Mikhail Mikhailov

The future becomes the present and the present becomes the past
Studying temporal structures in a parallel corpus of state treaties

Abstract
This paper examines how the verb tenses used in the different language versions of bilateral treaties correlate. It is commonly believed that such texts are very close in content and formal structure. However, differences in grammar make the direct copying of tenses impossible. The research dataset is a parallel corpus of State Treaties between Russia and Finland. The quantitative study of tenses in the Finnish and Russian versions of the documents shows significant differences. Although present tenses are common in both the Russian and the Finnish texts, past tenses are more frequent in the Russian texts. The absence of a future tense in Finnish and the absence of the perfect in Russian also cause differences. As a result, different language versions of the same treaty might present slightly different points of view. Understanding this problem is essential for improving the quality of multilingual documents that involve (multiple) translating and co-drafting.

1 Introduction
A bilateral political treaty is a bilingual (or sometimes multilingual) document that is a unity of two (or more) authenticated versions in different languages. These versions are in most cases the outcome of a long process: drafting the document in one language, translating it into another language, reviewing it, making additions and corrections, translating it back into the original, comparing and synchronizing the two versions, etc. As a result, all authenticated language versions of a treaty involve translation, even though, strictly speaking, there is no actual source text and no target text (see e. g. Schäffner 1997; Bunn-Livingstone 2002; Probirskaja 2009).

Preparing such multilingual documents often entails contributions from people with diverse professional, cultural and linguistic backgrounds. Nevertheless, the different language versions of the same treaty are expected to be identical in meaning and structure, irrespective of the languages in which they are drawn up (see the Vienna Convention on the Law of Treaties 1969 2005: Article 33). From a linguistic point of view, however, bilingual treaties cannot be identical, due to differences in language structure.
and in stylistic conventions; and because they only involve translation during certain stages of their production, such texts cannot be studied either as originals or as translations. Instead, we have to take into account the complexity of the whole process, which has to be explored from various points of view and by using different methods. In the present study, a corpus approach will be used. It will highlight the problematic areas that should be considered during the preparation of bilingual documents. In this way, the process of multilingual drafting and translation can be significantly improved.

In this paper, I examine bilateral state treaties between the USSR/Russia and Finland. My aims are to compare the temporal structures of the Russian and the Finnish versions of the documents, and to determine how close these structures are to each other. An annotated parallel corpus of state treaties is the source of my data. I use a quantitative analysis to obtain a general view of the findings, but I also study parallel concordances for a more detailed examination.

Specifically, I will show that verbal tenses are used in documents according to the logic of the text macrostructure. Of course, some of the dissimilarities across language versions might be unavoidable, e. g. the presence of certain grammar forms in one language and their absence in the other (cf. Jakobson 1959: 235–237). However, many other differences can be overcome, if technical writers and translators are made aware of the issues in question.

2 Text, time, and tense

Most texts have complex structure: a long text is assembled of parts, chapters, and sections, while a short text consists of paragraphs. A text’s macrostructure is composed according to its purpose, the information channel, and other external factors (see e. g. Dijk 1985). The elements of the macrostructure are realised in smaller units at the microstructure level – sentences – which in turn are made up of elements at the linguistic level – lexemes, syntactic constructions, and grammar forms (see e. g. Hatim 2015: 54–58). These units are usually arranged following certain patterns with the help of special lexical and grammatical markers (Salkie 1995: 91). Descriptions of events and states of affairs are placed in space and time, even if the subject of the text no longer exists or never even existed.

Vyvyan Evans explains the ideas of the present, past, and future in terms of “perceptual moment”, “memory system” and “anticipation” (Evans 2004: 188). This definition works when we approach the idea of time from the perspective of personal, everyday life experiences. When temporality is dealt with in more complex cases – when the individual cannot observe the events – the idea of time becomes much more complicated. To begin with, a text has its own time. A text may be written many years ago and discuss the events of that time, or it can tell about some events in a fictional future.

The concept of time is addressed in human languages using various aspects: duration, moment, instance, event, matrix, agent, measurement, and commodity (Evans
We perceive time through metaphors of all kinds. Lakoff and Johnson (1980: 7–9) discuss the metaphor ‘time is money’ among the first metaphors ‘we live by’: we spend time, the time spent is worth something, we have the time to spare, etc. The metaphor of space is another very typical way to get at least some understanding of time: most temporal prepositions are also used as locative prepositions (cf. in the house/in January, over the trees/over two hours, off the train/off season, etc.), and the same adjectives can be used to describe physical size and duration (long rope/long contract, short man/short life, etc.).

Any text is therefore linked to time and has a temporal structure. Temporal structures are complicated networks involving the events mentioned in the texts. Verbs together with lexical elements explicitly denoting time (yesterday, last Monday, earlier, etc.) function as markers of temporal structure. Verbs play a very important role in setting up the structural units at the micro level. Tense is a “grammaticalised location in time” (Comrie 1985: 9). Tenses can place the events in relation to the present moment (Now/In the Past/In the Future) or to other events (Before/After/At the same time) (see Comrie 1985: 2–13). Tenses form very complicated systems that vary greatly across languages (see e. g. Dahl 1985: 103–155; Jespersen 1992: 254–289). Tense should not be confused with time: e. g. the present tense can be used when referring to events in the past and in the future.

Despite the important role of verbal tenses, the recognition of the temporal structure of the text by the reader cannot be based on verbs alone; the process involves the context and background knowledge. For example, events can be mentioned in reverse order (cf. Comrie 1985: 27–28), but the reader is capable of reconstructing the order in which they took place (Galiotou 2002).

The matter becomes even more challenging when dealing with concrete texts, which can have very different temporal structure depending on their purpose. Egon Werlich’s (1976) typology of texts includes five idealised types: description, narration, exposition, argumentation, and instruction. Let us try to speculate what the temporal structure of these different text types would be like.

In narrative texts, grammatical tenses reflect the chronological sequences of events and their relations to the moment of narration or reading. We should bear in mind, however, that the order in which the events are presented is not necessarily chronological (see above) and that use of the present tense does not mean that the events take place at the moment of narration. The verb tenses correlate with time but do not correspond with it exactly.

In other text types, things are more complicated. Descriptive and expository texts describe states of affairs – they are static – so the tenses simply place the states in question in a certain moment of time. Argumentative texts show the advance of ideas and explain conclusions, and the actual tense forms are not connected to time at all, because the argumentation is not temporal. Instructions can be composed of imperatives, or commands can be hidden within present or future tenses.
One should remember that Werlich’s types are idealised: instances of such distinct types are uncommon. In the real world, texts – especially long texts – are usually a blend of two, three, four, or even all five text types. This means that if the text is not a clear-cut narrative, then the tenses of the verbs are not directly related to the time frame, because events – if there are any – do not make sequences, and sometimes it does not even matter whether they took place in the past or are taking place right now. I will now examine if this works in the texts that will be the subject of this paper – namely, political treaties.

3 Time flow in political treaties

Political treaties describe the relations, states, and processes that take place during the period of time when the agreement is valid. The preamble can mention events and processes that made the conclusion of the treaty possible and describe the process of the agreement coming into force. The body of the agreement text outlines the actions the parties are obliged to take after the treaty has been signed, and it determines the state of affairs maintained by the agreement.

The text of the treaty can be constructed in various ways depending on what point of view the writers choose. One can view the same events from different perspectives, and this can influence the verbal tense forms used throughout the document. Some possible scenarios are shown in Table 1. If the actual moment of signing the treaty is taken as the starting point, the events described in the preamble will be in the past, while the actions after signing the treaty will be in the future. Another option is to take the moment when the treaty comes into force as the starting point. In this case, the state of affairs will be in the present, while the preamble and the situation before the treaty comes into force will be in the past. It is also technically possible to present everything in the treaty as happening ‘now’. It seems, however, problematic to place the whole treaty in the past, because it would be against the essential idea of the agreement: a treaty is about to be followed or is being followed now. The text presumes that the treaty is in force: one cannot distinguish by any language features a treaty that is in force from another one that is not.

---

1 The texts of international treaties are written by large teams consisting of diplomats, politicians, lawyers, experts, technical writers, translators, copy-editors, etc. In this paper, I use the term writer to refer to this collective author of the document.
The future becomes the present and the present becomes the past

Studying temporal structures in a parallel corpus of state treaties

<table>
<thead>
<tr>
<th>States of affairs in the treaty</th>
<th>Time in the document</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scenario 1</td>
</tr>
<tr>
<td>State of affairs that made the</td>
<td></td>
</tr>
<tr>
<td>concluding of the treaty</td>
<td></td>
</tr>
<tr>
<td>necessary and possible</td>
<td>Now</td>
</tr>
<tr>
<td>The act of signing the treaty</td>
<td>Now</td>
</tr>
<tr>
<td>The state of affairs after</td>
<td></td>
</tr>
<tr>
<td>concluding the treaty</td>
<td>Now</td>
</tr>
</tbody>
</table>

Table 1: Possible time scenarios for a treaty

My key interest in this paper is to what extent the tenses used in the different language versions of the state treaties correlate. It is a common belief that the texts of language versions of the same treaty must be very close, both in content and in formal structure. However, as it was already mentioned, the drafting of documents is a very long and multi-phase process, and it is technically very difficult to ensure that the language versions are identical. Prieto Ramos (2011: 205) shows in his paper that certain differences between language versions are inevitable due to the time pressure, the complexity of processes, and the differences between languages.

Let us compare the Russian and the Finnish versions of the same treaty (see Table 2). The guidelines for compiling international treaties issued by the foreign ministries of Russia and Finland are entirely different (see Recommendations 2009; Treaty Manual 2017) and nevertheless the texts of the Russian and Finnish versions look very similar in structure and even in length: Russian version – 1,694 characters, Finnish version – 1,610 characters. However, the temporal structure of the texts is not so close: the Russian version is written in the present tense with some occasional present and past participles. The key clauses of the Finnish version are based for the most part on perfect tense, some clauses are in the present passive, and some phrases include past participles. The grammatical timelines of the two documents are therefore not identical, although this does not change their contents.
### Russian version

**ПРОТОКОЛ**

к Консульской Конвенции между Финляндской Республикой и Союзом Советских Социалистических Республик

1. Договаривающиеся Стороны соглашаются, что уведомление консула об аресте или задержании в иной форме гражданина государства, назначавшего <Part past> консула, упомянутое <Part past> в пункте 2 статьи 30 Консульской Конвенции между Финляндской Республикой и Союзом Советских Социалистических Республик от 24 января 1966 года производится <V pres> в течение одного-трех дней после ареста или задержания в зависимости от условий связи.

2. Договаривающиеся Стороны соглашаются <V pres>, что упомянутые <Part past> в пункте 3 статьи 30 Консульской Конвенции права консульского должностного лица посещать <Inf> и сноситься <Inf> с гражданином государства, назначавшего <Part past> консула, находящимся <Part past> под арестом или задержанным <Part past> в иной форме, предоставляются <V pres> в течение двух-четырех дней после ареста или задержания такого гражданина в зависимости от его местонахождения.

3. Договаривающиеся <Part pres> Стороны соглашаются <V pres>, что указанные <Part past> в пункте 3 статьи 30 Консульской Конвенции права консульского должностного лица посещать <Inf> и сноситься <Inf> с гражданином государства, назначавшего <Part past> консула, находящимся <Part past> под арестом или задержанным <Part past> в иной форме или отбывающим <Part pres> срок тюремного заключения, предоставляются <V pres> на периодической основе.

Настоящий Протокол является <V pres> неотъемлемой частью Консульской Конвенции между Финляндской Республикой и Союзом Советских Социалистических Республик от 24 января 1966 года.

Совершено <Part past> в г. Москва, 24 января 1966 года в двух экземплярах, каждый на

### Finnish version

Suomen Tasavallan ja Sosialististen Neuvostotasavaltain Liiton väliseen konsulisopimuksen liittyvää <Part pres> PÖYTÄKIRJA


2. Sopimuspuolet ovat sopineet <V perf> siitä, että konsulisopimuksen 30 artiklan 3 kohdassa mainittu <Part past> konsulivirkamiehen oikeus tavana <Inf> vangittuna tai missä muodossa tahansa pidätetynä <Part past> olevaa nimittäjävaltion kansalaista ja olla <Inf> yhteydessä hänen myönnetään <V pres pass> 2–4 päivän kuluessa tämän vangitsemisen tai pidätysen jälkeen, hänen olinpaikastaan riippuen.

3. Sopimuspuolet ovat sopineet <V perf> siitä, että konsulisopimuksen 30 artiklan 3 kohdassa mainittu <Part past> konsulivirkamiehen oikeus tavana <Inf> vangittuna tai missä muodossa tahansa pidätetynä <Inf> olevaa <Part pres> taikka vapausraangaistusta kärsivää <Part pres> nimittäjävaltion kansalaista ja olla <Inf> yhteydessä hänen myönnetään <V pres pass> toistuvasti.

Tämä pöytäkirja on <V pres> erottamaton osa Suomen Tasavallan ja Sosialististen Neuvostotasavaltain Liiton välillä 24 päivänä tammikuuta 1966 tehtyä <Part past> konsulisopimusta.

Tehty <Part past> Moskovassa 24 päivänä tammikuuta 1966 kahtena suomen- ja
Table 2: A Russian-Finnish treaty (FI_USSR_consular_treaty_protocol_1966) presented as an aligned parallel text with grammatical tags

Even a cursory look at the document in Table 2 shows how difficult it would be to manually compare verbal forms in parallel texts. One needs parallel aligned corpora with grammatical annotations to make it possible to obtain extensive data and to automate the process of comparing parallel texts.

4 The research data

The study was performed on a parallel corpus of Soviet/Russian–Finnish state treaties that is a subcorpus of a larger corpus titled “the Parallel Electronic corpus of State Treaties” (PEST; Mikhailov/Santalähti/Souma 2019). All treaties between the two countries concluded during the last hundred years (1918–2019) are included in the corpus. The great majority of the treaties are bilingual documents and have two authentic versions: one in Russian and one in Finnish. The remaining treaties were concluded in three languages and have a third authentic version in French (until the early 1940s) or English (later on). Currently, we have only the Russian and the Finnish versions in the corpus. Some documents have long appendices with lists, tables, or technical specifications that might skew the corpus towards various specialist areas (like chemistry, aviation, navigation, education, etc.), so the appendices of the documents have not been included in the corpus.

The corpus is small, because it is limited by the number of existing treaties. At the moment this paper was written, there were 228 pairs of texts in the corpus with about 250,000 words in Finnish and about 300,000 words in Russian. Expired treaties were also included, because the validity of the treaties is not relevant from the point of view of linguistics, translation studies, and language technologies. The texts were aligned at the sentence level, lemmatised, and annotated morphologically and syntactically by open universal dependency grammar parsers (LF Aligner, TreeTagger, Finnish-dep-parser, and MaltParser).

2 In the text of the article, documents are referred to by codes. The complete titles of the quoted documents are provided in the list of Sources in the Appendix.
The corpus is stored on the server puolukka.uta.fi (IT Centre for Science of Finland, CSC) and a separate user account is required to access it. Access can be granted to others upon request. Once completed, the corpus may be opened for larger-scale use within the research community. For corpus queries, the research group uses its own online corpus manager, TextHammer (puolukka.uta.fi/texthammer), which is specifically designed for working with parallel corpora. The software includes various tools, such as parallel concordances, frequency lists, N-grams, collocations, etc., and new functionalities are being developed.

5 Tense systems in Russian and Finnish

Before we go any further, however, it would be helpful to comment briefly on the verbal tenses in the two languages. As Russian and Finnish belong to different language families (the Slavonic group of the Indo-European languages vs. the Finno-Ugric group of the Uralic languages), it is not surprising that their tense systems are so unlike.

Russian verbs belong to one of two aspects: the perfective (which sees the situation as a single whole (Comrie 1976: 16)) or the imperfective (which refers to general facts, or to continuing or repeated events). Perfective verbs have two tense forms: the past and the future simple. Imperfective verbs have three tense forms: the past, the present, and the future complex (which is formed with the auxiliary byt’ ‘to be’ + infinitive). Gerunds of perfective verbs are in the past tense, while gerunds of imperfective verbs are in the present tense. Perfective verbs can only form past participles, while imperfective verbs form both present and past participles. The Russian language does not have a perfect aspect (which should not be confused with the Russian perfective). For details, see (RG 1980: §§1384–1387, 1490–1498).

Finnish verbs have four tenses: the present, the past, the perfect and the pluperfect (Karlsson 2018: 299–314; VISK 2008: §112). The last two tenses are relative in their meaning, but both mainly contain the idea of past time. The semantics of the Finnish perfect tense as it is presented by Finnish grammarians resembles to some extent the English present perfect (cf. Karlsson 2018: 305), but it is more connected with the past and is often translated into English using the English simple past (see Mikhailov/Cooper 2016: 165–172). Finnish does not have a future tense; instead, the idea of the future is expressed lexically or sometimes by means of certain special verbal constructions (Karlsson 2018: 313). Comrie suggests that in languages like Finnish the present tense can be described as ‘non-past’, because it combines the idea of a situation happening either ‘now’ or in the future (Comrie 1985: 45). Passive verb forms are very frequent in Finnish and have both past and present forms. Finnish temporal constructions are also formed with the E-infinitive, and with past and present participles (VISK 2008: §543).

---

3 Grammatical note. In Finnish grammar there are at least three infinitive forms: A-infinitive (the basic form of the verb), E-infinitive (in relative temporal clauses) and MA-infinitive (used in verbal phrases with a finite verb as a head). See Karlsson (2018: 343–360) for details.
It is obvious, therefore that the grammar systems of the two languages are so distant that their respective grammatical tagsets do not match exactly (e.g. one infinitive in Russian vs. three infinitives in Finnish). As a result, a number of technical solutions need to be devised to make comparison possible.

Bernard Comrie points out that “… in many linguistic works, especially traditional grammars, the term tense is rather misleadingly used to cover both tense and aspect” (Comrie 1985: 6). Furthermore, tenses can be ‘absolute’, i.e. linked to a deictic centre, or ‘relative’, i.e. linked to other situations in the text. It is thus very difficult to separate tense and aspect when using grammatical annotation based on the categories in traditional grammars.

6 Querying tenses from the corpus

Morphological annotation makes it possible to query the corpus on the basis of grammatical forms. Each parallel text was queried by segment, and all verbal forms (the finite forms, the participles and the gerunds, as well as gerund-like E- and MA-infinitives in Finnish) were extracted from the corresponding Russian and Finnish segments.

The results of the search were placed in a table (see the fragment in Table 3). Only one pair of verbs can be stored in each row, and so multiple rows were added for multi-verb segments. The Russian verbs were aligned with the Finnish, but these alignments can be considered more or less reliable only in cases of one-verb segments. Where there was more than one verb in the corresponding segment, the alignment of verbs was performed pairwise in the order of appearance (see Figure 1). If the word order in the segments was different, the alignments could go wrong. However, this does not create problems for the current study, because I was more interested in the general flow of time in the documents rather than in matching particular words. The whole work was performed automatically with the help of a PHP-script I created. The script queried the corpus database, arranged the results in the desired format, and saved them in a text file.
Figure 1: Aligning tenses in parallel segments

In order to facilitate the processing of the data and to make cross-language comparison possible, the actual names of the tenses were replaced by numbers, and thus weights were assigned to verbal forms. Like any abstract weights (cf. grading student’s essay, evaluating own physical/mental state, etc.), these are assigned empirically and express the extent of “pastness”, “presentness” or “futureness”.

- Past tense: $-1$
- Past participle, past gerund (Russian): $-0.5$
- Infinitive / No verb: $0$
- Present participle, present gerund (Russian), E-infinitive and MA-infinitive (Finnish): $+0.5$
- Present tense: $+1$
- Future tense (Russian): $+2$

I assigned the zero weight to the forms with no time reference. The forms referring to the past received negative weights and the forms related to the present or the future received positive values.

When telling about events or states of affairs, the author can explicitate them in full using constructions with finite verbs, e.g. *The police used tear gas against migrants*. If the same events are viewed as part of chain of events, they can be “downgraded” to gerunds or participles, e.g. *Amnesty International condemned the authorities for using tear gas*. The function of participles and gerunds is to denote events that are from the writer’s perspective less important than those denoted by the finite forms (this is usually the main reason for choosing between a finite form and a participle/gerund to present a situation in a text); therefore, I decided to assign to them lower values than those for
finite verbal forms. The Finnish perfect form received the value -0.5 because it has closer connections to the present moment than the simple past. There are two future tenses in Russian: simple and complex, and the value +2 was assigned to both (the Russian complex future consists of the simple future of byt’ and the infinitive of the main verb, e. g. budu rabotat’ ‘I will work’, hence +2 + 0 = 2). As has already been mentioned, there is no future tense in Finnish.

7 How to study the flow of time?

The data acquired from the corpus is in the form of a table. An extract of the table containing data collected from the first five segments of a random document is displayed in Table 3. The script that collects the search results converts the names of tenses into numbers (see previous section), compares the ‘weights’ of the forms in the Russian–Finnish verb pairs, and outputs the verb pairs found in the documents for the curious researcher (form: lemma).

<table>
<thead>
<tr>
<th>#</th>
<th>Code</th>
<th>Ru</th>
<th>Fi</th>
<th>Russian Word</th>
<th>Finnish Word</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FI-RF_access_notes_2006</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>Same time</td>
</tr>
<tr>
<td>2</td>
<td>FI-RF_access_notes_2006</td>
<td>+1</td>
<td>+1</td>
<td>свидетельствует:</td>
<td>свидетельствовать</td>
<td>Same time</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>esittää:</td>
<td>esittää:</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FI-RF_access_notes_2006</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
<td>Same time</td>
</tr>
<tr>
<td>4</td>
<td>FI-RF_access_notes_2006</td>
<td>+0.5</td>
<td>+1</td>
<td>ссылаясь:</td>
<td>еhdottaa:</td>
<td>Earlier</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ссылаться</td>
<td>еhdottaa:</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>FI-RF_access_notes_2006</td>
<td>+1</td>
<td>+1</td>
<td>предлагает:</td>
<td>tekevät:</td>
<td>Same time</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>предлагать</td>
<td>tehdä:</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: An extract from the data table

The comparison of weights assigned to the verbs in the corresponding segments makes it possible to decide whether the state of affairs described by them takes place earlier, later, or at the same time as the corresponding segment. Indeed, the tense is not the same as the time (see Section 2 of this paper) and verbal tenses are only indirectly connected to the time. Still, using past tenses makes the text more “past” and using more present tenses makes text more “present”. As result, the sum of the weights of a document can give an indication of the dominant time: a positive sum would mean that the state of affairs is treated as happening in the present or future, while a negative sum means that the document is generally placed in the past (e. g. summing up the weights of verbal forms in the Russian and Finnish versions of the treaty in Table 2 will yield +5

\[ \text{N/A = not applicable.} \]
for the Russian version and +1.5 for the Finnish one, which will mean that although both texts are in the present, the Russian version is 'more in the present').

The statistics collected in this way from the texts can be visualised by plotting on a diagram the flows of time in the different versions of the same documents, and thus comparison across language versions becomes possible. In Figure 2, the flows of time in four random documents are shown as examples.

The four example documents all concern nuclear energy, but the treaties were concluded at different time periods and are not of the same length. However, in the diagrams in Figure 3, the same pattern can be observed: the Russian versions of the documents seem to be balancing between the past, present, and future tenses, while the Finnish versions deal mainly with the present. It is true that Finnish has no future tense, but it does have the perfect tense and the simple past.

Figure 2: The time flow in the Finnish–Russian international treaties (blue: Russian, green: Finnish)

The barplots summing up the numbers for the whole data (Figure 3) confirm the observations from the separate texts. Finnish texts are either ‘timeless’ (i.e. no verb form
The future becomes the present and the present becomes the past
Studying temporal structures in a parallel corpus of state treaties

was detected) or ‘happening now’, while the Russian texts handle all time dimensions from the past to the future. The number of ‘timeless’ segments in the Finnish versions of the documents is almost three times higher (Russian: 7,360 vs Finnish: 17,971).

Figure 3: Time statistics for the whole corpus

The sum of points by document is usually positive, and Finnish documents tend to score higher. One possible explanation is that past tenses are more typical in Russian documents (cf. Figure 3). The Pearson correlation of the sums of points between language versions is very strong (0.912, p-value < 2.2e–16). This is not surprising: the strong correlation simply confirms that generally speaking, in corresponding versions past is past, and present is present.

However, the mapping of parallel segments demonstrates clearly that the time in the Russian segment is the same as in the related Finnish segment in only a third of cases. It is clear from Figure 4 that changes towards the past, changes towards the future, and no changes at all are almost equally probable.
Figure 4: The time in the segments of the Russian versions of the documents compared to the time in the corresponding Finnish segments

Still, the general picture can be misleading and the comparison of time-lines in random documents (Figure 5) shows that the variation of the proportions is great. Some documents tend to preserve the time-lines in the different language versions, whereas others do not.
Mikhail Mikhailov

The future becomes the present and the present becomes the past

Studying temporal structures in a parallel corpus of state treaties

Figure 5: Time comparison: five random documents
The collected data can be cross-tabulated, and thus it becomes possible to observe what kind of temporal correlations occur between the Russian and the Finnish language versions (Table 4). We can see that only in the ‘present tense’ (=1) and ‘timeless’ (=0) segments, there are no changes more than in 30% of the cases. The exact numbers are as follows: 1:1 – 5,149 of 10,938 (Finnish) and 8,356 (Russian); and 0:0 – 5,271 of 14,971 (Finnish) and 7,360 (Russian). There is a moderate Pearson correlation between the Russian and the Finnish tenses; the coefficient is 0.358 and it is highly significant, \( p < 2.2e^{-16} \).

<table>
<thead>
<tr>
<th></th>
<th>-1</th>
<th>-0.5</th>
<th>0</th>
<th>0.5</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>34</td>
<td>442</td>
<td>686</td>
<td>104</td>
<td>182</td>
</tr>
<tr>
<td>-0.5</td>
<td>75</td>
<td>458</td>
<td>3,774</td>
<td>892</td>
<td>1,734</td>
</tr>
<tr>
<td>0</td>
<td>40</td>
<td>160</td>
<td>5,271</td>
<td>897</td>
<td>992</td>
</tr>
<tr>
<td>0.5</td>
<td>27</td>
<td>114</td>
<td>2,255</td>
<td>872</td>
<td>1,444</td>
</tr>
<tr>
<td>1</td>
<td>42</td>
<td>189</td>
<td>2,400</td>
<td>576</td>
<td>5,149</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>60</td>
<td>585</td>
<td>227</td>
<td>1,437</td>
</tr>
<tr>
<td></td>
<td>232</td>
<td>1,423</td>
<td>14,971</td>
<td>3,568</td>
<td>10,938</td>
</tr>
</tbody>
</table>

Table 4: Correlations between Russian and Finnish tenses in the documents (Russian documents by rows, Finnish documents by columns)

It is important to bear in mind that the numbers in Table 4 may be inaccurate to some extent due to parsing and verb alignment errors. To gain a better picture, let us look at some examples of the temporal shifts that take place in the treaties.

**Future → Present**

(1) По всем иным вопросам суд определит порядок своей работы самостоятельно.

‘On all other questions the court will decide the order of its work independently.’

Kaikissa muissa suhteissa välimiesoikeus päättää itse menettelytavoistaan.

‘In all other issues the intermediary court decides independently on its ways of procedure.’

(FI_RF_investment_protocol_1996)

This is a very typical transformation for Russian–Finnish texts (1,437 cases in the corpus; see Table 4). As already mentioned, there is no future tense in Finnish; the present tense is used with non-past meaning. Some contexts can be identified as present or future (usually due to lexical temporal markers), while others are ambiguous and can be interpreted as ‘now or very soon’. In most cases, especially if the present tense is

---

5 The translations of examples are literal; their purpose is to give the general idea of the meaning.
acceptable in the Russian context (e. g. it is acceptable in example 1 above), the Russian future tense is translated into the Finnish present tense. In the Russian segment, the form of the simple future is dictated by the choice of the perfective verb opredelit’ (‘to decide, define’) of the verb aspect pair opredelit’/opredelat’. If the verb opredelat’, the imperfective pair of this verb, were chosen, the tense would be the present. In this particular context, both verbs can be used (sud opredelit future vs sud opredelat present).

**Future → Participle**

(2) О таких причинах будет сообщаться консулам.
‘On such reasons will be reported to the consuls.’

Тällaiset syyt on ilmoitettava konsulille.
‘Such reasons should be reported to the consul.’

This is another case with a future tense in the Russian text, albeit a less frequent one (227 cases; see Table 4). In the Finnish version, a construction with the third person present form of olla (‘to be’) + the present passive participle is used. This construction is referred to in Finnish grammar as the necessive construction (fi nesessivirakennet); it has a deontic modal meaning of necessity, obligation, or inevitability (VISK 2008: §1580). The Russian version with the compound future tense has the more general meaning of a future event. The future indicative, as opposed to other indicative tenses like the present and the past, cannot have the meaning of something that takes place in the real world, because the future is not yet taking place. When taken out of context, the Russian passage in example (2) can be interpreted as ‘very likely to happen’, ‘planned’, ‘desired’, or ‘obligatory’ (cf. RG 1980: part 1, § 1498). In the context of a bilateral treaty, the Russian future tense has the meaning of ‘the actions decided upon in the document’, and the degree of necessity or obligation can be interpreted by the reader. The Finnish version therefore contains a stronger obligation than the Russian one.

**Past → Perfect**

(3) Приостановление течения срока доставки учитывается на той железной дороге, на которой это произошло.
‘The suspension in the counting of the shipping time is taken into account on the railway, on which it happened.’

Kuljetusmääräajan keskeytyminen luetaan sen rautatien hyväksi, jolla keskeytyminen on tapahtunut.
‘The suspension in the counting of the shipping time is taken into account on that railway, on which the suspension has happened.’
Past → Past

(4) [...] в течение срока на рассмотрение претензии перевозчик уведомил заявителя о полном или частичном отклонении претензии.

‘[…] during the time for processing the claim for reimbursement, the carrier informed the applicant about the complete or partial rejection of the claim.’

 [...] kuljetusyhtiö ilmoitti korvausvaatimuksen käsitelyajan kuluessa korvausvaatimuksen esittäjälle vaatimuksen täydellisestä tai osittaisesta hylkäämisestä.

‘[…] the carrier informed the applicant during the time for processing the claim for reimbursement, of the complete or partial rejection of the claim.’

(FI_RF_railway_2015)

Example (3) demonstrates the third most frequent case, with the past tense in the Russian version and the perfect tense in the Finnish (442 occurrences in the data). It is much more frequent than the present in Russian versus the perfect in Finnish (189 occurrences) or the past in Russian versus the past in Finnish (34 occurrences). As already mentioned, the state of affairs in the document is usually placed in the present and all other activities that happened before are related to it and as a result demand the perfect tense in Finnish. In Russian, which has no perfect tenses, such situations are expressed by the past tense, as we see in example (3). Using the past in Russian as the equivalent for the past in Finnish (as in example (4)), which might seem the most natural pairing, occurs very rarely; only 34 cases were detected by the query in the whole dataset.

Present → Perfect

(5) [...] в случае, если это авиапредприятие не соблюдает законы и правила Договаривающейся Стороны, предоставляющей эти права […]

‘[…] in the case that this air carrier does not follow the laws and rules of the contracting party that grants these rights […]’

 [...] jos tämä lentoyhtiö on jättänyt noudattamatta oikeudet myöntäneen sopimuspuolen voimassa olevia lakeja tai määreyksiä […]

‘[…] in the case that this air carrier has left without fulfilling the existing laws and rules of the contracting party that grants these rights […]’

(FI_RF_air_traffic_1993)

In example (5), the perfect tense in the Finnish document is accompanied by the present tense in the Russian version. The Russian context is generic; the present is used in the meaning of something that may happen, while the Finnish version is setting up the situation as an example case. Using the past tense in the Russian version – as in example (3) – would have been quite correct.
8 Discussion

The quantitative study of the use of grammatical tenses in the Finnish and Russian versions of the bilateral treaties shows significant differences. Although the most typical tense form for both the Russian and the Finnish texts is the present tense, there are more ‘timeless’ segments in the Finnish documents, while the past tenses occur much more often in the Russian texts. The absence of a future tense in Finnish and the absence of perfect tenses in Russian also cause differences in text flow. The study has also demonstrated that there is no exact correspondence between tenses in the language versions of the same documents. The focus often shifts in the other version from the present to the past, from the present to the future, and even – in some rare cases – from the future to the past. It is important to remember, however, that the findings are based on morphosyntactic annotation only, they are sequences of verbal tenses and not reconstructions of timelines. To obtain these, semantic annotation of the corpus would be necessary. With semantically annotated texts it would be possible to use all temporal markers and get the sequence of events.

Different language versions of the same treaty may show the state of affairs as existing in the present, as passed, as desired, or as planned: in other words, the point of view might alter. The study of random contexts has confirmed this. The possible problematic zone might be the expression of modality, because it is very difficult to express the same extent of obligation, desirability, or prohibition in language versions. The differences between the texts are likely to be the consequences of the adaptation of the text to the conventions existing in the given language and the text register. These differences are the result of a compromise between the closeness of content and readability: texts that indeed have an identical structure and contents might be very difficult to read and comprehend. Slight differences in the flavour are, however, not critical for the overall meaning of the document.

It is not possible to find out from the corpus at what stages of work this desynchronizing in temporal structure takes place. The changes in tenses could be made by translators or it might be result of updating or revision of the documents. More detailed information can be obtained only by interviewing translators and technical writers.

The contrastive analysis of bilingual documents helps to highlight points of divergence between languages. Knowledge of these potential problems makes possible to avoid them on various stages of work with such documents: drafting, co-drafting, translating, editing, proof-reading, etc. It is a good translation practice to be aware of pragmatic differences in usage of tenses in order to express modality correctly.

All in all, the results of the present investigation were somewhat unexpected: it seems that there are even more differences in the uses of tenses in different language versions of state treaties than could have been foreseen. In other text types, especially in journalistic and literary translations, and in localization, the choice of deictic viewpoint and tense would no doubt be freer than in language versions of legal documents, which are required to be identical.
The method outlined here for mapping verbal forms in parallel texts is clearly also applicable to other genres, and it may be even more interesting for studying time shifts in translations of fiction texts. This may help in finding typical correspondences in the ways tenses are used in two languages, as well as in collecting statistics on use of tenses when translating temporal constructions. With a larger dataset it may even be possible to build a statistical model for tense correspondences in parallel texts.

References


Mikhailov, Mikhail; Robert Cooper (2016): *Corpus linguistics for translation and contrastive studies.* London/New York: Routledge


Studying temporal structures in a parallel corpus of state treaties


Appendix: Sources

FI_RF_access_notes_2006: Протокол в форме обмена нотами о доступе обслуживающего персонала Саймейского канала на участок автомобильной дороги между пунктом пропуска Брусничное и пунктом пропуска Нуйямаа / Tasavallan presidentin asetus Suomen ja Venäjän välillä Saimaan kanavan ja Malajy Vuotski-jaaarten huoltohenkilökunnan kulkupamenetelystä noottienvaihdolla tehdyyn pöytäkirjaa voimaansaattamisesta. [Protocol in the form of the exchange of notes on the access of personnel of the Saimaa channel to the motorway between Brusnitsnoje and Nujamaa]

FI_RF_air_traffic_1993: Соглашение между Правительством Российской Федерации и Правительством Финляндской Республики о воздушном сообщении /Suomen tasavallan hallituksen ja Venäjän federaation hallituksen välisen LENTOLIIKENNESOPIMUS. [Agreement between Russia and Finland on air traffic connections]

FI_RF_atomic_energy_1995: Соглашение об оперативном оповещении о ядерной аварии и об обмене информацией о ядерных установках / Suomen tasavallan hallituksen ja Venäjän federaation hallituksen välisen SOPIMUS ydinonnettomuutta koskevasta pikaisesta ilmoittamisesta ja ydinlaitoksia koskevasta tietojenvaihdosta. [Agreement on the provision of information about nuclear accidents and on the exchange of information on nuclear plants]

FI_RF_investment_protocol_1996: Протокол о внесении изменений в Соглашение между Финляндией и СССР о содействии осуществлению и взаимной защите капиталовложений / Pöytäkirja muutoksista Suomen Tasavallan hallituksen ja SNTL hallituksen väliseen sopimukseen sijoitusten edistämisestä ja molemminpuolisten suojelust. [Protocol on changes in the agreement on cooperation in protecting investments]

FI_RF_railway_2015: Соглашение о прямом международном железнодорожном сообщении / Sopimus suorasta kansainvälisestä rautatieliikenteestä. [Agreement on direct international railway connections]

FI_RF_railway_rules_2005: Условия перевозок пассажиров, багажа, товаров и грузов в российско-финляндском прямом железнодорожном сообщении / Suomen ja Venäjän rautatietyhdyssäliikenteen matkustajien, matkatavaroiden, kiitotavaroiden ja tavaroiden kuljetusehdot. [Conditions for the transportation of passengers, luggage and goods by railway]

FI_USSR_atomic_energy_1969: Соглашение о сотрудничестве в области использования атомной энергии в мирных целях / SOPIMUS yhteistyöstä atomienergian rauhanomaisen käytön alalla. [Agreement on cooperation in using nuclear energy for peaceful purposes]

FI_USSR atomic_energy_1987: Соглашение об оперативном оповещении о ядерной аварии и об обмене информацией о ядерных установках / Sopimus ydinonnettomuutta koskevasta pikaisesta ilmoittamisesta ja ydinlaitoksia koskevasta tietojenvaihdosta. [Agreement on informing on nuclear accidents and on the exchange of information on nuclear plants]

FI_USSR_consular_treaty_1966: КОНСУЛЬСКАЯ КОНВЕНЦИЯ между Финляндией и СССР / Suomen Tasavallan ja Sosialististen Neuvostotasavaltaan Liiton välinen KONSULISOPIMUS [Consular treaty between Finland and the USSR]

FI_USSR_consular_treaty_protocol_1966: ПРОТОКОЛ к Консультской Конвенции между Финляндией и СССР / Suomen Tasavallan ja Sosialististen Neuvostotasavaltaan Liiton väliseen konsulisoopimukseen liittyvä PÖYTÄKIRJA. [Protocol of the Consular treaty between Finland and the USSR]

FI_USSR_german_funds_article_1947: Протокол к статье 9 Соглашения о германских денежных средствах / Pöytäkirja saksalaisten rahavarojen käytöstä 9 artiklaa. [Protocol on article 9 of the Agreement of the funds of former Nazi Germany]
The future becomes the present and the present becomes the past

Studying temporal structures in a parallel corpus of state treaties

Mikhail Mikhailov is Professor of Translation Studies (Finnish and Russian) at Tampere University, Finland. He is one of the authors of the book *Corpus linguistics for translation and contrastive studies* (Routledge 2016). He compiles multilingual corpora and develops web-based corpus software. His research covers corpus-based translation studies with a particular focus on parallel and comparable corpora, terminological studies, and translation technologies.

E-mail: michail.mikhailov@tuni.fi
Neu bei Frank & Timme


**TRANSÜD. Arbeiten zur Theorie und Praxis des Übersetzens und Dolmetschens**

Herausgegeben von Prof. Dr. Klaus-Dieter Baumann, Dr. Susanne Hagemann, Prof. Dr. Dr. h.c. Hartwig Kalverkämper, Prof. Dr. Klaus Schubert


**Easy—Plain—Accessible**

Herausgegeben von Prof. Dr. Silvia Hansen-Schirra, Prof. Dr. Christiane Maaß


**Kommunikation – Partizipation – Inklusion**

Herausgegeben von Jun.-Prof. Dr. Bettina M. Bock, Prof. Dr. Ulla Fix, Prof. Dr. Nathalie Mälzer

Kerstin Alexander (Hg.): Mit Typografie und Bild barrierefrei kommunizieren. Forschungsstand und Studien. ISBN 978-3-7329-0584-3

**FFF: Forum für Fachsprachen-Forschung**

Herausgegeben von Prof. Dr. Dr. h.c. Hartwig Kalverkämper


**KKK: Kulturen – Kommunikation – Kontakte**

Herausgegeben von Prof. Dr. Dr. h.c. Hartwig Kalverkämper


**TTT: Transkulturalität – Translation – Transfer**

Herausgegeben von Prof. Dr. Dörte Andres, Dr. Martina Behr, Prof. Dr. Larisa Schippel, Prof. Dr. Cornelia Zwischenberger


Frank & Timme

Wittelsbacherstraße 27a, 10707 Berlin • Telefon +49 (0)30 88 66 79 11 • info@frank-timme.de, www.frank-timme.de