Contents | Author index | Subject index | Search | Home

The evolution of recent research on Catalan literature through the production of PhD theses: a bibliometric and social network analysis

Jordi Ardanuy and Cristóbal Urbano
Departament de Biblioteconomia i Documentació. Universitat de Barcelona. Edifici UB-Sants. C/ Melcior de Palau, 140, 08014 Barcelona, Catalonia, Spain.

Lluís Quintana
Departament de Filologia Catalana. Universitat Autònoma de Barcelona. Edifici B. Campus de la UAB. 08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain. .

Abstract

Introduction. This paper studies the situation of research on Catalan literature between 1976 and 2003 by carrying out a bibliometric and social network analysis of PhD theses defended in Spain. It has a dual aim: to present interesting results for the discipline and to demonstrate the methodological efficacy of scientometric tools in the humanities, a field in which they are often neglected due to the difficulty of gathering data.
Method. The analysis was performed on 151 records obtained from the TESEO database of PhD theses. The quantitative estimates include the use of the UCINET and Pajek software packages. Authority control was performed on the records.
Analysis. Descriptive statistics were used to describe the sample and the distribution of responses to each question. Sex differences on key questions were analysed using the Chi-squared test.
Results. The value of the figures obtained is demonstrated. The information obtained on the topic and the periods studied in the theses, and on the actors involved (doctoral students, thesis supervisors and members of defence committees), provide important insights into the mechanisms of humanities disciplines. The main research tendencies of Catalan literature are identified. It is observed that the composition of members of the thesis defence committees follows Lotka's Law.
Conclusions. Bibliometric analysis and social network analysis may be especially useful in the humanities and in other fields which are lacking in scientometric data in comparison with the experimental sciences.

CHANGE FONT

Introduction

The PhD thesis is one of the most important expressions of research in the humanities, and provides a fairly good representation of the personal interests of researchers. It also represents the culmination of the education of students, in which they are asked to present an original research project and demonstrate their research skills to a thesis defence committee composed of university lecturers, academics and specialists. PhD theses are therefore a good reflection of the predominant lines of work and research at universities and other academic institutions. Furthermore, as PhD candidates must have a supervisor and obtain the approval of the academic community represented by the thesis defence committee, PhD theses are a good source of information on the social structure of research (López Yepes 2002; Delgado *et al.* 2006). This aspect is especially interesting in some areas of the social sciences and in the humanities which, unlike the experimental sciences, involve practically no collaborative work (Hagstrom 1965; Bourke 1997; Molteni and Zulueta 2002; Ardanuy 2008).

In order to study the connections between the academic actors involved in PhD theses, we decided to use social network analysis. This is a set of methodological tools and techniques aimed at identifying social structures that appear when individuals or organisations interact. By analysing the way in which individuals relate, we can determine the structure of the network they form, and within it the groups and actors that hold privileged positions. Social network analysis has been used in many fields, including academic communication. In this context the relations established between the actors within the networks provide us with information on the importance and influence of academicians and the institutions to which they are affiliated; on the scientific academic collaboration and affinity between the researchers; on the coincidence and degree of regional and thematic ubiquity of academicians; and on the direct channels of transmission of academic ideas, of ways of knowing and doing (Newman 2001; Otte and Rousseau 2002; Watts 2004; Delgado *et al.* 2006).

The first bibliometric work using social network analysis was undertaken at the Institute for Scientific Information under the leadership of Henry Small, and was devoted primarily to building scientific maps from the co-citations contained in the documents that they indexed. However, the development of information technologies has facilitated the appearance of independent studies covering a variety of topics: scientific maps of a discipline (co-citation analysis) (McCain 1990); patent analysis (co-citation) (Leydesdorff 1995); scientific productivity (co-authorship) (Eaton *et al.* 1999; Otte and Rousseau 2002); the structure of the networks of scientific collaboration (co-authorship) (Newman 2001; Molina, Muñoz and Domènech 2002); comparisons between very close academic fields (co-citation) (Vivas and Urquijo 1999; Kreuzman 2001); journal maps (co-citation) (Leydesdorff 2004); maps of

disciplines of a country (the discipline or disciplines of each article) (Moya *et al.* 2006); and PhD thesis defence committees (relation and co-presence) (Delgado *et al.* 2006).

The research presented in this paper focuses on the actors and institutions involved in the academic study of Catalan literature. and it has a dual aim: to present interesting results for the discipline and to demonstrate the methodological efficacy of scientometric tools in the humanities, a field in which they are often neglected due to the difficulty of gathering data.

Catalan literature is that which is written in Catalan, a Romance language belonging to the Western Romance group that is proper to the Catalan-speaking countries (the traditional name for the Catalan-speaking territories of France, Spain and Italy, in addition to Andorra, where it is the only official language). As most of these territories are currently under Spanish sovereignty, the vast majority of the literary production and academic production on Catalan literature is carried out in Spain. Research on Catalan literature was subject to the highly negative effect of the political and social events that occurred in Spain for several decades in the mid-20th century. Therefore, only in recent years, with the institutionalisation of Catalan philology chairs at Spanish universities, has this field of study begun to be professionalised. A significant stock of knowledge has now been published, but little research has been done on it (Molas, 2001; Duran 2004; Ardanuy, Quintana and Urbano 2008).

Method

The main source of the data used in this paper was the TESEO database, which contains information on PhD theses defended and approved in Spanish universities since 1976. The theses defended in North Catalonia (France) were not included. The TESEO system, which depends on the Spanish University Coordination Council, has been accessible online since 1997, and is now accessible through the website of the Spanish Ministry of Science and Innovation. The database allows searches by author, title, abstract, supervisor, university, place of defence, academic year and descriptor.

In the absence of a category for Catalan literature, or for Catalan philology in general, we recovered the relevant documents by using a search strategy equivalent to Table 1.

In the 'Descriptor' field	(HISTORIA DE LA LITERATURA ESPAÑOLA) OR (HISTORIA DE LA LITERATURA) OR (LENGUA Y LITERATURA) OR (ANÁLISIS LITERARIO))
	OR
In the 'Title/Abstract' field	((LITERATURA CATALANA) OR (CATAL$) OR (LITERATURA VALENCIANA) OR (VALENCI$))
	($ is the symbol for truncation in this database)

Table 1: Search equation

Most of the documents were retrieved with the descriptors HISTORIA DE LA LITERATURA ESPAÑOLA, HISTORIA DE LA LITERATURA and the more general LENGUA Y LITERATURA. The search was completed with records of theses that contained in either their title or their abstract a reference to Catalan literature, or even to anything related to Catalan or the Catalan situation. This procedure generated a considerable number of irrelevant results, but there was no other way to draw up an exhaustive list of theses, an essential requirement to ensure our study has statistical value. The search was also extended to the Valencian language, the name currently given to the Catalan spoken in the Valencian community. When the irrelevant records had been eliminated, 151 remained, of which 90.1% corresponded to theses written in Catalan. All those dealing even partially with Catalan literature were included.

It should be noted that the quality of the TESEO database varies greatly, because it essentially depends on the rigour applied by the doctoral students, committees and universities when supplying the data. The limitations include lack of exhaustiveness, slow updating (which is why we set the end date of our study in 2003), incomplete records (especially with regard to the thesis supervisor) and a lack of authority control (Delgado *et al.* 2006).

In order to compensate for the lack of information in some fields, we consulted the catalogues of university libraries, and in four cases made personal queries by e-mail. Some theses on Catalan literature defended during the period were found to be absent from the database, but for reasons of internal coherence we were unable to use any data that did not come directly from TESEO. Another problem was that as there was no category for Catalan literature, it was impossible to be sure that a totally exhaustive retrieval of the relevant documents had been made. This problem is shared by literature in other languages.

The records of documents obtained from TESEO were exported to a bibliographic manager to perform authority control and from there to a relational database manager with which the one-dimensional statistical data were processed and represented graphically. For the network analysis, two types of network were used to study the relationships between the actors involved in the research. In the case of the members who co-participated in PhD thesis committees, the network used was Mode 1, in which all the elements formed part of a single group of protagonists. Mode 2, which corresponded to two differentiated sets, was used for the relationship between thesis supervisors and selected committee members. The numerical data of these networks were processed with UCINET to obtain several indicators (Degree, Closeness, Betweenness and weak transitivity relations) and later exported to the Pajek software for the graphic representations.

For the affiliation and main speciality of the committee members, a large number of directories and databases (mainly from the Catalan-speaking countries and Spain) were used (Ardanuy 2008: 183). In the results and discussion, in addition to the text the distributions of values are presented in tables to facilitate reading and interpretation.

Results and discussion

Chronological and gender distribution

The extraction of data on theses from TESEO on which our study is based covers the period 1976 to 2003. Figure 1 shows the chronological evolution of production, showing clearly that, after a rather hesitant beginning, there was a sharp increase in the defence of theses from 1984, with a peak in 1992. From then on, though there have been annual variations, the number of theses defended has fallen, as shown in the polynomial trend curve. This phenomenon seems particular to the field and is not found in other studies on PhD theses in Spain (Civera Molla and Tortosa Gil 2001; Lázaro and Torres 2002; Fernández Cano *et al.* 2003; Delgado *et al.* 2006) or in the totals for theses in TESEO, as can be seen in the trend graph of Figure 1.
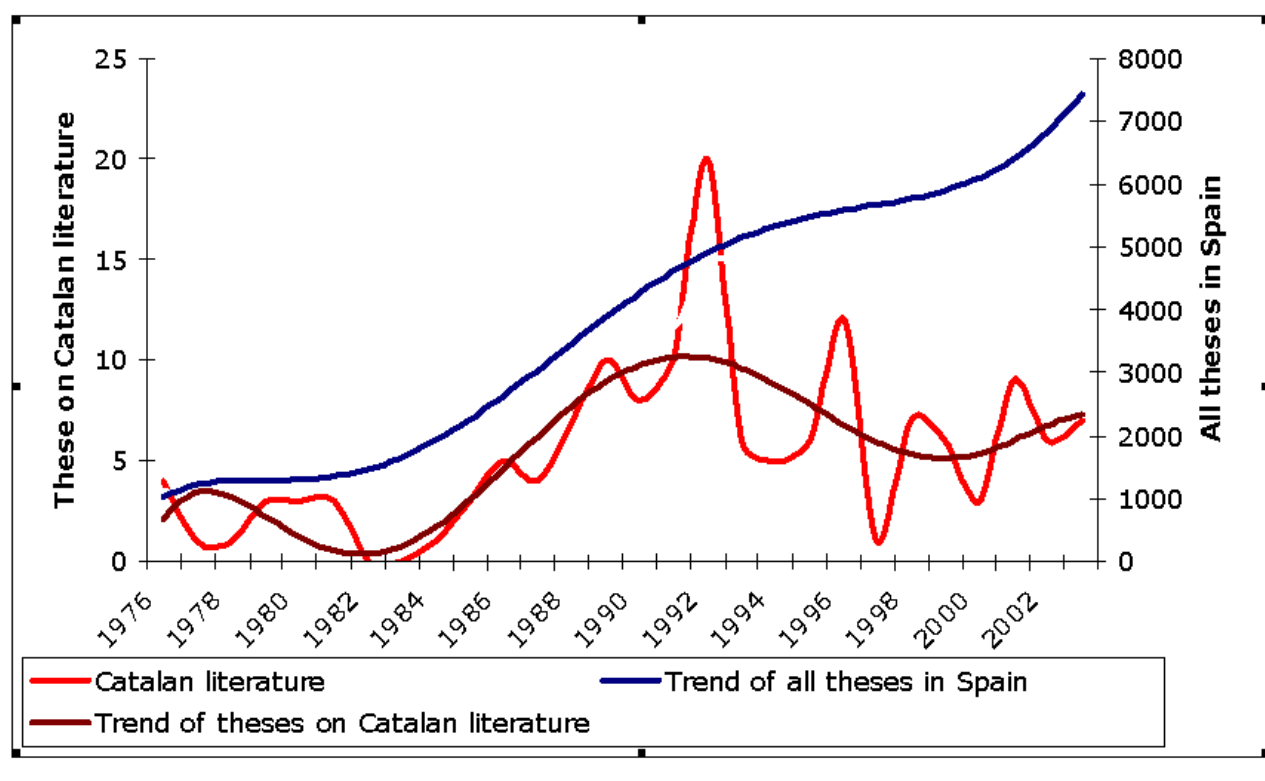
Figure 1: Chronological evolution of the theses on Catalan literature defended in Spain and all the theses contained in TESEO

The registration of new students in Catalan Philology began to fall in the late 1990s (Figure 2, right axis), and the number of university graduates in this discipline began to fall in 2000 (Figure 2, left axis). However, this decline in students fails to explain the fall in the number of theses defended, which was earlier and far more sudden.
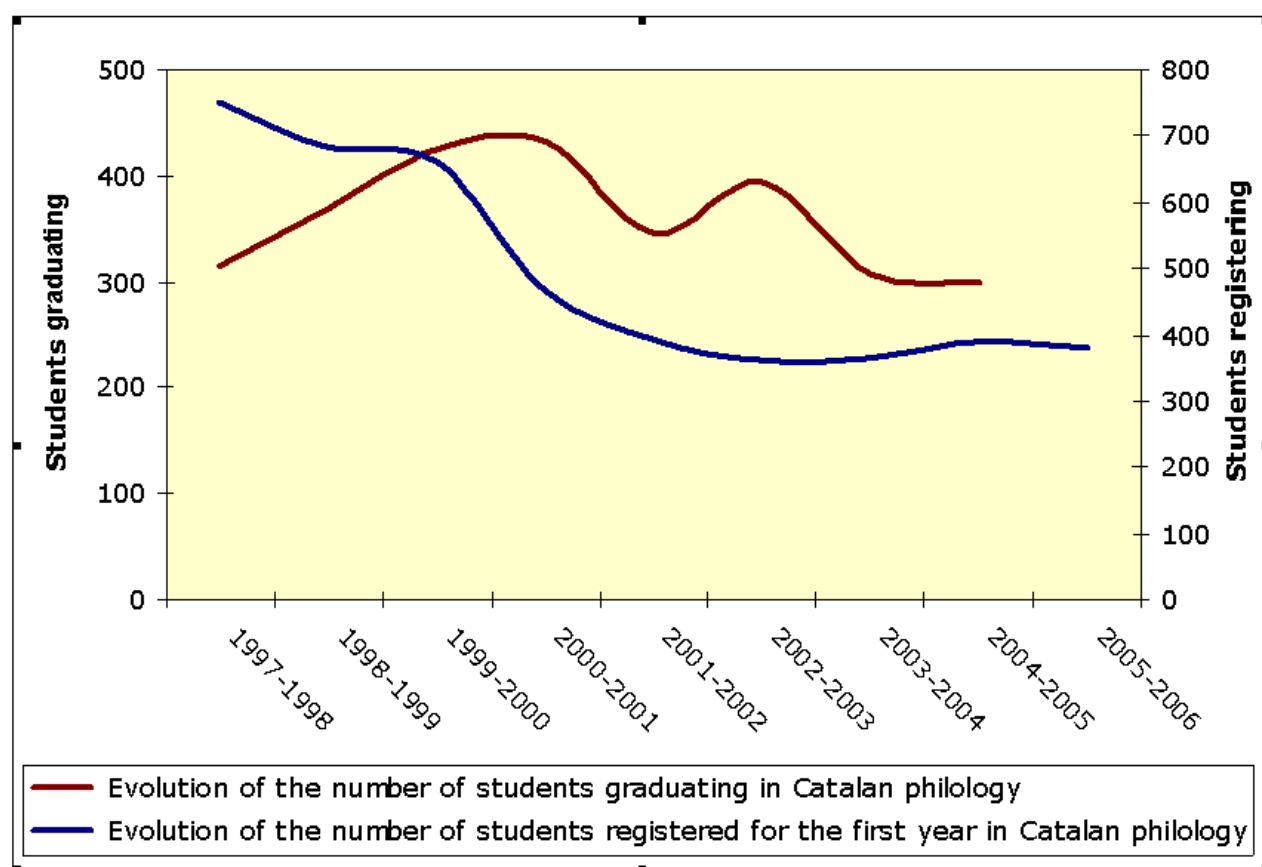
Figure 2: Evolution of the number of students graduating in Catalan philology and first-year students registering at Catalan universities, the Universitat de les Illes Balears and the Universitat de València Source: Commission for Universities and Research of the Generalitat de Catalunya (Government of Catalonia), Statistics Institute of the Balearic Islands and Universitat de Valencia.

Figure 3 shows how the phenomenon affected the Universitat de Barcelona (UB), the Universitat Autònoma de Barcelona (UAB), the Universitat de València (UV) and the Universitat de les Illes Balears (UIB). On the other hand, at the universities of Lleida (UdL) and Alacant (UA), created in the 1990s, the peak was in 1996. The explanation is to be found in the process of obtaining tenured teaching posts in Spain, for which a PhD is a requisite. A large proportion of the teaching staff of universities in the Catalan-speaking countries obtained their posts in the early 1990s, whereas in more recent universities the process took place a few years later. According to Molas (2001), this process has greatly limited the prospects of promotion for new researchers, because of the limits on numbers of university teaching staff, and the result has been a fall in the number of theses presented.
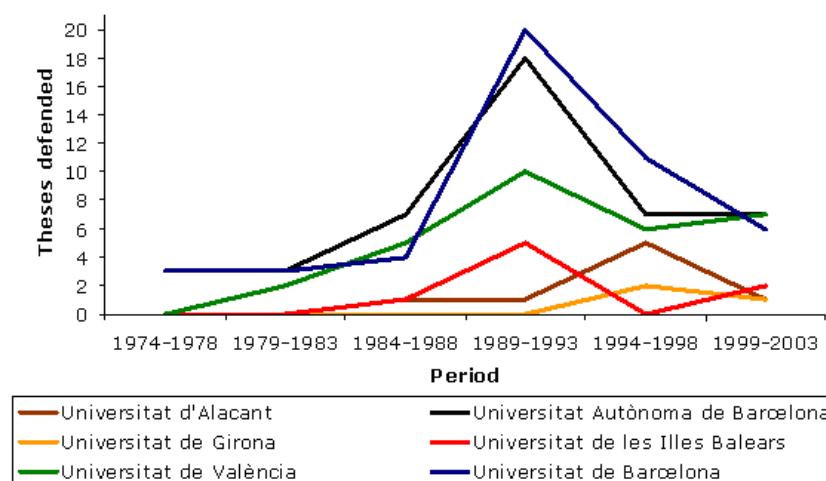

Figure 3: Chronological evolution of thesis defences by universities

Of the total of 151 theses extracted from TESEO, 96 were written by men (63.6%) and 55 by women (36.4%). Figure 4 shows the evolution of the

number of theses defended according to the sex of the author. The sharp peak shown around 1992 is associated with thesis defences by men in a discipline in which women have traditionally been a majority. We only have figures for the number of students of each sex graduating in Catalan philology from Catalan universities and the Universitat de les Illes Balears between 1998 and 2005 (Figure 5), which show an overwhelming predominance of women: never below 70% of the total number. Though the number of women graduates was three times higher than that of men in this period, men continued to present more PhD theses.
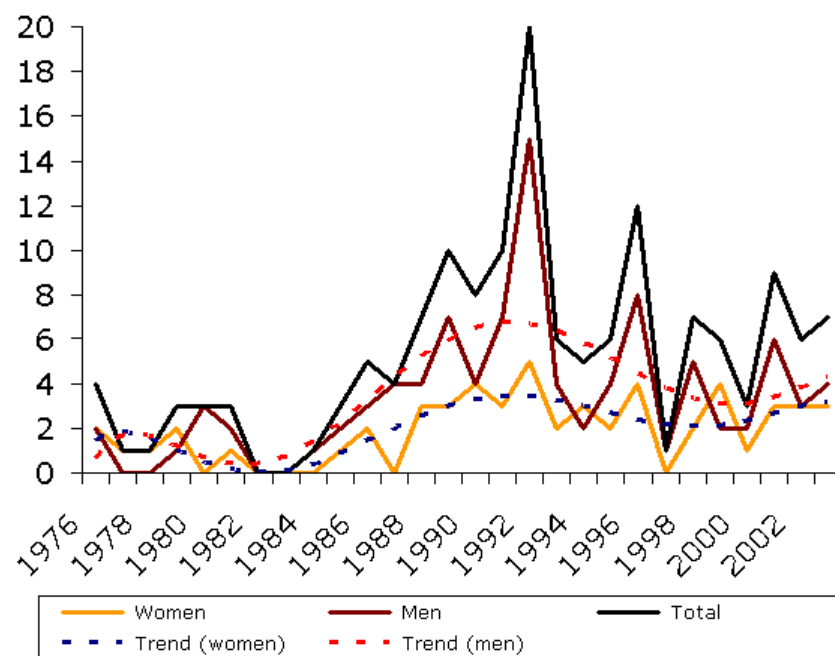

Figure 4: Evolution of the number of PhD defences according to the sex of the author
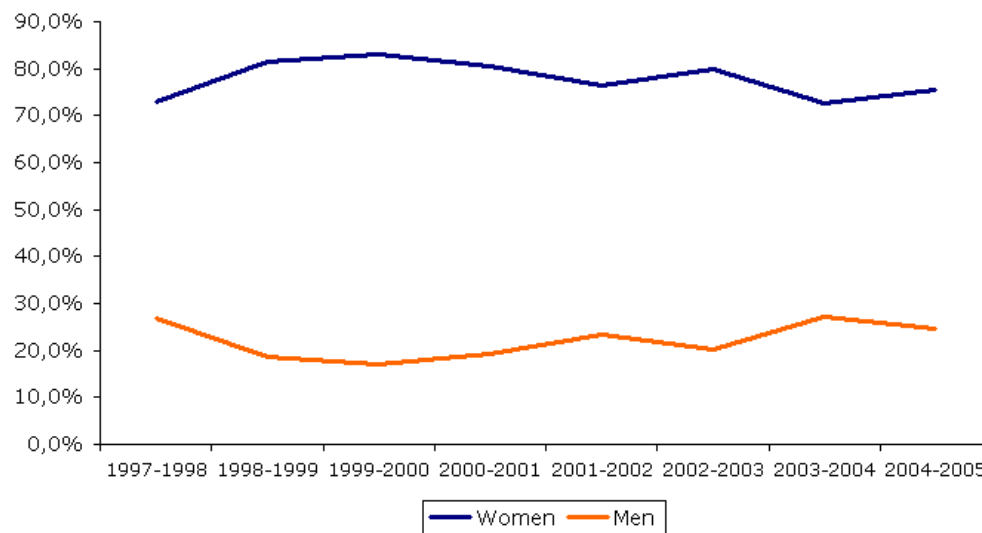

Figure 5: Percentage distribution of graduates according to sex

Although some PhD theses may have been produced in other areas of Catalan philology, or even the humanities, the reason why women have not produced a number of theses proportional to their weight in the number of students is probably to be found in the traditional socio-cultural stimulus for men to occupy academic posts in Catalan-speaking countries. This explains the greater dedication of men to producing PhD theses, and thereby overcoming the first hurdle in the academic hierarchy. As the peak of production in the early 1990s is associated with a period of creation of teaching posts in Catalan philology and the theses were defended mainly by men, it can be concluded that men were the main beneficiaries of the situation.

In order to study the evolution in more detail, the study period was divided into three periods of equal duration:

- 1976-1983 included 15 theses, representing 9.9% of the total;
- 1984-1993 included 74 theses, representing 49.0% of the total; and
- 1994 -2003 included 62 theses, representing 41% of the total

Universities involved

The theses detected in this study were defended at a total of thirteen different universities (Table 2). However, three universities represented 80.8% of the total: the Universitat de Barcelona (31.1%), the Universitat Autònoma de Barcelona (29.8%) and the Universitat de València (19.9%). A long way behind were the Universitat de les Illes Balears and the Universitat d'Alacant (5.3% each), and of the others only the Universitat de Girona had more than one thesis (2.0%). The four universities of other parts of Spain had only a token presence (less than 5%).

Territory	University	Abbreviation	No. of theses	% of theses	Accumulated % of theses
Catalan-speaking countries	Universitat de Barcelona	UB	47	31.1	31.1
	Universitat Autònoma de Barcelona	UAB	45	29.8	60.9
	Universitat de València	UV	30	19.9	80.8
	Universitat d'Alacant	UA	8	5.3	86.1
	Universitat de les Illes Balears	UIB	8	5.3	91.4
	Universitat de Girona	UdG	3	2.0	95.4
	Universitat Jaume I	UJI	1	0.7	98.0
	Universitat de Lleida	UdL	1	0.7	97.4
	Universitat Rovira i Virgili	URV	1	0.7	99.3
Rest of the Spanish State	Universidad Complutense	UCM	3	2.0	93.4
	Universidad Nacional de Educación a Distancia	UNED	2	1.3	96.7
	Universidad de Murcia	UM	1	0.7	98.7
	Universidade de Santiago de Compostela	USC	1	0.7	100

Table 2: Number and percentage of theses defended at Spanish universities

Figure 3 shows the chronological evolution of the six universities at which most theses were defended. The four universities that already existed in the 1980s (Universitat de Barcelona, Universitat Autònoma de Barcelona, Universitat de València and Universitat de les Illes Balears) showed a fairly similar behaviour, despite the difference in absolute terms in the volume of theses defended, which showed a peak in the period 1989-1993. The decline was so sharp at the Universitat de les Illes Balears that between 1992 and 1999 there were no records in TESEO of theses on Catalan literature at this university. The Universitat d'Alacant showed a small peak between 1994 and 1998, with five of its eight theses being defended in this period, and the three theses of the far more recent Universitat de Girona were defended in the second half of the 1990s.

Between 1974 and 1983, a period of political transition and recovery of freedom in Spain after the Franco regime, one thesis was defended at the Universidad de Murcia, six each at the Universitat de Barcelona and Universitat Autònoma de Barcelona, and two at the Universitat de València. Between 1984 and 1993, a period characterised by a consolidation of democratic freedom and the expansion of the university system, the number of theses defended rose to seventy-four. The Universitat Autònoma de Barcelona and the Universitat de Barcelona were the most active, with twenty-five and twenty-four theses, respectively, followed by the Universitat de València, with fifteen. The Universitat de les Illes Balears and the Universitat d'Alacant made their first contributions to the discipline, with six and two theses, respectively. One thesis was defended at the Universidad Nacional de Educación a Distancia and one at the Universidad Complutense of Madrid.

Finally, between 1994 and 2003 there was a decrease in the number of theses defended. The number of theses of the Universitat de Barcelona and the Universitat Autònoma de Barcelona fell to seventeen and fourteen, respectively, whereas the recently independent universities of Lleida, Tarragona and Girona (previously outposts of the larger universities) produced their first theses (three, one, and one, respectively), and the newly-created Universitat Jaume I of Castelló produced one. The production of the Universitat d'Alacant rose to six, whereas that of the Universitat de les Illes Balears fell to three. Outside the Catalan-speaking countries four theses were defended: two at the Universidad Complutense, one at the Universidad Nacional de Educación a Distancia and one at the Universidade de Santiago de Compostela, representing 6% of the total of the period and showing the lack of interest in Catalan literature in Spain outside the Catalan-speaking countries.

A map of the discipline and its trends: literary periods studied and regions, authors and works analysed

The study of Catalan literature and its history tends to be divided into four main areas (Molas 2001, Duran 2004): literary theory and comparative literature; medieval literature (up to the 15th century); modern literature (16th, 17th and 18th centuries and first third of the 19th century); and contemporary literature (from the second third of the 19th century to the present). When applied to the theses studied, this classification shows a clear predominance of contemporary literature (59%), followed by medieval (21%) and modern literature (15%). This distribution is in agreement with the results we obtained in another study based on the analysis of citations from journals of Catalan literature (Ardanuy *et al.* 2008), but the theses showed slightly more studies of modern literature, slightly fewer studies of contemporary literature, and the same proportion of studies of medieval literature. In combination with the chronological figures provided in Figure 6, these results seem to indicate that research on Catalan literature of the modern period is slowly growing. They are also in agreement with the more qualitative results offered by the *Reports de la recerca a Catalunya* (Reports on research in Catalonia) of the Institut d'Estudis Catalans (Institute for Catalan Studies) (Molas 2001, Duran 2004). However, the thematic distribution of theses is not homogeneous between the main universities. Whereas the Universitat de Barcelona and Universitat Autònoma de Barcelona concentrate overwhelmingly on contemporary literature, this is not the case at the Universitat de València, and at the Universitat de Girona this subject is surpassed by medieval literature (Figure 7).
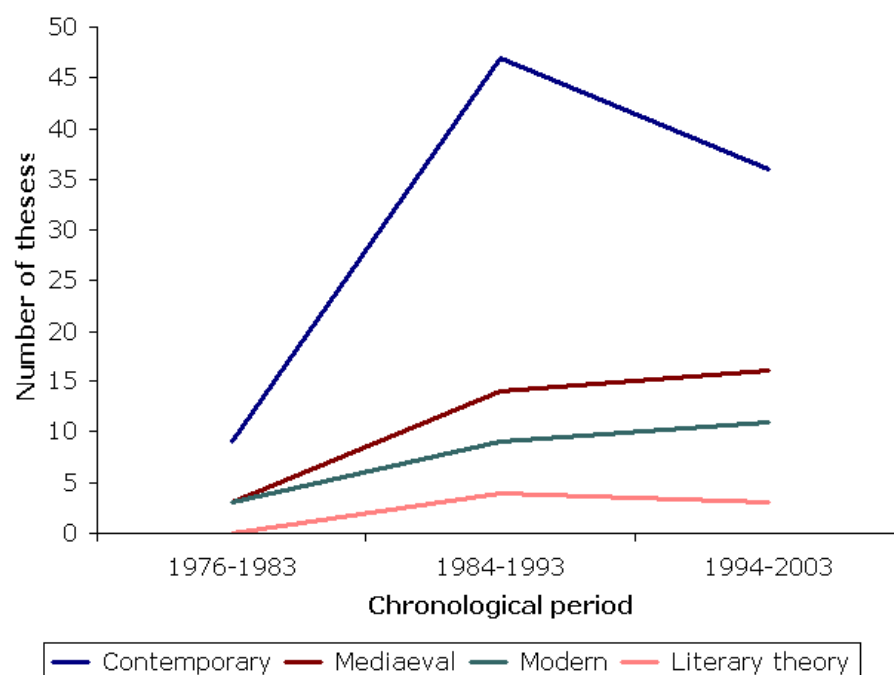
Figure 6: Chronological evolution of the distribution of theses according to the four main areas of study of Catalan literature

Figure 7: Number of theses defended in each area of study at the 5 universities with most theses on Catalan literature.

The theses dealing with a specific literary genre, normally in a specific geographic area and period in time, show a predominance of narrative, and in particular the novel (Table 3).

Genre		No. of theses
Narrative	Total	34
	General narrative	5
	Novel	28
	Short narrative	1

Poetry	12
Prose (*)	4
Theatre	10
(*) In the prose group we classify the works corresponding to writing that does not use verse but cannot be classified as narrative, such as essays and descriptions	

Table 3: Number of theses dealing with specific literary genres

Of the theses dealing with specific geographic areas (see Table 4), the highest number were on the Valencian Community (thirteen), followed by the Balearic Islands (six), whereas Catalonia only had three, in agreement with the results of the citation analysis (Ardanuy *et al.* 2008). These data confirm that the literature produced in Catalonia has been hegemonic, like French literature in France or German literature in Germany, and may explain why most of the theses on specific areas deal with literature from outside Catalonia

Region		No. of theses
Valencian Community	Total	13
	Valencian Community in general	8
	Castelló de la Plana	2
	City of Alacant	1
	Counties of Castelló	1
	City of Valencia	1
Balearic Islands	Total	6
	Balearic Islands in general	3
	Island of Mallorca	2
	Palma de Mallorca	1
Catalonia	Total	3
	Catalonia in general	1
	City of barcelona	2

Table 4: Number of theses dealing with specific geographic areas

Two theses each dealt with the medieval works *Tirant Lo Blanc* by Joanot Martorell and *L'Espill* by Jaume Roig, whereas the other works analysed were only dealt with in one thesis. Six theses dealt with author Josep Carner, and five with author Jacint Verdaguer.

Thesis supervisors

Thesis supervisors are essential to the success of the PhD projects presented to a thesis defence committee. They are normally outstanding figures in the academic community of the university at which the PhD thesis is presented, and show a high level of knowledge on the research subject. They enjoy academic influence based on their activity as researchers, the management of research groups and networks, their personal relationships with other academics and their presence in the academic media (journals, congresses, publishers, portals, etc.). In short, thesis supervision involves academic status.

In our period of study, the most outstanding figure was Joaquim Molas (Universitat de Barcelona), who supervised 34 theses (22.4% of the total), followed at some distance by Jordi Castellanos (Universitat Autònoma de Barcelona) with 10 theses, and Lola Badia (Universitat de Barcelona) and Antoni Comas (Universitat de Barcelona) with 6 each. In line with their main specialities, Joaquim Molas and Jordi Castellanos supervised theses on contemporary literature, Lola Badia supervised theses on medieval literature, and Antoni Comas supervised theses on modern and contemporary literature.

The study of thesis defence committee membership involves certain problems due to the legal changes introduced in the period of our study. One of the changes affecting the study is the fact that during certain periods it was possible for the thesis supervisor to also be a committee member. We excluded such cases from our calculations.

Another interesting topic is the genealogy created when the author of a thesis becomes, in turn a thesis supervisor, thus extending the influence of the senior academic. In our study, Joaquim Molas only supervised the thesis of one PhD student who went on to become a thesis supervisor: Jordi Castellanos. But Castellanos supervised ten theses (Figure 8). Therefore, in addition to supervising the greatest number of theses (thirty-four), Molas extended his influence through Castellanos to ten more researchers. The genealogies stemming from Molas-Castellanos represent approximately 30% of the whole production of theses in Catalan literature.

As can be seen in Figure 8, the longest genealogies are of three generations (four elements), two of Antoni Comas continued by Joan Alegret, one of Martí de Riquer continued by Lola Badia, and one of Joaquim Molas continued by Jordi Castellanos. It is striking that despite the large number of theses he supervised, Joaquim Molas only formed genealogies through Jordi Castellanos.
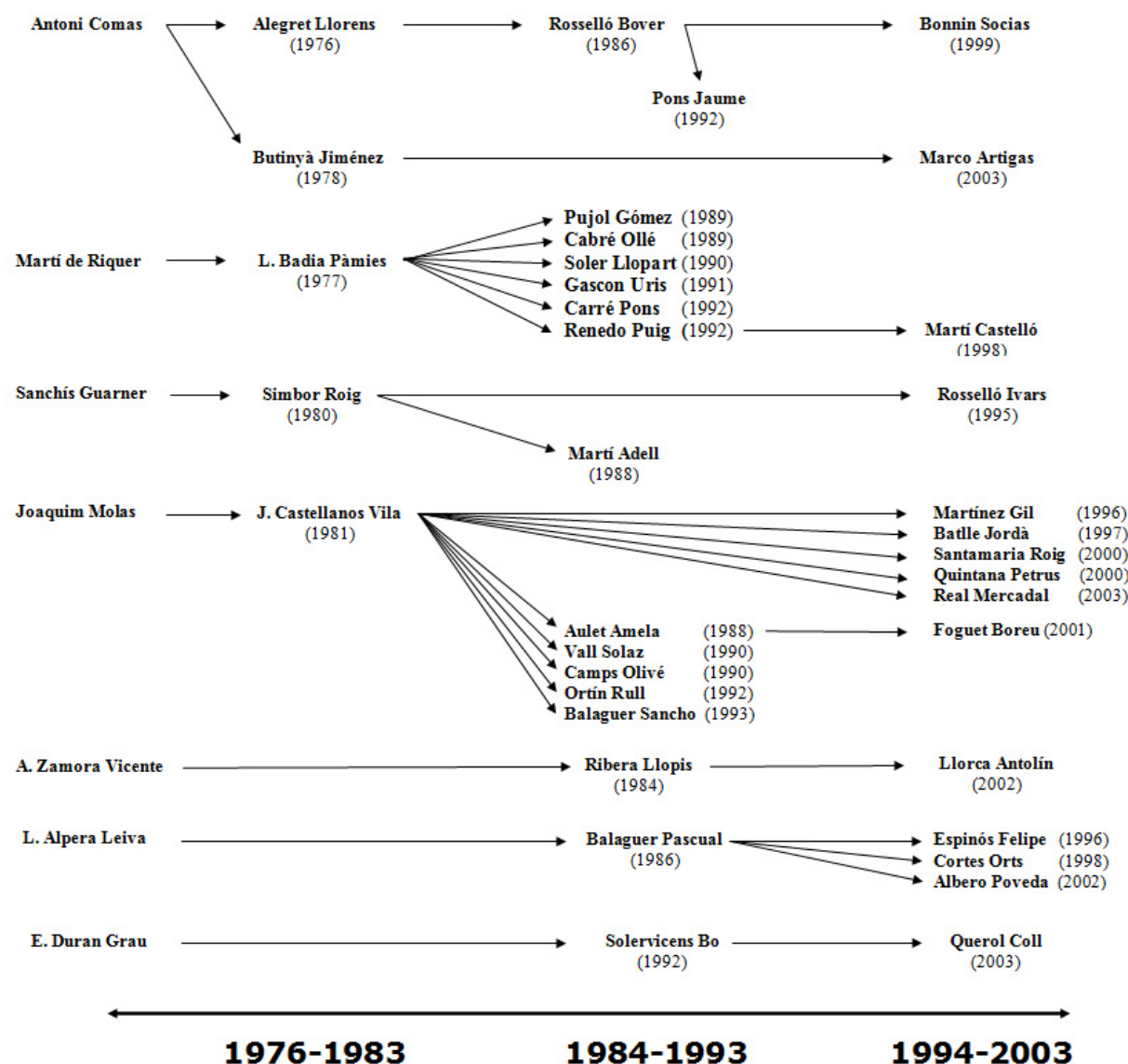
Figure 8: Genealogies generated from PhD theses.

Committee members

Participations

The total number of committee memberships was 743, fewer than the expected 755 (151 * 5) because in some cases they were not recorded in TESEO. These memberships corresponded to 265 different academics. Table 5 shows the academics who were members of most committees. Jordi Castellanos heads the ranking as a member of twenty-five committees, followed by Albert Hauf as a member of twenty.

Academic member	Total no. of memberships	Accumulated % of membership
Castellanos Vila, J.	25	3.4
Hauf Valls, A. G.	20	6.1
Pla Arxé, R.	19	8.6
Molas Batllori, J.	18	11.0
Blecua Perdices, A.	17	13.3
Mas Vives, J.	17	15.6
Gallén Miret, E.	16	17.8
Ferrando Francés, A.	15	19.8
Jorba Jorba, M.	15	21.8
10 academics between 8 and 14	110	36.6
5 academics with 7	35	41.3

4 academics with 6	24	44.5
9 academics with 5	45	50.6
17 academics with 4	68	59.8
27 academics with 3	54	70.7
34 academics with 2	68	79.8
150 academics with 1	150	100

Table 5: Academics with most participation on PhD thesis committees

The distribution of the number of committee members in relation to all participations on thesis committee follows a power law $f(k) = C/k^{\beta}$. Using the LOTKA program presented in the article by Rousseau and Rousseau (2000), we find that the value of C is 0.5981, whereas that of the exponent ß is 1.9718, which closely fits Lotka's numerical expression (1926).

Academic specialities

With regard to the main academic specialities of the committee members (Table 6 and Figure 9), Catalan literature represented half (50.7%), whereas the rest were from other literatures, Catalan language, history, literary theory and others. The percentage of committee members specialising in Catalan literature increased over time, from the initial 19.5% when specialists were lacking, to 55.7% in the period 1994-2003. There was a parallel fall in the proportion of committee members from Spanish literature and other specialities, with the exception of literary theory and comparative literature. As expected, the participation of specialists in Catalan literature has increased with their increasing availability.

Speciality	No. of specialists	% of specialists	Accumulated % of specialists
Catalan literature	377	50.7	50.7
Spanish literature	99	13.3	64.1
Romance Philology	54	7.3	71.3
Catalan Language	48	6.5	77.8
Other literatures	45	6.1	83.8
Literary theory and comparative literature	37	5.0	88.8
Other disciplines (philosophy, art, anthropology, etc.)	34	4.6	93.4
History	33	4.4	97.8
Other philology specialities	16	2.2	100

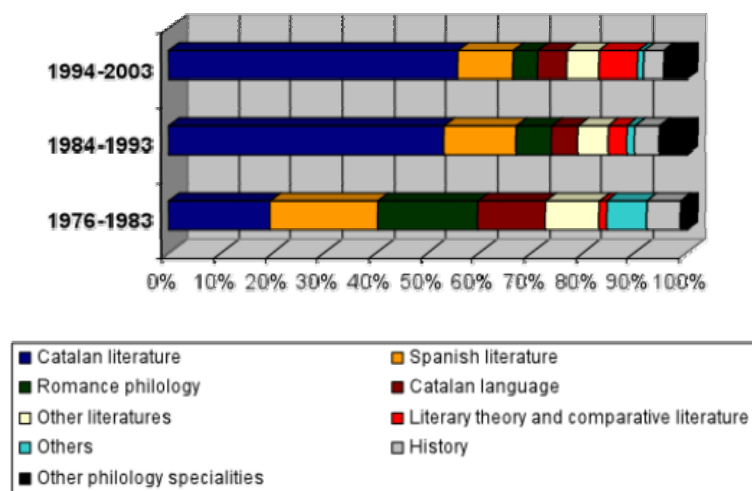**Table 6: Distribution of specialities of the committee members**

Figure 9: Percentage distribution of committee memberships by specialists in Catalan literature by periods.

Table 7 shows the number of memberships of specialists according to the main areas of Catalan literature. There is a clear predominance of contemporary literature, with 237 memberships (64%), which corresponds quite closely with the predominance of theses on contemporary literature.

Area of Catalan literature	Total no. of memberships	No. of memberships 1976-1983	No. of memberships 1984-1993	No. of memberships 1994-2003
Contemporary	237	10	115	112
Medieval	102	3	52	47
Modern	31	2	19	10

Table 7: Distribution of committee members in the three main areas of Catalan literature

Affiliation

The affiliation of the committee members shows a low level of internationalisation: in 90% of the cases the members came from the Catalan-speaking countries, in 7% from Spain and in only 3% from other countries. Table 8 shows the affiliation of the committee members in the five universities with the highest number of theses defended, (UA, UAB, UB, UIB and UV), and the total.

Affiliation of the committee member	University of thesis defence					
	UA	UAB	UB	UIB	UV	Total no. (*)
Universitat de Barcelona (UB)	7	54	124	4	19	216
Universitat Autònoma de Barcelona (UAB)	1	101	49	3	13	175
Universitat de València (UV)	7	10	11	4	65	102
Universitat Pompeu Fabra	0	23	12	1	2	41
Universitat de les Illes Balears (UIB)	3	3	9	18	4	38
Universitat de Girona	2	9	6	5	7	32
Universidad Complutense (Madrid)	3	5	1	0	1	18
Universitat d'Alacant (UA)	5	3	1	0	8	17
Universitat Rovira i Virgili	2	3	6	1	2	15
Universitat Jaume I	2	0	3	0	5	12
Universitat de Lleida	0	1	3	1	3	9
Other universities in the Catalan-speaking countries	0	2	2	0	3	7**
Other Spanish universities	3	3	2	1	6	34
British Universities	1	3	0	1	3	8
Other European universities	2	2	2	0	2	10
Universities of the USA	0	1	0	0	0	1
Spanish National Library	0	0	0	0	0	1
No data	1	1	2	0	2	7
Total	39	224	233	39	145	743
*Includes all the universities at which theses were defended. ** Includes the Universitat de Perpinyà (Via Domícia) in France.						

Table 8: Affiliation of committee members in the five universities with the highest number of theses defended

Our study also revealed a high proportion of committee members from the university at which the thesis was defended (Table 9). At the Universitat de Barcelona over half (53.2%) the committee members belonged to the same university; at the Universitat de les Illes Balears the figure was 46.2%, at the Universitat Autònoma de Barcelona 45.8%, at the Universitat de Valéncia 44.8%, and at the Universitat d'Alacant 12.8%.

The Universitat Autònoma de Barcelona and the Universitat de Barcelona showed comparable figures for receiving committee members from each other (24.1 and 21.0%, respectively). On the other hand, whereas the Universitat de Barcelona only received 4.5% of its committee members from the Universitat de València, the Universitat de València received 13.1% of its committee members from the Universitat de Barcelona.

The Universitat d'Alacant and the Universitat de les Illes Balears both received 17.9% of their committee members from the Universitat de València and 10.3% from the Universitat de Barcelona.

The regulations for forming thesis committees have changed over the years, though the number of members has remained stable at five. Between 1974 and 1983 the committee could be composed wholly of members of the same university. Between 1983 and 1985 the maximum number of committee members of the same university was four, whereas since 1985 it has been three. The university system has thus allowed high levels of participation by members of the same university as the PhD candidates. Of course, in the case of Catalan literature, this opportunity has been used quite frequently.

Affiliation of the committee member	University of thesis defence				
	UA	UAB	UB	UIB	UV
Universitat de Barcelona (UB)	17.9%	24.1%	53.2%	10.3%	13.1%
Universitat Autònoma de Barcelona(UAB)	2.6%	45.1%	21.0%	7.7%	9.0%
Universitat de València (UV)	17.9%	4.5%	4.7%	10.3%	44.8%
Universitat Pompeu Fabra	0.0%	10.3%	5.2%	2.6%	1.4%
Universitat de les Illes Balears (UIB)	7.7%	1.3%	3.9%	46.2%	2.8%
Universitat de Girona	5.1%	4.0%	2.6%	12.8%	4.8%
Universidad Complutense	7.7%	2.2%	0.4%	0.0%	0.7%
Universitat d'Alacant (UA)	12.8%	1.3%	0.4%	0.0%	5.5%
Universitat Rovira i Virgili	5.1%	1.3%	2.6%	2.6%	1.4%
Universitat Jaume I	5.1%	0.0%	1.3%	0.0%	3.4%
Universitat de Lleida	0.0%	0.4%	1.3%	2.6%	2.1%
Others	15.5%	4.5%	2.0%	5.2%	9.8%
No data	2.6%	0.4%	0.9%	0.0%	1.4%
Total	100%	100%	100%	100%	100%

Table 9: Percentages of affiliation of committee members at the five universities with the highest number of theses defended

Co-participation networks

The links that are established between committee members as a result of their co-participation during the whole study period (a study by chronological periods is available in Catalan in the PhD thesis of Jordi Ardanuy (2008)) constitute a disconnected symmetric graph formed by 265 nodes. The nodes or vertices (committee members) not connected to the others correspond to theses defended with little relation to the central trend of general studies of Catalan literature. These theses were excluded from the calculations, finally leaving 248 nodes, However, the adjacency matrix formed by these nodes gives rise to a network that is too complex to allow a minimally intelligible graphical approximation. Figure 10 shows an abbreviated version in which the edges corresponding to fewer than two co-participants have been eliminated, together with the committee members who were thus isolated. The thickness of the edge joining the nodes represents the intensity with which the pairs of members co-participated on different committees. We used the representation algorithm proposed by Kamada and Kawai (1989). This figure shows the most important academics in terms of participation on committees, because they are related by more arcs (Castellanos, Hauf, Plan Arxé, Mas Vives, Blecua Perdices, Gallén, Ferrando Francés, Lola Badia, Molas, etc.) and also identifies the cases of considerable coincidence on committees, thanks to the thickness of the edges with which they are joined.
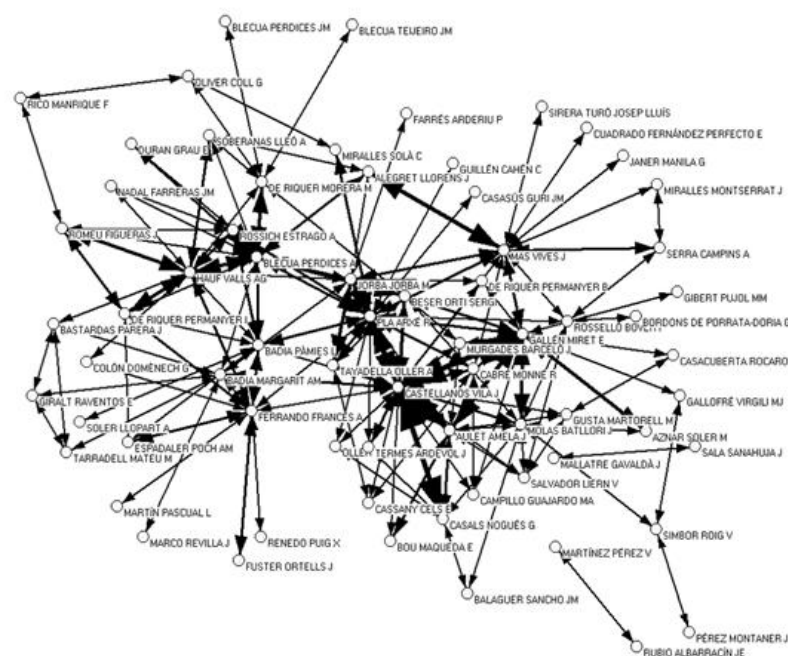

Figure 10: Network of co-participation relations on thesis committees between main nodes.

Most significant nodes

Social network analysis can be used to identify the nodes that occupy most significant roles. The degree indicator indicates the intensity of the direct union of an element with the others in the network, showing the local centrality of the person studied (Freeman 1977). Table 10 shows the academics with a high degree indicator, and the centrality share normalised to the sum total of degrees. Of course, the first places are occupied by the people with the highest number of thesis committee memberships.

Committee member	Degree indicator	Share	Number of committee memberships
Castellanos Vila, J.	100	0.036	25
Hauf Valls, A. G.	78	0.028	20
Pla Arxé, R.	76	0.028	19
Mas Vives, J.	67	0.024	17
Blecua Perdices, A	64	0.023	17
Gallén Miret, E.	63	0.023	16
Ferrando Francés, A.	59	0.021	15
Jorba Jorba, M.	59	0.021	15
Aulet Amela, J.	55	0.020	14
Molas Batllori, J.	54	0.020	14

Table 10: Degree of the committee members

Overall, the network shows an average centrality of 11.1, with an average deviation of 14.5. This gives a variation coefficient of 130.6%, showing the great distance between the members due to the enormous difference in committee memberships. However, if we compare the network with an equivalent one in number of nodes totally centralised through a hub vertex (100%), we obtain a level of 5.2%. This fairly low value indicates that, though there are many differences between the nodes, the network has a large number of vertices that occupy important places with regard to the number of relations established by co-participation on committees. In other words, it does not depend on a few centralising elements.

To highlight the general centrality in the network we used the closeness indicator of Beauchamp (1965). According to this indicator, in our network of committees, the members who occupy a more central place close to all the others slightly alter the ranking of the first ten in comparison with the indicator based on the degree of connectivity, and the first place is now occupied by Albert Hauf (Table 9).

Committee member	Closeness indicator
Hauf Valls, A, G,	51.7
Castellanos Vila, J,	51.2
Ferrando Francés, A.	49.9
Mas Vives, J.	49.2
Jorba Jorba M.	49.0
Pla Arxé, R.	49.0
Molas Batllori, J.	47.6
Badia Pàmies, L.	47.4
Rossich Estrago, A.	47.3
Beser Ortí, S.	47.2

Table 11. Closeness of the committee members

The importance of one node as a mediator between the others was measured with the betweenness indicator (Freeman *et al.* 1991).

Committee member	Betweenness indicator
Hauf Valls, A. G.	14.5
Ferrando Francés, A.	11.1
Castellanos Vila, J.	10.8
Mas Vives, J.	9.2
Molas Batllori, J.	6.7
Pla Arxé, R.	6.6
Alegret Llorens, J.	6.5
Jorba Jorba, M.	6.2
Badia Pàmies, L.	5.8
Beser Ortí, S.	5.0

Table 12: Betweenness of the committee members

In this case, the values shown in Table 12 are highly coincident with those of closeness, but not exactly the same, which shows their complementary nature. For greater clarity we offer in Table 13 the six actors who figure at the same time in the ten first places in the three measures of centrality

Committee member	Degree		Closeness		Betweenness	
	Value	Position	Value	Position	Value	Position
Hauf Valls, A.G.	78	2	51.7	1	14.5	1
Castellanos Vila, J.	100	1	51.2	2	10.8	3
Ferrando Francés, A.	59	7	49.9	3	11.1	2
Mas Vives, J.	67	4	492	4	9.2	4
Pla Arxé, R.	76	3	49.0	5	6.6	6
Jorba Jorba, M.	59	8	49.0	5	6.2	8

Table 13: Degree, Closeness and Betweenness of the most important committee members

These were the most important actors in the whole period of study from 1976 to 2003. The highest value of the degree indicator, which represents a committee member's level of direct communicative activity with others, corresponds to Jordi Castellanos. However, he loses this privileged position to Albert Hauf when the closeness to other nodes is measured, and is even surpassed by Antoni Ferrando in betweenness, i.e., the dependence of the network on his presence in order for there to be other members.

The use of several indicators jointly is revealing. For example, Alberto Blecua, who occupies sixth place in degree due to his high participation on committees (11.3% of the total) occupies 11th place in closeness and 12th place in betweenness. These more modest places are commesurate with the fact that he is a specialist in Spanish literature (rather than Catalan) and therefore, despite his importance due to his number of committee memberships, he is not a key figure in them because they are outside his speciality.

Most significant relations

The many co-participation relations between the vertices of the network hinder the appearance of clear structures or sub-networks corresponding to specific subjects or any other taxonomy. However, one of the tools for determining mutually exclusive groups in a non-binary network such as this one is to seek weak transitivity relations. In other words, to belong to the same group it must be fulfilled that, if there exists a relation between actor a and actor b, and between actor b and actor c, the direct relation between a and c must be of the same intensity; otherwise, the elements do not belong to the same triad (Hanneman and Riddle 2005). The result of this type of analysis is a group of members who are strongly linked, formed by Jordi Castellanos, Ramon Pla, Enric Gallen, Jaume Aulet, and Glòria Casals, all of whom are mainly devoted to contemporary literature.

Alternatively, links can be found between the members of the network through the Pearson linear correlation matrix. Here we performed the analysis for cases of at least five committee memberships, so that the coincidences make sense. The results show three main groups of vertices corresponding to researchers devoted mainly to contemporary, mediaeval and modern literature.

Network of supervisors and committee members selected

This graph is of the directed type, because the arcs are oriented from the vertices corresponding to thesis supervisors to the persons selected as committee members. Committee members who were also thesis supervisors were excluded. Figure 11 shows the main selection relations that are established. Again, for clarity the arcs of value 1 and the vertices that were isolated were eliminated. Thus, the only cases shown are those in which the thesis supervisor has used the same specialist as a committee member at least twice.

The network can be considered to have three components. The main one is that formed by Molas and Castellanos, who had Molas as his thesis supervisor. The fact that these vertices radiate a cluster of vertices in the diagram (especially from Molas) means that these academics repeated as committee members on the basis of their work as supervisors.

A second, far smaller, structure radiates from Antoni Comas at the top right of Figure 11. Finally, at the middle left, a small structure shows the relation between Sanchís Guarner and Peréz Durà.
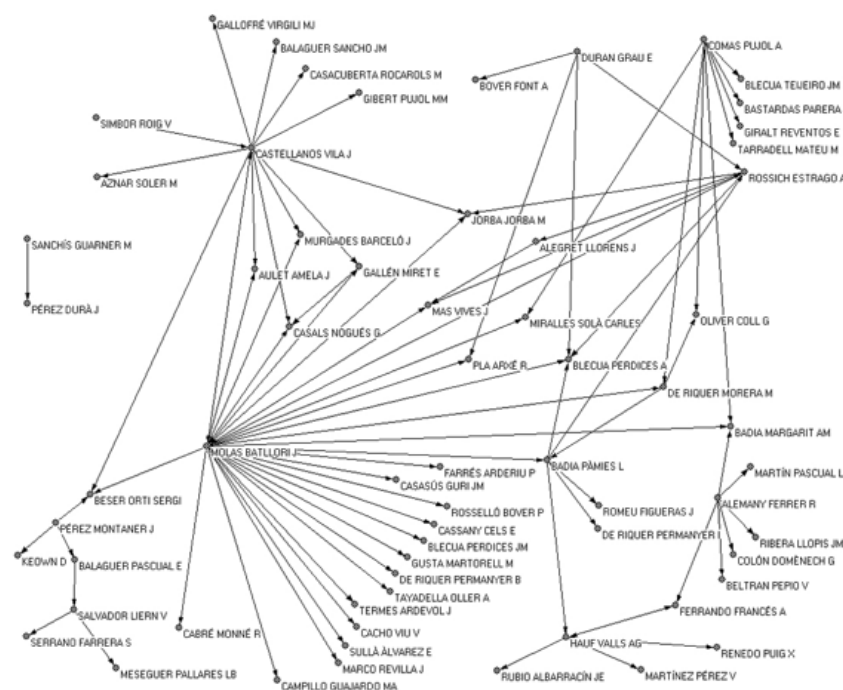

Figure 11: Network of thesis supervisors and selected committee members

Figure 12 shows the intensity of the relations between the most prolific thesis supervisors and the selected committee members during the whole study period. For the purpose of graphical clarity, the vertices least demanded by a specific supervisor were eliminated from the network. The result, centred on Joaquim Molas, is especially striking with regard to his relation with Castellanos. Molas selected Castellanos for sixteen theses, which is almost half of the thirty-four that he supervised according to TESEO. In turn, Castellanos selected Molas for three out of the ten theses he supervised, obviously a less significant value in view of the far smaller number of theses he supervised.
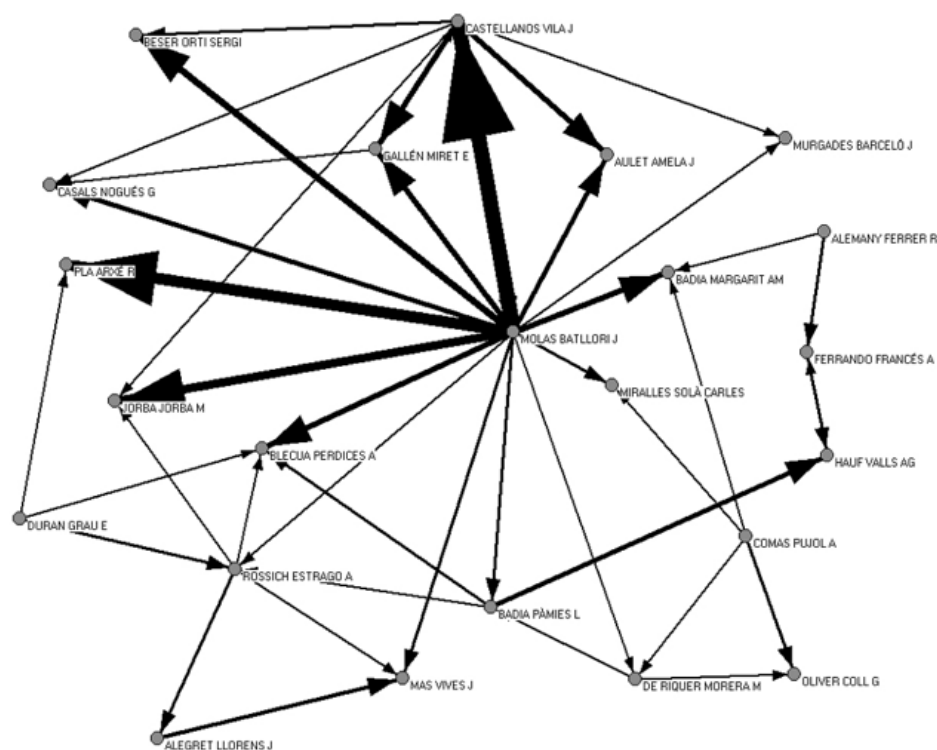
Figure 12: Network of most prolific thesis supervisors

Network of committee members selected and area of literature of the thesis

In this case the network links each committee member to the sub-area of Catalan literature in which the PhD thesis can be placed. Again, it was necessary to simplify the diagram (Figure 13) to show only the vertices that participated in a thesis of the specific subarea at least three times. This meant that literary theory and comparative literature, with a low proportion of the theses (5%) and almost no repetition of committee members, do not appear.

It is worth noting the two vertices that participated with a fairly high intensity in mediaeval, modern and contemporary literature: Albert Hauf, though his speciality is mediaeval literature, as shown in the graph; and Alberto Blecua Perdices who, though from a Spanish literature department, participated mostly in committees of modern and contemporary literature, and slightly less of medieval literature. On the other hand, Castellanos, Molas, Gallén, Pla Arxé and Aulet focused on their speciality of contemporary literature.
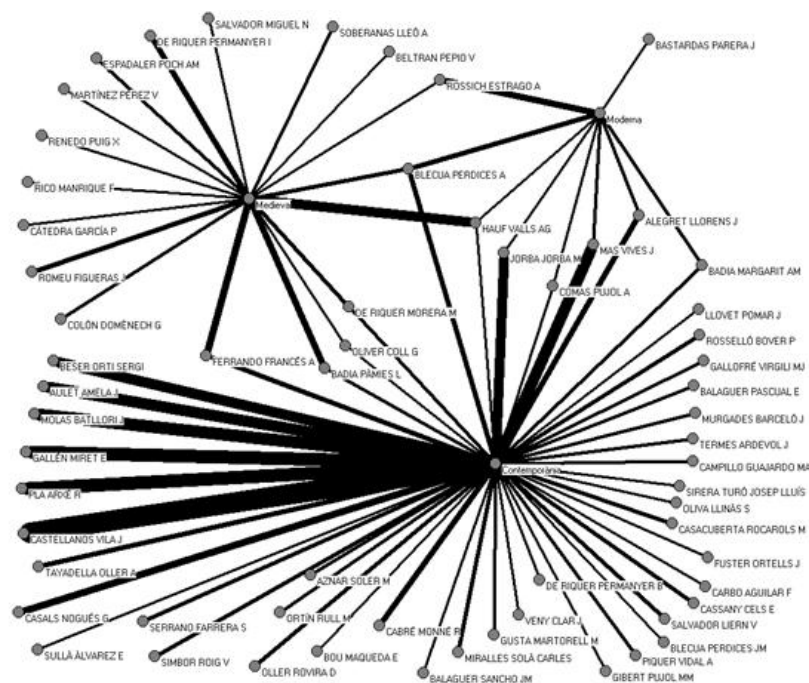
Figure 13: Network of relations between committee members and the area of literature of the thesis

Conclusion

Catalan literature is a discipline which, after an initial expansion with the restoration of democracy in Spain, has suffered a period of stagnation. The study detected a set of academic figures who were highly influential and the strong influence of members of the same university at which a thesis was defended. This is partly due to the small number of academics in the area but does not help to create an image of intellectual independence. Related to this is the almost negligible presence of academics from universities in other countries. Substantial support should be provided to remedy the current gender bias so that research work by women reflects their majority presence on Catalan philology courses.

The large amount of information obtained (only partially presented here: see Ardanuy 2008) confirms the extraordinary relevance of social network analysis techniques for bibliometrics. It can be used for data not only from citation analysis, but also from other bibliographic or administrative records. The information obtained is very relevant to the humanities in general, in which the Arts and Humanities Citation Index has a limited scope because it does not index monographs and shows many limitations with regard to non-English-speaking serial publications. It is also particularly relevant to the study of literatures of languages with fewer speakers, which are undervalued in comparison with languages with greater demographic strength such as French and Italian, just as the latter are in comparison with English.

References

- Ardanuy, J. (2008). *Anàlisi bibliomètrica de la producció científica en literatura catalana*. [Bibliometric analysis of the scientific production in Catalan literature.] Barcelona: Universitat de Barcelona. (Doctoral dissertation). Retrieved 11 June, 2009 from http://www.tesisenxarxa.net/TDX/TDX_UB/TESIS/AVAILABLE/TDX-0226108-124023//JAB_TESI.pdf.
- Ardanuy, J. Quintana, L. & Urbano C. (2008). Una aproximació bibliomètrica a la recerca en literatura catalana. [A bibliometric approach to research on Catalan literature.] *Els Marges*, No. 86, 41-66.
- Beauchamp, M. A. (1965). An improved index of centrality. *Behavioral Science*, **10**(2), 161-163.
- Bourke, P. (1997). *Discipline boundaries in the social sciences*. Canberra: Academy of the Social Sciences.
- Civera Molla, C. & Tortosa Gil, F.M. (2001). Estado de la investigación psicológica en España: el grado de Doctor y la investigación académica (1976-1998). [The state of psychological research in Spain: PhD and academic research (1976-1998)] *Papeles del psicólogo*, 79, 42-52.
- Delgado, E., Torres Salinas, D., Jiménez-Contreras, E. & Ruiz Pérez, R. (2006). Análisis bibliométrico y de redes sociales aplicado a las tesis bibliométricas defendidas en España (1976-2003): temas, escuelas científicas y redes académicas. [Bibliometric analysis of social networks applied to bibliometric theses defended in Spain (1976-2003): issues, scientific schools and academic networks.] *Revista Española de Documentación Científica*, **29**(4), 493-524.
- Duran, E. (2004). *Reports de la recerca a Catalunya. 1996-2002 filologia catalana: literatura*. [Reports of research in Catalonia. 1996-2002 Catalan philology: literature.] Retrieved 11 June, 2009 from http://www.iecat.net/recerca/reports/Publicats2Edicio/Rep_Filolo-literat-2.pdf
- Eaton, J.P., Ward, J.S., Kumar, A. & Reingen, P.H. (1999). Structural analysis of coauthor relationships and author productivity in select outlets for consumer behavior research. *Journal of Consumer Psychology*, **8**(1), 39-59.
- Fernández Cano, A., Torralbo, M., Rico, L., Gutiérrez, P. & Maz, A. (2003). Análisis cienciométrico de las tesis doctorales españolas en educación matemática (1975-1998). [Scientometric analysis of Spanish doctoral theses on mathematics education (1075-1998).] *Revista Española de Documentación Científica*, **26**(2), 162-176.
- Freeman, L.C. (1977). A set of measures of centrality based upon betweenness. *Sociometry*, **40**(1), 35-41.
- Freeman, L.C., Borgatti, S.P. & White, D.R. (1991). Centrality in valued graphs: a measure of betweenness based on network flow. *Social Networks*, **13**(2), 141-154.
- Hagstrom, W.O. (1965). *The scientific community*. New York, NY: Basic Books.

- Hanneman, R. A., Riddle, M. (2005). *Introduction to social network methods* Riverside, CA: University of California, Riverside. Retrieved 11 June, 2009 from http://faculty.ucr.edu/~hanneman/nettext.
- Kamada, T. & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, **31**(1), 7-15.
- Kreuzman, H.(2001). A co-citation analysis of representative authors in philosophy: examining the relationship between epistemologists and philosophers of science. *Scientometrics*, **51**(3), 525-539.
- Lázaro y Torres, M. L. (2002). La geografía a las puertas del tercer milenio a partir de las tesis doctorales leídas en los noventa. [Geography at the gates of the third millennium from doctoral theses read in the nineties.] *Estudios Geográficos*, **63**(246), 154-168.
- Leydesdorff, L. (1995). *The challenge of scientometrics: the development, measurement, and self-organization of scientific communications*. Leiden, The Netherlands: DSWO Press, Leiden University.
- Leydesdorff, L. (2004). Clusters and maps of science journals based on bi-connected graphs in 'Journal Citation Reports'. *Journal of Documentation*, **60**(4), 371-427.
- López Yepes, J (2002). Focos de investigación y escuelas científicas en documentación a través de la realización y dirección de tesis doctorales: el caso del Departamento de Biblioteconomía y Documentación de la Universidad Complutense de Madrid (1983-2001). [Foci of research and scientific schools in documentation through the implementation and direction of doctoral thesis: The case of the Department of Library and Information Science at the Complutense University of Madrid (1983-2001)] *Documentación de las Ciencias de la Información*, No. 25, 19-54.
- Lotka, A J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, **16**(12), 317-323.
- McCain, K.W. (1990). Mapping authors in intellectual space: a technical overview. *Journal of the American Society for Information Science*, **41**(6): 433-443.
- Molas, J. (2001). *Reports de la recerca a Catalunya: literatura catalana*. [Reports of research in Catalonia: Catalan literature.] Retrieved 11 June, 2009 from http://www.iec.cat/gc/digitalAssets/500_3680_1097660898265_filologia.pdf.
- Molina, J.L., Muñoz, J.M. & Domènech, M. (2002). Redes de publicaciones científicas: un análisis de la estructura de coautorías. [Networks of scientific publications: an analysis of the coauthorship structure.] *Redes.Revista hispana para el análisis de redes sociales*, **1**(3). Retrieved 11 June, 2009 from http://revista-redes.rediris.es/html-vol1/vol1_3.htm.
- Molteni, V, & Zulueta, M. A. (2002). Análisis de la visibilidad internacional de la producción científica argentina en las Bases de datos Social Science Citation Index y Arts and Humanities Citation Index de 1990-2000. [Analysis of the international visibility of scientific production in Argentina in the Social Science Citation Index and Arts and Humanities Citation Index databases from 1990-2000.] *Revista Española de Documentacion Científica*, **25**(455), 455-465.
- Moya, F., Vargas, B., Chinchilla, Z., Corera, E., González, A., Múñoz, F. J. & Herrero, V. (2006). Visualización y análisis de la estructura científica española: ISI Web of Science 1990-2005. [Visualization and analysis of the structure of Spanish science: ISI Web of Science 1990-2005] *El profesional de la información*, **15**(4), 258-269.
- Newman, M. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, **98**(2), 404-409.
- Otte, E. & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, **28**(6), 441-453.
- Rousseau, B., & Rousseau, R. (2000). LOTKA: a program to fit a power law distribution to observed frequency data. *Cibermetrics* **4**(1), Retrieved 11 June, 2009 from http://www.cindoc.csic.es/cybermetrics/articles/v4i1p4.html.
- Vivas, J. R. and Urquijo, S. (1999). Análisis de redes aplicado a un estudio bibliométrico en psicología de la conducta. *Iber Psicología*, **4**(1). Retrieved 11 June, 2009, from http://www.fedap.es/IberPsicologia/iberpsi4-1/vivas/vivas.htm.
- Watts, D. J. (2004). *Small worlds: the dynamics of networks between order and randomness*. (2nd ed.). Princeton, NJ: Princeton University Press.

How to cite this paper

Ardanuy, J., Urbano, C. & Quintana, L. (2009). "The evolution of recent research on Catalan literature through the production of PhD theses: a bibliometric and social network analysis" *Information Research*, **14**(2) paper 404. [Available at http://InformationR.net/ir/14-2/paper404.html]

Find other papers on this subject

Scholar Search   Google Search   Bing

Bookmark This Page

Contents | Author index | Subject index | Search | Home