

LANGUAGE IN INSTAGRAM CAPTIONS:  
LANGUAGE CHOICES BETWEEN ENGLISH AND FINNISH

Iida Källroos  
Bachelor's thesis

University of Jyväskylä  
Department of Language and Communication  
English  
May 2022

## JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistinen tiedekunta	Laitos – Department Kieli- ja viestintätieteiden laitos
Tekijä – Author Iida Källroos	
Työn nimi – Title Language in Instagram captions: language choices between English and Finnish	
Oppiaine – Subject Englannin kieli	Työn laji – Level Kandidaatin tutkielma
Aika – Month and year Toukokuu 2022	Sivumäärä – Number of pages 30 + 6 liitettä
Tiivistelmä – Abstract <p>Kielellinen näkökulma jää usein huomioimatta, kun puhutaan ilmaisesta sosiaalisen median alustasta Instagramista, jonka käyttö perustuu pääosin visuaalisen sisällön jakamiseen. Instagram-kuvatekstin ja kielivalintojen avulla on mahdollista konkreettisesti ilmaista itseään sekä viestiä muun muassa ajatuksiaan, mielipiteitään, asenteitaan sekä arvojaan. Tämän tutkimuksen painopiste oli tutkia kielivalintoja kaksikielisten englanti-suomi Instagram-tilien kuvateksteissä, ja näin ollen yleisesti lisätä tietämystä ajatuksista kielivalintojen takana.</p> <p>Yksilöt käyttäytyvät kielellisesti eri tavoin eri tilanteissa, ja siksi tämä tutkimus pyrki selvittämään, kuinka kielivalintoja Instagram-kuvateksteihin tehdään englannin kielen ja suomen kielen välillä, sekä sivusi myös kuinka identiteetti on näkyvillä näissä kuvateksteissä. Tutkimus oli suunnattu henkilöille, joiden Instagram-tilit ovat kaksikielisiä, ja jotka kirjoittavat kuvatekstejä sekä englanniksi että suomeksi. Tutkimuksessa käytettiin sekä määrällisiä että laadullisia tutkimusmenetelmiä, ja tiedonkeräys tapahtui verkkokyselyn avulla. Tutkimukseen osallistui 57 henkilöä.</p> <p>Tutkimuksen tulokset osoittavat, että tekijät Instagram-kuvatekstien kielivalintojen takana niin englannin kuin suomen kohdalla ovat hyvin samanlaisia. Täten valinnat tehdään hyvin tilannesidonnaisesti, huomioiden esimerkiksi yksilön omat arvot, mielipiteet ja mieltymykset. Useimmiten mainittuihin syihin kielivalintojen taustalla lukeutuu englannin kohdalla esimerkiksi seuraajakunnan huomioiminen ja tietyn ilmauksen käyttäminen, sekä suomen kohdalla kontekstisidonnaisuus ja luonnollisuuden tuntu. Lisäksi suurin osa vastaajista koki oman identiteettinsä olevan näkyvillä heidän Instagram-kuvateksteissään.</p>	
Asiasanat – Keywords Language choice, Instagram caption, identity, language identity, social media, bilingualism	
Säilytyspaikka – Depository JYX	
Muita tietoja – Additional information	

## Table of contents

1 INTRODUCTION .....	1
2 LANGUAGE ON INSTAGRAM.....	1
2.1 COMMUNICATING ON INSTAGRAM.....	1
2.2 BILINGUALISM .....	2
2.3 CODE-SWITCHING .....	3
3 DIFFERENT IDENTITIES .....	4
3.1 IDENTITY AS A CONCEPT.....	4
3.2 ONLINE IDENTITY.....	5
3.3 LANGUAGE IDENTITY .....	6
4 PRESENT STUDY .....	6
4.1 AIM OF THE STUDY AND RESEARCH QUESTIONS.....	6
4.2 STRATEGIES OF THE STUDY .....	7
4.3 DATA COLLECTION AND METHODS OF THE STUDY.....	8
4.4 ETHICS .....	9
4.5 DATA ANALYSIS .....	10
4.6 QUESTIONNAIRE.....	11
5 RESULTS AND DISCUSSION.....	13
5.1 RESPONDENTS' LANGUAGE PROFICIENCY AND LANGUAGE USE.....	13
5.2 LANGUAGE CHOICE IN INSTAGRAM CAPTIONS .....	16
5.2.1 REASONS BEHIND CHOOSING ENGLISH FOR ONE'S INSTAGRAM CAPTION.....	16
5.2.2 REASONS BEHIND CHOOSING FINNISH FOR ONE'S INSTAGRAM CAPTION.....	19
5.3 IDENTITY IN INSTAGRAM CAPTIONS .....	22
6 CONCLUSION.....	25
BIBLIOGRAPHY .....	27
APPENDIX.....	31
THE QUESTIONNAIRE .....	31

## 1 INTRODUCTION

In this thesis, the focus is on the lingual aspect of Instagram and how one chooses which language to use in their captions. Additionally, identity representation in Instagram captions is also discussed. Instagram is a free social media platform where one can share visual content. It was founded in 2010 and nowadays is under Meta, a company brand that owns for example Facebook and WhatsApp as well (meta.com). Instagram has evolved to one of the most downloaded applications and it has over a billion monthly users all around the world (cyberclick.net, about.instagram.com).

I chose to investigate Instagram because it is mostly viewed as an app where one is able to express themselves mainly visually. The lingual aspect of Instagram is not highlighted, though the maximum character count for a caption in Instagram, 2,200 characters (Quintly 2019), enables a concrete tool to extensively express one's thoughts, opinions, attitudes and values under each post. There have been studies on language use or choice in other social media platforms, especially Facebook (Dovchin 2017; Cunliffe, Morris & Prys 2013). Therefore, as Instagram can be viewed as an equally linguistic platform as other social media, the focus of this study is precisely focused on language choice as in the context of bilingual Finnish-English Instagram accounts' captions. Alongside this, as identity is closely related to language (Thornborrow, cited in Thomas & Wareing 2000: 121, cited in Nurhantoro & Wulandari 2017: 38) the present study aims at investigating how it is visible in these captions and choices. These phenomena are examined using a mixed methods approach, that is both qualitatively and quantitatively.

## 2 LANGUAGE ON INSTAGRAM

As stated, language on Instagram plays an important role in the present study. Therefore, in this section, different, common ways of language use on Instagram are described as well as some relevant previous research related to the topic is presented.

### 2.1 COMMUNICATING ON INSTAGRAM

The description of Instagram (about.instagram.com) states that the application allows all to explore a community where individuals are able to be themselves and share everything from daily moments to the highlights of life. According to Calkin (2015), Instagram has evolved to

a platform where one constructs a picture of self. They point out that Instagram can act as an incentive for many things, both positive and negative, such as self-doubt, self-suppression, self-discovery and change in general. It is argued that the conceptualization of self is socially developed through interaction that takes place in the social world (Cote 2006, Kroger 1989, cited in Calkin 2015: 2), while also the process of constructing and reconstructing our identities is constant since humans build themselves through experiences (Calkin 2015: 2).

While some elements of communication in computer-mediated environments stay the same, often individuals modify their ways of communicating due to the context. In other words, this social and linguistic variation means that individuals apply different linguistic styles, elements and features in different online settings (Barton & Lee 2013: 6, Biber & Conrad 2009: 4). This phenomenon is called register, e.g. “a variety associated with a particular situation of use” (Biber & Conrad 2009: 6), and it is an important term in the present study. From the linguistic perspective, Biber and Conrad (2006: 9) suggest that certain linguistic features often occur in a register since they suit both; the purposes as well as the situation of the register.

## **2.2 BILINGUALISM**

This thesis is limited to be focused on bilingualism, rather than multilingualism, since my research concentrates on studying the choice between two languages: English and Finnish. According to Edwards (2004: 8), bilingualism has earlier been defined as mastering two languages. Nowadays, however, the definition varies a lot and even mastering one sentence of a foreign language can be considered as bilingualism (Edwards 2004: 7-8). All bilinguals might in fact be “non-fluent” bilinguals, meaning that the second language is weaker than the first, and might never reach native language-like competence (Edwards 2004: 10-11). Additionally, Auer (2007: 26) states that bilingualism is precisely applied for expressing identities when interacting.

Bilingualism can be structured and explained in various ways. One important distinction when it comes to this study is additive and subtractive bilingualism. The former refers to both languages being valued and useful, while the latter marks a situation where one language is valued more than the other. (Edwards 2004: 10-11) The choice between the use of Finnish and English language in the context of Instagram, a worldwide and international network, offers a platform to investigate this distinction and how it applies to the choice between Finnish and the main language of an application, English.

Measuring bilingualism reliably is a complex mission. There are five (5) universally defined basic language skills, in other words, modalities; speaking, listening, reading, writing (Edwards 2004: 8-9) and sign language (Quer & Steinbach 2019). However, according to Edwards (2004: 8-9), these skills include many subsections and elements that all should be evaluated in determining bilingualism. Furthermore, there are no set categories for different levels of bilingual proficiency, and neither are there determined ways for how it should be measured. One recommended way is self-ratings, which are applied in this study as well, even though they rely on the assumption that an individual is able to self-report as accurately as possible. (Edwards 2004: 8-9) However, some indicators for language proficiency have been developed. According to the Common European Framework of Reference for Languages, CEFR, language proficiency can be divided into six (6) levels. The CEFR is based on self-ratings of how one is able to perform linguistically. (Council of Europe)

### **2.3 CODE-SWITCHING**

Code-switching refers to a bilingual person's use of several languages or dialects in the same conversation or sentence (Gardner-Chloros 2009: 4). Barman et al. (2014: 14) state that code-switching is a common product of bilingual and multilingual language use. Research on what causes code-switching has been done a lot, but according to Gardner-Chloros (2009: 4), bilinguals often tend to switch languages in order to communicate something more than the superficial meaning of words. Code-switching helps in understanding how identities are formed and how individuals express their identities (Gardner-Chloros 2009: 5).

Code-switching can be described as a linguistic repertoire an individual has. From this repertoire, individuals are able to choose the type of language that they consider to be suitable for different occasions. (Joos 1967, cited in Edwards 2009: 30) Code-switching is very natural, even unnoticeable in individuals' daily lives. People are used to naturally choosing different selections from their linguistic repertoire and using them in certain situations such as with family, with co-workers or with friends. (Joos 1967, cited in Edwards 2009: 30) Sociolinguistics see code-switching as mainly a spoken genre, but it does have a long history in written language as well (Gardner-Chloros 2009: 20), and therefore this observation can be applied to written language too.

Hinrichs (2016) studied language choice in adult Facebook users' networks. When it comes to choosing English over their native language, in this study German, it is due to an attempt of

maximising the audience of a certain post. Moreover, the study explains language choice in general as an action of pragmatic considerations. (Hinrichs 2016: 36) As Edwards (2009: 30) puts it, the selections of individuals' speech repertoires can be considered as portraying different aspects of their identities, and when used, highlighting different features about their identities in different situations. Also Sultana, Dovchin and Pennycook (2013: 700) point out that rather than examining code-switching as precisely mixed codes, it should be viewed as "integrated forms of stylization" with which individuals present different aspects of their identities in certain situations.

### 3 DIFFERENT IDENTITIES

Identity is an extremely complex concept that can be divided into several categories and subclasses. In this section, a brief overview of the different identity types that are useful for this thesis are presented. This is relevant since I aim at investigating how they are visible in the respondents' lingual choices in their Instagram captions. Moreover, I am interested in finding out how the respondents themselves view the relationship between their identity and their Instagram captions.

#### 3.1 IDENTITY AS A CONCEPT

Edwards (2009: 19) states that "the essence of identity is similarity". By this they refer to the sameness an individual portrays no matter what time or situation. In other words, identity can be used to refer to the fact that a person is not trying to be someone else, but is being true to oneself. (Edwards 2009: 19) According to Groebner (2004, cited in Edwards 2009: 16), identity as a word has various definitional nuances. Therefore, it can be used to refer to many different types of identities and situations. Because of this, it is important to define which kind of identity is being discussed. It can refer to for example the subjective sense of self an individual has, to classification identifiers to oneself or others, or to identifiers of group membership. Therefore, it is also easy to be misused. (Edwards 2009: 16) Additionally, according to social constructionism, identity is constantly changing based on different factors in different contexts (Georgalou 2017: 10). Therefore, as Ivanic (1998, cited in Georgalou 2017: 10) states, "identity is not socially determined but socially constructed". In this study, identity is discussed precisely as the subjective sense of who we are.

### 3.2 ONLINE IDENTITY

One aspect of identity is online identity. Here, the word “online” is used to refer to self-representation through communication that takes place on networked devices and technologies (Barton & Lee 2013: 7). According to Russel and Stutzman (2007, cited in Georgalou 2017: 13), online identity can be referred to as “the representation of one’s persona in a digital context”. Therefore, identities are created and recreated as individuals carry out different sorts of actions, such as post text or pictures, online (Sundén 2003, cited in Georgalou 2017: 13). According to Lee (2014: 91), identifying identity as a multilateral concept enables one to understand the dynamics of its self-presentation online. This is an important issue to consider when conducting the present study.

As Georgalou (2017: 13) points out, interaction online happens through discourse. People are able to express different aspects of their identities online since most forms of online interaction are largely based on text without physical contextual clues. Online, individuals are given the possibility to be what they want or what others expect them to be by carefully choosing for example the lingual aspects of interaction, such as which language to use and how. These aspects, however, may not often be as explicit and assertive as in the real world, even though individuals are able to utilise various multimodal resources when expressing their identities. (Lee 2014: 92-94) On the contrary, Zhao, Grasmuck and Martin (2008) suggest that identity construction in an anonymous setting differs significantly from a non-anonymous setting as it limits the “freedom of identity claims” and tends to obey social norms (Brennan & Pettit 2004, Douglas & McGarty 2001; cited in Zhao et al. 2008). Furthermore, Barton and Lee (2013, cited in Georgalou 2017: 14) suggest that since nowadays many social practices tie offline and online activities together, possible differentiation should rather be done between offline and online situational contexts, where the communication in fact happens.

Nurhantoro and Wulandari (2017) studied the use of English on Instagram and the influence it has on Indonesian users’ identity. According to them, most people do not want to show their real selves on social media, but rather seek acceptance and appreciation from others by portraying themselves as the subject of attention. Therefore, identity as a concept is harder to comprehend. (Nurhantoro & Wulandari 2017) Furthermore, Lee (2014: 93-94) points out that it is important to remember that online identity is always situated in a particular context and therefore it should as well be interpreted as referring to that context. When it comes to this study, the context is Instagram posts, and more specifically, the captions of Instagram posts.



### **3.3 LANGUAGE IDENTITY**

According to Thornborrow (cited in Thomas & Wareing 2000: 121, cited in Nurhantoro & Wulandari 2017: 38), language is a basic factor when it comes to one's identity determination. Sapir (1921: 7, cited in Edwards 2009: 53) defines language as "a purely human and non-instinctive method of communicating ideas, emotions, and desires by means of a system of voluntarily produced symbols". In addition, Popper (cited in Edwards 2009: 54) names storytelling as one of the principal characteristics of human language. Therefore, it is clear that language is something that is very unique and a significant trademark for humans in particular. Language and identity are strongly connected to each other since by language use one builds their identity (Hassan 2010, cited in Nurhantoro & Wulandari 2017: 38) and portrays it to the outside world.

The relationship between language and identity is fundamental; identities are formed through language, whereas identifications guide individuals in how to use their linguistic resources as well as how to evaluate others' (Pavlenko & Blackledge 2008, cited in Lauring 2008: 348). When it comes to English language use on Instagram, Nurhantoro & Wulandari's (2017: 43) study suggests that it has both positive and negative effects on non-native speakers of English, and that it should not be used to impress, but for purposes in which it is necessary. In Europe in general, the idea of language and collective identity being closely tied is central and very deeply bound in Europeans' language ideologies. (Auer 2007: 1-2) In other words, individuals express their identities through something that is unique to them, for example the Finnish language, which also furthers the feeling of togetherness.

## **4 PRESENT STUDY**

This section consists of discussion considering the aim of the study and the research questions. Moreover, I will discuss the strategies of the study, methods and data collection as well as ethics. Later on this section the questionnaire conducted in order to carry out the present study is also presented.

### **4.1 AIM OF THE STUDY AND RESEARCH QUESTIONS**

The present study is focused on language choice between English and Finnish in Instagram captions. Further, I will also discuss whether and how identity can be seen in individuals' captions. The aim is to gain a better understanding on which factors affect these language

choices and how identities are visible in them. What is more, this study sheds light on the lingual aspect of a major social media platform whose main focus is usually on visuality. Therefore, themes such as code-switching and bilingualism are also discussed. In order to investigate these phenomena, I have come up with one main research question and one subquestion:

1. Which factors affect the choice to use English or Finnish in Instagram captions?
2. How is identity visible in these Instagram captions?

Both of these questions are meant to gather information and help in understanding language choices and identity representation. The present study is mainly qualitative with quantitative features, and aims for theory formation, in which a generalisation from individual cases is formed in a sense of understanding and explaining the problem in question (Varto 1992: 101). When it comes to the research problem, it is simply to gain understanding on this specific and current phenomenon, that is language choice in Instagram captions, which has little previous research. Thus, the aim is not to provide solutions or solve anything since any specific or correct answers for these questions cannot be defined. Understanding both language use in a social media environment as well as investigating how individuals make language choices in certain situations is the main idea of the present study.

## **4.2 STRATEGIES OF THE STUDY**

There are always multiple ways to conduct a study, and therefore it is essential for the researcher to define and decide the right strategy to follow (Denscombe 2014: 3). The present study is observant, meaning it investigates connections between phenomena, typical magnitudes and frequencies (Nummenmaa, Holopainen & Pulkkinen 2019: 17). In more detail, the present study is qualitative with some quantitative features since it can benefit from both aspects. While qualitative data offers a wider viewpoint of the respondents' many views and reasons behind their language choice, quantitative data helps in recognizing connections and frequencies between different factors (Denscombe 2014: 257-258). Therefore, a mixed methods approach was utilised.

A mixed methods approach combines aspects of different strategies and helps in investigating and comparing different perspectives when it comes to a certain phenomenon (Denscombe 2014: 4). For this study, strategies that gather both qualitative and quantitative data were used.

In order to collect opinions and facts as well as measure aspects of a certain phenomenon, a survey was used. Moreover, phenomenology for understanding specific issues from someone else's view was utilised. (Denscombe 2014: 4) In addition, ethnography is included in the present study since it aims at interpreting social interaction within a certain culture (Denscombe 2014: 4), in this case Instagram. By utilising these specific strategies, the sort of diverse data needed for the present study could be gathered. In other words, the respondents' opinions and perspectives could be documented as well as understood and further analysed.

### **4.3 DATA COLLECTION AND METHODS OF THE STUDY**

The data for this study was collected through a questionnaire, as it can be used to discover various things such as opinions and feelings. As Valli (2017) points out, it is a very versatile data collection tool, i.e. a method (Denscombe 2014: 3), that can be used for many purposes and can be carried out in different ways. Considering the present study, a questionnaire allowed to gather both qualitative and quantitative data easily at the same time. In this case, the questionnaire was carried out online using the Webropol online survey tool, as it was the most efficient way to reach a vast amount of respondents and make it easy for them to answer on their own terms whenever, wherever and with whatever device.

Other methods were also considered before choosing an online questionnaire, interviews being the other main option. Interacting face-to-face with the respondents could have given more precise and in-depth information on the phenomenon. However, there would have been some challenges in arranging the interviews. First, scheduling the long interview sessions would have been difficult since people can be quite busy. Second, the Covid-19 pandemic situation would have made discussing face-to-face more complicated and therefore was not important enough to risk people's health. Doing the interviews online was also considered but was ruled out mainly due to plausible scheduling problems and possible technical difficulties. Considering all these points, an online questionnaire was quite an obvious choice as it was the easiest to carry out as well as the most efficient timewise. Additionally, it offers a wide picture of the phenomenon in question as well as the largest number of answers.

As Valli (2017: 3) states, the questions of the survey should be carefully planned to achieve a successful study in order to avoid misunderstandings and gain accurate results. It is important that the questions are objective in order to achieve an image of the respondents' own views (Valli 2017: 3). According to Denscombe's (2014: 166) qualifications for research

questionnaire, the questionnaire conducted for the present study provides respondents with the exact same set of questions, enabling consistency and precision. Additionally, the information comes directly from the respondents without any intermediary actors involved (Denscombe 2014: 166).

Both qualitative and quantitative data was collected in the questionnaire. The qualitative data is from open-ended questions, and therefore in the form of free-written text written by the respondents. Later, it is presented how these responses are classified in order to investigate them better. The quantitative data collected in the present study, on the other hand, can be divided into different set categories. First is nominal, which “represents something to which a name has been attached”, in this case gender and native language (Denscombe 2014: 250). Ordinal data, on the other hand, can be collected from Likert scales where respondents are able to position their response on a scale based on whether they agree with a statement or not (Denscombe 2014: 251, Marshall 2005: 133). Lastly, continuous data in this study refers to the respondents’ age since it is collected in sets of five to ten years (Denscombe 2014: 252). As quantitative data is based on mathematics and probability, it can further be presented in various statistical techniques (Denscombe 2014: 272).

#### **4.4 ETHICS**

When it comes to conducting research, ethics is an important issue to consider. Since anonymity and processing personal data is a major part of the present study, I have taken certain measures based on the University of Jyväskylä’s (jyu.fi 2021) instructions regarding the processing of personal data in conducting it. Additionally, Denscombe’s (2014: 310) four core principles for social research, i.e. the protection of the respondents’ interests, ensuring voluntary participation based on informed consent, avoiding deception and operation with scientific integrity, and compliance with the laws of the land, are implemented.

First, a privacy notice is presented before being able to answer the questionnaire. In this privacy notice, the aims and purposes for the present study are explained. Additionally, it is made clear that the results are used as data for the present study. Moreover, it is explained that participation is anonymous and voluntary. Lastly, respondents are not able to continue to the actual questionnaire before they have answered “yes” to a question that ensures they have read the instructions, are at least 18 years old as well as give permission to use the answers as research data. After this, background questions about one’s age, gender and native language were asked.

Even though one of the most important aspects of the survey is its anonymity, this information is needed in order to be able to analyse the results and make certain conclusions and divisions. Age and gender are a major factor when it comes to behaviour in social media (DNA 2021), and knowing the respondents' native language is relevant considering the topic of the research. In order to achieve anonymity, other questions included in the questionnaire were chosen with care. According to the minimisation principle (jyu.fi 2021), everything unnecessary as well as anything that could reveal the respondents' identity was left out.

When it comes to the collected data and its preservation, the data was preserved in Webropol, on paper and as a document on my computer, and was processed only by me as the conductor of the present study. Processing the data included printing it on paper in order to investigate the qualitative data thoroughly, and working with the quantitative data on a statistical software. Lastly, the data was disposed of immediately as the present study was finished and there was no need for it anymore.

#### **4.5 DATA ANALYSIS**

In order to successfully complete a research, the data must be analyzed properly (Gibson & Brown 2009: 1). However, as Gibson and Brown (2009: 1) state, analyzing the data is one of the most problematic stages of social research. In the present study, content analysis for qualitative aspects and statistical analysis for quantitative aspects were used.

Generally speaking, qualitative analysis searches for “general statements about relationships and underlying themes” (Marshall & Rossman 2006: 154, cited in Gibson & Brown 2009: 4). When it comes to content-based analysis, its purpose is to create a theoretical entity of the research material where it can be analyzed systematically as well as objectively (Tuomi & Sarajärvi 2009: 95). Qualitative content analysis, QCA, aims at systematically describing the meaning of the collected data, and the research questions determine the angle from which the data is analysed. (Schreier 2012: 3-4) One of its main strengths is that it provides a way of quantifying the written data. Furthermore, qualitative data in its original form can be hard to investigate, hence it should be first organized in a way that enables systematic and meaningful analysis. (Denscombe 2014: 284, 277) Therefore, certain aspects from the data are selected for the analysis (Schreier 2012: 3-4). In the present study, QCA is executed by first classifying the data based on frequency of certain features (Denscombe 2014: 283-284), which subsequently helps in analyzing and describing those in more detail (Gibson & Brown 2009: 31).

In general, quantitative data analyses should be based on objective laws rather than the researcher's values (Denscombe 2014: 272). More specifically, statistics apply mathematical principles to the conduct of a research (Thomas 2014: 1). Since the findings of statistical analysis are based on measured quantities, they offer a strong foundation for further description (Denscombe 2014: 272) that can be turned into useful knowledge (Thomas 2014: 9). In the present study, numeric data from closed-ended questions was analyzed by using statistical analysis. More precisely, the data was analyzed by organizing it in illustrative tables and cross tables, and further comparing the data in order to find the most significant results. Subsequently, these results can be presented concisely and understandably (Nummenmaa et al. 2019: 10) for example with different tables and figures.

Lastly, when it comes to analyzing the results, knowing the difference between making conclusions deductively or inductively is essential. Deductive reasoning stems from premises known to be true, and it aims at revealing information already included in the premises. Therefore, it does not reveal any new information, whereas inductive reasoning aims at precisely that. It is based on the assumption that new observations are similar to old, and that the future complies with the past. (Nummenmaa et al. 2019: 12) As the assumption is that the respondents answer the questionnaire based on what is true to them, the analysis leans on to the deductive side (Nummenmaa et al. 2019: 12).

#### **4.6 QUESTIONNAIRE**

The questionnaire consisted of eight (8) questions, excluding the first question considering consent, and the three (3) background questions considering age, gender and native language. Four (4) of the questions were multiple choice, offering the respondent alternatives from which they were able to choose one or more (Marshall 2005: 132), and providing accurate data and numbers for the researcher. Two (2) were Likert scale questions which provided the respondent with a scale where they could choose an option based on the level in which they agree with a statement (Marshall 2005: 133). Two (2) open-ended questions and a chance to elaborate some multiple choice answers were also included in order to gain a wider understanding of the qualitative side of the respondents' views and language use.

The intended respondents for the present study were people who write their Instagram captions using both English and Finnish. Therefore, there were no other criteria for participation than to

produce both of these languages at some level in one's Instagram captions. Responses were accepted for fifteen (15) days between February and March 2022. Altogether 57 people participated in the survey but five (5) responses were discarded since they did not meet this criteria. Therefore, 52 responses are considered in the analysis. Figure 1 shows the respondents' age whereas Figure 2 presents their gender. Lastly, in Figure 3 one can see how the respondents' native languages are divided.

Figure 1. The respondents' age

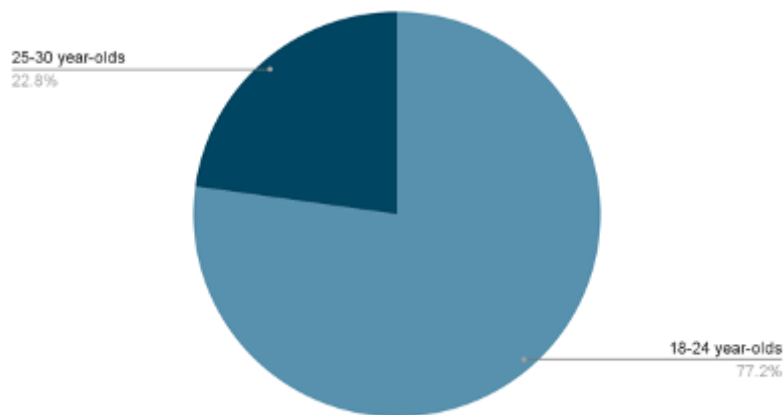


Figure 2. The respondents' gender

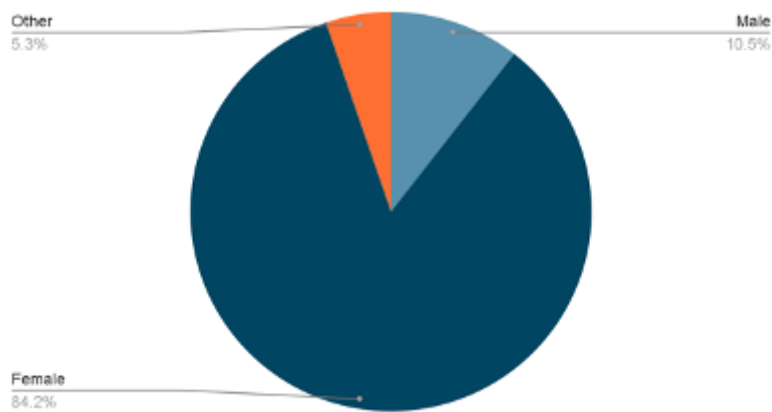
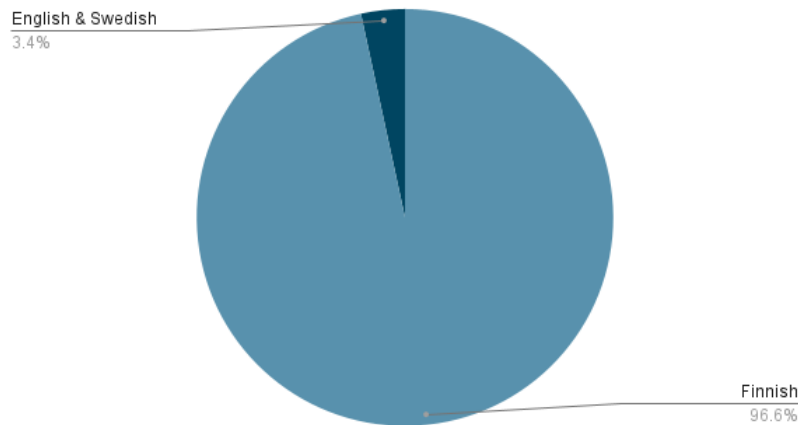


Figure 3. The respondents' native languages



Lastly, even though over 50 answers were gathered in the present study, the results cannot be generalized due to reaching only a certain group of respondents. Firstly, all the respondents are under 30 years old, and the vast majority of them are women. Therefore, all other age groups are left outside investigation and the male perspective on the phenomena is also very limited due to having only five (5) male respondents. Secondly, this study was designed to examine language choice only between English and Finnish, and therefore many views considering language choice in general and among all other languages were left out. Thirdly, even though the study was not targeted to only Finnish people, all the respondents reported Finnish as at least one of their native languages which again narrows the perspective on the phenomenon.

## 5 RESULTS AND DISCUSSION

This section includes discussion about the results of the questionnaire as well as comparison of the results and the theoretical background that was presented earlier. Next, the results are presented and further discussed alongside illustrative charts.

### 5.1 RESPONDENTS' LANGUAGE PROFICIENCY AND LANGUAGE USE

Before investigating the respondents' language choices, it is important to gain an understanding of their Finnish and English language proficiencies as well as how frequently these languages are used in their Instagram captions. Thus, the respondents were asked to evaluate their language proficiencies based on descriptions of six (6) Common Reference Levels, which can be divided into three (3) categories: basic user (A1 and A2), independent user (B1 and B2) and proficient user (C1 and C2) (CEFR 2020: 36). Next, the respondents' English and Finnish language proficiencies are presented in Tables 1 and 2 below.



Table 1. Respondents' English language proficiency

CEFR level	n	%
A1	0	0,0%
A2	0	0,0%
B1	0	0,0%
B2	4	7,7%
C1	25	48,1%
C2	23	44,2%

Table 2. Respondents' Finnish language proficiency

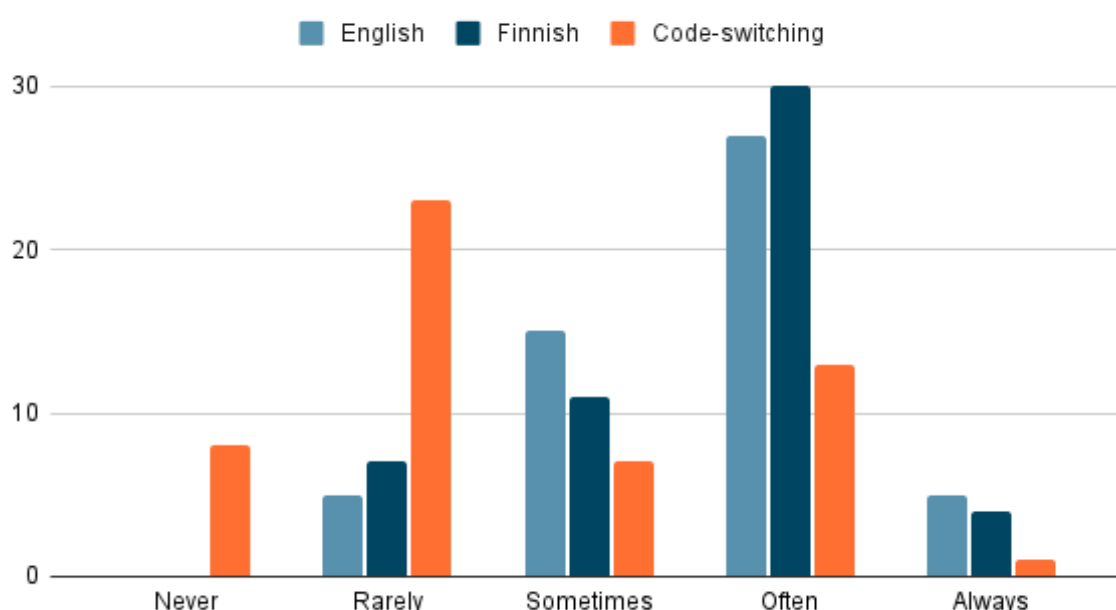
CEFR level	n	%
A1	0	0,0%
A2	0	0,0%
B1	0	0,0%
B2	0	0,0%
C1	2	3,8%
C2	50	96,2%

As the vast majority of the respondents have Finnish as their only native language (Figure 3), their Finnish language proficiency is obviously high (Table 2). However, notable is that nearly 90% of the respondents rated their English language proficiency on the two highest levels as well (Table 1). When considering the respondents' average age and native language, it could be inferred that they have received high quality English language education in Finnish schools. To demonstrate, English is the first foreign language Finnish-speaking pupils, that is nearly 90% of all the pupils, are taught at the beginning of third grade at the latest (Finnish National Agency for Education n.d.).

Moreover, the respondents reported the frequency of their Finnish and English language use in Instagram captions with a Likert scale, choosing one of five (5) time-bound options. As Figure 4 below shows, the differences between English and Finnish in each option are not as major. However, what is striking is that when it comes to both English and Finnish, 'often' was the

most frequent answer whereas ‘never’ was not answered at all. Thus, leaning onto Joos’ (1967, cited in Edwards 2009: 30) suggestion regarding linguistic repertoire, it can be suggested that the respondents use both of these languages in certain situations rather than completely excluding either of them from their repertoire when it comes to language use in Instagram captions. Moreover, these results can be investigated from the perspective of linguistic variation, also known as register. In other words, individuals choose the language that best suits their purposes and the situation in which they will use the language in. (Biber & Conrad 2009: 6)

Figure 4. Language use in Instagram captions (n)



Additionally, the majority of the respondents code-switch between English and Finnish in their Instagram captions never or rarely (Figure 4), which implies that individuals are more likely to choose one language over using both English and Finnish at the same time. The reasons behind this are not investigated more thoroughly in the present study since they are not the main subject of attention. However, as one respondent points out, after a translate function for captions was added to Instagram enabling translation from Finnish to English, their English language use has decreased. Therefore, it is important to notice that this sort of code-switching or bilingualism can be seen in captions even though one does nothing in order to enable it. Lastly, the possibility of code-switching in other languages is not considered here, and therefore including that aspect could affect the results.

## 5.2 LANGUAGE CHOICE IN INSTAGRAM CAPTIONS

The two open-ended questions in the questionnaire represent the qualitative side of the study. In order to get a wider outlook on the answers and observe them better, I created classes based on frequently appearing reasons for language choice and classified all of the answers into these classes. The classes help in recognizing the reasons for language choice which appear the most often in the present study. This way, small entities can be created and analyzed more thoroughly than one vast one, in order to make conclusions in a systematic and objective way (Tuomi & Sarajärvi 2009: 95). Majority of the answers include features from at least one of these classes, some fit into several classes, and a few do not fit in any of the classes. When it comes to analyzing these responses, the analysis is made from a deductive approach since even though the data is based on individuals' personal opinions and viewpoints, it is true to the respondents and therefore considered factual information. With a deductive approach, new information can be stemmed from information that already exists (Nummenmaa et al. : 12). Next, a closer look into the responses is taken alongside some excerpts from the data.

### 5.2.1 REASONS BEHIND CHOOSING ENGLISH FOR ONE'S INSTAGRAM CAPTION

The first one of the open-ended questions was: *Why do you use English in your Instagram captions instead of Finnish?* From the 52 responses, five (5) reasons for choosing English over Finnish differentiated the most frequently. According to these, I created five (5) classes which are presented alongside the number of responses which had features from these classes.

Table 3. Classification of reasons behind choosing English for one's Instagram caption

Class	Number of features in the responses
English feels more natural or allows easier expression than Finnish.	16
International followers/most followers understand English.	30
There is a certain English expression one uses.	20
English fits the context of the picture better than Finnish.	8
English sounds/is better/cooler than Finnish.	15

Out of all the 52 responses, the most common reason for choosing to use English over Finnish in one's Instagram caption was that everyone, i.e. all followers, would be able to understand it. Similarly to Hinrichs' (2016: 36) findings in his study, by choosing to use English over one's native language, one is able to reach a wider audience while at the same time including international followers. As the excerpts below illustrate, respondents rationalized that most Finnish followers are able to understand English as well, which means the rate of inclusion and understanding when it comes to the caption is higher. This in fact supports Hinrichs' (2016: 36) view on language choice as an act of pragmatic consideration, as well as Nurhantoro and Wulandari's (2017: 43) suggestion that non-native speakers of English should use English on Instagram when necessary rather than when trying to impress.

- (1) *"I have friends from abroad who do not speak Finnish. While Instagram offers translation, it usually fails because I write in dialect Finnish.--"*
- (2) *"-- I consider English 'the universal language of social media'."*
- (3) *"-- I'm trying to reach an international audience--"*

Respondents explained their language choice also by wanting to use a certain English expression, such as lyrics, quotes, jokes, words or memes, since they work better in English than in Finnish. As Joos (1967, cited in Edwards 2009: 30) states, individuals possess a linguistic repertoire from which they can choose the sort of language they want to use in specific situations. What is more, individuals can have certain expressions stored in their repertoires even though they do not master English. Therefore, as Edwards (2004: 7-8) suggests, they can portray a certain image of themselves to the outside world and be considered as bilinguals. Furthermore, as can be seen in excerpt 5 below, the difficulty of translating certain expressions accurately enough can affect one's language choice.

- (4) *"-- straight or modified quotes from music, books, etc. and since I consume most of my media in English, they will also be in English."*
- (5) *"-- not all English expressions can be translated into Finnish (for example "swipe for surprise" would be difficult and stiff to translate)."*

Additionally, it was reported that choosing English for one's caption is due to finding it a more natural or easier way of expression than Finnish. Some respondents think in English or prefer speaking it, hence their usage of English in Instagram as well. As Joos (1967, cited in Edwards 2009: 30) states, choosing which language to use can be very natural and unnoticeable.

Preferring English as the language to express oneself could be considered as an act of highlighting a certain part of one's language identity (Edwards 2009: 30), as can be seen in this response: *"--in some way I like to show my fluency in it and how comfortable I am when using English, even when it's not my main language"*.

Respondents also opt for English since it sounds better or cooler than Finnish, in general or in a certain situation. One respondent even mentioned that the appearance of English affects their language choice, as illustrated in excerpt 8 below. These views could be analyzed from the perspective of stylization and register, i.e. the linguistic means used in a certain situation. For example Sultana et al. (2013: 700) propose that using different languages should be considered as forms of stylization, and therefore individuals choose to portray different sides of their identities. Based on this perception, choosing to use English because of its coolness or superiority is due to seeking a certain style or register in one's caption; either appearance or content wise. Moreover, finding English somehow superior clearly refers to subtractive bilingualism, that is a situation where one language is valued more than another (Edwards 2004: 10-11).

- (6) *"I also feel like some sentences sound "dumb" in Finnish and way cooler in English."*
- (7) *"Sometimes English just sounds better, smoother or more "artistic" in the caption."*
- (8) *"-- monet sanat ja lauseet näyttää nätimältä enkuksi '- many words and sentences look prettier in English'"*

Lastly, according to some respondents, the context of the picture has an effect on the language they choose for their caption. As can be seen in the excerpts below, it is argued that if the topic, event, situation or people present in the picture are international, English suits the caption better. This view can also be supported by Hinrichs' (2016: 36) suggestion that language choice is in fact as simple as considering practical matters. For example, enabling understanding the caption for possible international people present in a picture can be quite an obvious choice when writing a caption.

- (9) *"I usually use English if the topic of the picture(s) is more international, e.g. from my travels, of my international friends etc."*
- (10) *"I mainly use English phrases like "see you later alligator" (since I have a picture of me holding an alligator), because the[y] fit the picture well and there is no Finnish equivalent."*

### 5.2.2 REASONS BEHIND CHOOSING FINNISH FOR ONE'S INSTAGRAM CAPTION

The second open-ended question was: *Why do you use Finnish in your Instagram captions instead of English?* In the 52 responses to this question, similarly to the responses of the first open-ended question, five (5) reasons were emphasized in the data. Therefore, the classes I created as well as the number of responses that fit them are presented next.

Table 4. Classification of reasons behind choosing Finnish for one's Instagram caption

Class	Number of features in the responses
Finnish feels natural or enables easier expression than English.	25
Most followers are Finnish or all Finnish followers do not understand English.	13
There is a certain Finnish expression one uses.	8
Finnish fits the context of the picture better than English.	15
Finnish is more suitable for longer, story-like captions and expressing emotions than English.	9

When it comes to choosing Finnish over English for one's Instagram caption, the reason mentioned the most often was that it feels the most natural or easy to use. This is no surprise, since all of the respondents report Finnish as at least one of their native languages (Figure 3). As discussed earlier, language and collective identity seen as very intertwined is deeply rooted in European's language ideologies (Auer 2007: 1-2). With this and the respondents' native language distribution (Figure 3) in mind, it is understandable that the results indicate one of the main reasons behind choosing Finnish being the natural feel of it. More specifically, identity expression through one's native language, which can be viewed as unique and personal, can also promote the feeling of togetherness within a community (Auer 2007: 1-2). Therefore, when Finnish individuals express themselves in Finnish for their presumably Finnish-dominant follower group, it could be concluded that they are likely to feel part of a community and therefore more understood and comfortable.

(11) *"I feel that I can express myself more precisely in it and be more playful and creative with the language."*

(12) *"Because Finnish is my native language and it would feel a bit pretentious, in a way, to use English--"*

Ensuring understanding for all Finnish followers was emphasized in the responses as well. Some respondents mentioned having relatives and older people with weaker English language competence as their followers. It is natural that individuals choose the language they find the most suitable for a certain situation, according to Joos (1967, cited in Edwards 2009: 30). Clearly, people want to take their different types of followers into account when choosing the language for their caption. Individuals include certain linguistic features and elements in these registers since in addition to the situation, they suit the individual's purposes as well (Conrad 2006: 9). Additionally, what also seemed to be behind choosing Finnish was that people avoid giving a certain image of themselves to their Finnish followers and people they personally know. To demonstrate, one respondent rationalized: *"--It might seem like I think I'm posting for a huge, multinational and -cultural audience, when in fact less than 10% of my followers are from some other country than Finland --"*. Having these "offline-based online relationships", i.e. anchored relationships (Zhao 2006, cited in Zhao et al. 2008), in one's follower group can change the way one communicates linguistically online. To specify, these anchored relationships can be originated for example from mutual friends, education, residence or pastime activities (Zhao et al. 2008). A similar observation considering acceptance and appreciation from followers appeared in Nurhantoro & Wulandari's (2017) study about English use and its influence on Indonesians.

(13) *"-- I would feel a little weird about using solely English, since most of my followers are Finnish-speaking as well. It just wouldn't make sense to me."*

(14) *"All of my followers are Finnish and some of them don't understand English that well.-- I feel like I would seem pretentious to my relatives and acquaintances if my captions were in English."*

(15) *"-- most of my followers are Finnish people that I know personally which is why it would be weird to always use English."*

Similarly to the factors behind choosing English, certain Finnish expressions were highlighted in some responses. Emphasis on Finnish found as a fun, descriptive language can be seen in the data, as the excerpts below show. For example, one respondent noted that they intentionally choose Finnish since they find English has become a more dominant language in social media: *"--I quite actively try to avoid using too much English, since the predominance of English in social media annoys me a little--"*. Additionally, as excerpt 16 illustrates, it was

stated that Finnish people should improve their Finnish skills when it comes to, for example, replacing English words with Finnish equivalents. These findings support Auer's (2007: 1-2) view on language seen as unique and therefore promoting togetherness.

(16) *"--I also think that Finnish people should improve their use of Finnish in all kinds of situations, for example learning how to replace English words with Finnish ones. That would mean inventing new words/phrases or using already existing words more."*

(17) *"I might use Finnish if I want to say something that won't convey the same meaning in English"*

(18) *"Finnisj [Finnish] for me is a "fun language" that can express a lot more than English. Some things just cannot be translated."*

The possible Finnish context also affects people's language choice, according to the collected data. For example, if pictures include Finnish holidays, landscapes or habits, they call for Finnish captions. Moreover, family and certain occasions were found more natural to post about in Finnish. This can be explained for example by social constructionism, which sees identity as constantly changing in different contexts (Georgalou 2017: 10). Therefore, individuals modify their means of online communication based on their contextually changing identities. In this case, it can be seen that as Auer (2007: 1-2) claims, individuals' language identities are closely bound to their Finnish identity which is further portrayed by expressing themselves through the Finnish language, as the excerpts below illustrate.

(19) *"-- Some things are linked to our language and culture, for example when we are celebrating first of may or participating to student events--"*

(20) *"-- I use it when I'm describing some events or habits that are normal in Finland. E.g. I write "Hyvää Juhannusta!" [happy midsummer] in Finnish because it is not an international holiday."*

(21) *"-- if I'm posting about family--"*

(22) *"-- Finnish instead of English in posts that are more 'official' e.g. graduation post because it feels more natural--"*

Lastly, many respondents rationalized that they tend to use Finnish when the captions are longer or cover a sensitive topic. Story-like captions about one's life are found to feel more natural in Finnish according to the respondents. What is more, respondents mentioned that it is easier to talk about serious subjects and express certain emotions and feelings in Finnish. Considering Finnish being at least one of all the respondents' native languages, this phenomenon could be traced to the respondents' language identities, e.g. their Finnishness, as well as storytelling



being one of the primary characteristics of human language (Auer 2007: 1-2; Popper, cited in Edwards 2009: 54). Moreover, individuals tend to express themselves through a collective language which further strengthens the feeling of togetherness. In conclusion, in an environment where one encounters Finnish and Finnish people, such as Instagram, social cohesion can be sought for example by using Finnish since it is a shared, unique way of expression for a certain group of people. (Auer 2007: 1-2)

(23) *"If the caption of a picture is a bit longer and it's something like a story related to my life, I will use Finnish."*

(24) *"Sometimes Finnish is the best way to express certain emotions, especially in my dialect.--"*

(25) *"-- I somehow feel like it [Finnish] makes the caption feel more authentic and sincere--"*

When it comes to additive and subtractive bilingualism, in the light of the responses that were gathered in this study, it could be suggested that both apply simultaneously. It is simply due to a certain situation whether either of them is highlighted. For example, as became evident in many excerpts, English is considered simply better suitable for some situations and Finnish for others. One respondent stated: *"Usually the language choice of my caption depends on many things; my mood on that particular day of "posting", the context of the picture posted and how the certain language functions with the aforementioned things."* Therefore, it is challenging, if even impossible, to define the situations where either of these languages would be better suitable since the language choice fully depends on the individual's personal, situation-bound opinions and thoughts at that exact moment.

### 5.3 IDENTITY IN INSTAGRAM CAPTIONS

Beside language choice in Instagram captions, the visibility of identity in the captions is another subject of interest in the present study. This aspect was investigated by four (4) Likert scale questions regarding expression in English and Finnish, consideration when it comes to linguistic choices in Instagram captions and visibility of language identity in Instagram captions. As can be seen in Table 5 below, the results show that the vast majority of the respondents state that their language identity is visible in their Instagram captions.

Table 5. Identity and self-expression in Instagram captions (%)

	My language identity is visible	I express myself best in my Instagram	I express myself best in my Instagram	I consider my linguistic choices when it comes to

	in my Instagram caption.	captions when using English.	captions when using Finnish.	my Instagram captions.
1=I disagree	1,9%	7,7%	0,0%	1,9%
2=I somewhat disagree	9,6%	17,3%	13,5%	0,0%
3=Neutral/I don't know	15,4%	23,1%	19,2%	1,9%
4=I somewhat agree	46,2%	26,9%	36,5%	26,9%
5=I agree	26,9%	25,0%	30,8%	69,2%

Table 6. Visibility of language identity: comparison between genders (n)

Gender	1=I disagree	2=I somewhat disagree	3=Neutral/I don't know	4=I somewhat agree	5=I agree
Male	0	2	1	0	2
Female	1	3	7	21	12
Other	0	0	0	3	0

As much as over 70% either somewhat agree or agree that their language identities are visible when it comes to their Instagram captions (Table 5). However, when comparing the results between male and female respondents (Table 6), it is revealed that the male responses are distributed on both ends of the Likert scale, whereas the vast majority of the female responses is situated on the positive side of the scale. This could partially be explained by women's greater use of Instagram in general (DNA 2021). What is more, considering all the respondents reporting Finnish as at least one of their native languages, it is no surprise that nearly 70% state expressing themselves the best in Finnish in their Instagram captions. However, the percentage on the same issue when it comes to expression in English is surprisingly high, being over 50%. As the respondents' English language proficiency is mostly on level C as Table 1 illustrates, it can be concluded that the better English skills one possesses, the easier they find expression in English.

As mentioned earlier, individuals build and portray their identity by language use, thus language and identity are closely tied together (Hassan 2010, cited in Nurhantoro & Wulandari 2017: 38). Therefore, as identities are easily expressed through text (Georgalou 2017: 13), which language to use can be seen as a reflection of one's identity and help in communicating

beyond the literal meaning of words (Gardner-Chloros 2009: 4). That is, individuals communicate different things in English and in Finnish, as became evident in sections 5.2.1 and 5.2.2, and similarly these languages represent certain aspects of their identities (Edwards 2009: 30). Based on the results, individuals use English and Finnish in different contexts when it comes to Instagram captions. Therefore, it can also be concluded that identity representation is different in English and in Finnish since individuals use these languages to communicate different things. Thus, identities become visible in multiple different ways when using English versus when using Finnish. As the excerpts below show, these ways are challenging to pinpoint since every caption and every language choice is dependent on a certain situation, feeling, need or goal, among other things. Moreover, these excerpts also illustrate Edward's (2009: 19) statement on identity meaning that a person is being true to oneself and not trying to pretend something else.

(26) *--using Finnish feels too personal sometimes--*

(27) *--Finnish captions don't always sound like my own voice--*

(28) *--it [Finnish] just feels more right.*

(29) *--easier to express my identity and personality [in Finnish]--*

However, what is interesting when it comes to the results is that even though it is challenging to define specific ways in which identity is visible in Instagram captions based on the responses, the vast majority of the respondents still report their language identities being visible in their Instagram captions as well as considering the language choices they make for these captions (Table 5). Additionally, on the contrary to Nurhantoro and Wulandari's (2017) findings on one's identity on social media being not as visible, these results indicate the opposite. Therefore, it could be inferred that even individuals themselves are not necessarily aware of the specific ways in which they express and build their identities in their Instagram captions even though they consider the language choices they make. However, the subject of interest here was specifically identity in relation to language in Instagram captions rather than the whole multifaceted concept of identity online. Additionally, the results would likely be very different if for example the respondents' average English language proficiency was lower. Lastly, identity being an extremely complex concept, understanding the connection between it and language on individuals' Instagram captions would require more profound research.

## 6 Conclusion

The present study investigated language choice between Finnish and English on Instagram captions, attaching identity to the phenomenon as well. These two phenomena were investigated by mixed methods which produced both qualitative and quantitative data. The results uncovered multiple factors behind language choice when it comes to individuals' Instagram captions, as well as revealed how individuals perceive their identity in relation to language being visible in those captions.

The most evident results when it comes to the present study are that there are multiple reasons behind choosing either English or Finnish for one's Instagram caption, and that these reasons are quite similar when it comes to both languages. Moreover, language choices were seen as more situation-bound rather than straightforward, and there can be multiple situational factors behind a single caption. In other words, the decisions simply depend on an individual's wants, needs, opinions and preferences at the moment of posting, as was stated in many responses. Therefore, it is extremely challenging to define a specific pattern behind language choice since every time it is based on individual, situation-bound factors and reasoning. Similarly, it can be summarized that it is a challenge to specifically define the ways in which individuals' identities become visible in their Instagram captions due to the constantly changing and context-bound nature of identity. However, it can be said that individuals build and portray their identities through how they choose to use language and communicate to others. Moreover, the results of the present study indicate that even though individuals might do these choices subconsciously, they still know their identities being visible.

In general, language choice and identity representation in Instagram captions should be investigated more in future studies since it is nowadays a major way of expressing oneself linguistically. Moreover, as English can be seen as a dominating language in social media, investigating the choices behind choosing it can reveal interesting viewpoints and more knowledge on language in social media in general. In addition, investigating language choice in relation to the languages individuals consume in their social media could unveil more of the processes behind language choice. Instagram being one of the most used social media platforms (cyberclick.net), this study only scratches the surface when it comes to language choice in Instagram captions. However, as these issues have not been thoroughly researched before, the

present study offers new insights and a primary standpoint on language choice between English and Finnish on Instagram captions, which again was the purpose of this study.

## Bibliography

Auer, P. (2007). *Style and social identities: Alternative approaches to linguistic heterogeneity*. Mouton de Gruyter.

Barman, U., Das, A., Wagner, J. & Foster, J. (2014). *Code mixing: A Challenge for Language Identification in the Language of Social Media*. Dublin City University, Dublin.

Biber, D. & Conrad, S. (2009). *Register, genre, and style* [online]. Cambridge University Press.

Calkin, M. M. (2015). *Making Pretty: Examining Contemporary Identity Construction through Instagram*. Communication Studies. San Francisco State University.

*Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. (2020). Companion volume. Council of Europe. 36, 177. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>

Council of Europe. <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions> (13 December, 2021)

Cunliffe, D., Morris, D. & Prys, C. (2013). Young Bilinguals' Language Behaviour in Social Networking Sites: The Use of Welsh on Facebook. *Journal of Computer-Mediated Communication* [online] 18 (3), 339–361. doi: <https://doi.org/10.1111/jcc4.12010>

Cyberclick.net: Top 10 Most Downloaded Apps of 2021 So Far. <https://www.cyberclick.net/numericalblogen/top-10-most-downloaded-apps-of-2020-so-far> (7 December, 2021)

Denscombe, M. (2014). *The good research guide: For small-scale social research projects* [online]. Open University Press.

DNA Digitaaliset elämäntavat -tutkimus. (2021). <https://www.stinfo.fi/data/attachments/00200/838ead53-d63a-4f2a-9d3e-db3845973aec.pdf>. (21 March, 2022).

Dovchin, S. (2016). The role of English in the language practices of Mongolian Facebook users: English meets Mongolian on social media. *English today* [online] 33 (2), 16-24.

doi: <https://doi.org/10.1017/S0266078416000420>

Edwards, J. Foundations of Bilingualism. In Bhatia, T. K. & Ritchie, W. C. (2004). *The handbook of bilingualism*. Blackwell Publishing. 7-31.

Edwards, J. (2009). *Language and identity: An introduction*. Cambridge University Press.

Gardner-Chloros, P. (2009). *Code-switching*. Cambridge University Press.

Georgalou, M. (2017). *Discourse and identity on Facebook: How we use language and multimodal texts to present identity online*. Bloomsbury Academic.

Gibson, W. J. & Brown, A. (2009). *Working with qualitative data*. SAGE.

Hinrichs, L. Modular repertoires in English-using social networks: A study of language choice in the networks of adult Facebook users. In Squires, L. (2016). *English in Computer-Mediated Communication*. De Gruyter Mouton, 17-42.

Inha, K. (n.d.). *Introducing an earlier start in language teaching: Language learning to start as early as in kindergarten*. Finnish National Agency for Education. oph.fi.

[https://www.oph.fi/sites/default/files/documents/introducing\\_an\\_earlier\\_start\\_in\\_language\\_teaching.pdf](https://www.oph.fi/sites/default/files/documents/introducing_an_earlier_start_in_language_teaching.pdf). (4 April, 2022)

Instagram: About. <https://about.instagram.com>. (6 November, 2021)

University of Jyväskylä. Instructions for students. <https://www.jyu.fi/en/university/data-privacy/tietosuojahjeet/students>. (30 March 2021)

Lauring, J. (2008). Rethinking Social Identity Theory in International Encounters: Language Use as a Negotiated Object for Identity Making. *International Journal of Cross Cultural Management: CCM* [online] 8 (3), 343-361. doi: <https://doi.org/10.1177/1470595808096673> (11 November 2021)

Lee, C. Language choice and self-presentation in social media: the case of university students in Hong Kong. In Seargeant, P. & Tagg, C. (2014). *The language of social media: Identity and community on the internet*. London: Palgrave Macmillan, 91-111.

Marshall, G. (2005). *The purpose, design and administration of a questionnaire for data collection*. Radiography 11 (2), 131-136.

Meta.com: About. <https://about.facebook.com/meta/> (7 December 2021)

Nummenmaa, L., Holopainen, M. & Pulkkinen, P. (2019). *Tilastollisten menetelmien perusteet*. Sanoma Pro Oy.

Nurhantoro, T. S. & Wulandari, N. (2017). *Exploring the use of English in Instagram and its influence on the user's identity*. Journal of English Language and Education [online] 3 (2), 37.

Quintly.com: Instagram study 2019. <https://info.quintly.com/instagram-study-2019/view#length> (28 March 2022)

Squires, L. (2016). *English in Computer-Mediated Communication*. De Gruyter Mouton.

Tuomi, J. & Sarajärvi, A. (2009). *Laadullinen tutkimus ja sisällönanalyysi* (6. uud. laitos.). Tammi.

Sultana, S., Dovchin, S. & Pennycook, A. (2013). *Styling the periphery: Linguistic and cultural takeup in Bangladesh and Mongolia*. Journal of Sociolinguistics 17 (5), 687-710.

doi: <https://doi.org/10.1111/josl.12055>

Thomas, S. (2014). *Basic statistics* [online]. Alpha Science International Ltd.

Valli, R. (2017). *Creating a questionnaire for a scientific study*. International Journal of Research Studies in Education 6 (4), 15-27. doi: <https://doi.org/10.5861/ijrse.2016.1584>

Varto, J. (1992). *Laadullisen tutkimuksen metodologia*. Kirjayhtymä.

Vogt, W. P., Vogt, E. R., Gardner, D. C. & Haeffele, L. M. (2014). *Selecting the right analyses for your data: Quantitative, qualitative, and mixed methods*. The Guilford Press.



Zhao, S., Grasmuck, S. & Martin, J. (2008). *Identity construction on Facebook: Digital empowerment in anchored relationships*. *Computers in human behavior*, 24(5), 1816-1836. doi: <https://doi.org/10.1016/j.chb.2008.02.012> (30 March, 2022)

Quer, J. & Steinbach, M. (2019). *Handling Sign Language Data: The Impact of Modality*. *Frontiers*. doi: <https://doi.org/10.3389/fpsyg.2019.00483> (1 April, 2022)

## Appendix

### THE QUESTIONNAIRE

(\* required an answer)

**1. I have read the instructions, I ensure I am at least 18 years old and give permission to use my answers as data in the Bachelor's thesis mentioned above. \***

Yes

**2. Age \***

18-24 years old

25-30 years old

31-40 years old

41-50 years old

51+ years old

**3. Gender \***

Male

Female

Other

Prefer not to say

**4. Native language(s) \***

---

---

---

---

---

**5. According to the descriptions of CEFR levels below, what is your English language proficiency? \***

Picture from CEFR2020 page 177 (<https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>)

Interaction	A1	A2	B1	B2	C1	C2
<b>Oral interaction</b>	<p>I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate and help me formulate what I am trying to express.</p> <p>I can ask and answer simple questions in areas of immediate need or on very familiar topics.</p>	<p>I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities.</p> <p>I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.</p>	<p>I can deal with most situations likely to arise while travelling in an area where the language is spoken.</p> <p>I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).</p>	<p>I can interact with a degree of fluency and spontaneity that makes regular interaction with users of the target language quite possible.</p> <p>I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.</p>	<p>I can express myself fluently and spontaneously without much obvious searching for expressions.</p> <p>I can use language flexibly and effectively for social and professional purposes.</p> <p>I can formulate ideas and opinions with precision and relate my contribution skilfully to those of others.</p>	<p>I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms.</p> <p>I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.</p>
<b>Written and online interaction</b>	<p>I can post short, simple greetings as statements about what I did and how I liked it, and can respond to comments in a very simple way.</p> <p>I can react simply to other posts, images and media.</p> <p>I can complete a very simple purchase, filling in forms with personal details.</p>	<p>I can engage in basic social interaction, expressing how I feel, what I am doing or what I need, and responding to comments with thanks, apology or answers to questions.</p> <p>I can complete simple transactions such as ordering goods, can follow simple instructions and can collaborate in a shared task with a supportive interlocutor.</p>	<p>I can interact about experiences, events, impressions and feelings, provided I can prepare beforehand.</p> <p>I can ask for or give simple clarifications and can respond to comments and questions in some detail.</p> <p>I can interact with a group working on a project, provided there are visual aids such as images, statistics and graphs to clarify more complex concepts.</p>	<p>I can interact with several people, linking my contributions to theirs and handling misunderstandings or disagreements, provided the others avoid complex language, allow me time and are generally co-operative.</p> <p>I can highlight the significance of facts, events and experiences, justify ideas and support collaboration.</p>	<p>I can understand the intentions and implications of other contributions on complex, abstract issues and can express myself with clarity and precision, adapting my language and register flexibly and effectively.</p> <p>I can deal effectively with communication problems and cultural issues that arise by clarifying and exemplifying.</p>	<p>I can express myself in an appropriate tone and style in virtually any type of interaction.</p> <p>I can anticipate and deal effectively with possible misunderstandings, communication issues and emotional reactions, adjusting language and tone flexibly and sensitively as appropriate.</p>

 A1 A2 B1 B2 C1 C2

**6. According to the descriptions of CEFR levels below, what is your Finnish language proficiency? \***

Interaction	A1	A2	B1	B2	C1	C2
<b>Oral interaction</b>	<p>I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate and help me formulate what I am trying to express.</p> <p>I can ask and answer simple questions in areas of immediate need or on very familiar topics.</p>	<p>I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities.</p> <p>I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.</p>	<p>I can deal with most situations likely to arise while travelling in an area where the language is spoken.</p> <p>I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).</p>	<p>I can interact with a degree of fluency and spontaneity that makes regular interaction with users of the target language quite possible.</p> <p>I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.</p>	<p>I can express myself fluently and spontaneously without much obvious searching for expressions.</p> <p>I can use language flexibly and effectively for social and professional purposes.</p> <p>I can formulate ideas and opinions with precision and relate my contribution skilfully to those of others.</p>	<p>I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms.</p> <p>I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.</p>
<b>Written and online interaction</b>	<p>I can post short, simple greetings as statements about what I did and how I liked it, and can respond to comments in a very simple way.</p> <p>I can react simply to other posts, images and media.</p> <p>I can complete a very simple purchase, filling in forms with personal details.</p>	<p>I can engage in basic social interaction, expressing how I feel, what I am doing or what I need, and responding to comments with thanks, apology or answers to questions.</p> <p>I can complete simple transactions such as ordering goods, can follow simple instructions and can collaborate in a shared task with a supportive interlocutor.</p>	<p>I can interact about experiences, events, impressions and feelings, provided I can prepare beforehand.</p> <p>I can ask for or give simple clarifications and can respond to comments and questions in some detail.</p> <p>I can interact with a group working on a project, provided there are visual aids such as images, statistics and graphs to clarify more complex concepts.</p>	<p>I can interact with several people, linking my contributions to theirs and handling misunderstandings or disagreements, provided the others avoid complex language, allow me time and are generally co-operative.</p> <p>I can highlight the significance of facts, events and experiences, justify ideas and support collaboration.</p>	<p>I can understand the intentions and implications of other contributions on complex, abstract issues and can express myself with clarity and precision, adapting my language and register flexibly and effectively.</p> <p>I can deal effectively with communication problems and cultural issues that arise by clarifying and exemplifying.</p>	<p>I can express myself in an appropriate tone and style in virtually any type of interaction.</p> <p>I can anticipate and deal effectively with possible misunderstandings, communication issues and emotional reactions, adjusting language and tone flexibly and sensitively as appropriate.</p>

 A1

 A2

 B1

 B2

 C1

 C2

### 7. Language choices in Instagram captions

Please evaluate your English and Finnish language use in your personal Instagram captions from a scale of 1 to 5.

	1=Never	2=Rarely	3=Sometimes	4=Often	5=Always
I use English in my Instagram captions *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use Finnish in my Instagram captions *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I code-switch between English and Finnish in my Instagram captions (code-switching refers to the use of several languages or dialects in one sentence) *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 8. Why do you use English in your Instagram captions instead of Finnish? \*

---



---



---



---



---

### 9. Why do you use Finnish in your Instagram captions instead of English? \*

---



---



---



---



---

### 10. Language identity in Instagram captions \*

Evaluate your language identity in your personal Instagram captions from a scale of 1 to 5.

	1=I disagree	2=I somewhat disagree	3=Neutral/I don't know	4=I somewhat agree	5=I agree
I express myself best in my Instagram captions when using English *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	1=I disagree	2=I somewhat disagree	3=Neutral/I don't know	4=I somewhat agree	5=I agree
I express myself best in my Instagram captions when using Finnish *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I consider my linguistic choices when it comes to my Instagram captions *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My language identity is visible in my Instagram captions *	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**11. How often do you post on your personal Instagram feed? \***

- Once a day or more
- Weekly
- Monthly
- Few times a year
- Once a year or less
- I don't post on my Instagram feed

**12. Why do you use Instagram? Elaborate if you want. \***

- To keep in touch with family, friends and acquaintances \_\_\_\_\_
- To follow celebrities, influencers etc \_\_\_\_\_
- To get inspiration \_\_\_\_\_
- For work or school purposes \_\_\_\_\_
- To share content \_\_\_\_\_
- For some other reason, what? \_\_\_\_\_

**13. If you still have something you would like to add considering the topic of**

**this questionnaire or your answers, please feel free to do so here.**

---

---

---

---

---