

TZOS: an Online Terminology Database Aimed at Working on Basque Academic Terminology Collaboratively

Izaskun Aldezabal, Jose Mari Arriola, Arantxa Otegi

HiTZ Basque Center for Language Technologies - Ixa, University of the Basque Country UPV/EHU
{izaskun.aldezabal, josemaria.arriola, arantza.otegi}@ehu.eus

Abstract

Terminology databases are highly useful for the dissemination of specialized knowledge. In this paper, we present TZOS, an online terminology database to work on Basque academic terminology collaboratively. We show how this resource integrates the Communicative Theory of Terminology, together with the methodological matters, how it is connected with real corpus GARATERM, which terminology issues arise when terms are collected, and future perspectives. The main objectives of this work are to develop basic tools to research academic registers and make the terminology collected by expert users available to the community. Even though TZOS has been designed for an educational context, its flexible structure makes possible to extend it also to the professional area. In this way, we have built IZIBI-TZOS which is a Civil Engineering oriented version of TZOS. These resources are already publicly available, and the ongoing work is towards the interlinking with other lexical resources by applying linking data principles.

Keywords: terminology databases, academic and professional terminology, Communicative Theory of Terminology, open linguistic data

1. Introduction

Terminology databases are important for many Natural Language Processing (NLP) applications (Maslias, 2017; Du et al., 2020), for instance, they are usually applied in the fields of machine translation, information retrieval, information extraction and text categorization. In addition to the role they play in the field of NLP applications, terminology databases provide users with effective and reliable access to terminology quickly (Cabr , 1993); therefore, the importance of such resources in the dissemination of specialized knowledge is undeniable. In the same vein, Faneca (2021) states that the practical purpose of terminology is to solve problems arising from the specialized communicative processes. Moreover, there is an urgent need to find a more efficient approach to terminology work as language workers acknowledge that terminology is a backbone of professional communication (Gornostay and Vasiljevs, 2014).

In this work, we present TZOS (Online System for Terminology Service)¹, which is an online terminology database to work on Basque academic terminology. TZOS is based on the Communicative Theory of Terminology (CTT) (Cabr , 1999; Cabr , 2001) since we agree on conceiving terms and words as activations of different semantic features of a lexical item that take place in different contexts of use, in contrast to the General Theory of Terminology (GTT) of W ster (1968). W ster established the principles for systematizing terminology work based on an onomasiological perspective, starting from the systematic ordering of concepts in specialized domains, and thereafter assigning a term in each language for each concept. This approach is based on the idea that the systematic organization of concepts into a specialized domain is universal and language-independent. After all, the goal is to ensure the univocity of terms in professional communication, crucial at the international level.

However, this perspective involves both a very restrictive subject of analysis and a way of working (Cabr , 2002), even more so in the case of languages that are in the process of standardization and development of

terminological elaboration, which is the case of Basque. As stated in Zabala (2019), natural development and self-regulation of terminology and phraseology happen inside specialized discourse communities, since such communities are the natural environment for the real usage of those linguistic elements. Hence, they have to be taken into account when describing the terminology. In the case of Basque, specialized discourse communities started growing, and terminology work and linguistic control of academic texts became more and more institutionalized since the creation of the Government of the Basque Country. Nevertheless, as disregarding real usages may hinder natural development, not only institutionalized language planning is needed. That is why when describing and providing specialized uses of words in our work, users/experts of the specialized area are considered. Currently, in the *Terminologia Sareak Ehunduz* (TSE, Weaving Terminology Networks) program (Zabala et al., 2018), which started in 2010 at the University of the Basque Country intending to describe the Basque terminology used at the university, teachers of the subjects of the degrees are considered experts and the unit of study is the subject they teach (e.g, the “Environmental Pollution Treatment” subject from the “Environmental Sciences Degree”). In our view, the subject gives the needed context to measure and decide about the termhood of a designation as well as the relevance for listing it in a terminological dictionary related to the subject. The subject also offers a semantic perspective to define the term or add details to the definition. Based on that idea, TZOS is thought mainly to store uses related to university education. However, we will show that it is flexible enough to extend it also to the professional area.

The design of TZOS presented in this paper is, therefore, a consequence of all these considerations, and it is integrated into a platform called GARATERM (Zabala et al., 2013) that provides the necessary resources for carrying out the terminological work according to the adopted theoretical and methodological approaches. Besides, we can benefit from using online platforms, 135 which offer the possibility of regular updates and more

¹<https://tzos.ehu.eus/>

effective collaboration via the Internet (Kitanovic et al., 2021).

Finally, the adoption of Semantic Web technologies and the Linked Data paradigm has been adopted by the need to ensure the construction of a resource that is at the same time interoperable, shareable, and reusable by the scientific community. In that way, we have represented the terminology of TZOS as Linguistic Linked Open Data (LLOD). We have designed a prototype for a Resource Description Framework (RDF) (Cyganiak et al., 2014) representation of TZOS terminology, modeled according to the *Ontolex-lemon* model.

The paper is organized as follows. After the introduction, the methodological foundations and GARATERM platform are described in Section 2. In Section 3 the terminology database TZOS is explained in depth. In Section 4 we show how we have extended TZOS for dealing with the Civil Engineering terminology. Section 5 is devoted to linking the tool to other resources utilizing the LLOD approach. And finally, in Section 6, we draw some conclusions and suggest avenues for future developments.

2. Methodological Foundations

The methodological foundations of the terminology obtained and compiled into TZOS are described in detail in Zabala (2019). Zabala describes it as an *active* description methodology because the authors/experts of the texts are involved in describing the terminology that they use, and they have the cognitive and pragmatic knowledge to organize terms and variants correctly.

In order to carry out such an active description, the GARATERM platform is used. GARATERM is a robust work environment composed of different modules that deal with all the processes involved in the active corpus-driven description of Basque academic terminology:

- GARATERM (Zabala et al., 2013): academic corpus which is composed of documents that the teachers of the University of the Basque Country (UPV/EHU) create for their teaching activity such as exams, exercises, practical works, programs, notes, and presentation slides. The natural environment of terms is made up of the texts in which they are used.
- Erauzterm (Gurrutxaga et al., 2005): automatic term extractor. Term candidates from the texts of each expert are extracted from the GARATERM corpus by Erauzterm and are offered as a list in the GARATERM platform so that the experts can validate them. Then, they provide the corresponding terms in other languages (mainly Spanish, English, and French). In this way, we obtain multilingual term glossaries of the subjects taught at the university. Since its creation, Erauzterm has been enriched with various available lexical and terminological resources that have been integrated to improve the term extractor and facilitate the term identification to the members of the GARATERM community.
- TZOS: online terminology database to store terms and bridge the gap between term extraction/elaboration with descriptive purposes

and real corpora by offering a framework for connecting the terms with the corpora examples of GARATERM.

The development of this platform, the interconnection between its modules, and the enrichment of the extractor with other lexical and terminological resources have been the result of a continuous refinement carried out within the TSE program. Previously, the TZOS term entries were not bound to the corpus, the automatic extractor was not so accurate and it was not integrated into the platform.

It is worth highlighting that both GARATERM and TZOS have been recently added to the Euskalbar toolbar (add-on available for Firefox and Chrome) which provides dictionary consults and offers a complete list of many referential resources to work in Basque, for example, Euskalterm², the Basque Public Term Bank managed by the Basque Government. Euskalterm stores the terminology that the Basque Government aims to promote and encourage. It is updated and maintained following the methodology and criteria approved by the Terminology Commission of the Basque Government and it offers the normalized status of the terms. Hence, when consulting one term of TZOS in Euskalbar, it is possible to see at the same time its information regarding the normalized status in Euskalterm, as long as the term is there.

3. TZOS

TZOS is an online database to work on terminology collaboratively. It is connected to the GARATERM corpus to bridge the gap between term extraction/elaboration with descriptive purposes and real corpora. Both the formal representation of the terms based on the TBX standard and the main functionalities of the online system have been described in Arregi et al. (2013).

3.1 Terms in TZOS

As stated in the previous section, TZOS is the result of a continuous process aimed at creating a multilingual dictionary specialized in academic terminology, with CTT as its basis. According to CTT, we conceive terms as elements from natural language found in specialized texts. A term in fact is a binary relationship between a linguistic form or designation (one or more words) and the concept (the meaning).

In TZOS, the terms are classified according to the Unesco nomenclature for fields of science and technology³; more precisely, all the terms are classified into 24 different fields. A given linguistic form can have more than one term-entry depending on the subject subfield it is used because we assume the meaning is different depending on the subfield. For instance, as shown in Figure 1, the word *agindu* is actually considered five different terms, and they are referred to as five terminological entries. On the one hand, *agindu* as a noun has at least four different meanings depending on the corresponding subfields: (1) *order* (Nutritional sciences); (2) *prescription*

²<https://www.euskadi.eus/web01-apeuster/en/ac36aEuskaltermWar/publiko/erakutsiBilaketa?locale=en>

³<https://skos.um.es/unesco6/00/html>

(Pharmacology); (3) *instruction, sentence, command* (Computer technology); and (4) *instruction* (Instrumentation technology). On the other hand, it is used as a verb in the Law philosophy subfield.

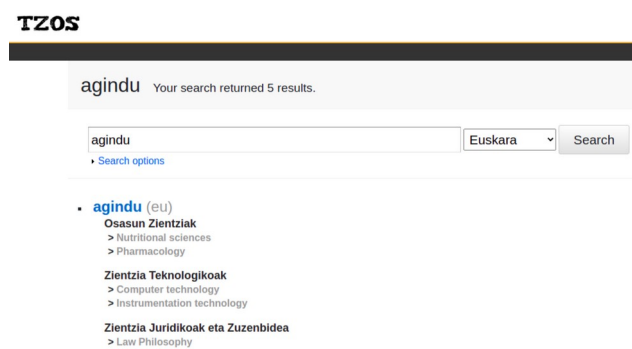


Figure 1: Search results for the Basque word *agindu* on the online TZOS interface.

Following with the previous example, *instrukzio* (*instruction*) and *sententzia* (*sentence*) are variants of *agindu* (*command*), which can be proposed by either the same author or another different one. When term lists corresponding to different authors are put together in the TZOS database, terminological hetero-variation is evidenced. Note that, as Zabala (2019) states, the description of the terminology obtained is less dependent on other languages than that obtained based on concept structures and term lists elaborated for other languages. In this way, our goal is to collect and make visible as many as possible denominative variants used in academic texts, to help the authenticated harmonization and stabilization of terms.

It could happen that when an expert proposes a new term, an equivalent term –equal spelling within the same subfield– already exists in TZOS. In that case, the system automatically shows a collision message and it is up to the expert to decide if it is or not a real collision. By analyzing the information already stored in TZOS, the expert will decide: (a) to include the proposed term as a new one –in that case, both terms will become homographs–; (b) to include it as a synonym –a denominative variant– of an already stored term; or (c) to discard the proposed term because the denomination has been already stored for the same concept.

When adding a new term to the database, in addition to the field/subfield and variants of terms, other features such as equivalent(s) in other language(s) (at least one), definition(s), semantic category, and hypernym(s) among others are also considered for the description of the term. On the other hand, in order to record the uses, in addition to the name of the expert, their origin is also specified, that is, the faculty or school and the degree where the subject is taught (and hence, the term is used).

TZOS currently contains 136,540 term-entries grouped into 24 fields and 308 subfields, all of them collected by 180 authors/experts.

Nowadays our interest is focused on describing variants of different types (orthographic, morphosyntactic, lexical, and discursive). More recently we have started adding

definitions as well as semantic tags in some terms within the TSE program.

3.2 Connection with GARATERM Corpus

TZOS is connected to the GARATERM corpus and, therefore, by clicking on a term in Basque you can see the contexts in which this term is found in GARATERM. This is called *in vivo* terminology, while in other reference corpus an *in vitro* approximation is presented. In other words, TZOS shows the terminology used by professors in their subjects and not as other terminology databases, the terminology that language normalizers want to promote, for instance, in the previously mentioned Basque Public termbank Euskalterm. Nevertheless, TZOS may be very useful for the groups and institutions involved in the normalization of terminology, since it allows them to look up real usage of terminology and to consider the data when making decisions about the terms.

As we have mentioned before, TZOS has been built up for descriptive purposes, consequently, descriptive terminology work requires significant text-corpora to be collected and systematized. GARATERM corpus nowadays collects spontaneous (uncorrected) and controlled (corrected) texts, allowing observation of the actual use made in university classrooms by lecturers and professors at the UPV/EHU. But it is also prepared for other texts (articles, specialized blogs...), as well as multimedia texts. These texts are organized by fields and subfields, as well as by textual genres. The data can therefore be obtained based on these parameters, which is essential for the description of the specialized records. GARATERM contains 21,762,846 words from the 2010-2021 period collected by 839 authors/experts.

As addressed in Section 2, this corpus is used not only to make linguistic searches and analysis, but also to extract real terminology. Terms are extracted by using the term extractor Erauzterm which is integrated into the same platform. With this infrastructure, we are gathering the real terminology used in the different subjects, that is the terms used in the natural context of the subjects and without any experimental interference. Once we have collected a significant amount of terms, we will be able to describe the terminology variation.

3.3 User Profiles and Term Development

As described in Arregi et al. (2013), five types of users are defined in TZOS: visitors, terminologists, supervisors, correctors, and administrators.

Terminologists, supervisors, and correctors are the responsible ones for the development of terms, and they work collaboratively as they form a kind of commission to monitor the terms. First, the terminologist proposes terms, then the supervisor ensures the suitability of the proposed terms, and, finally, the corrector verifies whether the proposed terms are correct regarding orthography and orthotypography. Figure 2 depicts this process, that is, it shows how we deal with a term marking its developing status (started, ongoing or consolidated) automatically. The value of the status is shown in the term-entry.

The term state cycle ends when it is consolidated, so the last step is the responsibility of the corrector, but, if needed, they may return the term to previous status.

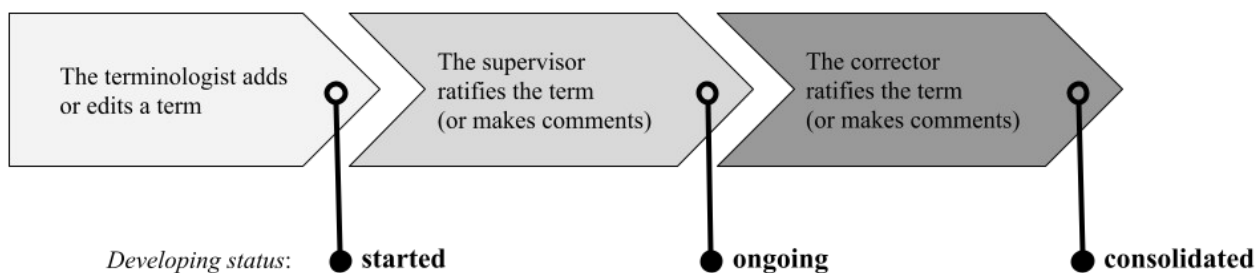


Figure 2: The working progress of a term, showing its developing status.

4. IZIBI-TZOS, a Civil Engineering Oriented TZOS

Within the same framework for describing terminology and focusing on one specialized area which is low-sourced in Basque, the Civil Engineering, we have built the so-called IZIBI-TZOS⁴, a Civil Engineering oriented version of TZOS.

Although TZOS was thought and aimed above all at dealing with academic works, its flexible structure and terminology classification by semantic fields (where the Civil Engineering is a subfield of the Technological Sciences), allow us to include more subfields within the Civil Engineering subfield. These subfields are those defined for the Civil Engineering profession, represented by the tree of concepts (Bridges and Structures, Building according to Loe 38/19995nv, Energy and Industry Areas, Environment, Geotechnics and Foundations, Health and Safety, Hydraulic, Ports and Coasts, Roads, Transportation, Urban Works and Services, Urbanism). That is, the IZIBI-TZOS database gathers both the academic area and the specific area of Civil Engineering in the same resource via the semantic classification tree. This can be seen in Figure 3, where the word *hargune*, together with its English equivalent *catchment*, is being added in the Hydraulic subfield of Civil Engineering, created specifically for IZIBI-TZOS.

In addition, new origins relating to the professional area such as companies and administrative entities, have also been added.

The first preliminary version of IZIBI-TZOS was created by making a copy of the general TZOS taking into consideration all the terms that were already included in the Civil Engineering subfield as well as the terms that are used in the subjects taught in the Civil Engineering degree which are 82,359 in total. We decided to make a copy because of two reasons: (1) to maintain the purely academic approach of the general terminological classification of TZOS (without adding a subclassification that only concerns the professional field of Civil Engineering); and (2) to offer the future professional civil engineering community an independent database that will facilitate the management.

Nowadays, we are working with experts from both the academic and professional areas to enrich the preliminary version and form the commissions based on the profiles explained in the previous Section 3.3: the experts will work on the profile they consider most suitable for them

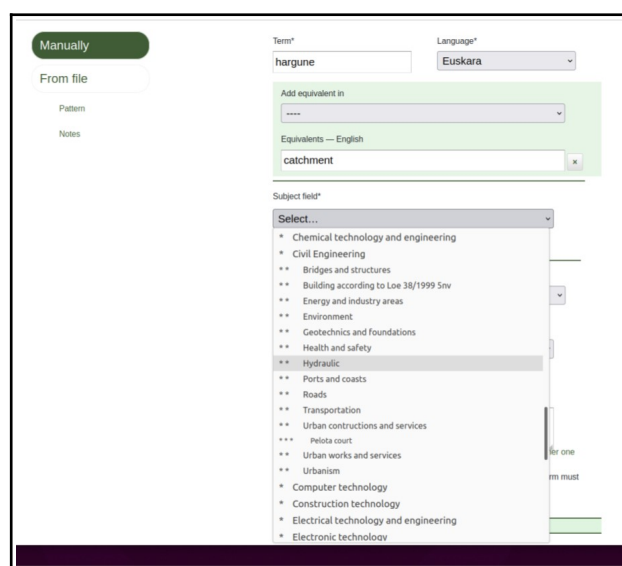


Figure 3: Adding a term in the “hydraulic” subfield of the Civil Engineering subfield.

(either *terminologist* or *supervisor*), and a linguist will be assigned for doing the final linguistic revision (*corrector*). Following the explained process for developing terms, the terminologist will start adding a term. If such a term is already included in the preliminary IZIBI-TZOS, both terms will collide and one of the following decisions should be made: (a) add the term for considering it as a different concept, (b) although considering the same concept, add some new information (such as definition, examples, variants...), or (c) consider it as the same concept, in which case the application would automatically store the author as another user.

This idea has been carried out through an agreement between the UPV/EHU and the College of Civil, Channels and Ports Engineers of the Basque Country, which is currently involved in a plan aimed at achieving greater use of Basque (or Basquisition). In this sense, IZIBI-TZOS will be a great resource to get a whole description of the terminology in Civil Engineering for future harmonization and normalization, as well as to build a community of experts in Civil Engineering. It will also be a good opportunity to open new lines of work, such as comparing the uses at university and professional areas, dealing with the most difficult aspects when creating terms in the area, and analyzing the most usual linguistic structures in the area.

⁴ <http://izibi-tzos.ehu.eus/login/>

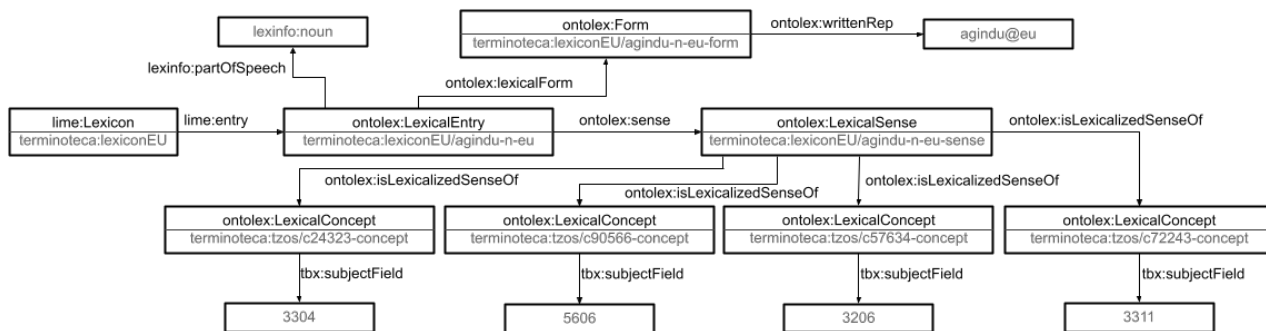


Figure 4: A TZOS entry modeled with *Ontolex-lemon*.

5. TZOS as Linguistic Linked Open Data

Although TZOS makes a significant contribution to the educational community as a resource that can be used as an online reference tool, we aim to make much more useful and accessible, even to other communities, by applying linked data principles to it. According to Chiarcos et al. (2013), applying linked data principles to linguistic resources has several advantages. For example, the use of comparable formalisms (formats, query languages...) and vocabularies to represent data from different types of resources makes them more interoperable, accessible and reusable for applications of different disciplines as they can be processed in a uniform way.

In that way, we plan to represent the terminology of TZOS as Linguistic Linked Open Data (LLOD). We have designed a prototype for a Resource Description Framework (RDF) (Klyne and Carroll., 2004) representation of TZOS terminology, modeled according to the *Ontolex-lemon* model. The *Ontolex-lemon* model is currently the most used model to represent lexical data as linked data (McCrae et al., 2017). The core of the model offers classes to encode the main elements of lexical descriptions (lexical entries, their senses, their relation to external general ontologies, or their inflected forms, among other aspects). Through a series of modules, other information such as translations or terminological relations of lexical entries can also be described. Broadly, a collection of terms for each language covered by TZOS is modeled as a *lime:Lexicon*. Thus, we have 5 lexicons, namely for Basque, Spanish, English, French and Latin. Additionally, we have an extra lexicon for symbols. In each of these lexicons each term (or symbol) is represented as *ontolex:LexicalEntry*. In order to refer to a particular concept evoked by a lexical entry or term the *ontolex:LexicalConcept* class is used. Lexical entries are linked to lexical concepts by means of an intermediate class called *ontolex:LexicalSense*, which captures a specific sense or meaning that each single entry has. An important property of a lexical concept is the *tbx:subjectField* which indicates a subfield in which this concept is used.

Each entry and concept in TZOS source data has a unique numeric identifier. We have changed the URIs of lexical entries adopting the same URI naming strategy of Terminoteca RDF (Bosque-Gil et al., 2016) to facilitate the integration of TZOS in this collection of interlinked

multilingual terminologies in the future; using the same URI pattern makes it possible to merge terms that come from different terminologies, linking them at the lexical entry level. In this approach, the URI of a lexical entry, form and sense is composed by their written representation, part of speech and language identifier; for example, *lexiconEU/agindu-n-eu* is the URI of the Basque noun *agindu* and the URI for the English noun instruction is *lexiconEN/instruction-n-en*. The URIs of lexical concepts preserve the numerical source identifiers (e.g. *tzos/c24323-concept*).

Figure 4 shows an example of a lexical entry of the Basque lexicon (the same one used in Figure 1). The shown lexical entry is the noun *agindu* (which could be *instruction*, *sentence*, *command*, *order* or *prescription* depending on the subfield), which is linked with four different concepts, each one from a different subfield. These subfields are specified using UNESCO identifiers like 3304, 5606, 3206 and 3311 which are the identifiers for the Computer Technology, Philosophy of Law, Nutritional Sciences and Instrumentation Technology, respectively. Besides, the entry is linked to its corresponding syntactic category through *lexinfo:partOfSpeech* property.

TZOS includes many synonyms or variants (orthographic, morphosyntactic, lexical and discursive) in the same language, and the *vartrans* module is used to model them. Two lexical senses from the same lexicon (thus, same language) that are linked to the same concept are related through the class *vartrans:LexicoSemanticRelation*. As there is no directionality in these variants or synonyms (there is not any source or target term), the property *vartrans:relates* is used in this relation. Figure 5 illustrates the representation of two Basque variants (*agindu* and *instrukzio*) which are linked to the c24323 concept. In such a multilingual terminology the modeling of the translations of terms from Basque to other languages are modeled with the *vartrans* module. More specifically, the *vartrans:Translation* is used to represent the translation between two lexical senses from different lexicons. Only translations from Basque to other languages have been modeled since the Basque lexicon is taken to be the core of the resource. Figure 6 is included as an example of an English translation (*command*) of the Basque entry *agindu*. The two translation senses are linked with the source and target properties to the *vartrans:Translation* element. We add an additional property, *vartrans:relates*, to this element too.

As commented before, we have a lexicon for Latin lexical entries which includes the scientific names of some terms.

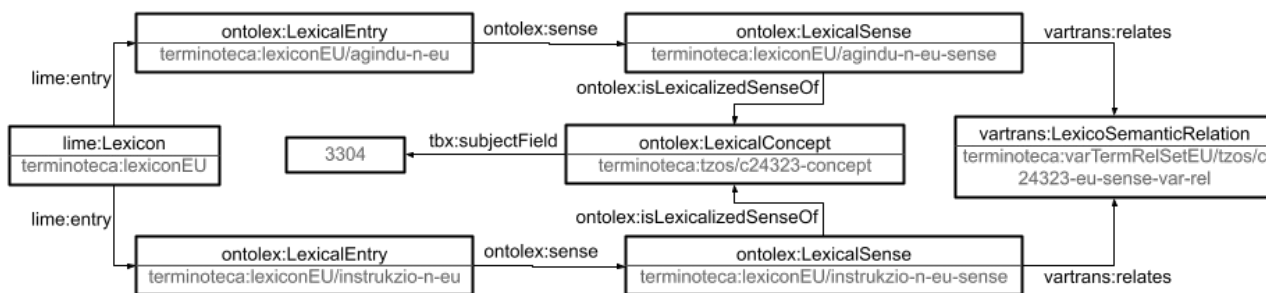


Figure 5: Two variants of the c24323 concept modeled with *vartrans*.

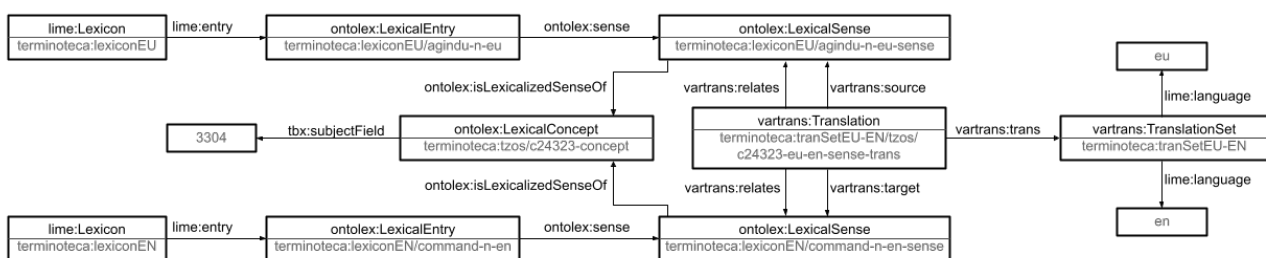


Figure 6: An English translation example of the Basque entry *agindu* (*command*).

These Latin terms are modeled as terminological variants through the *vartrans:LexicoSemanticRelation* class linking two senses from different languages (Basque and Latin), and capturing the specific type of variants (*lexinfo:internationalScientificTerm*) by the property *vartrans:category*.

This work of modeling is ongoing and we plan to add definitions and information about semantic categories to this representation in the future. Once completed, we will include it in the so-called LLOD cloud⁵. The LLOD cloud is a collaborative effort and it contains a number of monolingual and multilingual linguistic resources, like machine readable dictionaries, lexicon, thesauri and corpora. Moreover, we aim to include the RDF representation of TZOS in Terminoteca RDF, which is a collection of interlinked multilingual terminologies in Spain (Bosque-Gil et al., 2016).

This RDF representation of TZOS terminology is publicly available and can be downloaded from <https://www.ixaeus/node/12983>.

6. Conclusion

We have presented the online terminology database TZOS as a system for collaborative terminology work of real academic terminology. It combines a complete terminological formalization, with a robust online terminology service tool. Our main contributions are the following: (1) to allow sharing terminology resources and facilitating discussion groups to compensate for the lack of fluid communication networks in Basque among experts/teachers; (2) to contribute to the dissemination of terminology standardized by normalizing institution; and

(3) to monitor texts and terminology used in academic communication in Basque.

We have also opened a new line of work, extending TZOS to deal with the terminology of a professional area, specifically the Civil Engineering area, which has few terminological resources in Basque. By bringing together academic and professional experts, we help to form the community of the area and offer an interesting resource for the future harmonization and standardization of the terms as well as future research topics.

Moreover, we have worked on interlinking TZOS to other resources in the LLOD cloud to make it more useful and accessible to communities other than the educational one.

The future challenge of TZOS involves the incorporation and codification of semantic tags or classes (as hyperonyms) to enrich the term-entry. The intention is that professors and lecturers will group the terms used in their subjects into such semantic tags. However, this work is in its beginnings. Besides, we plan, on the one hand, to extend the collection and processing of texts necessary to make the corpus GARATERM representative, and, on the other hand, to obtain and process relevant and significant amounts of oral texts to integrate into GARATERM.

To summarize, we can state that TZOS and GARATERM are the keys to having a working active framework for describing and developing real academic terminology. The challenges for the future include integrating the results of TZOS and GARATERM with the agents and entities responsible for the normalization of Basque, such as Euskaltzaindia (The Royal Academy of the Basque Language), Commission of Terminology of the Basque Government, and UZEI (Basque Lexicography and Terminology Center).

⁵<http://linguistic-lod.org/lod-cloud>

7. Acknowledgements

This work is partially supported by the Basque Government (IXA excellence research group IT1343-19 and DeepText KK-2020/00088) and by the Spanish Ministry of Science, Innovation and Universities (DeepReading RTI2018-096846-B-C21, MCIU/AEI/FEDER, UE; HARTAvas PID2019-109683GB-C22; and DOTT-HEALTH/PAT-MED PID2019-106942RB-C31).

8. Bibliographical References

- Arregi, X., Arruarte, A., Artola, X., Zabala, I., and Lersundi, M. (2013). TZOS: An On-Line System for Terminology Service. In *Actualizaciones en Comunicación Social. Actas XIII Simposio Int. de Comunicación Social (Santiago de Cuba, 2013)*, pages 400-404. Centro de Lingüística Aplicada.
- Bosque-Gil, J., Montiel-Ponsoda, E., Gracia, J., and Aguado-de-Cea, G. (2016). Terminoteca RDF: a Gathering Point for Multilingual Terminologies in Spain. In *Proceedings of TKE 2016 the 12th International conference on Terminology and Knowledge Engineering*, pages 136-146.
- Cabré, M.T. (1993). La terminología: Teoría, metodología, aplicaciones. Editorial Antártida/Empúries.
- Cabré, M.T. (1999). La terminología. Representación y comunicación. Barcelona: IULA, Universitat Pompeu Fabra.
- Cabré, M.T. (2001). Sumario de principios que configuran la nueva propuesta teórica. In M.T. Cabré & J. Feliu (Eds.), *La terminología científico-técnica*, IULA, Universitat Pompeu Fabra, pages 17-26.
- Cabré, M.T. (2002). Terminología y lingüística: la teoría de las puertas. *Estudios de Lingüística del Español (ELiEs)*, vol. 16.
- Chiarcos, C., McCrae, J., Cimiano, P., and Fellbaum, C. (2013). Towards open data for linguistics: Lexical Linked Data. In Alessandro Oltramari, Piek Vossen, Lu Qin, and Eduard Hovy (eds.), *New Trends of Research in Ontologies and Lexical Resources*. Springer.
- Cyganiak, R., Wood, D., Lanthaler, M. (2014). Resource Description Framework (RDF) 1.1: Concepts and Abstract Syntax: W3C Recommendation 25 February 2014.
- Du, J., Alexantris, C., and Yu, P. (2020). Comparative Research on Terminology Databases in Europe and China. *Human Interaction, Emerging Technologies and Future Applications II. IHIET 2020. Advances in Intelligent Systems and Computing*, vol 1152. Springer, Cham.
- Faneca, C. R. (2021). Fundamentos teóricos y principios metodológicos para la creación de bases de datos terminológicas. El caso de BTERAD dentro del campo de las enfermedades raras del aparato digestivo. *Panace@: Revista de Medicina, Lenguaje y Traducción*, 32 (53), 109-117.
- Gurrutxaga, A., Saralegi, X., Ugartetxea, S., and Alegria, I. (2005). Erauzterm: euskarazko terminoak erazteko tresna erdiautomatiko. Mendebalde Kultur Alkartea, IX. Jardunaldiak: Euskera zientifiko-teknikoa.
- Gornostay, T. and Vasiljevs, A. (2014). Terminology Resources and Terminology Work Benefit from Cloud Services. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*.
- Kitanovic, O., Stankovic, R., Tomasevic, A., Skoric, M., Babic, I., and Kolonja, L. (2021). A Data Driven Approach for Raw Material Terminology. *Applied Sciences*, 11, 2892.
- Maslias, R. (2017). Converting the European Terminology Database IATE into the World's Largest Multilingual Data Space. In *Terminological Approaches in the European Context*. Paola Faini (ed.). Cambridge Scholars Publishing, pages 13-20.
- McCrae, J.P., Gil, J.B., Gràcia, J., Bitelaar, P., Cimiano, P. (2017). The OntoLex-Lemon Model: development and applications. In *Proceedings of the 5th Biennial Conference on Electronic Lexicography (eLex)*.
- Wuster, E. (1968). The Machine Tool. An Interlingual dictionary of basic concepts. London: Technical Press.
- Zabala, I., Lersundi, M., Leturia, I., Manterola, I., and Santander, G. (2013). GARATERM: euskararen erregistro akademikoen garapenaren ikerketarako lan-ingurunea. In *Ugarteburu terminologia jardunaldiak (V). terminologia naturala eta terminologia planifikatua euskararen normalizazioari begira*, pages 98-114.
- Zabala, I., Aldezabal, I., Aranzabe, M.J., Arriola, J.M., Gonzalez-Dios, I., and Lersundi, M. (2018). Corpus-driven Terminology Work for Describing Basque Academic Terminology: the Weaving Terminology Networks programme (TSE programme). In *Proceedings of EAFT Summit Poster session*.
- Zabala, I. (2019). The elaboration of Basque in academic and professional domains. In L. Grenoble, P. Lane, U. Røyneland, (Eds.), *Linguistic Minorities in Europe Online*. The Gruyter Mouton, pages 1-15.