Loss, Gain, and Chromosomes

Readability and Translation Shifts in Medical Information for Families of Children with 10q25/10q26 Deletions

Author: Kristina Runyeon-Odeberg
Supervisor: Fredrik Heinat
Examiner: Jukka Tyrkkö
Academic Term: Spring 2019
Subject: English
Level: Advanced
Course Code: 4EN31E
Abstract
The purpose of this thesis for the degree of Master, one year, is to investigate the translation of medical information from British English into Swedish. The analysis is concerned with readability and terminology. In specific, the areas of investigation are how readability of the material compares to other studies, what translation strategies or sources a translator may use, and what kind of translation shifts the terms analyzed undergo.

The source text material consists of some six thousand words from a leaflet called *10q25 and 10q26 Deletions* from Unique, a charity organization based in Great Britain, which welcomes families located worldwide as members. The theory in this thesis is based on previous research within the focus areas of readability, terminology, translation strategies, and translation shifts. Examples include Frege (1948 [1892]), Ogden and Richards (1923), Flesch (1948), McLaughlin (1969), Vinay and Darbelnet (1995), Deléger et al (2010), Kolahi and Shirvani (2012), and Acar and İşisığa (2017).

The text material does not fulfill the professional recommendations regarding readability, which confirms observations from previous studies. Overall, there is a tendency towards an improved readability level in the translation, which contradicts one bilingual study (Kolahi and Shirvani) but partially confirms another (Acar and İşisığa). Of the terms identified for analysis, 49.4 % have been found to require translation strategies sorting under oblique translation, where the predominant method is transposition. A significantly higher number of terms have been found in term banks and corpora than indicated in a previous study.

Key Words
medical information, medical terminology, readability, translation strategies, translation shifts

Acknowledgments
Heartfelt thanks to:
Elisabeth Jansson Ek, Vanja Vinter, and Sara Torres Löndahl, who have been excellent peer reviewers, and whose constructive comments have been very helpful.

Claire Andersen at Unique, who has patiently answered questions on topics including, but not limited to, the leaflet that forms the ST of this thesis.

Cecilie Fonnesbech, text supervisor at Linnaeus University Library, without whose encouragement this thesis would not have been completed.
Paul McShane at Aktiv Ortopedteknik Malmö, who shed some well-needed light on orthopedic terms that were hard to find.

My wonderful family, which consists of my husband Hans and our sons Samuel and Markus; you are the wind beneath my wings. I love you all to the moon... and back. —Wifey/Mom
Table of Contents

1 Introduction .......................................................................................................................... 1
  1.1 Aim and Scope ................................................................................................................ 2

2 Material and Method .......................................................................................................... 3
  2.1 Procedure for Translation ............................................................................................... 3
    2.1.1 Readability ................................................................................................................ 4
    2.1.2 Terminology ............................................................................................................... 4
    2.1.3 Translation Strategies ............................................................................................... 4
  2.2 Procedure for Analysis .................................................................................................... 5
    2.2.1 Readability ................................................................................................................ 6
    2.2.2 Terminology ............................................................................................................... 6

3 Theoretical Background ..................................................................................................... 7
  3.1 Readability ...................................................................................................................... 8
    3.1.1 Previous Research ..................................................................................................... 8
    3.1.2 Readability Tests ...................................................................................................... 11
  3.2 Terminology .................................................................................................................. 13
    3.2.1 Previous Research ................................................................................................... 15
  3.3 Translation Strategies: Sourcing ..................................................................................... 16
    3.3.1 Previous Research ................................................................................................... 17
  3.4 Translation Strategies: Shifts .......................................................................................... 18
    3.4.1 Previous Research ................................................................................................... 20

4 Analysis ............................................................................................................................... 21
  4.1 Readability ...................................................................................................................... 21
    4.1.1 Text Properties and Test Results ............................................................................. 21
    4.1.2 Flesch Reading Ease vs SMOG ............................................................................... 24
  4.2 Terminology .................................................................................................................. 26
    4.2.1 Sourcing .................................................................................................................... 27
    4.2.2 Distribution ............................................................................................................... 28
  4.3 Translation Strategies ..................................................................................................... 29
  4.4 Examples of Shifts .......................................................................................................... 31
    4.4.1 Compound Nouns ..................................................................................................... 32
    4.4.2 Modified Nouns ........................................................................................................ 34
    4.4.3 Simple Nouns ........................................................................................................... 37
    4.4.4 Potential False Friends ............................................................................................ 38
    4.4.5 Adaptation ................................................................................................................. 38
  4.5 Summary of Findings ...................................................................................................... 39
5 Conclusion.................................................................40
  5.1 Readability..........................................................40
  5.2 Terminology.......................................................41
  5.3 Translation Strategies.........................................41
  5.4 Further Work.....................................................41
Sources...........................................................................43
  Primary......................................................................43
  Secondary...................................................................43
  Online Dictionaries/Term Banks................................49
  Parallel texts and Corpora.........................................50

Appendices

Appendix A: Investigation of the Source Text..........................A1
  Background Information...........................................A1
  Context......................................................................A1
  Source Text Analysis................................................A2
1 Introduction

On the subject of translation, Tytler (1999 [1791]: 1) wrote: “There is perhaps no department of literature which has been less the object of cultivation than the Art of Translating.” The same author has written the following three laws of translation (ibid., 14):

“I. That the Translation should give a complete transcript of the ideas of the original work.

II. That the style and manner of writing should be of the same character with that of the original.

III. That the Translation should have all the ease of original composition.”

These three laws are relevant to medical information to this day, as there are guidelines regarding readability (related to laws II and III), cultural aspects (related to law I) and who should be trusted with translating medical information (also related to law I). The International Medical Interpreters Association (IMIA), a professional organization for interpreters in speaking and writing, defines professional standards in this particular area (IMIA, 2011). Part of the organization’s work encompasses having issued guidelines that discuss the profile of a medical translator. Txabarriaga (2009: 3) writes: “To [have a medical text translated by anyone but a professional medical translator] is to risk releasing erroneous information, with all its implications, legal and otherwise.” These implications may well be serious: a patient or guardian misunderstanding the information may lead to problems such as less than optimal care (Cooke et al., 2000: 119) or even adverse effects (Dalton-Oates, 2017: 234–235). Since medical information needs to be correct, to the point and easy to understand, the purpose of this thesis is to investigate readability and terminology issues related to the translation of medical writing.

The text material for this thesis consists of medical information directed at laymen in the form of a leaflet, 10q25/26 Deletions (Unique, 2013) for parents of affected children. To date, the British English version, which forms the source text (henceforth: ST) of this thesis has been handed to parents in Sweden when their children have been diagnosed with the deletion. The speaking and reading proficiency of English among many Swedish parents might be good to excellent; however, the situation may be very different when they are in a state of stress. The material for this thesis is the first translation from English to Swedish of a Unique disorder guide. It is one of around 200 guides written by communicators; some of them happen to be parents of children with chromosome disorders, and some of them have also acquired some sort of medical or biological expertise. The medical information on these disorders that was offered to families before the guides were produced

---

1This expression is used throughout and refers to written information.
was either non-specific or vague (Andersen, 2019). This has been addressed in the guides by including summaries of, and references to, published medical literature alongside results from surveys and testimonials from parents in the form of direct quotations. The result is a guide that tells parents how life can be with a 10q25/26 deletion, as it has information on long-term effects, not only early childhood, as the case tends to be with medical literature (Andersen, 2019).

Since the guide content is directed at laymen, medical terms will need explanation in the target text (henceforth: TT) as well as in the ST. This means mixing different kinds of medical language, which can be confusing at first sight, but parents/readers may meet professionals who insist on using the preferred medical term over a layman’s expression. Parents may also want to research the information further, in which case it is almost certain that the preferred term is one of the references used, as in (1).

(1) A squint (strabismus), where one or both eyes can turn inwards, outwards or upwards, is the most common vision problem noted by researchers and by almost all Unique families. (ST, 350) En skelnning (strabismus), där ett öga eller båda ögon kan vara vända inåt, utåt eller uppåt, är det vanligaste synproblemet som dokumenterats av forskare och nästan alla Unique-familjer. (TT, 350)

In this particular example, the first strategy that comes to mind is to use the layman term of squint, but a reader who is native in Swedish may well think squint in (1) means kisa, ‘to peer with partially closed eyes’, rather than skela, ‘to be cross-eyed’ (Merriam-Webster, 2019; Linguee, 2019; Karolinska Institutet, 2019). Hence, the medical term and its explanation are both present. The text contains a mix of scientific medical language and generic medical language, concepts that are further defined at the beginning of Chapter 3.

1.1 Aim and Scope
The purpose of this thesis is to analyze translation strategies used in medical information. As there are professional directives regarding terminology (Txabarriaga, 2009: 8) and readability (ibid., 6) in medical writing, both areas are suited for investigation on a general level. The specific aim is to find answers to the following research questions:

- How does the readability level compare to that of similar medical information, what methods have been used previously for investigating readability, and how do they perform?
- What kind of translation strategies are used, and where may a translator locate suitable terms?
- What kind of translation shifts may be observed among the terms included in the analysis?
2 Material and Method

The material for this thesis consists of approximately six thousand words from a leaflet about the medical and developmental implications of living with a 10q25/26 deletion. For more information on the source text, please refer to page A1 in Appendix A. In this section, the method of translation and analysis of the ST and TT are discussed.

2.1 Procedure for Translation

Anyone who translates a text needs to be aware of its nature. When it comes to scientific information in general and medical information in specific, a translator either needs personal expertise or know where to find it (Chesterman and Wagner, 2014 [2002]: 70). Sometimes, the expert is the author of the ST, sometimes, it is a terminologist (Kletke and Rydell, 2016: 97). On this note, O’Neill (2016: 69) writes that “all medical professionals learn the language of medicine during their education and training” (ibid.). However, “writing is not usually part of the medical curriculum” (ibid., 77). She has conducted a survey (ibid., 71), which showed that medical translators use the following strategies to gain expertise (ibid., 72):

- study medicine, or take medical courses
- work in a profession related to medicine (hospital secretaries and pharmacists included)
- take a translation course with emphasis on medicine
- ask a close relative, or a friend/acquaintance, who is a medical professional
- research a personal illness and be in contact with medical professionals for that reason

The survey has shown that due to possible knowledge gaps, a bureau professional (O’Neill, 2016: 75) favors the team approach, and a medical professional agrees (ibid.). Txabarriaga (2009: 5), on the other hand, writes that only a professional translator should undertake the task of translating medical information (ibid.). Quality assurance should be an ongoing process (ibid., 7). On another note, Nisbeth Jensen and Korning Zethsen (2013: 37) write about medical professionals being keen to adhering to expert language to the point of translating word for word, but they also stress the point of lay-friendliness (ibid., 32). A worst-case scenario would be that no one is qualified to translate medical information. However, Pietrzak (2015: 317) writes: “Medical [information] is a specific genre where terminological problems are surmountable provided that translators find the term they need and have some conceptual knowledge about it”. Furthermore, in the case of the text material for this thesis, a charity organization is very likely to welcome help with translation on a voluntary basis.
First, the source text was analyzed according to a text model of Hellspong and Ledin (1997: 49–). This analysis is available in Appendix A on page A2. Then, the ST was translated using Trados Studio 2019 and SDL Multiterm 2019; the latter helped keep track of the potential terms intended for analysis as they were encountered.

2.1.1 Readability
Special care was taken to use as simple words as possible during the translation to preserve the readability level between the ST and TT according to the directives of the IMIA (Txabarriaga, 2009: 6), while preserving the message and and conveying the ideas as well (Tytler, 1999 [1791]: 14). The tone of voice was an additional issue to take into consideration and observe, since Vinay and Darbelnet (1995: 17) argue that “[i]f at all possible, translators must preserve the tone of the text they translate” (ibid.).

2.1.2 Terminology
Concerning terms, one of the subject matter experts of the ST (Unique, 2013) is located at Karolinska Institutet. Therefore, the Swedish version of Medical Subject Headings, Svensk MeSH (Karolinska Institutet, 2019; henceforth: MeSH), was one of the choice locations to search for translations of medical terms encountered. Another choice term location was SNOMED CT. SNOMED stands for Systematized Nomenclature of Medicine; it is the recommended term source of diagnoses according to the Board of Directors of the American Medical Informatics Association (AMIA) (Awaysheh et al., 2018: 17). Some terms were related to growth and development; neither MeSH nor SNOMED CT contains any such information. An online encyclopedia on the subject of psychology (Egidius, 2019) proved to be a resource in this area. Two special dictionaries/reference books on medical terms (Lindskog, 2016; Malmquist and Lundh, 2016) had information on etymology and concepts, which also proved useful.

2.1.3 Translation Strategies
Any translator needs to consider what strategies to follow during the translation process, and terms are no exception. The use of already existent terms in the target language is the first choice for a translator (Kletke and Rydell, 2016: 97), which is mainly what has been done in this thesis. Kletke and Rydell point to the importance of explaining terms used (ibid., 99), but there has not been any need for adding any further explanations than those already present in the ST. Furthermore, the procedure for choosing terms in the TT includes studying relevant sources and asking experts
(ibid., 97). If the TT term is created from an ST term, its context in the TT (and other texts) needs to be investigated (ibid.).

One possible strategy for locating terms is to use parallel texts (Chesterman and Wagner, 2014 [2002]: 73). Five parallel texts were identified that could be of use in the case of the leaflet (Annerén, 2016; Botella, 2013; Kraniofaciala föreningen, 2006; Socialstyrelsen, 2018; Tornkvist, 2017), but they only covered a small amount of the content (single terms). Ågrenska, a foundation and educational/recreational center for rare diagnoses, could only supply the information that 10q25/26 deletions are so rare that there was no material around. No Unique guides have been translated into Swedish, and scientific articles are written in English rather than Swedish, since they tend to be submitted to international medical journals. However, it was possible to find terms by delving into related disciplines with the knowledge that children with 10q25/26 deletions undergo surgery, so searching texts on surgery helped. For webbed neck (ST, 208), for example, there was an article on plastic surgery treatment (Chaput et al., 2013: 1), yielding pterygium colli, which led to halsvingar (Lindskog, 2016: 507).

2.2 Procedure for Analysis

When the translation was available in a draft version, readability formulas were investigated and the terms for analysis were categorized; the average words per sentence and characters per word were also determined. Regarding readability, scholars seem to agree to disagree on which formula to use for medical information. To use four examples, Wang et al. (2013: 508) consider the SMOG formula (McLaughlin, 1969: 639) to be best suited, while Williamson and Martin (2010: 1824) use the Flesch and the Flesch-Kincaid Grade Level Index formulas to evaluate readability. Baker, Wilson, and Kars (1997: 156) seem to have settled for the Flesch-Kincaid Grade Level Index only. Royal, Sheats, and Kedrowicz (2018: 59), however, use readabilityformulas.com to evaluate texts and state that “[e]ach formula is well-documented in the research literature and experts agree the formulas are highly accurate” (ibid.). The matter needs ample clarification. As for bilingual studies, Kolahi and Shirvani (2012: 349) have used the Fog Index to analyze how readability differs regarding translations from English to Persian. Acar and İşisığ (2017: 48) have compared readability between English and Turkish using the Flesch Test of Reading Ease and an adapted version of it for Turkish, the Ateşman Reading Ease formula (ibid., 50). In Chapter 3, all the readability formulas that Royal, Sheats and Kedrowicz use are listed and explained, since they include the methods used in all the other studies as well.
2.2.1 Readability
The recommendations of IMIA (Txabarriaga, 2009: 6) state that medical information to patients should target the sixth to eighth grade reading levels, and that the translator should adhere to the reading level of the ST (ibid., 6) in the TT. However, in the United States, the readability target level is set at below the fifth grade according to some (Hill-Briggs, Schumann, and Dike, 2012: 294), or between the fourth and sixth grade according to others (Royal, Sheats and Kedrowicz, 2018: 59).

The Flesch Test of Reading Ease (Flesch, 1948: 229) and the SMOG formula (McLaughlin, 1969: 639) were used to scrutinize a snippet of the text as two examples of readability tests. The Flesch Test has been found to be the most widely used reading level assessment (Wang et al., 2013: 508). The ST and TT were analyzed in their entirety according to a consensus reading (My Byline Media, 2019), but the main focus was comparing the performance of the Flesch and SMOG formulas, respectively.

2.2.2 Terminology
Merriam-Webster (2019) defines a term as “a word or expression that has a precise meaning in some uses or is peculiar to a science, art, profession, or subject” (ibid.). However, Kletke and Rydell (2014: 96) write that if a word is accompanied by defined concepts in its context, it can be identified as a term. Both methods of identification are used in this thesis. To narrow the number of terms down, the decision was made to focus on three categories of terms:

- conditions/complications
- medical/occupational/educational aids
- treatments/procedures

When a term was not found in an online resource, defined concepts in its context were investigated, such as in the following passage: “Some children need support (such as a standing frame, walking frame, support boots, a supportive Lycra ‘second skin’ and/or leg braces) while learning to walk” (ST, 289). Since the word support is possible to find as a defined concept (Merriam-Webster, 2019), it is a reasonable assumption that standing and walking frames, support boots, Lycra second skins and leg braces are specially crafted aids aimed at assisting people who are learning (or re-learning) to walk. In this thesis, they were all considered to be terms and treated as such.

Terms encountered in the text belonging to the aforementioned categories were found to be all nouns. These were divided into subcategories for the sake of analysis, though. Some consider the discipline of morphology, the study of formation of words, meaningful when it comes to terms.
However, when the majority of the terms involved are encountered in term banks and possible to verify in a medical corpus, there are other areas of investigation to consider. One of these is translation shifts, which has been one of the focus areas of this thesis. The identified terms were found to belong to the following different categories:

- **simple nouns**, consisting of one word and not possible or meaningful to analyze affixes or break down into morphemes (Bauer and Huddleston, 2017 [2002]: 1624),
- **modified nouns**, called thus for the sake of analysis, because a premodified noun in the ST often wound up being a post-modified noun in the TT or shifted to another category;
- **compound nouns**\(^2\) (Bauer and Huddleston, 2017 [2002]: 1646), a special case of modified nouns, where a noun modifies another noun or two nouns are used in sequence;
- **[potential] false friends**\(^3\) (Vinay and Darbelnet, 1995: 68), placed in a separate category for the sake of analysis.

An investigation was made of the shifts observed among the terms encountered. This investigation will be discussed in Chapter 4, where the quantitative analysis primarily shows the distribution of terms before and after translation and which translation strategies were used. There is also an investigation on how different types of identified terms shift between categories. The qualitative analysis lists examples of shifts according to the theories of Vinay and Darbelnet on different translation strategies. Examples of deviations from recommendations are discussed as they appear.

### 3 Theoretical Background

Whosoever is translating medical information needs to be aware that medical language is not a uniform concept. The text material for this thesis combines several variants, as it addresses laymen while providing them with medical terms for the purpose of their further education if necessary. Nyman (1996: 16) writes about the following three variants of medical language:

- scientific medical language, used by professionals in articles and lectures
- generic medical language, used by professionals when communicating with patients
- colloquial medical language, used by professionals speaking to colleagues

*Scientific medical language* needs to convey a single meaning (Nyman, 1996: 16); hence, it uses terms such as “gastro-oesophageal (GO) reflux” (ST, 183) or “hydronephrosis” (ST, 324). *Generic medical language* needs to be clear and informative to the patient (Nyman, 1996: 17); this involves

\(^2\)Sometimes called noun compounds or noun-noun compounds, but a corpus search reveals 178 hits for compound noun (Davies, iWeb) against six for noun compound (ibid.) and three for noun-noun compound (ibid.).

\(^3\)Called faux amis in V & D, but the English translation is used throughout this thesis.
using generic words such as “bladder” (ST, 326) or “heart condition” (ST, 328). Finally, colloquial medical language only needs to be understood between colleagues (Nyman, 1996: 17), which could involve the following question: “Now where is that femur?” – referring to a patient needing an upper leg X-ray. These three types of medical language may well co-exist, but they do affect the readability of written information, as do the terms used. For further information on medical language, see page A1 in Appendix A.

Medical information is not a uniform concept, either. It comprises instructions for taking medicine, informational leaflets about diseases aimed at both children and adults, and consent forms for procedures involving some risk to the patient, such as surgery. Regardless of the purpose or the target group, however, some sort of assessment is required of the complexity of the text, and one such available yardstick is readability.

3.1 Readability

The readability methods discussed in this thesis center mainly on the length of words (measured in syllables or characters) and sentences to determine either a corresponding reading grade level or a certain score, thus indicating if a text is easy or hard to read. There is dissent on the recommended reading grade level of medical information, which will be discussed next.

3.1.1 Previous Research

Three examples indicate that readability of medical information does not adhere to the recommended reading grade levels. There seems to be some disagreement on what the grade level should be: eighth (or fifth) grade\(^6\) (Baker, Wilson, and Kars, 1997: 156), between fourth and sixth grade (Royal, Sheats, and Kedrowicz, 2018: 58), or sixth grade (Williamson and Martin, 2010: 1824). To add to the confusion on the subject, Txabarriaga (2009: 6) recommends a reading [sic] level between the sixth and eighth grade. The TT level should also adhere to the level of the ST (ibid.). However, there is no recommendation which method to use.

The reading levels found by using the methods of these previous studies have ranged between the 14\(^{th}\) and 16\(^{th}\) grade level (Baker, Wilson, and Kars, 1997: 156) using the Flesch formula, between the 5\(^{th}\) and 11\(^{th}\) grade level\(^7\) (Royal, Sheats, and Kedrowicz, 2018: 59) using the Flesch-Kincaid grade level, and between the 2\(^{nd}\) and 11\(^{th}\) grade level (Williamson and Martin, 2010: 1825–1826),

---

\(^{4}\)This is an example from personal experience (author’s note).

\(^{5}\)Photon/proton radiotherapy and chemotherapy fit into this category, to name two other examples.

\(^{6}\)All grade levels mentioned correspond to a school grade in the United States.

\(^{7}\)Most of the grade levels should technically be expressed in letters rather than numbers, but in the name of uniformity within the paragraph, the numerical version is used.
using seven different tests and a consensus reading. There could be several reasons for this. Nisbeth Jensen and Korning Zethsen (2013: 37) write that a medical professional is keen on adhering to expert language, but “the importance of lay-friendliness cannot be emphasized enough” (ibid., 32), indicating that explanations need to be present in medical information directed at laymen. Latour and Woolgar have concluded that “[scientific] writing [is] not so much a method of transferring information as a material operation of creating order” (2013: 245), something that is relevant to a leaflet that has the purpose of bringing order to a disorder (Andersen, 2019). However, there is another problem to consider: scientific terms can be both complicated and polysyllabic. Using the terms may well mean raising the reading level above the desired target grade. Yet another possibility is the non-uniform nature of medical information: explaining genetics is a different type of text from giving instructions on how to take medicine, although in an ideal world neither would be more complex than the other.

When it comes to determining the readability of a text, one of the oldest formulas still in use today is the Flesch Test of Reading Ease, which has been in use since 1943. At readabilityformulas.com (My Byline Media, 2019), there are newer alternatives along with the original Flesch test: Flesch-Kincaid, the Gunning Fog Index, the Coleman-Liau Index, the Automated Readability Index, The Linsear Write Formula, and the SMOG formula. These are all listed and their calculations are shown in the next section, although they are by no means a complete list of all the yardsticks in use. In Sweden, there is another readability formula in use, läsbarhetsindex (lix), ‘readability index’ (My Byline Media, 2019), which is based on the number of long words (above 6 characters), the number of periods (including colons) and the total number of words counted. The reasons for not using lix in this thesis are twofold: (1) lix yields a number, not a grade level like most of the readability formulas discussed in this thesis, and (2) to be able to compare the results found in this thesis with those of previous studies, using the same yardstick(s) is beneficial.

Kolahi and Shirvani (2012) have compared readability for the translation of English to Persian of five textbooks on translation (ibid., 349) using the Gunning Fog Index. The reasons stated for the use of the Fog Index are that it is suited for texts written “above college reading level such as magazines, newspapers, and textbooks” (ibid.). Furthermore, it is stated that the formula can be used for both English and Persian without any adaptation, and that it is suitable for texts containing jargon (ibid.). While medical information contains jargon, it should not be above the college reading level, and it is somewhat surprising that newspaper and magazine texts are listed as complex from a readability perspective, given that their target group is the entire reading population, not only those who have many years of education. The results found are that the TT is less readable than the ST.
showing an average difference of 3.7 grade levels for Persian (20.1) compared to English (16.4) (Kolahi and Shirvani, 2012: 354). Scores above 17 are considered beyond the “danger line” (ibid., 350). The results are considered to be due differences between the nature of English and the nature of Persian along with the often-employed translation strategy of explicitation (ibid., 357), namely the procedure of adding information in the TT which is only implied in the ST (Vinay and Darbelnet, 1995: 170). However, when it comes to financial texts, which are also rich in jargon, Loughran and McDonald (2014: 1658–1659) criticize the Fog Index on certain grounds, although they have used it to investigate the readability of financial reports (ibid.). These grounds include tendencies for long sentences to contain fewer complex words, due to their findings that short words are necessary to link complex words together (ibid.). Furthermore, they claim that most readers of financial reports will understand words containing three syllables and more, such as corporation, agreement, and company (ibid., 1645). In the case of the material used for this thesis, terms are either explained or appear with the expressions of laymen, so the ability to comprehend should not be an issue; yet, the readability results are outside the recommended range (Txabarriaga, 2009: 6) for the ST and the TT.

Acar and İşişağ (2017: 48) have also studied the difference in readability between two languages. Their study covers technical and scientific texts in English (SL) and Turkish (TL) using the seven-formula consensus reading and the Ateşman Reading Ease formula (ibid.), which is the only readability formula available for Turkish texts. However, this specific formula is an adaptation of the Flesch Test of Reading Ease (ibid., 50). Their findings include a slight improvement of readability between their chosen STs and TTs (ibid., 52), but these observations are not explained.

According to some scholars, “no unified explanatory model of text readability has been produced so far” (François 2015), and with so many different formulas still in use which focus on the length of words or the length of sentences, the author of this thesis can only agree. Wang et al. (2013: 508) propose to use the Simple Measure of Gobbledygook (SMOG) in medical texts, but Kolahi and Shirvani (2012: 349) state that the Fog Index is suited to texts rich in jargon, although Gunning (1969: 11) claimed that “[the Fog Index is] a simple warning system” and that there is no formula that can guarantee good writing (ibid.). Having said that, it is time to find out more about which readability tests have been used in previous studies cited in this thesis.

---

8 The study in question also covers comprehensibility, measured by readers, but that concept is not discussed in this thesis.
3.1.2 Readability Tests

*Flesch Test of Reading Ease*. First published in 1943 (My Byline Media, 2019), the Flesch Test of Reading Ease is still one of the most widely used readability tests around (Hartley, 2016: 1523).

The Flesch test of reading ease uses the following formula:

\[
R.E. = 206.835 - (1.015 \times ASL) - (84.6 \times ASW)
\]

where

- \( R.E. \) = Reading Ease
- \( ASL \) = Average Sentence Length, and
- \( ASW \) = Average Syllables per Word.

**Formula 1.** The Flesch test of reading ease (Flesch, 1948: 229).

To achieve the reading grade level, early users needed to look in a conversion table. This may be the reason why Kincaid revised the Flesch formula (Kincaid et al., 1975: 14) to show a grade level (GL) according to the following:

\[
GL = 0.39 \times \left( \frac{\text{words/sentence}}{\text{sentence}} \right) + 11.8 \times \left( \frac{\text{syllables/word}}{\text{word}} \right) - 15.59
\]

**Formula 2.** The Flesch-Kincaid grade level index (Kincaid et al., 1975: 14).

This formula is known as the “Flesch-Kincaid Index, Flesch-Kincaid Grade Level Score, Flesch-Kincaid Scale, Flesch-Kincaid Score, Flesch-Kincaid Readability Score, Flesch-Kincaid Readability Statistics, Flesch-Kincaid Grade Level Index, Flesch-Kincaid Readability Index, Flesch-Kincaid readability equation...” (My Byline Media, 2019). In this thesis, it is called Flesch-Kincaid Grade Level Index, to distinguish it from the Flesch Test of Reading Ease.

*The (Gunning) Fog Index* (My Byline Media, 2019) was developed by Gunning, a textbook publisher (ibid.), who wrote that it is a “simple warning system” (Gunning, 1968: 12). It yields a grade level (GL). The formula to determine the Fog Index is as follows:

\[
GL = 0.4 \times (ASL + PHW)
\]

where

- \( ASL \) = Average Sentence Length; divide the number of words by the number of sentences
- \( PHW \) = Percentage of Hard Words, meaning words consisting of \( \geq 3 \) syllables per 100 words.

**Formula 3.** The Gunning Fog Index (My Byline Media, 2019).

*The Coleman-Liau Index* (Coleman and Liau, 1975: 283) was developed according to the following principle: “There is no need to estimate syllables since word length in letters is a better predictor of readability than word length in syllables” (ibid.), and for the purpose of using
computers to calculate the score instead of counting by hand. The calculation to determine the Coleman-Liau index is:

$$CLI = 0.0588 L - 0.296 S - 15.8$$

where

\[ L = \text{the average number of letters per 100 words, and} \]
\[ S = \text{average number of sentences per 100 words.} \]

**Formula 4.** The Coleman-Liau Index (My Byline Media, 2019).

*Automated Readability Index* (Senter and Smith, 1967: 8). Originally developed for use in real time on typewriters (ibid.), it has been evaluated against other formulas (Kincaid et al., 1975) and stayed in use. The formula for the reading grade level is:

$$GL = 0.50 \times \left( \frac{w}{s} \right) + 4.71 \left( \frac{c}{w} \right) - 21.43$$

where

\[ GL = \text{[reading] grade level,} \]
\[ w = \text{word (or words),} \]
\[ s = \text{sentence, and} \]
\[ c = \text{characters (originally: strikes on a typewriter).} \]

**Formula 5.** The Automated Readability Index (Senter and Smith, 1967: 8).

The *Linsear Write Formula* (Khosrow-Pour, 2017: 1503) was developed by the U.S. Air Force and intended for technical manuals. The procedure for determining the Linsear score is: first, find a 100-word sample. For this sample, do the following (My Byline Media, 2019):

Identify the easy words, meaning words containing two syllables or less. Simple words such as an, the, and others are included. Mark them with a “1”. Then, find the hard words, meaning those consisting of three or more syllables, and label those with “3”. The calculation to perform is:

$$\frac{\text{number of easy words} \times 1 + \text{number of hard words} \times 3}{\text{number of sentences}}$$

**Formula 6.** The Linsear Write Formula (My Byline Media, 2019).

If the result from the calculation is > 20, divide by 2; if it is ≤ 20, subtract 2 and divide by 2. The number received is a reading grade level.

According to Wang et al. (2013: 508), the readability formula found to perform the best in the discipline of medical information is the Simple Measure of Gobbledygook (SMOG) (McLaughlin, 1969: 639; Wang et al., 2013: 506). The *smog formula* is concerned with syllables over sentences.
3 + \sqrt{\frac{\text{number of words containing } \geq 3 \text{ syllables}}{\text{number of sentences counted}}} \times 30 / \text{number of sentences counted}

Formula 7. The SMOG readability formula (McLaughlin, 1969: 639)

Ten sentences at the beginning of the text, ten in the middle, and ten at the end of the text are counted, but the formula allows for counting fewer sentences as well.

There is some level of systematization in SMOG, since not all of the text is used, but it does not single out words in the way that the Flesch Test of Reading Ease does, for example. The same goes for the Linsear Write formula, which only investigates a 100-word sample.

According to all of these formulas, however, three-syllable words are to be avoided to keep the readability at an acceptable level if the Flesch, Flesch-Kincaid Grade Level, Gunning Fog, SMOG and Linsear methods are used. The remaining two, Coleman-Liau and Automated Readability Index, are concerned with characters per word, but a long word would tend to contain more syllables. For any method, the average sentence length should also be kept down. Hill-Briggs et al. (2012: 294) write that to achieve a reading level below fifth grade in order to reach readers with low literacy, the following should apply (ibid.):

- the sentence length should be less than 15 words
- the active voice should be the predominant form, with less than 5 % in passive voice
- common words should be used, avoiding or minimizing the use of polysyllabic words (more than two to three syllables).

The two major text ingredients that affect readability are long sentences, and so do polysyllabic words. This leads to the subject of terminology, which will be discussed next.

3.2 Terminology

In medical information, misinformation or the lack of information can lead to serious adverse effects (Dalton-Oates, 2017: 234–235). Even a minor error may lead to a reader drawing the wrong conclusions, and while the interpretation may eventually be close to the truth, it may not be possible to backtrack and understand exactly what the correction should have been (O’Neill, 1998: 71). On the other hand, Kerna (2018: 1) writes: “What may seem foreign, impractical or intimidating to the layperson is quite meaningful to the healthcare practitioner or medical researcher.” To communicate with professionals, the reader needs to be made aware of the correct medical terminology.

An example, although it is not a term as such, is the proper name of Llanfairpwllgwyngyllgogerychwyrndrobwllllantysiliogogoch, a town on the island of Anglesey just off the coast of northwestern Wales. With 58 characters and 19 syllables, it will affect any desired readability test result adversely.
While terminology may cause problems for some and create understanding for others, in any kind of scientific (or even non-fiction) text, there is a set of terms (Kletke and Rydell, 2014: 93), and medical information is no exception. Kletke and Rydell go on to say that first, the terms already in use for a set concept are to be used, and to consider the criteria of a good choice of term (ibid., 97). Terms in the leaflet serve both to inform and ease the worry, because by knowing the terms, parents can communicate with medical staff and find out more. This could be done in any order: either they search for information and talk to medical staff afterwards or they manage by communicating first (Unique, 2013).

However, terms should not come about lightly; they should be constructed with logic in mind. German philosopher and logician Frege contributed to the model of terms by distinguishing between two types of meaning, namely what the thought content, or sense, of a specific word was, and what reference it implied (Frege 1948 [1897]: 209). An example is the planet Venus, of which Frege writes that the reference is the same, but not the sense: “The referent of ‘evening star’ would be the same as that of ‘morning star,’ but not the sense” (ibid., 210). Venus can sometimes be observed in the morning, sometimes in the evening; thus, it has the designation Morning Star and Evening Star, respectively.

Ogden and Richards (1923: 11), however, discussed language use by means of what is called the semiotic triangle, where the thought or reference, at the top of the triangle, denotes a person’s experience, or recollection, of the referent, which invokes the thought, which is pictured at the right base corner of the triangle. On the left base corner is the symbol, the word that brings the referent to mind in the process of a person’s thoughts (ibid.).

Figure 1. The Ogden-Richards semiotic triangle (Ogden and Richards, 1923: 11).

14 (50)
The attentive reader will note that there is no base in the Ogden-Richards triangle; this is because the relationship between symbol and referent is considered not to be indirect (Ogden and Richards, 1923: 11–12). However, there are cases where the triangle will be complete, for example with onomatopoeic words (ibid., 12), where the symbol imitates the referent. Terms are the symbols of the referent; they need to create a thought in the reader’s mind.

Ingo (2007: 86–87) writes that the task of translation encompasses the task of transferring meaning from source language (SL) to target language (TL); it is also done after a thorough analysis of the SL meaning (ibid., 87). There is a clear distinction between the denotation of a word, or its basic meaning, and its connotations, or what associations the word evokes (ibid.). But to translate terms also requires a source of knowledge, and where to find them.

3.2.1 Previous Research
Peters, Qian, and Ding (2018: 99) have performed a study on translating medical terms from English to Chinese. The first issue to note is the extensive borrowing of terms from Greek and Latin (ibid.); in this particular study, there is the added challenge of the SL and TL belonging to different language groups (ibid.). According to Peters, Qian, and Ding (ibid., 101), “[f]or translators, as for bilingual terminographers, the challenge of finding appropriate [translation equivalents] for medical terms may be there even when there is a set of pre-existing alternatives for the same medical terms, i.e., heteronymy” (ibid.). This situation has been encountered in the text material for this thesis. There is also the phenomenon of polysemy (ibid., 102), where a term has been borrowed from another discipline and is given a new meaning by means of a modifier. As an example, the word radiation is given, which can apply to solar radiation, atomic radiation from uncontrolled nuclear power (ibid.), or even radiation therapy (Karolinska Institutet, 2019), which is an accepted synonym of radiotherapy (Peters, Qian, and Ding, 2017: 102). The results of this study indicate that the Chinese translation equivalents contain more information, which could be restated as they incorporate explicitation. Laymen receive the explanation with the term and there is no need for the paraphrasing that may have had to accompany the English term (ibid., 109). The corpus approach is favored; however, it is worth noting that there is no access to term banks such as SNOMED (ibid.), leaving the reader with an open conclusion: what if there had been access to term banks? Would the study have yielded different results? A study that has used both term banks and corpora is discussed in Section 3.3.1.

Sometimes, the baseline in the semiotic triangle is dashed or dotted (for example in Campbell et al., 1998: 423).
3.3 Translation Strategies: Sourcing

Ingo (2007: 107) writes that even special dictionaries may not include new terms, and that in difficult situations, a translator may consult term banks (ibid.). During the course of translation and analysis for this thesis, encountered terms have mainly been found in term banks rather than special dictionaries, although the special dictionaries used (Egidius, 2019; Lindskog, 2016; Malmquist and Lundh, 2016) have served well as complementary sources. Two recognized term banks within the field of medical terminology are SNOMED CT and MeSH. As SNOMED CT is the recommended term source of diagnoses according to the Board of Directors of the American Medical Informatics Association (AMIA) (Awayssheh et al., 2018: 17); it is structured in hierarchies so that the translator may understand the context of a term. For example, a search for an occurrence such as sandal gap will yield with what syndrome it is associated, that it is a congenital foot anomaly, etc. (SNOMED CT, 2019). MeSH (Karolinska Institutet, 2019) places medical terms within a hierarchy, too; however, the returned results are presented with the Swedish (in this case: TL) and English (in this case: SL) definitions in sequence so that readers will know the context of the term. A search for hypoplastic left heart syndrome will produce the definitions in both the SL and TL and list broader or narrower terms in a hierarchy (ibid.)

If a term is not available in any of the above sources, Ingo suggests following certain procedures of forming terms in Swedish (2007: 107–108). Those used in this thesis, either as part of the translation process or as already existent terms, are the following:

**Affixes.** Swedish uses both prefixes (an affix at the beginning of a word) and suffixes (an affix at the end of a word) during word formation. Examples include obalans, ‘imbalance’, utradera, ‘eradicate’ (prefixes), and kanonad, ‘cannonade’ entreprenad ‘outsourcing’ (suffixes). There is also the possibility of an infix (an affix within a word), which is not common in English, but a colloquial (albeit cleaned-up) example is abso-freaking-lutely! (Ingo, 2007: 107).

**Compounding.** In this case, a compound word will represent the meaning; examples include motorvärmare, ‘motor preheater’, and formpressad, ‘moulded’ (Ingo, 2007: 108). While no compounding has been used in this thesis to create new terms during the translation process, compounds are present among the terms in both languages.

**The creation of a phrase.** Examples include filosofie kandidat, ‘Bachelor of Arts’, volt framåt med dubbel skruv, literally ‘forward somersault with two twists’ (Ingo, 2007: 108). Instead of using a single word, there are entire phrases: *Bachelor of Arts* is a noun phrase (NP) with a post-modifier,
of Arts, after the head, *Bachelor*.* Forward somersault with two twists* is another NP, where the head of *somersault* is premodified by *forward* and postmodified by *with two twists*.

The creation of new terms may not be necessary; Lušicky and Wissik (2015: 12–13) include these strategies in “translation-oriented terminology work” (ibid.). The procedures used in this thesis are listed below and will be addressed in Section 4.3.2.

**Collecting resources**, by which is meant consulting standards, dictionaries, term banks, experts, corpora... (ibid.,16).

**Term extraction and term selection**, which can be done manually by reading the texts and selecting appropriate terms, or by using software tools for the purpose (ibid., 20).

**Terminological research in both the source and target language**, which means researching the context and description in both the SL and TL, documenting the term and its definition along with grammatical information, sources, etc. (ibid., 20–21).

3.3.1 Previous Research

Another study makes use of both corpora and term banks. Deléger et al. (2010: 152) have investigated the translation of terms present on *Medline Plus*, a health information site intended for laymen (ibid.), from English to French by using a twofold approach. The material used, which was around in October of 2009, had 1,982 entries, of which 848 were preferred terms and 1,134 were synonyms, used to reach the preferred terms. The knowledge-based approach consists of the following procedure: four term banks have been consulted, where SNOMED is one and MeSH is another (ibid., 153). Alongside this knowledge-based approach, parallel medical corpora have also been consulted (ibid.). The results indicate that term bases provide the most translations; of 611 *Medline Plus* terms found in term banks, there was an acceptance rate of 67.6 % (ibid., 154). Using medical corpora yielded 143 translated terms of which 71.3 % were accepted (ibid.). The total number of accepted terms was 435, or 51.3 %. However, the authors make a note of corpora being more likely to contain terms that are oriented to patients (or laymen) rather than professionals (ibid.). Morin et al. (2010: 1:1), however, have investigated the process of searching for terms using comparable, not parallel, corpora (ibid., 1:2). The reason stated for this is that parallel corpora are often translated, and as such, the vocabulary has been affected by the ST (ibid.). Moreover, terms may well be phrases and consist of more than one word, which means that their translations are not likely to consist of the same parts as their SL versions (ibid., 1:4). The term *plantation énergétique* in French is used as an example: its English counterpart is *fuel plantation*, where the parts *fuel* and
énergétique would not match as individual translation pairs (ibid.). Furthermore, it is stated that using comparable corpora in connection with bilingual dictionaries have yielded translations that received a high acceptance rate (ibid., 1:16). But it is time to address another phenomenon which affects translation, including terms: shifts.

3.4 Translation Strategies: Shifts

The concept of translation shifts was introduced by J. C. Catford (1965: 73) and defined as “departures from formal correspondence in the process of going from the SL to the TL” (ibid.). Vinay and Darbelnet (1995) have described the same phenomenon, but in different terms. According to Vinay and Darbelnet, there are two strategies for translation, direct translation (Vinay and Darbelnet, 1995: 31) and oblique translation (ibid., 36–41); the shifts according to Catford belong to the latter category. There are three direct translation strategies: borrowing, calque and literal translation, and four oblique translation strategies: transposition, modulation, équivalence (idiomatic translation) and adaptation.

Another strategy for translating medical terms, synonym replacement, is listed here, too, because it comes from a more recent and specific study related to medical terminology (Abrahamsson et al., 2014: 57). All of the strategies listed here are mentioned alongside examples in Chapter 4.

**Borrowing** (Vinay and Darbelnet, 1995: 31–32) is the direct transfer of a word from the ST to the TT. Examples include glasnost, perestroika, and sushi (ibid.). There is an example of borrowing in Section 4.4.1.

**Calque** (Vinay and Darbelnet, 1995: 32–33) transfers an expression from the ST to the TT; there are no such examples discussed in this thesis. However, calque sometimes leads to problems in a translation; instances where such problems arise are when a false friend (Vinay and Darbelnet, 1995: 68) appears. This is when an expression in the ST has a similar expression in the TT which either means something completely different or simply does not work, which is a pitfall for a translator. Sections 4.4.1 and 4.4.4 include examples of potential false friends.

**Literal translation** (Vinay and Darbelnet, 1995: 33–35) is the literal translation of an expression, or ‘word-for-word’, as in the following example (ibid.):

“English ST: I left my spectacles on the table downstairs.
French TT: J’ai laissé mes lunettes sur la table en bas.”
Transposition (Vinay and Darbelnet, 1995: 36) is the replacing of “one word class with another without changing the meaning of the messsage” (ibid.), for example a verb for a noun, as in the following example (ibid.):

“Dès son lever...” : As soon as he gets/got up...
As soon as he gets up...
: Dès son lever...
: Dès qu’il se lève...

There is an example of this phenomenon in Section 4.4.2.

Modulation (Vinay and Darbelnet, 1995: 36) is a procedure considered justified (ibid.) when literal translation, while grammatically correct, appears “unsuitable, unidiomatic or awkward in the TL” (ibid.). Modulation may be either obligatory or optional; in the latter case, it is “linked to preferred structures of the two languages” (ibid.), as in the following example (ibid.):

“it is not difficult to show” : il est facile de démontrer...

Examples of modulation from the ST/TT pairs are discussed in Sections 4.4.2 and 4.4.3.

Équivalence (Vinay and Darbelnet, 1995: 38) is the practice of describing a situation where there is no exact expression in the TT for the ST. Suitable candidate expressions for the use of équivalence include idioms, as in the following example (ibid.):

“comme un chien dans un jeu de quilles” : like a bull in a china shop"
(‘like a dog in a game of skittles’)

This strategy has been criticized because Vinay and Darbelnet (1995: 35) mentioned a bilingual dictionary of accepted expressions; this would be a utopia (Panou, 2013: 2). However, this is a central part of translating literary as well as non-fictional works.

According to Jakobson “there can be no full equivalence between two words” (1959: 114), but this does not mean that translation is impossible; it simply means that different languages use different definitions and include more or less in them. Panou (2013: 2) notes that there are other strategies that can be used according to how they are suited. Examples of different strategies will be discussed throughout the qualitative analysis section of Chapter 4, while examples of équivalence, or idiomatic translation, appear in Section 4.4.1.

Adaptation (Vinay and Darbelnet, 1995: 39–40) is a strategy used when there are cultural differences reflected in the ST and TT, since there is no situation in the TT that can be explained by literal translation of the ST. According to Vinay and Darbelnet, the Tour de France might be a French adaptation of what cricket is to the English (ibid.). There is an example of adaptation in Section 4.4.5.
**Loss, gain, and compensation.** “[T]wo languages do not report one and the same situation in the same way.” (Vinay and Darbelnet, 1995: 169–170), which usually leads to loss. A loss may well be compensated (ibid., 198) at some point by introducing gain, (ibid., 170), namely adding information. There are examples of loss, gain or compensation in Sections 4.4.1 and 4.4.2.

The practice of *explicitation* (Vinay and Darbelnet, 1995: 170) is that of adding information in the TT which is only implied in the ST. There is an example of explicitation in Section 4.4.3.

Abrahamsson et al. (2014: 57) have studied the phenomenon of *synonym replacement*, meaning when terms from other languages are replaced by Swedish equivalents: “For synonym replacement to be a meaningful method for text simplification, there must exist synonyms that are near enough not to change the content of what is written” (ibid.) The material for study was compiled by building a corpus consisting of articles from *Läkartidningen*, a Swedish medical journal (ibid., 59). The results were that by synonym replacement, the number of unique terms decreased, but the readability was deemed slightly more difficult by one readability method and slightly simpler by another (ibid., 61). The first readability method used, *LIX*, singles out words that contain more than six characters and also takes sentence length into account. However, the other, *Ordvariationsindex*, ‘Word Variation Index’, or *OVIX*, is concerned with the number of unique words compared with the total number of words only, not their length (ibid., 60). A few examples of synonym replacement include *osteoporosis* → *benskörhet*, ‘brittleness of the bones’ (ibid., 59) and *stroke* → *slaganfall*, ‘strike-seizure’ (ibid., 61). The latter is one of the examples of how the readability level is more difficult with synonym replacement but the word variation is prone to decrease. There is an example of synonym replacement from the ST/TT in Section 4.4.2.

### 3.4.1 Previous Research

In one study found, Ndwayamato Madadzhe and Mashamba (2014: 51) highlight shifts in practice in their research around translating medical terms from English to Tshivenda (ibid.). They have found that literal translation will end up being incomprehensible in Tshivenda (ibid., 53), and that the language is suited for idiomatic translation out of respect (ibid., 54). There is also *transposition*; in Tshivenda, the word for *sick*, *lwala*, is a verb, not an adjective (ibid., 55). The situation of no equivalent expression in the TL, or *zero equivalence*, often makes translators turn to *borrowing* (ibid., 54). Apart from the lack of term resources, they have also found that a translator needs to know about cultural matters (ibid., 56), if the TT is to be understandable. As an example, they have used the word-for-word translation *Ṉowa yawe a i fari*, ‘her snake does not catch’ when the actual
meaning is *she cannot conceive* (ibid., 53). An example of where adaptation is needed appears in Section 4.4.5.

Pietrzak (2015) has made comparisons between discharge summaries in English and Polish (ibid., 319). She states that in English, descriptions tend to be more descriptive and less formal, while Polish uses short, concise sentences without verbs (ibid., 321). According to Pietrzak, “it is of primary importance for the translator to ensure that the text, which—in medical environment—must be read quickly but thoroughly, is as familiar as possible” (ibid., 329). While the text material of this thesis may not be read quickly, it still needs to be familiar, as a stressed parent will not get through the text otherwise, and the element of recognition cannot be underestimated.

### 4 Analysis

This chapter addresses readability before translation strategies and translation shifts, as readability is concerned with the translation in its entirety and the latter two categories involve examining the translation in more detail.

#### 4.1 Readability

In order to understand more about readability issues, the basic properties of the texts have been investigated, namely the average number of words per sentence and the average numbers of characters per word. These properties may be investigated by using either a word processor or an online service (My Byline Media, 2019). After this investigation, two of the formulas are compared and examined in more detail.

#### 4.1.1 Text Properties and Test Results

Before showing the readability scores, the basic properties of the ST and TT respectively will be discussed, as they form relevant background information for readability scores. The number of total words, the number of total characters, the average number of words per sentence were determined (My Byline Media, 2019) for the ST and TT respectively, along. They show the following results:

<table>
<thead>
<tr>
<th>Text investigated</th>
<th>Number of total words</th>
<th>Number of total characters</th>
<th>Average number of words/sentence</th>
<th>Average number of characters/word</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>6,046</td>
<td>35,988</td>
<td>18</td>
<td>4.5</td>
</tr>
<tr>
<td>TT</td>
<td>5,696</td>
<td>35,507</td>
<td>15 (+ 20 %)</td>
<td>4.7 (+ 4 %)</td>
</tr>
</tbody>
</table>

*Table 1. Word/sentence counts for the ST and TT.*
This shows that the TT is shorter than the ST, both in terms of the number of words and characters. Kolahi and Shirvani (2012: 357) state that “translated texts contain more words than their originals” (ibid.), which is not the case regarding the material of this thesis. One of the reasons for this is that it has been possible to streamline the text somewhat and avoid repeated statements in close proximity. While streamlining does not exclude the use of explicitation, the word and character length of the TT still contradicts the statement of Kolahi and Shirvani. According to some scholars, however, shorter sentences tend to contain a higher percentage of longer words (Loughran and McDonald, 2014: 1658–1659), which makes it improbable that both components actually measure readability (ibid., 1659). Their study found “a positive correlation between complex and common words” (ibid.); when a complex word appeared, there would be common words around it for the purpose of linking and clarification. While the TT is sure to contain explicitation, there are passages where it has been possible to streamline the TT compared to the ST, such as in (2):

(2) Some children continue to have articulation difficulties, with consonant sounds often proving more difficult, and/or their speech may sound ‘slurred’.

Many children continue to have articulation difficulties, with consonant sounds often proving more difficult, and/or they may stutter or have ‘slurred’ speech. (ST: 255, 259)

En del barn har fortsatt svårt med artikulationen, där konsonantljuden oftast visar sig vara svårare.

Barnen kan stamma eller sluddra.

(TT: 255, 259)

At a word count of 20 compared to 42, the TT avoids repetition in close proximity. Regardless of the number of words per sentence, the number of syllables per word, and the average sentence length, this is too small a sample to investigate by itself. François (2015) states that it is “common knowledge” that no text sample should be shorter than 100 words, which also happens to be the sample size used in The Linsear Write Formula (My Byline Media, 2019). Graesser et al (2004: 199) state that for the Flesch Test or the Flesch-Kincaid Grade Level formulas to work, a text should be longer than 200 words (ibid.).

However, Kolahi and Shirvani (2012: 357) suggest that an increased level of reading difficulty can be attributed to explicitation. While there are instances of explicitation in the TT, the overall length is shorter, which in itself does not affect readability at all. The average length of sentences is shorter in the TT, which should generally affect readability positively. The average number of characters per word, which is slightly higher in the TT, will affect readability negatively; however, again on an average, both stay below five characters per word. This means that the reasons for some methods scoring the TT as being more difficult to read than the ST are found elsewhere. Keeping the basic
text properties in mind and using the seven consensus readability calculations on readabilityformulas.com (My Byline Media, 2019), these are the readability scores:

<table>
<thead>
<tr>
<th>Text for assessment</th>
<th>Flesch Reading Ease</th>
<th>F-K Grade Level Index</th>
<th>Fog Index</th>
<th>Coleman-Liau</th>
<th>SMOG</th>
<th>Automated Readability Index</th>
<th>Linsear Write Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>56.9</td>
<td>9.7</td>
<td>12.8</td>
<td>9</td>
<td>9.5</td>
<td>8.7</td>
<td>11.5</td>
</tr>
<tr>
<td>TT</td>
<td>48.1</td>
<td>10.2</td>
<td>12.6</td>
<td>10</td>
<td>9.6</td>
<td>8.1</td>
<td>9.1</td>
</tr>
</tbody>
</table>

Table 2. Summary of the readability scores (My Byline Media, 2019).

A higher Flesch Reading Ease number indicates that the text is considered more reader-friendly. All the other numbers above refer to U.S. grade reading levels. A level between two integers, such as 9 and 10, follow the usual rounding rules in mathematics: it is rounded upward for an odd number \( x \) if it is equal to or more than \( x.5 \) and downward if it is below \( x.5 \), or equal to \( x.5 \) for an even number. Therefore, the Flesch-Kincaid Grade Level Index score for the TT above is considered to be 10\(^{th}\) grade, whereas the results from using the SMOG formula also puts the reading level at 10\(^{th}\) grade, even though it is 9.6. The consensus reading is based on all seven formulas, and it was grade level 10, or fairly difficult to read, for the ST; the reading age was found to be between 14 and 15 years (My Byline Media, 2019). For the TT, the corresponding grade level was one level lower, or 9, and the age of the target readers was found to be 13 to 15 years old (ibid.).

Examples of research studies comparing results between languages include Kolahi and Shirvani (2012), and Acar and İşisağ (2017). Research studies comparing readability within the English language, namely Baker, Wilson, and Kars (1997); Royal, Sheats, and Kedrowicz (2018); and, finally, Williamson and Martin (2010), have found that readability levels are higher than recommended for medical texts in the SL, which can be confirmed by the investigations carried out for this thesis. However, the differences in readability encountered in this thesis contradict the findings of Kolahi and Shirvani, as well as those of Acar and İşisağ (2017: 52). The Fog Index is slightly lower for the TT; Kolahi and Shirvani (2012: 357) have found their TT samples to be on an average 3.7 grade levels above that of the ST. The study of Acar and İşisağ (2017: 52) has found that the Flesch number result is higher for the TT, indicating that it is easier to read, whereas the Flesch results for the material of this thesis indicates that the TT is harder to read. All individual methods used in this thesis that are most affected by longer words (Flesch, Flesch-Kincaid, Coleman-Liau, and SMOG) have scored the TT higher than the ST. Those concerned with sentence length (Fog Index, ARI and Linsear Write Formula) mostly score the TT lower than the ST. However, there is a difference of only 0.1 using SMOG, hardly pointing to a significant difference. To
understand more about how readability formulas work, a sample of the ST/TT will be investigated in some more detail in the next section.

4.1.2 Flesch Reading Ease vs SMOG

Two of the formulas used to determine readability will be investigated in more detail. These are the Flesch Test of Reading Ease (Flesch, 1948: 229) and the SMOG formula (McLaughlin, 1969: 639). Published in 1943, the Flesch test of Reading Ease, although revised, is still one of the most widely used readability tests around (Hartley 2016: 1523).

In the case of the text material for this thesis, compounds in the ST consisting of two or more words in sequence will end up being a single, often longer, compound word in the TT, making the Flesch score of reading ease state a higher level of reading difficulty simply based on the inherent properties of the SL and TL, as in (3) on the next page. François (2015) states that it is “common knowledge” that no text sample should be shorter than 100 words. For the sake of comparison, a sample from the ST/TT of approximately that size is used in (3):

(3) A number of parents also had unusual findings when undergoing ultrasound scans:

- two babies had a two vessel umbilical cord; a kidney anomaly was detected in one baby; increased fluid around the baby was seen in two pregnancies; ureteral reflux was detected in one baby; and one baby was diagnosed with hydronephrosis.

There are also two cases in the published medical literature in which urinary tract anomalies were detected on prenatal scans (Scigliano 2004; Unique).

Only one family known to Unique discovered the 10q25.2q26.13 deletion before their baby was born.

An amniocentesis was performed after a nuchal scan showed an increase in fluid at the back of the neck. (ST, 148–152)

The Flesch Reading Ease of the sample is 39.4 for the ST, which is in the “difficult to read” span, corresponding to a score of 30 to 49 (My Byline Media, 2019) and 32.9. The score of the TT is also deemed difficult at a score of 32.9 (ibid.), although this value is closer to the “very confusing” span, which sets in at a score of 29 or less. A person who is at the same reading level of the ST and the TT
should have no problem with either text; while the words in the TT are simply longer, they are not more complex. Yet, there is a notable difference in the Flesch Reading Ease score; based on this small sample, this test is not suited for the text at hand. Table 2 shows the same tendency: the difference between the ST and the TT is a score of 56.9 for the former and 48.1 for the latter. The original Flesch Test uses a systematic choice of words and takes only their length into account. These two testing criteria are used by Hartley (2016: 1525) as two reasons to state that it is time to abandon the method. Hartley then proposes using methods that take the age, ability and discipline of the reader into account – or a combination of all three (ibid.). All other readability formulas used on readabilityformulas.com, except for the Flesch Test of Reading Ease, account for the reader’s age (Flesch-Kincaid, Fog Index, Coleman-Liau, SMOG, ARI, Linsear Write Formula), but none of them take the ability and discipline of the reader into account.

Ability and discipline may well be needed, as any of the above formulas, including SMOG (McLaughlin, 1969: 639) is affected by the inherent properties of the ST and TT. The readability level according to SMOG, using (2) above, shows that the ST receives an index of 11.3, or eleventh-grade reading level (My Byline Media, 2019), and the TT received an index of 10.6, which is also rounded to the eleventh-grade level (ibid.). While the TT is seen as less complicated according to the SMOG formula, it is within the same school grade. The average syllable length of the TT, which is 1.90 syllables, is slightly larger than that of the ST, where the corresponding number is 1.76 syllables. This may seem strange at first sight, but this is where the nature of two different languages comes in alongside the nature of readability formulas.

While Kolahi and Shirvani (2014: 357) have used neither Flesch nor SMOG for investigating the readability levels in a translation study, one comment in particular is worth noting. “Since the variables in the Gunning Fog Index are language dependent, this nature may affect the results obtained by the readability formula” (ibid.). In Section 3.2.1, the nature of readability formulas was discussed and it was established that for the Fog Index and other methods, three-syllable words or above should be avoided in order to make texts easier to read – in any language. The readability formulas used in this thesis, including the Fog Index, are all concerned with the length of words and sentences, and evidently, a 4% increase in word length makes a difference in the readability score.

To try to find out why this slight increase in word length causes such an effect, an cross-linguistic examination is in order. This is done using a particular group of terms among those identified for analysis: the group of compound nouns, meaning nouns pre- or post-modified by other nouns. In English, these compounds can be formed by putting a deverbal noun either in the pre-modifying or post-modifying position (Bauer and Huddleston, 2017: 1654), such as in learning difficulties.
Other formation methods include the placement of a deverbal noun in the post-modifying position, such as in pregnancy complications (ST, 144). In Swedish, learning difficulties will be inlärningssvårigheter (literal translation). Pregnancy complications can be translated as either komplikationer under graviditeten, ‘complications during pregnancy’, or graviditetskomplikationer (literal translation), respectively. The translation graviditetskomplikationer with 120 hits (Språkbanken) is more common than komplikationer under graviditeten, which gets 44 hits (Språkbanken), but from a readability perspective, the latter choice is preferred. Swedish compound nouns can often be formed as deflections, where the first noun acts as a prefix, connected to the second by the use of an s (Teleman, Hellberg and Andersson, 2010[1999]: 43), as in the case of inlärningssvårigheter and graviditetskomplikationer. The source language uses two words with 2 + 4 syllables for an average word length of 3.00 syllables (learning difficulties) or 3 + 4 syllables for an average word length of 3.5 syllables (pregnancy complications). However, the target language has to use a single word for the former expression, inlärningssvårigheter, with seven syllables. The alternate translation, which was used for the latter, komplikationer under graviditeten, uses 5 + 2 + 5 syllables for an average of 4 syllables; having gone for the most common occurrence (Språkbanken) would have resulted in a single, nine-syllable word, which goes against the recommendations of Hill-Briggs et al. (2012: 294) to avoid polysyllabic words (ibid.). It follows that if polysyllabic words cannot be avoided, at least the count of syllables should be kept down.

The difference in the nature of the two languages becomes even more evident when comparing ST words such as mitral valve prolapse with mitralklaffprolaps (ST/TT, 334). Neither of these are compound nouns, but the difference in the number of syllables between the source and target language is still there. Any readability test concerned with syllables will consider the ST term to be three words with an average length of 5/3, or 1.67, syllables, while the TT counterpart gets a five-syllable rating. Many of the polysyllabic words in the ST/TT are scientific terms, which are hard to avoid in a medical text, and they will be discussed in Section 4.3; however, the question of translation strategies, concerning the sourcing and the distribution of terms, needs to be addressed.

4.2 Terminology
When faced with terms, medical or otherwise, a translator may have to turn to many sources to have access to as complete a basis as possible. Term banks may or may not contain the entirety of the terms found in the ST; the same goes for corpora. In this section, the sourcing and categorization of terms included in the analysis is discussed.
4.2.1 Sourcing

To find terms in medical information, a translator will most probably need to consult multiple sources. In the study of Deléger et al. (2010: 154) that investigated the translation of medical text from Medline Plus, it was reported that for 848 terms requiring translation, 67.6% of 611 terms found in term banks were accepted, while the terms found in corpora yielded 143 translated terms of which 71.3% were accepted (ibid.). In total, 435 out of 848 preferred terms were possible to find using the combined approach. As Medline Plus consists of medical text with the target group of laymen, the study is relevant to the material investigated in this thesis. However, the investigation here has chosen only 79 terms, and Table 3 shows the sources of their translations. The percentages in Table 3 apply to the number of terms found and used in the TT. The acceptance rate is that of the translator, who has “some conceptual knowledge” (Pietrzak, 2015: 317) of medical terms but is not a medical expert or a chief librarian, both of which were accessible to Deléger et al. (2010: 154) for both manual translation and validation. No professional validation has been performed within the scope of this thesis, but the term entries used in the translation have been investigated to make sure they were part of a relevant context (Kletke and Rydell, 2016: 97).

<table>
<thead>
<tr>
<th>Knowledge-based (term banks, dictionaries)</th>
<th>Corpus-based</th>
<th>Expert-based</th>
<th>Other (inflected, derived etc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>67/79 (84.8 %)</td>
<td>65/79 (82.3 %)</td>
<td>5/79 (6.3 %)</td>
<td>1/79 (1.3 %)</td>
</tr>
</tbody>
</table>

Table 3. Distribution of terms found in different locations.

The only medical corpus that has been used for translation and analysis is Korp (Språkbanken), which includes a corpus built from text appearing in Läkartidningen, a periodical for medical professionals, while the term banks used are MeSH and SNOMED CT.\(^\text{11}\) Läkartidningen was the basis of the corpus used by Abrahamsson et al. for study material (2014: 57). Moreover, it was possible to find the terms above in a relevant context in the SL (Kletke and Rydell, 2016: 97). The number of terms found in both term banks or dictionaries, as well as the number found within a medical corpus, exceeds the fraction of terms found by Deléger et al. (2010: 154). There are at least four possible explanations for this. The first is that as they are all related to conditions, complications or special aids, the chosen terms are likely to appear both in term banks and in a medical periodical. Terms from these categories are all used in both professional medical language and popularized medical language. The second may be that no medical professionals have validated the TT. Had there been professional validation, the percentage of final choices used in the TT might have been lower. The third reason for the significant difference from Deléger et al. (ibid.) may be the lack of a

\(^{11}\)SNOMED CT is the choice location for terms regarding diagnoses according to the Board of Directors of the AMIA (Awaysheh et al., 2018: 17).
A bilingual medical corpus, parallel (Deléger et al., 2010) or comparable (Morin et al., 2010). The fourth may be that the Swedish medical corpus (Språkbanken) was used after term banks had been consulted to verify terms, not to find them (Deléger et al., 2010: 154).

Conditions, as stated above, are likely to appear both in term banks and corpora. However, had there been a term such as ceterizine (the active ingredient in some allergy medications), it would have been more likely to appear in a term bank, not a medical periodical. In the medical periodical it might appear as one of the brand names used to market and sell it, namely Zyrlex. Having looked at the sources of terms, it is time to investigate the distribution of them.

When terms are not found in term banks or corpora, having access to an expert in the area of investigation is another possible strategy (Chesterman and Wagner, 2014 [2002]: 70). This strategy proved to be beneficial, as can be seen in Section 4.4.1.

4.2.2 Distribution
In the ST, 79 terms belonging to the categories described in Section 2.2.2, Terminology, have been identified (conditions or complications; medical, occupational and educational aids, treatments/procedures) and the distribution of them according is pictured in Figure 2.

The distribution of the terms identified in the ST are as follows: there are 29 compounds of two or more nouns, which are called compound nouns in this thesis. There are eleven nouns consisting of one word, not possible to break down into smaller morphemes, known as simple nouns, and 34 modified nouns. Moreover, there are five potential false friends (Vinay and Darbelnet, 1995: 68), which are also worth bringing to attention, as they can form pitfalls for translators.

---

12 In the US, one of the brand names used for the same active substance is Zyrtec.
13 The name has been assigned for translation purposes, even though this applies to the ST. Pre-modified nouns from the ST may end up being post-modified in the TT.
Of the terms identified, there are fourteen that have identical or near-identical translations; hence, they are not discussed in depth here. Examples of near-identical translation pairs include *karyotype/karyotyp* (ST/TT, 12), *fundoplication/fundoplikation* (ST/TT, 187), and *centromere/centromer* (ST/TT; 51, 59), where the only differences between the SL and TL term consist of adhering to the language-specific spelling conventions. Since the potential *false friends* (Vinay and Darbelnet, 1995: 68) are identical or near-identical as well, they are not considered from a shift standpoint. This leaves 60 terms to analyze from a quantitative perspective.

### 4.3 Translation Strategies

Vinay and Darbelnet (1995: 279) wrote: “The relationships between languages can be classified according [to] the frequency with which oblique translation methods have to be used” (ibid.); if the text material of this thesis is of any indication, the terms analyzed from the source and target language differ considerably.

Of the ST and TT terms identified for analysis in this thesis, 39 out of 79, or 49.4 %, have required translation strategies that sort under *oblique translation* (Vinay and Darbelnet, 1995: 36–41), where *transposition* (ibid., 36), *modulation* (ibid., 36), *équivalence or idiomatic translation* (ibid., 38), *adaptation* (ibid., 39), *loss, gain, compensation* (ibid., 169–170, 198) and *explicitation* (ibid., 170) are all represented. From Figure 3, it is clear that transposition is by far the most common shift encountered among the terms analyzed; no other strategy comes close. The rest of the terms have been possible to solve using *direct translation* (ibid., 88).

---

14The mathematically skilled and/or alert reader will notice that the total number of translation strategies used exceeds 79. This is because terms may undergo more than one shift, which is shown in the qualitative analysis (Section 4.5).
Since it was possible to investigate how the terms chosen for analysis shifted during the translation (which all ended up being nouns), this has been illustrated using a bar chart.

The bar chart in Figure 4 shows that modified nouns occur most commonly in the ST (34), while compound nouns are most common in the TT (31). A slight increase in the number of simple nouns can be seen, from four in the ST to six in the TT. There is also an adjective in the TT, whereas there were no adjectives in the ST terms identified for analysis. The reader should keep in mind that these bar charts include the twenty identical and near-identical translations along with the potential *false friends*. However, the different groups also undergo shifts between the ST and TT, as is shown in Figure 5.
Of the 22 compound nouns in the ST, 16 of them, or 72.7%, are still compound nouns in the TT, but with the difference that they often consist of two or more words in the SL and a single, longer word in the TL. Four of the compound nouns, or 18.2%, have turned into modified nouns, and two of them, or 9.1%, have become simple nouns. Of the simple nouns, three (or 75.0%) have remained in the same category in both the ST and TT, whereas one (or 25.0%) has become a compound noun in the TT. It is in the modified noun group that the most notable difference appears: 19 of them, or 55.9%, are still modified nouns, but as many as 13 of them, or 38.2%, have become compound nouns. One of them has become a simple noun and another has turned into an adjective, accounting for 5.9% together. However, to illustrate what these charts say in more detail, examples from the text are required, which leads to the subject of observed shifts among the terms included in the analysis.

4.4 Examples of Shifts
There were two alternatives when categorizing the shifts: either splitting the segments, thus including only one instance of a shift, or keeping them together, allowing for different observations. The latter choice has been made, meaning that a compound noun will be discussed alongside a modified noun, for example. The first category to be addressed is compound nouns, followed by modified nouns, simple nouns, potential false friends, and adaptation in turn.
4.4.1 Compound Nouns

Of the 22 compound nouns identified in the text, 20 of them, or 90.9%, contain what is known as a *noun attribute* (Ingo, 2007: 187), meaning that the first noun modifies the second. This is the case in (4), where there are two examples of them:

(4) Like most other chromosome disorders, having parts of chromosome 10 missing may increase the risk of *birth defects*, developmental delay and *learning difficulties*.

De flesta kromosomrubbningar kan medföra en ökad risk för *medfödd missbildning* samt fysisk och *intellektuell funktionsnedsättning*.

Both compound nouns turn into modified nouns in this example, and the noun modifiers both turn into adjectives, so there is *transposition* (Vinay and Darbelnet, 1995: 36) involved on the word level. As for the word *defect*, it was *borrowed* (Vinay and Darbelnet, 1995: 31–32) into Middle English from the Latin *défectus*, which means ‘failure, absence, lack, weakness’ (Merriam-Webster, 2019). *Learning difficulties* of the ST also turns into a modified compound noun in the TT, but the most notable shift is the transposition of the noun modifier.

In (5), there are two compound nouns, one which is a potential *false friend* (Vinay and Darbelnet, 1995: 68) and another which has a potentially ambiguous translation. The third term, a simple noun, is mentioned here as well, since it belongs to the same segment.

(5) Development: *hand-eye co-ordination* and *dexterity* (fine motor skills) and *self care*  

(ST, 303)  

Utveckling: *öga-hand-koordination* och *manuell färdighet* (finmotorik) samt *personlig vård*  

(TT, 303)

The potential *false friend* (Vinay and Darbelnet, 1995: 68) is *hand-eye co-ordination*, which would be literally translated as *hand-öga-koordination*. However, a visit to *Psykologilexikon* (Egidius, 2019) shows that the term is *öga-hand-koordination*. It is possible to use the same order as in English as in Swedish (Språkbanken). However, Kletke and Rydell (2016: 97) recommend using the preferred term, which is exactly what has been done. The term shifts slightly from being a premodified compound to being a compound.

As for *dexterity*, *dexter* means ‘right’ or ‘right side’ (Merriam-Webster, 2019; Malmquist and Lundh, 2016: 76), which becomes a noun by adding the suffix (Ingo, 2007: 107) -*ity*, which means ‘quality, state, degree’ (Merriam-Webster, 2019); the corresponding suffix in Swedish is -*itet* with the same meaning (Mattsson, Wendt, and Rosqvist, 2019). The *symbol* (Ogden and Richards, 1923: 11) of *dexter* invokes the *thought* and in turn *referent* of right (side) to the reader (ibid.). The first known use of the word is indeed *adroitness*, ‘mental skill or quickness’ (Merriam-Webster, 2019)
with *droit* from the French meaning ‘right’ and the noun suffix -ness means ‘state, condition, quality, degree’, much like -ity. Egidius (2019) lists the meaning *händighet*; however, this could be misunderstood as someone working with home improvement. SNOMED CT has *manuell färdighet* instead. Loss, gain and compensation (Vinay and Darbelnet, 1995: 169–170, 198) are involved. While *färdighet* shows loss, the premodifier *manuell* adds gain and compensation at the same point in the TT.

For the translation of *self care*, Egidius (2019) lists *egenvård*, the expression could have been considered a potential false friend (Vinay and Darbelnet, 1995: 68). This is a term which means both performing health care tasks and tasks that prevent the need for health care (Karolinska Institutet, 2019). Tasks such as brushing teeth and getting dressed, while preventing deteriorating dental health, for example, are hardly health care tasks, though; however, both sort under *personlig vård*. SNOMED lists *personlig vård* (SNOMED CT), although the English version is *personal care*. However, from the English and Swedish descriptions at SNOMED (ibid.), it is apparent that the symbol (Ogden and Richards, 1923: 11) invokes the same referent.

Sometimes, the selection of terms (Lušicky and Wissik 2015: 20) involves the use of dictionaries and corpora to make the final decision regarding the appropriate choice, along with research in both languages (ibid., 20–21). Although Egidius (2019) has been the prime source of checking terms related to growth and development, some of them have been potential or actual false friends. In the case of *self care/egenvård* discussed in (4), the tasks described are more related to *personlig vård*, ‘personal care’ than *egenvård*, ‘self care’. *Personlig vård* may be found in Korp, and its context is relevant (Språkbanken; Kletke and Rydell, 2016: 97).

One of the segments deals with orthopedic technology and contains five terms, all appearing in (6):

(6) Some children need support (such as a standing frame, walking frame, support boots, a supportive Lycra ‘second skin’ and/or leg braces) while learning to walk. (ST, 289)

Apart from the orthopedic expertise necessary to confirm the translation (McShane, 2019), which is known as *collecting resources* (Lušicky and Wissik, 2015: 15), SNOMED CT (2019) can verify the translations used, since there are entries for *gånghjälpmedel* and ortoser. The Swedish term, ortoser, is also worth mentioning. It comes from the Greek *orthosis*, meaning “straightening” (Lindskog, 2016: 450). This has been translated using *borrowing* (Vinay and Darbelnet, 1995: 31–32), which means that the SL word is transferred to the TL. However, the English term, *leg braces*, uses the
strategy of \textit{équivalence} (Vinay and Darbelnet, 1995: 38–39), in this case applied to describing an object rather than a situation.

Another compound noun (though in a single word) that shows \textit{équivalence} is found in (7):

\begin{itemize}
  \item \begin{itemize}
    \item (7) A small number of children have been reported with \textit{hyperreflexia}; overactive reflexes which can result in twitching and can also affect fine motor skills.
    \end{itemize}
\end{itemize}

\begin{itemize}
  \item Några få barn har också visat sig ha \textit{onormala} eller \textit{överdrivna reflexer}. Det kan resultera i ryckningar och även påverka finmotoriken. (TT, 303)
\end{itemize}

Hyperreflexia is built from a prefix that is Greek in origin, \textit{hyper-}, meaning ‘over’ (Lindskog, 2016: 271) and Latin \textit{reflexia}, meaning ‘to bend back’ (ibid., 273). Merriam-Webster (2019) also lists the prefix \textit{hyper-}, meaning ‘excessive, over-, above, beyond’ (ibid.) and confirms the Greek etymology, via Latin (ibid.). MeSH lists \textit{onormala reflexer} (Karolinska Institutet, 2019) as the preferred term, which is an \textit{idiomatic translation} (Vinay and Darbelnet, 1995: 38–39). This could be misinterpreted as slow or non-functioning reflexes, especially since the prefix \textit{hyper} is not present in the TT, so the \textit{overactive} from the ST has to be present by way of \textit{compensation} (Vinay and Darbelnet, 1995: 198) in the TT.

The term formation in Swedish comes from the \textit{creation of a phrase} (Ingo, 2007: 208), albeit a short noun phrase with the head \textit{reflexer}, ‘reflexes’ with the adjective modifier(s) of \textit{onormala} ‘abnormal’ or \textit{överdrivna} ‘exaggerated’.

\subsection*{4.4.2 Modified Nouns}

A special case of the modification of nouns, namely premodification (Ingo 2007: 188) is a common occurrence in English. However, only one of them is hyphenated (ibid., 187), so it is not discussed in detail. In (8), a premodifier in the ST has generated a shift in the TT.

\begin{itemize}
  \item \begin{itemize}
    \item (8) Children with a 10q25/6 deletion are typically slow to reach their \textit{developmental motor milestones}.
    \end{itemize}
\end{itemize}

\begin{itemize}
  \item För de flesta barn med en 10q25/26-deletion är den \textit{motoriska utvecklingen} långsam. (TT, 303)
\end{itemize}

The term \textit{motorik}, ‘motor function’, is available (Egidius, 2019); it is part of the explanation of other terms related to fine and gross motor functions (ibid.). The modifier \textit{developmental} is an inflicted adjective from the noun \textit{development} (ibid.). There are also \textit{developmental milestones} (SNOMED CT). Hence, \textit{developmental motor milestones} can be considered a term.
While the literal translation of *developmental motor milestones* is *milstolpar i den motoriska utvecklingen*, two other translation strategies appear here; the first is *transposition* (Vinay and Darbelnet, 1995: 36). The second strategy is *loss* (ibid., 169–170), in this case omitting milestones, but there is no *compensation* (ibid., 198). There is another example of a shift in (9):

(9) Many children walk with a wide gait and trip easily... (ST, 291)    Många barn går bredbent och har lätt för att snubbla... (TT, 291)

According to MeSH (Karolinska Institutet, 2019), *gait* means ‘the way of walking’. A *literal translation* (Vinay and Darbelnet, 1995: 33–35) would be *bred gång* (Linguee), which is not idiomatic; therefore, this term is subject to a shift. Here, *transposition* (Vinay and Darbelnet, 1995: 36) has been used; the shift from a premodified noun to an adjective takes place to form the word *bredbent*. Mattsson, Wendt, and Rosqvist (2019) has *gå bredbent*, literally ‘walking with wide legs’; that is, walking with the legs apart, or a wide gait.

Normally, a term should carry a single meaning in the translation (Kletke and Rydell, 2016: 98). From this it is possible to derive that a single term should be used in the TT. In one case, that principle has been abandoned, namely for *genetic material*. In (10), the first occurrence of the term fits the definition found in MeSH, *gener* (Karolinska Institutet, 2019).

(10) A 10q25 or 10q26 deletion means that the cells of the body have a small but variable amount of *genetic material* missing from one of their 46 chromosomes – chromosome 10. (ST, 14)    En 10q25- eller 10q26-deletion betyder att cellerna i kroppen har en liten men varierande mängd *gener* som saknas från en av deras 46 kromosomer: kromosom nr 10. (TT, 14)

The term *gener* displays *loss* (Vinay and Darbelnet, 1995: 169–170) without *compensation* (ibid., 198). The very next ST segment, shown in (11), used the same term. However, when *genetic material* was used in the specific meaning of DNA, there is the problem of a mismatch with the ST:

(11) For healthy development, chromosomes should contain just the right amount of *genetic material* (DNA) – not too much and not too little. (ST, 15)    För att en människa ska utvecklas normalt behöver kromosomerna innehålla precis rätt mängd *genmaterial* (DNA) – inte för mycket och inte för lite. (TT, 15)

A combined approach, used by Deléger et al (2010: 154), shows that *genmaterial* is a possible choice (SNOMED CT). In this particular segment, there is a *literal translation* (Vinay and Darbelnet, 1995: 33–35); in all other instances in the text, *loss* (Vinay and Darbelnet, 1995: 169–170) without
any gain or compensation (ibid., 198) is used. The very same term will represent different kinds of medical language in (12).

(12) Every chromosome contains hundreds to thousands of genes which may be thought of as individual instruction booklets (or recipes) that contain all the genetic information telling the body how to develop, grow and function.  
    (ST, 20)

Varje kromosom innehåller hundratals till tusentals gener som kan ses som specifika instruktioner (eller recept) och innehåller all genetisk information som säger till kroppen hur den ska utvecklas, växa och fungera.  
    (TT, 20)

Here, genes are compared to a book that both contains instructions and reads them out loud, with the purpose of explaining a complex concept and bringing it down to the level of anyone in general, and a parent under stress in particular (Andersen, 2019). In this example, it is possible to see how more than one kind of medical language (Nyman, 1996: 16) is represented. The term pair genes/gener represent scientific medical language, while which may be... instruction booklets/som kan ses som... instruktioner represents generic medical language, intended to explain the term.

There is, however, no colloquial medical language (ibid.).

Another premodified noun, medical literature, undergoes a shift. This is an expression that appears throughout the text; one instance of it is shown in (13):

(13) The information in this leaflet is drawn partly from the published medical literature.  
    (ST, 4)

Informationen i den här broschyren är delvis hämtad från den medicinska fackliteraturen.  
    (TT, 4)

The ST segment after the one shown in (13) states that medical literature means articles. Published is probably there to tell readers that the articles have gone through all the steps required for publication and are not subject to modification at the time of publication of the leaflet. The literal translation (Vinay and Darbelnet, 1995: 33–35) would be publicerad medicinsk litteratur. Here, the prefix fack- has been added, instead showing the strategy of gain (Vinay and Darbelnet, 1995: 170), combined with loss (Vinay and Darbelnet, 1995: 169–170). Moreover, there is partial transposition (Vinay and Darbelnet, 1995: 36), since it adds another adjective prefix in the TT to a noun that has none in the ST.

In (14), there is an example of synonym replacement:

(14) One child in the medical literature was reported to have hypoplastic optic discs...  
    (ST, 357)

Ett barn hade förminskade synnervspapiller...  
    (TT, 357)
Merriam-Webster (2019) defines *hypoplasia* as “a condition of arrested development in which an organ or part remains below the normal size or in an immature state” (ibid.). Hypoplastic is the inflected adjective (ibid.). Lindskog (2016: 276) defines *hypoplasi* as ‘incomplete or arrested development of a body part, not completely developed, shrinkage’. Förminskade, ‘shrunken’, conveys the same *sense* (Frege, 1948 [1897]: 209) as *hypoplastisk*, but the *reference* (ibid.) might be seen as slightly different to some, since förminskad could also mean that the modifier describes something that was once ‘full-sized’ but is not.

For the sake of comparison, here is an instance of a similar construction which has no synonym replacements but a modulation:

(15) ...a *hypoplastic left heart*... (ST, 335)  
    *hypoplastisk vänsterkammare*... (TT, 335)

In (15), *hypoplastic* has not been replaced by a TT synonym. However, there is *modulation* from the whole to the part (Vinay and Darbelnet, 1995: 36), because *vänsterkammare*, ‘left chamber’ or *left ventricle* (Merriam-Webster, 2019) is not all of the *left heart*; there is *vänster förmak*, ‘left atrium’ (Merriam-Webster) as well.

4.4.3  Simple Nouns

Shifts are also encountered among the simple nouns, alongside other translational challenges.

Consider (16):

(16) Seizures are not a major feature of 10q25/6 deletions. (ST, 343)  
    Anfall (kramper) är inte vanligt förekommande bland de som har 10q25/26-deletioner. (TT, 343)

A *seizure* could imply an epileptic or a *febrile seizure*, which translates to *feberkramp* (Karolinska Institutet, 2019), which means that in this case, there is *polysemy* (Peters, Qian, and Ding, 2017: 102) or multiple meanings. This segment also adds explicit information that is only implied in the ST, making i an example of *explicitation* (Vinay and Darbelnet, 1995: 170). In (17), a simple noun with no obvious definitions has undergone a shift on the phrase level:

(17) He currently knows about 150 signs and his *vocalisation* is improving a lot  
    (ST, 265)  
    Han kan ungefär 150 tecken just nu och han gör stora framsteg med *ljudningen* (TT, 265)

Neither MeSH nor SNOMED CT has any translation for *vocalisation*, but Merriam-Webster (2019) lists *vocalize* as an intransitive verb meaning ‘to utter vocal sounds’ (ibid.) Moreover, MeSH (Karolinska Institutet, 2019) lists *Animal Vocalisation* as *Djurens ljud*. The Swedish verb *ljuda*
means ‘to make a sound’ (Mattsson, Wendt, and Rosqvist, 2019). The nominalized version of it is *ljudning* (ibid.). Using this term has caused a shift on the phrase level. A noun phrase followed by a verb phrase (ST) is replaced by a noun phrase followed by a verb phrase and a preposition phrase (TT). There is *transposition* on the phrase level and also *modulation* (Vinay and Darbelnet, 1995: 36); the literal translation would have been *hans ljudning förbättras mycket*. This particular *modulation* (ibid.) is a perspective shift from the passive voice in the ST to the active voice in the TT. Korp verifies the use of *ljudning* (Språkbanken).

### 4.4.4 Potential False Friends

False friends have been discussed as pitfalls for translators, and in (18), there is a potential *false friend* present:

(18) More rarely, and not seen in any Unique children, complex heart conditions have been found including Fallot’s *tetralogy*... (ST, 335)

Det är ovanligt att man ser komplicerade hjärtfel som behandlas kirurgiskt, såsom Fallots *tetradi*... (TT, 265)

The potential false friend is *tetralogi*, known to mean ‘a sequence of four’ in Swedish (Mattsson, Wendt and Rosqvist, 2019). Here, however, *tetradi*, ‘a set of four’, is listed as the preferred choice in technical, or medical, language (ibid.). MeSH confirms this (Karolinska Institutet, 2019). The presence of several terms for a similar *reference* (Ogden and Richards, 1923: 11), a set or sequence of four, might cause problems for a translator.

### 4.4.5 Adaptation

Finally, one of two examples of *adaptation* (Vinay and Darbelnet, 1995: 39–40) is included. It is discussed because part of translation strategy is to adhere to the style of the original text (ibid., 17):

(19) “He still finds certain sounds difficult but he can be understood and ‘talks the hind legs off a donkey’!” (ST, 183)

”Han har fortfarande svårt med vissa ljud, men det går att förstå honom och han är en riktig praktvärn!” (TT, 183)

Borrowing would not be a recommended working strategy, as Ndwayamato Madadzhe and Mashamba (2014: 54) have pointed out. Translating this phrase requires knowing that it is an *idiom*, or the result might be quite confusing. This was shown by Ndwayamato Madadzhe and Mashamba (2014: 53), where a snake was used in Tshivenda to convey the meaning of a woman not being able to conceive. The idiom here, however, *talking the hind legs off a donkey*, can mean ‘to talk for a very long time’ (Merriam-Webster, 2019) or to ‘talk incessantly’ (Lexico, 2019). The Swedish
expression *pratkvarn* (Mattsson, Wendt, and Rosqvist, 2019), literally ‘talk windmill’, describes a similar situation. An alternate strategy is to use *pratmakare* (ibid.), ‘talkwright’, but readability also needs to be considered, where *pratkvarn* has two syllables and *pratmakare* four, making it sort under polysyllabic words. Polysyllabic words are mentioned as a factor that affects readability negatively in Section 3.1.2 (Hill-Briggs et al., 2012: 294).

4.5 Summary of Findings

Four out of seven readability yardsticks used in this chapter take the length of words into account over the length of sentences, while the remaining three are constructed to measure the length of sentences. Readability can be deemed more difficult when a text is rich in terms, which tend to be polysyllabic. There is no clear consensus on which method to use, and the professional medical translator/interpreter recommendations do not include any information on which method is to be preferred, either. It is possible that due to the nature of the SL and the TL, respectively, different formulas should be used, as there are unexpected effects observed: this thesis contradicts both bilingual studies cited here (Kolahi and Shirvani, 2012; Acar and İşisag, 2017). The bilingual studies in question have used the same method, or an adaptation of the same method, to determine the readability level of both the ST and the TT, however. The average sentence length of the TT is considerably shorter than that of the ST (-20%), but a 4% increase in the average length of words has affected the methods concerned with choosing difficult words considerably. For the text material of this thesis, the style of both the ST and TT may make the text more difficult to read, even if sentences are the element of focus, as sentences with terms along with their corresponding layman expressions tend to be longer. The number of terms found in term banks is only slightly higher than that of the TL medical corpora, but the latter has been used for verification only, and the studies on term banks and corpora (Deléger et al, 2010; Morin et al., 2010) have both used corpora, parallel or comparable, to source terms, not verify them. The greatest number of shifts encountered can be attributed to *transposition*, although *direct translation* is the most common strategy used. This may well be coupled to the fact that the percentage of terms found exceeds that of Deléger et al (2010). The two instances of *adaptation*, of which one appears, have required considering the pitfalls of direct translation (Ndwayamato Madadzhe and Mashamba, 2014).
5 Conclusion

Early on in this thesis, it was established that the aim was to investigate readability, terminology and translation strategies in a medical text. The focus areas were how the readability level compares to similar medical texts, what methods were used, how they performed, and how terms were sourced. Finally, translation shifts encountered among terms were investigated.

5.1 Readability

The results of the investigations made show that neither of the readability levels of the ST or the TT are ideal, although the ST level compares favorably to those found in other readability studies. Methods used were some of, or all of, the following formulas: Flesch, Flesch-Kincaid, Fog Index, Coleman-Liau, SMOG, ARI, and the Linsear Write formula (My Byline Media, 2019). Clearly, there is dissent on which formula to use, even within the field of medical information. While Acar and Işisağ found improved readability in the TT using an adaptation of the ST yardstick for the TT, Kolahi and Shirvani found the readability levels to be more difficult for their TT using the Fog Index. While a parameter mentioned in professional guidelines needs to be taken into account, the presence of polysyllabic medical terms requires a formula that is suited to such conditions. Considering the readability outcomes for the text material of this thesis, very few, if any, of the readability formulas used can be recommended, because the TT does not necessarily use more complex words; the words simply tend to contain more syllables, which is part of the nature of the TL. One of the reasons why the IMIA does not recommend a specific formula for determining the readability level may be found in the observed differences between texts in two languages. As Loughran and McDonald (2014) have pointed out, longer words may not necessarily be the root cause of a text being hard to read, confusing or even impossible to comprehend. In the source material, the terms present are explained in layman-friendly language. Yet, readability formulas claim that the text readability level is too complex for the target audience according to the professional recommendations. Gunning (1969: 11), the creator of the Fog Index, has written that there is no formula that can guarantee good writing (ibid.), and good writing does not necessarily equal good readability. Adhering to the tone of voice and applying target reader review, namely addressing the issue of comprehensibility, would both be better indicators of a successful translation than readability.
5.2 Terminology
A common issue with terms is knowing where to find them. A medical translator should search all available sources of terms: term banks; corpora – parallel or comparable, but firstly bilingual ones; specialized, and sometimes even general, dictionaries. When those sources fail to produce a term, it can be helpful to search a related discipline for more information, which has been of some help during the translation of the text material. However, sometimes it is even better to find and consult an expert on the subject. To scrutinize all sources critically is a must for avoiding translator pitfalls, such as potential false friends, which were encountered during the analysis. The context of terms in both the SL and TL definitely need investigating to make sure that the translation does not confuse the reader.

5.3 Translation Strategies
The investigation of shifts among terms has revealed that literal translation is the most common translation strategy, but that transposition is the most common type of translation shift among the terms investigated. The second and third most common oblique translation strategies are gain and modulation. Gain may well be caused by translators favoring explicitation. Since incorrect information may lead to adverse medical effects, loss is not a favored strategy in a medical publication; in fact, a beginner in translation may even be afraid of loss and avoid it at all costs for this very reason. Literal translation has been identified as a potential problem when there are cultural aspects involved, but this is a method that scholars warn against using when constructing new terms. If there is an established term which happens to use literal translation, consider the recommendation of using the preferred term; it is sound. However, when cultural aspects surface, such as in the case of idioms, a translator should be watchful of using literal translation and instead look into the oblique translation strategy of adaptation.

5.4 Further Work
As the readability investigation has not shown any formula to be superior, comparing more than one guide would be helpful in further research. A total of 199 guides to date remain to be translated into Swedish for Unique. As for further investigations regarding terms, this thesis has investigated philosophy and shifts over word formation. In texts where there are very few terms to be found, word formation would be a relevant area of investigation. One suggestion of combining terminology and readability (or even comprehensibility) is to work with the medical language, make it fully colloquial so that the text creates a more uniform impression, and create a section of terms at the end of the guide.
One method of translation not present in this thesis was the use of combined bilingual and medical corpora, which might have been helpful to make up for the scant availability of parallel texts. Had it been possible to access them, bilingual specialized corpora would also have been a source of idiomatic phrases around the terms. However, term banks have performed according to expectation from the previous combined study, as the highest percentage of terms for this thesis came from term banks. The percentage of terms identified from a medical corpus is higher than the combined study. This may be due to the TL corpus used in this thesis helping to verify terms rather than to identify them. Another possible explanation is that the TT did not undergo expert revision before being finalized for this thesis. However, terms, like any other type of word, are likely to appear with standard collocations, and corpora would be helpful in that respect, too, especially if they are bilingual.
Sources

Primary

Secondary


Andersen, C. (Unique). 2019. Re Maj Hultén. [Email; contents shared with permission.]


McShane, P. (Svenska Ortopedtekniska Laboratoriet). 2019. Översättning. [Email; contents shared with permission.]


**Online Dictionaries/Term Banks**


Parallel texts and Corpora


Appendix A: Investigation of The Source Text

This appendix contains background information on the source text and an analysis, both of which have led to decisions regarding the translation strategy, but neither are used in the theoretical section.

Background Information
The Unique leaflet that forms the source text of this thesis (Unique, 2013) was written to overcome the problem of either non-specific or vague information being offered to families of a child born with a 10q25/26 deletion (Andersen, 2019). Other reasons why the leaflet was written include to relieve stress, to possibly aid recall, and to be kept for later as a source of tips (ibid.). Families are often eager to share their stories in order to find research studies in which they can participate, and they value the stories from other families (ibid.). Having ready access to medical information and feeling informed also adds a sense of feeling included (Gerard et al., 2017: 2). The opposite, namely not having the medical explanations needed serves as a possible stressor (Hulsman, Smets, Karemaker, and de Haes, 2011: 421).

A leaflet is also possible to read at a time when the reader is open to processing the information (Andersen, 2019). Comments from parents who have already been through the stages the reader has yet to experience serve as information from patients (Laughey, Sangvik Grandal, and Finn, 2018: 664). While a leaflet in itself is no substitute for an encouraging nod from a physician (ibid., 667), the effect of reading a hopeful testimonial cannot be underestimated. This leads to the concept of context.

Context
According to Hellspong and Ledin (1997: 49), readers can only understand the purpose of a text when they know under what conditions it has been produced, which is why the context (ibid.) of a text is important.

First, there is the situational context (Hellspong and Ledin, 1997: 50), meaning who is the sender (ibid., 53), namely Unique, and its receiver (ibid., 52). The leaflet has many receivers. Parents of children with 10q25/26 deletions make up the primary target group. A secondary target group is preschool/school teachers, who tend to get questions from other parents in the child’s group or class when the child is absent for surgery and rehabilitation, for example. Teachers also need to be aware of at least the learning difficulties children with 10q25/10q26 deletions face (Unique, 2013), not to mention possible personality traits, the need for peace and quiet, and how to deal with the reactions of other children in the same group/class to deviating appearance of children with 10q25/10q26 deletions.
deletions. In short, teachers benefit from reading the brochure because they might understand their pupils and their possible challenges better. Second, the *inter-textual context* (Hellspong and Ledin, 1997: 53) determines the relationship of the text to other texts and the conventions it must follow; in this case, the ST is a compilation of findings from medical literature and survey results (Unique, 2013).

Unique is a charity organization; as such, it depends on grants and donations, and chromosome disorders are not a common occurrence. Therefore, Unique identifies a “critical mass of families with a disorder” (Andersen, 2019) before considering the compilation of a guide, where medical literature is used and affected families are a part of the process by participating in surveys and providing comments before publication (ibid.). The 10q25/10q26 leaflet mentions that 69 members of Unique have a pure 10q25/10q26 deletion (Unique, 2013: 2); these members are located worldwide.

In April of 2019, there were close to 200 guides concerning disorders of chromosomes (Unique, 2019), and some of these had been translated to mostly Dutch, French, German, and Russian. There were two Danish translations of guides, but none in Swedish at all. The cost of translating these to major languages would take resources from grants and donations; therefore, organizations such as Unique rely on voluntary work. No guides were available in Swedish during the production of this thesis.

Third, the *cultural context* (Hellspong and Ledin, 1997: 58) is hard to define; Unique is based in Great Britain but the leaflet is aimed at families around the world, due to the rare nature of 10q25/26 deletions (Unique, 2013). The text tends to avoid making cultural references.

With the context clear, the source text is now ready for analysis.

**Source Text Analysis**

First, there is need for clarification on what non-fiction language is, because it is not a uniform concept. Nuopponen and Pilke (2016: 59) write that non-fictional communication can be classified according to the following:

<table>
<thead>
<tr>
<th>Sender</th>
<th>Receiver</th>
<th>Manner of Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layman</td>
<td>Layman</td>
<td>Generic language</td>
</tr>
<tr>
<td>Professional</td>
<td>Layman</td>
<td>Popularized language</td>
</tr>
<tr>
<td>Professional</td>
<td>Professional</td>
<td>Jargon</td>
</tr>
</tbody>
</table>

*Table 4. Different kinds of non-fiction communication.*
The leaflet on 10q25/26 deletions does not use the language from simply one section of the communication model in Table 4. The safe assumption to make is that professionals are communicating with laymen, as there are bound to be more parents among geneticists than geneticists among parents. In order to produce leaflets, medical literature is researched (Unique, 2013); terms from said literature are kept so that parents who wish to research the subject by reading scientific articles will recognize them (ibid.). Hence, the leaflet contains both generic medical language and scientific medical language or even jargon.

The *perspective* of the text (Hellspong and Ledin, 1997: 156) is tied to Unique as an organization, but Unique adds to findings of medical literature with information from surveys to show the bigger picture: “evidence at Unique does not support this...” (ST, 352). This particular example also shows *metatextuality* (Hellspong and Ledin, 1997: 193), indicating a critical standpoint from experience. However, the text does not contain any warnings (ibid., 193), and the text does not argue a standpoint. There are, however, reservations: “Every person with a 10q25/6 deletion is unique and so each person will have different medical and developmental concerns” (ST, 192).