

## CHAPTER TWENTY

# GRADUATE TEACHING OF SPECIALIZED REGISTERS IN A LANGUAGE IN THE NORMALIZATION PROCESS: TOWARDS A COMPREHENSIVE AND INTERDISCIPLINARY TREATMENT OF ACADEMIC BASQUE<sup>1</sup>

Igone ZABALA, Itziar SAN MARTIN, Mikel LERSUNDI and Agurtzane ELORDUI  
(University of the Basque Country)

### Abstract

Teaching specialized registers requires a previous description of the registers based on corpora that contain the linguistic resources that are employed by specialized users in academic and professional contexts. However, in the case of languages in the normalization process such as Basque, it is specially difficult to build corpora that reflect the spontaneous speech productions of real specialized communication. Moreover, the description of the linguistic uses that we may detect in such corpora are insufficient for a meaningful and interactive teaching of specialized registers: the speech patterns that need to be described and taught to the students are unstable, changing and often far from optimal, precisely because the registers are being developed. This paper describes a pilot scheme aimed at integrating the research and didactics of specialized registers together with the development and normalization of the registers. The pilot scheme includes the monitorization of the professors' real usage, the identification of the linguistic characteristics, the diagnosis of the achievements and deficiencies in the lexical and discursive development of such registers, the preparation of a retraining program to stress the deficiencies that are detected and, finally, the creation of networks to help identify and fix the terminology and phraseology that arises in the use. After the introduction, the paper contains a section on the communicative skills for academic and professional purposes that the University of the Basque Country (UPV/EHU) includes in the new study programs that are being developed. Section 3 is a brief description of the basics of the project GARATERM and of its implementation through several actions that have been designed to monitorize and dynamize the development of academic Basque.

### 1. Introduction

Nowadays nobody questions that teaching languages for academic and professional purposes requires previous work in describing such registers by employing representative corpora from the specialized fields (Biber 1988, 1994, 2006; Swales 1990, 2001, 2004; Parodi 2005, 2007; Cabré and Gómez de Enterría 2006, among others). However, mere description of such registers by linguists does not suffice in the case of languages that are in the process of normalization like Basque. At least two features differentiate Basque from normalized languages. On the one hand, the standardization process is too recent and it is not completed yet. On the other hand, Basque has recently incorporated into the academic field and it has little presence in professional areas. In fact, the standard variant is neither completely implemented nor functionally adapted to the uses of specialized discourse areas, since there has been no time to completely develop and fix the phraseology and terminology of the specialized fields (Elordui and Zabala 2005, Zabala and Elordui 2006). As a consequence, the speech patterns that need to be described and taught to the students are unstable, changing and often far from optimal. In addition, to get representative corpora of real specialized uses entails great difficulties for most of the fields. Considering all these factors, we have reached the conclusion that a didactic model that pursues a meaningful and interactive type of teaching must necessarily consider the actual sociolinguistic state of both the language and of the specialized discourse communities that use the language.

In this paper we present a pilot scheme integrated in the project GARATERM that was carried out during the year 2008, as well as its implementation in the academic courses 2009-2010 and 2010-2011 in various faculties of the University of the Basque Country. The project GARATERM takes into account the actual sociolinguistic factors that are conditioning the development of academic and professional registers, and it assumes the idea that the role of the professor of Basque for academic and professional purposes cannot be restricted to teaching specialized linguistic elements based on the already existing specialized corpora<sup>2</sup>. Just like in the case of normalized languages, it is absolutely necessary to identify the discourse communities that employ Basque in academic and professional contexts as well as to describe their actual usage of the language; however, the specialized registers in Basque are still being created, and hence, it is important to make a diagnosis of the achievements and deficiencies in the area of the lexical and discursive development. This is essential in order for the students to become active agents in the lexical and discursive elaboration and development of the language in their areas.

---

<sup>1</sup> This work is related to three other projects: RICOTERM-3, sponsored by the Ministerio de Ciencia e Innovación MICINN of the Spanish Government (HUM2007-65966-C02-02), EBALUA (EHU 08/53), which has received financial support from the University of the Basque Country and the project HIZLAN (DIPE08/16), supported by the Provincial Council of Biscay.

<sup>2</sup> The only specialized corpus in Basque available for the public is the technical-scientific corpus (*ZT corpora*) created by Elhuyar and the IXA group. This corpus has 8,5 million words from different technical and scientific fields. The characteristics of the corpus are in the following address: <http://www.elhuyar.org/hizkuntza-zerbitzuak/EN/ZT-Corpusa>.

The plan that we will describe below follows the theoretical and methodological tenets of the Communicative Theory of Terminology (Cabr e 1999, 2001). This theory claims that any work on terminology or on managing specialized linguistic elements requires a previous description of the way these elements function in specialized texts<sup>3</sup>. Our view is, therefore, opposite to other prescriptive and corrective approaches that are based on the knowledge of the literary tradition and on general grammar<sup>4</sup>. It is relevant to point out the importance that the study of variation has in the framework, which we adopt in our work. The corrective-prescriptive approaches impose a too strict control on the language and, as a consequence, they many times ignore the direct correlation that exists between the lexical and discursive development of a language and the increases in the functional variation. From our viewpoint, the grammatical mistakes that we may find in the diagnosis have to be taken into account, but nevertheless, we consider that special attention needs to be paid to other aspects, such as: a) the identification of new uses and new resources<sup>5</sup> that the specialists create in order to fulfill the needs of specialized communication, b) the study of functional variation depending on the areas and the degrees of specialization of texts and genres<sup>6</sup> and c) the detection of asystematic variation due to problems for fixing new resources of the language.

We have adopted the theoretical and methodological principles described above in order to carry out our pilot scheme, which includes several tasks, such as the collection of texts produced by university teachers, the diagnosis of incorrect uses, the description of new linguistic uses motivated by specialized communication, terminology extraction and the setting up of networks that enable diffusing and fixing terminology and the new specialized uses. In order to carry out the different stages of the pilot scheme, we have tested various tools for corpus processing, for automatic analyses of linguistic elements and for the extraction and management of terminology. This experience is a step forward in the integration of several semiautomatic tools, resources and products in the platform GARATERM with the aim of combining the research, the development and the didactics of academic registers. The plan requires a comprehensive and interdisciplinary approach that includes the participation of linguists, experts in language technologies and specialists in various subject areas.

## 2. Academic and professional communication skills in the new study plans of the University of the Basque Country

The adaptation to the requirements of the new European Space of Higher Education (ESHE) entails adopting a new methodological framework that views university teaching as a process in which students progressively approach, incorporate and integrate in the professional community of their area. The development of the skills for communicating among experts and professionals is considered as fundamental for the integration of the students in the professional community of experts that they eventually aim to belong. This new view is clearly reflected in the so called White Books of the degrees, which stress the importance of the communicative skills in academic and professional life. In fact, the communicative skills are amongst the most important in the lists of both the general skills and of the specific skills of the degrees.

With respect to the study plans of the University of the Basque Country, the reference documents for the design of the new degrees emphasize the importance of the communicative skills. This is so to the extent that professors that teach in any of the three languages used at the university (Spanish, Basque and English) are being offered the possibility to participate in seminars (tutored by professors of the Department of Basque Philology) that help reflect on the way they can include some of the skills related to specialized communication in their programs<sup>7</sup>.

Apart from the above-mentioned initiatives, the *Guiding Plan for Basque of the University of the Basque Country* approved in 2007<sup>8</sup> is intended to compensate for the difficulties that may arise in the process of acquiring communicative skills in Basque due to the fact that it is not normalized. Thus, all the new degrees will keep 12 credits for elective subjects to teach specialized communication in Basque. These new subjects replace the existing elective subjects called *Technical Basque* that are being offered in all the current study plans and which are the result of the former 1990 *Guiding Plan for Basque*. The new subjects will also be offered by professors from the Department of Basque Philology.

All these changes in the study plans have encouraged a profound reflection about the didactics of Basque in specialized uses that some professors of the department of Basque Philology had already started during the last decade. The professors that participated in such debates agreed on the fact that, as in the case of other languages, the didactics of Basque in specialized communicative contexts must be preceded by a description of the registers that are to be taught and learned.

This new approach makes complete sense in the current sociolinguistic context: experts have increasingly used Basque in their academic and professional areas and hence, the number of specialized productions that can be compiled in a corpus and described is also growing. The times are over where the role of the professor came down to describing the grammatical rules and the system of word formation mechanisms in Basque and to promoting the coining and diffusion of lexical and discursive resources –that were often created *in vitro*– and whose aim was to promote the use of Basque in specialized communicative

<sup>3</sup> The methodological premises that have been mentioned for the Communicative Theory of Terminology are also adopted by other current theories of terminology such as the Socioterminology (Gaudin 1993; Boulanger 1991) and the Sociocognitive Theory of Terminology (Temmerman 2000).

<sup>4</sup> Some thoughts on these two types of initiatives for Basque can be found in Zabala and Elordui (2006). Flowerdew (2004) makes a similar reflection for English.

<sup>5</sup> We intentionally distinguish between the activation of new specialized uses of already existing linguistic elements and the creation of new specialized elements: both are necessary processes for the lexical and discursive development of the language in specialized fields.

<sup>6</sup> For a study of variation of terminology in Basque see Elordui and Zabala (2005).

<sup>7</sup> The materials of these seminars can be accessed in the following address: <http://www.ehu.es/PAT/competencias.htm>

<sup>8</sup> The *Guiding Plan for Basque* in the UPV/EHU can be checked in the following address: [http://www.euskara-errektoeordetza.ehu.es/p267-content/es/contenidos/informacion/plan\\_informa/es\\_plan/adjuntos/plan\\_director.pdf](http://www.euskara-errektoeordetza.ehu.es/p267-content/es/contenidos/informacion/plan_informa/es_plan/adjuntos/plan_director.pdf)

contexts. Nowadays, we may state that the main goal is the description of the experts' actual use of the language, without ignoring that the description and monitorization of these uses must go hand in hand with a process of diagnosis and dynamization.

### 3. The implementation of the project GARATERM: monitorizing and dynamizing the development of Basque in academic contexts

Last year, during our research within the project GARATERM, we detected the components that should be contained in an initiative aimed at integrating the research and the teaching of specialized registers together with the development and normalization of the registers. As it could not be otherwise, the new language technologies would play a key role in this integration process that we planed to start.

To begin with, any corpus building prior to the description of specialized registers in Basque must consider the fact that such incipient registers are inherently unstable. Thus, in order to get a valid characterization of specialized registers we need corpora of the monitor-type<sup>9</sup>, which allow us to consider the diachronic component in our analysis. Our study previous to the monitorization of the lexical and discursive development of the specialized registers reveals that the use of such registers is mostly found in academic contexts and often restricted to oral texts. Moreover, whereas the communication networks between professors and students are very active, communication between experts that use Basque in their area of specialization is still very limited<sup>10</sup>.

Fortunately, due the growing use of the Information and Communications Technologies (ICT) in teaching, most professors create materials in digital format (programs, notes, presentations, protocols for laboratory experiments, etc.). In our view, these types of materials reflect best the actual use of the language in academic contexts, better than many publications that are specifically created in Basque for teaching or for popularization purposes. In fact, the latter are mostly elaborated by translators and they are controlled by entities that offer language services that control the texts from a corrective and prescriptive view. Such sociolinguistic type of controls are seldom based on systematic descriptions of the real uses that linguistic communities make. However, these communities are precisely the potential users of the publications aforementioned. Thus, when studying the specialized discourse in Basque it is extremely important to consider the controlled or spontaneous characteristic of the texts that are included in the corpora, specially when we make comparative studies.

The monitorization and the dynamization processes of the development of academic Basque are complementary, but each contains its own objectives, which are described below:

On the one hand, the main goals of monitoring the university professors' use of Basque are the following:

- a) To make visible the academic uses of Basque that are presently invisible for corpus creation and also for the codification of standard Basque and for the normalization of terminology.
- b) To create a monitor corpus which includes texts from various specialized fields, from different genres and levels of specialization and of various degrees of spontaneity.
- c) To exploit the corpus monitor with the following aims: c1) the characterization of the specialized discourse from various areas in Basque, c2) the diagnosis of the progress and the difficulties in the lexical and discursive development of academic registers, and c3) the study of how sociolinguistic control over academic registers has an influence in the development of such registers.

The main objectives of the dynamization of the lexical-discursive development of academic Basque are the following:

- a) To contribute to improve the quality of the texts that university professors create. This will dynamize the implementation and functional adaptation of standard Basque at the university, since students will mostly get the linguistic input that they need to develop their communication skills in academic Basque from the productions of their professors.
- b) To describe the linguistic resources (terminology, phraseology, stylistic options and discourse strategies) employed in the production on university texts..
- c) To analyze the variation in terminology and phraseology detected in corpora with the aim of determining the causes of such variation, which may be: functional reasons, trouble in fixing newly created lexical and discursive units, sociolinguistic control or non-systematic intervention.
- d) To create networks and platforms for helping communication among university professors. This would enable the diffusion, discussion and fixation of the terminology and phraseology.
- e) To edit and diffuse the products (glossaries, phraseological dictionaries, style cards...) that may contribute making the linguistic resources that have been created more visible and fixed.

We employ several language technology tools that help achieve all those aims. These tools are currently being used independently from each other, but our goal is to integrate them in a single environment. This will speed up the processes, which in turn will entail that we have more texts and data to consider and work with. We will employ the tools *Dokumentu Biltegia* (DB) (Ezeiza 2010) and *Corpusgile* (Areta et al. 2006) for the compilation of texts, for the formal and linguistic processing of the corpus and for the analysis of linguistic elements included in the tools. For the extraction of terminology we

<sup>9</sup> According to the classification in Sinclair (1991) monitor corpus is a type of corpus which is a growing, non-finite collection of texts. Monitor corpus reflects language changes in a constant growth rate of corpora, leaving untouched the relative weight of its components (i.e. balance) as defined by the parameters. The same composition schema should be followed year by year, the basis being a reference corpus with texts spoken or written in one single year.

<sup>10</sup> According to the information gathered by Artetxe (2010), out of the data collected for 1340 professors in the INGUMA database 408 professors (30,4 %) do not contain any production in Basque.

will use the tool *Erauzterm* (Alegria et al. 2004). Finally, we will use *Terminologia Zerbitzurako on line Sistema* (TZOS) (Arregi 2010) for the creation of networks for the diffusion and discussion of terminology among university professors.

With regards to the dynamization of the academic registers at the university, we need to mention the workshops for linguistic retraining that are being held since academic year 2008-2009. From the present year on these seminars will be named *Terminologia Sareak Ehunduz* (Weaving Terminology Networks) and they will be included amongst the programs pertaining to the Vice-chancellorship of Quality and Innovation of the University of the Basque Country. These workshops are the testing ground for the tools and besides, they can be considered as the embryo for the implementation of the project GARATERM. The work unit of the professors that participate in the workshops is a subject that they typically teach. We collect the different texts that the professors use for the subject (syllabus, exams, notes, protocols for experiments, etc.) and we process them. The texts that we collect are also used for making linguistic diagnosis of the teachers' use and for the extraction of frequent combinations of words and terms. The diagnosis of the texts is used to design a retraining plan, which includes terminological discussions by using the tool TZOS. During and after the retraining workshops, we offer the professors the chance to revise their texts. In the end, we get a corpus of texts that contains the documents of the subjects that the professors use in class. We organize these documents by degree and also by subject and we also elaborate terminological glossaries that derive from the extraction and discussion of terminology contained in the material of the subject. Finally, the linguists participate in elaborating style cards from the information obtained from the analysis of the texts. All these will be offered to the university community through the platform GARATERM. The following picture shows the processes and tools that have been included in the implementation of the project GARATERM.

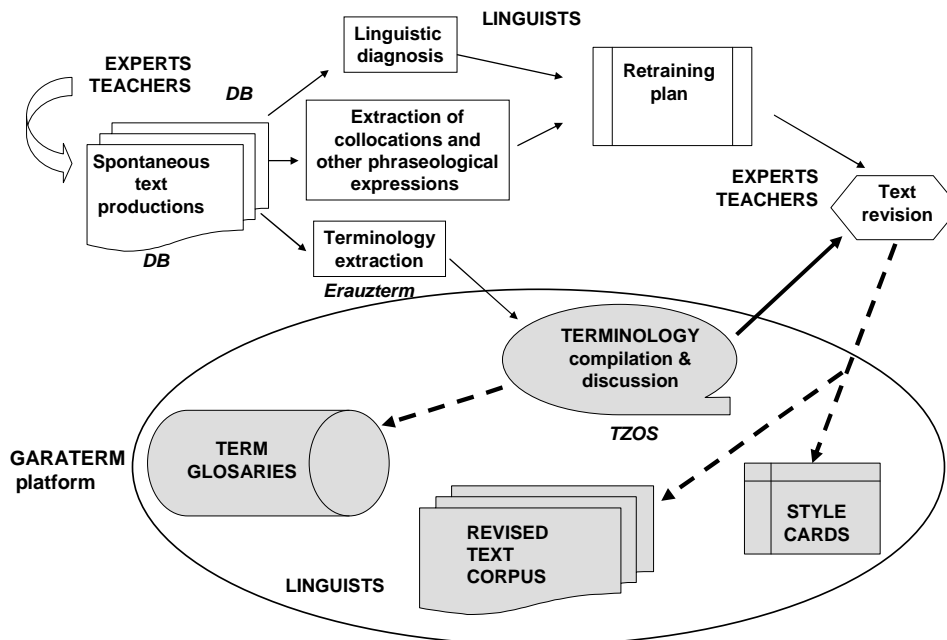


Fig. X-1. Implementation of GARATERM project: monitorization and dynamization of the development of Basque in academic fields

## REFERENCES

- Alegria, Iñaki, Antton Gurrutxaga, Pili Lizaso, Xabier Saralegi, Sahats Ugartetxea, and Ruben Urizar. 2004. "An Xml-Based Term Extraction Tool for Basque." *4th International Conference on Language Resources and Evaluations (LREC)*. Lisbon.
- Areta, Nerea, Antton Gurrutxaga, Igor Leturia, Ziortza Polin, Rafa Saiz, Iñaki Alegria, Xabier Artola, Arantza Diaz de Ilarraza, Nerea Ezeiza, Aitor Sologaitoa, Aitor Soroa, and Andoni Valverde. 2006. "Structure, Annotation and Tools in the Basque ZT Corpus." In *Language Resources Evaluation Conference (LREC)*.
- Arregi, Xabier, Ana Arruarte, Xabier Artola, Mikel Lersundi, Gotzon Santander, and Joseba Umbelina. 2010. "TZOS: Terminologia Zerbitzurako On-lineko Sistema." In *Ugarteburu Terminologia Jardunaldiak: Euskararen garapena esparru akademikoetan*, edited by Pello Salaburu and Xabier Alberdi. Bilbao: University of the Basque Country Press.
- Artetxe, Karmele. 2010. "Komunikazio zientifiko euskalduna; azterketa historikoa eta bibliometrikoa." PhD diss., University of the Basque Country.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, Douglas. 1994. "Using register-diversified corpora for general language Studies." In *Using large corpora*, edited by Susan Armstrong, 180-201. Cambridge: The MIT Press.
- Biber, Douglas. 2006. *University language. A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- Boulanger, Jean-Claude. 1991. "Une lecture socioculturelle de la terminologie." *Cahiers de Linguistique Sociale* 18: 13-30.

- Cabré, Maria Teresa. 1999. "Elementos para una teoría de la terminología: hacia un paradigma alternativo." In *La terminología. Representación y comunicación*, edited by Maria Teresa Cabré, 69-92. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, Maria Teresa. 2001. "Sumario de principios que Configuran la nueva propuesta teórica." In *La terminología científico-técnica*, edited by Maria Teresa Cabré and Judit Feliu, 17-26. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, Maria Teresa, and Josefa Gómez de Enterría. 2006. *La enseñanza de los lenguajes de especialidad. La simulación global*. Madrid: Gredos.
- Elordui, Agurtzane, and Igone Zabala. 2005. "Terminological Variation in Basque. Analysis of Texts of Different Degrees of Specialization." *SKY Journal of Linguistics* 18: 71-91.
- Ezeiza, Joseba. 2010. "DB (Dokumentu Biltegia): corpus akademikoak sortzeko eta kudeatzeko azpiegitura teknologikoa." In *Ugarteburu Terminologia Jardunaldiak: Euskararen garapena esparru akademikoetan*, edited by Pello Salaburu and Xabier Alberdi. Bilbao: University of the Basque Country Press.
- Flowerdew, Lyne. 2004. "The argument to using English specialized corpora to understand academic and professional languages." In *Discourse in the Professions. Perspectives from corpus Linguistics*, edited by Ulla Comnor and Thomas Albin Upton, 11-33. Amsterdam: John Benjamins.
- Gaudin, François. 1993. *Pour une socioterminologie: des problèmes pratiques aux pratiques institutionnelles*. Rouen: Publications de l'Université de Rouen.
- Parodi, Giovanni. 2005. *Discurso especializado e instituciones formadoras*. Valparaíso: Ediciones Universitarias de Valparaíso.
- Parodi, Giovanni. 2007. "El discurso especializado escrito en el ámbito universitario y profesional: Constitución de un corpus de estudio." *Revista Signos* 40 (63): 147-179.
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Swales, John. 1990. *Genre Analysis: English in Academic and Research Setting*. Cambridge: Cambridge University Press.
- Swales, John. 2001. "EAP-related linguistic research: an intellectual history." In *Research Perspectives on English for Academic Purposes*, edited by John Flowerdew and Mathew Peacock, 42-55. Cambridge: Cambridge University Press.
- Swales, John. 2004. *Research genres. Explorations and applications*. Cambridge: Cambridge University Press.
- Temmerman, Rita. 2000. "Why traditional terminology theory impedes a realistic description of categories and terms in the life sciences." *Terminology* 5, 1: 77-92.
- Zabala, Igone, and Agurtzane Elordui. 2005. "Euskara zientifikoaren normalizazioa: hizkuntzalarien ekarpena." In *Euskera zientifiko-teknikoa: normalizaziotik homologaziora*, edited by Jesus Mari Txurruka, 15-54. Bilbao: Mendebalde Kultura Alkartea.
- Zabala, Igone, and Agurtzane Elordui. 2006. "Specialized Discourse and the Linguistic System of Basque." In *Andolin gogoan. Essays in honour of Professor Eguzkitza*, edited by Beatriz Fernández and Itziar Laka, 895-914. Bilbao: University of the Basque Country Press.