

Chapter 2

History in Keywords

2.1 Introduction

In her presentation about the 10-year history of corpus-based translation studies, Laviosa (2003) divided this period into three stages. The first one, labelled by her as “the dawn”, comprises the years 1993–1996, when the idea of using corpus methods for descriptive translation research was introduced by Baker (1993, 1995, 1996). The second stage covers the time period starting from 1996, when Laviosa-Braithwaite (1996) published her Ph.D. thesis, and continuing until the end of the decade. These years were characterised by an intensive study on translation universals (e.g. Olohan 2000; Mauranen and Kujamäki 2004). During the third stage, including the first years of the new millennium, the range of the research field expanded. From the perspective of the present study it is interesting that Laviosa mentioned as one of the new fields *corpus-based critical translation studies*. This term was used by her for corpus-based translation studies that have common ground with critical linguistics and critical discourse analysis.

Now that another ten years have elapsed it is apparent that the interest towards translation universals did not fade away in the beginning of the new millennium. Conversely, as Zanettin (2012: 23) puts it, “[t]he search for regularities in corpora of translations has been mostly carried out with the aim of investigating universals.” He adds that some studies with more specific aims have been carried out as well, and mentions two scholars—Williams and the author of the present study. Williams (2009) examines Spanish translations of the verb ‘report’ in biomedical research articles. Kemppanen’s (2004, 2008) studies concern ideology-bound lexis in Russian-Finnish translations compared to original Finnish texts on the political history of Finland.

The present study continues work with ideological features of lexis. The aim of the study is twofold: first, to introduce different methods for corpus-based research with an orientation towards the historical study of lexical features of translations,

and, second, to discuss how the corpus-based approach can be combined with qualitative methods (cf. Laviosa 2002: 89).

The corpus analysis focuses on describing ideology-specific features in different sub-corpora by using corpus linguistic applications—keywords, key-keywords, word clusters and concordances. In addition, the paper provides an example of using corpus methods together with a qualitative analysis of narrative structure. The narrative analysis is based on Greimas's (1980) model of narrative structures, which is applied for analysing contexts of certain ideology-bound keywords.

The material of this study comprises a Finnish-language corpus referred to as *The Comparable Corpus of History Texts*.¹ It consists of 0.5 million words of texts translated from Russian into Finnish and 0.5 million words of non-translated texts (non-fiction books on political history). The corpus of translated texts consists of different sub-corpora reflecting different circumstances of text production: texts from the Soviet Era (282,938 words), the Perestroika Era (38,830 words) and the post-Soviet Era (193,313 words). The analysis focuses on the two bigger corpora—the Soviet Era translations dating from the 1970s and the post-Soviet translations from the 1990s. The reference corpus consists of non-translated history texts (537,154 words) written in Finnish by Finnish historians and published between 1971 and 1999.

2.2 Studying Ideology-Bound Lexis

2.2.1 *Ideology as a Research Object*

The starting point of the present study is the assumption that every text fulfils a certain *ideological function* (Kemppanen 2005: 187–188). This assumption is based on the ideas of Vološinov 1972 ([1930]: 18), also cited by Petrilli (1992: 252), on the existence of an ideological function attached to every verbal sign. From the perspective of quantitative study on representations of ideology, this kind of approach appears problematic. What is the point in counting manifestations of a phenomenon, if it is said to be present everywhere? In addition, as Vološinov (1972 [1930]: 18) states, the same linguistic items of a language can be used for verbalising any ideology. Every verbal sign gets its actual ideological meaning in the social context where it is produced and received by the language users. Still, it has been claimed that texts differ from each other in the explicitness of representing their ideological features (Fairclough 1997: 26). A political speech expresses values laden in the text itself more clearly than, for example, a text book in physics. The

¹*The Comparable Corpus of History Texts*. Compiled at the University of Joensuu (nowadays the University of Eastern Finland) within the research project *Translated Finnish and Translation Universals. A Corpus-based Study* (1997–2001) led by Prof. Anna Mauranen.

latter—maybe considered “ideologically neutral” by most readers—supports the power of science and technical culture just by being compulsory reading in schools.

Despite the fuzziness of the concept of ideology, some scholars have carried out studies to identify its markers. These studies come from the area of critical discourse analysis—both in linguistics and translation studies. Among them could be mentioned studies carried out by the following authors: Fowler and Kress (1979), Thompson (1994 [1990]), Fairclough (1995), Knowles and Malmkjær (1989), Hatim and Mason (1997) and Puurtinen (1998).

The markers of ideology differ in their applicability to corpus-based study. The most suitable issues for quantitative research are e.g. tautology in a text (cf. Sowarka 1987) or translators’ solutions concerning repetition (cf. Knowles and Malmkjær 1989). Hunston (2002: 109) stresses that ideological stances can be identified in texts with several methods: analysis of keywords, lexical patterning, semantic prosody and repetitions.

However, it must be added that corpus linguistic analysis does not deal just with the use of disconnected words. A multidimensional study of texts including the study of collocations, word clusters and keywords in larger contexts makes it possible to get a clearer view of manifestations of such a complex phenomenon as ideology. Some scholars emphasise the existence of two levels of observations. Hatim (1999, cited in Laviosa 2000: 160) stresses that corpus-based analysis enables us to tap into both what is *in the text* and what is *from the text*. ‘In the text’ means the language itself, which can be analysed by text analysis. ‘Of the text’ refers to the text as a whole, including its overall influence on the level of ideology. Baker (1999: 282) emphasises the importance of viewing the results from the micro-level analysis in a social and ideological context. She sees focusing on formal structures and ignoring the social and ideological context as one of the reasons for a negative attitude towards the role of linguistics in translation studies in the 1980s and 1990s. The same idea of combining different approaches has been expressed by Olohan (2004: 149) as well. She sees it fruitful to link together corpus data and sociocultural studies by relating the results of corpus analysis to the production and reception of translations.

2.2.2 *Keywords, Keyness and Ideology*

The present study introduces how *keyword analysis* can be used for identifying ideology-bound lexical features. The keyword method requires the use of a comparable corpus. In our case the material consists of translated and non-translated texts on the political history of Finland.

Scholars in corpus studies have different views on the focal issue of *keyness*. Scott and Tribble (2006: 55) state that “many languages use the metaphor *key* to identify people, places words, ideas as important; the term is used in text databases for retrieval and in practice it does not seem to be defined. *Keyness* is, therefore, a quality which is intuitively obvious.” Scott and Tribble stress that we should talk

about keyness as a primarily textual phenomenon suggesting what is important in this concrete text (or a set of texts), what the text is all about. Scott and Tribble stay at the textual level, but they mention some attempts by other researchers to go further and make claims about keyness on the cultural level. Williams (1976) compiled a dictionary of keywords in British culture. This contribution was based on intuitive considerations about the keyness of English words. Stubbs (1996, 2001) applied this idea to corpus-based identification of cultural keywords. Actually, the idea of defining keyness on the cultural level is included already in Firth's (1969 [1957]) concept of *focal* or *pivotal words* referring to 'sociologically important words'. He used this term in his article "The technique of semantics", written as early as 1935. It is interesting that Firth (1968 [1957]: 180) refers to this concept also in his article "A synopsis in linguistic theory, 1930–1935" and uses the term *key words* as a synonym for *focal words*. Firth's *technique of semantics* can be considered as the model for current corpus-based studies of keywords. This project on textual analysis included both a formal and sociological component. Firth developed a method for defining the contextual distribution of words. He emphasises that words should be analysed in their actual contexts (Firth 1969 [1957]: 10; Stubbs 1996: 165–166).

Applying corpus methods to monolingual comparable corpora enables us to study manifestations of ideology. They can be analysed as proportional linguistic features of two corpora. The main notion of our study, *keywords*, is well suited for this purpose. A list of keywords reflects a relationship between an analysed text (or a set of texts) and a reference corpus. As Scott (1998: 65) puts it: "Keywords are those whose frequency is unusually high in comparison with some norm". The word *norm* refers here to a frequency norm, which is used for defining the quantitative features of the text corpus in focus. When examining translations, the researcher usually takes the position that non-translated texts stand for the "normal", "typical" or "domestic". For example, corpus studies on translation universals, the so called T-universals, deal with notions such as simplification, untypical lexical patterning and under-representation of target-language-specific items (Chesterman 2011: 177). All of these suggested universals characterise translations as something "untypical" compared to the norm of non-translated texts.

When analysing translations and non-translations from the point of view of ideology, it must be taken into account that the juxtaposition of "typical" and "untypical" is also present here. The concept of ideology has been approached by many scholars from an evaluative perspective. Ideological meanings have been understood as something negative (Fairclough 1997: 66–67) and the term *ideology* has been used for characterising only the ideology of others. Referring to Marx and Engels, this viewpoint of the concept has been labelled as *Die deutsche Ideologie* (Kellner 1998: 67–74). Using quantitative markers of ideology, like corpus-driven lexical features in our study, makes it possible to avoid a one-sided position in defining what is "ideological" in a text. In translation research this means that the analysis can be focused on lexical features of non-translated texts produced in the

ideological framework of the target culture as well. Comparison of statistical information on language use in translated and non-translated texts may reveal regularities in non-translated texts that might remain unnoticed in research with an orientation only towards “untypicalities” of translations.

2.3 Combined Methods: Keywords and Narrative Analysis

This section presents a method which combines two different approaches: a corpus-based, quantitative analysis and an *actantial analysis*, a qualitative analysis of narrative structures in the texts. The analysis starts from generating a keyword list from the corpus of translations by using the corpus program *WordSmith Tools* (Scott 2014). The list functions as a basis for further analysis. The next step consists of comparing the keywords in corpora representing different historical periods. After this, keywords are examined in their contexts: firstly, by analysing frequent word clusters and secondly keywords in narrative structures.

2.3.1 Keyword List as a Starting Point

A multi-method corpus analysis starts with retrieving a keyword list from a comparable corpus. The list used in this study indicates the keywords in translated history texts. The list illustrated below (Table 2.1) actually shows the so called *key-keywords* (Scott 1998)—the words that are key in more than one text (labelled in the table as *key files*).² The key-keywords tool excludes from a list words that characterise a certain text, but are rare in other texts. For example, the frequency of a certain personal or geographical name can be high in one book, where this name is relevant, but in a set of texts representing the whole genre this lexical item may show very low frequency. The words appearing repeatedly as keywords in several texts are of greater interest from the perspective of possible markers of ideology.

Table 2.1 presents facts about the corpus of translated history texts. The words in this list can be classified into several semantic groups. The words *Suomi* ‘Finland’, *Neuvostoliitto* ‘Soviet Union’, *Venäjä* ‘Russia’ and *Leningrad* are related with each other by their geographical or political character. The first three are used for naming the two neighbouring countries, which were closely linked through foreign policy. *Leningrad* refers to a big city in the Western part of the Soviet Union/Russia. In the context of political history this proper name has a “priming

²This key-key-word list originates from an earlier study (Kemppanen 2008), where the lemmatised version was in use. The remainder of the present study is based on the non-lemmatised version of the corpus.

Table 2.1 Keywords in translations^a

Position	Word	Number of “key files”	Number of “key files” %
1	Suomi ‘Finland’	6	54.55
2	ja ‘and’	6	54.55
3	Neuvostoliitto ‘Soviet Union’	5	45.45
4	minä ‘I’	3	27.27
5	tuo ‘that’	3	27.27
6	vuosi ‘year’	3	27.27
7	Leningrad	3	27.27
8	tuollainen ‘that kind of’	2	18.18
9	hän ‘he/she’	2	18.18
10	sota ‘war’	2	18.18
11	sopimus ‘agreement, treaty’	2	18.18
12	suhde ‘relationship, relations’	2	18.18
13	hallitus ‘government’	2	18.18
14	Venäjä ‘Russia’	2	18.18
15	yhteistyö ‘co-operation’	2	18.18
16	välinen ‘between’	2	18.18
17	ystävyyss ‘friendship’	2	18.18

^aKeyness has been calculated via log-likelihood

effect” that makes the reader expect information about the strategic importance of the city. In addition, the word *hallitus* ‘government’ could be placed in this word group. In a political text this lexical item can be used for referring to a state that the text is about: *The government/the government of Finland/Finland does not allow something*.

The second category can be built around the semantic field ‘relationship between actors’. This group includes three words with a meaning of a positive relationship (*sopimus* ‘agreement, treaty’, *yhteistyö* ‘co-operation’ and *ystävyyss* ‘friendship’) and one referring to a negative relationship (*sota* ‘war’). In addition, the word *suhde* ‘relationship, relations’ itself belongs to the group. It can be associated with either a positive or a negative relation to something or somebody.

The third group comprises the following words: *tuollainen* ‘that kind of’, *välinen* ‘between’, *minä* ‘I’, *hän* ‘he, she’, *tuo* ‘that’ and *ja* ‘and’. These lexical items differ from those mentioned above as far as their semantic function is concerned. They all are used as function words (i.e. grammatical words), while the words in the previous categories pertain to content words (semantic words). At first glance function words may appear neutral from the point of view of verbalising ideology. Still, their contexts reveal, that these words can be used for building ideology-bound constructions together with the content words as well.

The only word in the list that was left outside the mentioned categories is the word *vuosi* ‘year’. It is of course logical that this word is typical in history texts

because of their time-orientation, but it is surprising that when two sub-corpora of history texts are compared the word is still in the list. Why is it much more common in translations? A look at the concordances of the word *vuosi* ‘year’ shows that one of the main reasons for the difference in frequencies between the comparable corpora lies in the common use of the phrase *vuoden* [genitive of ‘*vuosi*’] *1948 sopimus* ‘the agreement of the year 1948’ in translated texts (192 occurrences). This item is a lexical calque of the Russian expression *Dogovor 1948 goda* appearing only twice in non-translated texts. In Finnish this document is labelled with the acronym expression *YYA-sopimus* ‘The Agreement of Friendship, Cooperation, and Mutual Assistance’, also known as the *YYA Treaty*. This phrase occurs in non-translated texts 130 times, which implies the importance of this concept in the foreign policy between Finland and the Soviet Union, regardless of how it is named. The YYA Treaty has been considered as one of the most obvious signs of the Soviet influence on Finnish foreign policy.

2.3.2 Historical View on Keywords

The next step of the analysis is to look closer at the different layers of time in the key word list above. The corpus of translated texts is categorised into three blocks according to the year of publication. Five of the eleven texts represent the Soviet Era, two of them the time of transition, Perestroika, and four texts belong to the post-Soviet time. This study concentrates on comparing the texts produced before Perestroika and after the collapse of the Soviet Union.

Table 2.2 introduces a list of keywords which has been made of the five translations from the Soviet Era. Unlike in the previous section, a non-lemmatised version of the corpus is used, so in some cases several word forms of the same

Table 2.2 Keywords in the Soviet Era corpus

	Key word	Freq.	%	RC. Freq.	RC. %	Keyness
1	<i>Suomen</i> ‘Finland’, gen. ^a	4929	2.25	5203	0.97	1765.42
2	<i>ja</i> ‘and’	9992	4.56	15,723	2.92	1204.47
3	<i>Neuvostoliiton</i> ‘Soviet Union’, gen.	1965	0.90	2010	0.37	740.44
4	<i>sopimuksen</i> ‘agreement, treaty’, gen.	481	0.22	288	0.05	372.87
5	<i>Paasikivi</i> (proper name), nom.	467	0.21	274	0.05	369.28
6	<i>Neuvosto-Venäjän</i> ‘Soviet Russia’, gen.	308	0.14	125	0.02	328.94
7	<i>kanssa</i> ‘with’	1258	0.57	1538	0.29	323.78
8	<i>rauhan</i> ‘peace’, gen.	370	0.17	237	0.04	267.58
9	<i>suhteiden</i> ‘relations’, gen. pl.	365	0.17	242	0.05	247.46
10	<i>ystävyyden</i> ‘friendship’	149	0.07	24		246.62

^aThe case of each word form is indicated in the table. Plural indicated with *pl.*, singular not marked

lexical item appear in the list. When the material represents the lexis in an agglutinative language like Finnish, non-lemmatised material provides a lot of information about morphological features in language use. The table indicates the frequency of each key word in the corpus of the Soviet Era material. The first column on the left provides the frequency in absolute numbers, the second one shows the proportion of each word form as a percentage, the third shows frequency in the reference corpus of non-translated history texts, the fourth shows the proportion in the reference corpus and the last column gives the so called *keyness value*. It reflects how strongly the frequency of a word in the study corpus differs from its frequency in the reference corpus. The bigger the value, the bigger is the difference in frequencies between the comparable materials. The *KeyWord*-function generates a word list demonstrating the words with the highest keyness values. The corpus program automatically carries out the comparison of the frequency lists in focus, cross-tabulates the frequencies and performs the statistical tests. Word frequencies with maximal proportional difference in the comparable corpora draw a word into the keyword list (Scott 1998: 71).

Table 2.2 demonstrates the top of the keyword list extracted from the Soviet Era corpus. The total number of keywords generated amounted to 488 items, but even this short sample from the list gives a clear picture what the books on history in this material are about. The main characters in the story are the acting states—*Finland and the Soviet Union (Soviet Russia)*. They have undoubtedly good *relations with* each other: they have made a mutual *agreement* and live in *friendship* and *peace*. One of the main architects of this foreign policy was *Paasikivi*, the President of Finland. This discourse is familiar to all who are acquainted with the clichés of the foreign policy between the Soviet Union and Finland. This policy was at its peak during the 1970s, precisely the time the translated books were published. The analysis of the keywords shows that although Finland has been accused of “Finlandisation”—for practising a foreign policy that flatters the policy pursued by the Soviet Union, there is still a difference in Finnish and Soviet discourse when verbalising these ideas in history writing. If the discourses had been alike, there would not have been so many keywords with this “friendship prosody” in the list.

Most of the words in the list (8 of 10) are semantic words, which stresses the differences in the sub-corpora concerning the content. The two grammatical words *ja* ‘and’ and *kanssa* ‘with’ match with the semantics of the content words referring to positive relations between actors. It is obvious that these grammatical items can be used for combining the names of the actors and positive-laden content words into longer clusters with clearly represented ideological meaning. An extract from the concordance (Figure 2.1) gives an idea how the grammatical word *kanssa* ‘with’ functions as a part of a larger lexical element with positive content: *rauhanneuvotteluista Neuvostoliiton kanssa* ‘peace negotiations with the Soviet Union’, *diplomaattista dialogia Neuvostoliiton kanssa* ‘diplomatic dialogue with the Soviet Union’, *suhteiden normalisointia Neuvostoliiton kanssa* ‘normalisation of relations with the Soviet Union’, *ystävyyys ja yhteistyö Neuvostoliiton kanssa* ‘friendship and co-operation with the Soviet Union’, *Yya-sopimus Neuvostoliiton kanssa* ‘the Agreement of Friendship, Cooperation and Mutual Assistance with the

Concord

File Edit View Compute Settings Window Help

N	Concordance	Set	Tag	Word	#	nt	#	Pos	#	Pos	#	t	#	os	File	%
1	rauhanneuvotteluista Neuvostoliiton kanssa. Tässä yhteydessä Tanner tekee				8 755	532	100%	0	18%	0	8%				skt/bartenje.tbl	18%
2	diplomaattista dialogia Neuvostoliiton kanssa ehtyneenä eikä edes yrittänyt				8 242	500	24%	0	17%	0	7%				skt/bartenje.tbl	17%
3	suhteiden normalisointia Neuvostoliiton kanssa. Toukokuussa 1940 perustettu				9 081	553	100%	0	19%	0	9%				skt/bartenje.tbl	18%
4	on ystävyys ja yhteistyö Neuvostoliiton kanssa, on vankasti taattu vuoden 1948				37 420	235	73%	0	70%	0	0%				kt/ikomisar.tbl	69%
5	Yy-a-sopimus Neuvostoliiton kanssa sallii Suomen esiintyä				50 389	033	24%	0	94%	0	4%				kt/ikomisar.tbl	94%
6	Hyvä naapuruus ja yhteistyö Suomen kanssa lujittuu edelleenkin.»				37 703	121	78%	0	77%	0	7%				skt/bartenje.tbl	77%
7	aktiiviseen yhteistyöhön Suomen kanssa sen talouselämän eri aloja ja				45 188	721	47%	0	85%	0	5%				kt/ikomisar.tbl	84%
8	rauhan ja ystävyys suhteet Suomen kanssa. Samalla se osoitti itsehillintää ja				7 852	527	100%	0	15%	0	5%				kt/ikomisar.tbl	15%
9	allekirjoitettiin rauhansopimus Suomen kanssa. Rauhansopimusta valmisteltiin				29 511	812	100%	0	55%	0	5%				kt/ikomisar.tbl	55%
10	ja hedelmällistä yhteistyötä Suomen kanssa. Kuten Neuvostoliiton				39 923	814	100%	0	00%	0	0%				uskts/mimov.tbl	100%

Figure 2.1 Concordance of the word *kanssa* ‘with’

Soviet Union’, *hyvä naapuruus ja yhteistyö Suomen kanssa* ‘good-neighbourliness and co-operation with Finland’, *aktiiviseen yhteistyöhön Suomen kanssa* ‘active co-operation with Finland’, *rauhan ja yhteistyön suhteet Suomen kanssa* ‘relations of peace and co-operation with Finland’ *allekirjoitettiin rauhansopimus Suomen kanssa* ‘peace agreement with Finland was signed’, *hedelmällistä yhteistyötä Suomen kanssa* ‘fruitful co-operation with Finland’.

The keyword list in the Table 2.3 looks very different to the previous one. It illustrates the material of translated history texts representing the post-Soviet Era.

The keyword list shown in the table does not give any picture of the basic theme or themes of the texts. However, it suggests a different genre or at least another perspective to the narrated content. This can be inferred from the range of pronouns and verb forms. Most of them are verbalisations of the first person singular. The lemma *minä* ‘I’ is represented in the list by three different case forms *minulle* ‘to me, for me (allative)’, *minua* ‘me, of me (partitive)’ and *minut* ‘me (accusative)’. This paradigm continues with verb forms of the first person singular: *olen* ‘I am’, *olin*

Table 2.3 Keywords in the post-Soviet corpus

	Keyword	Freq.	%	RC. Freq.	RC. %	Keyness
1	ja ‘and’	10,868	3.85	15,723	2.92	497.09
2	minulle ‘to me, for me’, allative	339	0.12	48		474.03
3	en ‘not’, negative, 1st pers. sg.	391	0.14	81	0.02	470.29
4	minua ‘me’, partitive	246	0.09	22		391.54
5	sanoi ‘he/she said’	452	0.16	163	0.03	391.14
6	olen ‘I am’	296	0.10	59	0.01	362.22
7	olin ‘I was’	243	0.09	28		326.58
8	ntenä, ending used with certain numerals, ‘on the 4th, 5th etc. of’	153	0.05	0		326.58
9	minut ‘me’, accusative	180	0.06	19		274.86
10	on ‘he/she is’	2543	0.90	3123	0.58	267.22

‘I was’ and *en* ‘not (negative in 1st pers. sg.)’.³ All the listed word forms suggest that the texts on history in this sub-corpus favour a narrative in first person singular. Also over-presentation of the word *sanoi* ‘he/she said’ refers to claims of single human actors which are missing in the Soviet Era materials, where the focus is on the state level, manifested in the names of acting states and the president of the state.

A special case in the list is the abbreviation *ntenä* that is connected with a colon to certain ordinal numbers, e.g. *4:ntenä* ‘on the 4th of’ and *5:ntenä* ‘on the 5th of’. It is clear that this keyword does not belong to the same group of markers as those mentioned above. This item reflects clearly the influence of a lexical pattern in the source texts and can be traced to the Russian way of indicating dates: *4-ogo oktjabrja* ‘on the 4th of October’. The Finnish construction is a calque of the Russian one. This construction is possible in Finnish, but as the keyword list shows, it does not occur at least in this field of language use. The fact that the ending *ntenä* is totally missing in the reference corpus confirms a result from earlier studies (Kemppanen 2008) showing the over-representation of the names of the months in Russian-Finnish translations of history texts. The Finnish convention of expressing dates in history books is based on using numbers and dots, e.g. *4.10. 1941*, instead of attaching a suffix to the number and writing the name of the months with letters.

The fact that a keyword can be recognised as a clear manifestation of interference instead of being identified as an ideology-bound element distinguishing the two comparable corpora of texts is present in all keyword research with an orientation on the analysis of ideology in texts. The researcher is repeatedly forced to ask the question: “Is this ideology or interference? Must this lexical feature be interpreted as influence on the cultural level or purely on the linguistic level?” (cf. Kemppanen 2006).

The only keyword that is common to both lists is the word *ja* ‘and’ at the top of the lists. As was stated earlier, the key words tell what the text is about. This “aboutness” comes mainly from the semantic picture transferred by the content words in a keyword list. However, in many cases it may appear even more interesting to find out the reasons why a grammatical word has achieved a position in a list. The coordinating conjunction *ja* ‘and’ is a fascinating case, because it is one of the most frequent words in any Finnish text. Why does it appear as a keyword in history texts translated from Russian, and both in Soviet and post-Soviet translations? The earlier considerations on the semantic matching of the item *ja* ‘and’ to the Soviet discourse now appear in a new light. This question will be analysed more closely in the next section.

³In the Finnish language the negative word is conjugated for person. The verb form appearing with the negative remains unconjugated.

2.3.3 *Keywords in Context: Clusters*

The next step in this keyword study consists of a more detailed analysis of the contexts where a certain keyword has been used. As the main tools for more advanced analyses of words, corpus researchers mention mostly the terms *collocations*, *clusters* and *lexical patterning* (Olohan 2004: 148–149; Zanettin 2012: 130–146). Other terms have also been used. Jantunen (2004: 19) states that *collocation* refers mostly to the co-occurrence of two words and the term is often distinguished from *multi-word strings* (Mauranen 2000: 120), *lexical bundles* (Biber and Conrad 1999), *multi-word items* (Nelson 2000) and *clusters* (Kenny 2000: 99), used for labelling sections of three or more words. All these issues refer to a tendency of words to appear repeatedly together. Firth defines *collocations* as “actual words in habitual company” (Firth 1968 [1957]: 182). Sinclair (1991: 170) puts this more clearly by stating that a *collocation* is a tendency of two or more words to occur close each other. From the perspective of the present study it is worth mentioning that the relevance of lexical patterning has been recognised by scholars in critical linguistics. Heikkinen (1999: 163) stresses the role of language patterns in verbalising ideological meanings. According to him, the patterns arise repeatedly in a certain context as a result of so called *collocation schemes* where one element forces the other to keep company with the first one.

The patterns building up around the keywords show their *semantic prosody*. Louw (1993: 157) defines *semantic prosody* as “a consistent aura of meaning with which a form is imbued by its collocates”. According to Sinclair (1998: 20), semantic prosody is the reason why a language user selects a certain lexical item. The item is attached with positive or negative evaluation. A particular prosody that occurs recurrently in a text or a set of texts manifests a system of values—an ideology.

Examining collocations and multi-word strings allows us to identify lexical elements which manifest an ideology. Stubbs (1996: 169) considers the study of lexical patterns very important in investigating the problem of language and ideology. He claims that “fixed and semi-fixed expressions (collocations, catch phrases, clichés and idioms) encode cultural information”. Empirical evidence based on this kind of analysis provides information about “how the culture is expressed in lexical patterns”. The importance of large-scale context searches is stressed when the research object deals with culture-bound issues, such as ideology.

The idea of a multi-step keyword analysis is to examine the lexical features of a certain set of texts from various viewpoints. The study of the context where the keywords in focus appear can be carried out by analysing word clusters. The next two tables given below illustrate the most frequent 3-word clusters including the keyword *ja* ‘and’ in the two corpora of translated history texts: the texts from the Soviet Era and those from the post-Soviet period. The keyword *ja* ‘and’ was chosen as an object for further study, because it occurs as a keyword in both corpora. The analysis will show the apparent differences between the two corpora.

Table 2.4 Clusters with the word *ja* ‘and’ in the Soviet Era translations

	The cluster and the case(s) of the conjugated words	Number of clusters
1	<i>Neuvostoliiton ja Suomen</i> ‘Soviet Union and Finland’, gen.	707
2	<i>ja Suomen välisen</i> ‘between... and Finland’, gen./gen sg.	135
3	<i>Suomen ja Neuvostoliiton</i> ‘Finland and the Soviet Union’, gen.	134
4	<i>ja Suomen välisten</i> ‘between... and Finland’, gen./gen. pl.	90
5	<i>ja Suomen välisissä</i> ‘between... and Finland’, gen., inessive pl.	68
6	<i>että Neuvostoliiton ja</i> ‘that... the Soviet Union and’, gen.	52
7	<i>Neuvostoliitto ja Suomi</i> ‘the Soviet Union and Finland’, nom.	48
8	<i>Suomen ja Venäjän</i> ‘Finland and Russia’, gen.	48
9	<i>ja Suomen väliset</i> ‘between... and Finland’, gen./nom. pl.	45
10	<i>Venäjän ja Suomen</i> ‘Russia and Finland’, gen.	41
11	<i>rauhan ja turvallisuuden</i> ‘peace and security’, gen.	40
12	<i>ystävyyden yhteistyö ja</i> ‘friendship, co-operation and’, nom.	40
13	<i>ja keskinäisestä avunannosta</i> ‘and mutual assistance’, elative	37
14	<i>Englannin ja Ranskan</i> ‘England and France’, gen.	33
15	<i>ja Suomen välinen</i> ‘between... and Finland’, gen.	31
16	<i>Neuvosto-Venäjän ja Suomen</i> ‘Soviet-Russia and Finland’, gen.	31
17	<i>ja Suomen välillä</i> ‘between... and Finland’, gen.	30
18	<i>ystävyyden ja yhteistyön</i> ‘friendship and co-operation’, gen.	30
19	<i>ja keskinäisen avunannon</i> ‘and mutual assistance’, gen.	28
20	<i>Repolan ja Porajärven</i> ‘Repola and Porajärvi’, gen. (geographical names)	28
	Number of clusters in total	1696

The list of clusters shown in Table 2.4 gives a clear picture of the discourse in the Soviet Era corpus. It repeats the lexical set provided in the list of individual keywords. The coordinating conjunction *ja* ‘and’ links, firstly, names of states co-operating for achieving common aims—e.g. *the Soviet Union and Finland*, and, secondly, nouns referring to positive action, such as *friendship, co-operation, mutual assistance, peace and security*. Some of the clusters do not fully show the co-operating partners or the coordinated actions, but the cluster elements give a hint of the missing part, e.g. *ja Suomen välillä* ‘between ... and Finland’. In the context of the other clusters the story as whole can be guessed. It is clear that in many cases the keyword *suhteet* ‘relations’ is somewhere outside the borders of the cluster. Even those clusters with the element ‘between’ can be seen as implicit markers of “the ideology of friendship”. The last cluster in the list *Repolan and Porajärven* consists of two geographical names combined with the word *ja* ‘and’. This string of words refers to a region that was offered by the Soviet Union as a compensation for its territorial claims elsewhere in Finland. These geographical names are thus politically value-laden as well.

The list of clusters including the word *ja* ‘and’ retrieved from corpus of the post-Soviet translations (Table 2.5) has some common features with the list of

Table 2.5 Clusters with the word *ja* and' in the post-Soviet Era translations

	The cluster and the case(s) of the conjugated words	Number of clusters
1	<i>Neuvostoliiton ja Suomen</i> 'the Soviet Union and Finland', gen.	135
2	<i>Suomen ja Neuvostoliiton</i> 'Finland and the Soviet Union', gen.	106
3	<i>Englannin ja Ranskan</i> 'England and France', gen.	55
4	<i>Venäjän ja Suomen</i> 'Russia and Finland', gen.	40
5	<i>Suomen ja Venäjän</i> 'Finland and Russia', gen.	33
6	<i>ja Suomen välisten</i> 'between ... and Finland', gen./gen. pl.	31
7	<i>ja sanoi että</i> 'and said that'	28
8	<i>ja että se</i> 'and that it'	22
9	<i>ja sen jälkeen</i> 'and after that'	22
10	<i>Saksan ja Neuvostoliiton</i> 'Germany and the Soviet Union', gen.	21
11	<i>ja niin edelleen</i> 'and so on'	18
12	<i>Englanti ja Ranska</i> 'England and France', nom.	18
13	<i>ja ennen kaikkea</i> 'and first of all'	17
14	<i>Neuvostoliittoa vastaan ja</i> 'against the Soviet Union and', partitive	16
15	<i>ja lisäsi että</i> 'and added that'	14
16	<i>ja Suomen välisen</i> 'between ... and Finland', gen./gen. pl.	14
17	<i>ja Neuvostoliiton välillä</i> 'between ... and the Soviet Union', gen.	13
18	<i>ja ennen muuta</i> 'and primarily'	13
19	<i>ja Ranskan kanssa</i> 'with ... and France', gen.	13
20	<i>että Suomen ja</i> 'that ... of Finland and', gen.	12
	Number of clusters in total	484

clusters based on the Soviet Era translations, but mostly differs from it. Among the shared features should be mentioned the cluster at the top of the list—*Neuvostoliiton ja Suomen* 'the Soviet Union and Finland'—that names the main actors in the foreign policy of the neighbouring countries. As in the other corpus, Finland and Russia (or the Soviet Union) are represented in several clusters. Both of these acting states occur in four clusters with the element *ja* 'and'. The clusters with at least one of them together account for 10 positions in the list.

The repetition of the Soviet discourse is obvious, but even so there are some novelties in the post-Soviet list. In addition to the two neighbouring countries, there have appeared some extra states that co-operate with some other state. *France* occurs in three clusters, *England* in two, *Germany* in one. "The border discourse" has extended into a broader European space. Other newcomers in the list that should be mentioned are markers of the individual-oriented genre of history instead of the earlier official state-ruled terminology. The following items belong to this group of clusters: *ja sanoi että* 'and said that', *ja sen jälkeen* 'and after that', *ja niin edelleen* 'and so on', *ja ennen kaikkea* 'and first of all', *ja lisäsi että* 'and added that', *ja ennen muuta* 'and primarily'.

In addition to the differences in content of the keyword lists retrieved from the two corpora, the difference shows up in the frequencies of clusters. The total number of the clusters with the element *ja* ‘and’ in the top 20 list of the Soviet Era is 1696 (5.99 per 1000 words). The corresponding frequency in the post-Soviet corpus is 484 (2.50 per 1000 words). These numbers demonstrate that the translations from the Soviet Era are characterised by much more repetition.

2.3.4 *Keywords in Narrative Structures*

A more detailed keyword analysis can be conducted by applying larger contexts where individual keywords are examined. Our study uses as a method the so called *actant analysis*, which was developed by Greimas (1980) for studying narrative texts. This approach is based on Vladimir Propp’s (1928) more complex model that was used for analysing Russian fairy tales. The choice of the method is justified by the character of history texts: they are textual realisations of narratives. According to Dray (1989: 111–112), narrativity is considered as one the focal features of historical discourse. Greimas and Courtés (1982: 5) define *actant* as something “that accomplishes or undergoes an act, independently of all other determinations”. The relevance of the actant model for analysing manifestations of ideology is supported by the claim that “ideology can be defined as an actantial structure which actualises the values that it selects within axiological systems” (Greimas and Courtés 1982: 149).

Greimas’s actant analysis is based on recognising the *actants*—the semantic roles of the actors in a narrative. The model consists of three opposed pairs of actants. The main concept is represented by the actants of *subject* and *object*. The subject is the acting hero in a narrative, the one who does something. The object is the goal towards which the action of the subject is oriented, something that is desired by the subject. Another pair of actants is formed by *helper* and *opponent*. The helper is understood as an actor giving assistance to the subject and trying to reach a common aim with the subject. The opponent, on the contrary, refers to the villain in a narrative. The opponent aims at preventing the subject from reaching the goal. Typical of every story is the double narrative trajectory: one for the hero and the other for the anti-subject (the villain). The third pair of actants in Greimas’s model consists of *sender* and *receiver*. The receiver acts as an initiator of the action carried out by the subject. The receiver stands for the one who benefits from the subject’s action (Greimas 1980: 201–207; Greimas and Courtés 1982: 371). The study of manifestations of the last mentioned pair is disregarded in the present analysis.

The actant analysis is conducted in the present study using a corpus-driven approach. The semantic roles are defined in the contexts where the keyword *Neuvostoliitto* ‘the Soviet Union’ occurs. The analysis is based on the use of concordances of the word *Neuvostoliitto* ‘the Soviet Union’, and it is examined in the two corpora representing different periods of time. The word occurrences are resorted according to the first word on the right (and on the left as well), which

arranges the co-text elements in alphabetical order and helps to observe regularities in the neighbouring lexical items. Concordances provide raw material for analysing narrative structures. They can be used for demonstrating manifestations of *the semantic prosody*—“the consistent aura of meaning with which a form is imbued by its collocates” (Louw 1993: 157)—hidden in the context of the keyword in focus. According to Sinclair (1998: 20) and Stubbs (2001: 65), semantic prosody reveals attitudes of the language user towards something represented in the text.

The keyword *Neuvostoliitto* ‘the Soviet Union’ represents one of the main actors in the narrative of the political history of Finland. The narrative structure in the two sets of texts looks quite different. The history texts dating from the Soviet Era present the Soviet Union as a hero-subject, despite the fact that the texts discuss the history of Finland. An interesting detail is that in these texts the subject actant meets the helper actant. The Soviet Union co-operates with Finland and helps it to deal with foreign policy threats. The object of acting is to reach peace and harmony between the two neighbouring countries.

Example 1 **Suomi** oli ensimmäinen teollisesti kehittynyt kapitalistinen maa, jossa ryhdyttiin **Neuvostoliiton** teknisen avun turvin rakentamaan atomivoimaloita. (**Finland** was the first industrially developed capitalistic country that started to build nuclear power plants with the help of **the Soviet Union**.) (Bartenjev and Komissarov 1977)

The role of the Soviet Union as a helper is illustrated in example 1. The helper-subject, the leading socialistic state, assists a capitalistic country to get technically developed. Finland has acted very reasonably: it is the first country in the capitalistic block whose leaders have understood to start co-operation with the Soviet Union.

The lexical item verbalising the main subject of the Finnish history *Suomi* ‘Finland’ occurs, naturally, in the close environment of the keyword *Neuvostoliitto* ‘the Soviet Union’. Although the latter has taken the place of the hero-subject and helper, Finland is still needed for constructing a story.

Example 2 *Samaan aikaan kun Suomen hallitus julisti uskollisuuttaan puolueettomuudelle, Suomen kenraalikunta kehitti sotilaallista oppia, jonka lähtökohtana oli yksi ainoa mahdollisuus – sota Neuvostoliittoa vastaan.* (At the same time as the Government of **Finland** proclaimed its loyalty to neutrality, the generals of **Finland** were developing a military doctrine that was based on one single possibility—a war against **the Soviet Union**.) (Komissarov 1974)

Example 2 introduces the other semantic role of Finland in addition to the co-operating hero—the role of an opponent. A typical feature for a narrative is the earlier mentioned double trajectory. This means that there are two different narratives in a text at the same time. The writer takes a position on the side of the hero-subject who is trying to reach a certain object. He builds the text from the point of view of the hero. This story is accompanied by another narrative, the one of an opponent, also labelled as *anti-subject* or *opponent-subject* (Greimas and Courtés 1982: 371).

Actually, this example demonstrates the double narrative trajectory very clearly. The first item *Suomen*, referring to Finland, displays the role of the hero-subject. It represents the discourse of Finland's official foreign policy. The opponent-subject is verbalised by the second, identical lexical *Suomen* element with the meaning 'Finland'. The Soviet discourse points out the enemies of the friendship policy, in this case the Finnish generals.

The translations of texts written after the collapse of the Soviet Union show how the semantic roles in the narratives have changed. The keyword *Neuvostoliitto* 'the Soviet Union' appears as a verbalisation of the subject, but its status has changed. It is no longer the hero of the story and not a helper. The subject makes mistakes, it could be described as an anti-hero, maybe even a victim-hero. However, the Soviet Union is not described clearly as an opponent-subject either.

Example 3 *Neuvostoliiton* hallitus teki pahan virheen, kun se vuoden 1940 joulukuussa ei asettunut tukemaan... *Suomen* liittoutumista Ruotsin ja Norjan kanssa maiden puolueettomuuden turvaamiseksi Saksan mahdollisilta loukkaamisyrityksiltä. (The Government of **the Soviet Union** made a bad mistake when it did not in December 1940 take a position in supporting... **Finland's** association with Sweden and Norway for the sake of guaranteeing neutrality in these countries against the possible attempts of violation by Germany.) (Sinitsyn 1995)

Example 3 gives briefly an idea of the role of the Soviet Union as a subject in a narrative. The Soviet Union represents the old regime in the post-Soviet reality and acts in the role of an anti-hero. Its decisions are criticised and, on the contrary, Finland's efforts in reaching the goal of neutrality are highlighted. This discourse would have been impossible in the texts of the Soviet Era, where there is no criticism towards one's own government, and instead of mentioning Finland's negotiations on neutrality, the author of the text would have stressed Finland's contacts with Nazi Germany—acting in the role of the opponent.

The last example (4) in principle places the Soviet Union in the category of the opponent-subject. Nevertheless, the role is not defined by the writer of the story. The narrator transfers the responsibility from the Soviet Government to the public opinion that defined the role of the Soviet Union as an invader. Finland is introduced here as a victim-hero.

Example 4 *Talvisodan* ensi päivästä saakka koko maailman yleinen mielipide julisti *Neuvostoliiton* hyökkääjäksi ja asettui *Suomen* puolelle. (From the first days of the Winter War the public opinion of the whole world declared the Soviet Union as an invader and took sides with Finland.) (Sinitsyn 1995)

This example clearly shows that both acting states have taken a new role compared to the texts of the Soviet Era. The Russian authors of the newer books on the history of Finland have more freedom in choosing their perspective. That is why the translations also reflect a broader range of evaluations on Finnish history than the Soviet Era translations could do. The latter, instead, provided stereotype roles of Finland and the Soviet Union.

2.4 Conclusions and Discussion

The aim of the study was to introduce a corpus-based keyword analysis of translated history texts and combine it with a study of the narrative structure of these texts. The corpora, consisting of translations from two time periods with distinct features of the prevailing text production environment, showed differences in keywords. The corpora from the Soviet Era were characterised by frequent repetition of keywords referring to the two main actors in the narratives of the Finnish political history—Finland and the Soviet Union. In addition, these texts focus on verbalising the semantic field of positive relations between the main actors. Contents such as, ‘friendship’, ‘co-operation’, and ‘agreement’ are typical of the keyword list retrieved from the Soviet Era corpus.

The post-Soviet corpus of translations presents a keyword list with a range of grammatical words instead of reoccurring content words typical of the Soviet Era texts. The post-Soviet translations give more space for individual use of words, and they are distinguished from the texts in the other corpus by their references to the first person singular and to human actors in general. The texts from the Soviet Era are characterised by the action of states, armies, governments and other official issues.

The keywords were examined in larger contexts by using cluster analysis. The study was focused on the clusters that shared the only common keyword *ja* ‘and’ in the keyword lists of both translation corpora. The cluster analysis revealed the same kind of differences between the two corpora as the keyword lists. The Soviet Era clusters repeat the contents familiar from the keyword list. The coordinating conjunction *ja* ‘and’ links the verbalisations of positive relation between the actors into clusters, such as *Neuvostoliitto ja Suomi* ‘the Soviet Union and Finland’, *ystävyyden ja yhteistyön* ‘friendship and co-operation’ and *rauhan ja turvallisuuden* ‘peace and security’.

The next part of the study, the actant analysis, made it possible to show how the keywords are linked together in a narrative structure. Combining the information taken from specific statistical data with qualitative considerations enables a link to be built between the two levels of ideology: the linguistic markers and the social context. The latter is needed for the interpretation of the markers.

The results of the actant analysis make it clear that there exists a difference between the narrative structures in the Soviet Era and post-Soviet Era corpora. The Soviet Era texts introduce a co-operating hero—*Finland and the Soviet Union*—reaching the goal of friendship. The semantic role of the Soviet Union is the subject and the helper at the same time. Finland, especially its right wing and capitalistic forces, fulfils another role as well—the role of the opponent. The post-Soviet texts illustrate a more complicated picture of actants, where the Soviet Union can appear as an anti-hero and Finland as a victim of the action of the Soviet Union.

The method of combining corpus statistics and qualitative analysis give new challenges for further research. The linkage between the statistical features of translated texts and the reception of translations appears especially interesting. Mäkisalo and Kempainen (2010) carried out a reception test where the subjects

were asked to rank Russian-Finnish translations of the political history of Finland and non-translated texts of the same genre in terms of the foreignness versus the domesticity of the texts. It was hypothesised that the concept of keyness could be used for operationalising the concepts of foreignisation and domestication. Despite the fact that in this particular study there was no significant correlation between the results of the corpus statistics and the reception test, it is clear that approaches that combine different methods provide new perspectives for corpus-based translation and linguistic studies.

References

- Baker, Mona. 1993. Corpus linguistics and translation Studies. Implications and applications. In Text and Technology. In Honour of John Sinclair, eds. Mona Baker, Gill Francis, and ElenaTognini-Bonelli, 233–250. Philadelphia and Amsterdam: John Benjamins.
- Baker, Mona. 1995. Corpora in translation studies: An overview and some suggestions. *Target* 7 (2): 223–243.
- Baker, Mona. 1996. Corpus-based translation studies: The challenges that lie ahead. In Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager, ed. Harold Somers, 175–187. Amsterdam: John Benjamins.
- Baker, Mona. 1999. The role of corpora in investigating the linguistic behaviour of professional translators. *International Journal of Corpus Linguistics* 4(2): 281–298.
- Bartenjev, T., and J. Komissarov. 1977. *Kolmekymmentä vuotta hyvää naapuruutta (Thirty Years of Relations of Good Neighbourhood)*. Trans. U. Heino. Helsinki: Otava.
- Biber, Douglas, and Susan Conrad. 1999. Lexical bundles in conversation and academic prose. In Out of Corpora. Studies in Honour of Stig Johansson, eds. Hilde Hasselgård, and Signe Oksefjell 181–190. Amsterdam: Rodopi.
- Chesterman, Andrew. 2011. Translation universals. In Handbook of Translation Studies, ed. Yves Gambier, and Luc van Doorslaer, vol. 2, 175–179. Amsterdam and Philadelphia: John Benjamins.
- Dray, William H. 1989. *On history and philosophers of history*. Leiden, New York, København and Köln: E.J. Brill.
- Fairclough, Norman. 1995. *Critical discourse analysis: The critical study of language*. London: Longman.
- Fairclough, Norman. 1997. Miten media puhuu? (Media Discourse). Trans. V. Blom, and K. Hazard. Tampere: Vastapaino.
- Firth, J.R. 1968 [1957]. A synopsis of linguistic theory, 1930–55. In Selected Papers of J.R. Firth 1952–59, ed. Frank R. Palmer, 168–205 London: Longmans.
- Firth, J.R. 1969 [1957]. The technique of semantics. Papers in Linguistics 1934–1951. Reprinted. London: Oxford University Press, 7–33.
- Fowler, Roger, and Gunther Kress. 1979. Critical linguistics. In *Language and control*, ed. Roger Fowler, Robert Hodge, Gunther Kress, and Tony Trew, 185–213. London: Routledge and Kegan Paul.
- Greimas, Algirdas J. 1980. Strukturaalista semantiikkaa (Sémantique structurale. Recherche de méthode). Trans. E. Tarasti. Helsinki: Gaudeamus.
- Greimas, Algirdas J., and Joseph Courtés. 1982. Semiotics and Language: an Analytical Dictionary (Sémiotique: Dictionnaire raisonné de la théorie du langage). Trans. L. Crist, et al. Bloomington: Indiana University Press.

- Hatim, Basil. 1999. The cultural and the textual in the way translation studies has evolved. Paper presented at the Seminar held at the Centre of Language and Linguistics of the University of Salford, 24 March 1999.
- Hatim, Basil, and Ian Mason. 1997. *The Translator as Communicator*. London and New York: Routledge.
- Heikkinen, Vesa. 1999. Ideologinen merkitys. Kriittisen tekstintutkimuksen teoriassa ja käytännössä. (Ideological Meaning in the Theory and Practice of Critical Text Analysis). Helsinki: Suomalaisen kirjallisuuden seura.
- Hunston, Susan. 2002. *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Jantunen, Jarmo H. 2004. Synonymia ja käännössuomi. Korpusnäkökulma samanmerkityksisyyden kontekstuaalisuuteen ja käännöskielen leksikaalisiin erityispiirteisiin. (Synonymity and Translated Finnish. A Corpus-based View of Contextuality of Synonymous Expressions and Lexical Features Specific to Translated Language). University of Joensuu Publications in the Humanities 35. Joensuu: University of Joensuu.
- Kellner, Douglas. 1998. Mediakulttuuri. (Media Culture: Cultural Studies, Identity and Politics between the Modern and the Postmodern). Trans. R. Oittinen, et al. Tampere: Vastapaino.
- Kemppanen, Hannu. 2004. Keywords and ideology in translated history texts: a corpus-based analysis. *Across Languages and Cultures* 5(1): 89–106.
- Kemppanen, Hannu. 2005. Avainsanoja, ideologiaa ja lojaaliutta: venäjistä suomennettujen historiatekstien korpusanalyysi (Keywords, ideology and loyalty: a corpus-based analysis of history texts translated from Russian into Finnish). In Käännössuomeksi. Tutkimuksia suomenosten kielestä, eds. Anna Mauranen and Jarmo H. Jantunen. Tampere Studies in Language, Translation and Culture. Sarja A, vol. 1, 185–208. Tampere: Tampere University Press.
- Kemppanen, Hannu. 2006. Ideologija ili interferencija? Korpusnyj analiz russko-finskih perevodov i original'nyh finskih tekstov po političeskoj istorii Finljandii (Ideology or interference? A corpus-based analysis of Russian-Finnish translations and original Finnish texts on the political history of Finland). Universitetskoe perevodovedenie. Sedmye Fëdorovskie čtenija. Vypusk 7:221–232. Sankt-Peterburg: Peterburgskij gosudarstvennyi universitet. Filologičeskij fakul'tet.
- Kemppanen, Hannu. 2008. Avainsanoja ja ideologiaa: käännettyjen ja ei-käännettyjen historiatekstien korpuslingvistinen analyysi (Keywords and Ideology: A Corpus-based Analysis of Translated and Non-Translated History Texts). University of Joensuu Publications in the Humanities 51. Joensuu: University of Joensuu.
- Kenny, Dorothy. 2000. Lexical hide-and-seek: Looking for creativity in a parallel corpus. In *Intercultural Faultlines. Research Models in Translation Studies I. Textual and Cognitive Aspects*, ed. Maeve Olohan, 93–104. Manchester, UK and Northampton MA: St. Jerome Publishing.
- Knowles, Murray, and Kirsten Malmkjær. 1989. Translating ideology: Language, power and the world of the Tin Soldier. In *English Language Research Journal LR Journal 3: Language and Ideology*, ed. Murray Knowles, and Kirsten Malmkjær, 205–241. Birmingham: University of Birmingham.
- Komissarov, Juri. 1974. *Suomi löytää linjansa. Neuvostoliittolaisen historian tutkijan näkemys Suomen tasavallan vaiheista (Finland Finds Its Course. The Point of View of A Soviet History Scholar on the Phases of the Finnish Republic)*. Helsinki: Otava.
- Laviosa, Sara. 2000. TEC: A resource for studying what is “in” and “of” translational English. *Across Languages and Cultures* 1(2): 159–177.
- Laviosa, Sara. 2002. Europe in the making in translational and non translational English. In *CTIS Occasional Papers*, volume 2, ed. Keith Harvey, 85–95. Manchester: Centre of Translation and Intercultural Studies, UMIST.
- Laviosa, Sara. 2003. Corpus-based translation studies: Where does it come from? Where is it going? A paper read at the conference Corpus-based Translation Studies Research and Applications Conference, University of South Africa, Pretoria, 21–25 July 2003.
- Laviosa-Braithwaite, Sara. 1996. The English Comparable Corpus (ECC): A Resource and a Methodology for the Empirical Study of Translation. Department of Language Engineering,

- Volume I. A thesis submitted to the University of Manchester Institute of Science and technology for the degree of Doctor of Philosophy.
- Louw, Bill. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In *Text and Technology*. In Honour of John Sinclair, eds. Mona Baker, Gill Francis and Elena Tognini-Bonelli, 157–176. Philadelphia and Amsterdam: John Benjamins.
- Mäkisalo, Jukka, and Hannu Kemppanen. 2010. Frekvenssejä ja reseptiota: Kaksi näkökulmaa kotouttamiseen ja vieraannuttamiseen (Frequencies and reception: two perspectives on domestication and foreignisation). In *MikaEL—The Publications of the Annual Finnish Symposium in Translation and Interpreting Studies*, 4: 1–14. <http://www.skitl.net/mikael/vol4.php>. Accessed 25 June 2013.
- Mauranen, Anna. 2000. Strange strings in translated language. A study on corpora. In *Intercultural Faultlines. Research Models in Translation Studies I. Textual and Cognitive Aspects*, ed. Maeve Olohan, 119–141. Manchester: St. Jerome.
- Mauranen, Anna, and Pekka Kujamäki (eds.). 2004. *Translation universals—Do they exist?*. Amsterdam and Philadelphia: John Benjamins.
- Nelson, Mike. 2000. A corpus-based study of business English and business English teaching materials. Unpublished thesis. University of Manchester: Manchester. <http://users.utu.fi/micnel/thesis.html>. Accessed 25 June 2013.
- Olohan, Maeve. 2000. *Introducing corpora in translation studies*. London and New York: Routledge.
- Olohan, Maeve. 2004. *Introducing corpora in translation studies*. London and New York: Routledge.
- Petrilli, Susan. 1992. Translation, semiotics and ideology. *TTR* 5(1): 233–264.
- Propp, Vladimir. 1928. *Morfologija skazki (Morphology of the Folktale)*. Leningrad: Academia.
- Puurtinen, Tiina. 1998. Syntax, readability and ideology in children's literature. *Meta*, XLII 4: 524–533.
- Sinitsyn, Jelisei T. 1995. *Vaiettu totuus: salaisen agentin todistajalausunto (The Untold Truth: A Witness Statement of a Secret Agent)*. Trans. E. Adrian. Helsinki: Otava.
- Scott, Mike. 1998. *WordSmith Tools Manual*. Version 3.0. Oxford: Oxford University Press. <http://www.lexically.net/wordsmith/version3/manual.pdf>. Accessed 25 June 2013.
- Scott, Mike. 2014. *WordSmith Tools Manual*. Version 6.0. Oxford: Oxford University Press.
- Scott, Mike, and Christopher Tribble. 2006. *Textual patterns. Key words and corpus analysis in language education*. John Benjamins: Amsterdam and Philadelphia.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, John. 1998. The lexical item. In *Contrastive Lexical Semantics*, ed. Edda Weigand. Amsterdam Studies in the Theory and History of Linguistic Science. Series IV, Current Issues in Linguistic Theory, 1–24. Amsterdam and Philadelphia: John Benjamins.
- Sowarka, Bernhard H. 1987. Die Ideologiehafteigkeit kognitiver Strukturen in Texten. In *Textanalyse als Kognitionskritik? Möglichkeiten und Grenzen ideologiekritischer Inhaltsanalyse*, ed. Peter Vorderer, and Norbert Groeben, 46–136. Tübingen: Narr.
- Stubbs, Michael. 1996. *Text and corpus analysis*. Computer-assisted Studies of Language and Culture. Oxford: Blackwell.
- Stubbs, Michael. 2001. *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.
- Thompson, John B. 1994 [1990]. *Ideology and modern culture: Critical and social theory in the era of mass communication*. Reprinted. Cambridge: Polity Press.
- Vološinov, V.N. 1972 [1930]. *Marksizm i filosofija jazyka. Osnovnye problem sociologičeskogo metoda v nauke. (Marxism and the Philosophy of Language. Fundamental Problems of the Sociological Method in the Science of Language)*. Janua Linguarum. Series Anastatica, 5. 2nd edition. The Hague and Paris: Mouton.
- Williams, Raymond. 1976. *Keywords: A vocabulary of culture and society*. London: Fontana/Croom Helm.

- Williams, Ian A. 2009. A corpus-based study of Spanish translations of the verb ‘report’ in biomedical research articles. *Meta* 54(1): 146–160. <http://id.erudit.org/iderudit/029798ar>. Accessed 25 June 2013.
- Zanettin, Federico. 2012. *Translation-driven corpora. Corpus resources for descriptive and applied translation studies*. Manchester, UK and Kinderhook (NY), USA.

Textual and Contextual Analysis in Empirical Translation
Studies

Laviosa, S.; Pagano, A.; Kemppanen, H.; Ji, M.

2017, VIII, 157 p. 17 illus., 5 illus. in color., Hardcover

ISBN: 978-981-10-1967-8