Heli Leena Marjukka Teitto

HUMAN REFERENTS IN SUBTITLES
A Study on Personal Pronouns and Proper Nouns in Translated and Original Finnish

UNIVERSITY OF EASTERN FINLAND
Philosophical Faculty
Foreign Languages and Translation Studies
Pro Gradu Thesis
November 2010
The aim of this study is to examine the use of human referents in subtitles. The main emphasis is on two types of human reference: personal pronouns and proper nouns. Their use in subtitles is compared with other texts translated into Finnish (translated Finnish) and texts originally written in Finnish (original Finnish). The special characteristics of subtitles give reason to expect that the use and frequency of personal pronouns and proper nouns is different in subtitled Finnish when compared with translated or original Finnish. Translated and subtitled Finnish are considered separate in this study. The main theory of this study centres on the earlier studies on personal pronouns in translations.

The hypothesis is that subtitles use fewer personal pronouns because of the limited space and time. In Finnish, grammatical person can be expressed in various ways. In standard Finnish personal pronouns are often omitted because the verb form is enough to indicate person. Overuse of personal pronouns is considered as feature belonging to spoken Finnish. Directly addressing people by name is less common in Finnish than in English, for example, and therefore it can be expected that original Finnish contains less proper nouns than translated Finnish. Subtitled Finnish may use proper nouns more to avoid the ambiguous third person pronouns, because subtitles must be clear to understand.

The material used in this study is compiled from two corpora: the Corpus of Translated Finnish and The Finnish Broadcasting Company corpus of subtitles. From the latter, subtitles of two episodes of Gilmore Girls, an American television series, were chosen. From the former were selected altogether eight text excerpts, four of which were translated from English to Finnish and four original Finnish. These were further divided into dialogue and narrative, because the subtitles examined in this study consist of dialogue. The material was analysed with the help of WordSmith Tools, which is a programme for corpus analysis. The language of the material adhered to the norms of standard Finnish with a few deviations. Colloquialisms occurred mainly in the vocabulary.

The analysis shows that personal pronouns occur more frequently in translated and subtitled Finnish than in original Finnish. Dialogue in general contains more personal pronouns than narrative. When examining dialogue only, personal pronouns are the most common in original Finnish and least common in subtitled Finnish, which supports the original hypothesis. Dialogue consists mostly of first and second person pronouns whereas in narrative third person pronouns are clearly more common. In subtitles the difference is not as great, which indicates that first and second person pronouns are omitted if possible. Comparing the results of this study to those of earlier ones shows that personal pronouns occur more frequently in the material of this study.

The frequency of proper nouns does not vary as much in the material as the personal pronouns. Examining the undivided material shows that proper nouns are most frequent in original Finnish and least frequent in subtitles. In dialogue, proper nouns are the most common in subtitles, which supports the original hypothesis. Examining the use of proper nouns as vocatives shows that such use is clearly more common in translated Finnish than in original or subtitled Finnish.

The most common case forms of personal pronouns and proper nouns, the nominative and genitive case, are also examined. The results indicate that there are some differences in their use in the material, for example personal pronouns were used as genitive modifiers more often in translated Finnish than in original or subtitled Finnish.

The results of this study are not definitive, but they indicate that the need for human referents in texts varies little although more research is necessary. Personal pronouns are used less in subtitles than in other translations or original Finnish, and proper nouns are used more. Genre seems to affect especially the use of personal pronouns. There are also differences in how personal pronouns and proper nouns are used, for example the use of personal pronouns in subtitles is not related to creating a feeling of colloquial language use as in original Finnish.
TABLE OF CONTENTS

1 INTRODUCTION ..................................................................................................... 1

2 SPECIAL CHARACTERISTICS OF SUBTITLES.................................................... 4
   2.1 Introduction......................................................................................................... 4
   2.2 Subtitles and dialogue ......................................................................................... 5
   2.3 Limitations .......................................................................................................... 7
   2.4 Language ............................................................................................................ 9

3 PERSONAL PRONOUNS ....................................................................................... 10
   3.1 Expressing grammatical person ......................................................................... 10
   3.2 Earlier studies .................................................................................................... 13
      3.2.1 Introduction ................................................................................................ 13
      3.2.2 The SG1 pronoun ....................................................................................... 14
      3.2.3 The SG2 pronoun ....................................................................................... 17

4 PROPER NOUNS .................................................................................................... 19

5 MATERIAL AND METHOD .................................................................................. 22
   5.1 Material ............................................................................................................. 22
   5.2 Method .............................................................................................................. 24

6 ANALYSIS ............................................................................................................. 26
   6.1 Human Reference .............................................................................................. 26
      6.1.1 Undivided material ..................................................................................... 26
      6.1.2 Dialogue and narrative ................................................................................ 28
      6.1.3 A note on the use of non-standard language ................................................ 31
   6.2 Personal pronouns ............................................................................................. 33
      6.2.1 Undivided material ..................................................................................... 33
      6.2.2 Dialogue ..................................................................................................... 34
      6.2.3 Narrative .................................................................................................... 36
      6.2.4 Distribution of grammatical person ............................................................. 36
      6.2.5 Comparison with earlier studies .................................................................. 41
      6.2.6 Use and structures ...................................................................................... 45
      6.2.7 Corpora-specific use ................................................................................... 50
   6.3 Proper nouns ..................................................................................................... 52
      6.3.1 Undivided material ..................................................................................... 52
      6.3.2 Dialogue and narrative ................................................................................ 53
      6.3.3 Vocative expressions .................................................................................. 54
      6.3.4 Grammatical case ....................................................................................... 57
      6.3.5 Corpus-specific structures ........................................................................... 59

7 CONCLUSION........................................................................................................ 62

RESEARCH MATERIAL .......................................................................................... 69

WORKS CITED ......................................................................................................... 69

APPENDIX

FINNISH ABSTRACT
1 INTRODUCTION

Human reference in translations is a subject that has interested many researchers. The studies have, however, mainly concentrated on prose. One popular area of research has been centred on the problem of translating the English third person singular “he” or “she” into Finnish *hän* (see e.g. Auvinen 2005 or Rivinoja 2004). Because Finnish lacks gender-marked pronouns, the third person singular may cause problems when translating into Finnish, if it has been employed as a stylistic convention (Rivinoja 2004: 77). Studies on proper nouns have concentrated on their translation (see e.g. Kauhanen 2004)

Human reference is especially vital in dialogue, where it is used to establish the relationship between the speaker and addressee and their relationship to the person or persons referred to in the conversation. However, hardly any research has been done on the use of human reference in subtitles, a type of audiovisual translation. Jääskeläinen (2007: 116-128) has discussed the need for more studies on audiovisual translations. She points out that, although some research has been done on the subject, there are still several basic questions that require answering. One such question is how the illusion of colloquial language is created in subtitles (Jääskeläinen 2007: 124).
Subtitles are a relatively new field of research. In small countries, such as Finland, most foreign television programmes are subtitled, and subtitles form a major part of the people’s daily reading (Vertanen 2007: 149; Jääskeläinen 2007: 116-117), which makes them a field of study in need of more attention. Subtitles can be seen as having similarities with popular fiction, for example, because both tend to receive fairly little attention from researchers. Mauranen questions this trend of ignoring the so-called popular genres and argues that

Yet the question of translation in popular genres is interesting; popular texts are read much more widely than academic texts, and thereby can be expected to exert more influence on the language as a whole. Nevertheless, they have not received as much attention as more prestigious genres within translation studies, or in corpus linguistics. (Mauranen 2002)

Current technology has made examining subtitles much simpler than it was, and the compilation of a corpus from the subtitles of The Finnish Broadcasting Company (YLE) has offered researchers an access to a wide range of different subtitles.

The main inspiration for this study are the results Mauranen and Tiittula (2005) gained by examining the use of the SG1 pronoun minä (‘I’) in texts. Their results are presented in section 3.2.2. Other earlier studies on the use of personal pronouns have similarly concentrated on only one or two personal pronouns. This study aims to examine their use as a whole. Proper nouns have been included in this study to offer a point of comparison for personal pronouns. The use of proper nouns and personal pronouns in subtitles will be examined and compared with texts that have been translated into Finnish and with texts that were originally written in Finnish with no translating involved. These will be referred to as translated Finnish (TF) and original Finnish (OF) henceforth. The material of this study is compiled from The Finnish Broadcasting Company corpus of subtitles and the Corpus of Translated Finnish.
According to Jääskeläinen (2007: 127), the special characteristics of subtitles may give a reason to assume that the language used in subtitles differs from original Finnish as well as translated Finnish. The language of subtitles will subsequently be referred to as subtitled Finnish (SF). Jääskeläinen’s suggestion forms another point of interest in this study: how the use of proper nouns and personal pronouns in subtitles differ, if at all, from translated or original literature and especially from the dialogue in them.

The hypothesis of this study is that the condensation causes variation in the way persons are referred to in subtitles when compared with other kinds of translations and texts originally written in Finnish. Subtitled Finnish and translated Finnish are treated as separate types in this study, despite the fact that the examined subtitles are, in fact, translations. It is likely that personal pronouns are used less in subtitled Finnish than in translated or original Finnish in order to save space, and proper nouns may be favoured instead of the ambiguous third person pronouns. Other possible differences lie in the structures in which the personal pronouns and proper nouns occur. Both subtitled Finnish and translated Finnish may favour structures uncommon to original Finnish.

Section 2 presents the special characteristics of subtitles, including the relationship between subtitles and other written dialogue. In section 3, personal pronouns and their use in Finnish is discussed, and results of earlier studies examining the use of personal pronouns in literature will be presented. Section 4 covers briefly the basic theory behind proper nouns. The material and method used in this study are introduced in Section 5, and the analysis in Section 6. The conclusion of this study is discussed in section 7. The abbreviations used in this study are explained in the appendix.
2 SPECIAL CHARACTERISTICS OF SUBTITLES

2.1 Introduction

Subtitles belong to the genre of audiovisual translation, which differs in various aspects from other types of translations, such as prose. According to Immonen (2005: 167), audiovisual translation has developed its own unique language. In audiovisual translation the importance of the target language and culture is emphasized more than in other types of translation.

Especially subtitles have their own special features that set them apart from other genres. In subtitles, the emphasis is on the relationship between the subtitle lines and the image on the screen, because in a situation where several characters speak, the only way to identify the speaker is the correct timing of the subtitle line (Vertanen 2007: 150-151).

This relationship with the image on the screen also defines how much clarification the viewers need in order to follow the events, which then in turn affects how the characters are referred to, among other things. For example, there is no need to address the participants of a conversation by their name in subtitles, if it is clear who is speaking to
whom (Vertanen 2007: 152). The use of proper nouns is discussed more in depth in section 4.

2.2 Subtitles and dialogue

In subtitles, the speech of the characters on screen is condensed into text, which is shown at the bottom of the screen usually in one or two lines (Immonen 2005: 167; Vertanen 2007: 151-152). Even though subtitles are speech turned into text, it is worth noting that the dialogue in television dramas is rarely natural, but rather based on a script. Due to this artificial nature and the limitations of subtitling, the relation of subtitles to other written dialogue in literature is of special interest in this study. The relationship between written dialogue and spoken language is complex. According to the Oxford English Grammar,

Midway between the spoken and written language are written representations of speech. These appear in plays, which are read privately or are spoken by actors, and in the dialogue to be found in novels and stories. The representations attempt to imitate, to a greater or lesser extent, characteristics of normal conversation. (Greenbaum 1996: 368)

Subtitles, too, are written representations of speech existing between the spoken and written language: in general, subtitles are written in standard Finnish, but occasionally they imitate the phenomena occurring in spoken language to some extent (Vertanen 2007: 153).

How commonly non-standard features occur in texts varies between translated and original Finnish. Juva (1998: 50) has noted that, although colloquial expressions are
becoming more frequent in texts written in original Finnish, the dialogue in translations still uses standard Finnish. Juva (1998: 53) points out that creating an illusion of colloquial language is more important than recreating all features of spoken language in a text.

The various features related to non-standard language use has been examined by Nevalainen (2003). His study concentrated on the features used in written Finnish to create an illusion of colloquial language. The aim of his study was to see whether there were any fundamental differences between the language use in translated and original Finnish. According to Nevalainen (2003), a general assumption is that translations tend to be more conservative, in which case they contain less colloquial expressions than a corresponding text written originally in Finnish.

The results of Nevalainen’s (2003) study show that in contemporary fiction both standard language and colloquial features are used, and although standard language is in the majority, the colloquial features are used to establish a character’s personality, for example. According to him, the common ways to create a feeling of colloquial language are by using non-standard pronunciation and vocabulary, avoiding long and abstract words, and by adding pronouns, among other things (Nevalainen 2003: 4).

Nevalainen’s (2003) results suggest that, as Juva (1998) has also noted, translators are more wary of using colloquialisms than authors writing in Finnish, which is why translated fiction is more conservative. Nevalainen does, however, emphasize that the corpora used in his study are not parallel and therefore not necessarily well-suited for examining non-standard language use.
An important observation in Nevalainen’s (2003) study is that the source text does not necessarily contain enough stimuli (e.g. features that clearly classify the speech as representing a dialect) for translators to use colloquialisms. The same may apply to subtitles: there may not also be enough stimuli in the speech of the characters for the translator to add dialectal features in the subtitles. Possible deviations from this would be television programmes in which the characters speak with very strong accents or a specific dialect, or programmes for young viewers in which the characters are made to use mā and sā (the dialectal variants of ‘I’ and ‘you’) in subtitles, perhaps in order to appeal more to the perceived main audience. As discussed more in detail in the section 2.4, subtitles should be understood on the first reading, which lessens the use of colloquialisms in subtitles.

2.3 Limitations

Translators usually have alternative ways for translating a certain part of a text, although the common expectation is that they should remain as loyal as possible to the source text. In subtitles no such freedom of choice exists, because there are only a certain number of characters that fit on the screen, and only a certain amount of time is allotted for each line.

The character limit for subtitles varies on different channels. On the channels of The Finnish Broadcasting Company YLE, the limit was thirty characters per line until autumn 2006, meaning that there were altogether sixty characters available in a two-line subtitle (Vertanen 2007: 151). This limit still applies to the subtitles analysed in this
study. The current average character limit on the channels of YLE is thirty-three. Because the space available for the translation is limited, not everything can be translated. The parts of dialogue vital for the plot are given the first priority in the subtitles (Vertanen 2007: 152.)

In addition to the character limits, the time for each line shown on the screen is also restricted. The synchronization between the subtitles and the image on the screen is vital for the viewer’s understanding (Vertanen 2007: 152). As mentioned earlier, subtitles, like all audiovisual translation, are connected to the image and sound: what is shown on the screen, and even the sound effects and music are important in constructing a meaning (Immonen 2005: 165). Vertanen (2007: 150-151) points out that the subtitles are not complete by themselves, meaning that they can only be understood in conjunction with the image on the screen.

How long a subtitle line is shown on the screen is also limited by how fast an average viewer can read the subtitles. According to Vertanen (2007: 151-152), an average Finnish television viewer needs four to five seconds to read a full-length two-line subtitle and two to three seconds for reading a one-liner. If the lines change too fast, the viewer will not have enough time to read them, but a subtitle line remaining on the screen too long will also have a disruptive influence on the viewer’s enjoyment of the programme.
2.4 Language

The purpose of this study is to examine the language used in subtitles by examining personal pronouns and proper nouns. As mentioned in the introduction, one area of interest is how the feeling of colloquial language is created in subtitles. In general, subtitles are written in standard Finnish; they follow the conventions and established ways of writing as well as the norms, structures and rules of the language (Immonen 2005: 166). YLE TV1, the channel on which the subtitles examined in this study were broadcast, is seen as a quality channel, and its subtitles generally follow the standard language use (Vertanen 2007).

Subtitlers do not always follow the recommendation of using standard Finnish. Immonen (2005), for example, points out that the subtitles should convey an impression of the way a character speaks, if the source language represents a dialect or a subculture. Vertanen (2007: 153) emphasises the fact that each line in subtitles should be as clear and easy to read as possible. Using slang or colloquial expressions may draw the viewers’ attention away from what is being said, or even hinder their understanding. However, Vertanen (2007) also considers it important for subtitles to convey the way a character speaks, and to achieve that some colloquial expressions or slang may be used. Nevertheless, the message should be relayed on the first reading, because usually there is not enough time to read the line a second time (Vertanen 2007: 153).
3 PERSONAL PRONOUNS

3.1 Expressing grammatical person

Expressing the grammatical person is an essential aspect of any language, according to Mauranen and Tiittula (2005). The grammatical person is especially vital in establishing and maintaining the relationships between the speaker and the addressee, but can sometimes be a significant challenge for a translator (Mauranen and Tiittula 2005: 35).

In Finnish there is a wide variety of ways to express the grammatical person: personal pronouns and their dialectal variants, as well as suffixes, such as the possessive suffix. There are three possibilities for person: first, second and third person. First person is used by the speaker to refer to herself, second person is used to refer to the individual or individuals being addressed and third person is used to refer to an individual or individuals who are neither the speaker nor the listener (Hakulinen et al 2004: §716).

The grammatical person is usually expressed through the verb and the subject being in congruence or following the so-called rule of agreement (Hakulinen 1979, Karlsson 1999), for example minä luen (‘I read’) but sinä luet (‘you read’). In other words, “a verb must agree with its subject in number and person” (Baker 1995: 47). In English,
grammatical person is mainly expressed with personal pronouns, because only the third person singular has its own verb form (“I read” and “you read” but “she reads”). In Finnish, the grammatical person can be seen in genitive structures in addition to verb forms. Both the first and second person have their own possessive suffix for plural and singular, for example *minun kirjani* (‘my book’) but *meidän kirjamme* (‘our book’). Third person has the same possessive suffix for both plural and singular: *hänen kirjansa*, *heidän kirjansa* (Hakulinen et al 2004: §95).

The abovementioned congruence causes repetition: when the first and second person pronouns are used as the subjects in a sentence, they are in a way redundant and not always strictly necessary, because the verb form in itself is enough to indicate who or what the subject of the sentence is (Hakulinen et al 2004: §1268; Hakulinen 1979: 552-553). With the third person the situation is slightly different and the pronoun cannot be omitted as easily as the first and second person pronouns without risking a misunderstanding or making the utterance seem unidiomatic.

If the third person pronoun is omitted, the referent of the utterance becomes ambiguous. There is a tendency in Finnish language to avoid directly addressing individuals (Yli-Vakkuri 2000, quoted in Nuolijärvi and Tiittula 2001: 585). There are two main types of structures in which the referent is unspecific: passive (so-called fourth person) and generic third person (Mauranen and Tiittula 2005: 35, 52). Although these structures do not have a clear referent, according to Mauranen and Tiittula (2005) they are conventional in expressions, which are usually interpreted as referring to first person singular and plural, but depending on the context they can also be used to refer to other grammatical persons.
The first and second person pronouns (*minä* ‘I’, *sinä* ‘you, sg.’, *me* ‘we’, *te* ‘you, pl.’) refer to the participants of a speech act, the speaker and the addressee. These speech act pronouns are deictic, and their referent changes depending on who is speaking (Hakulinen et al 2004: §716). The third person pronouns (*hän* ‘he/she’, *he* ‘they’) refer to the person or persons who are neither the speaker or the addressee, and thus the referent does not necessarily change when the speaker changes (Hakulinen et al 2004: §716).

The SG2 pronoun *sinä* can also be used in generic structures (see Auvinen 2005), but the use is mostly confined to spoken language. The personal pronoun *te* (‘you’, pl.) is occasionally used to address an individual (Hakulinen et al 2004: §1273) as the formal or more polite form of address instead of the more informal *sinä* (‘you’, sg.). According to Nuolijärvi and Tiittula (2001: 586), this use has become less common in Finnish over the years.

Hakulinen (1979: 554) considers the excessive use of personal pronouns as a foreign feature in Finnish and points out that the redundancy lessens considerably if the first and second person pronouns are used only when there is a need to emphasize the speaker or the addressee. Translators should constantly be aware of the choices they make; in Finnish there is less need for overt coding of grammatical person than in English (Mauranen and Tiittula 2005).

In Finnish, personal pronouns are usually perceived as a feature belonging to spoken language. According to Hakulinen (1979: 553-554), one reason for this may be the need for clarity in fast-changing speech situations. Written language does not have similar
requirements, and therefore using the first and second person pronouns is often perceived as placing too much emphasis on the speaker or the addressee: omitting the pronoun may be the more neutral choice (Hakulinen 1979: 553-554).

In spoken Finnish, the demonstrative pronoun se (‘it’) can be used to refer to a person without any derogatory connotations, and in certain cases demonstrative pronouns tämä (‘this’), tuo (‘that’) and se can be used to refer to persons in written texts as well (Hakulinen et al 2004: §720). Thus expressing grammatical person is not purely tied to the use of personal pronouns only. Despite this, demonstrative pronouns have been excluded from this study and the main focus is on the standard forms of personal pronouns.

3.2 Earlier studies

3.2.1 Introduction

An overview of earlier studies is necessary to emphasize a few basic differences occurring in the results of this study, which are discussed in detail in section 6. Mauranen and Tiittula’s study has been the main inspiration for this study. Although the earlier studies presented in the following section concentrate only on the SG1 and SG2 pronouns, they were nevertheless helpful in forming the basis for the theory and hypothesis behind this study. The following sections present the main discoveries made in the studies and their results, which will be compared with the ones gained in this study in section 6.2.5.
3.2.2 The SG1 pronoun

The use of the SG1 pronoun has been examined by Mauranen and Tiittula (2005). They examined the frequency of the SG1 pronoun in translated and original Finnish by comparing its frequency in texts with the help of the FECCS-corpus (Finnish-English Contrastive Corpus Studies) and Finde-corpus. The FECCS-corpus contains texts translated from English to Finnish and vice versa, and the Finde-corpus contains texts written originally in Finnish, among other things. Their material contained fiction translated from German and English into Finnish, which was compared with corresponding texts originally written in Finnish as well as with the English and German source texts of the translations.

Mauranen and Tiittula’s (2005) study showed that the SG1 pronoun was used more in translations than in original Finnish. According to their results, original Finnish had 5.9 instances of the SG1 pronoun per 1 000 words, whereas translated Finnish had 15.0 instances. In both original and translated Finnish, the SG1 pronoun was commonly used in dialogue (Mauranen and Tiittula 2005: 47).

According to Mauranen and Tiittula (2005), in large text collections of translated and original Finnish, the two can be separated from each other on the basis of how frequently pronouns are used. According to them, this is especially true in the case of first and second person singular which are used noticeably more in translations than in original Finnish. Original Finnish contains relatively more third person pronouns than translations (Mauranen and Tiittula 2005).
In Finnish, using the SG1 pronoun together with the verb instead of just the corresponding verb form or a zero subject structure is an overt way of indicating grammatical person (Mauranen and Tiittula 2005). In translated Finnish, the SG1 pronoun was used to imitate spoken language or in idiomatic phrases. Mauranen and Tiittula (2005) criticised this because according to them using the pronoun may emphasize the person or have an opposite effect than intended and make the text seem more like standard language. They point out that in spoken language *minä* is not as common as the dialectal *mä*, for example.

As a whole, the SG1 pronoun was used less in original Finnish than in translated Finnish. Grammatical person was expressed explicitly with a personal pronoun more often in translations than in original Finnish. Original Finnish used more verb forms to indicate person, for example. Thus, according to Mauranen and Tiittula (2005), translated Finnish was more explicit than original Finnish in expressing grammatical person. Mauranen and Tiittula (2005) suggest that translators’ choices cause certain indigenous structures of the target language to be underrepresented in translated texts.

Based on Mauranen and Tiittula’s (2005) results, a small-scale study (Teitto 2008) compared the use of the SG1 pronoun in the subtitles of 43 episodes of the television series *Gilmore Girls* to the light reading and fiction subcorpora of the *Corpus of Translated Finnish* containing both original and translated Finnish. The results showed that the difference between original, subtitled and translated Finnish is not as clear-cut as Mauranen and Tiittula’s (2005) results may lead one to expect. In fiction, the instances of the SG1 pronoun formed 15.0 instances per 1 000 words for translated fiction and 14.7 for original fiction. The light reading corpora had more variation with
18.5 instances per 1 000 words in the translated light reading and 13.3 in the original light reading. The subtitles had 17.0 instances of the SG1 pronoun per 1 000 words. Subtitled and translated Finnish generally used the SG1 pronoun more than original Finnish, but especially in the fiction corpora, the difference is small.

However, it should be noted that, in both studies, the subtitles were compared with corpora which did not consist of pure dialogue. As Mauranen and Tiittula (2005: 47) point out, most personal pronouns appear in dialogue. Subtitles of television dramas consist of dialogue only, whereas fiction and light reading are a mixture of narrative and dialogue, with dialogue generally forming a smaller part of the text than narrative. When this is taken into consideration while examining the results in Teitto’s (2008) study, the frequency of the SG1 pronoun in the subtitles appears fairly low.

In conclusion, Mauranen and Tiittula’s (2005) results showed that the SG1 pronoun is clearly more common in translations than in original Finnish. There was no difference in its use in translated and original Finnish; it was often used either to emphasise the speaker or in idiomatic structures. In translated Finnish using the SG1 pronoun without a clear function was nevertheless more common (Mauranen and Tiittula 2005). Teitto’s (2008) study on subtitles showed that subtitled and translated Finnish resemble each other because they both use the SG1 pronoun more than original Finnish. However, the difference between original and translated Finnish was not as clear-cut as in Mauranen and Tiittula’s (2005) study. The suitability of the material used as a point of comparison in Teitto’s (2008) study was also questioned because it did not consist purely of dialogue.
3.2.3 The SG2 pronoun

Mauranen (2002) has also studied the use of the SG2 pronoun. In her study, the material consisted of popular non-fiction texts from the Corpus of Translated Finnish, which were written originally in Finnish or translated from English to Finnish. The corpora were compared with the help of the Key Word tool in WordSmith Tools, which is a programme developed for corpus analysis (Scott 1998). Mauranen’s (2002) results showed that the SG2 pronoun and related verb forms formed the top of the list. The instances of the SG1 pronoun did not have such clear differences in the material (Mauranen 2002).

A concordance search showed that in popular non-fiction, original Finnish contained 60 instances of the SG2 pronoun per 100,000 words and the translated Finnish contained clearly more instances with 631 per 100,000 words. In academic texts, the difference was not as large but still clear: there were 26 instances of the SG2 pronoun per 100,000 words in original Finnish and 126 in translated Finnish. Most instances of the SG2 pronoun came from the so-called self-help books, in which original Finnish contained 173 instances of the SG2 pronoun per 100,000 words and translated Finnish 1,408 instances (Mauranen 2002).

The case forms had a few differences mainly in relation to nominative case, which was “the second most frequent in the originals, but ranked only fifth in the translations” (Mauranen 2002). According to Mauranen (2002), the SG2 pronoun does therefore not appear to be used particularly often in subject position compared to its other uses in the translations.
Inspired by Mauanen’s study on the SG2 pronoun, Auvinen (2005) has studied the
generic use of second person singular in Finnish texts. Although the generic second
person is a feature occurring in Finnish, the use of the personal pronoun in the structure
may have been affected by English. Auvinen used the Corpus of translated Finnish as
the research material.

In her study, Auvinen (2005) ran a concordance search in the corpora to find all
occurrences of the SG2 pronoun sinä. The number of occurrences was 3 900 instances
of the SG2 pronoun per million words in translated Finnish and 2 019 in original
Finnish. The dialectal form sää had 132 occurrences per million words in translated
Finnish and 552 in original Finnish. Even when the occurrences for both the standard
and dialectal form are added together, it is clear that the SG2 pronoun in Auvinen’s
(2005) study is noticeably more common in translated Finnish that in original Finnish.

In conclusion, the occurrences and uses of the SG2 pronoun were similar to the SG1
pronoun. The SG2 pronoun appeared to be clearly more common in translations than in
original Finnish in both of the studies, but the genre of the texts in the corpora seemed
to affect the frequency of the SG2 pronoun.
4 PROPER NOUNS

Proper nouns are names usually consisting of one or more words (Hakulinen et al 2004: §596). Their main function is to name an entity in order to distinguish it from others of the same type, for example a person, a place or a product. The purpose of a proper noun is to identify its bearer, but since proper nouns are not descriptive, it is difficult to know what their bearer is like without context. Therefore the context in which the proper noun appears or the person who utters it defines to whom or what a proper noun refers (Hakulinen et al 2004: §596-598).

The line between proper nouns and common nouns is not always clear: according to Hakulinen et al (2004: §553), proper nouns and common nouns can be used similarly and share similar functions (for example äiti, ‘mother’, can function as both a common noun and proper noun). In this study common nouns used as proper nouns are excluded, although they may be an interesting area of future research.

No qualitative studies seem to exist on the possible differences in the use and amount of proper nouns in texts written in different languages. There is also a lack of studies on proper nouns which would compare their use in translated texts with texts written originally in Finnish. Current research on proper nouns is concentrated on their
etymology, or in the case of translation studies, on translating the names (see e.g. Kauhanen 2004). Some studies have concentrated on the possible changes occurring when a personal pronoun is replaced with another noun (including proper nouns) in translations (see e.g. Rivinoja 2004).

Proper nouns were nevertheless chosen as the other subject for this study in order to provide a point of comparison for personal pronouns. Both categories share a few similar features. They can be used as a way to address the participants directly in conversation. In addition to this, proper nouns can be replaced by pronouns without changing the structure of a sentence. Replacing proper nouns with pronouns is a way to avoid repetition.

Addressing a person by their name is seen as something that varies between different cultures. A general opinion is that Finnish uses fewer proper nouns than for example English (Nuolijärvi and Tiittula 2001: 592). Vertanen (2007: 152) observes that proper nouns are very common in American polite address in television series and recommends leaving out unnecessary names from subtitles because of the character limit. Nuolijärvi and Tiittula (2001: 581, 592) point out that, although direct address is said to be avoided in Finnish, the way in which the participants are addressed or referred to depends on several other factors, such as the participants’ relationship, in addition to cultural norms.

Nuolijärvi and Tiittula (2001: 592) line out different ways for the speaker to address the listener or listeners: by name, with a pronoun, by title and name or with generic third person. How grammars define the use of proper nouns differs slightly between Finnish and English, especially when proper nouns are used to address the participants of a
conversation. According to Hakulinen et al (2004: §1077), proper nouns are used to
direct the addressee’s attention to the speaker. The use varies depending on how many
participants there are in a conversation. In a conversation with multiple participants,
proper nouns are used to indicate to whom the message is meant and is usually related
to commands or questions requiring an answer, for example. In a conversation between
two participants, addressing a person by name when no reply is expected is rare.
Hakulinen et al (2004: §1077) consider this as special use, which is explained by
cultural or personal differences.

The above-mentioned definition differs from the one given in English grammars.
Oxford English Grammar (Greenbaum 1996), for example, uses the term vocative or
vocative expressions and defines them as

optional additions to basic sentence structures that are used to address the person
or persons spoken to, either to single them out from others or to maintain some
personal connection with them (Greenbaum 1996, emphasis added).

The definitions emphasize a cultural difference: whereas in English proper nouns seem
to be used to maintain personal connection, similar use in Finnish is seen as special use
which deviates from the norm. The definitions give a reason to expect that the use of
proper nouns as vocatives varies between original and translated Finnish. It is worth
noting that the above-mentioned definitions do not necessarily represent how proper
nouns are actually used. Nuolijärvi and Tiittula (2001) emphasise that addressing the
participants in a conversation is vital for creating relationships between them in Finnish
as well.
5 MATERIAL AND METHOD

5.1 Material

Studies utilizing corpora are fairly common. According to Mauranen and Tiittula (2005: 37) comparing target language texts with each other gives information about the kind of linguistic characteristics that separate translations from other texts written in the same language. The research material of this study is compiled from two different corpora. The first corpus used for compiling material is The Finnish Broadcasting Company corpus of subtitles (YLE-korpus), which is a collection of digital research material of translated subtitles compiled by Mäkisalo and Tirkkonen-Condit (2005). From the corpus, the Finnish subtitles of two episodes of Gilmore Girls, a long-running American television series, were selected. The subtitles from the two episodes form the corpus of subtitled Finnish, which consists of 7 294 words (table 1).

The Gilmore Girls was chosen as research material because of the series’ fast-paced dialogue, which is one of the defining features of the series and can pose an extra challenge for the translator. Gilmore Girls is a mixture of drama and comedy, and concentrates around the lives of Lorelai Gilmore and her daughter Rory in Stars
Hollow, a small fictional town in Connecticut. Both of the episodes were broadcast in 2003 on YLE TV1.

The basis of comparison for the subtitles were compiled from the Corpus of Translated Finnish (Käännössuomen korpus), which is a collection of digital research material of both translated and original Finnish. The corpus is divided into several subcorpora, which consist of five different literary genres: academic texts, children’s books, fiction, popular science and light reading. Corpus of Translated Finnish contains roughly four million words of original Finnish and approximately the same amount of words of English to Finnish translations.

For this study, extracts have been selected from altogether eight books from the fiction and light reading subcorpora. Fiction usually contains a lot of dialogue (Mauranen and Tiittula 2005: 47), and light reading can be expected to be similar. The effect of genre on the results is not examined except when comparing the results with those gained in earlier studies. Four of the extracts are translated from English to Finnish and the other four are original Finnish. In each group two of the extracts are written in first person and two in third person. The extracts of translated Finnish form a corpus of 27,994 words, and the extracts of original Finnish a corpus of 27,319 words (table 1).

The two above-mentioned corpora have also been divided into dialogue and narrative in order to provide a better point of comparison for the subtitles. Only direct speech is considered as dialogue, because in this study the subtitles examined are by default direct speech; indirect speech is considered as narrative. Reporting clauses (such as minä sanoin, ‘I said’) have been included in the narrative. In problematic cases (i.e. where the
line between dialogue and narrative was unclear) punctuation and context were used as defining factors. The translated Finnish dialogue (TFD) forms a corpus of 8 432 words and the original Finnish dialogue (OFD) forms a corpus of 6 988 words. The corpus of translated Finnish narrative (TFN) consists of 19 512 words and the corpus of original Finnish narrative (OFN) of 20 331 words (table 1).

<table>
<thead>
<tr>
<th>TABLE 1. Number of words in each corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undivided</td>
</tr>
<tr>
<td>Translated Finnish</td>
</tr>
<tr>
<td>Original Finnish</td>
</tr>
<tr>
<td>Subtitles</td>
</tr>
</tbody>
</table>

It is worth emphasising that subtitled Finnish and translated Finnish are considered as separate in this study, although the subtitles examined are translated from English to Finnish and thus are translated Finnish. The main purpose of this study is to examine how subtitles are related to other kinds of translation and to original Finnish. Examples presented in this study have been taken from the corpora and translated. In the case of subtitled and translated Finnish, the translations are based on the target text with no reference to the original.

5.2 Method

First, a small-scale sample of the material was examined, which showed that proper nouns and personal pronouns form a large part of nouns with human referent. The WordSmith Tools programme (Scott 1998), which is especially designed for corpus
analysis, was then used to analyse the different corpora. The two main tools used are the Wordlist and the Concordancer.

When examining the proper nouns, first an alphabetical wordlist was created separately for each corpus, and the list was then used as a guide for running concordance searches into the corpora in order to find the proper nouns and their context. In the case of personal pronouns, a concordance search for each personal pronoun and their inflections was made separately into each corpus. Context was used to eliminate noise: for example the essive case of the demonstrative pronoun *se* (*sinä iltana* ‘on that evening’) is the same as the SG2 pronoun *sinä* (‘you’).

The different case forms (for example the adessive case: *minulla, Askolla*; the genitive case: *minun, Askon*) as well as other suffixes, such as particles (for example *minullako, meidänkin*), are taken into consideration when running the searches. The dialectal variants of personal pronouns, such as *mä* (‘I’), were not included in the results, because their use was almost non-existent in all of the corpora (see section 6.1.3 on the use of non-standard language in the material). In proper nouns, there were a few instances where it was unclear whether the name referred to a human or to something else: for example “Norton” was used both as the name of a company and as a family name. The context was used to determine whether the referent was human or not, and instances of proper nouns with non-human referent were excluded. The results for each corpus were then compared with each other and analysed.
6 ANALYSIS

Section 6.1 gives a general overview of proper nouns and personal pronouns in the material. They will then be examined separately in section 6.2, which concentrates on personal pronouns, and in section 6.3, which centres on proper nouns.

6.1 Human Reference

6.1.1 Undivided material

Examining a small-scale sample of the material used in this study shows that nouns with a human referent form a large part of the corpora, although their amount varies slightly between the individual corpora. Approximately 14 to 15 per cent of the words used in the OFN, TFN and TFD corpora are nouns with a human referent. The SF corpus contains the least amount of them with approximately 12 per cent and the OFD corpus the most with 20 per cent. The high amount of nouns with human referent in the OFD may be partially explained by the fact that the extracts in the small-scale sample are from the beginning of the texts, where it is likely that more human referents are used
when creating the setting. Although the same applies to the other extracts as well, the frequency of the occurrences may even out in a larger sample.

Two of the largest subgroups were chosen for a closer examination: personal pronouns and proper nouns. Even in the OFD corpus where the amount of nouns with human referent is the highest, the personal pronouns and proper nouns clearly form the largest part of the words used. They were also chosen because they can be defined clearly, which makes it simpler to sort them and to eliminate noise. Other commonly occurring nouns with human referent in the corpora are demonstrative pronouns and common nouns describing family ties, for example.

The personal pronouns and proper nouns form approximately one tenth of the words used in the corpora.

**TABLE 2. The percentage of personal pronouns, proper nouns, and personal pronouns and proper nouns added together**

<table>
<thead>
<tr>
<th></th>
<th>pronouns %</th>
<th>proper nouns %</th>
<th>added %</th>
</tr>
</thead>
<tbody>
<tr>
<td>OF</td>
<td>5.0</td>
<td>4.8</td>
<td>9.8</td>
</tr>
<tr>
<td>TF</td>
<td>7.1</td>
<td>4.1</td>
<td>11.2</td>
</tr>
<tr>
<td>SF</td>
<td>6.9</td>
<td>3.2</td>
<td>10.1</td>
</tr>
</tbody>
</table>

The instances of personal pronouns and proper nouns in the material suggest that the need for personal reference is fairly constant in texts, regardless of whether the text is a translation or not. The study of the small-scale sample showed that there is not much variation in the amount of personal reference in general in the corpora. The same applies even more clearly to proper nouns and personal pronouns; there is little variation
in their amount when added together (see table 2). The results also indicate that the use of personal pronouns is dependent on the use of proper nouns and vice versa: for example, a corpus with a high frequency of personal pronouns contains fewer proper nouns.

Although the need to use personal pronouns and proper nouns seems fairly constant, there are some differences. The TF corpus uses personal pronouns and proper nouns the most with 11.2 per cent of the corpus consisting of them, while the OF corpus has the least amount of them with 9.8 per cent. The SF corpus is situated in between with 10.1 per cent \( (\chi^2 p = .000) \). Only the difference between the OF and SF corpus is statistically insignificant \( (\chi^2 \) with Yates’ correction \( p = .545) \), indicating that subtitled Finnish is closer to original Finnish than to translated Finnish.

6.1.2 Dialogue and narrative

The aforementioned trend of constancy remains the same when the material is divided into narrative and dialogue (see Table 3). For example, 5.0 per cent of the OF corpus consists of personal pronouns and 4.8 per cent of proper nouns. Added together they form 9.8 per cent of the words used in the corpus. In the OFD corpus, personal pronouns form 7.4 per cent of the corpus and proper nouns 2.4 per cent, but added together their percentage is the same as in the undivided material.

The same trend can be seen in the OFN corpus as well: 4.1 per cent of it is personal pronouns and 5.7 per cent proper nouns. Both the TFD and TFN corpus have slightly
more variation when compared with the frequency of proper nouns and personal pronouns in the undivided corpus. Personal pronouns form 8.1 per cent and proper nouns 2.3 per cent of the TFD corpus adding up to 10.4 per cent. In the TFN corpus, 6.7 per cent of the words are personal pronouns and 4.9 per cent proper nouns, adding up to 11.6 per cent.

### TABLE 3. The percentage of pronouns, proper nouns, and both added together when examining material divided into dialogue and narrative.

<table>
<thead>
<tr>
<th></th>
<th>pronouns %</th>
<th>proper nouns %</th>
<th>added %</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFD</td>
<td>7.4</td>
<td>2.4</td>
<td>9.9</td>
</tr>
<tr>
<td>TFD</td>
<td>8.1</td>
<td>2.3</td>
<td>10.4</td>
</tr>
<tr>
<td>SF</td>
<td>6.9</td>
<td>3.2</td>
<td>10.1</td>
</tr>
<tr>
<td>OFN</td>
<td>4.1</td>
<td>5.7</td>
<td>9.8</td>
</tr>
<tr>
<td>TFN</td>
<td>6.7</td>
<td>4.9</td>
<td>11.6</td>
</tr>
</tbody>
</table>

The instances of proper nouns and personal pronouns in the narrative and dialogue corpora show that their distribution in the corpora is very similar to the undivided material. Translated Finnish still contains the most instances of proper nouns and personal pronouns with 10.4 for dialogue and 11.6 for narrative, and original Finnish the least with 9.9 for dialogue and 9.8 for narrative. The SF corpus remains between the two with 10.1 per cent of the corpus consisting of proper nouns and personal pronouns (table 3).

Comparing the dialogue corpora shows that the difference between them is insignificant ($\chi^2 p = .556$) and remains so in pairwise comparisons between the TFD and SF corpora, the OFD and SF corpora as well as the SFD and TFD corpora ($\chi^2$ with Yates’ correction $p = .500, p = .748$ and $p = .304$ respectively). On the basis of this, it seems that the
amount of proper nouns and personal pronouns is fairly constant in dialogue regardless of it being translated, subtitled or original, although there is some slight variation between the corpora.

In the comparison made between the two narrative corpora, the difference is significant ($\chi^2$ with Yates’ correction $p = .000$). If the SF corpus is included in the comparison, the difference between the three remains significant ($\chi^2 p = .000$), although comparing subtitles with narrative is not very practical due to their inherently different nature. Only comparing the OFN corpus to the SF shows that the difference is insignificant ($\chi^2$ with Yates’ correction $p = .520$); comparison between the TFN and the SF corpora shows that the difference is significant ($\chi^2$ with Yates’ correction $p = .000$).

A few differences become apparent when examining the distribution of proper nouns and personal pronouns separately. The TFD corpus uses personal pronouns the most (8.1 per cent) and proper nouns the least (2.3 per cent). The percentage of proper nouns in the TFD and the OFD corpora is fairly similar with 2.3 per cent and 2.4 per cent respectively, but the SF corpus differs from them both with 3.2 per cent of the words consisting of proper nouns. Similarly, the SF corpus uses personal pronouns the least with 6.9 per cent. The instances of personal pronouns and proper nouns are examined separately in the following sections.

The results indicate that subtitled Finnish seems to differ from translated Finnish in regard to both proper nouns and personal pronouns. Subtitled Finnish differs from original Finnish as well, but there are more similarities between the two than between subtitled and translated Finnish. It is worth noting that the difference between original
Finnish and subtitled Finnish remains statistically insignificant when comparing the added together instances of personal pronouns and proper nouns, regardless of the OF corpus being undivided, narrative or dialogue. This suggests that subtitled Finnish is closer to original Finnish than to translated Finnish in regard to the use of personal pronouns and proper nouns.

There are nevertheless enough differences between the OF, TF and SF corpora to indicate that, as Jääskeläinen (2007) suggested, the special characteristics of subtitles may affect them in a way which causes them to differ from original Finnish as well as translated Finnish. The differences become clearer when examining the proper nouns and personal pronouns separately in the following sections.

6.1.3 A note on the use of non-standard language

The material of this study contains very few, if any, dialectal variants of personal pronouns. No dialectal variants appear at all in the TF corpus, and even in the OF corpus their frequency is very low. In all of the corpora, the language mainly follows the rules of standard Finnish, but there are a few deviations. One such aspect is using genitive constructions (a possessive pronoun and a noun) without the possessive suffix as in examples 1 and 2. The standard expressions would be *minun lävitseni* and *meidän kertomuksemme*.

   ‘You could walk right through me and you wouldn’t even notice it.’
2. Onko tämä meidän kertomus?
‘Is this our story?’

Another non-standard structure, which occurs commonly in spoken language, is using passive verb forms with the PL1 pronoun as in example 3. The standard and grammatically correct expression would be *me luimme*.

3. Ei, ei. Me... luettiin. Joo, me luettiin.
‘No, no. We… were reading. Yeah, we were reading.’

In the SF corpus, there are no deviations from standard Finnish in the instances of personal pronouns. Both possessive suffix and correct verb forms are always used. In the TFD corpus, the non-standard use is rare and connected to the speech of young characters, as in the example 3 above.

It is clear that, despite the few instances of non-standard language use, the language in the corpora mostly adheres to the rules of standard Finnish. Nevalainen (2003) discovered in his study that the feeling of colloquial language in texts was created by avoiding long and abstract words and by favouring colloquial pronunciations and vocabulary. The material of this study showed that word choices were employed to certain extent to create the feeling of colloquial language as in examples 4 and 5. In example 4, *telkkari* is the more colloquial version of *televisio*, and in example 5 the more standard term for *kämppä* would be *asunto*, for example.

4. Minun telkkarini räjähti.
‘My TV exploded.’
5. Onko sinulla kämpässäsi mitään juotavaa?
‘Do you have anything to drink at your place?’

The word choices mainly represent the more common, everyday vocabulary instead of representing a specific dialect. It could be then argued that, instead of using dialects or very colloquial expressions, the writers rely on other methods for creating the feeling of colloquial language in dialogue.

6.2 Personal pronouns

6.2.1 Undivided material

Overall, personal pronouns form a significant portion of the words used in all of the corpora (table 4). In the TF corpus, 7.1 per cent of the words are personal pronouns. The same is true for 5.0 per cent of the OF corpus. The frequency of personal pronouns in the SF corpus is closest to translated Finnish with 6.9 per cent of the corpus consisting of personal pronouns (see table 4).

<table>
<thead>
<tr>
<th></th>
<th>Undivided %</th>
<th>Dialogue %</th>
<th>Narrative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translated Finnish</td>
<td>7.1</td>
<td>8.1</td>
<td>6.7</td>
</tr>
<tr>
<td>Original Finnish</td>
<td>5.0</td>
<td>7.4</td>
<td>4.1</td>
</tr>
<tr>
<td>Subtitles</td>
<td>6.9</td>
<td>6.9</td>
<td>-</td>
</tr>
</tbody>
</table>
Comparing the instances of personal pronouns in the three corpora shows that the difference between them is statistically significant ($\chi^2 p = .000$). If the comparison is made between the SF corpus and the TF corpus, the difference is insignificant ($\chi^2$ with Yates’ correction $p = .648$), which suggests that subtitled Finnish is closer to translated Finnish than to original Finnish in regard to the use of personal pronouns, unlike the results gained by comparing the added together instances in section 6.1. Original hypothesis of subtitles using less personal pronouns seems invalid, because although the SF corpus contains less personal pronouns than the TF corpus, it still contains noticeably more of them than the OF one.

According to Mauranen and Tiittula (2005), when examining large masses of texts containing both translations and original texts, it is possible to distinguish them from each other on the basis of the frequency of pronouns. Their observation is valid when looking at the instances of personal pronouns in the undivided material: translated and subtitled Finnish both have more personal pronouns than original Finnish.

6.2.2 Dialogue

As noted earlier (see section 2.2), subtitles consist purely of dialogue. Thus comparing them with corpora containing both narrative and dialogue does not offer a valid point of comparison. Dividing the material into narrative and dialogue creates a few differences to the above-mentioned results for undivided material.
In general, dialogue contains a clearly higher proportion of personal pronouns than narrative. In translated Finnish, 8.1 per cent of the dialogue consists of personal pronouns, whereas the same is true for 7.4 per cent of the original Finnish dialogue. The narrative corpora use fewer personal pronouns than any of the dialogue corpora with 6.7 per cent for translated Finnish and 4.1 for original (table 4).

The TFD corpus contains the largest amount of personal pronouns. In the undivided material, the amount of personal pronouns in subtitled Finnish is quite similar to that of translated Finnish. In the divided material, they differ greatly: 6.9 per cent of the SF corpus is personal pronouns, which is clearly less than in the OFD or TFD corpora.

Comparing the amount of instances shows that the difference is statistically significant \((\chi^2 p = .021)\) when comparing all three corpora. The same is true when comparing the SF and TFD corpora \((\chi^2 \text{ with Yates' correction } p = .006)\) but comparisons between the SF and OFD corpora and between the OFD and TFD corpora are insignificant \((\chi^2 \text{ with Yates' correction } p = .244 \text{ and } \chi^2 \text{ with Yates' correction } p = .144 \text{ respectively})\).

Personal pronouns form a clearly smaller portion of subtitled Finnish than of original or translated Finnish. Especially the difference between subtitled Finnish and other translated Finnish dialogue is clear. The results support the original hypothesis; subtitled Finnish uses fewer personal pronouns than original or translated Finnish. The possible reasons for this are the lack of space and time which limit the translator’s choices, as outlined in section 2.3.
6.2.3 Narrative

Narrative contains fewer personal pronouns than dialogue. In the TFN corpus, 6.7 per cent of the words are personal pronouns. The OFN corpus uses personal pronouns the least with 4.1 per cent ($\chi^2$ with Yates’ correction $p = .000$). As discussed in section 6.1.1, the use of personal pronouns and proper nouns are dependent on each other. Thus the lower frequency of personal pronouns in the narrative is at least partially explained by the higher frequency of proper nouns in the material (5.7 per cent in the OFN corpus and 4.9 per cent in the TFN corpus).

The amount of personal pronouns in the TFN corpus is fairly close to their amount in the SF corpus, although the comparison is not truly valid due to the nature of subtitles. The difference between the SF and TFN corpora is insignificant ($\chi^2$ with Yates’ correction $p = .446$), but the difference is significant in a comparison between the three corpora as well as between the TFN and SF corpora ($\chi^2 p = .000$ and $\chi^2$ with Yates’ correction $p = .000$ respectively).

6.2.4 Distribution of grammatical person

As shown above, there is not always a clear difference in the frequency of personal pronouns between two different types of corpora. Nevertheless, dialogue and narrative are clearly different from each other in regard to the use of personal pronouns, which is best illustrated by examining the distribution of grammatical person in the material. The instances of personal pronouns in all of the corpora were divided into two groups on the
basis of which person (first, second or third) they represent. The first group consists of all the instances of the first and second person, and the second group consists of third person only (table 5). Both plural and singular forms are included.

TABLE 5. The distribution of the plural and singular forms of the first and second person, as well as the third person as a percentage of all the instances of personal pronouns in the corpora

<table>
<thead>
<tr>
<th></th>
<th>undivided</th>
<th></th>
<th></th>
<th>dialogue</th>
<th></th>
<th></th>
<th>narrative</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st, 2nd</td>
<td>3rd</td>
<td></td>
<td>1st, 2nd</td>
<td>3rd</td>
<td></td>
<td>1st, 2nd</td>
<td>3rd</td>
<td></td>
</tr>
<tr>
<td>Translated Finnish</td>
<td>40.9</td>
<td>59.1</td>
<td>77.1</td>
<td>22.9</td>
<td>21.9</td>
<td>78.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Finnish</td>
<td>53.2</td>
<td>46.8</td>
<td>91.7</td>
<td>8.3</td>
<td>29.3</td>
<td>70.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtitles</td>
<td>59.4</td>
<td>40.6</td>
<td>59.4</td>
<td>40.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The basis for the division is the different uses of first, second and third person: as mentioned in section 3.1, first and second person pronouns are so-called speech act pronouns, which refer to the participants of a speech act, whereas third person refers to the individual or individuals who are neither the speaker or the addressee (Hakulinen et al 2004: §716).

The type of the text affects how personal pronouns representing different grammatical person are distributed. In the undivided material, each corpus contains a similar amount of the first and second person and the third person (table 5). In the TF corpus, 40.9 per cent of the personal pronouns are first or second person pronouns and 59.1 per cent third person. In the OF corpus, 53.2 per cent of the personal pronouns consist of first or second person and 46.8 per cent of third person.
In the SF corpus, first and second person pronouns form 59.4 per cent of the personal pronouns and third person 40.6 per cent. As it was shown earlier, subtitled Finnish uses personal pronouns less than other dialogue. Examining the distribution of the first and second and the third person pronouns shows that the proportion of the first and second person is relatively higher in the SF corpus than in the other corpora, which seems to contradict the original hypothesis of subtitles using less first and second person pronouns.

Dividing the TF and OF corpora into narrative and dialogue changes the distribution. In general, the personal pronouns in the dialogue consist mainly of the first and second person (77.1 per cent for the TFD and 91.7 for the OFD), whereas in the narrative, the majority are third person pronouns (78.1 per cent for the TFN and 70.7 per cent for the OFN).

Comparing the SF corpus with the OFD and TFD corpora shows that the SF corpus does not have such a clear-cut difference between the use of the first and second person pronouns and the third person pronouns. The subtitles use the first and second person pronouns less (59.4 per cent) and the third person pronouns more (40.6 per cent) than either original or translated Finnish.

The results give further support for the original hypothesis: subtitles have relatively less first and second person pronouns and relatively more third person pronouns when compared with other dialogue, which indicates that personal pronouns are omitted if possible. The high proportion of third person pronouns when compared with other dialogue is also partially explained by this, because they cannot be omitted as easily.
The results seem to disprove the hypothesis in which it was expected that subtitles would use the third person less in order to avoid ambiguity, but if the relatively high frequency of proper nouns in the SF (see section 6.3) is taken into account while examining the results, the hypothesis seems plausible. The amount of third person pronouns in the SF corpus seems higher because of the omitted first and second person pronouns.

In addition to examining how the first and second person and third person are distributed in the material, examining the ratio of each personal pronoun separately further illustrates the above-mentioned influence that the narrative and dialogue have on the use of personal pronouns.

The SG1, SG2 and SG3 pronouns are generally the most common ones in the material. Examining the undivided material shows that the proportion of the SG3 pronoun is very high with 50.7 per cent in the TF corpus and 42.4 per cent in the OF one. In the SF corpus, the SG3 pronoun constitutes 37.2 per cent of the personal pronouns used. Comparison with the frequency of third person pronouns (table 5) in general emphasises that only a small portion of them are PL3 pronouns.

Examining the material divided into narrative and dialogue again changes the distribution of personal pronouns (table 6). In the TFD and OFD corpora, the frequency of the SG1 and SG2 pronouns in the corpora is proportionally higher than that of the SG3 pronoun. In the SF corpus, however, the SG3 pronoun forms the largest proportion of the personal pronouns used in the corpus with 37.2 per cent, which is nearly twice as
much as in the TFD corpus (19.6 per cent) and five times as much as in the OFD corpus (7.3 per cent).

TABLE 6. Proportions of each personal pronoun from all of the personal pronouns in the corpora

<table>
<thead>
<tr>
<th></th>
<th>SG1</th>
<th>SG2</th>
<th>SG3</th>
<th>PL1</th>
<th>PL2</th>
<th>PL3</th>
</tr>
</thead>
<tbody>
<tr>
<td>OF</td>
<td>30.3</td>
<td>15.5</td>
<td>42.4</td>
<td>4.9</td>
<td>2.5</td>
<td>4.4</td>
</tr>
<tr>
<td>TF</td>
<td>24.7</td>
<td>8.3</td>
<td>50.7</td>
<td>4.9</td>
<td>2.9</td>
<td>8.4</td>
</tr>
<tr>
<td>SF</td>
<td>26.7</td>
<td>24.0</td>
<td>37.2</td>
<td>5.0</td>
<td>3.8</td>
<td>3.4</td>
</tr>
<tr>
<td>OFD</td>
<td>41.5</td>
<td>35.0</td>
<td>7.3</td>
<td>8.7</td>
<td>6.5</td>
<td>1.0</td>
</tr>
<tr>
<td>TFD</td>
<td>33.4</td>
<td>23.5</td>
<td>19.6</td>
<td>12.0</td>
<td>8.2</td>
<td>3.2</td>
</tr>
<tr>
<td>OFN</td>
<td>23.3</td>
<td>3.4</td>
<td>64.3</td>
<td>2.6</td>
<td>0.0</td>
<td>6.5</td>
</tr>
<tr>
<td>TFN</td>
<td>20.2</td>
<td>0.3</td>
<td>67.0</td>
<td>1.2</td>
<td>0.2</td>
<td>11.1</td>
</tr>
</tbody>
</table>

The SG1 pronoun is used the least in the SF corpus with 26.7 per cent and the most in the OFD corpus with 41.5 per cent, while the TFD corpus is situated in-between with 33.4 per cent. The frequency of the SG2 pronoun does not vary as much: it constitutes 24.0 per cent of the personal pronouns in the SF corpus, 23.5 per cent in the TFD corpus and 35.0 per cent in the OFD corpus.

In narrative, the SG3 pronoun is the most common one, and there are hardly any second person pronouns. The few instances of the SG2 pronoun in the OFN (3.4 per cent) come from a letter that addresses the reader directly and could, in fact, be considered as dialogue. The instances of the SG2 pronoun in TFN (0.3 per cent) are discussed more in depth in section 6.2.7. The amount of the first person pronouns in the narrative is most probably affected by the fact that some of the extracts in the corpora have first-person narration.
The results show that subtitled Finnish differs from original and translated Finnish, especially when examining dialogue. It should be noted that the authors’ style and the topics the characters discuss in the excerpts may cause some variation in the results: for example, if the characters always refer to third-person parties by proper nouns instead of using third person pronouns.

6.2.5 Comparison with earlier studies

The earlier studies on the use of personal pronouns in literature have been discussed in section 3.2. Although they concentrate on the SG1 and SG2 pronouns, they still offer a useful point of comparison for the results gained in this study.

Mauranen and Tiittula (2005) note in their study that texts written originally in Finnish and texts translated into Finnish can be distinguished from each other on the basis of personal pronouns: the SG1 and SG2 pronouns are used noticeably more in translations. According to them, original Finnish contains more third person, but it is not specified whether this refers to third person in general, or just personal pronouns or verb forms.

To compare the results with those of the earlier studies, it is necessary to examine the proportion of the SG1 and SG2 pronouns of all the words used in the corpora. The SG1 pronoun forms 1.75 per cent of the words used in the TF corpus and 1.5 per cent in the OF corpus. In the undivided material, the SG1 pronoun is the most frequent in the SF corpus with 1.85 per cent (table 7).
TABLE 7. The percentage of the SG1 pronoun in the material of this study and in Mauranen and Tiittula’s results (M&T)

<table>
<thead>
<tr>
<th>SG1</th>
<th>undivided</th>
<th>M&amp;T</th>
<th>narrative</th>
<th>dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>1.75</td>
<td>1.50</td>
<td>1.34</td>
<td>2.70</td>
</tr>
<tr>
<td>OF</td>
<td>1.50</td>
<td>0.59</td>
<td>0.95</td>
<td>3.09</td>
</tr>
<tr>
<td>SF</td>
<td>1.85</td>
<td>-</td>
<td>-</td>
<td>1.85</td>
</tr>
</tbody>
</table>

In Mauranen and Tiittula’s (2005) results, the SG1 pronoun constituted 0.59 per cent of the words in original Finnish, and translated Finnish used the SG1 pronoun almost three times as much with 1.5 per cent. The instances of the SG1 pronoun are more common in the material used in this study than in Mauranen and Tiittula’s (2005), and although the SG1 pronoun is more common in the TF corpus than in the OF one, the difference between them is not as clear. The situation is opposite when examining dialogue: the OFD corpus contains more SG1 pronouns than the TFD corpus with 3.09 per cent and 2.7 per cent respectively. In the OFN corpus, the frequency of the SG1 pronoun is 1.34 per cent and in the TFN corpus 0.95 per cent.

TABLE 8. The percentage of the SG2 pronoun in the material of this study and in Auvinen’s results (Auvinen)

<table>
<thead>
<tr>
<th>SG2</th>
<th>undivided</th>
<th>Auvinen</th>
<th>narrative</th>
<th>dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>0.59</td>
<td>0.39</td>
<td>0.02</td>
<td>1.90</td>
</tr>
<tr>
<td>OF</td>
<td>0.77</td>
<td>0.20</td>
<td>0.14</td>
<td>2.60</td>
</tr>
<tr>
<td>SF</td>
<td>1.66</td>
<td>-</td>
<td>-</td>
<td>1.66</td>
</tr>
</tbody>
</table>

The SG2 pronoun forms 0.59 per cent of the TF corpus and 0.77 per cent of the OF. In the SF corpus, it is used clearly the most with 1.66 per cent (table 8). Auvinen’s (2005) results showed that translated Finnish used the SG2 pronoun almost twice as much as
original Finnish with 0.39 per cent and 0.20 per cent respectively. Similar to the results of the SG1 pronoun, the SG2 pronoun occurs more frequently in the material of this study than in Auvinen’s.

According to Mauranen and Tiittula (2005), the SG2 pronoun should be more frequent in the TF corpus than in the OF corpus as in Auvinen’s results. In the material of this study, the SG2 pronoun is more common in original Finnish, if subtitled Finnish is excluded. This applies to dialogue narrative as well: the SG2 pronoun forms 2.6 per cent of the words used in the OFD corpus and 1.9 per cent in the TFD corpus. In narrative, the frequency of the SG2 pronoun is lower than in Auvinen’s (2005) study with 0.02 per cent for TFN and 0.14 per cent for OFN.

The use of third person in the material of this study does not follow Mauranen and Tiittula’s (2005) observation of original Finnish using it more: in the TF corpus, the third person pronouns form a larger portion of the personal pronouns than in the OF corpus. When examining their portion of all the words in the corpora, 4.19 per cent of the TF corpus consists of third person pronouns whereas the same is true for 2.32 per cent of the OF corpus, if both singular and plural forms are taken into account.

Only the SF corpus seems to follow Mauranen and Tiittula’s (2005) observation: the SG1 pronoun constitutes 1.85 per cent of the words in the SF corpus and the SG2 pronouns 1.66 per cent of the corpus. The amount of the SG2 pronoun is clearly higher in the SF corpus than in the OF corpus or even in the TF one, and the amount of the SG1 pronoun is higher as well, but the difference is not as great as in the SG2 pronoun. Comparing the SF corpus to the dialogue corpora shows that the SF corpus has fewer
instances of the SG1 and SG2 pronouns than the OFD and TFD corpora (see Tables 7 and 8). At the same time the SF corpus uses more third person pronouns (2.81 per cent) than the OF corpus.

The relatively high frequency of the SG1 and SG2 pronouns in the OFD corpus deviates from the expected. According to Juva (1998), dialogue written originally in Finnish tends to have more non-standard features than translated dialogue. Nevalainen (2003) has noted the same phenomenon and points out that translations are more conservative in their use of colloquial expressions because translated dialogue is tied to the translator’s notion of “proper” language.

Mauranen and Tiittula’s (2005) study also showed that the SG1 pronoun was often used to create a feeling of colloquial language in the text. On the basis of this, it seems that personal pronouns are used as a kind of colloquialism in dialogue, even if it does not contain any other clear colloquial or dialectal features. The results of this study indicate that similar use does not occur in subtitles, most probably because of the restraints created by the limited time and space.

The comparison with earlier studies emphasises the importance of genre when examining personal pronouns. Most of the earlier studies have used corpora consisting of several genres or academic texts as their material. The material of this study, as outlined in section 5.1, consists of fiction and light reading which were considered as the most suitable genres for a comparison with subtitles. With texts representing another genre the results may be very different.
6.2.6 Use and structures

The results of this study indicate that there are certain structures which are used more in translations than in original Finnish and vice versa. Some features appearing in original Finnish do not appear in translated Finnish at all. Examining the distribution of grammatical cases in the material illustrates the differences to a certain extent.

According to Hakulinen et al (2004: §1227), the most common grammatical cases are nominative, partitive and genitive: added together they form two thirds of all grammatical cases used in texts. In this study, the use of nominative and genitive will be examined more closely. Hakulinen et al (2004: §1228) write that, for personal pronouns, the most common case is nominative with nearly half of the instances. The same applies to the material of this study; the nominative case is the most common one for personal pronouns in all of the corpora.

TABLE 9. Proportion of nominative and genitive personal pronouns of all the instances of personal pronouns in each corpora

<table>
<thead>
<tr>
<th></th>
<th>nominative %</th>
<th>genitive %</th>
</tr>
</thead>
<tbody>
<tr>
<td>OF</td>
<td>58.4</td>
<td>15.9</td>
</tr>
<tr>
<td>TF</td>
<td>56.9</td>
<td>19.5</td>
</tr>
<tr>
<td>SF</td>
<td>41.6</td>
<td>14.5</td>
</tr>
<tr>
<td>OFD</td>
<td>63.1</td>
<td>13.5</td>
</tr>
<tr>
<td>TFD</td>
<td>52.8</td>
<td>18.6</td>
</tr>
<tr>
<td>OFN</td>
<td>55.5</td>
<td>17.4</td>
</tr>
<tr>
<td>TFN</td>
<td>59.1</td>
<td>20.0</td>
</tr>
</tbody>
</table>
Nominative personal pronouns form 41.6 per cent of the personal pronouns in the SF corpus, which is the least in all of the corpora. This seems to support the earlier observation of subtitled Finnish using fewer personal pronouns. The nominative case is the most common case for subject in a sentence, and because the verb form is often enough to indicate the grammatical person, it is possible to omit personal pronouns appearing in the subject position. The small proportion of nominative case indicates that personal pronouns are often omitted if possible.

Nominative personal pronouns are the most frequent in the OFD corpus with 63.1 per cent, and the TFD corpus is situated between the SF and the OFD corpora with 52.8 per cent. The two narrative corpora have a fairly high proportion of nominative as well with 59.1 per cent for the TFN corpus and 55.5 per cent for the OFN corpus, but it is worth noting that, if compared with dialogue, the proportion of nominative is not much larger in narrative even though most personal pronouns in it are third person ones.

In Mauranen and Tiittula’s (2005) study, the nominative SG1 pronoun was used commonly in both original and translated Finnish. Their observations of its use apply to the results of this study as well, including the other speech act pronouns and subtitled Finnish. According to Mauranen and Tiittula (2005), the nominative case was especially common in dialogue, which contained more colloquial or idiomatic expressions than narrative. Example 6 can be considered as an idiomatic expression where omitting the SG1 pronoun is not possible, unless the word order and subsequently the meaning of the sentence is slightly changed. In example 7 the personal pronoun forms an idiomatic expression together with the demonstrative pronoun se (‘it’).
6. En minä sinulta rahaa pyydä
   ‘I’m not asking you for money’

7. Sinä se etsit minut, koska sinusta Paris oli vähän töykeää.
   ‘You are the one who sought me out because you thought that Paris was a bit rude.’

In Mauranen and Tiittula’s (2005) study, the SG1 pronoun in nominative case commonly appeared in structures where it was in some way emphatic, as in the following example:

8. -- ja kun hän tekee mitä sitten tekeeekin. Minä olen täsmälleen samassa jamassa kuin ennenkin.
   ‘-- and when he does whatever he does. I will be in the exact same situation as before.’

Another common type of use was in structures indicating a contrast between the speaker and the listener or some sort of reference group as in example 9, or in places where the speaker wanted to convey a common attitude as in example 10 (Mauranen and Tiittula 2005).

   ‘I’m not a good person. Your father was a good person.’

    ‘Yeah, I know that you shouldn’t read Reader’s Digest if you wish to pass yourself off as educated. I do not read it at the hairdresser’s.’

The function of the SG1 pronoun was the same in both translated and original Finnish, but according to Mauranen and Tiittula (2005), the translations contained more cases in which the function of the pronoun was more difficult to interpret. Example 11 illustrates such use. In the sentence, both of the SG1 pronouns could be omitted, although latter
could be argued to be emphatic. Similar instances occur in all of the corpora but more frequently in the TF corpus.

11. Tosiaan, minä taisin unohtaa kertoa siitä, mutta minä pyysin lankoni Zürichistä paikalle.
‘Indeed, I must have forgotten to tell you about it, but I asked my brother-in-law from Zurich to come here.’

The examples above illustrate possible reasons for the high frequency of nominative personal pronouns in the OFD corpus. Mauranen and Tiittula (2005: 47) also note that, in general, using the SG1 pronoun together with the verb form is more common in spoken language than in written, which is closely tied to the Nevalainen’s (2003) results on dialogue imitating features of spoken language.

The genitive is the second most common case for personal pronouns in the material. In the TF corpus, 19.5 per cent of the personal pronouns are in genitive, whereas the same is true for 15.9 per cent of the OF corpus (table 9). Examining the narrative corpora shows that in the TFN corpus, 20.0 per cent of the personal pronouns are genitive and 17.4 per cent in the OFN corpus. Both of the narrative corpora have a relatively high proportion of genitive case, but it is most likely tied to the use of third person in narrative.

More interesting is the variation of the genitive personal pronouns in dialogue. They are the most common in the TFD corpus with 18.6 per cent. In the OFD corpus they constitute 13.5 per cent of the personal pronouns. The SF corpus is close to the OFD one with 14.5 per cent. In the material of this study, genitive personal pronouns seem to have two main functions. The most common type is the use of personal pronouns as a
genitive modifier as in examples 12-14. In the TFD corpus, the genitive modifiers form 72 per cent of the genitive personal pronouns, and in the OFD one 69 per cent. In the SF corpus, genitive modifiers form 44 per cent of the genitive personal pronouns used.

12. Olen varma, että se kävi sinun mielessä äsken.
   ‘I’m sure that you thought about that just now.’

13. Tiedättekö, että teidän silmänne säihkyvät tässä valaistuksessa?
   ‘Did you know that your eyes shine in this light?’

14. Onko sinulla hauskaa minun siskoni kanssa?
   ‘Are you having fun with my sister?’

In general, the genitive modifier is followed by a possessive suffix. The only exceptions to this can be found in the OFD corpus (example 12, where the standard would be sinun mielessäsi), and it is a feature generally associated with non-standard language use.

The second most common use for genitive personal pronouns in this material is in structures indicating requirement, as in examples 15 and 16. Here as well there is a difference between subtitled Finnish and translated or original Finnish. In the SF corpus, 45 per cent of the genitive personal pronouns occur in structures indicating requirement, whereas the same is true for 22 per cent in the TFD corpus and 24 per cent in the OFD corpus. This suggests that subtitles may use more verbs which require a genitive subject.

15. Minkä takia minun pitää olla näin tyhmä.
   ‘Why do I have to be this stupid.’

   ‘If the doctor thinks that you should stay, then you will stay.’
The frequency and use of genitive personal pronouns seems to vary between the corpora. The results of this study show that the frequency of personal pronouns as genitive modifiers varies greatly between subtitled Finnish and translated or original Finnish. Future studies concentrating on the use of possessive suffix, for example, may be useful in order to examine the reasons behind the variation.

6.2.7 Corpora-specific use

The PL2 pronoun *te* can be used instead of the SG2 pronoun *sinä* to indicate politeness. As mentioned in section 3.1, this use has become less common (Nuolijärvi and Tiittula 2001: 586). There are only a few instances of such use in the corpora, and all of them occur in the TFD corpus (examples 17 and 18).

17. Mutta ettehän te ole ollenkaan veljenne näköinen
   ‘But you do not look like your brother at all’

18. Te taas olette aika lailla sisarenne näköinen
   ‘You on the other hand look quite a lot like your sister’

The choice of using the polite form of address stands out in the translation. The context does indicate that there is a difference in socioeconomic status between the speakers, but in Finnish the polite form of address is not commonly used to indicate this, because the conversation takes place in an informal setting and there is no great age difference between the speakers. The translator’s choice of using the polite form makes the text seem slightly archaic.
As mentioned in section 6.2.4, the use of the SG2 pronoun in the narrative differs from the expected. In general, the SG2 pronoun is used very little in narrative. The OFN contains a few SG2 pronouns, which all occur in a letter addressing its reader directly.

More interesting are the instances of the SG2 pronoun in the TFN corpus. They occur in a passage which is written in a style similar to stream of consciousness. In the passage the speaker is giving orders to herself. Despite this, the use of the SG2 pronoun (as well as the second person verb forms) can be claimed to be generic rather than referring to the speaker.

19. -- sinä nouset, työnnyt ulos, -- silmäsi sumenevat mutta sinä räpyttelet, olet lukevinasi Harvey Nicholsin mainoksia --
   ‘-- you stand up, push outside, -- your eyes blur but you blink, pretending to read Harvey Nichols’ ads --’

According to Auvinen (2005), generic use of the SG2 pronoun is more common in translated texts. The TFN corpus also contains the only instance in all of the corpora where the narrator addresses the reader directly by using the PL2 pronoun:

20. Älkää ikinä uskoko, jos joku väittää teille, ettei votkalla ole hajua; suljetussa komerossa sillä kyllä on.
   ‘Never believe if someone tells you that vodka does not smell; in a closed cupboard it certainly does.’

The above-mentioned use may be related to the author’s style of writing, but the PL2 pronoun could have been omitted from the translation.
6.3 Proper nouns

6.3.1 Undivided material

Proper nouns form the second largest group of nouns with a human referent after personal pronouns (see section 6.1). The amount of personal pronouns, as discussed in section 6.2, varied fairly much between the different corpora. The amount of proper nouns varies as well, but the difference is in several cases less clear. Proper nouns form 4.1 per cent of the TF corpus and 4.8 per cent of the OF. The subtitles use proper nouns the least with 3.2 per cent (table 10). Comparing all three shows that the difference is statistically significant ($\chi^2 p = .000$), and remains so in pairwise comparisons ($\chi^2$ with Yates’ correction $p = .000$).

<table>
<thead>
<tr>
<th></th>
<th>whole</th>
<th>dialogue</th>
<th>narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>translated</td>
<td>4.1</td>
<td>2.3</td>
<td>4.9</td>
</tr>
<tr>
<td>original</td>
<td>4.8</td>
<td>2.4</td>
<td>5.7</td>
</tr>
<tr>
<td>subtitles</td>
<td>3.2</td>
<td>3.2</td>
<td>-</td>
</tr>
</tbody>
</table>

As mentioned earlier, addressing a person by name is not used as much in Finnish as for example in English (Nuolijärvi and Tiittula 2001), which makes the amount of proper nouns in the OF corpus a deviation from this: it could be expected that original Finnish would use proper nouns less than translated Finnish. In the material of this study, there are clearly more proper nouns in the OF corpus than in the TF or SF ones. However, it is important to note that in dialogue the frequency of proper nouns in the OFD is similar to TFD.
6.3.2 Dialogue and narrative

Examining the material divided into dialogue and narrative shows that, unlike personal pronouns which were most common in dialogue, proper nouns are used the most in narrative. In the TFN, 4.9 per cent of the corpus consists of proper nouns, whereas the same is true for 5.7 per cent of the OFN corpus.

The percentage of proper nouns in the OFN is surprisingly high, but it is possible that, due to the relatively small size of the corpora, the high amount is related to a single author’s style of writing. Similar to the undivided material, the difference is statistically significant if comparing the TFN, OFN and SF corpora ($\chi^2 p = .000$), and remains so even in pairwise comparisons between them (TFN-OFN $\chi^2$ with Yates’ correction $p = .001$; TFN-SF $\chi^2$ with Yates’ correction $p = .000$; OFN-SF $\chi^2$ with Yates’ correction $p = .000$).

Proper nouns are used clearly less in dialogue, and form 2.3 per cent of the words used in the TFD corpus and 2.4 per cent of the words in the OFD one. In the SF corpus the percentage is the highest with 3.2 per cent ($\chi^2 p = .003$). The original hypothesis of this study was that subtitles were expected to contain more proper nouns than translated or original Finnish instead of the more ambiguous third person pronouns. The high frequency of proper nouns in the SF corpus supports this.

A comparison between the OFD and SF corpora as well as between the TFD and SF corpora shows that the differences between them are significant ($\chi^2$ with Yates’ correction $p = .014$ and $p = .002$ respectively), but between the OFD and the TFD
corpora insignificant ($\chi^2$ with Yates’ correction $p = .628$). Thus it could be said thatsubtitled Finnish contains proper nouns clearly more than original Finnish or translated
Finnish. Both the OFD and the TFD have similar frequencies, although direct address is
Therefore the use of proper nouns as vocatives is examined more closely in the
following section.

The high proportion of proper nouns in narrative is related to their use. Proper nouns
can be seen as having similarities with the third person pronouns; they both are used to
refer to an individual or individuals outside of speech situation, which makes its use less
common in dialogue. In the case of subtitles, the use of proper nouns may require an
analysis together with the image, because Finnish lacks gender-specific pronouns and
subtitles require clarity. Also, similar to third person pronouns, proper nouns cannot be
omitted.

6.3.3 Vocative expressions

Addressing a person by their name is generally seen as something that varies between
different cultures. The theory section of this study outlined two slightly different
definitions of the so-called vocative expressions (see section 4). According to
Hakulinen et al (2004: §1077, §1651), proper nouns are used to direct the addressee’s
attention to the speaker and to indicate to whom the message is meant. Addressing a
person by name when no reply is expected is rare and Hakulinen et al (2004: §1077)
consider it as special use. The definition given in the Oxford English Grammar
(Greenbaum 1996) was similar, except that proper nouns are also used to maintain some personal connection with the addressee.

On the basis of these definitions, one could expect that the use of proper nouns as vocatives would vary between original and translated Finnish. Examining the material seems to support this; there are a few clear differences in how vocatives are used for addressing participants in dialogue. In all of the three dialogue corpora, the vocatives most commonly appear in the final position in a sentence (examples 22-27), and the number of such instances varied greatly. The proper nouns also occurred in an initial position (example 20) or medial position (example 21) in a sentence, but these instances were less frequent and had no clear differences.

21. Irene, sinun vuorosi laittaa aamiaista
   ‘Irene, your turn to make the breakfast.’

22. Varo, Ville, ne kiipeävät sinun vaatteisiisi
   ‘Watch out, Ville, they are climbing into your clothes.’

Examining vocative expressions occurring in the final position shows clear differences between translated and original Finnish. Instances of proper nouns that were not clearly a part of an utterance are excluded because of the nature of the material: when divided into narrative and dialogue, it is difficult to know whether proper nouns are separate entities or parts of an utterance separated by narrative.

The results indicate that the use of proper nouns as vocative expressions is more common in translated Finnish than in original Finnish. The proper nouns used as vocative expressions at the end of sentences form 21.6 per cent of the instances of
proper nouns in the TFD corpus whereas the same is true for 7.6 per cent of the OFD corpus. In the SF corpus such structures are used clearly the least with 3.9 per cent.

23. Hetkinen, Kirk. Seiso tässä ja katso, ettei se tule ulos. ‘A moment, Kirk. Stand here and watch that it doesn’t come out.’

24. Tule toistekin, Mimi. ‘Come again, Mimi.’

25. Siellä on juuri nyt täysi sekasotku, Art. Meillä tehdään inventaariota. ‘It’s a complete mess in there right now, Art. We’re taking inventory.’

26. Syöpäs nyt murosi, Allison. Meidän täytyy lähteä ihan pian. ‘Eat your cereal, Allison. We must leave very soon.’

27. Teet mikä on tehtäväksesi annettu. Eikös se niin ole, Asko? ‘You will do what you were meant to. Isn’t that right, Asko?’

28. Minun ollut ikävä sinua, Emma. ‘I have missed you, Emma.’

The examples above illustrate the most common types. Vocative expressions most commonly appear in short questions (example 26), orders or requests (examples 22, 23, and 25) and greetings. Instances such as these form approximately one half of all the instances of vocative expressions occurring in the final position. Example 24 is one of the several instances in the TFD corpus where using a proper noun does not seem to have any real need. Such vocative expressions are equally common in all three corpora, but they stand out more in the TFD corpus because they occur in it almost thrice as frequently as in the other corpora. Example 27 shows how proper nouns are used in sentences expressing emotion towards the listener.

It should be noted that the higher number of instances in the TFD may be partially explained by the small size of corpora, which is why vocatives may be worth more
study in the future. It is also probable that the use of proper nouns when the speaker addresses others is not something that a translator would pay special attention to. It is likely that a translator will not omit a name from the translation, even though it then stands out in the translation if compared with texts originally written in Finnish.

6.3.4 Grammatical case

Examining the distribution of grammatical cases shows that their frequency in the proper nouns differs from that of personal pronouns (see section 6.2.4). In personal pronouns, the distribution varied between the corpora without a clear pattern, whereas with proper nouns the distribution is more closely tied to whether the corpus they occur in consists of translated or original Finnish.

<table>
<thead>
<tr>
<th></th>
<th>nominative %</th>
<th>genitive %</th>
</tr>
</thead>
<tbody>
<tr>
<td>OF</td>
<td>66.5</td>
<td>24.1</td>
</tr>
<tr>
<td>TF</td>
<td>75.7</td>
<td>15.6</td>
</tr>
<tr>
<td>SF</td>
<td>54.8</td>
<td>27.2</td>
</tr>
<tr>
<td>OFD</td>
<td>68.2</td>
<td>21.8</td>
</tr>
<tr>
<td>TFD</td>
<td>74.7</td>
<td>15.5</td>
</tr>
<tr>
<td>OFN</td>
<td>66.2</td>
<td>24.4</td>
</tr>
<tr>
<td>TFN</td>
<td>75.9</td>
<td>15.7</td>
</tr>
</tbody>
</table>

Examining the instances shows that nominative proper nouns are the most common ones in the material. The TF corpus has relatively more nominative proper nouns than the OF or SF corpora with 75.7 per cent to 66.5 and 54.8 per cent respectively.
Separating the material into dialogue and narrative has little effect on the values and causes only minor variation (table 11).

The genitive is the second most common case for proper nouns. In the TF corpus it constitutes 15.6 per cent of proper nouns and in the OF corpus 24.1 per cent. Dividing the material into narrative and dialogue does not cause significant variation in the portion of genitive proper nouns either. In the SF corpus, they form 27.2 per cent of all the proper nouns in the corpus.

Unlike with personal pronouns, there is no clear variation between the corpora in the frequency of proper nouns as genitive modifiers. They occur slightly less in the OFD corpus (73 per cent) than in the SF corpus (81 per cent) or the TFD corpus (77 per cent), and is slightly more common in narrative with 81 per cent for the TFN corpus and 84 per cent for the OFN corpus. Unlike genitive personal pronouns, genitive proper nouns do not occur almost at all in structures indicating requirement, and therefore the high portion of proper nouns in the OF and SF is not explained by verbs which require a genitive subject.

The results indicate that the frequency of grammatical cases may be different in subtitled Finnish. Examining the results shows the OF corpus uses genitive relatively more and nominative relatively less than the TF corpus. If the distribution of grammatical cases in the OF corpus is taken as the norm, their distribution in the SF corpus seems hypercorrect, because it has nominative relatively less and genitive relatively more than the OF corpus and (table 11).
6.3.5 Corpus-specific structures

Examining the sentences in which proper nouns occur shows that there are features that appear only in certain corpora. Some of the differences found in the material are culture-related rather than unidiomatic structures caused by translating. One such feature is related to the use of honorific in the text. In all of the corpora there are instances where a character is addressed or referred to by using an honorific (see example 28-33) and the surname. However, there are differences in the types of titles that are used. In the TF corpus, the most common titles used are titles such as herra and rouva (examples 28 and 29):

29. Hyvä huomenta, herra Partridge
‘Good morning, Mr Partridge.’

‘It was Jonathan Sellers from Mrs Edelman’s English class.’

Examining the use of proper nouns in the OF and SF corpora shows that they do not follow the above-mentioned trend. There are hardly any instances similar to those occurring in the TF corpus. Instead, if a character is referred to or addressed by title and name in the OF and SF corpora, the titles are occupation-related ones as in examples 30-33.

31. -- talousjohtaja Aava arveli ja aikoi sulkea vastaanottimen.
‘-- CFO Aava thought and was about to close the television.’

32. Terveisiä toveri Janajeville ulkoministeri Veijo Salmelta
‘Greetings to comrade Janajev from foreign minister Veijo Salmi’

33. Tunnetko republikaaniedustaja Doug Osen Kaliforniasta?
‘Do you know Republican Congressman Doug Ose from California?’
The results suggest that using a title and surname to address an individual is not a feature that commonly appears in texts imitating spoken Finnish. It is also worth noting that no names in the material have been translated or domesticated, which may indicate that names are a part of creating the setting in a literary text. The use of honorific is something that is likely to be affected by each author’s individual style, and also by the subject matter of the texts: works concentrating on interaction between family members would have no need for using honorific, for example.

As mentioned earlier, the genitive is the second most common case for proper nouns. Similar to personal pronouns, they are often used as genitive modifiers. In addition to indicating ownership, genitive constructions are used to indicate family relations or belonging to a group, as in the following examples:

35. Meitä vastapäätä Billin äiti paineli silmiään nenäliinalla ja nyyhkytti rajusti.
   ‘Across us Bill’s mother dabbed her eyes with a handkerchief and sobbed violently.’

36. Arttu-veljen vaimo Pipsa oli pukeutunut nuorekkaaseen, ihonmyötäiseen trikoomekkoon.
   ‘[My] brother Arttu’s wife Pipsa was dressed in a youthful skin-tight tricot dress.’

37. Irvistelisi muuten taas Ella-tädille tai repisi Kähkösten pojalta tukan pästä irti.
   ‘Otherwise [she] would again make faces at Aunt Ella or tear the hair off the head of Kähkönen’s son.’

38. Näyttääköhän Christopherin ja Sherryn vauva minulta?
   ‘I wonder if Christopher and Sherry’s baby will look like me?’
In the OF corpus there are instances of genitive structures where a genitive form of a surname is used before the first name, as in example 38 and 39. Similar use does not appear in the TF corpus or the SF corpus, suggesting that the feature is indigenous to Finnish.

   ‘At the other end of the line was probably master Rantala. Or Lähde’s Eetu.’

   ‘That’s what I thought. Mäkelä’s Pekka. I lived next door to you as a kid.’

Another feature, which only appears in the OFD corpus, is the use of demonstrative pronouns together with proper nouns, as in examples 40 and 41.

41. Vieläkö se Sinikka on lypsyllä?
   ‘Is that Sinikka still milking the cows?’

42. Et kai häirinnyt sitä Siltalan tyttöä?
   ‘You didn’t bother that Siltala’s girl, did you?’

In the examples, the use of the demonstrative pronoun indicates that the person mentioned has been the subject of conversation before or the conversation refers to an earlier exchange between the participants (Hakulinen et al 2004: §1414).

In general, dialogue seems to prefer shorter noun phrases than narrative, which is especially seen in the subtitles. Subtitled Finnish preferring shorter noun phrases follows the results of Hirvonen’s (1992) study on the condensation in subtitles. The largest difference in the length of noun phrases is between narrative and dialogue.
7 CONCLUSION

The results of this study are not definitive, and there is a need for more research with larger and perhaps more varied corpora than the ones used in this study. Due to the small size of especially the dialogue and subtitle corpora, the results may be distorted because one text may contain more of certain features than the others because of each author’s individual style of writing. It could also be argued that, unlike in this study, the subtitles should be studied together with the image on screen, because as Vertanen (2007) emphasised, they are not complete by themselves.

It is also important to note that the genre of the television series from which the subtitles are taken from is likely to have an effect on the results. Examining the subtitles of a documentary, for example, would most likely yield different results. *Gilmore Girls* is a mixture of drama and comedy and very dialogue-oriented, which is likely to affect the use of proper nouns and personal pronouns. Genre will quite likely affect the use of non-standard language as well. Television shows or movies centring on subcultures or rural communities may have subtitles that contain more non-standard features, because the speech of the characters is more likely to contain more stimuli for the translator to use non-standard features (Nevalainen 2003).
The results of this study indicated that the genre affects the use of personal pronouns in texts as well. The comparison made between the results of this study and those gained in earlier ones on the use of the SG1 and SG2 pronouns showed that the personal pronouns were more common in the material of this study, which were compiled from the fiction and light reading subcorpora of the *Corpus of Translated Finnish*.

For the above-mentioned reasons, examining and comparing subtitles and other texts representing different genres may be worth future study. Another possible area of future research may lie in comparing the source text to the target text, or in the case of subtitles, comparing the original dialogue with the subtitles. Especially examining the difference in the use of proper nouns in texts from different cultures may be an interesting area of research.

Although the results of this should be considered as tentative ones, they nevertheless show that subtitled Finnish differs from both original and translated Finnish: there are differences in how much personal pronouns and proper nouns are used, and also in the structures they appear in.

Examining a small-scale sample of the material used in this study showed that the amount of human reference is fairly constant. This was supported by examining the percentage of proper nouns and personal pronouns in the whole material: added together, proper nouns and personal pronouns formed approximately ten per cent of the words used in the corpora. The amount of human referents was slightly more varied, and therefore examining all human referents in a larger sample than the one examined in this study may yield different results.
The material used as point of comparison was divided into dialogue and narrative in order to offer a better point of comparison for the subtitles. Dividing the material showed that the frequency of personal pronouns in dialogue and in narrative is different, and the division also affected the use of proper nouns. The TF corpus uses personal pronouns and proper nouns the most, the OF corpus the least and the SF corpus is situated between the two.

Examining the instances of personal pronouns and proper nouns separately indicates that, as Jääskeläinen (2007) suggested, translated Finnish differs from original Finnish, and subtitled Finnish differs from both of them. This difference is best illustrated by examining dialogue. The amount of personal pronouns is smaller in the SF corpus than in the other dialogue corpora, thus supporting the original hypothesis of subtitled Finnish omitting personal pronouns if possible. Subtitled Finnish also contains more proper nouns than original or translated Finnish.

Both translated and original Finnish dialogue consist mostly of first and second person pronouns with only a small portion of third person pronouns. Subtitled Finnish does not have such a clear difference, and although first and second person are more common in it, the portion of third person pronouns is large in comparison with the other dialogue corpora. Third person pronouns cannot be omitted as easily as first or second person ones, which increases their proportion. The higher frequency of proper nouns in the SF corpus also indicates that third person pronouns may be replaced with proper nouns in subtitled Finnish in order to avoid ambiguity caused by the lack of gender-marked third person pronouns.
Comparing the results of earlier studies on the SG1 and SG2 pronouns to the results of this study emphasise a few important differences. In general, the results of this study are quite different or even opposite which, as mentioned earlier, indicates that genre as well as the division to narrative and dialogue has a strong influence on the use of personal pronouns.

According to Mauranen and Tiittula (2005), the SG1 pronoun occurs less frequently in original Finnish. Their observation is true in the material of this study when examining undivided material. If examining dialogue only, the frequency of the SG1 pronoun is higher in original Finnish than in translated Finnish. The same applies to the SG2 pronoun as well. Subtitled Finnish contains less SG1 and SG2 pronouns than other dialogue.

The results indicate that personal pronouns are used to create the feeling of colloquial language in texts. The use of nominative speech act pronouns follows Mauranen and Tiittula’s (2005) results for the use of the nominative SG1 pronoun. Nominative first and second person pronouns were commonly used to imitate colloquial language, and occurred in idiomatic expressions, for example. All of the corpora contained instances in which the function of the personal pronoun was not clear, but where they could be considered as imitating colloquial language. This is further supported by the high frequency of nominative personal pronouns in the OFD corpus, because the nominative is a common case for subject. It is important to note that there are hardly any clear markers of dialectal language in the material of this study, and certain common non-standard features do not appear either or occur only in certain corpora.
How the feeling of colloquial language is achieved in subtitles is not truly answered in this study, although it was one of the questions that this study set out to answer. The results indicate that, unlike in the original Finnish dialogue, the feeling of colloquial language in subtitled Finnish is not achieved with the use of personal pronouns. Although there are some colloquial word choices, the subtitles used as the material of this study contain no clear dialectal features. The low frequency of personal pronouns in subtitles indicates that they are often omitted, if possible. This is also supported by the low frequency of nominative personal pronouns in the SF corpus.

Proper nouns were used the most in original Finnish and the least in subtitled Finnish when examining the undivided material. In dialogue, original and translated Finnish contained a similar amount of proper nouns, which deviated from the expected, because addressing people by name when no reply is expected is considered as a special use in Finnish (Hakulinen et al 2004: §1077). Examining the use of proper nouns as vocatives showed that such use is almost thrice as common in the translated Finnish dialogue than in original Finnish, and the least common in subtitled Finnish. In most cases the direct address served a specific purpose, but there were also instances where the function was not clear. These seemed especially common in the translated Finnish dialogue, because it contained vocatives the most. The results indicate that there may be differences in how proper nouns are used in different languages: although the amount of proper nouns is similar in translated and original Finnish, direct address is less common in original Finnish. The high frequency of proper nouns in subtitled Finnish indicates that proper nouns may be used to avoid the ambiguous third person pronouns. It is also likely that the high portion of proper nouns is related to the fact that the subtitles examined in
study are from a dialogue-based television drama. However, this area would require more research.

This study also examined the two most common grammatical cases in personal pronouns and proper nouns, which again indicated that translated, original and subtitled Finnish use them differently. The use of nominative personal pronouns has already been discussed above. In proper nouns, the nominative case forms a clearly larger portion of the proper nouns used than in the personal pronouns, and its amount varies less. The genitive case formed a larger portion of the cases used in personal pronouns in translated Finnish than in subtitled or original Finnish, which gave reason to examine the use of personal pronouns as genitive modifiers. Examining them showed that personal pronouns as genitive modifiers are clearly the least common in subtitled Finnish and that their amount in translated and original Finnish dialogue is similar. However, in the OFD corpus genitive modifiers occurred without possessive suffix, which is a feature colloquial language. In the personal pronouns, another common use of genitive was in structures indicating requirement (minun pitää mennä ‘I have to go’). Such structures were common in subtitles but used less in translated Finnish. In proper nouns such use was rare, and their use as genitive modifiers did not vary as much as is the personal pronouns.

Differences between subtitled, original and translated Finnish can also be seen by examining features occurring only in certain corpora. The results indicate that these features or the lack of them may mark a text as a translation. For example, the polite use of the PL2 pronoun te instead of the more informal SG pronoun sinä in a situation where there is no clear need for it made the text seem archaic. Another such marker was
related to a possible cultural difference: when using a title together with a name, translated Finnish had more instances of titles such as *herra* (‘Mr’) whereas both original and translated Finnish contained more occupation-related titles (*talousjohtaja* Aava ‘CFO Aava’). In the case of subtitled Finnish, it is likely the use is related to the greater need for clarity in subtitles.

As was mentioned in the introduction, the special characteristics of subtitles gave reason to assume that they may differ from other translations and original Finnish. This study supports that hypothesis. The results show that the use of personal pronouns and proper names in subtitled Finnish is situated between translated Finnish and original Finnish, but in several cases subtitled Finnish seems closer to original Finnish than to translated Finnish.
RESEARCH MATERIAL

*Corpus of Translated Finnish* (Käännössuomen korpus). Digital research material of translated Finnish. Compiled in Translated Finnish and Translation Universals – project at The Savonlinna School of Translation Studies in the University of Joensuu. Savonlinna, 1997–.


WORKS CITED


APPENDIX

Abbreviations used:

OF: Original Finnish
PL1: plural 1. person, me
PL2: plural 2. person, te
PL3: plural 3. person, he
SG1: singular 1. person, minä
SG2: singular 2. person, sinä
SG3: singular 3. person, hän
SF: Subtitled Finnish
TF: Translated Finnish
FINNISH ABSTRACT


Suomen kirjakielessä peroonapronominit jäävät usein pois, koska yleensä lauseen subjektin voi päätellä verbin taivutusmuodosta. Runsas peroonapronominen käyttö liitetään yleensä puhekieleen. Ruututekstien piirteisiin kuuluu pyrkimys huoliteltuun yleiskieleen, mutta ne eroavat silti erityispiirteiltään muista käänöksistä, sillä käänä tunnovalintoja rajoittavat sekä käytettävissä oleva tila että aika. Joissakin tapauksissa puhekielisykyttä ei kuitenkaan voi jättää pois ilman, että elokuvan tai sarjan tyyli tai tunnelma muuttuu, esimerkiksi erilaisissa amerikkalaisia alakulttuureja kuvaavissa elokuvissa.
Tutkielman hypoteesi on, että ruututekstien ominaisrajoitteet saavat tekstittäjät jättämään tarpeettomat personapronominit pois. Erisnimien käyttö puolestaan eroaa kulttuurien välillä, ja suora puhuttelu on suomessa melko harvinaista. Tämän perusteella voi siis olettaa, että erisnimiä käytetään supisuomessa vähemmän kuin käännössuomessa. Ruututekstien ominaispiirteet antavat syyn olettaa, että erisnimiä suosittaisiin ruututeksteissä monitulkintaisten kolmannen persoonan pronominien sijaan, jolloin erisnimien yleisyys on suurempi kuin muissa käännöksissä tai supisuomessa.


Koska materiaalina käytetyn ruututekstin koostuvat pelkästään dialogista, myös muu materiaali on jaettu dialogiin ja kerrontaan. Vain suora esitys lasketaan dialogiksi. Persoonamuotoiset verbit sekä murremuotoiset personapronominit on jätetty tutkimuksen ulkopuolelle. Erisnimistä on rajattu pois erisnimen lailla käytetyn substantiivin (esimerkiksi äiti). On olennaista huomioida, että käännössuomea ja
ruututekstejä käsitellään tässä tutkielmassa erillisinä, vaikka myös ruututekstit ovat käänössuomessa.


Persoonapronominen muodostavat suuriman osan henkilöihin viittaavista ilmauksista. 

Ruututekstit ja muu käännössuomi sisältävät lähes yhtä suuren määrän persoonapronomineja, ja ne ovat kyseisissä korpuksisissa selvästi yleisempiä kuin supisuomessa. Dialogi sisältää selvästi enemmän persoonapronomineja kuin kerronta. 

Tarkasteltaessa pelkkää dialogia persoonapronominen ovat yleisempiä käännössuomessa ja supisuomessa kuin ruututeksteissä, mikä tukee alkuperäistä hypoteesia.


Tutkielman tulosten vertailu persoonapronominen aiemmin tehtyihin tutkimuksiin osoittaa, että tämän tutkielman materiaalissa persoonapronominen ovat yleisempiä. Maurasen ja Tiittulan (2005) mukaan käännössuomen ja supisuomen voi erottaa toisistaan persoonapronominen käytön perusteella, sillä heidän mukaansa käännössuomessa käytetään selvästi enemmän yksikön ensimmäisen ja toisen persoonan pronominen. Tässä tutkielmassa tulos on osin päinvastainen. Koko
materiaalissa yksikön 1. persoonan ja 2. persoonan pronominit ovat yleisempiä käännössuomessa, mutta ero ei ole niin suuri kuin aiemmat tutkimukset antaisivat olettaa. Dialogissa tilanne on pääinvastainen, sillä kyseisiä persoonapronomineja käytetään eniten supisuomen dialogissa ja ruututeksteissä vähiten. Vertailu aiempien tutkimusten kanssa nostaa esiin genren mahdollisen vaikutuksen persoonapronominien yleisyyteen, sekä sen, miten selvästi persoonapronominien käyttö dialogissa eroaa kerronnasta.

Supisuomessa yksikön ensimmäisen ja toisen persoonan pronominin käyttö on yllättävän yleistä, mikä liittyy todennäköisesti puhekielisyyden luomiseen. Ruututeksteissä persoonapronomineja ei käytetä samalla tavalla puhekielisyyden vaikutelman luomiseen, ja niissä on muutenkin vain hyvin vähäinen määrä yleiskielestä poikkeavia ilmauksia.


Kun tarkastellaan pelkästään erisnimien käyttöä puhuttelussa, käy ilmi, että nimeltä puhuttelu on lähes kolme kertaa yleisempää käännössuomen dialogissa kuin
supisuomen. Ruututeksteissä kyseinen käyttö on vähäisintä. Valtaosalla puhutteluista on jokin tietty tarkoitus, mutta on myös tapauksia, joissa puhuttelulla ei ole selkeää funktiota. Tällaiset tapaukset erotuvat käännössuomen dialogista, koska nimeltä puhuttelu on siinä paljon yleisempää kuin muissa korpuksissa.


Genetiivimuotoiset persoonapronominnit ovat suhteessa yleisimpiä käännössuomessa, ja vähiten niitä käytetään ruututeksteissä. Dialogia tarkasteltaessa tilanne säilyy muuten samana, mutta genetiivin osuus supisuomen dialogin persoonapronomineista on hieman ruututekstejä vähäisempi. Koska possessiivisuffixi riittää usein ilmaisemaan omistajaa, persoonapronominien käyttöä genetiivimääritteenä tutkittiin. Niiden käyttö on selvästi
vähäisintä ruututeksteissä, kun taas supisuomen ja käännössuomen dialogissa genetiivimääritteellä yleisyyys on lähes sama. Toinen yleinen genetiivin käyttötapa on pakkoa ilmaseissa rakenteissa (minun pitää mennä), jotka ovat yleisiä ruututeksteissä, mutta vähemmän käytettyjä muussa käännössuomessa.

Genetiivimuotoiset erisnimet ovat yleisiä ruututeksteissä ja supisuomessa ja niitä käytetään vähiten käännössuomessa. Erisnimien käyttössä genetiivimääritteellä ei ole samalla tavalla eroja kuin persoonapronomineissa. Niitä ei myöskään käytetä pakkoa ilmaseissa rakenteissa kuten persoonapronomineja.

Materiaalin analyysin yhteydessä nousi esiin rakenteita, jotka esiintyivät vain tietyissä korpuseissa. Yleisesti ottaen korpusten kieli seuraa huoliteltua yleiskieltä, mutta poikkeuksena tästä on muun muassa genetiivirakenteen käyttö ilman possessiivisuffiksia, mitä esiintyy pelkästään supisuomen dialogissa. Käännössuomessa käytetään muutamassa kohdassa te-pronominia kohteliaana puhutteluna, mikä vaikuttaa arkaaiselta.

Eräs mahdollisesti kulttuurieroihin liittyvä seikka on erisnimen ja tittelin käyttö henkilöihin viitattaessa, mikä poikkeaa supisuomessa ja käännössuomessa. Käännössuomessa käytetään tittelit herra tai rouva, kun taas supisuomessa suositaan esimerkiksi työhön liittyviä nimikkeitä (talousjohtaja Aava). Ruututekstit seuraavat supisuomen käytäntöä. Lisäksi materiaalissa on muutamia suomen kielelle ominaisia rakenteita, joita ei ollut käännöksissä lainkaan, esimerkiksi demonstratiivipronominin käyttö nimen yhteydessä (se Sinikka).