

Translation Challenges in a Multilingual Society: Legal Translation in Belgium

Dr Hendrik J. Kockaert

Dr Gabriele Sauberer

Prof. Dr Frieda Steurs

Lessius/KULeuven University
Department of Applied Language Studies
Sint-Andriesstraat 2
B-2000 Antwerp

TermNet
Mooslackengasse 17
1190 Vienna
Austria

Abstract

Purpose

This paper proposes a system architecture for a context¹-sensitive terminology management system (TMS) in the domain of legal LSP. The TMS is expected to meet the needs of the translation services in the Belgian Federal Public Service (FPS) of Justice in that our TMS will allow to search, (re)use and edit each term in its customized archetypical phraseological context. In a second part, we propose to link this TMS to EN 15038, the EU standard on translation services, which has been developed to streamline translation projects aiming to deliver more efficient multilingual LSP documentation.

Method

In a first step, automatic knowledge acquisition algorithms will be developed to mine context-sensitive terminological information from on-line legal LSP corpora. In a second step, a feed-back loop between interface developers and expert users will be set up to optimize user access to the contextualized terminological knowledge base during the actual translation process. To meet these requirements, there will be an intense co-operation between terminologists, the developers of the knowledge extraction algorithms, database and software engineers, and expert users.

Results

The TMS is expected to display an interface offering the look-and-feel of a standard desktop CAT-tool. The actual application will run remotely on a server in a cloud-based architecture which will be available via an on-line interface. As a collaborative working environment, the tool will allow extensive and real-time interactive user management.

Conclusion

This paper analyses the challenges legal translators are facing in Belgium. At present, translation-oriented authoring and translation proper of legal documentation are being carried out on an ad-hoc basis. An R & D project has been set up to design a TMS which is expected to deliver correct and consistent legal terminology and collocations in both source and target languages.

Key terms: LSP translation, context-sensitive terminology management, translation environment tools

1. Legal Translation in the Belgian Context

Before discussing the actual challenges legal translators are facing in the Belgian FPS of Justice, we first need to draft some basic background information. Translating legal documents for the federal government in Belgium involves mainly an exclusively linguistic enterprise, because the federal legal system is identical in the language communities (Herbots, 1987, p. 813). Constitutionally, Belgium recognises Dutch, French and German as its three official languages.² According to the law on the use of official languages in legislative documents,³ all of them (laws, regulations, Acts, Ministerial Orders, and Royal Decrees) were to be published in the official Bulletin of Acts, Orders and Decrees (*Belgisch Staatsblad*) in French and Dutch. To accommodate German as well, the Belgian Constitution was modified on 21 April 2007,⁴ in order to assure that all legislative documentation is to be translated into the three constitutionally determined official languages.

At the national level, the government architecture in Belgium is known to be a federal government which is composed of subnational or *federated* entities such as communities, regions, language areas, municipalities, and provinces (Crikemans, 2006).⁵ The most politically significant federated entities are the regions (Brussels-Capital Region [Brussels], the Flemish Region [Flanders], and the Walloon Region [Wallonia]), which accommodate four linguistic areas respectively: the bilingual French-Dutch area, the Dutch language area, the French language area and the German language area (which has specific language facilities for francophone inhabitants). These subnational entities appropriated more and more exclusive competences through successive constitutional State reforms from 1970 until 2001 (Erk, 2003, p. 2008). This process of steadily shifting competences to the regional levels parallels a growing publication volume of their own legislation in their respective regional languages (Ordinances of the Brussels-Capital Region, Flemish Decrees, Decrees of the Walloon Region). In this context, translation has shifted from an exclusively linguistic exercise to expert-based legal comparison, i.e. comparing legal concepts, which do not always coincide with the concepts of the federal State's legal framework. The purely linguistic exercise of legal translation at the level of the federal government has consequently developed into a combined linguistic and legal comparison exercise when it comes to translate the legislative documents of the different linguistic areas, each of them governed by different governmental entities.

At a supranational level, Belgium (including all its subnational entities) is at the same time an EU member State, which encompasses adopting an increasing number of EU directives with EU-specific legislation, hence EU-specific terminology. Accordingly, translating EU legal documentation into Dutch, French and German involves a legal analysis of the EU directives before *transposing* them into the different national law systems of the EU member States, hence the Belgian legislation. This transposition exercise generates a risk of infecting the national law LSP with imported terminology proper to the EU legislation, which is delivered by Dutch translators, who originate both from Belgium and The Netherlands.

Another issue arises when it comes to analyse source and target arrests of the Constitutional Court: parallel to the United Nations Convention on the Law of Treaties⁶, these arrests are authenticated in the three official languages, which implies that source and target texts cannot be differentiated, and are equally authoritative.

Moreover, the target version (translated) is to be considered an authentic text having the same legal status as the source document.

1.1 Current Practices in the Translation Department of the FPS of Justice

We refer briefly to a survey which was set up by Kockaert, Vanallemeersch & Steurs (2007) to gauge the current translation practices in the translation department of the FPS of Justice. It appeared that individual term bases and translation memories (TMs) are developed ad hoc by most translators. Subsequently, a well-planned synergy between these term bases has been sporadic, which hampers wide-ranging harmonized and consistent terminology work. Yet, legal texts are known to be linguistically characterized by recurrent and standard expressions, and legal affairs are known to rely necessarily on a clear-cut and coherent phraseology (Hendrickx, 2002, p. 115). This dichotomous situation easily leads to inadequate terminology and translation, and needs to be remedied by an efficient terminology management which focuses on terminological consistency, relinquishing at the same time linguistic and legal ambiguity. A crucial starting point for leading such an enterprise to a sustainable result is the retrieval of *legal terms*, extracted and aligned from multilingual legal documentation. Fortunately, the Belgian e-government publishes the legislative documents (Council of State, Conciliation and Settlements Court, ...) free of charge on its websites.

Based on the current practices in the FPS of Justice, the need to remedy the current situation seems legitimate and this paper will try to put forward some pragmatic solutions.

2. Multilingual LSP Resources: Terms, Units of Meaning and Translation Units

When we refer to the currently used instruments of multilingual legal terminology, a first given seems to be that translators and terminologists are ploughing the sands by consulting print and on-line legal dictionaries, when they start their quest for the exact term. Legal translation is a domain in which dictionaries are of limited value. The reason for this is that most dictionaries look for the convenient "one-to-one equivalent" translation, and very seldom if ever provide an explained translation (Cairns & McKeon, 1995, p. 197). Especially in legal LSP, it is important to *explain* the exact meaning of a term, which Teubert & Cermáková (2007, p. 17) define actually as a *unit of meaning*, which is a broader unit of text containing key words, and which we fully understand only when it is embedded in a unique context. Consequently, LSP term bases should not only display single terms, but also this kind of units together with their textual environment. On the same lines, Cairns & McKeon (1995) support our idea that a suitable legal TMS needs to offer explained translation, or in other words, translation of customary and standard expressions in their textual or collocational environments. In the practice of legal translation, offering translated terms in their specific collocational context, which often seems to be customized by inheritance and tradition, seems essential (Dop, 2003, p. 4; Kjær, 1990, p. 26). It is precisely this kind of information which is needed in legal translation. Consequently, a legal TMS will be considered a success only if each particular term can be found in its customized, standard or traditionally used environment.

When we extend the relevance of broader meaning units to the area of translating LSP, we follow Teubert & Cermáková (2007, p. 115) and focus on broader *translation units*, which are ideally translated into a unique target equivalent. Referring to our specific translation project, we need to search and define key terminological units, embedded in their collocation profiles, which are part of both the broader source and target meaning units respectively.

2.1 Example

The example below has been selected as a relevant sample from the corpus which was set up by Corten (2010) from Arrests of the Constitutional Court between January 2006 and December 2009, which are all available on line. Thanks to this sample, we were able to illustrate the current translation practices described above, and to commence an adequate remedial road map. The following scenario has been observed as one of the most frequent ones: context-independent terminological inconsistency (Figure 1).

The terminological unit *un litige pendant* has been registered in 76 arrests (Corten, 2010). In 38 of them, the unit has been translated as *hangend geschil*. Alternative Dutch variants were the following: *hangend rechtsgeding* (4 arrests), *geschil dat aanhangig* (3 arrests), *hangend geding* (3 arrests), *hangend rechtsgeschil* etc. (as can be seen below in Figure 1). When we focus on the contexts of *un litige pendant*, we observe that the unit at discussion appears frequently in the following collocational context: *le litige pendant devant le juge a quo*.

The Dutch equivalents for this unit are the following: *het voor de verwijzende rechter hangende geschil*, *het geschil voor de verwijzende rechter*, *het geschil dat hangende is voor de verwijzende rechter/dat voor de verwijzende rechter hangende is*, *het geschil dat voor de verwijzende rechter aanhangig is*, *het voor de verwijzende rechter aanhangige geschil*, *het geschil dat voor de verwijzende rechter aanhangig is gemaakt* en *het voor de verwijzende rechter hangende geding*. This means that we pinpointed at least 7 differences in the parallel arrests, which are translated in Dutch. More, the observed equivalents are context-independent, which leads us to conclude that these differences need to be harmonised in order to comply with the requirements of terminological and translational consistency.

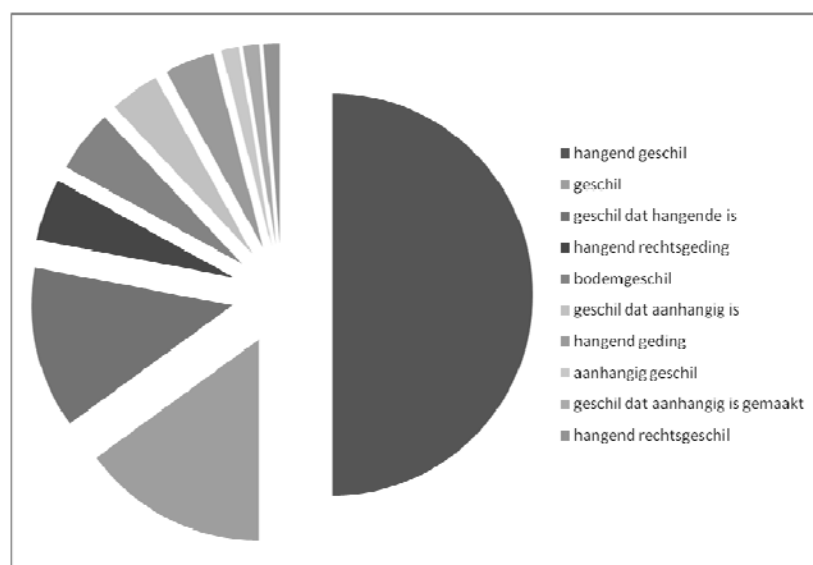


Figure 1: Dutch equivalents for *un litige pendant*

3. Context-conditioned Legal Language

Having put forward that legal language is characterized by context-dependent units of meaning and units of translation, we need to discuss further the nature of the contexts as such. Although a comprehensive discussion on this theme is not within the scope of this paper, we put forward some clarification in the following lines, which will allow us to have a better idea.

Let us first survey some major views on *phraseology*, which can be paralleled with the notion of context we are referring to in this paper. Besides, it emerges that these views do not always correspond. Burger, Buhofer & Sialm (1982) qualify phraseologies as being semantically irregular, and syntactically fixed and irregular, as living independently of non-linguistic context, e.g. legal or administrative documentation. Term phraseology, as it is called by Picht (1989) is characterized by free word combinations based on semantic markers and distinguishers. Kjær (1990) prefers naming recurrent legal phrases as context-conditioned word combinations, which are syntactically regular, and which are composed of segments that have inherent meanings. For the purpose of this study, we prefer following Kjær (1990, p. 26) by adopting her point of view that recurrent legal expressions are “typically reproduced in certain recurrent contexts” and are “characterized by a context-dependency [...] in legal language”. With context dependency, Kjær refers to the context as well as the situation⁷, combining the institutionalized and inherited use of archetypical legal language and non-linguistic factors. Among the latter, Kjær (1990, p. 26) refers firstly to laws that explicitly prescribe the use of such or such term in legal documentation, e.g. the German Civil Procedure Act. In this context, translators are not allowed to simplify idiosyncratic formulae with a view to making their texts more accessible to the public. Secondly, not respecting the usual procedural steps by modifying the routine word strings might weaken the legal effect, or the intention of the document might be misunderstood. Reproducing the often sibylline legal expressions, hence reproducing linguistic fossilization, fosters, surprisingly, disambiguity and continuity of Law.

Adopting Kjær allows us at the same time to adopt the ISO 1087-1 standard and to apply its definition of a *terminological phrase*: “Phrase in which at least two elements (terms) are syntactically linked and thus form an utterance with subject-field-specific content; the internal coherence of the elements is based on their conceptual combinability”.

To turn back to our TMS project, we claim that one crucially needed key feature is accommodating a means to fine-tune extraction modalities for legal documents so that terms can be extracted together with their contextual and situational environment from on-line or electronically saved documentation. One of the nuts-and-bolts of a legal TMS should be the replacement of typical term extraction techniques by context-based extraction of a given terminological unit and its translation.

4. System Architecture Requirements for a Context-sensitive TMS

Generally, LSP translation is known to be heavily dependent on terminology, and legal translation, as has been discussed above, is more in particular dependent on context-based terminological units. With a view to deliver these cornerstones of legal translation, it seemed justified to propose a TMS architecture which is

hoped to satisfy this need. Comparing database architectures for a legal TMS with any other TMS seems a basic first step in our process to achieve a manageable purpose-built TMS for legal documentation.

Generally, legal databases should in the first place comply with the regular TMS requirements, including the following key features: indexing⁸, term extraction, hierarchical data structure of the fields, unlimited number of languages, flexible search functionalities, data import and export, concept modelling facilities, bibliography management, and data categorization. In order to maximize interoperability of our TMS, we follow the ISO 12620 standard. The descriptive fields which are deemed necessary for our TMS in the light of the above-depicted peculiarities of legal LSP, will include the following: an expert source field (including an oral source referring to legal experts' advices), a technical definition field, a possibility of drawing relations between concepts, definition source, context, context source, legal area (subnational, national or supranational), type of legal documents (Arrests, Ministerial Orders, Royal Decrees, Directives, Ordinances, etc.). These fields are ISO 12620 compatible, or are at least compatible with the core structure of the standard.

On the basis of ISO 12620, which specifies mechanisms for creating, selecting and maintaining data categories, as well as an interchange format for representing them, the TC37 Technical Committee *Terminology and other language and content resources* adopted in 2008 the TBX framework which has been published as the ISO 30042 standard. This standard is designed to support different hierarchically ordered architectures of terminological data, such as the above mentioned descriptive fields, which are also listed and hierarchically ordered in ISO 12620, and to facilitate data interoperability.

TBX is designed to generate differently formatted terminological data, or data categories, which reflect the descriptive fields in different term base structures. To maximize interoperability of the actual terminological data, TBX also provides a default set of data categories based on ISO 12620, which are commonly used in terminological databases. However, changes to the default set of data categories can be applied within the TBX framework to support specific user requirements. What a legal TMS should add to the TBX core structure, is a means to fine-tune the extraction tool for legal documents so that context-dependent terminological units and phraseologies can be extracted and listed. One of the underlying nuts-and-bolts features of a legal TMS, should be a context-extraction and aligning tool, replacing a mere term extraction device. Descriptive fields should include: term status (preferred, recommended, standardized, deprecated, rejected), advice from legal experts, e.g. on accepted polysemy dependent on different legal situations (subnational, national, supranational), and different terms dependent on different LSP contexts.

6. Term-Wise Knowledge Platform

A State-funded R & D project, called *TermWise Knowledge Platform*, was set up by Lessius and the University of Louvain. This project is expected to lead to the design of a TMS suitable to cater for the above expressed needs and will operate in the following way. The TermWise Knowledge Platform will basically follow a two-step approach for implementing the above system architecture requirements and bringing the proposed contextualized TMS to the user. In a first step, automatic knowledge acquisition algorithms will be deployed, to mine context-sensitive terminological information from the existing (large) corpus of

translations in the Belgian legal domain. In a second step, a feed-back loop between interface developers and expert users will be set up to optimize user access to the contextualized terminological knowledge base during the actual translation process. The user interface is expected to be integrated in the familiar and mostly available CAT tools. Additional features may include: automatic analysis of the user's source text for known translations and term usage in context. This added functionality should run either in the background or be intuitively accessible, so that users can gradually and easily learn to integrate this functionality in their translation process.

There will be an intense co-operation between terminologists, the developers of the knowledge extraction algorithms, database and software engineers, and expert users. The interface will have *the look and feel* of a standard desktop CAT-tool, but the actual application will run remotely on a server in a cloud-based architecture that will be available via an on-line interface. As a collaborative working environment, the tool will allow for extensive user management that keeps track of which user made which contribution to the database. As in traditional CAT tools, the editing interface will display panes for both source text and translation with automatic suggestions of translations from the knowledge base. However, unlike in standard CAT tools, we envisage translation suggestions below the segment level, i.e. on a more refined level of terminological units and (potentially non-continuous) broader translation units referred to earlier. Translation suggestions will be displayed context-dependently, meaning that translations which are more probable on the basis of context are suggested first. The reliability degree of the translations will be indicated either by a colour code or by different underlining styles.

Terms and broader units in both the source text and the (suggested) translations will then be clickable to access the knowledge base in a pop-up window. Users will be able to browse through the different translations available in the knowledge base, the source of the suggested translations ranked by authoritativeness, and examples of idiomatic uses of a clicked term in context. Possible alternative terms based on semantic similarity (rather than previous translations) will also be available. The users will also be able to provide additional information to the database themselves. These can take the form of informal comments on a translation, but also editing formally structured fields in a term base (in an ISO 12620 compliant manner). The functionality of the CAT-tool will be iteratively improved and customized to user needs in a test run at the translation department of the Belgian FPS of Justice (See Heylen, Kockaert & Steurs, 2010).

7. EN 15038 Compliance

At this stage of the TermWise Knowledge Platform project, we limit ourselves to focus on the areas in which the translation department of the FPS of Justice will be compliant with the EN 15038 standard: human resources management, professional competences of the translators, revisers and reviewers, technical resources, quality management, which all pertain to the entire translation process.

7.1 Human Resources Management, Professional Competences of the Translators, Revisers and Reviewers

In the above titled area, the EN 15038 compliance is confirmed on the account of the stringent hiring procedures for all staff members in the translation department which are controlled by a selection procedure organised at the national level by governmental selection services (SELOR). Pre-hiring procedures include screening the necessary qualifications and experiences, depending on the actual job descriptions (revisers and reviewers).

7.2 Technical Resources, Quality Management, and the Translation Process

The FPS of Justice is equipped with up-to-date translation and terminology software, which is suitable for performing translation assignments in accordance with the pre-set EN 15038 quality criteria and a TermWise Knowledge Platform-based system architecture. The translation process compliance criterion is worth some additional attention given the scope of this paper. The TermWise Knowledge Platform is expected to cater for adequate pre-translation based on earlier completed and in-process translation projects, stocked in the TermWise Knowledge Platform TMS. After having completed the pre-translation, the staff will be able to translate the assignment with the support of the translation memory and the TMS, which will help the translators produce the actual translation in compliance with the following EN 15038 criteria: checking the translation by the translator, reviser and reviewer, checking the proofs, and releasing the assignment.

7.3 Quality Management

As mentioned earlier, we propose to link our project of legal translation to the quality criteria of the EN 15038 standard on translation services. We refer to the EN 15038 standard because it allows us to engage in a step-by-step translation process which runs from computer assisted pre-translation through translation proper, up to the inevitable revision and review stages before delivering a target text of high quality. Implementing the compliance criteria of the entire translation process leads indirectly to a standardized and sustainable quality management. In the context of legal LSP, we chose to refer to both the generally required quality management issues, and to the extra quality assurance regulations issued by the EU Court of Justice, which state that a translation, without being literal, needs to render the content of the source text, faithfully and fully, without any error, inconsistency or omission (Anckaert, 2006, p. 23). Furthermore, they stress on the need to maintain the typical jurisprudential style of arrests and conclusions. In general, we deduct from the quality assurance regulations that legal translation is to be based on consistency, correctness, customized or standardized terms and phrases, all of which is to be topped by the archetypical style of legal language.

8. Conclusion

In this paper, we presented a comprehensive and performing translation environment tool, which is geared towards an efficient translation process with a view to producing precise LSP target documentation, based on correct and consistent terminological units, embedded in their contextual and/or situational collocation profiles. Moreover, it is hoped that translation assignments in the translation department of the FPS of Justice will be produced along a real-time expert-based, interoperable and default translation process in compliance with the EU standard on translation services. Further research and experiments in other LSP domains will eventually benefit from the spin-offs of this TermWise Knowledge Platform.

References

- Anckaert, L. (2006). Kwaliteitsgaranties inzake vertalen bij het hof van justitie van de Europese Gemeenschappen. In Y. Vanden Bosch (Ed.), *Recht & Taal. Rechtszekerheid voor de anderstalige rechtzoekende* (pp. 17-26). Antwerpen/Oxford: Intersentia.
- Burger, H., Buhofer, A., & Sialm, A. (1982). *Handbuch der Phraseologie*. Berlin/New York: De Gruyter.
- Cairns, W., & McKeon, R. (1995). *Introduction to French Law*. Cavendish Publishing Ltd.
- Corten, L. (2010). Kwaliteitscontrole van juridische vertalingen: casestudy op basis van arresten van het Grondwettelijk Hof [Quality control of legal translations: case study on arrests of the Constitutional Court], Master's Thesis under the supervision of Hendrik J. Kockaert & Frieda Steurs, Lessius/KULeuven University.
- Criekemans, D. (2006, June). How subnational entities try to develop their own "paradiplomacy": the case of Flanders (1993-2005). Paper presented at the International Conference *Challenges for Foreign Ministries: Managing Diplomatic Networks and Optimising Value*, Geneva: DiploFoundation (Malta). Retrieved May 30, 2010, from <http://www.diplomacy.edu/Conferences/MFA/papers/criekemans.pdf>.
- Delisle, J., Lee-Jahnke, H., & Cormier, M-C. (Eds.). (1999). *Terminologie de la traduction. Translation Terminology. Terminología de la traducción. Terminologie der Uebersetzung*. Amsterdam/Philadelphia: John Benjamins.
- Dop, A. (2003). Common law versus civil law. Rechtsfamilies en de praktijk van het juridisch vertalen. *Linguaan*, 3, 2-8.
- EN 15038. (2006). *Translation services - Service requirements*. Technical Committee CEN/BT/TF 138, European Committee for Standardization.
- Erk, J. (2003). "Wat We Zelf Doen, Doen We Beter"; Belgian Substate Nationalisms, Congruence and Public Policy. *Journal of Public Policy*, 23 (2), 201-224.
- Hendrickx, K. (2002). *Taal- en formuleringsproblemen in de regelgeving: de taalopmerkingen in de adviezen van de Raad van State*. Unpublished PhD Thesis, KULeuven University.
- Herbots, J. H. (1987). La traduction juridique (Un point de vue belge). *Cahiers de Droit*, 28, 813-844.
- Heylen, K., Kockaert, H., & Steurs, F. (2010, August). The TermWise Knowledge Platform: an efficient translation and terminology management suite for legal translation in Belgium. Paper presented at the TKE 2010 Conference - Presenting terminology and knowledge engineering resources on-line: models and challenges, Fiontar, Dublin City University.
- ISO 12620. (2009). *Terminology and other language and content resources - Specification of data categories and management of a Data Category Registry for language resources*. Geneva: International Standards Organisation.
- ISO 30042. (2008). *Systems to manage terminology, knowledge and content - TermBase eXchange (TBX)*. Geneva: International Standards Organisation.

- Kjær, A. L. (1990). Context-conditioned word combinations in legal language. *Journal of the IITF*, 1 (1-2), 21-32.
- Kockaert, H. J., Vanallemeersch, T., & Steurs, F. (2007). Term-based Context Extraction in Legal Terminology: a case study in Belgium. In Á. Fóris, & J. Puszta (Eds.), *Current Trends in Terminology, International Conference on Terminology. Terminologia et Corpora 4*. Szombathely: Berzsenyi Dániel College.
- Picht, H. (1989). Fachsprachliche Phraseologie. In C. Laurén, & M. Nordman (Eds.), *Special Language: From Human Thinking to Thinking Machines* (pp. 89-109). Philadelphia: Clevedon.
- Teubert, W., & Cermáková, A. (2007). *Corpus Linguistics: A short introduction*. London: Continuum.

Notes

¹ The *context* refers to the “linguistic environment of a lexeme that contributes to the specifications of its relevant meaning in order to extract its sense” (Delisle, Lee-Jahnke & Cormier, 1999, p.129).

² Constitution – Coordinated text of 17 February 1994.

³ *Wet van 31 mei 1961 betreffende het gebruik der talen in wetgevingszaken, het opmaken, bekendmaken en inwerkingtreden van wetten en verordeningen.*

⁴ *Wijziging van de wet van 31 mei 1961 betreffende het gebruik der talen in wetgevingszaken, het opmaken, bekendmaken en inwerkingtreden van wetten en verordeningen.*

⁵ Constitution – Coordinated text of 17 February 1994, Art. 1.

⁶ Vienna 23 May 1969, Article 33.

⁷ The *situation* refers to the “sum of the extralinguistic elements surrounding the production of a text” (Delisle, Lee-Jahnke & Cormier, 1999, p.78).

⁸ The indexing feature refers to the possibility of creating term lists for each language in alphabetical order.