# STUDENTS' DATA MINING SKILLS IN SECOND-YEAR UNDERGRADUATE TRANSLATION

Mari Pakkala-Weckström

University of Helsinki

## Abstract

This paper discusses the different research methods that second-year undergraduate translation students employ for their homework, and particular attention is paid to the relationship between the genre and function of the source text, and the variety of sources used by the translators. The material of the study consists of the students' own work reports (or translation commentaries) on five homework assignments representing different genres: a tourist brochure, a popularized science column, two food recipes, an EU report, and a page from an art history textbook. The results strongly suggest that in their second year of translator training, the students have already acquired rather versatile data mining skills that depend on the type of assignment they are working on. While the Internet was clearly the most important source of information, other sources were used quite creatively as well, and in their use of Internet sites, the students showed they were able to do selective and precise searches.

Key words: translator training, translation competence, data mining skills

# 1. INTRODUCTION

The aim of this paper is to trace the various methods and sources that university translation students in an undergraduate program use for research or data mining for their homework translation assignments. I am particularly interested in the relationship between the genre and function – or *skopos* (Vermeer 1989, p. 173–174) – of the source text, and the sources and search methods that the students have reported using in producing the corresponding target text.

From the point of view of translation, genre is a key element. To quote Nord (1997, p. 54): "If a target text is to be acceptable as a representative of a target-culture genre, the translator has to be familiar with the conventions that the target text is to conform to." This, for the translator trainee, means that the translation process should start with conducting a translation-oriented analysis of the source text (see Schäffner 2002, p. 5; also Nord 2005, p. 1). The analysis involves decisions about searching for information in different ways and from different sources, and developing this skill in their trainees is one of the key areas translation teachers aim to address when planning their courses.

The data for this study was gathered within the subject of English Translation at the University of Helsinki in Finland during a second year English–Finnish translation course, (2

h/week × 14 weeks) in 2012.[1] The course is a continuation of the first-year translation course, where the students were taught the basics of translation competences, including information mining and its documentation. During the course under scrutiny in this paper, the students individually translated ten different texts of varying length as homework,[2] and along with their translations, they submitted reflective commentaries on the translation process. The translation commentaries are a compulsory part of every homework assignment; the students are instructed to briefly analyze the source text in terms of genre, type, and function, and to document the sources and methods they have used in the translation process, and to reflect upon the most challenging parts of their work (see e.g. Fox 2002; Pakkala-Weckström 2009 & 2011; Orlando 2011 & 2012). The students usually had a whole week to complete the homework: the assignments were given in class, and uploaded to Moodle at the beginning of the following class.

The texts translated during the course represent a variety of genres and text types, since the course (Professional Translation from English into Finnish II), as mentioned above, was the second of two general translation courses which precede more

---

[1] The syllabus structure of English Translation underwent a major change in 2013. The Professional Translation II from English to Finnish was replaced by a more specialized course titled 'Academic texts'. This course comprises a more limited generic variety of homework, which makes it somewhat less suitable for a study of this kind.

[2] In-class assignments were also included in the course structure, so the students were not assigned homework every week.

specialized, i.e. LSP translation courses (cf. Tagnin 1995). All assignments for the course were chosen so that they would develop certain skills and include various challenges[3] for the trainees – both structural, to develop linguistic skills, and lexical, to develop search skills. In this study, the focus is on the latter. At the end of the course, the students were evaluated based on portfolios containing all their assignments with the teacher's comments, their own commentaries, and a self-review of the course and their development during it (cf. Johnson 2003; Kelly 2005; Pakkala-Weckström 2011, Eskelinen & Pakkala-Weckström 2015).

For the purposes of this study, I have chosen five texts given as homework assignments, since including all ten assignments completed by the students would not have been possible considering the scope of this paper. The texts have been chosen so that, in terms of genre, they offer the widest possible range of the course material. I shall analyze the data both quantitatively and qualitatively, dividing the types of sources the students have used into several categories (e.g. dictionaries, Internet sources, personal contacts, parallel texts), and also by comparing which particular texts have benefitted most from which sources and methods. My research hypothesis is that translating certain genres would seem to benefit from using a wide variety of

---

[3] I use the term 'challenge' instead of 'problem' (see Nord 1997, p. 64) intentionally, however, the reasoning behind this choice, based on several years of experience as a translator trainer, falls out of the scope of this paper.

sources more than others, and that for particularly specialized genres, it is more beneficial to dig deeply into a narrower array of information. Obviously, my data here is limited as to the number of students as well as translated texts, so no universal conclusions can be drawn from it.

I shall first discuss the concept of translation competence, specifically from the point of view of the importance of developing the students' data mining skills. I shall then proceed to give a description of the five homework assignments discussed in this paper, and the challenges related to each of the texts. The analysis of the students' translation commentaries will be presented next, followed by the conclusions.

## 2. THE SIGNIFICANCE OF DEVELOPING STUDENTS' DATA MINING SKILLS

Translation competence is a multifaceted concept; it can be defined, for example, as a combination of the following areas: language competence, textual competence, subject competence, cultural competence, and transfer competence (Neubert 2002, p. 6; for a slightly different description see Beeby 2002, p. 186–187). Another model of competence is suggested by Kelly (2005, p. 32–33); she proposes a list of seven competence areas:

● Communicative and textual competence

● Cultural and intercultural competence

143

- Subject area competence

- Professional and instrumental competence

- Attitudinal or psycho-physiological competence

- Interpersonal competence

- Strategic competence

In this paper, the focus is on the development of data mining or research skills. In Neubert's classification, these belong under the heading of *subject competence*; he states that "translators don't know everything and they need not know everything but they must know where to look for it and where to find it" (2002, p. 9). In Kelly's model, these skills are included in the area of *professional and instrumental competence* as the "use of documentary resources of all kinds, terminological research, information management for these purposes" (2005, p. 32). The EMT (2009) model ("wheel of competence") includes a section *information mining*, which comprises the following:

- Knowing how to identify one's information and documentation requirements
- Developing strategies for documentary and terminological research (including approaching experts)
- Knowing how to extract and process relevant information for a given task (documentary, terminological, phraseological information)
- Developing criteria for evaluation vis-à-vis documents accessible on the internet or any other medium, i.e. knowing how to evaluate the reliability of documentary

sources (critical mind)
- Knowing how to use tools and search engines effectively (e.g. terminology software, electronic corpora, electronic dictionaries)
- Mastering the archiving of one's own documents

According to Mackenzie (2004, p. 32), the translator needs to (a) analyze the translation situation, (b) search and research information, (c) produce and evaluate his/her product. In my opinion, among the most crucial competencies for a translator trainee at this stage are the abilities to carry out research and manage information; thus Mackenzie notes that "knowing where to look for information, whom to consult, and how to classify and systemize information sources, i.e. information management, is part and parcel of this competence" (2004, p. 34–35). This kind of information management (cf. also Englund Dimitrova 2002, p. 76) is one of the essential skills that the Professional Translation courses I and II at the University of Helsinki aim at, and the results of this study suggest that the students were going in the right direction.

Even though there does not seem to be any disagreement on the importance of developing adequate research skills as a central part of becoming a professional translator, the question of how students actually achieve this has not been widely discussed. Eskelinen and Garant conducted a survey on students' views concerning various issues on the use of ICT in translator training at the University of Helsinki, and noted, among other things, that the use of the Internet in data mining appears to be making the

use of other sources all but extinct (2011, p. 38). However, their respondents were students participating in two more specialized translation courses (Professional Translation III, divided into LSP groups of e.g. Law and Administration, Technical Translation, Business, Biology and Medicine); the results of the current study indicate that 'real world' sources such as books or people are still useful for certain types of translation assignments.

# 3. THE HOMEWORK ASSIGNMENTS AND THEIR PARTICULAR CHALLENGES

In this study, I am focusing on reported data mining documentation in the commentaries on the five homework translation assignments selected for the purposes of this paper (see 1 above). Altogether, the material gathered consisted of 95 translations and translation commentaries by 19 students, all native speakers of Finnish translating from their B-language. Not every translation assignment was handed in by every student, so the amount of translations and commentaries per assignment varies slightly. In the commentaries of 32 translations, sources were not clearly listed, despite the instructions to do so, making the total number of commentaries analyzed here 63.

As mentioned above, the students of English Translation at the University of Helsinki are taught, from the very beginning of their studies, that one of the central competences of a translator

is his/her ability to find information that is both reliable and relevant to the task at hand. To encourage the students in their research for each homework translation, they are asked to list their sources in the commentaries they submit with every assignment.[4] Judging by the source material for this study, this seems to work tolerably well, since the commentaries often contain not only the sources but also excuses for not having more of them (most often lack of time). However, in many cases, the students did not provide a comprehensive list; they often left out what could be called routine searches, such as online dictionaries, or Google, and concentrated on what they found to be the most important sources in their process.

Below, I shall introduce the texts in the order in which they were given to the students, and briefly discuss the specific challenges each text presented for the translators with regard to data mining. The texts varied in length but were chosen so as to not exceed ca. 300 words.

## 3.1. Tourist Brochure – Hampton Court

The second assignment during the course was to translate the first 3 pages and the inside of the front cover of a tourist brochure

---

[4] The scope of this paper does not allow for a detailed description of the assessment system for the courses, but the students are aware that the commentaries they compile for each assignment also have an impact on their final grades (cf. Eskelinen & Pakkala-Weckström 2015).

booklet from Hampton Court, London. The booklet was a souvenir from a visit to London some years ago, and at least at the time, though versions in several different languages were available, a Finnish one was not. The brief given to the students was that they should translate the text so as to produce a Finnish version of the brochure to be sold on site at Hampton Court, and that the layout should not be interfered with. The section given as homework contains a welcome letter by the Director of Hampton Court, a Table of Contents, and a summary of the most important sights ("What to see"), introducing six different routes within the Palace buildings, courtyards and cloisters, and the gardens. The overall genre of the text was a combination of a brochure and a tourist guide (cf. Torresi 2010, p. 101), and its function was to give information and instructions for tourists visiting Hampton Court Palace.

The main challenge – or even a translation problem (cf. Nord 1997, p. 64; see also Orozco & Hurtado Albir 2002, p. 380) – from the point of view of data mining in the Hampton Court brochure is that it contains a multitude of proper names and cultural references: names of historical people and periods, names of rooms and places, names of works of art, etc. Some of these names – for example, those of regents and periods – have official Finnish translations, which obviously have to be used, such as *Henry VII – Henrik VII,* or *Baroque – barokki.* In the case of most names, however, informed decisions must be made by the translators (cf. Torresi 2010, p. 104). These decisions must take the genre and function of the text into account: a

brochure which is sold on site to provide information and even directions to the visitor whose command of English is perhaps not very good.

One such challenge was how to treat the name of the series of nine paintings by Andrea Mantegna, given in the source text in English as *Triumphs of Caesar* and displayed at the Lower Orangery at Hampton Court. Since the paintings are of Italian origin and have no official Finnish name, the rule of thumb would be to use the original Italian name with a Finnish translation; however, if the brochure's function is to serve as a guide to help the tourist navigate on site, should not the English name be included as well? And how about the names of the six routes described in the brochure – should the English names be preserved somehow to facilitate navigating them? And if some of the source language is left on the target language version of the brochure, does something else need to be condensed in order to keep the layout intact?

## 3.2. Popular science – A *Scientific American* column

Assignment number 3 was a short column under the heading "Ask the experts" from the popular science magazine *Scientific American* (Issue December 2007). As Byrne (2012, p. 106) puts it, "popular science publications are intended to serve as a bridge between scientists and the general public," and its aim is to "improve the public's understanding of science and to increase

public engagement with science." The question posed to an expert in this issue was "How do short-term memories become long-term memories?" and the brief given to the students was to translate the text as if it would appear in the Finnish popular science magazine *Tiede* (in English, 'Science'). The Finnish magazine has a similar column, although it usually comprises several questions with short answers, whereas the *Scientific American* often only answers one question with more detail. However, the style and register of the columns can be considered to correspond to one another.

One of the main challenges (the others mainly having to do with complex sentence structures) in the translation process had to do with the vocabulary of the column: a multitude of terms related to the brain and its functions. Translating these terms included careful consideration of the (imagined) audience of the target text – i.e. what is the readership of the Finnish magazine? In Finnish, there are often parallel scientific terms, usually a loan word and a translated word, which are then used depending on the addressee; loan words for more selected audiences such as the scientific discourse community, and translated ones for general audiences (see e.g. Pirttimaa 2007). In this particular column, such terms included *neuron,* for example, which can be translated as *neuroni* or *hermosolu* ('nerve cell'). The translators needed to come to a decision as to which direction to go with these terms, i.e. whether the readers of the *Tiede* magazine would prefer to be addressed as specialists or laymen. In their translation commentaries, many students reported having given

150

this particular challenge a lot of thought, and the solutions varied.

## 3.3. Food recipes – the case of two soups

The fourth homework assignment was to translate two soup recipes. The recipes were quite straightforward instructions to make vegetarian soups; however, producing fluent and idiomatic recipes that fulfill the generic expectations of the target language audience is far from being a straightforward translation process.

The text function of recipes is instructive, and to achieve a translation that can meet this function, several things need to be taken into account. First of all, there are measurements to consider: not only do Imperial units need to be transferred into metric ones (in this case, both were given in the ST so it only remained to leave out the irrelevant ones), but a fluent translation also takes into consideration e.g. available package sizes. Some ingredients may not be at all available on the target culture food shelves – the first of the two soups was garnished with double cream which is not even sold in Finland; a Finnish recipe would use sour cream, crème fraîche, or perhaps even whipped cream for the same purpose.

In addition, there may be different conventions both in the actual working methods and in how things are described: e.g. in Finnish recipes the list of ingredients does not usually make a distinction

between chopped or whole ingredients, but the instruction part of the recipe will the start with instructions for chopping the ingredients. In British recipes, on the other hand, the list of ingredients often contains items such as 'onions, finely chopped'[5]. Finally, there can be lexical pitfalls such as the proper translation for *cumin* in the second recipe; it is often confused with Finnish *kumina,* but the correct translation is *jeera* or *roomankumina – kumina* is actually *caraway seed* in English. This is a typical case of 'false friends' (e.g. Baker 2011, p. 22), but as the two spices are quite different in nature, it is important for a translator to recognize this particular pitfall. (cf. Pakkala-Weckström 2014.)

In conclusion, while food recipes may seem quite innocent as translation tasks on the surface, they do often contain some hidden depths. As with any instructive text, the purpose of a food recipe is to help the reader to achieve a certain result, and the language is of secondary importance; however, a translation with actual faults – such as translating *cumin* with *kumina* – will not achieve this goal. A translation with odd or clumsy choices will perhaps not hinder an experienced cook from producing the course he/she is aiming for, but it will probably make them stop and wonder for a while. A fluent food recipe translation will not

---

[5] This is not a subject that, to my knowledge, has been studied systematically; however, both my own and my students' observations after careful analyses of various parallel texts suggest that there is a difference in conventions. In assessing the translations, however, the students were not penalized for leaving the structure of the ingredients section intact.

do this; it will enable the user to concentrate on the actual matter at hand – the cooking.


## 3.4. European Union internal report

As the sixth homework assignment, the students were given an internal customer satisfaction report by the European Commission's Directorate-General for Translation, dated 12/2008. This report was acquired during a Finnish Translation teaching staff visit to the DGT in March 2009. The text was relatively short and structurally straight-forward; however, it was rich with percentages and comparisons: the report was obviously a breakdown of a questionnaire distributed among the users of DGT customers.

For the students, the main challenges regarding this homework assignment involved being systematic in how they expressed all the figures in the text, and in addition to that, familiarizing themselves with typical EU-jargon. In the Translation Studies at the University of Helsinki syllabus, EU-translation is included, so one of the reasons for the inclusion of this particular report was to introduce students to this genre.

## 3.5. Art History textbook

The ninth (i.e. second-but-last) homework assignment was a page from an art history textbook, giving an introduction to the renaissance period. This text was a paper copy originally

received as a hand-me-down from a colleague, so unfortunately, no original source for it can be credited. Like the tourist brochure described above (3.1.), this text contained several names of historical people, places, and works of art. The function – and possibly also the audience – of a textbook, however, is very different from a brochure, and this should be taken into consideration when translating the text. The commission was to translate the text as a test translation for the publisher who would be publishing the textbook series, and was looking for translators for their translation team, so this would add another dimension to the imagined audience.

Although I have included this text in this study because of the versatility of sources the students reported in their commentaries, the challenges related to this particular assignment were not primarily related to data mining; they were mainly stylistic, structural, and lexical. A particularly interesting challenge which went unnoticed by many of the translators, was that of positioning the translation: when a text discussing the Renaissance brings up northern Europe, to which direction should we in Finland be pointing at, as the areas mentioned actually lie south of our country?

Yet another challenge had to do with using one's common sense: if an English text gives the origin of a foreign-based word (*renaissance*) and notes that it is French for 'rebirth,' what should be done when the target language has a direct transfer for the French term (*renessanssi*)? Surely the reference to French

should be left out and just the meaning (i.e. the Finnish calque, *uudelleensyntyminen*) for 'rebirth' given as explanation?

# 4. ANALYSIS: DATA MINING FOR THE ASSIGNMENTS

In this section, I shall briefly discuss the instructions given to the students during the first Professional Translation courses about how to do research on a given topic, i.e. how to develop their data mining skills, and then proceed to analyze their performance as reported in their translation commentaries.

## 4.1. Background for data mining

During the first two translation courses, the students are taught the basics of doing research and data mining: they are pointed to reliable dictionaries and information sites, warned against suspect sources, and against being over-confident – if they are 99% sure about something, that clearly suggests they need to check it anyway. They are also encouraged to browse parallel texts, i.e. texts representing the genre of the source text, but originally written in the target language. Parallel texts will serve as guides to the students in producing idiomatic target texts, but the students will also need to be careful in evaluating their reliability.

Pakkala-Weckström, M. (2015). Student's Data Mining Skills in Second-Year Undergraduate Translation. Current Trends in Translation Teaching and Learning E, 2. 139–170.

## 4.2. Categorizing the sources

I have divided the sources listed in the students' translation reports into two main categories: Internet sources and other sources. These two were then divided into various sub-categories, which will be further detailed below.

### 4.2.1. Internet sources

The Internet sources were divided into five sub-categories: (1) information sites, which comprise library sites, online dictionaries, and for example Kotus, the Institute for the Languages of Finland; (2) Google (used for collocations); (3) Google Images; (4) Wikipedia, and (5) a diverse category I have called 'Other sites'. This last sub-category comprised parallel texts (e.g. for the recipes), official sites (Hampton Court, EU) and any other sites listed in the commentaries.

Table 1: Internet sources

| Assignment | Hampton Court | Memory | Recipe | EU report | Art History |
|---|---|---|---|---|---|
| Information sites | 12 | 9 | 7 | 7 | 5 |
| Google | 4 | 1 | 1 | - | - |
| Google Images | 1 | - | - | - | - |
| Wikipedia | 8 | 6 | 2 | - | 11 |
| Other sites | 14 | 13 | 12 | 4 + 18* | 9 |

\* Official EU-webpages

## 4.2.2. Other sources

The category of 'other sources' covers everything mentioned in the commentaries that lies outside the Internet: visits to the library, browsing one's own or one's parents' bookshelf, and person sources, i.e. contacting experts or calling friends or relatives for information. In 'person sources,' I have also included the instances when the students have reported relying on their own expertise in a given field.

Table 2: Other sources

| Assignment | Hampton Court | Memory | Recipe | EU report | Art History |
|---|---|---|---|---|---|
| Library | 2 | 2 | - | - | 2 |
| Own bookshelf | 1 | 5 | 12 | - | 5 |
| Person sources | 3 | 4 | 6 | - | 2 |

In all the reports, only six library visits were mentioned, and none for the recipe translation or the EU assignment. In contrast, the students did browse their (or their parents') bookshelves quite often: for the Memory text, old high school psychology course books were mentioned five times, and for the recipes, 12 students consulted their own or their parents' cook books. The quote from a student report below demonstrates how this affected the translation process (my translation from Finnish):

> I noticed that in Finnish recipes, the list of ingredients is really just a list of ingredients needed for the dish. So there are no instructions, like there were in the source text. So, I decided to remove the instructions from the list and add them to actual instructions. I also moved the portion size to the very beginning according to the parallel text.

The sub-category of 'person sources' also proved quite interesting: the students reported eight instances where they had contacted friends or relatives for information. The Hampton Court assignment commentaries contained one mention of a student calling their mother who had several London tourist guides. In the Memory text reports, one student had consulted their mother who is a geriatrist, three others mentioned friends who were (1) a psychology student, (2) a doctor, (3) interested in psychology. In the recipe translation reports, one student mentioned calling both their mother and grandmother, both having a lot of experience in cooking, and another had consulted their father for the same reason. The following is a quote from the report of the former (my translation from Finnish):

> In the text, it was instructed to puree the soup in a blender or a food processor. However, my mother and grandmother both agreed that in Finland, one would typically puree something with a hand blender. Since a blender and a food processor are very similar and could be replaced by one another, I replaced the food processor with a hand blender in the translation.

As mentioned above, I also counted the students' personal experience in the category of 'person sources': when translating the Hampton Court brochure, one student reported having benefitted from their own background of working in a museum, another had visited Hampton Court in the past. In the recipe

translation commentaries, three students justified some of their solutions by their own previous experience in cooking, one reported having experimented with the recipes of the assignment. The quote below illustrates the effect of personal experience in the translation process (my translation from Finnish):

> I also replaced the blender with a hand blender since they are more common in Finland, and a food processor is pretty much the same as a blender, and removed the bit about returning the soup to the saucepan, because I thought it was superfluous to a Finnish reader.

## 4.3. Genre and data mining

There seems to be a connection between the genre of the source text and the variety of sources consulted by the translators. In Table 3, the sources listed have been broken down according to text and source category.

Table 3: Genre and data mining

| Text/ source | Info sites | Google | Google Images | Wiki-pedia | Other sites | Lib-rary | Own books | Person sources | Not listed | Total (not include not listed) |
|---|---|---|---|---|---|---|---|---|---|---|
| Hampton Court | 12 | 4 | 1 | 8 | 14 | 2 | 1 | 3 | 5 | 45 |
| Memory | 14 | 1 | - | 6 | 13 | 2 | 5 | 3 | 5 | 44 |
| Recipe | 7 | 1 | - | 2 | 12 | - | 12 | 6 | 6 | 40 |
| EU | 7 | 1 | - | - | 4 + 18* | - | - | - | 6 | 30 |
| Art History | 5 | - | - | 11 | 9 | 2 | 5 | 2 | 10 | 34 |
| Total | 45 | 7 | 1 | 27 | 70 | 6 | 23 | 14 | 32 | 193 |

* Official EU-webpages

In Table 3 above, several interesting observations can be made. First of all, the largest numbers of listed sources, as well as the

lowest numbers of commentaries not listing sources, can be found in the Hampton Court Palace brochure and the Memory text. This might indicate that, at the start of the term (also, having a new teacher and a new course), the students take their homework quite seriously. This is perhaps supported by the largest number of 'no sources indicated' in the reports of the Art History homework, which was the students' second-but-last assignment. These figures, however, may reveal more about the students' tight schedules at the end of the academic year than their actual working methods – but they might, of course, also indicate that towards the end of the course, the students were feeling more confident about data mining. Secondly, the Hampton Court and Memory texts also show the widest variety of searches, both online and in the 'real world'; however, the recipe translation and the Art History texts do not fall far behind.

In contrast, the narrowest array of sources was listed for the EU report; nobody reported consulting Wikipedia, the library, or personal sources, but every student who listed their sources (18/19) mentioned having browsed official EU pages.

Finally, when looking at the total numbers for each of the categories, 'Other sites' is on top of the list with 70 mentions: 18 of these were official EU pages, and the other sites mentioned were usually quite specifically informative: e.g. official EU pages, Hampton Court Palace home pages, or, in the case of the recipe assignment, parallel texts. 'Information sites,' including online dictionaries, received 45 mentions; this figure probably is

not very accurate, since the students, when asked, admitted that they often left dictionary searches out of their list of most important sources, considering using dictionaries more of routine than actual hunting for information. Wikipedia was mentioned 27 times, most often in connection with the Art History text (11) and the Hampton Court brochure (8), both of which – as mentioned above – contained several references to historical people or places, and works of art or literature. In the class discussions on translating these particular texts, it also transpired that the students found Wikipedia helpful in the actual translation process, as in the left-hand bar there is a list of the language versions of the article. If a Finnish version was available, it was used to find out whether a useful translation to some challenging item could be found.

The third largest number of searches (23) was reported in the 'real world' – browsing one's own or one's parents' bookshelves. Particularly popular were cookbooks and old schoolbooks. Person sources were mentioned 14 times altogether, and they showed considerable variety from personal experience in the kitchen or workplace to 'harassing' relatives or friends who might possess the required knowledge. There were seven mentions of students having used Google for finding the right collocation, and six reports of actual footwork in a library. Google Images was mentioned only once, for the Hampton Court Palace brochure.

# 5. CONCLUSIONS

Judging by their own commentaries, the second-year students' data mining skills appear quite versatile. Obviously, the Internet provides them with the easiest access to different sources (various Internet sources were listed in the reports 150 times, while the category of 'other' sources had 43 listings). The variety of sites listed seems to suggest that in their second year, students already possess some knowledge of where and how to find the most reliable information on the Internet. For example, in the case of the recipe translations, students were perhaps somewhat apprehensive about parallel texts online; in class discussions, it became apparent that they did not feel they could be certain that Internet recipes were not translations – which may well be the case with the growing popularity of food blogs.

However, the students were also quite capable of looking for information in the 'real world' – they turned to their old school books, and readily consulted friends and family they considered experts in the fields of psychology, cooking, or art history. They also reported how they had used their own personal experience in various capacities to meet the data mining challenges posed by the assignments.

Interestingly, the genre of the source text did seem to play a key role in the data mining process: the EU report, which was directed to a particularly specific audience (inside the DGT), guided the students to a very limited and specific direction, i.e. EU Internet pages. Similarly, specific searches were undertaken

when translating the recipes – also texts directed to a specific audience: parallel texts were most often mentioned in connection with this assignment, both online and by browsing bookshelves. However, it was clear that even though recipes are texts for specific audiences, the definition of that audience was familiar to most of the students. The texts intended for more general audiences, i.e. the brochure, the scientific column, and the art history textbook produced less specialized search strategies, such as the use of Wikipedia.

It would be interesting to compare these results with an analysis of first-year students' translation reports, or with reports from more advanced courses, or even with how professional translators search for information. Such a comparison would provide translator trainers with actual data on how the students' data mining competences develop through and after their studies. It would also be interesting to compare the actual quality of the students' translations with the amount of groundwork they have reported; however, such comparisons were, unfortunately, out of the scope of this paper. What can be concluded, however, is that more detailed instructions on writing their translation commentaries would hopefully result in better documentation of the students' data mining processes, and thus provide more information on the development of their data mining competence.

Pakkala-Weckström, M. (2015). Student's Data Mining Skills in Second-Year Undergraduate Translation. Current Trends in Translation Teaching and Learning E, 2. 139–170.

# References

Baker, M. (2011). *In Other Words. A coursebook on translation.* 2nd edition. London and New York: Routledge.

Beeby, A. (2002). Evaluating the Development of Translation Competence. In Schäffner, C. & Adab, B. (eds.). *Developing Translation Competence*. Amsterdam/Philadelphia: John Benjamins. 185–198.

Byrne, J. (2012). *Scientific and Technical Translation Explained. A Nuts and Bolts Guide for Beginners.* London and New York: Routledge.

Englund Dimitrova, B. (2002). Training and educating the trainers. A key issue in translators' training. In Hung, E. (ed.). *Teaching Translation and Interpreting 4*. Amsterdam and Philadelphia: John Benjamins. 73–82.

Eskelinen, J. & Garant, M. (2011). ICT practices among undergraduate English translation students at the University of Helsinki. In Garant, M. and Pakkala-Weckstrom, M. (eds). *Current Trends in Training European Translators and Interpreters – a Selection of Seminar Papers from ESSE10. Current Trends in Translation Teaching and Learning IV. Translation Research IX.* University of Helsinki, Department of Modern Languages. 21–42.

Pakkala-Weckström, M. (2015). Student's Data Mining Skills in Second-Year Undergraduate Translation. Current Trends in Translation Teaching and Learning E, 2. 139–170.

Eskelinen, J. & Pakkala-Weckström, M. (2015). Assessing translation students' learning of professional competences. Paper presented at New Perspectives in Assessment in Translator Training: Bridging the Gap between Academic and Professional Assessment. 4 September 2015, Univeristy of Westminster.

Fox, O. (2002). The use of translation diaries in a process oriented translation teaching methodology. In Schäffner, C. & Adab, B. (eds.). *Developing Translation Competence*. Amsterdam and Philadelphia: John Benjamins. 115–130.

Gambier, Y. et al. (2009). Competencies for professional translators, experts in multilingual and multimedia communication. EMT expert group, DGT, Brussels. http://ec.europa.eu/dgs/translation/programmes/emt/key _documents/emt_competences_translators_en.pdf (retrieved September 14, 2015).

Johnson, J. E. (2003). Learning through portfolios in the translation classroom. In Baer, B. J. & Koby, G. S. (eds.). *Beyond the Ivory Tower. Rethinking translation pedagogy.* Amsterdam and Philadelphia: John Benjamins. 97–116.

Kelly, D. (2005). *A Handbook for Translator Trainers.* Manchester: St. Jerome.

Pakkala-Weckström, M. (2015). Student's Data Mining Skills in Second-Year Undergraduate Translation. Current Trends in Translation Teaching and Learning E, 2. 139–170.

Mackenzie, R. (2004). The competencies required by the translator's roles as a professional. In Malmkjaer, K. (ed.). *Translation in Undergraduate Programmes.* Amsterdam and Philadelphia: John Benjamins. 31–38.

Nord, C. (1997). *Translating as a Purposeful Activity. Functionalist Approaches Explained.* Manchester and Kinderhook: St. Jerome.

Nord, C. (2005). *Text Analysis in Translation. Theory, Methodology and Didactic Application of a Model for Translation-Oriented Text Analysis.* 2nd edition. Amsterdam: Rodopi.

Orlando, M. (2011). Evaluation of Translations in the Training of Professional Translators: At the crossroads between theoretical, professional and pedagogical practices. In *The Interpreter and Translator Trainer* 5(2). 293–308.

Orlando, M. (2012). Training of Professional Translators in Australia: Process-Oriented and Product-Oriented Evaluation Approaches. In *Global Trends in Translator and Interpreter Training*, Hubscher-Davidson S. and Borodo M. (eds.), Continuum, London. 197–216.

Orozco, M. & Hurtado Albir, A. (2002). Measuring Translation Competence Acquisition. In *Meta,* XLVII 3. 375–402.

Pakkala-Weckström, M. (2015). Student's Data Mining Skills in Second-Year Undergraduate Translation. Current Trends in Translation Teaching and Learning E, 2. 139–170.

Pakkala-Weckstrom, M. (2009). Learning the Ropes. How First-Year Translation Students Develop their Translator Identity. In Gaballo, V. (ed.). *English Translation Studies: Methodological Perspectives*. eum x Translation Studies. 95–106.

Pakkala-Weckstrom, M. (2011). Students' working methods & motivation in teaching University-level translation. In Garant, M. and Pakkala-Weckstrom, M. (eds). *Current Trends in Training European Translators and Interpreters – a Selection of Seminar Papers from ESSE10. Current Trends in Translation Teaching and Learning IV. Translation Research IX.* University of Helsinki, Department of Modern Languages. 87–100.

Pakkala-Weckström, M. (2014). The food recipe as a genre – What challenges can it pose to a translator? In Tyrkkö, J., Timofeeva, O. and Salenius, M. (eds.). *Ex Philologia Lux: Essays in Honour of Leena Kahlas-Tarkka.* Helsinki: Société Néophilologique. 327–342.

Pirttimaa, H. (2007). Lääketieteen termit – sanakirja lääketieteen suomen vaalijana. *Kielikello* 4/2007.

Schäffner, C. (2002). Editorial. In Schäffner, C. (ed.). *The Role of Discourse Analysis for Translation and in Translator Training.* Clevedon, Buffalo, Toronto, and Sydney: Multilingual Matters. 1–8.

Pakkala-Weckström, M. (2015). Student's Data Mining Skills in Second-Year Undergraduate Translation. Current Trends in Translation Teaching and Learning E, 2. 139–170.

Tagnin, S. (1995). Students' research for translation. In Dollerup, C. & Appel , V. (eds.). *Teaching Translation and Interpreting 3. New Horizons.* Amsterdam and Philadelphia: John Benjamins. 163–169.

Torresi, I. (2010). *Translating Promotional and Advertising Texts.* London and New York: Routledge.

Vermeer, H. J. (1989). Skopos and commission in translational action. In Chesterman, A. (ed.). *Readings in Translation Theory.* Helsinki: Finn Lectura. 173–187.