

Adaptation to the EHEA in the training of professionals in automatic natural language processing

Lorraine Baqué and Àngels Catena

Department of French and Romance Languages and Literatures
Faculty of Philosophy and Arts
Universitat Autònoma de Barcelona

Abstract

Globalisation has meant an exponential increase in multilingual communication. The technological advances associated with the democratisation of access to information and global knowledge in a multitude of languages require new multilingual instruments that enable us to manage and organise the flow of written and audiovisual documents that are being produced. In order to meet this rising need of multilingual professionals specialised in language training for applications in the different areas of the language industry, we must propose an advanced, eminently multidisciplinary and trans-linguistic education. However, in order to achieve this goal, we must design new teaching strategies that enable us to manage the diversity of initial educational levels and contact languages.

General area of interest of this innovation

The goal of the innovation presented in this article was to improve the management of diversity (both linguistic and initial educational levels) of students in a European Master's in automatic natural language processing (NLP), but beyond the speciality in which it is applied, it might be of interest to any professors who in their respective fields must confront interdisciplinary or multilingual diversity in the classroom, and both manage and foster it.

1. Objectives

The impact of the new information and communication technologies in today's society is linked to widespread access to multilingual information sources that also enable contents generated locally to be disseminated worldwide. This is true to such an extent

that, for example, in its 2002 *Rapport de discussion sur la stratégie d'innovation du Canada*, the government of Canada proposed a strategy called «prioritising the language industries in the national educational strategy» (<http://www.innovationstrategy.gc.ca/gol/innovation/site.nsf/fr/in02346.html>).

For this reason, in a globalised context, languages have become an essential element, and the training of multilingual professionals specialised in language training for applications in the different areas of the language industry requires an advanced, eminently multidisciplinary and trans-linguistic education

This is the educational profile that defined the design of the European Master's in Automatic Natural Language Processing (NLP), developed in conjunction with other European universities in terms of both the contents of the programme and the teaching methodology. However, in order to achieve this goal we had to design new teaching strategies that would enable us to manage the diversity of different initial educational levels (related to knowledge of both technology and language) and contact languages. This need, which was already accounted for in the initial design of the Master's, became even more clear when the programme was implemented. Indeed, the students that signed up for this programme came from a wide range of degree programmes, while they were also specialised in a wide variety of languages. In order to carry out the prior diagnosis before beginning the project which served as a basis for the improvement plan, we first analysed the distribution by languages and by field of speciality of all the students in the first two years that the Master's was implemented.

With regard to prior education, if we divided the degrees into two large groups of language specialities and technology specialities, we can see that despite the fact that the majority of students came from the languages (in the broad sense, including some specialities in natural language processing but with a predominance of languages), at the start 28.2% of the students in the the NLP Master's came from technology-oriented degree programmes (IT, engineering, etc.). This was considered quite favourable for the programme, but it required an assessment of how to manage this diversity.

With regard to the languages of the group which was used to make the initial diagnosis, all the students were at least trilingual, but they had fourteen different native languages: Catalan, Spanish, French, English, German, Italian, Russian, Polish, Romanian, Portuguese, Arabic, Japanese, Hungarian and Greek. And, if we bear in mind only the first three languages of each, we saw that the students could work in 24 different language combinations, which is extremely rich for creating multilingual instruments that were to be made. Nevertheless, there was no common language among all the students registered.

As a result, the project's goals included an interest in facilitating the creation of digital and multimedia teaching resources in order to increase the out-of-classroom support actions by the use of ICT and to ensure the quality of the programme in terms of the results of the learning acquired, and in particular students' attainment of the skills related to multi-

lingualism and multidisciplinary, given the fact that these aspects are not just relevant for the existent target but are also core for the profile of a Master's education.

2.2. Description of the project

2.1. Point of departure

The project presented in this article is part of a sweeping process of methodological reflection on the European Master's in NLP after the first two years it was implemented. The working group on the project organised the work into several different phases and goals.

First of all, we made a self-evaluation of the degree programme as a whole, which ended by drawing up a corresponding report for AQU [Catalan Agency for University Quality Assurance] based on the results of the first two courses on a variety of points, including the academic results and satisfaction of students and professors.

With regard to the academic results, what particularly stood out was that the number of students who did not show up for the evaluation of several courses rose in the second course of the Master's programme compared to the first course and compared to the previous figures from the doctoral programme. Likewise, we also noticed a certain improvements in students' academic performance compared to previous years.

With regard to the second point, we could see that the degree of satisfaction among the students (3.78 points over 5) and the faculty (3.3 points over 4) was quite high overall. However, both groups negatively rated the items related to the teaching spaces and, in the case of the faculty, there was a great deal of dissatisfaction with the external dissemination of information about the programme (1.8 points over 4).

This self-evaluation, which also included a qualitative evaluation, led to a diagnoses of the strong and weak points of the Master's and an improvement plan, which took the following factors into account:

1. The lack of teaching resources for blended learning. Even though the faculty in the programme had made many efforts to adapt to the ECTS methodology, the results of this undertaking had mainly resulted in a significant improvement in the teaching-learning activities in the educational programme, not just the conceptual contents but also the more procedural and attitudinal contents. In contrast, there was a clear and major lack of digital materials that would enable the type of activity done in class to be supplemented by independent learning activities using the new technologies. The creation of this type of digital and multimedia resource was also aimed at offsetting the weak points mentioned by the students on the equipment in the teaching labs.
2. The need to foster exchanges of teaching experiences among the faculty of the different partner universities in order to encourage the convergence of teaching methodologies used at the different universities that were part of the consortium.

3. The wisdom of validating the competences designed (cf. AQU, Tools). In the subsequent coordination meetings, the absence of external validation mechanisms (by academics and professionals) became clear, in terms of both the relevance of this type of academic programme and the effective attainment of the competences it aims to convey.

In the project to improve teaching quality, we worked on these diverse aspects of improvement, although here we only present the points in which the methodological efforts of the faculty team were more heavily required: the design of functionalities for the virtual campus, the creation of multimedia material for independent, semi-independent or collaborative study among peers, and the design of pedagogical activities to foster students' multilingualism and multidisciplinary.

3. Methodology

From the organisational standpoint, the project presented in this article had to set up a series of actions, means and resources aimed primarily at implementing a virtual environment that would foster the support of learning, encourage the development of teaching materials with ICTs and design pedagogical activities that would foster multilingualism and multi-disciplinary.

First, to resolve the difficulties generated by the initial multilingualism and multi-disciplinary, we believed that we had to create a specific virtual space for the Master's programme that would facilitate interaction and the exchange of documents and materials among students and between students and professors, in particular for the students and professors not just of the UAB but also, in a later phase, for all the partner universities. Therefore, with the help of the Office of Interactive Autonomous Teaching (OAID) of the UAB, we designed and made a teaching platform that has a public area presenting the Master's and a restricted area that has the usual functionalities of a virtual campus (teaching materials, file management, turning in projects, email, news and notices), yet also functionalities such as forums in large and small groups, for collaborative work and for self-evaluation. We should stress that this restricted area is managed by the programme itself, which enables us to create groups and subgroups, create transversal, inter-course activities and include the contributions by the other partner universities.

Secondly, based on the existence of a teaching platform tailor-made for the Master's programme, the faculty of the Master's started to draw up specific materials for the different courses, using and taking advantage of the new information and communication technologies. Indeed, when we started the project, this type of material only existed for one of the Master's courses. We have tried to ensure that this material follows the principles of fostering student autonomy, that it includes not just contents but also educational activities aimed at resolving problems and collaborative work with peers. We should also stress that support materials were developed for transversal learning in several different courses within the Master's programme.

In order to foster the multilingualism of the students and minimise the difficulties they might encounter in understanding classes, the literature and the materials, conducted two specific actions. First of all, we decided to use the GALANET teaching platform for inter-comprehension among the Romance languages (<http://www.galanet.eu>), which includes a space of language resources in Spanish, French, Italian and Portuguese, along with an independent learning area to help students understand an unfamiliar related language, as well as to help speakers of different related languages to understand each other. We also held an interactive training session among groups of students from four different languages from around the world which asked them to complete a task together (write a press release on a subject agreed upon in all four languages) collaboratively among the different language groups of the students, which involves understanding written and oral documents in all four languages.

Likewise, we encouraged the inclusion of a learning activity in some of the Master's courses that consists of each student searching for, for each of the courses involved, an introductory bibliography of three articles in three different languages (native language, a second language and an unknown language but one that is related to the student's languages), making an oral summary of them in the language of their choice (Catalan, Spanish, French, Italian, Portuguese or English) in front of their classmates, explaining it in such a way that they can understand it. The students have to take notes and make a written summary (in any of the same languages) of all the articles based on their classmates' presentations (in different languages).

Finally, as mentioned above, multidisciplinary is present at three levels in the NLP Master's: in the discipline of automatic natural language processing itself, in the profile of the training, which includes both linguistic and technological contents and procedures, and in the students' educational background.

This is a considerable wealth, but it also means resolving the difficulties derived from such diversity. The goal is to empower students in each of the specialities in the skills of the other, as well as to train them to work on multidisciplinary teams to resolve the same problem cooperatively, that is, not just working in groups but also finding compatible procedures.

In order to foster this multidisciplinary not just in the contents but also in the practical aspects, which is essential in the working world, we conducted two actions.

First, we encouraged the creation of educational activities that affect more than one course in order to lower the atomisation of the contents and procedures, and secondly we designed multidisciplinary problems that the students have to resolve in small groups made up of students from educational backgrounds in both technologies and languages.

4. Results

The results of the project to improve the teaching quality were mainly seen in three areas: first, in the quantity (and quality) of the teaching materials to support learning and peer

cooperative activities and activities for working groups, secondly in the increase in the students' functional multilingualism, and finally in fostering multidisciplinary.

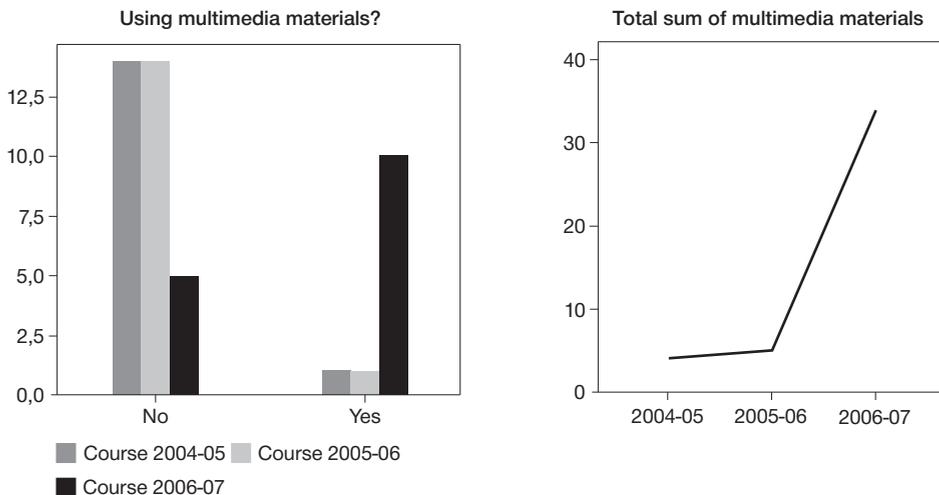
The results we can present here affect the changes that took place between the first two courses in the NLP Master's (2004-06) and the third course (2006-07), in which some of the innovations mentioned above were applied.

4.1. Quantity (and quality) of the teaching materials to support learning

We can evaluate the quality and quantity of the multimedia teaching materials to support learning according to the following criteria: a) the existence (or not) of materials for each of the courses in the programme; b) the total number of materials in all the courses combined; c) the existence (or not) of activities with an NLP approach for each course; and d) the existence (or not) of group work activities.

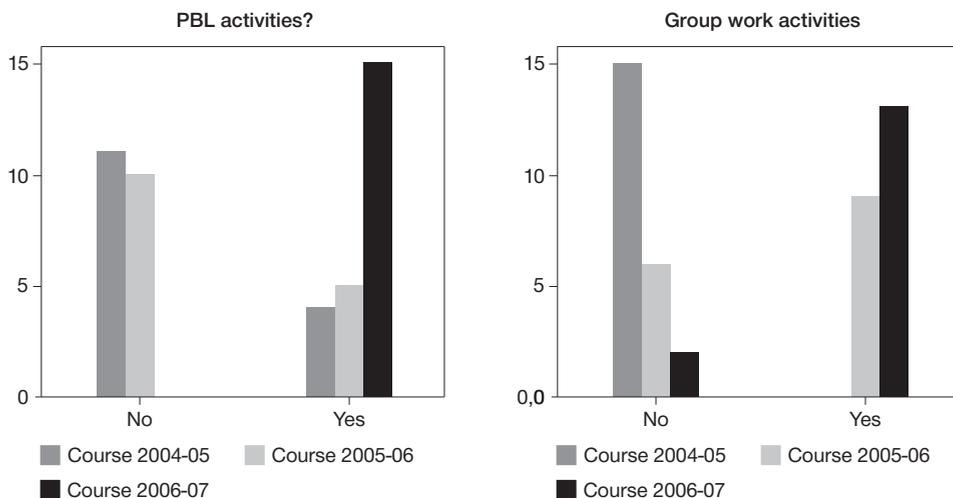
As can be seen in Figure 1, the evolution over the three-year period is considerable in terms of both the number of courses that have multimedia support activities (one until 2005-06 vs. ten in 2006-07) and the total number of materials (fewer than ten until 2005-06 vs. more than 30 in 2006-07).

Figure 1. Number of courses with multimedia support activities and total number of materials.



With regard to the presence or absence of pedagogical activities involving problem-solving or group work, we can also see that they have at least tripled in academic year 2006-07, to the extent that all the courses now have NLP activities and more than 85% include group work activities.

Figure 2. Number of courses with NLP activities and group work activities.



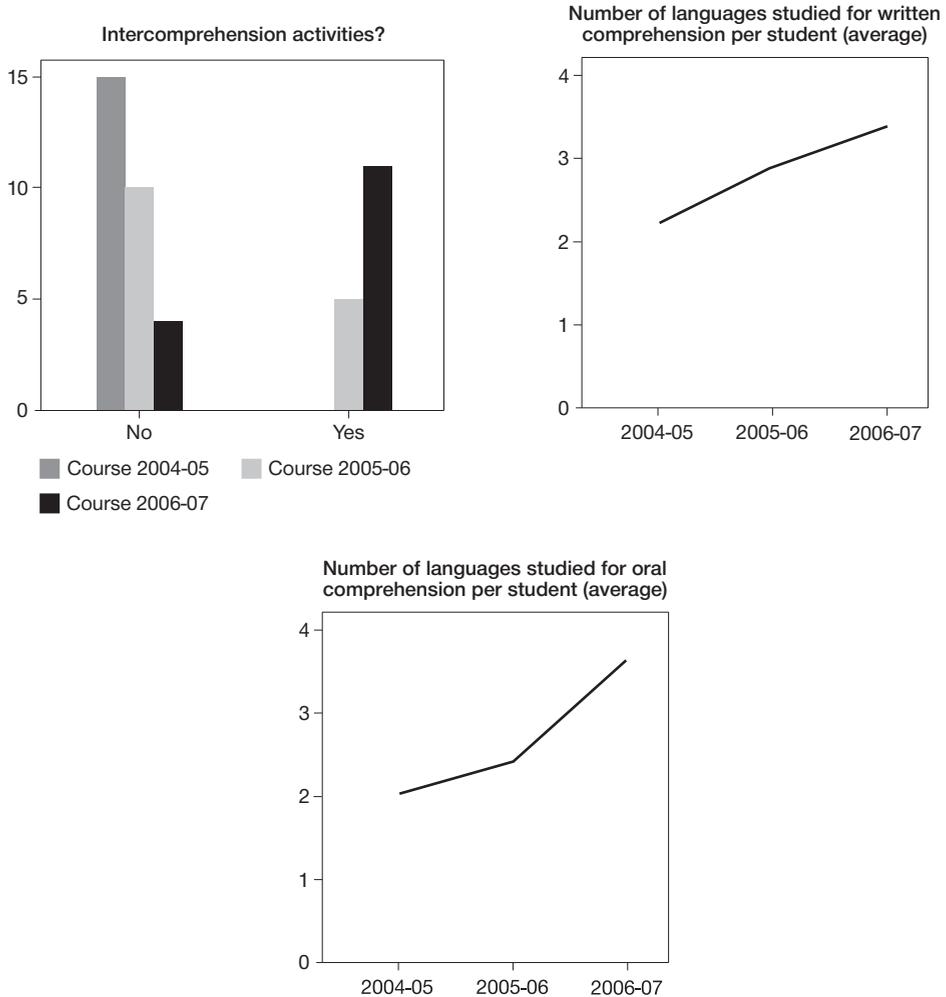
4.2. Increase in students' functional multilingualism

In view of our students' linguistic diversity, we have chosen two indicators on improvement in the innovation: a) the number of modules that include specific activities to foster multilingualism (generally speaking, by means of activities that encourage mutual comprehension among related languages) and b) not the number of languages our students are capable of understanding orally and in writing, but the number of languages in which each student on average has had to do the evaluated oral and written comprehension activities. The results can be seen in Figure 3. The number of courses that include specific activities has multiplied twofold in the past academic year, and the number of languages that each student worked on in written comprehension rose significantly (Sig.=.000) in recent years (2004-05: 2.21; 2005-06: 2.90 and 2006-07: 3.39). With regard to oral comprehension, we have seen an even more pronounced rise (2004-05: 2.00; 2005-06: 2.39 and 2006-07: 3.62; Sig.=.000).

4.3. Fostering students' multidisciplinary

In order to evaluate the effect of the actions undertaken to foster students' interdisciplinarity, we adopted the indicator on the number of courses that have included specific activities involving working in interdisciplinary groups. In the majority of cases, this involved problem solving which requires both technological and language skills at the same time. The results are shown in Figure 4, in which we can see that more than 85% of the courses have included this type of specific activity.

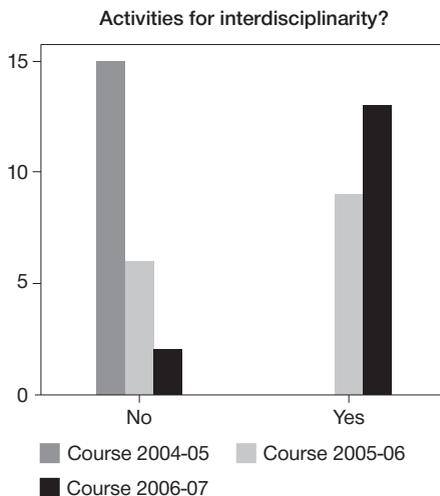
Figure 3. Number of courses with specific activities to foster multilingualism and number of languages used by each student on average in written and oral comprehension.



5. Conclusions

The preliminary results show some interesting headway from the project to improve the teaching quality for the NLP Master's programme. First of all, we can see that the faculty has significantly bolstered the amount of intra- and inter-course teaching materials and has generally included teaching activities involving problem-solving, interdisciplinary group work and students' functional multilingualism. By the end of the academic year, the professors who had included these aspects into their teaching practice believed that when they finished the programme the students were able to read litera-

Figure 4. Number of courses with specific activities to foster students' interdisciplinarity.



ture in at least four different languages, make oral presentations in public in languages other than their own, and help to pose and solve practical problems with people from fields quite different to their own.

Just as in the preceding years, at the end of the academic year we administered a student satisfaction survey. The most highly rated items (all over 4.1/5) after the availability of the faculty, were, in order, the items on: a) working in interdisciplinary and international teams; b) how the professors managed the multilingualism; c) the teaching materials; and d) the use of a specific syllabus. These results were confirmed by the responses to a qualitative question on the survey, namely «What did you like the best about the NLP Master's?» The aspects that students mentioned the most often were, in order of importance, «the new way of teaching» (79% of the students), the availability of the faculty (58%), the linguistic diversity and overcoming the initial language challenges (43%), the multidisciplinary approach (42%) and the availability of teaching materials for everyone (29%), in addition to other more minority responses.

These results show that if the entire team of professors includes in their teaching transversal training activities specifically designed to manage multilingualism and the heterogeneity of students' educational backgrounds by collaborative peer assignments, methodological tools and virtual resources, the difficulties observed initially can be considerably overcome, and interaction and the personal and educational enrichment of the students can be attained, which is shown in a high degree of satisfaction. However, we should also stress that this requires the involvement not just of the team of professors but also the group of students.

References

- BAQUÉ, L.; LE BESNERAIS, M. and MASPERI, M. (2003). «Entraînement à la compréhension orale des langues romanes: quelques repères pour faciliter la prise en compte de la matière phonique», *LIDIL*, 28, 137-152.
- TEN THIJE, J. D.; ZEEVAERT, L. (2007). *Receptive multilingualism: linguistic analyses, language policies and didactic concepts*. Amsterdam: John Benjamins Publishing,
- DEGACHE, Christian (2003). «Intercompréhension en langues romanes. Du développement des compétences de compréhension aux interactions plurilingues, de Galatea à Galanet», *LIDIL*, 28, 95-108.

Interesting links

- <http://www.galanet.eu> [1-3-2003]
- <http://www.innovationstrategy.gc.ca/gol/innovation/site.nsf/fr/in02346.html> [1-3-2003]
- <http://www.eurocomcenter.com/index2.php?lang=fr> [1-3-2003]

Keywords

Managing diversity, multilingualism, plurilingualism, intercomprehension

Financing

This project was financed by the AGAUR programme on improving the teaching quality at Catalan Universities (MQD) for 2006 (official state gazette no. 4558 dated the 25th of January 2006).

Supplementary materials on the CD-ROM

Demonstration of the GALANET website: virtual tour of the phases in the process of publishing a press file.

Project leader

Lorraine Baqué
Departament French and Romances Language and Literatures
Faculty of Philosophy and Arts
Universitat Autònoma de Barcelona
lorraine.baque@uab.cat

Presentation of the working group

The authors of the project belong to the laboratory on Phonetics, Lexicology and Semantics (*fLexSem*) of the Department of French and Romance Languages and Literatures at the Universitat Autònoma de Barcelona. Each one in their speciality (prosody and lexicography) has been interested for some time in the new teaching-learning approaches for foreign languages, and more specifically the intercomprehension of Romance languages.

Members of the project

Xavier Blanco

Department of French and Romance Languages and Literatures

Faculty of Philosophy and Arts

Universitat Autònoma de Barcelona

xavier.blanco@uab.cat

Àngels Catena

Department of French and Romance Languages and Literatures

Faculty of Philosophy and Arts

Universitat Autònoma de Barcelona

angels.catena@uab.cat

Marta Estrada

Department of French and Romance Languages and Literatures

Faculty of Philosophy and Arts

Universitat Autònoma de Barcelona

marta.estrada@uab.cat

Dolors Català

Department of French and Romance Languages and Literatures

Faculty of Philosophy and Arts

Universitat Autònoma de Barcelona

dolors.catala@uab.cat

Jorge Manuel Baptista

Department of Classical and Modern Arts

Universidade do Algarve (Portugal)

jbaptista@ualg.pt

Krzysztof Bogacki

Department of Linguistics

Institute of Romance Languages and Literatures

Uniwersytet Warszawski (Poland)

kbogacki@gmail.com

Max Silberztein

UFR de Sciences du Langage

Université de Franche-Comté (France)

max.silberztein@univ-fcomte.fr

