Batumi Shota Rustaveli State University
Faculty of Humanities
Department of European Studies

With the right of manuscript

Modeling of the Analysis of Legal Text Translation in Parallel Corpus

Specialty - Translation Studies

Khatuna Diasamidze

Presented for the academic degree of Doctor of Philology
Dissertation

Annotation

Batumi – 2021

Introduction

Corpus research has become quite relevant in modern translation studies today. The relationship between translation and the corpus dates back to the last century. Over the years, with the development of computer technology, corpus-based translation studies has expanded and gradually covered different areas, forming different types of parallel corpora. In particular, "parallel corpus means the placement of completely identical texts in two or more languages in a specially designed computer program and the determination of their equivalence based on comparison and analysis" (Yang, 2002: 29). The parallel corpus is of particular importance for translation studies, as it unites source and target languages at the level of both texts and sentences and words, simultaneously presenting them to the translator and facilitating their work.

On the other hand, there were national corpora created for specific purposes, such as, for example, English or Georgian language national corpora.

They are composed of academic linguistic resources and play a great role in the research development in the digital humanities determining the linguistic or stylistic features of this or that text. There are translation corpus dictionaries, e.g., the large English-Georgian dictionary www.dict.ge created under Professor T. Margalitadze's leadership at the TSU Lexicography Center, which is based on the principle of corpus

systematization of definitions. There are also several fields (medical, legal, political, etc.) corpora.

Gradually, the common objects of research were revealed, which led to the creation of parallel corpora on the basis of the above-mentioned corpora, which is especially important for translation studies.

Legal corpora today are not finally established and widespread, which is due to the complexity and peculiarities of the field itself. It is this fact that led us to use the methodology of parallel corpus linguistics and teleological approaches to the interpretation of the law to research pages 1-4 of Chapters 1-24 of the EU-Georgia Association Agreement and made us decide to create a Georgian-English parallel legal corpus. This attempt is another step forward in modern Georgian corpus linguistics, which is in its infancy. It can be assessed as a minor parallel legal corpus.

When creating the corpus, we face many challenges, such as skills in modern technologies, human resources, but the most difficult among them is data collection. It would be quite difficult to obtain bilingual matching legal texts if we take into account the privacy of a given style.

Eventually, we opted for the Association Agreement, which is translated into many languages by professional translators and the texts are also non-confidential, on the contrary, available to both professionals and any interested person.

Short description of the dissertation:

A) Aim of the research: The research aim was to create a precedent for modeling the analysis of the translation of legal texts in a parallel corpus.

Accordingly, we have defined several research objectives:

1. Integrate translation research methods with corpus linguistics theory and practice methods;

2. Primary, contradictory non-machine analysis of bilingual texts, which includes several interrelated processes: a. Establishment of syntactic, morphological, terminological qualitative data, b. Identify translation problems; C. Defining structuring and marking models based on the above analysis;

3. Similar placement and manual marking of a parallel corpus of bilingual texts;

4. Performing technical work for processing marked texts in the translation analyzer;

5. Further analysis of statistical indicators obtained by the analyzer parallel lexical data, and parallel structured texts in the field of translation studies.

Following the objectives of the research, in the next stage of study of the theoretical foundations of corpus linguistics, a corpus of empirical material was created, which consists of parallel legal texts in Georgian and English languages.

The parallel corpus can contain a text and its translation at the same time, allowing us to create an individual model of translation analysis of legal texts for the parallel corpus of English and Georgian texts.

We conducted the research in automatic and semi-automatic mode(s) after pre-processing of the empirical material. We have used pragmatic, semantic, and terminological approaches in the research.

B) Urgency, novelty and significance of the issue:

The urgency, novelty and importance of the issue are determined by (1) the object of the research and (2) the methodology of the research.

The creation of linguistic resources for marked and annotated texts has become a major scientific direction for language learning and research over the last decade. Technologically developed collections for the scientific analysis of linguistic materials provide information about natural languages. The parallel corpus is of particular interest in this regard. The translation corpus should enable the researcher to solve a specific research issue. In other words, a nonspecific corpus researcher can select and use corpus texts for different types of contradictory research.

Adequate translation of legal terms is a general challenge since there are conceptual distinctions between the terms in the SL and TL languages. Therefore the conceptual differences of the terms could potentially lead to misinterpretation of the law.

Another novelty of the research is the technological database created in the research process as an adapted platform for specialized language corpora. It ensures the development of a parallel corpus of the Georgian language in general. It is also a unified work platform for interested professionals, as is based on our own research model.

C) Research methods and its novelty

We used only part of the research texts for non-machine analysis (Titles 1-4, pp. 1–278). In accordance with point 2 (c) of the research task, we identified translation problems in the section selected for analysis and defined the following four categories of translation inaccuracies:

1. Use of calques and the resulting intellectual ambiguity, in particular, calques used in translation of syntactic constructions and the resulting intellectual ambiguity;

2. Use of calque in translation of a definite lexeme;

3. Improper use of loans instead of Georgian equivalents;

4. Adequacy of translation of legal terms.

Therefore, their analysis is our sub-task in the present research.

In the next stage of elaboration of the theoretical bases of corpus linguistics, we defined the marking units corresponding to the objectives of our research from the TagSet of linguistic marking compiled by the head of the community for mechanical analysis of Georgian and English texts. The methodology for the study was defined by the following classification: 1) Identification of research linguistic and translation units - during the manual processing process; 2) Determining the model of statistical analysis - during the manual processing process; 3) Predictive analytics of terms - based on the results obtained by the corpus software analyzer; 4) Qualitative translation analysis - based on the results obtained with the corpus software analyzer. In the next stage,

we started the manual processing of the bilingual texts selected for the research, which can be described as a preliminary analytical process for machine analysis. Marked texts were placed in the software analyzer, based on the obtained data we conducted a scientific research in the field of translation studies.

In the next stage, based on the theoretical foundations of corpus linguistics, we defined the marking units corresponding to the objectives of our research from the TagSet for linguistic marking compiled for the machine analysis of Georgian and English texts by the dissertation supervisor.

At the next stage, we started the manual processing of the bilingual texts selected for the research, which can be described as a preliminary analytical process preparing for the machine analysis. The marked texts were placed in a software analyzer of BSU (www.corpus.bus.edu.ge). Based on the data obtained, we conducted a scientific study in the field of translation studies.

We obtained the following categorization of the research methodology:

1) Pattern Recognition;

2) Predictive Analytics;

3) Visualization of analytical data in the analyzer

4) Statistical analysis;

5) Qualitative analysis;

The use of computer technology in the development of linguistics naturally gave rise to the need to create parallel corpora and to study research by the new methodology. At

different stages of the field's development, translation research directions were based on contrasting literary studies and linguistics, and from the 1990s, the study of translation in the context of cultural studies began. Despite the variability of the urgency of the research directions, the working principle remains unchanged for the formation of valuable scientific conclusions: Extensive empirical material: search, collection and analysis of the target texts for original research. The research uses a specialized research corpus search program, through which it is possible to find terms, open contexts, determine contextual meanings, reverse analysis of definitions in a short period of time.

Language and translation: In the era of globalization, language develops mainly through translation. Naturally, during the last twenty years, it has been influenced by such new concepts and their terms that did not exist before, not only in the thesaurus of the Georgian language but also in the field of concepts of our worldview. Given that the correct establishment of new concepts is the mission of translators, they have the greatest responsibility in the development of language, as their terms only transfer them to our field of concepts and they are embedded in our thesaurus. If we imagine this process as a continuous social chain of the use of new concepts established in language and the language itself, the mistranslations enter the discourse of the mass media and create not only ambiguous but faulty syntactic constructions as well.

Structure and volume of the research: The research includes an introduction, four chapters, eleven subchapters, a

conclusion, references, internet sources and appendices. The research contains 174 printed pages.

Summary

Chapter I. Corpuslinguistics as an Innovative Scientific Direction - consists of two subchapters:

1.1. The chapter reviews corpus linguistics as a scientific field, as a relatively new direction of humanities, which arose on the basis of a combination of two fields of science: humanities and informatics. It, as a kind of "symbiosis", developed quite rapidly in parallel with the development of theoretical linguistics and informatics, and soon became an interdisciplinary scientific field. The question of whether corpuslinguistics is an independent scientific discipline and if so, what is the difference between corpuslinguistics and traditional linguistics causes differences of opinion in the scientific circles.

The issue is based on the fact that corpus linguistics in the International Classification (DDC - Dewey Decimal Classification, one of the most common international classifications of systematization of scientific fields, developed by American researcher Melville Davey and UDC - Universal Decimal Classification (created by Paul Otlet and Henry Fontaine (1895) as an independent scientific discipline. "Linguistics has its own code (4 in DDC, 8 in UDC), and computer science has its own code (in 004 DDC and in UDC as well) ”(Tandashvili, 2014: 13).

Corpus linguistics is a relatively new field today and is still in the process of research and development, so its definitions are also heterogeneous. We have considered the explanations proposed by different researchers that have been formulated from different viewpoint We have tried to conduct a comparative analysis of several of them to enable us to reconcile the most acceptable and relatively exhaustive explanations at our discretion. It is also noteworthy that in the study of most theoretical works in corpus linguistics we came to the conclusion that they were written by practicing corpus linguists, which is proof that in interdisciplinary sciences the line between theory and practice has been erased. If in the early period linguistic studies were mainly limited to theory and were they introduced into practice only later onn(i.e. the top-down process), the theoretical work of corpus linguistics relies directly on practical studies (bottom-up process). Moreover, the theoretically proved material is still intended for practical use ("bottom-up-down").

Let us return to the definitions of corpus linguistics proposed by different researchers and their analysis:

Part of the researchers consider corpuslinguistics to be a method that allows us to study linguistic problems based on digital texts. In order to conduct research, it is necessary to digitize the written and oral samples of the language and annotate them according to the parts of speech. For the most part, some aspects of research require the processing of large amounts of data, which requires a more detailed, automated annotation (Computer Linguistics Portal

http://www.computerlinguistik.org/portal/
portal.html?s=Korpuslinguistik).

From the above definitions of corpus linguistics it is clear how difficult it is to agree on its unambiguous determination, although based on them we can gather individual points which, in turn, determine the purpose of corpus linguistics, namely:

Corpus linguistics relies primarily on authentic linguistic resources, 2) digitized resources, 3) uses large volumes of empirical material.

1.2. The definition of the corpus - When talking about corpus linguistics, we consider it relevant to analyze the concept of 'corpus' itself. The word "corpus" (pl. "corpora") is of Latin origin and means "body" in Georgian. It seems strange what it has to do with linguistics. If we go deeper, the 'corpus' in this case is associated with one whole. It is no coincidence that, just as the body is associated with a living creature, the corpus is a kinetic set of digitized elements of language, not just an inactive, "lifeless" static set of linguistic elements.

Let us consider the term "corpus." It has several meanings in English. In linguistics, we understand the term 'corpus' as a collection of texts that will be used for linguistic analysis. "The language observed in the building is natural" (Togenin - Bonnell, 2001:2). Sinclair defined the corpus as "a collection of naturally occurring linguistic texts that are characteristic of a state or a diversity of languages" (Sinclair, 1991:171). In 2005, he came up with a more accurate definition of corpus, because "corpus is a collection of texts in digital form, selected according to external criteria and, as far as possible,

representing language diversity as a source of data" (Sinclair, 1991: 171).

Like any research, methods are of paramount importance in case studies. Because the corpus relies on Big Data, it is impossible to process them using traditional research methods, it even exceeds the common intellectual capabilities of many researchers. It becomes necessary to use special computer programs. Corpus linguists (S. Shreibman, D. Beri, M. Tandashvili, R. Khalvashi, and others) characterize big data according to various features, of which we consider the following to be especially important for corpus research: volume, velocity - speed of data processing , variety, veracity, and variability.

The same corpus linguists recognize different methods of big data analysis, of which we will single out the following: Data Mining, Crowdsourcing - categorizing and enriching data with the participation of many people on the basis of the public offering; Data Fusion and Integration, Machine Learning, Pattern Recognition, Simulation Modeling, Spatial Analysis, Statistical Analysis, Visualization of analytical data.

In the present research, we primarily use the Data Mining method. It is noteworthy that the term data mining (which belongs to G. Piatetsky-Shapiro, 1989), like Big Data, was created by analogy with minerals. Translated from English it means mining. Indeed, the colossal amount of information will provide the material needed for the study. Through data mining, it is possible to find, process, and calculate previously less-studied, often unknown, practically usable, and interpretable data.

Our goal is not just to classify existing groups, but to identify new groups and patterns in large databases. This is the stage of machine processing of the knowledge discovery process in databases. The latter allows us to generate, verify and clarify hypotheses.

In data mining, we rely on modern digital methods and traditional theoretical disciplines that are widely used by practicing or theoretical corpus linguists, such as Databases - a set of electronic information, digital objects, through which we search and process data automatically; Artificial Intelligence - a feature of an automated system that has certain functions of human intelligence; Machine Learning - specially designed computer algorithms that can automatically improve (smarten) the work process, which is programmed in it from the beginning; Algorithmization - using pre-compiled algorithms to solve a given task via a computer. Data mining is a rather long and time-consuming process that involves several stages. Corpus linguists distinguish between different stages, from which we will distinguish the following main ones: Discovery, during which certain patterns are revealed by the deductive method - from the private to the general; Validation - confirmation that we have created the right product; Verification - at this point we confirm that we have created the product as we planned; Predictive Modeling - the patterns detected at this stage are used to predict unknown significance; Forensic Analysis - at this stage we identify the anomalies found in the patterns and explain them (Khalvashi, 2018: 38). This research covers the above stages of data mining, the

analysis of which and the final result will be presented in the following chapters.

The next, no less time-consuming and hard process of data mining is text encoding, i.e. structuring the extracted data. Marking is considered to be the most common means of structuring data in the corpus. Marking can be interpreted as notifying, therefore, it involves the use of tags that are attached to words and phrases. Different forms of tags are used. Visually the tags contain vowels and surrounding symbols in the form of angular parentheses (<...>). Tags are usually represented in pairs, for example: <n> and </n>. The first is considered as the initial tag and the second as the finite. The text - the analytical material, which is referred to as the content - is placed between the initial and final tags. The content in the tags is written in both large and small vowels, for example: <n> or <N>, although this does not significantly differentiate them from each other. Tags usually have a hierarchical structure and are embedded in the text. In our research, we developed special tags, the use of which was based on the labeling of analytical texts for a pre-created software analyzer.

Consequently, any marking can be considered a means of interpreting the text. Through some interference with the text, we load additional encoded information into it, which, in turn, allows us to analyze it, conduct research, compare it with other texts. Marking is therefore an important element of corpus linguistics, as it is a key means of structuring a text.

As we have mentioned, tagging is quite hard, time-consuming and, at the same time, interesting process. It takes place in several stages.

Taking into account the above considerations, the first step in tagging is to select a scheme and model according to the task at hand, i.e. what kind of analysis we intend to make in the text, such as cultural characteristics, gender balance, or, in case of our research topic, the problem of translation equivalence. A set of tags should be selected considering how relevant it is both to the topic of the text and the research work set. For example, in the case of our research, during automatic and semi-automatic processing of the material to be inserted in the corpus, we perform the phasing in three stages: 1. Morphological analysis, 2. Syntactic analysis, 3, Semantic analysis. The following chapters discuss specific examples, their analysis and conclusions.

Chapter II. Proved methodology for the research of legal parallel texts based on foreign models - consists of four sub-chapters:

2.1. Translation quality standards and methodological recommenddations of the Directorate-General for Translation of the European Commission - since the research direction of our dissertation concerns English and Georgian texts of the Association Agreement and the translation equivalence analysis, we consider it advisable to study translation quality standards and methodological recommendations which in the

future, after association, will become a compulsory manual for standardization of translation quality in Georgia;

The Directorate-General for Translation of the European Commission imposes quality management frameworks in which translation must meet the needs of a certain client, other European Commission partner organizations or users and meet their expectations (DGT Translation Quality Guidelines, 2015:1). Special methodological recommendations have been developed that define the objectives of the translated text, the expected risks, and the specific quality requirements. DGT itself offers assistance to translators and editors to ensure goal matching.

All official languages have the same official status, requirements, and tasks that apply equally to each of them. This means that in 23 cases out of 24 official languages, the requirements and tasks must be fulfilled through translation, thus the translator is the ”de-facto” creator of the text.

A similar level of translation is referred to as "institutional translation" or "multilingual legislation" (DGT Translation Quality Guidelines, 2015:2). Consequently, translators are the "institutional voice" of their languages and create equally authentic texts. Thus, the existing metho-dological recommendations are intended not only for creators of source texts but also for translators and proofreaders. The main requirement for quality is the creation of texts that are read like the source text in all languages.

However, if we go deeper, the complete similarity of the translation can be achieved only marginally, since 24 to 23

texts are translations and rely on the source text. It is, therefore, necessary to define the degree of this dependence on the formula recommended by the Directorate-General for Translation and the similarity of translation mean: "as far as possible" in practice. In addition, it may be different in different languages, since the methodological instructions are not the same in different texts and languages.

The Directorate-General for Translation distinguishes between two types of quality control (DGT Translation Quality Guidelines, 2015:35):

• Revision, which means the bilingual study of the target and source texts to determine their relevance to the objectives set;

• Review, which means a monolingual study of the target text to determine its relevance to the objectives.

We will deepen the methodological guidelines for legal documenttation related to the following types of documents:

1. EU legal acts: treaties, regulations, directives, decisions, recommendations, opinions, international agreements;

2. Documents used in administrative and judicial proceedings and investigations, such as, crime, antimonopoly, state subsidy, or other cases, for example, lawsuit, official letters, counterclaim, etc.;

3. Procurement documentation and funding programs, tenders, grant applications and contracts;

4. Employment Certificates, Competition and Testing Documents.

Each translated text must comply with the basic principles set out in international standards (ISO 17100) and the quality requirements of a professional translation.

As for EU legal acts, they have legal power, namely: they define rights, obligations, the expected consequences of the law. The reader, whether an ordinary citizen or a representative of the court, must be fully confident in the reliability of the information contained in the document and act accordingly. All language versions of legal acts must be equally authentic. They must contain the same content in all languages and have the same legal power. It is not allowed to have more or fewer advantages for different citizens, companies or member states, which will be due to translation errors or inconsistencies between the versions available in different languages.

The terminology should be mutually corresponding, in particular, the same term without synonyms and additional formulations should be used in the legal act as well as in any major or parallel act.

The creation of a new term is permissible only as a result of conceptual analysis, which leads to a teleological understanding of the term or phrase.

As for the conceptual area of the term, it should remain unchanged. Taking into account the specific nature of EU law, we must take caution when using concepts and terminology specific to the legal system of an individual nation.

Understandable and correctly translated laws are the undisputed prerequisite for a democracy in action. Versions in all languages must comply with general legal principles, such

as equality before the law (law must be equally accessible and comprehensible to all) and credibility of the law (it must be possible to predict the expected results) (DGT Translation Quality Guidelines; 2015:6).

2.2. **Acquis Communautaire** - Parallel corpus in 20+ languages - According the aims and objectives of our research, we considered it appropriate to discuss the corpus of the Acquis Communautaire Association Agreement, as the Association Agreement with Georgia as a legal document will be part of it; Within the framework of the dissertation, the object of our research is the English and Georgian texts of the mentioned association agreement and the analysis of the translation equivalence.

The Acquis unites EU treaties and laws (directives, regulations, decisions), declarations and resolutions, international agreements and court's decisions. It also includes Joint Acts of the Governments of the Member States of the European Union in judicial and internal affairs (freedom, security and the judiciary), as well as Joint Acts of Foreign and Security Policy.

The Acquis is currently translated into 22 official EU languages, which exist alongside the source text as parallel texts.

Nowadays, as far as we know, JRC-Acquis is the largest among the existing multilingual corpora, which includes more than twenty languages. It is available in XML format and contains selected texts and ratio information for more than 190 language pairs.

An additional feature of the JRC-Acquis is that it is possible to classify texts according to subject characteristics using a special classification system - EUROVOC Thesaurus (a classification system that includes more than 6000 hierarchical groups). Knowledge of the subject characteristics of the text makes it possible to create a list of specific terminology, as well as test the document classification software and the automatic indexing system.

2.3. **Translation of the Association Agreement in Serbia** - In this research, we consider it appropriate to consider the corpus linguistic methods used by researchers in various EU countries to translate and corpus-linguistic processing of the Association Agreement; In particular, for our research, it is advisable to discuss M. Vukcevic's research article entitled: EU directives and standards on translation and their application in Serbia, a non-EU country undergoing a stabilization and association process. M. Vukcevic writes that a translation error in the Association Agreement with the EU cost the country 20 million euros.

We considered that this fact required a detailed analysis in our topic in order to (1) take into account the experience of Serbia and avoid its negative result, a financial precedent; (2) once again raise the issue of the need to develop a proper strategy for translating legal documentation terminology;

As in the case of Serbia, the sequence of translation editing causes problems, as evidenced by the analysis of the researcher Berteloot (Berteloot, 1999:101). The unfortunate consequences of the procedure for working on an association

agreement, namely the low quality of translation, scholars believe, is that linguistic editing did not precede legal editing, which is highly necessary because once jurists decide to convey this or that structure more accurately, linguistic change or correction is almost impossible (Pescatore, 1999:93).

2.4. **Anglo-Slovenian Corpus of the Association Agreement with the European Union** - Many countries can be mentioned when talking about the problems occurred during the translation of the Association Agreement; It is worth analyzing countries with languages spoken in a small area, highlighting the issue of exact equivalence and ways to solve it. One of such examples is Slovenia. The Slovenian-English parallel corpus officially includes 15 original texts and their translations, for a total of 1 million words. Their digital field terminology is sorted by corpus methodology and consists of different sub-corpora. The largest (33%) terminological base in the corpus structure includes legal documents and related documents related to Slovenia's association with the European Union.

While analyzing the Slovenian corpus, the scholars came to the conclusion that it was necessary to create a separate corpus for the directly contextual meanings of the phrases in the association agreement. What does this mean throughout our dissertation? The fact is that the approach of Slovenian scientists to the meaning of the term and phrase confirms the correctness of our view of the issue, because they indirectly but still focused on the teleological meaning of the term and the phrase when they raised the issue of the need to create a

separate corpus. This corpus was composed not of the lexical but of the contextual meanings of this or that term or phrase, which they acquired directly in the Association Agreement; These meanings can be understood only through their teleological analysis and understanding, that is, through the contextual meaning acquired in the direct association of legal terms and phrases. This will make the job easier for linguists as well as terminologists and translators. Thus, creating electronic databases of phrases, terminology, is primarily beneficial for them.

Chapter III. **Teleological approaches as method of interpretation of legal texts** - consists of three subchapters:

3.1. **The notion of teleology**; **Teleological interpretation of the text** - Teleology originates from the Greek word "telos" (last, ending). Every word, like a living body, has a certain "end" or purpose, function. Some scholars associate the term "teleology" with the name of the philosopher Christian Wolf. The latter viewed it as "part of a philosophy that reveals the end or aims of things" (Glasersfeld, 2009:17). Teleology considers and relies on the best, highest value, perfection to which the world aspires. This is the main purpose of the existence of the universe: the universe and everything in it, material and nonmaterial, is constantly evolving and changing for the better.

Most philosophers from the second half of the nineteenth century admitted the existence of an ultimate goal and discussed in their works the advantages of teleological analysis

(Wevel, Herschel, and others). The study of purposeful behaviors and the ultimate goal became the cornerstone of their research, thus explaining the purposeful behavior of humans and other living beings (and not only living).

Particularly interesting and important in the research process in the parallel mode of the Association Agreement was the teleological analysis, which ensured that the existence of each word could be explained for a specific purpose. We have repeatedly mentioned that every word used in translation, like any living being, must have a certain function. From a teleological point of view, one of the main features of the legal documentation is that each term must be encoded from the very beginning with certain information and, therefore, it must have specific goals and purposes. With this in mind, the primary task of the translator should be not just to translate legal terminology, but to select words for the target text that encode similar information to the words used in the original. Moreover, they should have the same goals and purpose. Given the large number of translation inaccuracies identified by the software analyzer in our research, we are convinced that a large number of legal terms are translated from lexical viewpoint, neglecting the purpose encoded in it.

The judge must interpret the law, for this he can not rely on the individual words contained in it, which may have different meanings and cause intellectual ambiguity. The best solution is a teleological interpretation, a kind of philosophical approach based on the aims and directions of this or that law and not by the meaning of individual words or phrases.

Teleological interpretation can be defined as "the method used by the court when interpreting legislative acts in terms of set goals, values, legal, social and economic objectives."

In order to better present the teleological interpretation of the law, we consider it appropriate to discuss R. Jacobson's theory of interpretation (Beridze, 2018:16). It distinguishes between three types of interpretation of a linguistic sign, of which we will distinguish between the interpretation of the law and the interpretation of verbal signs or the interpretation of verbal signs in the same language with other signs.

Drafting laws is the prerogative of the legislature, and the judge has the full right to interpret them. According to the teleological interpretation, he must reconcile the understanding of the law and the specific case under consideration. He interprets the law in his native language, i.e. there is an internal translation.

In the case of international legal acts, such as the Association Agreement we have analyzed, no less important is the inter-linguistic interpretation, or the translation itself. It is necessary to maintain the equivalence between the original and the source text so that the judge can correctly determine the purposes encoded in it and adapt it to the specific case under consideration, i.e. to carry out a teleological interpretation of the law. In the teleological study of legal parallel texts, we have been guided by the view that the specific purpose coded in each term serves to direct human consciousness towards the rule of law, and its ultimate goal is the formation of a just society. Therefore, judging from a teleological point of view, we can

conclude: laws are created not just to know, but to obey, to live in a law-abiding, just society, which is the supreme goal of society. According to all the above mentioned, every mistake or inaccuracy made in the translation, the selection of an inappropriate equivalent of this or that term, is not just a translation error, but there is a danger of misinterpretation of the law. Especially when it comes to such a level of intergovernmental legal document as the translation of an association agreement. According to this agreement, Georgia should become part of the Common European Legal Area, so the goal encoded in each legal term translated into Georgian should correspond to the goal encoded in the original term.

3.2. **Peculiarities of Pragmatic Interpretation of the Text** - Since teleology is closely related to pragmatics, in this subchapter we considered it appropriate to analyze the peculiarities of pragmatic interpretation of the text. While talking about the pragmatic meaning, we discussed the views of various scholars on the term pragmatics itself. It is widely believed that its modern use is associated with the philosopher Charles Morris (Levinson, 1983:2). He singled out pragmatics as one of the directions of semiotics (the science of signs) as a study of the relation of interpreters of signs.

According to Morris, pragmatics is a very broad concept that studies the ongoing psychological, biological and social processes in the functioning of signs. Such an understanding of pragmatics is very broad, and in sharing this position it would have included aspects of language that are now being studied

by the independent disciplines of linguistics: psycholinguistics, sociolinguistics, and others. On the other hand, the frameworks of pragmatics have been narrowed, which is associated with the name of the philosopher Carnap (Levinson, 1983:2). In his opinion, if the interlocutor is openly mentioned in the study of the meaning of the expression, in such a case the analysis is carried out in the context of pragmatics. The position that pragmatics is the study of aspects of language that inevitably require reference to the language user leads to a much narrower understanding of the term. According to Levinson (Levinson, 1983:7), such an explanation of pragmatics would make it impossible to make a thorough study of the use of language. He argues that if the existence of a trichotomy (syntax, semantics, pragmatics) is important for linguistics, then it is necessary to establish an appropriate framework for pragmatics so that linguistic research covers essential aspects of context. Context should include time and space options for the personalities of individuals involved in communication, their beliefs, knowledge, education, and individual intentions. Thus, it is important to find a definition of pragmatics that sets the appropriate research framework for this discipline and enables a full and consistent study of its meaning.

According to one of the definitions suggested by Yule (Yule, 2000:12), pragmatics is the study of meaning conveyed by the speaker or writer and interpreted by the listener or reader. Consequently, pragmatics is thought of as the analysis of an intention embedded in an expression expressed by an individual, and not as an analysis of the words or phrases

contained in that expression itself. According to the following definition of pragmatics proposed by Levinson, this discipline is the study of language from a functional point of view, and it is an attempt to explain the connection of linguistic structures with extralinguistic causes and factors. Levinson himself offers the following definition: Pragmatics is the study of the interrelationships between language and context that are grammatically or codified in the structure of language.

There are certain expected or suitable circumstances called successful communication conditions. These conditions of appropriate or successful communication are necessary for individuals involved in communication to interpret the verbal action correctly and appropriately. Some of these terms are, of course, common to any kind of linguistic communication. For example, it is a prerequisite that the speaker and the listener understand each other, for example, they must speak the same language. It is also essential that both individuals have hearing skills and so on. In some cases, the verbal action will be unsuccessful if the action is not uttered by the proper person. for example: –You are adjudged guilty! (You are guilty!)

If this expression is not uttered by the judge in a proper environment, in this case during the trial in the courtroom, the expression will certainly convey information, but it is completely deprived of illocutory power.

3.3. Context and the problem of translation - According to pragmatic viewpoints of communication, context is constantly expanding, although empirically it is delimited (it

has boundaries). According to Stalnaker (Stalnaker, 1999:27), every presumption uttered by the speaker is a delimitation of context as background information. However, according to Sperber and Wilson (Sperber, Wilson, 1986:57), context is delimited by the empirical fact that some opinions in the mind of the speaker are activated by the activation of other additional opinions and associations. Sharing this opinion means admitting that the more details are in the context, the greater is the ability to better understand and comprehend verbal action. This means, as mentioned above, the analysis of verbal action in accordance with the whole world experience, which, of course, we find impossible. Dependence on the context of meaning is discussed by language philosophers in terms of the analysis of index, demonstrative, descriptive words, and whole propositions. Context dependence must have some source, which is clearly seen in the fact that the use of language is mostly situational. In fact, the user of the language is responsible for using the language in the context of any action. Such situational character of verbal action requires a limited context. If the performance and comprehension of a verbal action according to any context is a situational event, we believe that this verbal action should be evaluated according to the same context. Moreover, in order for a verbal action to be evaluated, the context itself must be delimited, that is, it must be clearly defined what belongs to the context and what does not. A comprehensive context fails to provide situational character and, as J. Austin explains that he makes verbal action

"an ideal expression that would be suitable for any situation, for any purpose."

Chapter IV. Annotation, structuring, tagging and corpus research of legal parallel texts - consists of two subchapters:

4.1. **Modeling, annotation, structuring and tagging of English and Georgian legal parallel texts** - a technological database created in the research process as an adapted open platform for specialized language corpora. It ensures the development of a parallel corpus of the Georgian language in general. It is also a unified working platform for interested professionals, equipped with our own model of research.

The research uses a specialized corpus research program, through which it is possible to find terms, open contexts, determine contextual meanings, reverse analysis of definitions in a short period of time.

Like any other research, corpus research methods are of paramount importance in case studies. Because the corpus relies on Big Data, it is impossible to process them using traditional research methods, it even exceeds the common intellectual capabilities of many researchers. It becomes necessary to use special computer programs. In this research, of the various methods of data analysis, we primarily use the Data Mining method. Data acquisition was rather hard and time-consuming process, which we carried out in various stages tested in corpus linguistics: Discovery, during which we tried to reveal certain patterns by the deductive method - from the

private to the general; Validation - to prove that we have created the right product; Verification - at this point we have confirmed that we have created the product as we planned; Predictive Modeling - we used the patterns identified at this stage to predict unknown significance; Forensic Analysis - at this stage, we have identified and explained the anomalies found in the patterns.

The next, no less time-consuming process of data mining was text encoding, or structuring the extracted data, for which we used marking - the most common means of structuring data in the body. As already mentioned, marking (mark-up) can be defined as notifying, therefore, we used tags for marking, which we attached to words and phrases. We have developed different forms of tags that visually contain Latin letters and surrounding symbols in the form of angular parentheses (<...>). They are represented in pairs, for example: <lg> and </lg>. The first is considered as the initial tag, and the second as the finite. The text - analytical material, i.e. content - is placed between the initial and final tags. The contents of the tags are written in lower case, for example: <lg>. The tags we developed are hierarchical in structure and built into the text, using which we based the labeling of analytical texts on a pre-created software analyzer.

As we have mentioned, tagging is quite hard, time-consuming and, at the same time, interesting process. We carried it out in several stages.

According to the above considerations, in the first stage of tagging we tried to select the scheme and model according to

the task, i.e. considering what kind of analysis we wanted, in particular, in the case of our research topic - the problem of translation equivalence in parallel texts. We have compiled a set of tags considering how relevant it is to both the topic of the text and the research project. For example, in our case, during the automatic and semi-automatic processing of the material to be included in the corpus, we performed tagging in several stages: 1. morphological analysis; 2. syntactic analysis; 3, semantic analysis; 4. Teleological analysis. Below are specific examples, their analysis and conclusions.

In our research, we prefer forms of visualization such as graphs and tables. They are widely used to describe the structured data needed to build a body, as well as the word/tag cloud, which is a fairly common form of text data visualization today.

The technical format of the corpus:

The body is encrypted in XML format, in accordance with the Basic Text Encryption Recommendations (TEI P4). The corpus consists of two parts: texts and definitions.

The texts are grouped according to languages (English and Georgian). Texts in one language form a single TEI body consisting of the text itself and the TEI title (called a header). The latter provides comprehensive information about the corpus. Each text in turn contains a TEI title and its own text consisting of a separate paragraph.

The TEI text is accompanied by a double definition for each language pair, although the text is not fully presented but contains extracts from a separate paragraph.

In the research we used the experience shared by our foreign colleague, namely: In October 2019, Professor Ekaterina Lapshinova-Koltunski, the University of Saarland (Germany), visited the Translation and Interdisciplinary Research Center at the Faculty of Humanities of Batumi Shota Rustaveli State University. Prof. Lapshinova-Koltunsky presented the corpus explanations by various authors in the form of a presentation. We compared monolingual and multilingual corpora. In bilingual and multilingual corpora, the texts are sorted by register and genre.

With the help of Professor Lapshinova, using a fragment of the Georgian text of the Association Agreement and several tags, we created a sample for annotation in the MMAX2 program, called "project". When we open this folder, another folder called MMAX-terminology appears on the screen, which has the same structure as all other MMAX projects (such as Basedata, Markabls, etc.)

Open the program MMAX2 and find the above folder (MMAX-terminology) using the search field File> Load in the MMAX2 window, the file AA-geo.mmax will appear, when opening it we will see the Georgian text of the Association Agreement in XML format and the annotation program will start. It should be noted that the MMAX2 program covers almost all languages of EU member or candidate countries, except Georgian. This was the first successful attempt to place a Georgian text in the program.

4.2.**Annotated, structured and tagged parallel legal texts corpus research in automatic and semi-automatic modes**

At the first stage of the research, we tagged colloquialisms, calques, foreign words, terms (legal and economic), syntactic calque, misinterprettation of the passive voice construction in the English text of the Georgia-EU Association Agreement (Chapters 1-4, Article 1-276) and the Georgian translation with tautology, spelling, translation error, abbreviation, lexical addition. See example:

The Parties <lng_clc> commit themselves </lng_clc> to the <lg_trm> rule of law </lg_trm>, <lng_clc> good governance </lng_clc>, the <lng_clc> fight against corruption </lng_clc>, the fight against the various forms of transnational <lg_trm> organized crime and terrorism </lg_trm>, the promotion of sustainable development, effective multila-teralism and the fight against the proliferation of <lg_trm> weapons of mass destruction </lg_trm> and their <lng_trm> delivery systems < / lng_trm>. This commitment constitutes a <lng_clc> key factor </lng_clc> in the development of relations and cooperation between the Parties and contributes to regional peace and stability.

მხარეები <lng_clc>ვალდებულებას იღებენ</lng_clc> უზრუნველყონ <lg_trm>კანონის უზენაესობა</lg_trm>, განახორციელონ <lng_calq>კარგი

მმართველობა</lng_calq>, <lng_clc>ებრძოლონ კორუფციას</lng_clc> და სხვადასხვა ფორმის <soc_fwd> ტრანსნაციონალურ</soc_fwd> <lg_trm>ორგანიზებულ დანაშაულსა და ტერორიზმს</lg_trm>, ხელი შეუწყონ <lng_calq> მდგრად განვითარებას</lng_calq>, ეფექტიან <soc_fwd>მულტილატერალიზმს</soc_fwd და <synt_calq>იბრძოლონ <lg_trm>მასობრივი განადგურების იარაღის</lg_trm> და მისი <lng_trm>მი-წოდების საშუალებების</lng_trm> გავრცელების წინააღმდეგ</synt_calq>. ეს ვალდებულებები წარმოადგენს მხარეებს შორის ურთიერთობებისა და თანამშრომლობის განვითარების <lng_calq><lng_clc>ძირითად ფაქტორს</lng_clc></lng_calq> და ხელს უწყობს მშვიდობასა და სტაბილურობას რეგიონში.

According to the goals of our dissertation we have used the following tags to sign different linguistic constructions and terms:

<lng_clc>…</lng_clc> colloquialisms
<lng_calq>…</lng_calq> calque
<soc _fwd>…</soc_fwd> foreign words
<lng_trm>…</lng_trm> terms
<lg_trm>…</lg_trm> legal terms
<ecn_trm>…</ecn_trm> economic terms
<synt_calq>…</synt_calq> syntactic calque

<lng_mpv>…</lng_mpv> misinterpretation of the passive voice construction

<st_taut>…<st_taut> tautology

<lng_spl>…</lng_spl> spelling

<tr_err>…<tr_err> translation error

<lng_abr>…</lng_abr> abbreviation

<tr_lex.add>…</tr_lex.add> lexical addition

The tags were created by Khatuna Beridze, Associate Professor of the Department of European Studies, Batumi Shota Rustaveli State University. The tagging process requires special care and attention. We carried out it in several stages: 1. Processing and comparing the English and Georgian texts of the Association Agreement in parallel, which was mainly of an introductory nature; 2. Marking of source and target texts in parallel mode, using the same tag - the first stage of marking; 3. Re-review of already marked texts - marking. In most cases we had to repeat this steps several times because even the slightest inaccuracy in the marking process was a problem - the software analyzer could not perceive the data we provided. After a long and laborious marking process, we converted English and Georgian texts from Word format to XML format. To do this, we used a special program NOTEPAD ++, the advantage of which is that it is perceptible to the software analyzer, in addition, it makes it easier to work if any tag is entered incorrectly, for example, missing a fraction or more than the normal spacing between characters. NOTEPAD ++ finds the error itself, reddens it, and we no longer have to re-read the entire paragraph or chapter to correct the error. See Figure 1.2:

Figure 1. English tagged text in NOTEPAD ++:

Figure 2. Georgian tagged text in NOTEPAD ++

At the next stage, we used a special program created in cooperation with the Translation and Interdisciplinary Research Center and the Technological Department of Batumi Shota Rustaveli State University to check the marked texts. See Figure 3:

Figure 3. Picture of corpus analyzer:

To use this program, you first need to register as a user. We enter the program referring the username and password of the user. See Figure 4:

Figure 4. Program picture:

First of all, we upload English and Georgian texts converted to XML format in the program. Then in the special field - add tags, we enter the tags we have created in advance. Select the desired text already entered in XML format, select the theme in the field, select the desired tags and click the send button. After all, the program automatically searches and provides statistics that we can use to contrast the bilingual material processed in our research. The program sorts English and Georgian texts in parallel mode and, most importantly, automatically provides statistics for tagged terms for quantitative research methodology. See a screenshot showing the data obtained from the statistical analyzer of the bilingual text sorted in parallel mode by the program:

Figure 5. Structured texts

We obtained the following results by placing the tagged texts in a special search program (see sample):

Figure 6. Search results

As for the analyzer results:

We have identified and analyzed a number of cases of translation inaccuracies and errors, namely: calque - 47 examples; Syntactic calque - 18 cases; Misuse of foreign words and borrowings - 42 examples; Terms - 230, which can be used to compile a terminological dictionary of Georgia-European Union associate agreement.

Conclusion

The aim of this research was to set a precedent for modeling the analysis of the translation of legal texts in a parallel corpus, as corpus research is particularly relevant in modern translation studies. We opted for the Association Agreement, the contradictory analysis of the original and target texts of which led us to the following conclusions:

- We have created an individual model of translation analysis of legal texts for the parallel corpus of English and Georgian texts;

- The translation corpus created within our research allows the researcher to solve a specific research issue in a short time, is able to select and use corpus texts for different types of contrasting research;

- This parallel corpus is a valuable linguistic resource that provides the translation process with a rich vocabulary and is a kind of bank of ready-made phrases for the translator, simplifies and accelerates its work;

- Searching for terms, opening contexts, determining contextual meanings, contrasting analysis of definitions is

possible in a short period of time through the specialized search program used in the research;

- Following the objectives of the research, we automatically examined several categories of translation inaccuracies through a parallel corpus and obtained their statistics in a short time;

- If in the early period linguistic research was mainly limited to theory and then introduced into practice (top-down process), our corpus-linguistic work relies directly on practical research (bottom-up process). Moreover, the theoretically approved material is still intended for practical use;

- Taking into account the translation inaccuracies and errors detected by the software analyzer within our research will help: B) bring the Georgian translation of the Association Agreement closer to the basic principles set out in international standards and translation quality requirements, as the existing language version of all EU legal acts must be equally authentic; C) to bring the Georgian version of the agreement closer to the general legislative principles, such as equality before the law and credibility of the law; D) avoid potential risks of litigation, including financial or political damage; E) reduce difficulties and problems in interpreting the law or introducing the law at the national level;

- Among the practical values of the legal parallel corpus created within the framework of our research, the following should be noted: it can be used for a) statistical machine translation, b) bilingual lexical or semantic resources; C) for

dictionaries or anthologies; D) for testing and teaching multilingual information retrieval software;

- What is most important, the translation of the terms identified by the software analyzer within the framework of this research can be used to compile a terminological dictionary of the Georgia-European Union Association Agreement.

- We consider the fact that this is an attempt to become part of the JRC-Acquis, the largest multilingual corpus available today, which includes more than twenty languages of the EU member states, to be the greatest practical value of the English-Georgian parallel corpus of the Association Agreement;

- We especially mention the fact that the Legal Parallel Corpus we have created is another step forward, a prerequisite for joining the JRC-Acquis: we have already converted the English and Georgian transliterated texts of the Association Agreement into XML format, while the EU candidate languages are now available only in Word format, which JRC-Acquis program can not accept;

- This study will make it possible to reduce the number of translation errors or inaccuracies made by the software analyzer, which is sometimes very important, as a legal document of interstate importance similar to the Association Agreement may contain particularly important information, such as nuclear proliferation. In such a case, the risk of a fatal outcome caused by a translation error can be reduced;

- Another practical value of our research is the function of the software analyzer to sort bilingual texts in parallel and,

most importantly, to give us statistics of tagged terms automatically;

- It is especially noteworthy that the results of the research can be verified remotely through this program. We can boldly state that the joint and greatest result of corpus linguistics and informatics has been achieved - we can get the result of the research in the shortest time;

- We believe that the Georgian-English parallel corpus created within the framework of our dissertation is one of the steps forward and an attempt to establish a Georgian language corpus among the multilingual corpora in the world and make it available to interested specialists.

- By obtaining statistical indicators obtained by the analyzer, a parallel terminological dictionary and parallel structured texts, we have achieved the goal of our research in terms of software and created manual marking of legal text translation, processing in a software analyzer, and further English scientific research and modeling.

- As part of our research, we have determined that the role and function of a software analyzer is extremely important for qualitative translation research, as automatically obtained quantitative indicators allow for systematic conduct of research and solution of a specific research issue in a short time.

Bibliography:

1. Beridze Khatuna, Translatology, Batumi, 2018

2. Beridze Khatuna, Gvarishvili Zeinab, Corpus Research of Acquis Legal Terms, BSU publishing, 2014

3. Tandashvili Manana, Purtskhvanidze Zaqaria, Corpuslinguistic Paradigm in Linguistics, Frankfurt linguistic circle, 2014

4. Khalvashi Ramaz, Digital Humanities, BSU publishing, Batumi, 2018

5. Altenberg Bengt, The English-Swedish Parallel Corpus, Lund, Sweden, 2001

6. Aristotle, Wicomachean Ethics, book II

7. Austin N. How to do Things with Words; Oxford University Press, London, 1962

8. Berry M. David, Understanding Digital Humanities, New York, Palgrave Macmillan, 2012

9. Coward A. Ellis J. Language and Materialism, London 1976

10. Cruse A. Meaning in Language; An Introduction to Semantics and Pragmatics, II ed. Oxford University Press, Oxford, 2004

11. DGT Translation Quality Guidelines, EU, Brussels,2015

12. EC, DGT, European commission and Directorate General for Translation, Brussels, Belgium, Education Publishing House, 2002

13. Eisselle Andreas, Yu Chen, A Multilingual Corpus from United Nation Documents, Valetta, Malta, 2010 2010

14. Erjavec Tomaz, The English-Slovene ACQUIS corpus, Slovenia, 2017

15. Evans Dubley, Developments in English for Specific Purposes, Cambridge, 1998

16. Gazdar Gerald, Speech Act Assignment in Elements of Discourse Understanding, Cambridge University Press, Cambridge, 1981

17. Glasersfeld E.V. Teleology and the Concepts of Causation; 2009

18. Grice M.P. Logic and Conversation, Cambridge, 1967

19. Ide Nancy, Veronis Jean, Multilingual Text Tools and Corpora, Kyoto, Japan, 1994

20. Johnson M.R. Aristotle on Teleology, Oxford, 2005

21. Kaalep H.Jaan, Comparing Parallel Corpora and Evaluating their Quality, Tartu Univ. Estonia, 2009

22. Katz J.J. Propositional Structure and Illocutionary Force; New York,1977

23. Kijaczko St. Kant`s Pragmatic Teleology; Opole University, 2004

24. Koehn Philipp, A Parallel Corpus for Statistical Machine Translation, Phuket, Thailand, 2005

25. Lapshinova-Koltunski, E.“Variation in translation: Evidence from corpora.” In New directions in

corpus-based translation studies (Translation and Multilingual Natural Language Processing 1), ed. by Fantinuoli, C. and Zanettin, F., 71-93. Berlin: Language Science Press. 2015.

26. Leech G. Exploration in Semantics and Pragmatics; Amsterdam, John Benjamins,1980

27. Lenaerts K.,Guttierrez-Fons J.A. Methods of Interpretation and the European Court of Justice; European University Institute, 2013

28. Levinson S.C. Pragmatics; Cambridge University Press, Cambridge, 1983

29. Lewis D. Scorekeeping in a Language Game; Journal of Philosophical Logic, issue 8, pp.339-359,1979

30. Margalitadze, T. Once Again Why Lexicography Is Science. Lexikos 28 (AFRILEX-reeks/series 28: 2018): 245-261 http://lexikos. journals.ac.za; https://doi.org/10.5788/28-1-1464

31. Mayr E. The Idea of Teleology; University of Pennsylvania, 2014

32. Miller Vaughne, The EU`s Acquis Communautaire, 2011

33. Miller Vaughne,The EU`s Acquis Communautaire, International Affairs and Defence Section,2011

34. Partington Alan, Aims, Tools and Practices of Corpus Linguistics, University of Bologna, 2006

35. Sbisa M. Speech Acts in Context Language and Communication; vol.22, 2002,pp.421- 436

36. Scherer Carmen, Corpuslinguistics, Un. Winter, Heidelberg, 2006

37. Schreibman Suzan, Unsworth John, Siemens Ray, A New Companion to Digital Humanities, Oxford, Blackwell, 2016

38. Sinclair John, Corpus, Concordance, Collocation, Oxford University Press, 1991

39. Sinclair John, The Automatic Analysis of Corpora, Directions in Corpus Linguistic p.379- 397, Berlin, 1992

40. Sinclair John, Trust the Text.Language,, Corpus and Discourse, London, Routledge, 2004

41. Sperber D. Wilson D. Relevance, II ed. Blackwell, Oxford, 1999

42. Stalnaker R. Context and Content,Oxford University Press, Oxford,1999

43. Steinberger Ralf , An Overview of the EU`s highly multilingual parallel corpora ,Ispra, Italy, 2014

44. Tahira Naseem, Snyder Benjamin, Multilingual Part-of-Speech Tagging: Two Unsupervised Approaches, 2009

45. Tognini-Bonelli Elena, Corpus Linguistics at Work, John Benjamins Publishing, 2001

46. Trung D.K. Beyond Ab. Maslow`s Theory of Human Motivation; Philosophy on the net, 2016

47. Van Der Esch B. The Principles of Interpretation Applied by the Court of Justice of European Communities; Fordham International Law Journal, 1991

48. Vukčević Miodrag, EU directives and standards on translating and interpreting and their implementation in Serbia as a non-EU country and a country in the Stabilization and Association Process (SAP)

49. Yang Huizhong, An Introduction to Corpus Linguistics, Shanghai, Foreign Language Education Publishing House, 2002

50. Yu Chen, Kay Martin, Interesting Data for faster and Better Statistical Translations, Boulder, Colorado, 2009

51. Yule G. Introduction to Language Study; Oxford University Press, 2000

52. Steinberger Ralf, The JRC-Acquis: A multilingual parallel corpus with 20+ languages,Italy, 2011

53. Maud Ehrmann, Turchi Marco, Building a Multilingual Named Entity-Annotated Corpus, Hissar, Bulgaria, 2011

54. Wang Dapeng, The problems and prospect analysis of corpus-based translation studies, Bohai Univ. China, 2013

55. Yun-Chuang Chiao, Kraif Olivier, Evaluation of Multilingual Text Alignment Systems, Genoa, Italy, 2006

(http://europa.eu.int/)

http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexplus!prod!CELEXnumdoc&numdoc=CELEXCODE&lg=LG.

http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:CELEXCODE:LG:HTML.

http://ccvista.taiex.be/Fulkrum/CCVista/$LG/$CELEXCODE-$LG.doc.

http://nl.ijs.si/telri/Vanilla/

http://www.meta-net.eu/whitepapers/key-results-and-cross-language-comparison.),

http://www.meta-net.eu/whitepapers/key-results-and-cross-language-comparison.

http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2003:0001:0031:EN:PDF.

http://www.euromatrix.net

http://www.euromatrixplus.net/)

(http://europa.eu.int/)

http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexplus!prod!CELEXnumdoc&numdoc=CELEXCODE&lg=LG.

http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:CELEXCODE:LG:HTML.

http://ccvista.taiex.be/Fulkrum/CCVista/$LG/$CELEXCODE-$LG.doc.

http://nl.ijs.si/telri/Vanilla/

http://www.meta-net.eu/whitepapers/key-results-and-cross-language-comparison.),

http://www.meta-net.eu/whitepapers/key-results-and-cross-language-comparison.

http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2003:0001:0031:EN:PDF.

http://www.euromatrix.net

http://www.euromatrixplus.net/)

Https://dictionary.cambridge.org

http://www.lawteacher.net/free-law-essays/constitutional-law