kanta.Fi, February 18. <https://www.kanta.fi/en/mykanta>.
- Kela. 2025. “Kelan asiakasviestinnän kielilinjaus.” April 3. <https://www.kela.fi/documents/d/guest/asiakasviestinnan-kielilinjaus-2025-pdf>.
- Kervinen, Mari. 2020. “Digipalvelujen selkeä ja ymmärrettävä kieli: sisällöntuottajien näkemyksiä kielellisestä saavutettavuudesta.” Master’s Thesis, University of Helsinki. <http://urn.fi/URN:NBN:fi:hulib-202011264610>.
- Koponen, Maarit, and Mary Nurminen. 2020. “Konekäännös tiedon saavutettavuuden edistäjänä ja esteenä.” In *Saavutettava viestintä: yhteiskunnallista yhdenvertaisuutta edistämässä*, edited by Maija Hirvonen and Tuija Kinnunen. Gaudeamus.

- Kotus. n.d. “Hyvä virkakieli - parempi hallinto.” Kotimaisten kielten keskus - Kielitieto. Accessed September 23, 2024. [https://www.kotus.fi/kielitieto/hyva\\_virkakieli](https://www.kotus.fi/kielitieto/hyva_virkakieli).
- Kulkki-Nieminen, Auli. 2010. “Selkoistettu uutinen. Lingvistinen analyysi selkotekstin erityispiirteistä.” PhD diss, Tampere University. <https://urn.fi/urn:isbn:978-951-44-8093-5>.
- Laakso, Satu. 2021. “Translated English Online Patient Information on Finnish Public Hospital Websites : Usability and Equality in Focus.” Master’s Thesis, University of Vaasa. <https://osuva.uwasa.fi/handle/10024/12344>.
- Leskelä, Leealaura. 2019. *Selkokieli: saavutettavan kielen opas*. Kehitysvammaliitto ry.
- Leskelä, Leealaura, and Eliisa Uotila. 2020. “Selkokieli saavutettavan viestinnän välineenä.” In *Saavutettava viestintä: yhteiskunnallista yhdenvertaisuutta edistämässä*, edited by Maija Hirvonen and Tuija Kinnunen. Gaudeamus.
- Maaß, Christiane, and Silvia Hansen-Schirra. 2022. “Removing Barriers: Accessibility as the Primary Purpose and Main Goal of Translation.” In *Translation, Mediation and Accessibility for Linguistic Minorities*, edited by María Pilar Castillo Bernal and Marta Estévez Grossi. Frank & Timme. <https://doi.org/10.26530/20.500.12657/54058>.
- Määttä, Simo K., Tuija Kinnunen, Svetlana Probirskaja, and Päivi Kuusi. 2022. *Improving Communication with Migrants for Crisis Preparedness in Finland: Lessons Learned from COVID-19*. University of Helsinki. <http://hdl.handle.net/10138/350703>.
- Matamala, Anna, and Carla Ortiz-Boix. 2016. “Accessibility and Multilingualism: An Exploratory Study on the Machine Translation of Audio Descriptions.” *TRANS: Revista de Traductología*, no. 20: 20. <https://doi.org/10.24310/TRANS.2016.v0i20.2059>.
- Ministry of Finance = Valtiovarainministeriö. 2017. “Saavutettavuusdirektiivi.” February 17. <https://vm.fi/-/saavutettavuusdirektiivi>.
- Ministry of Finance = Valtiovarainministeriö. 2024. “Digitaalisen viranomaisviestinnän ensisijaisuus.” March 28. <https://vm.fi/digitaalisen-viranomaisviestinnan-ensisijaisuus>.
- Ministry of Justice = Oikeusministeriö. 2018. *Selvitys yhdenvertaisuuden toteutumisesta sosiaali- ja terveystalveluiden valinnanvapauskokeiluissa. Painopisteenä kielellinen saavutettavuus*. Oikeusministeriön julkaisu No. 33/2018. With Laura Jauhola, Olli Oosi, Salla Rausmaa, and Janika Keinänen. <http://urn.fi/URN:ISBN:978-952-259-710-6>.
- Nielsen, Jakob. 1992. “Finding Usability Problems through Heuristic Evaluation.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York), CHI ’92, June 1, 373–80. <https://doi.org/10.1145/142750.142834>.
- Nielsen, Jakob. 1994. *Usability Engineering*. Interactive Technologies Series. Academic Press.
- Nielsen, Jakob. 2005. “Lower-Literacy Users: Writing for a Broad Consumer Audience.” Nielsen Norman Group, March 13. <https://www.nngroup.com/articles/writing-for-lower-literacy-users/>.
- Nielsen, Jakob, and Rolf Molich. 1990. “Heuristic Evaluation of User Interfaces.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York), CHI ’90, March 1, 249–56. <https://doi.org/10.1145/97243.97281>.
- Non-Discrimination Act (1325/2014) = Yhdenvertaisuuslaki (1325/2014). <https://www.finlex.fi/fi/laki/ajantasa/2014/20141325>.

- Nurminen, Mary, and Maarit Koponen. 2020. "Machine Translation and Fair Access to Information." *Translation Spaces* 9 (1): 150–69. <https://doi.org/10.1075/ts.00025.nur>.
- OECD. 2024a. *Do Adults Have the Skills They Need to Thrive in a Changing World?: Survey of Adult Skills 2023*. OECD Skills Studies. OECD Publishing. <https://doi.org/10.1787/b263dc5d-en>.
- OECD. 2024b. "Survey of Adults Skills 2023: Finland." December 10. [https://www.oecd.org/en/publications/survey-of-adults-skills-2023-country-notes\\_ab4f6b8c-en/finland\\_f8ab67af-en.html](https://www.oecd.org/en/publications/survey-of-adults-skills-2023-country-notes_ab4f6b8c-en/finland_f8ab67af-en.html).
- OED. 2024. "datum (n.), sense 1.a." In *Oxford English Dictionary*. December. <https://doi.org/10.1093/OED/2249415276>.
- OED. 2025. "preparation (n.), sense 1.2." In *Oxford English Dictionary*. September. <https://doi.org/10.1093/OED/3487262967>.
- Palojoki, Sari, and Riikka Vuokko. 2025. *Kanta Information System Services: Assessment of the Development of Digitalisation in Healthcare 2010–2024*. 2025:8. Reports and Memorandums of the Ministry of Social Affairs and Health. Ministry of Social Affairs and Health. <https://urn.fi/URN:ISBN:978-952-00-4197-7>.
- Pasquier, Martial. 2018. "Communication of and for Public Services." In *The Palgrave Handbook of Public Administration and Management in Europe*, edited by Edoardo Ongaro and Sandra Van Thiel. Palgrave Macmillan UK. [https://doi.org/10.1057/978-1-137-55269-3\\_23](https://doi.org/10.1057/978-1-137-55269-3_23).
- PLAIN = Plain Language Action and Information Network. n.d. "Why Use Plain Language?" Plainlanguage.Gov. Accessed October 21, 2024. <https://www.plainlanguage.gov/about/benefits/>.
- Prime Minister's Office = Valtioneuvoston kanslia. 2023. *Avoimesti ajassa - Valtionhallinnon viestintäsuositus 2023*. Valtioneuvoston kanslian julkaisu, 2023:11. Valtioneuvoston kanslia. <https://urn.fi/URN:ISBN:978-952-383-373-9>.
- Pym, Anthony, Nune Ayvazyan, and Jonathan Maurice Prioleau. 2022. "Should Raw Machine Translation Be Used for Public-Health Information? Suggestions for a Multilingual Communication Policy in Catalonia." *Just. Journal of Language Rights & Minorities, Revista de Drets Lingüístics i Minories* 1 (1–2): 1–2. <https://doi.org/10.7203/Just.1.24880>.
- Rathore, Farooq Azam, and Fareeha Farooq. 2020. "Information Overload and Infodemic in the COVID-19 Pandemic." *JPMA. The Journal of the Pakistan Medical Association* 70(Suppl 3) (5): S162–65. <https://doi.org/10.5455/JPMA.38>.
- Rello, Luz, Ricardo Baeza-Yates, Laura Dempere-Marco, and Horacio Saggion. 2013. "Frequent Words Improve Readability and Short Words Improve Understandability for People with Dyslexia." *Human-Computer Interaction – INTERACT 2013* (Berlin, Heidelberg), Lecture Notes in Computer Science, vol. 8120: 203–19. [https://doi.org/10.1007/978-3-642-40498-6\\_15](https://doi.org/10.1007/978-3-642-40498-6_15).
- Schuster, Michal. 2012. "Language Accessibility of Signage in Public Settings: A Case Study of a Health Care Service." *Southern African Linguistics and Applied Language Studies* 30 (3): 311–24. <https://doi.org/10.2989/16073614.2012.739411>.
- Selovuo, Kari. 2019. *Saavutettavuusopas*. 1st ed. Kari Selovuo.

- Statistics Finland = Tilastokeskus. 2024. "Nykyisen tasoinen nettomaahanmuutto ylläpitäisi väkiluvun kasvua." October 24. <https://stat.fi/julkaisu/cln1i9lg94af80bw1rxn0njfg>.
- Statistics Finland = Tilastokeskus. 2025. "Vieraskielisten määrä ylitti 600 000 henkilön rajan vuoden 2024 aikana | Tilastokeskus." April 4. <https://stat.fi/julkaisu/cm1jg8tr20lco07vwvoif9s6i>.
- Suojanen, Tytti, Kaisa Koskinen, and Tiina Tuominen. 2015. *User-Centered Translation. Translation Practices Explained*. Routledge.
- Suominen, Riitta. 2019. "Virkatekstin käytettävyys: Pöytäkirjasta verkkotekstiksi ja sosiaalisen median päivitykseksi." PhD diss, Tampereen yliopisto. <https://urn.fi/URN:ISBN:978-952-03-1174-2>.
- THL. 2024. "Maahanmuuttaneiden terveys- ja sosiaalipalvelut." Maahanmuutto ja kulttuurinen moninaisuus, July 4. <https://thl.fi/aiheet/maahanmuutto-ja-kulttuurinen-moninaisuus/maahanmuutto-ja-hyvinvointi/maahanmuuttaneiden-terveys-ja-sosiaalipalvelut>.
- Tirronen, Helena, Nina Smolander, and Tutta Tanntari. 2020. "Johdanto." In *Terveyspalvelut Tutuiksi: Mateas - Maahanmuuttajat Terveyspalveluiden Asiakkaina*, edited by Helena Tirronen, Nina Smolander, and Taija Karhe. Tampereen ammattikorkeakoulu.
- Torkkola, Sinikka, Helena Heikkinen, and Sirkka Tiainen. 2002. *Potilasohjeet ymmärrettäviksi: opas potilasohjeiden tekijöille*. Tammi.
- UN. 2006. *Convention on the Rights of Persons with Disabilities and Optional Protocol*. <http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf>.
- UN. n.d. "Universal Declaration of Human Rights." United Nations, United Nations. Accessed February 20, 2025. <https://www.un.org/en/about-us/universal-declaration-of-human-rights>.
- UNESCO. 2007. *Report by the Director-General on the Execution of the Programme Adopted by the General Conference: Intersectoral Mid-Term Strategy on Languages and Multilingualism*. 176 EX/INF.15. UNESCO, Executive Board, 176th, 2007. <https://unesdoc.unesco.org/ark:/48223/pf0000150335>.
- Vipunen. n.d. *Yliopistojen tutkinto-opiskelijoiden kansainvälinen liikkuvuus*. VIRT-Opintotietopalvelu. Accessed October 4, 2024. [https://vipunen.fi/fi-fi/\\_layouts/15/xlviewer.aspx?id=/fi-fi/Raportit/Yliopistojen%20tutkinto-opiskelijoiden%20kansainv%C3%A4linen%20liikkuvuus%20-%20vuosi.xlsb](https://vipunen.fi/fi-fi/_layouts/15/xlviewer.aspx?id=/fi-fi/Raportit/Yliopistojen%20tutkinto-opiskelijoiden%20kansainv%C3%A4linen%20liikkuvuus%20-%20vuosi.xlsb).
- W3C. 2021. "Use Clear, Unambiguous Formatting and Punctuation." Supplemental Guidance to WCAG 2, April 29. <https://www.w3.org/WAI/WCAG2/supplemental/patterns/o3p06-format-punctuation/>.
- W3C. 2023. "Web Content Accessibility Guidelines (WCAG) 2.2." October 5. <https://www.w3.org/TR/WCAG22/>.
- W3C. 2024a. "Cognitive Accessibility at W3C." Web Accessibility Initiative (WAI), April 25. <https://www.w3.org/WAI/cognitive/>.
- W3C. 2024b. "Introduction to Understanding WCAG 2." WCAG 2.2 Understanding Docs, August 27. <https://www.w3.org/WAI/WCAG22/Understanding/intro>.
- W3C. 2024c. "Introduction to Web Accessibility." Web Accessibility Initiative (WAI), March 7. <https://www.w3.org/WAI/fundamentals/accessibility-intro/>.

- W3C. 2024d. “WCAG 2 Overview.” Web Accessibility Initiative (WAI), March 7. <https://www.w3.org/WAI/standards-guidelines/wcag/>.
- W3C. 2025a. “W3C Accessibility Guidelines (WCAG) 3.0.” W3C Working Draft, September 4. <https://www.w3.org/TR/wcag-3.0/>.
- W3C. 2025b. “WCAG 3 Introduction.” Web Accessibility Initiative (WAI), September 4. <https://www.w3.org/WAI/standards-guidelines/wcag/wcag3-intro/>.
- Web Accessibility Directive = Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the Accessibility of the Websites and Mobile Applications of Public Sector Bodies (Text with EEA Relevance), CONSIL, EP, 327 OJ L (2016). <http://data.europa.eu/eli/dir/2016/2102/oj/eng>.
- WHO. 2023. “Disability.” World Health Organization, March 7. <https://www.who.int/news-room/fact-sheets/detail/disability-and-health>.
- WHO. 2024. “Health Literacy.” <https://www.who.int/news-room/fact-sheets/detail/health-literacy>.

# SUOMENKIELINEN LYHENNELMÄ

## ”WE INVEST IN THE INTELLIGIBILITY OF LANGUAGE”: ENGLANNINKIELISEN KANTA.FI-SIVUSTON KIELELLINEN SAAVUTETTAVUUS

### Johdanto

Digitaalisten palveluiden saavutettavuus on aihe, joka on noussut toistuvasti esiin julkisissa keskusteluissa erityisesti viimeisen vuosikymmenen aikana. Saavutettavuuden tavoitteena on varmistaa, että erilaiset tuotteet ja palvelut ovat käytettäviä ja ymmärrettäviä kaikille, riippumatta yksilöiden kyvyistä ja mahdollisista toimintaesteistä (ISO 2018). Tavoitteen saavuttamiseksi on olemassa monenlaisia ratkaisuja, ja saavutettavuus riippuukin lopulta aina yksilöstä ja hänen tarpeistaan: esimerkiksi elokuvasta saadaan sokeille käyttäjille saavutettava kuvailutulkkauksen avulla, kun taas kuurot ja heikkokuuloiset käyttäjät voivat seurata juonta tekstitysten ansiosta. Pohjimmiltaan saavutettavuudessa onkin kyse tasa-arvon toteutumisesta ja erityisesti vammaisten henkilöiden yhdenvertaisuuden varmistamisesta.

Tässä tutkielmassa keskitytään kielelliseen saavutettavuuteen, eli siihen, kuinka helposti kuka tahansa käyttäjä voi ymmärtää kielellistä materiaalia ja siten hyödyntää sitä tavoitteidensa saavuttamiseksi. Kielellinen saavutettavuus on olennainen osa saavutettavuutta, mutta se on jäänyt vähemmälle huomiolle lainsäädännössä ja ohjeistuksissa sekä tutkimuksissa. Käsite liittyy läheisesti ymmärrettävyyteen, jota on tutkittu laajastikin, mutta kielellisessä saavutettavuudessa näkökulma on yhteiskunnallisempi: kyse on siitä, kenellä on pääsy tietoon ja kyky ymmärtää se, ja lopulta siitä, kuka voi osallistua yhteiskunnan toimintaan (Heikkinen 2012, 43). Oikeus tietoon onkin yksi perustavanlaatuisista ihmisoikeuksista, ja kielenkäytöllä on keskeinen rooli sen toteutumisessa: jotta tiedon voidaan katsoa olevan saavutettavaa ja oikeuden tietoon toteutuvan, täytyy käyttäjällä olla paitsi pääsy tietoon, myös kyky ymmärtää ja sisäistää tieto siten, että sitä voi käyttää (Suojanen ym. 2015, 7).

Kielellinen saavutettavuus ja tiedon ymmärrettävyys ovat erityisen olennaisia silloin, kun kyseessä on elintärkeä aihe – kuten terveydenhuolto. Tämä pro gradu -tutkielma pyrkii vastaamaan kysymykseen siitä, kuinka kielellisesti saavutettava Suomen terveydenhuoltojärjestelmän keskeisten Kanta-palveluiden verkkosivusto kanta.fi on. Tarkemmin tutkielma keskittyy kansalaisille suunnatun osion englanninkieliseen käännökseen. Kanta-palvelut ovat kokoelma digitaalisia sosiaali- ja terveystalvueluita, jotka ovat käytössä koko maassa, ja joiden pääasiallinen tarkoitus on säilöä tietoja potilaasta ja hänen terveydentilastaan (Kanta Services 2024d, 2025a). Tästä syystä olisikin erityisen tärkeää, että tieto palveluista olisi saatavilla ja saavutettavissa kaikille palveluita käyttäville, kognitiivisista kyvyistä ja kielitaidosta riippumatta. Tutkimuskysymystä lähestytään heuristisen arvioinnin keinoin: arvioin kymmenen kanta.fi-sivustolta satunnaisesti valitun tekstin kielellistä saavutettavuutta luomieni heuristiikkojen avulla ja pyrin arvioinnin perusteella muodostamaan käsityksen tekstien ja koko sivuston kielellisestä saavutettavuudesta.

## **Kielellinen saavutettavuus**

Tässä tutkielmassa käsitellään verkkosivuston kielellistä saavutettavuutta, jonka määrittelen oikeusministeriön (2018, 19) sosiaali- ja terveystalvueluihin keskittyvään määritelmään pohjaten seuraavasti: kuinka hyvin kuka tahansa verkkosivuston käyttäjä, riippumatta hänen henkilökohtaisista ominaisuuksistaan tai kyvyistään, voi ymmärtää sivuston kielellisen materiaalin ja siten hyödyntää sitä tavoitteidensa saavuttamisessa. Kielelliseen materiaaliin lukeutuu kaikki verkkosivuston kielen avulla välittämä tieto käyttöliittymäteksteistä ohjeisiin ja videoihin, mutta tässä tutkielmassa keskityn ainoastaan kirjalliseen materiaaliin ja analysoin sivustolla julkaistuja informatiivisia ja ohjeistavia tekstejä rajaten esimerkiksi käyttöliittymätekstit aineiston ulkopuolelle. Erityisesti ohjeistavia ja informatiivisia tekstejä luetaan usein jotakin tiettyä tarkoitusta varten, ja tavoitteiden saavuttaminen onkin keskeinen osa kielellisen saavutettavuuden määritelmää: käyttäjän tulee kyetä paitsi ymmärtämään tekstin sisältö, myös yhdistelemään sitä muun tiedon kanssa ja käyttämään sitä tavoitteidensa saavuttamiseksi (Hirvonen ym. 2020, 24).

Digitaalisten tuotteiden ja palveluiden, kuten verkkosivustojen, saavutettavuudesta määrätään lainsäädännöllä, joista Suomessa keskeisimpiä ovat EU:n saavutettavuus- (2016/2102) ja esteettömyysdirektiivit (2019/882) sekä niihin pohjautuva kansallinen lainsäädäntö. Vuonna 2016 julkaistu saavutettavuusdirektiivi (2016/2102) vaatii viranomaisten verkkosivustoja ja

mobiilisovelluksia täyttämään W3C:n julkaisemiin verkkosisällön saavutettavuusohjeisiin (WCAG) pohjautuvat saavutettavuusvaatimukset. Näiden mukaan saavutettava verkkosisältö on havaittavaa, hallittavaa, ymmärrettävää ja toimintavarmaa (W3C 2023). Vaikka ymmärrettävyys onkin yksi WCAG:n pääperiaatteista, ei kielenkäytölle ole määritelty tarkkoja vaatimuksia tai ohjeistuksia, jotka varmistaisivat tiedon ymmärrettävyyden, vaan fokus on pääasiassa teknisissä ratkaisuisissa (Leskelä ja Uotila 2020, 233). Koska kielen ymmärtäminen ja muut kognitiiviset taidot sekä niihin liittyvät vaikeudet voivat vaihdella paljonkin yksilöiden välillä, on yleistettävien ohjeistusten laatiminen hankalaa (Selovuo 2019, 32). W3C pyrkii kuitenkin parantamaan kielellisen ja kognitiivisen saavutettavuuden huomioimista, ja se onkin yksi kehitteillä olevan WCAG 3:n päätavoitteista – ohjeistuksen on tarkoitus valmistua standardiksi muutaman vuoden kuluttua (W3C 2024a; 2025b). Kielenkäyttöön liittyviä vaatimuksia on myös esteettömyysdirektiivissä (2019/882), joka laajentaa saavutettavuusdirektiivin vaatimuksia ja tuo niiden piiriin useampia tuotteita ja palveluita sekä julkiselta että yksityiseltä sektorilta; muun muassa e-kirjojen, pankkipalveluiden sekä verkkokauppojen tulee täyttää direktiivin vaatimukset. Kielellinen saavutettavuus on siis aiheena ajankohtainen, sillä siihen liittyvän lainsäädännön voi olettaa laajenevan lähivuosina.

Vaikka saavutettavuuslainsäädäntö onkin suhteellisen uutta ja kehittyä jatkuvasti, eivät vaatimukset sen taustalla ole varsinaisesti uusia. Saavutettavuuden perimmäinen tavoite on erityisesti vammaisten henkilöiden yhdenvertaisen kohtelun ja ihmisoikeuksien toteutumisen varmistaminen, ja vaatimus yhdenvertaisuudesta sisältyy jo Suomen perustuslakiin (731/1999). Lisäksi vaikkei kielellistä saavutettavuutta tämänhetkisellä lainsäädännöllä suoraan vaaditakaan, on ajatus siitäkin kuitenkin jo olemassa, sillä perustuslain (731/1999) nojalla jokaisella on oikeus saada ja ymmärtää tietoa. Kielellä onkin merkittävä rooli yhdenvertaisuuden ja lopulta ihmisoikeuksien toteutumisen varmistamisessa: jos tieto on saatavilla vain sellaisella vieraalla kielellä, jota käyttäjä ei ymmärrä, tai jos käytetty kieli on liian vaikeaselkoista, ei oikeuden tietoon voida katsoa toteutuvan (Hirvonen ym. 2020, 19–20; Koponen ja Nurminen 2020, 304). Tutkielman fokuksesta huolimatta on myös huomioitava, ettei kielellinen saavutettavuus yksin takaa saavutettavuutta, vaan myös muiden saavutettavuuden osa-alueiden tulee olla kunnossa, jotta esimerkiksi verkkosivustoa voidaan sanoa saavutettavaksi; oli tieto kielellisesti kuinka saavutettavaa tahansa, ei se ole käyttäjälle aidosti saavutettavaa, jos hän ei pääse siihen käsiksi (Hirvonen ym. 2020, 20).

Kielellistä saavutettavuutta arvioidessa on huomioitava kielenkäytölle ominainen vaihtelu eri tilanteissa ja käyttötarkoituksissa, ja sen vaikutus kohdeyleisön tarpeisiin. Tutkimusaineistoni tekstit ovat koko kansalle suunnattua viranomaisviestintää, jonka tulisi olla kaikkien saatavilla ja ymmärrettävissä, ja joka siksi on usein kielellisesti suhteellisen yksinkertaista (Leskelä 2019, 93–94; Kotus n.d.). Viranomaisviestinnän ymmärrettävyyden yhteydessä puhutaankin usein selkeästä tai saavutettavasta yleiskielestä, jolla viitataan kaikille yhteisen yleiskielen muotoon, jossa on kiinnitetty erityishuomiota saavutettavuuteen (Leskelä 2019, 94). Tämä on myös se, mihin tässä tutkielmassa keskitytään. Saavutettava kieli voidaan kuitenkin nähdä myös yläkäsitteenä eritasoisille saavutettaville kielimuodoille, kuten keinotekoisesti luodulle selkokielelle, joka on suunniteltu erityisesti niille, joilla on vaikeuksia ymmärtää yleiskieltä (93). Selkokielen käytöllä on merkittävä vaikutus kielelliseen saavutettavuuteen, mutta aiheen tarkempi tarkastelu jää tutkielman rajauksen ulkopuolelle.

Saavutettavat ratkaisut on lähtökohtaisesti aina suunniteltu vammaiset ja muut niitä tarvitsevat henkilöt keskiössä, mutta ratkaisut usein hyödyttävät kaikkia – on esimerkiksi todettu, että monet verkkopalvelun saavutettavuutta parantavat muutokset parantavat myös palvelun yleistä käytettävyyttä (Selovuo 2019, 15). Näin on myös kielellisen saavutettavuuden kohdalla: vaikka sen huomioiminen palveleekin erityisesti heitä, joilla on vaikeuksia ymmärtää yleiskieltä esimerkiksi kognitiivisten vaikeuksien vuoksi, on kielellisesti saavutettava teksti kaikille käyttäjille nopeampi ja helpompi ymmärtää, erityisesti jos sen aihe on käyttäjälle entuudestaan tuntematon (Leskelä 2019, 57). Kognitiiviset kyvyt myös heikkenevät iän myötä ja voivat vaihdella päivänkin sisällä esimerkiksi väsymyksen vuoksi, joten jokainen tarvitsee kielellistä saavutettavuutta jossakin vaiheessa elämäänsä (102).

Kielellinen saavutettavuus hyödyttää myös niitä käyttäjiä, joiden lukutaito on syystä tai toisesta heikko, kuten esimerkiksi kielenoppijoita (Jylhä 2020, 67). Vaikka Suomessa lukutaito on yleisesti korkealla tasolla, on maahanmuuttajataustaisen väestön lukutaito kantaväestöä selkeästi heikempi (OECD 2024a, 85). Tämä on merkittävä huomio, sillä Suomen väestöstä yli kymmenen prosenttia on vieraskielisiä (Tilastokeskus 2025). Koko kansalle suunnatussa viranomaisviestinnässä olisikin tärkeää huomioida myös tuo osa väestöstä tarjoamalla tietoa kielellisesti saavutettavassa muodossa sekä mahdollisimman monella kielellä (Koponen ja Nurminen 2020, 307). Käännökset ovatkin olennainen osa saavutettavuuden varmistamista, mutta aihepiirin tarkempi käsittely on tämän tutkielman rajauksen ulkopuolella.

Viranomaisia veloitetaan tietyissä tilanteissa tarjoamaan käännös- ja tulkkauspalveluita, mutta laajempaa vieraskielistä viestintää ei lainsäädännöllä vaadita (Hallintolaki 434/2003). Kuitenkin suurin osa viranomaistahoista julkaisee vähintään osan verkkosisällöstään myös englanniksi, mihin kannustetaan myös valtionhallinnon viestintäsuosituksessa (Valtioneuvoston kanslia 2023, 11). On kuitenkin huomioitava, että merkittävä osa englanninkielisten tekstien käyttäjistä ei puhu kieltä äidinkielenään tai välttämättä edes sujuvasti (Kela 2025a, 8). Tämän vuoksi olisikin erityisen tärkeää panostaa paitsi suomen- ja ruotsinkielisten tekstien saavutettavuuteen, jolla voidaan tukea kielten oppimista, myös englanninkielisen materiaalin selkeyteen, jotta mahdollisimman moni voisi ymmärtää sen sisällön. Lisäksi erilaiset konekääntimet kuten Google Translate ovat hyvin suosittuja työkaluja erityisesti nuoremman väestön keskuudessa, ja ne tuottavat yleensä laadukkaampia ja oikeampia käännöksiä, kun lähtöteksti on kielellisesti saavutettava (Pym ym. 2022, 74; Koponen ja Nurminen 2020, 315–6).

## **Kanta.fi**

Tämän tutkielman tarkoituksena on selvittää, kuinka kielellisesti saavutettava kanta.fi-verkkosivuston englanninkielinen käännös on. Sivusto tarjoaa laajasti tietoa Kanta-palveluista sosiaali- ja terveysalan ammattilaisille ja palveluiden kehittäjille, mutta myös kansalaisille, joiden tietoja palveluissa säilötään – tässä tutkielmassa keskitynkin juuri kansalaisille suunnattuihin teksteihin. Kanta-palvelut ovat kokoelma digitaalisia palveluita, joiden pääasiallinen tavoite on sujuvoittaa terveyttä ja hyvinvointia koskevaa asiointia (Kanta Services 2025c). Suurin osa palveluista on suunnattu sosiaali- ja terveysalan ammattilaisille, jotka tallentavat kansalaisia koskevia tietoja palveluun, jotta muut hoitoon osallistuvat ammattilaiset ympäri maan voivat hyödyntää niitä (Kanta Services 2024d; 2024b). Kansalaiset voivat tarkastella itsestään järjestelmään tallennettuja tietoja sekä hoitaa joitakin terveydenhuoltoon liittyviä asioita, kuten uusia reseptejä, OmaKanta-palvelussa, johon kirjaututaan kanta.fi-sivuston kautta (Kanta Services 2025d).

Koska Kanta-palvelut ovat olennainen osa suomalaisia sosiaali- ja terveystaloudellisia palveluita, olisi erityisen tärkeää, että tieto palveluista ja sinne talletetuista tiedoista olisi saavutettavaa kaikille. Terveys ja terveydenhuolto ovat osa jokaisen elämää, ja kielellä on niissä merkittävä rooli: kun tieto ei ole saavutettavissa, ei käyttäjä välttämättä ole tietoinen oikeuksistaan ja pahimmillaan se voi jopa vaarantaa potilasturvallisuuden (Oikeusministeriö 2018, 20). Lisäksi

terveydenhuollon konteksti itsessään on usein stressaava, mikä heikentää kognitiivisia kykyjä ja lisää siten tarvetta kielelliselle saavutettavuudelle (Leskelä 2019, 109). Kanta.fi valikoituikin tutkimuskohteeksi juuri palveluiden keskeisyyden sekä terveydenhuollon tärkeyden vuoksi.

Suomessa oikeus terveydenhuoltoon ei riipu yksilön kielitaidosta: kaikilla, joilla on kotikunta Suomessa, sekä kaikilla lapsilla on oikeus käyttää terveystalvueluita kansalaisuudesta riippumatta (THL 2024). Lisäksi aikuisilla turvapaikanhakijoilla sekä paperittomilla aikuisilla on oikeus kiireelliseen terveydenhuoltoon (THL 2024). Olisikin ensisijaisen tärkeää, että myös tieto terveydenhuollosta ja sen palveluista olisi kaikkien saatavilla kielitaidosta riippumatta. On myös tärkeää huomioida, että maahan juuri muuttaneet eivät todennäköisesti tunne suomalaista terveydenhuoltojärjestelmää, jolloin mahdollisen kielimuurin lisäksi koko toimintaympäristö on uusi; tämä lisää stressiä ja korostaa entisestään kielellisen saavutettavuuden merkitystä. Viranomaisten on terveydenhoitotilanteissa huolehdittava käännös- ja tulkkauspalveluista, mutta vaatimus ei koske verkkosivustoilla olevia tekstejä. Eri viranomaisten verkkosivustot ovatkin tyypillisesti saatavilla suomen ja ruotsin lisäksi vain englanniksi, ja näin on myös kanta.fi:n kohdalla. Englanninkielisen käännöksen tarkoitettujen käyttäjien kielitaito vaihtelee, joten olisi erityisen tärkeää huolehtia käytetyn kielen saavutettavuudesta. Näistä syistä aineistoni koostuukin kanta.fi:n englanninkieliseltä sivustolta poimituista teksteistä.

Tutkimusaineistoksi valikoin englanninkieliseltä kanta.fi-sivustolta satunnaisesti kymmenen erilaista tekstiä. Kukin teksti on julkaistu omalla sivullaan, josta poimin analysoitavaksi vain varsinaisen sisällön, eli esimerkiksi erilaiset navigaatioelementit rajasin aineiston ulkopuolelle. Aineisto on esitelty taulukossa sivulla 35. Jokaiselle tekstille on myös annettu koodi, jota käytän myöhemmin viitatessani teksteihin.

## **Heuristinen arviointi**

Tässä tutkielmassa on käytetty tutkimusmenetelmänä heuristista arviointia. Kyseessä on alun perin käytettävyydetutkimuksen alalla kehitetty menetelmä, jonka tavoitteena on löytää arvioitavasta tuotteesta käytettävyydetongelmia (Nielsen 1994, 155). Menetelmän ytimessä ovat heuristiikat eli erilaiset listat ohjeistuksista, säännöistä tai periaatteista, joita tuotteen tulisi noudattaa (Suojanen ym. 2015, 77). Ongelmia etsitään vertailemalla tuotetta heuristiikkoihin, ja arvioinnin lopputuloksena on lista löydöksistä ja niiden rikkomista heuristiikoista (80).

Heuristisen arvioinnin etuna ovat sen nopeus, helppous ja tehokkuus: arvioinnin tekemiseen ei tarvita erityisiä työkaluja tai tiloja, eikä sen suorittaminen yleensä vaadi kovinkaan suurta ajallista panostusta arvioijilta (Suojanen ym. 2015, 93). Arvioijat ovat tyypillisesti joko tutkittavan tuotteen ja siihen liittyvän alan ja/tai käytettävyyden asiantuntijoita. Tutkimusten perusteella arvioijia olisi hyvä olla kolmesta viiteen, jotta suurin osa ongelmista löytyy (Nielsen 1992, 373). Toisaalta menetelmä perustuu täysin asiantuntija-arvioijien subjektiivisiin mielipiteisiin, sillä loppukäyttäjät eivät ole osa heuristista arviointiprosessia (Nielsen 1994, 160). Onkin todennäköistä, etteivät arvioijat löydä ainakaan kaikkia niitä ongelmia tai puutteita, jotka saattaisivat loppukäyttäjää häiritä. Tämä loppukäyttäjien puute on myös tämän tutkielman heikkous: koska en itse kuulu kanta.fi:n englanninkielisten tekstien pääasialliseen kohdeyleisöön, on hyvin todennäköistä, etten onnistu löytämään kaikkia niitä seikkoja, jotka saattavat vaikeuttaa tekstien ymmärtämistä. Lisäksi toimin ainoana arvioijana, mikä myös vaikuttaa löydösten määrään ja niiden luotettavuuteen.

Tässä tutkielmassa heuristista arviointia käytetään kuitenkin analyysin välineenä. Tarkoituksena ei siis ole tuottaa kattavaa listaa kaikista kanta.fi:n kielellisen saavutettavuuden ongelmista, vaan muodostaa kokonaiskuva sivuston kielellisestä saavutettavuudesta heuristisen arvioinnin tulosten avulla. Koska arvioinnin tuloksia ei siis käytetä sellaisenaan esimerkiksi ongelmien korjaamiseen, ei kaikkien puutteiden löytäminen ole välttämätöntä; yksittäiset puuttuvat löydökset eivät vaikuta merkittävästi kokonaiskuvaan. Toisaalta tulosten luotettavuutta arvioidessa on myös huomioitava se mahdollisuus, että minulta jäi löytämättä joitakin merkittäviäkin ongelmia. Useamman arvioijan mukaan ottaminen sekä loppukäyttäjien osallistaminen kielellisen saavutettavuuden arvioimiseen olisivatkin kiinnostavia jatkotutkimusaiheita.

Tätä tutkielmaa varten loin kielellisen saavutettavuuden ohjeistuksen, joka perustuu kandidaatintutkielmassani luomaani kriteeristöön kielellisen saavutettavuuden arvioimiseksi (Jantunen 2020, 11–13). Tähän tutkielmaan parantelin ja laajensin kriteeristöä aiheeseen liittyvien tutkimusten ja uusien ohjeistusten pohjalta: selkeytin joidenkin ohjeiden muotoilua sekä lisäsin viisi uutta ohjetta. Lisäyksien tavoitteena on tuoda paremmin esiin monissa ohjeissa jo implikoitu vaatimus yksiselitteisyydestä ja abstraktiuden välttämisestä, sillä niillä on merkittävä vaikutus kielelliseen saavutettavuuteen (Suojanen ym. 2015, 55). Lisäksi uudet ohjeet huomioivat erikseen linkkitekstit, ei-tekstuaalisen sisällön tekstivastineet sekä viittaukset tekstin ulkopuolisiin käyttöliittymäelementteihin. Nämä erityisesti näkövammaiset

käyttäjät huomioivat lisäykset ovat asioita, joita lainsäädännöllä itse asiassa jo vaaditaankin, sillä ne ovat osa WCAG:n onnistumiskriteereitä (W3C 2023).

Ohjeistus on kandidaatintutkielmani kriteeristön tapaan tarkoitettu toimimaan sekä heuristiikkana tekstin kielellisen saavutettavuuden arviointia varten että ohjenuorana tukemaan saavutettavan tekstin kirjoittamista. Sitä ei siis ole kustomoitu täysin tämän tutkielman tarpeita varten, vaan pyrkimyksenä oli kirjoittaa ohjeista mahdollisimman yleisiä, ottamatta tarkemmin kantaa esimerkiksi käytettyyn kieleen tai tekstilajiin. Ohjeet on jaettu neljään kategoriaan sen perusteella, mihin tekstin ominaisuuteen ne pääasiassa liittyvät, ja ne on esitelty alla taulukoissa 1–4. Kullekin ohjeelle on lisäksi annettu koodi, jota käytän viitatessani ohjeisiin myöhemmin.

### Taulukko 1. Kielioppi ja oikeinkirjoitus

Koodi	Ohje
G1	Noudata käyttämäsi kielen tai variantin kielioppi- ja oikeinkirjoitussääntöjä johdonmukaisesti.
G2	Kirjoita aktiivissa, ellei passiivin käyttöön ole erityistä syytä (passiivin käyttö voi olla perusteltua esimerkiksi tilanteissa, joissa tekijä on tuntematon tai epäoleellinen).
G3	Käytä yksinkertaisia virkerakenteita. Esimerkiksi englanniksi ja suomeksi kirjoittaessa suosi subjekti-verbi-objekti-rakennetta.
G4	Käytä verbien aikamuotoja johdonmukaisesti.
G5	Käytä verbejä kuvaamaan tekemistä. Esimerkiksi suomeksi käytä "he tutkivat" -rakenteita ja vältä "he suorittivat tutkinnan" -rakenteita.
G6	Varmista, että pronominiin viittaussuhteet ovat selkeitä. Pronominin sijaan voit toistaa subjektin, jos se tekee lauseesta helpommin ymmärrettävän.
G7	Vältä monitulkintaisuutta. Käytä selkeitä, yksiselitteisiä virkerakenteita ja anna käyttäjälle riittävästi tietoa. Kirjoita esimerkiksi päivämäärät auki pelkkien numeroiden käyttämisen sijasta.

### Taulukko 2. Sanasto

Koodi	Ohje
V1	Käytä yleisiä, jokapäiväisiä sanoja. Vältä ammattisanastoa ja muita termejä, joilla on erikoistunut merkitys. Korvaa pitkät ja epätavalliset sanat lyhyemmillä, yleisessä käytössä olevilla sanoilla.
V2	Selitä mahdollisesti vaikeasti ymmärrettävät sanat, jos niitä on pakko käyttää.
V3	Vältä lyhenteitä. Selitä kaikki epätavalliset lyhenteet ja akronyymit, kun ensimmäisen kerran käytät niitä.
V4	Käytä nimiä, termejä ja nimikkeitä johdonmukaisesti. Viittaa yhteen asiaan aina samalla nimityksellä.
V5	Vältä abstraktia ja kuvainnollista kieltä. Tarjoa konkreettisia esimerkkejä ja linkitä abstraktit käsitteet oikean maailman asioihin, jotka ovat käyttäjälle todennäköisemmin tuttuja.

### Taulukko 3. Rakenne

Koodi	Ohje
S1	Jäsentele teksti loogisesti. Voit esimerkiksi esittää asiat tekstissä siinä järjestyksessä, kun ne tapahtuvat.
S2	Jaa teksti loogisiin kokonaisuuksiin. Käytä väliotsikoita selkeyttämään tekstin rakennetta.
S3	Käytä kuvaavia otsikoita. Varmista, että otsikot vastaavat niitä seuraavien tekstikappaleiden sisältöä.
S4	Yhtenäisyyden lisäämiseksi tuo esiin tekstin eri osien, kuten virkkeiden ja tekstikappaleiden, välinen looginen suhde. Käytä esimerkiksi sidosteisuutta lisääviä sanoja kuten "lisäksi", "mutta" ja "toisaalta".

S5	Kerro tärkein asia ensin.
S6	Käytä luetteloita sen sijaan, että kirjoittaisit pitkiä virkkeitä, joissa on monia pilkuin eroteltuja sanoja tai lausekkeita.

#### Taulukko 4. Informaatiothiety ja tiedon selkeys

Koodi	Ohje
I1	Kerro käyttäjälle kaikki, mitä hänen tarvitsee tietää, mutta älä mitään ylimääräistä. Ohjaa käyttäjä tarvittaessa muualle lisätiedon pariin.
I2	Käsittele yhdessä tekstikappaleessa vain yhtä aihetta. Aloita uusi tekstikappale aina, kun esittelet uuden aiheen tai näkökulman.
I3	Käsittele vain yksi asia kussakin virkkeessä.
I4	Anna linkeille lyhyet, kuvaavat ja yksilölliset nimet ja käytä niitä johdonmukaisesti. Jos linkki johtaa tiedostoon tai erikieliseen materiaaliin, kerro siitä linkkitekstissä.
I5	Tarjota tekstivastineet kaikelle ei-tekstuaaliselle sisällölle kuten kuville, paitsi jos sisältö on puhtaasti koristeellista. Noudata saavutettavan kielen ohjeita. Tekstivastineen tulisi palvella samaa tarkoitusta ja välittää sama tieto kuin ei-tekstuaalinen sisältökin.
I6	Älä käytä pelkästään visuaalisia piirteitä kuten kokoa, muotoa, väriä tai sijaintia viitatessasi tekstin välittömän kontekstin ulkopuolisiin elementteihin, jotka käyttäjän tulee löytää.

Muodostaakseni käsityksen tutkimusaineiston kielellisestä saavutettavuudesta pyrin löytämään mahdollisimman monet ohjeistusta rikkovat kohdat aineistosta käymällä jokaisen tekstin läpi useamman kerran vertaillen niitä ohjeisiin. Vaikka onkin todennäköistä, etten löytänyt jokaista ongelmaa, uskon huolellisen läpikäynnin tulosten riittävän tämän tutkielman tarpeisiin. Heuristisen arvioinnin jälkeen analysoin löydöksiä ja niiden yleisyyttä tunnistaakseni tyypillisimmät ongelmat ja tehdäkseni päätelmiä aineiston ja mahdollisesti koko kanta.fi-sivuston kielellisestä saavutettavuudesta.

#### Keskeiset löydökset kanta.fi-sivuston kielellisestä saavutettavuudesta

Analyysin tavoitteena oli muodostaa käsitys aineiston kielellisestä saavutettavuudesta ja tunnistaa siinä tyypillisesti esiintyvät ongelmat. Vaikka analyysin pohjana toimikin heuristinen arviointi, keskittyvät tulokset yksittäisten ohjeiden tai ongelmien sijasta niiden kautta tunnistettaviin laajempiin teemoihin, joihin liittyviä löydöksiä oli kaikissa teksteissä. Löydöksissä korostuivat erityisesti monitulkintaisuus ja epäselvyys, epäjohdonmukaisuus sekä informaation määrään liittyvät ongelmat. Tällaisten kategoriarajat ylittävien teemojen löytyminen oli täysin odotettavissa, sillä kielen eri ominaisuudet ja rakenteet (ja siten niitä käsittelevät ohjeet) vaikuttavat väistämättä toisiinsa, eikä niitä ole mahdollista täysin eristää – eikä se ole tarkoituksenmukaistakaan.

Monitulkintaisuus on hyvä esimerkki ohjerajat ylittävästä ongelmasta: vaikka aiheelle oli oma ohjeensa (G7), vaikuttavat siihen monet muutkin kielelliset seikat, eikä kaikkia löydöksiä voi yksiselitteisesti luokitella vain yhtä ohjetta rikkoviksi. Ohjeet myös vaikuttavat toisiinsa

monella eri tasolla, sillä esimerkiksi aineistossa usein käytetyn *jotakin tehdään jonkun toimesta* -passiivirakenteen (G2) voidaan katsoa monimutkaistavan lauserakenteita (G3) ja siten aiheuttavan monitulkintaisuutta (G7). Kielioppi ja oikeinkirjoitus -kategorian ulkopuolelta myös esimerkiksi informaation puute (I1) ja abstrakti kieli (V5) vaikuttivat negatiivisesti tekstien selkeyteen. Monimutkaiset virkerakenteet ja epäselvät tai monitulkintaiset merkitykset lisäävät väärinymmärrysten riskiä ja vaikeuttavat tekstin ymmärtämistä; jos pelkkä virkkeiden tulkitseminen kuluttaa merkittävän osan kognitiivisesta kapasiteetista, ei sitä välttämättä riitä enää sisällön ymmärtämiseen.

Epäselvyyteen ja tekstin ymmärtämiseen vaadittavaan kognitiiviseen kuormaan vaikuttaa myös toinen tuloksissa korostunut teema: epäjohdonmukaisuus. Toisin kuin yksiselitteisyydelle, johdonmukaisuudelle ei ollut erillistä ohjetta, mutta ajatus siitä sisältyy koko ohjeistukseen, jota tulisi lähtökohtaisesti noudattaa johdonmukaisesti – joskin sujuva teksti vaatii myös jonkin verran joustavuutta. Johdonmukaisuus myös mainittiin erikseen monissa ohjeissa; esimerkiksi verbien aikamuotojen johdonmukaisuutta vaadittiin ohjeessa G4, johon liittyviä ongelmia olikin teksteissä jonkin verran. Suurin osa näistä ja muista kielioppiin ja oikeinkirjoitukseen liittyvistä johdonmukaisuuslöydöksistä oli kuitenkin luonteeltaan lieviä, eivätkä yksinään todennäköisesti aiheuta vakavia saavutettavuusongelmia.

Sanastoon liittyvissä löydöksissä sen sijaan epäjohdonmukaisuus nousi suuremmin esiin, ja ongelmat olivat luonteeltaan vakavampia. Sanaston johdonmukaisuudella, eli yhteen asiaan viittaamisella aina samalla nimityksellä (V4), on merkittävä rooli tekstin kielellisen saavutettavuuden varmistamisessa; yhdessä nimityksessä pysyminen varmistaa sen, että käyttäjä ymmärtää mihin viitataan. Aineistossa esiintyi toistuvia ongelmia tämän suhteen, sillä esimerkiksi palveluntarjoajiin viitattiin usealla eri nimikkeellä, jotka vaihtelivat jopa saman kappaleen sisällä. Lisäksi jotkin rakenteeseen liittyvät epäjohdonmukaisuudet, kuten otsikot, jotka eivät vastanneet niiden alla olevien tekstikappaleiden sisältöä (S3), heikensivät tekstien ymmärrettävyyttä ja selkeyttä.

Myös tekstin rakenteella on suuri vaikutus sen kielelliseen saavutettavuuteen. Rakenteen selkeydessä oli jonkin verran eroja eri tekstien välillä, mikä oli odotettavissakin: suhteellisen lyhyet ja yksinkertaiset ohjetekstit (kuten T2, T3 ja T4) olivat kokonaisrakenteeltaan selkeämpiä, kun taas pidemmissä, sivuston osioihin johdattelevissa teksteissä (kuten T1, T7 ja T8) oli enemmän rakenteellisia ongelmia – erityisesti hajanaisuus (S1) ja heikko sidosteisuus (S4) tekstin eri osien välillä korostuivat näissä teksteissä. Tekstitason rakenne oli kuitenkin

pääasiassa selkeä suurimmassa osassa aineiston teksteistä (S1), minkä lisäksi kaikki tekstit oli jaoteltu pienempiin, väliotsikoituihin kokonaisuuksiin (S2). Kun aineiston tekstien rakennetta tarkasteli tekstikappaleiden ja virkkeiden tasolla, nousi esiin erilaisia ongelmia: lauseiden tai virkkeiden välinen yhteys ei aina ollut täysin selkeä (S4) ja käyttäjälle olennaisinta asiaa ei aina esitetty ensimmäisenä (S5). Näihin liittyvät löydökset usein linkittyivät myös monitulkintaisuuteen (G7) ja puutteelliseen tietoon (I1).

Tiedon määrään (I1) liittyvät löydökset olivatkin kenties aineiston yleisimpiä, osin siitä syystä, että riittävän tiedon tarjoaminen linkittyy hyvin vahvasti muihin ohjeisiin ja niihin liittyviin löydöksiin; puutteellinen tieto voi esimerkiksi aiheuttaa monitulkintaisuutta tai abstraktiutta ja siten heikentää tekstin ymmärrettävyyttä. Myös selittämättä jätetyt erikoisanat (V2) ja muut epätyypilliset sanat (V1) voidaan yhdistää puutteelliseen tietoon. Vaikkei niihin liittyviä löydöksiä ollut suuria määriä, voi jo yksi vaikea sana estää käyttäjää ymmärtämästä virkettä tai jopa koko tekstiä, minkä vuoksi olisikin tärkeää, että teksteihin valittaisiin yleisiä sanoja ja vähemmän tunnetut käsitteet selitettäisiin auki.

Kaikkea tietoa ei tietenkään voi yhteen tekstiin sisällyttää, minkä vuoksi informatiivisissa teksteissä on usein paljon linkkejä, joita seuraamalla käyttäjä voi löytää lisätietoa tekstissä käsitellyistä aiheista (I1). Tätä strategiaa hyödynnettiin myös aineistossa laajalti, mutta siitä huolimatta jokaisessa tekstissä oli useita kohtia, joissa selite tai linkki lisätietoon olisi ollut tarpeen kielellisen saavutettavuuden varmistamiseksi. Koska aineisto koostui käännettyistä teksteistä, osa näistä löydöksistä selittynee käänösprosessilla; monissa teksteissä oli löydöksiä, joissa jokin terveydenhuoltojärjestelmän osa (kuten lääkkeiden suorakorvaus apteekissa) tai organisaatio olisi kaivannut lisäselitettä, jotta järjestelmää heikosti tunteva käyttäjä voisi ymmärtää mistä on kyse. Vaikka näistä voi olla hyötyä myös suomenkielisessä tekstissä, ovat suomalainen terveydenhuoltojärjestelmä ja sen palvelut todennäköisemmin tuntemattomampia englanninkielisen käännöksen yleisölle, ja tämä ero yleisöjen taustatiedoissa olisi tärkeää huomioida käänösprosessissa.

Myös analyysin kenties kriittisin löydös voi ainakin osin selittyä käänösprosessilla: kansalaisille suunnattu OmaKanta-palvelu on saatavilla vain suomeksi ja ruotsiksi, mutta tästä ei ole ainuttakaan mainintaa paitsi aineistossa, ei missään muuallakaan kanta.fi-sivustolla. Tämän voi olettaa johtuvan siitä, että maininnalle ei suomen- tai ruotsinkielisellä sivustolla ole tarvetta, eikä sitä ole syystä tai toisesta lisätty käännökseen. OmaKanta-palvelun nimi tosin on käännetty, ja palveluun viitataan englanninkielisellä sivustolla systemaattisesti sen

käännetyllä nimellä *MyKanta* – tästä voi helposti muodostua virheellinen käsitys siitä, että palvelu itsessäänkin olisi käännetty. Se, ettei OmaKanta-palvelua ole käännetty englanniksi on jo itsessään mainitsemisen arvioinen saavutettavuusongelma, vaikka OmaKanta ja käännösversioiden saatavuus ovatkin tämän työn rajauksen ulkopuolella.

OmaKantaan kuitenkin kirjaututaan kanta.fi-sivuston kautta, ja sivustolla on myös runsaasti tietoa ja ohjeistusta palveluun liittyen, myös englanniksi. Ohjeiden saatavuus englanniksi on toki positiivista saavutettavuuden kannalta, sillä niiden avulla myös suomea tai ruotsia taitamaton käyttäjä voi mahdollisesti käyttää OmaKantaa. Koska nuo ohjeet voivat olla joidenkin käyttäjien ainoa keino hyödyntää tätä keskeistä palvelua, olisi erityisen tärkeää huomioida niiden selkeys ja ymmärrettävyys. Lisäksi kaikissa OmaKantaan liittyvissä yhteyksissä olisi välttämätöntä mainita siitä, ettei palvelu ole saatavilla englanniksi, sillä nyt käyttäjät, jotka kirjautuvat OmaKantaan englanninkielisen kanta.fi-sivuston kautta päätyvät varoittamatta suomenkieliselle sivulle, jota ei ole lainkaan saatavilla englanniksi. Ennakoimaton kielen vaihtuminen voi hämmentää varsinkin käyttäjiä, joilla on kognitiivisia vaikeuksia ja on lisäksi erityisen ongelmallinen ruudunlukijaohjelmistoja käyttäville.

Tyypillisesti verkkosivuston saavutettavuus ruudunlukijaohjelmistojen käyttäjille yhdistetään pääasiassa teknisiin ratkaisuihin, mutta myös sisällöllisillä ja kielellisillä valinnoilla on merkittävä vaikutus siihen. Informaatiotiheys ja tiedon selkeys -kategorian kolme viimeistä ohjetta kattavatkin erityisesti ruudunlukijaohjelmistojen käyttäjiä hyödyntäviä kielellisen saavutettavuuden tekniikoita: linkkitekstit (I4), tekstivastineet (I5) sekä viittaukset käyttöliittymäelementteihin (I6). Ohjeeseen I6 liittyviä löydöksiä ei tosin ollut aineistossa lainkaan, mutta koska viittauksia käyttöliittymäelementteihin oli ylipäätään teksteissä hyvin vähän, ei tästä voida tehdä mitään yleisempiä päätelmiä ohjeeseen liittyen.

Linkejä ja linkkitekstejä (I4) sen sijaan oli aineistossa runsaasti, ja niihin liittyvistä ongelmista yleisimmät johtuivat ylimääräisestä tiedosta linkkiteksteissä. Jotta ruudunlukijaohjelmistoja käyttävät voivat tehokkaasti navigoida linkkien välillä, tulisi linkkitekstien olla uniikkeja ja lyhyitä sekä sisältää vain kontekstin kannalta olennaiset seikat. Erilaiset johdattelevat lausekkeet kuten *Read more* ('Lue lisää') tulisi siis yleensä jättää linkkitekstin ulkopuolelle, mitä ei aineistossa aina ollut tehty. Lisäksi muutama linkkiteksti oli suomeksi ja osasta puuttui maininta siitä, että linkki johtaa suomenkieliselle sivulle, mikä voi taas selittyä käännösprosessilla.

Myös ei-tekstuaalisen sisällön, kuten kuvien, tekstivastineilla (I5) on merkittävä rooli kielellisen saavutettavuuden varmistamisessa erityisesti ruudunlukijaohjelmistoja käyttäville. Kuvien tekstivastineiden tulisi välittää sama tieto kuin kuvassa ja koristeellisten kuvien tekstivastineiden tulisi olla tyhjiä, jotta ruudunlukijat jättävät ne huomioimatta. Tätä periaatetta noudatettiin pääosin hyvin, joskin muutamalla koristekuvalla oli tekstivastine ja yhdeltä tietoa välittävältä kuvalta se puuttui. Merkittävämmän ongelman muodostivat tekstivastineet, jotka eivät vastanneet kuvan sisältöä: kuvissa ja tekstivastineissa mainituista vuosiluvuista päätellen tämä johtui siitä, että kuvat oli päivitetty, mutta tekstivastineita ei. Jotkin tekstivastineet olivat myös pitkiä eivätkä noudattaneet kielellisen saavutettavuuden periaatteita; esimerkiksi virkkeiden välinen yhteys (S4) jäi monesti epäselväksi.

Kokonaisuudessaan aineistosta löytyi suuri määrä erilaisia kielellisen saavutettavuuden ongelmia. Suurin osa näistä löydöksistä oli luonteeltaan lieviä, mutta vaikeivat ne yksin välttämättä juurikaan heikentäisi saavutettavuutta, vaikuttavat ne kokonaiskuvaan, ja suuri määrä pieniä puutteita voikin lopulta aiheuttaa vakaviakin saavutettavuusongelmia. Lisäksi jo yksittäiset vakavat ongelmat kielellisessä saavutettavuudessa voivat pahimmillaan estää koko tekstin ymmärtämisen. Tuloksia tarkastellessa on kuitenkin myös huomioitava, että vaikka ne keskittyvät heuristiselle arvioinnille tyypilliseen tapaan teksteissä esiintyneisiin ongelmiin, monia ohjeita myös noudatettiin hyvin eikä yhtäkään ohjetta rikottu aina. Esimerkiksi nominaalimuotoisia verbejä (G5) ja lyhenteitä (V3) käytettiin vain vähän, tekstikappaleet olivat lyhyitä (I2) ja luetteloita hyödynnettiin runsaasti (S6). Löydösten perusteella tekstejä ei kuitenkaan voi pitää kielellisesti saavutettavina kaikille käyttäjille.

## **Päätelmät**

Tämän tutkielman tavoitteena oli selvittää, kuinka kielellisesti saavutettava englanninkielinen kanta.fi-verkkosivusto on ja minkälaisia saavutettavuusongelmia sivustolla esiintyy. Lähestyin tutkimuskysymystä suorittamalla heuristisen arvioinnin kymmenelle sivustolta satunnaisesti valitulle tekstille ja analysoimalla arvioinnin tuloksia kokonaiskuvan muodostamiseksi. Analyysin perusteella voi todeta, etteivät aineiston tekstit ole kielellisesti saavutettavia. Todennäköisesti käyttäjät, joilla on hyvä englannin kielen taito ja korkeintaan lieviä kognitiivisia vaikeuksia eivät koe suuria vaikeuksia tekstien ymmärtämisessä, erityisesti jos suomalainen terveydenhuoltojärjestelmä on heille jo tuttu. Se ei kuitenkaan riitä, vaan tekstien tulisi, erityisesti niiden konteksti ja tärkeys huomioiden, olla saavutettavia kaikille – myös

niille, joiden kieli- tai lukutaito taikka kognitiiviset kyvyt ovat heikompia. On tärkeää, että kaikilla on tieto palveluista, joihin heillä on oikeus sekä esimerkiksi siitä, mihin ja miksi heidän tietojaan tallennetaan.

Kaikki aineiston tekstit olivat kielelliseltä saavutettavuudeltaan suhteellisen samalla tasolla, ja samankaltaiset ongelmat toistuivat kaikissa teksteissä. Luonnollisesti jonkin verran vaihtelua eri tekstien välillä esiintyi, sillä esimerkiksi tekstin rakenteeseen liittyvät ongelmat olivat tyypillisempiä pidemmissä teksteissä ja ei-tekstuaalista sisältöä oli vain osassa teksteistä. Jokaisessa tekstissä oli kuitenkin löydöksiä kaikista ohjekategorioista ja löydösten suhteellinen määrä eri tekstien välillä oli suunnilleen sama. On tietysti huomioitava, että suoritin heuristisen arvioinnin yksin, enkä itse kuulu aineiston kohderyhmään. Useamman arvioijan sekä oikeiden käyttäjien mukaan ottaminen voisivat olla mielenkiintoisia jatkotutkimusaiheita, ja niiden avulla olisi mahdollista saada tarkempaa tietoa saavutettavuusongelmista. Tämän tutkielman tarkoitukseen tulokset ovat kuitenkin riittävät.

Tulokset olivat linjassa sen kanssa, mitä odotinkin löytäväni. Vaikka kanta.fi:n saavutettavuusselosteessa todetaankin heidän panostavan kielen ymmärrettävyyteen, on tuo ajatus ilmaistu tavalla, joka ei ole kielellisesti saavutettava: ”We invest in the intelligibility of language” (Kanta Services 2025b). Vastaavasti suomenkielisessä saavutettavuusselosteessa todetaan, että ”verkkosisällöt kielenhuolletaan ennen julkaisua” – *kielenhuoltaa* ei ole suomen kieleen vakiintunut verbi ja sen käyttö voi vaikeuttaa lauseen ymmärtämistä.

Löydösten johdonmukaisuudesta voi päätellä, että samankaltaisia ongelmia esiintyy todennäköisesti myös sivuston muissa teksteissä, eikä englanninkielistä kanta.fi-sivustoa siis voi pitää kielellisesti saavutettavana kaikille käyttäjille. Lisäksi analyysi antaa osviittaa siitä, minkälaisia ongelmia sivustolla todennäköisesti erityisesti esiintyy, ja tuloksia voisikin käyttää tukena sivuston kielellisen saavutettavuuden parantamiseksi. On kuitenkin pidettävä mielessä, että kielellinen saavutettavuus ei yksin riitä, vaan myös muut saavutettavuuden osa-alueet tulee huomioida sivuston toteutuksessa. Lisäksi eri kieliversioiden tarjoaminen lisäisi kielellistä saavutettavuutta. Sivuston kielellisen saavutettavuuden parantaminen olisi kuitenkin erityisen tärkeää, jotta voidaan varmistua siitä, että mahdollisimman moni ymmärtää tekstien tarjoaman tiedon ja osaa toimia sosiaali- ja terveydenhuoltojärjestelmässä – lopulta kyse on kuitenkin myös heidän ihmisoikeuksiensa toteutumisesta.