

Grey literature on bilingualism in Belgium

J. Schöpfel

INIST-CNRS and research groupe "Document Numérique & Usages"

Abstract

Because of the town Antwerp host of the GL9 conference and because of our own former scientific experience we selected the field of bilingualism for a study on the importance of grey literature in social sciences and humanities. The study is meant to be an additional contribution to comparable scientometric analyses on the distribution of types of publications in different scientific domains.

Bilingualism, the learning and use of two or more languages, is a linguistic, social, educational and psychological reality for many people and most countries. In the heart of Europe, Belgium, a country with two cultural and linguistic populations, with immigration, international business and institutions, is particularly confronted with this reality. Reaction to and part of its multicultural society, Belgium developed since many years a significant interdisciplinary research activity in the field of bilingualism.

The particularity of our study is twofold: First, the research on bilingualism is interdisciplinary, at the crossroad of linguistics, sociology, psychology and educational sciences, each domain presenting its own vectors of publication and communication. Second, while most of the previous studies on the importance of grey literature are citation analyses, our study is based on search results from databases, catalogues, open archives and search engines.

Notes on the author

Joachim Schöpfel obtained his Ph.D. in psychology from the Hamburg University in 1992. At present he is head of the e-publishing and document supply department at INIST-CNRS, lecturer in information science at the Nancy University and member of the research group "Document Numérique & Usages" of the Paris 8 University. He published on grey literature, document delivery, digital libraries, usage statistics and professional development.

Contact address

Dr. Joachim Schöpfel, INIST-CNRS, 2 allée du Parc de Brabois, 54519 Vandœuvre-lès-Nancy, France. E-Mail: schopfel@inist.fr

1. On the (relative) importance of grey literature

Grey literature has a role of its own as a means of distributing scientific and technical information, and professionals insist on its importance for two main reasons: research results are often more detailed in reports, doctoral theses and conference proceedings than in journals, they are distributed in these forms up to 12 or even 18 months before being published elsewhere, and some results simply aren't published elsewhere.

But how do researchers use grey literature? One way of evaluating this is to analyze the citations given in their publications. Schöpfel & Farace 2007 provide some empirical evidence based on a synthesis of different scientometric studies (see table 1).

Field	Grey literature citations (in %)
Soil science	14%
Biology	5-13%
Veterinary medicine	6%
Psychiatry (addiction)	1%
Psychology	3%
Engineering Sciences	39-42%
Economics	9-17%
Sociology	7-9%
Education Science	14-19%

Table 1: Part of grey literature in different scientific domains

The relative importance of grey literature is largely dependent on research disciplines and subjects, on methodological approaches, and on sources used. In some fields, especially the life sciences and medical sciences, there has been a traditional preference for conventional distribution media (journals), while in others, such as agriculture, aeronautics and the engineering sciences in general, grey literature resources tend to predominate¹.

A small study of bibliographic references published in four different journals in library and information sciences revealed that grey literature accounts on average around 20% of all sources used. Even so, citations to grey material vary widely between different papers from 0% to 50% and more, depending at least in part on subject areas and methodologies. Most of the non-commercial documents can clearly be identified as unpublished material, doctoral or master theses, proceedings, reports, legal texts, working papers, lecture notes, websites, and even posted messages. But especially

¹ See also Schöpfel et al. (2005) and Farace et al. (2006).

for some conference proceedings and other serial resources, the identification of the exact nature of the cited material poses difficulties.

In order to add more empirical evidence on the position of grey literature in scientific publishing, especially in Social Sciences and Humanities, we carry out an alternative approach that is complementary to the citation analyses mentioned above.

We select a scientific field in Social Sciences and Humanities with a specific relationship to the hosting town of the conference GL9 (bilingualism in Belgium), we define a recent time period (2000-2007), we investigate with usual search tools the scientific production (databases, catalogues, search engines, repositories), and finally we evaluate the part of grey literature of this production. Our leading questions were:

- (1) What is the scientific production on bilingualism in Belgium?
- (2) How important is grey literature for this scientific activity?
- (3) What are the particular features of grey literature in this field?

The following section provides a global overview on Belgium and on scientific research on bilingualism, together with some general features related to publications. The third section describes the methodological approach. The 4th section presents the empirical data, the 5th section discusses the limits of the study and resumes the major results, and the 6th section concludes with some apparent features of this sample of grey literature and with two additional questions about quality and access.

2. Bilingualism as an object of scientific study

In our global village, learning and using two or more languages is an everyday reality of many people and most countries. Among the most important reasons for this reality are immigration, international commerce and business, linguistic and cultural minorities. In the heart of Europe, Belgium presents two more particularities: the nation is composed of two cultural and linguistic populations, the French-speaking Walloons in the South and the Dutch-speaking Flemish in the North, while its capital, Brussels, is also the capital of Europe with important international institutions, such as the Commission of European Communities, the European parliament and the NATO headquarter².

Verbal behaviour is linked to economic, social, historical, political, psychological, educational and linguistic factors. Thus, the scientific regard on bilingualism is manifold and depends on the specific background and methodology of the scientist: the psychologist will study the influence of motivation, attitude and cognitive development, the social scientists will focus on identity, role and social group, the linguist will analyse cross-linguistic semantics and syntactic, the educational researcher will evaluate the output of bilingual programs, a psychiatric specialist will publish case

² See for instance Swenden & Jans (2006).

studies on pathological verbal behaviour, and a historian will try to understand the cultural, political and economic roots of bilingualism. In many cases, research on bilingualism is conducted by interdisciplinary teams, developing original psycholinguistic, sociolinguistic, psychopedagogical or even neuro-linguistic approaches³.

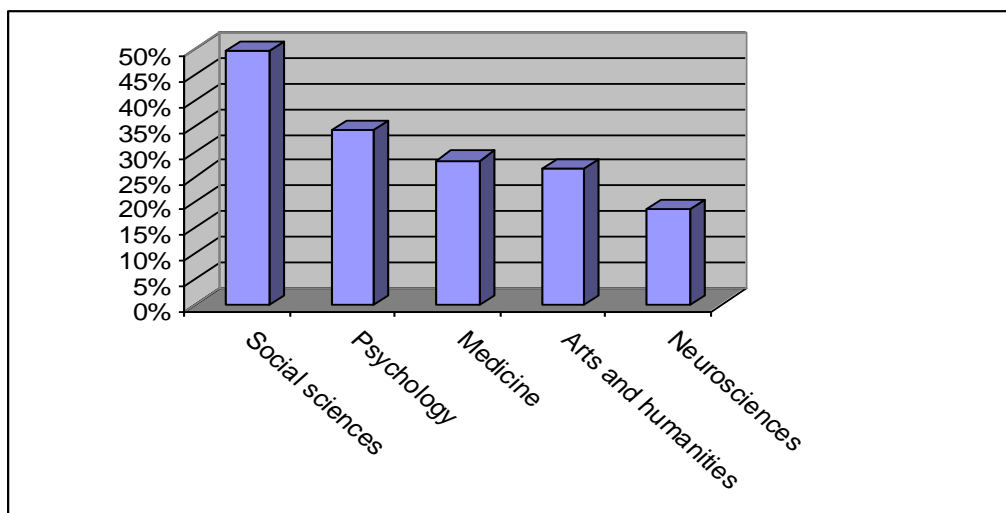


Figure 1: Publications on bilingualism 2000-2007 – scientific domains (Scopus)

50% of recent literature in the Scopus database (from 648 items published between 2000 and 2007) is from social sciences, while psychology, medicine, arts and humanities and neurosciences account for 20-35% (see figure 1).

Figure 1 shows also that articles on bilingualism are partly published in journals indexed in more than one scientific discipline.

In spite of the universal character of bilingualism, only some countries developed a significant amount of scientific research on bilingualism.

Another analysis of the Scopus database reveals that 7 countries produce 85% of the recent literature on bilingualism (see figure 2).

Not surprisingly, more than the half of the indexed literature is published by American or Canadian scientists. All the same, Belgium with 4% ranks at nearly the same level as France (6%) or Germany (5%), which confirms the importance of its scientific production in this field. These results from only one database may not be exhaustive but they seem representative enough.

³ See for instance some highly cited studies on bilingualism such as Portes & Schauflier (1994), Fabbro (2001), Leseaux & Siegel (2003) and Hartsuiker et al. (2004).

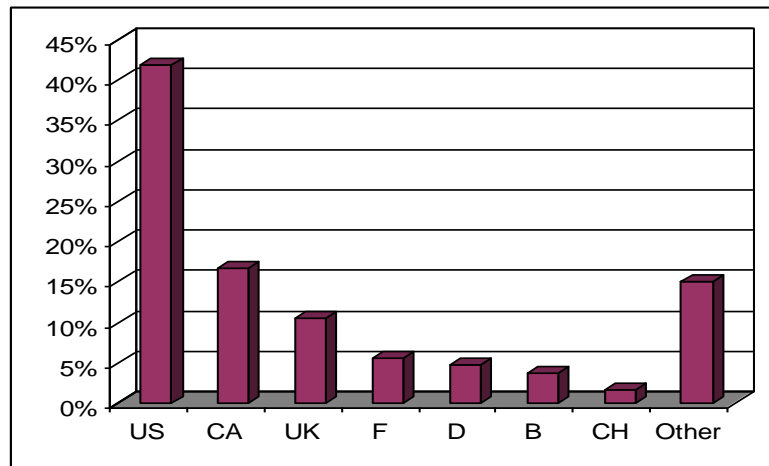


Figure 2: Publications on bilingualism 2000-2007 – countries (Scopus)

3. Methodology

Bibliographic searches in scientific databases and library catalogues and with STI and general search engines were carried out to identify the Belgian scientific production on bilingualism. The following databases were selected:

SCI Science Citation Index (ISI Web of Knowledge)
SSCI Social Science Citation Index (ISI Web of Knowledge)
A&HCI Arts and Humanities Citation Index (ISI Web of Knowledge)
Scopus (Elsevier)
IBSS International Bibliography of the Social Sciences (CSA)
ERIC (CSA)
PAIS International (CSA)
Social Services Abstracts (CSA)
Dissertation Abstracts (ProQuest)
PsycARTICLES (Ovid)
PASCAL (INIST)
FRANCIS (INIST)
ISD (INIST)

Together they cover the core of scientific production (WoK, Scopus, the INIST databases, Diss. Abstr.) and more specific domains related to bilingualism (social science, humanities). The databases were accessed through the national CNRS SS&H portal "BiblioSHS"⁴ or via the local academic campus. Only regret: for technical reasons the *Linguistics Abstracts* database could not be accessed.

The Belgian academic library catalogues that reflect the major national scientific holdings were accessed via the union OPAC *Libis*, via the ILL network *IMPALA* or directly on the universities' web sites (Leuven, Ghent, Brussels etc).

⁴ <http://biblioshs.inist.fr>

Complementary searches were undertaken with three mostly used generalist and STI search engines:

Scirus (Elsevier)
Google Scholar
Google

The *Directory of Open Access Repositories*⁵ (OpenDOAR) was used to identify and search in Belgian open archives. We searched also directly on different institutional web sides (university home pages) for institutional and/or thematic open repositories.

The study was undertaken between May and November 2007.

The search criteria were:

1. Concepts of "bilingualism" in titles, keywords and abstracts (in English, French or Flemish)⁶.
2. The mention of Belgium in titles, authors' address, authors' affiliation or other institutional information.
3. Limitation to recent publications from 2000 to 2007.

The search strategy was rather iterative than linear, alternating between databases and catalogues, repositories and search engines in order to complete affiliations and references, confirm free access and check for other publications from identified authors or co-authors.

Based on the search results two file sheets were created, one with information on authors (first and last name, institution/affiliation, domain and sub-domain), another with information on publications (first author, presence of co-authors, short title, language, year of publication, type of publication, free access on the web, deposit in repository). Data analyses combined information from both tables.

4. Results

The three leading questions for the data analysis are:

- (1) What is the scientific production on bilingualism in Belgium? Who publishes in the field of bilingualism, where are the main academic and research structures in this field, which are the dominating topics and approaches?
- (2) How important is grey literature for this scientific activity? What is the part of theses, reports and conference proceedings compared to the overall production of results of the research on bilingualism in Belgium?
- (3) More specifically, what are the particular features of grey literature in this field? Special attention is paid to the importance of digital information available on the Web, on academic or personal pages or in institutional repositories and other open archives.

⁵ <http://www.opendoar.org>

⁶ E.g. bilingual, bilingualism, bilingue, bilinguisme, tweetalig, tweetaligheid, second language...

Based on empirical evidence, the following sections try to provide at least partial responses.

4.1. Publications: number, year, language and category

Search results in databases, catalogues and repositories and on the web allow for the identification of 164 different publications on bilingualism in Belgium or written by Belgian scientists on this topic, with an average production of 20 publications per year (see figure 3). Considering that 2007 publications are only incompletely recorded in databases and catalogues so far, an overall growth of around 20% can be observed from 2000 to 2007 that is consistent with the general tendency of an annual increase of 2-3% for overall scientific production.

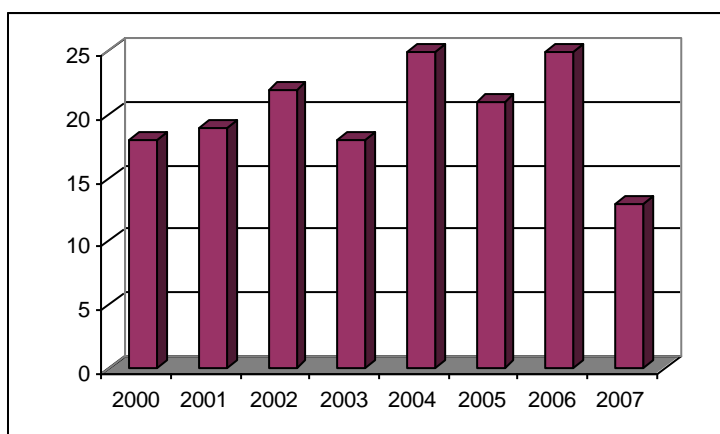


Figure 3: Publication year (sample)

85% of the identified documents are published in English, 13% in French and 5% in Flemish; 2 documents are in Spanish, one in German and another in Italian.

Most of the documents are published as articles in scientific journals (see figure 4).

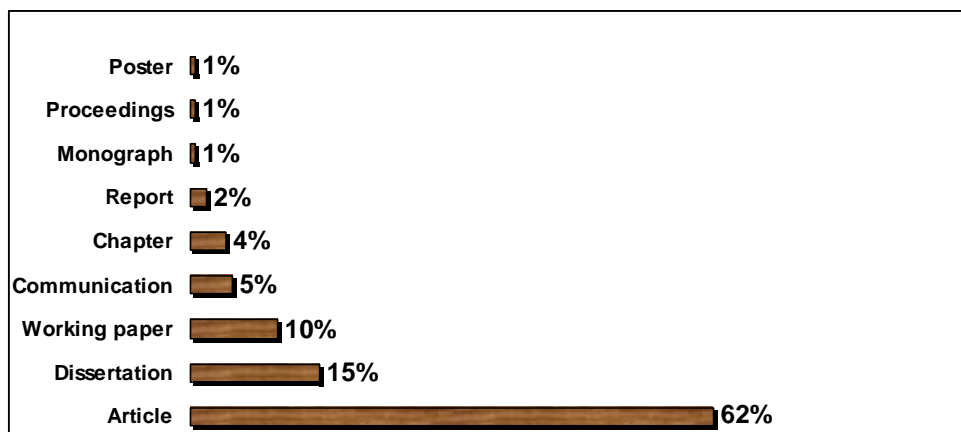


Figure 4: Publication category (sample)

30% are dissertations (e.g. doctoral or master theses), working papers or communications held in conferences. The rest – less than 10% - are book chapters, reports, a monograph, a book with a conference proceedings and a poster.

4.2. Authors, institutions and domains

The 164 publications are written by 81 first authors to whom we added 5 co-authors clearly identified as being affiliated to Belgian academic institutions. Most of the first authors have published one or two documents on bilingualism between 2000 and 2007 but three researchers from the Ghent University occupy a more important and noticeable position: Jan Blommaert with 27 publications, Marc Brysbaert with 9 publications and Robert Hartsuiker with 6 publications.

Affiliation of first author	Number of publications
Ghent University	75
Leuven Katholieke Universiteit (KUL)	13
Antwerp University	11
Brussels Free University (ULB)	10
Lessius Hogeschool Antwerp	5
Kempen Hogeschool	4
Liège University	4
Brussels Research Center for Multilingualism	2

Leuven Catholic University (UCL)	2
Hasselt University	2
Brussels Royal Academy of Science and Arts	1
Namur University	1
Not identified or non Belgian	34

Table 2: Affiliation of first author (sample)

The first authors are affiliated to 12 Belgian scientific institutions, even if some of them are also affiliated to institutions in other countries, such as London University (see table 2).

Obviously the most important Belgian institution for research on bilingualism is the Ghent University with nearly 50% of all publications of the 2000-2007 sample. Four academic campus – Ghent, Leuven (KUL), Antwerp and Brussels (ULB) – cover together 2/3 of the scientific production on bilingualism. 20% of the publications can not be clearly linked to an university or another Higher Education structure or research center, or are written by an author affiliated to a non Belgian institution.

For each publication the major scientific domain was derived from the authors' affiliation, from database indexing and from the topic of publication⁷ (see figure 5).

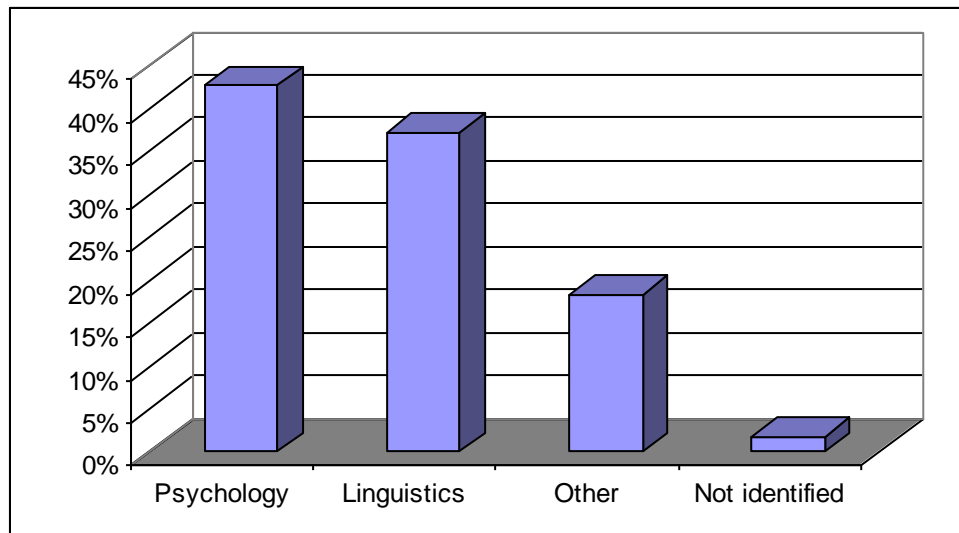


Figure 5: Scientific domains of publications (sample)

Psychology and Linguistics cover together 80% of the sample. The remaining 20% are from other disciplines, e.g. Education, Communication,

⁷ E.g. title, abstract, author's key words

Neurosciences, Political Sciences, Economics and Business, Literature, History and even Informatics.

Nevertheless, we can distinguish more specific and/or interdisciplinary approaches such as Cognitive Science, Experimental Psychology and Neuropsychology for the Psychology discipline, Ethnography and African Language and Cultures for Linguistics, and Psycholinguistics at the intersection of both disciplines. Some typical subjects of publications:

- Cross-cultural communication in international schools
- The role of schooling
- Multinational and polyethnic Politics entwined
- Aphasia in a bilingual brain
- Forward and backward number translation
- Ethnography, discourse and hegemony
- Influence of the language-specific orthographic representations
- Bilingual visual word recognition
- Cross-language comedy
- Bilingualism and the development of executive control
- Understanding bilingual memory
- Size of the cross-lingual masked phonological priming
- Cross-linguistic priming
- Universals of language maintenance
- Pathophysiology of language switching
- Multilingualism: an fMRI study
- Investigating narrative inequality

Compared to the international production (content of the Scopus database, see figure 1), there seem to be significantly less Belgian publications on bilingualism in Social and Medical Sciences but as we did not match up our indexing to the Scopus index, this may be at least partly biased by our method.

4.3. Features of grey literature

In our sample, 55 documents (34%) correspond to the criteria of the New York definition of grey literature, e.g. "which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers i.e. where publishing is not the primary activity of the producing body" (GL6, 2004).

Nearly half of these grey documents (45%) are dissertations (master or doctoral theses), another 31% are unpublished working papers, and 15% are communications (see figure 6).

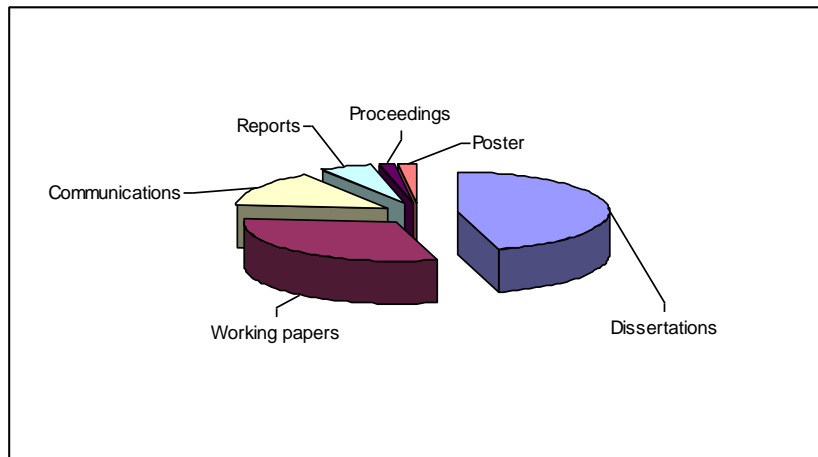


Figure 6: Typology of grey documents (sample)

A closer look on these documents reveals some differences with the “white” publications:

Production: The average number of grey literature is not increasing but stable (around 6-8 publications per year) or slightly decreasing (only 2 identified grey documents in 2005).

Language: The part of non English publications is significantly higher than of articles. While 95% of the published articles and book chapters are in English, the percentage of English grey literature is only 64%, e.g. one third of grey documents are written in French (18%), Flemish (13%) or other languages.

Domains: The part of disciplines other than Linguistics or Psychology is slightly higher (24%) than for articles (18%).

Institutions: In contrast, there seems to be no significant difference regarding to scientific institutions, e.g. no one of them produce more or less grey literature compared to white output.

Authors: The grey documents were published by 38 first authors (47% of all first authors). 24 of them (63%) didn’t publish anything else during the observed time period.

Grey to white: Based on the comparison of titles, we can suppose that in at least 8 cases (15%) it exists a direct relationship between grey and white publications, the grey document (working paper, communication, dissertation) preceding one, two or even four years the publication of an article or a chapter on the same subject.

4.4. Open access to grey documents

For each of the 55 grey documents internet searches were performed, firstly with the mentioned search engines, then with the OpenDOAR search facility, finally directly in some major Belgian academic repositories, especially from

the institutions to which the first authors are affiliated. The results are shown in figure 7.

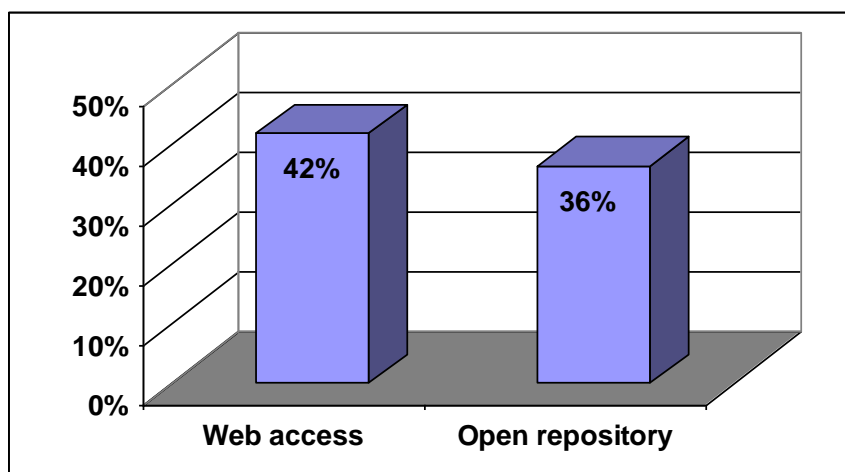


Figure 7: Web access to grey documents (sample)

23 documents (42%) can be accessed on the Web, most of them (20 documents or 36%) are deposited in an open, mainly institutional repository.

Compared to the different categories of grey documents, a tendency or bias can be discerned: In our sample, reports and working papers are more often accessible on the web while dissertations are less. Maybe that this last tendency is related to our choice to include master theses in the sample that are less often deposited on an academic web sites than doctoral theses. In some rare cases access to documents seems to be lost.

5. Discussion

The presented results are limited by some factors related to the methodology and/or to the subject of the study, e.g.

Language: Belgian academics publish in at least three languages, English, French and Flemish. As a matter of fact, international scientific databases are strongly biased and contain mainly English publications. Even if we tried to counterbalance this bias by searching directly in Belgian catalogues and repositories, searching in Flemish is more difficult than in French or English, even for a German-speaking professional. So some Flemish publications may simply have been overlooked.

Affiliations: Analysis of affiliations was limited to the first author. Nevertheless, 56% of the publications of the sample are co-authored. To gather an in-depth and more comprehensive look on the scientific production of the Belgian academic and research institutions, affiliations of the co-authors should be evaluated.

Scientific fields: Differing from the Scopus database, we only indexed the main scientific domain for each publication which introduces a bias to the major disciplines, e.g. Psychology and Linguistics.

Interdisciplinary: As mentioned above, bilingualism is an object of interdisciplinary research interest. In some "borderline" cases it is not obvious to decide whether a given publication is part of research on bilingualism or should be excluded from the sample, bilingualism being only a marginal topic of limited interest.

The "Belgian aspect" of research: The sample contains (mostly) publications from Belgian authors as well as (some) studies on bilingualism in Belgium published by non Belgian authors. In the second case attention was paid that the publication places this aspect in the centre of interest. This double approach allows a more comprehensive analysis but may introduce some "fuzziness" in the evaluation of affiliations and open access.

Repositories: Some of the visited institutional repositories provide rather limited search facilities. The alternative approach, browsing or searching "from outside" based on OAI metadata protocol harvesting, may exclude some publications on bilingualism from the sample.

In spite of these limitations, the main results of the study seem to be sufficiently significant and can be resumed as follows:

(1) Scientific production on bilingualism in Belgium: More than 80 academics and researchers published between 2000 and 2007 on bilingualism in Belgium. The annual output is about 20 publications, with a slightly increasing tendency. The main research institutions are the Ghent University and, to a lesser extent, the universities of Leuven (KUL), Antwerp and Brussels (ULB). Most research is performed in Psychology and Linguistics, but the interdisciplinary character of most publications is apparent.

(2) Importance of grey literature: One third of the publications on bilingualism are grey, nearly half of them are master or doctoral theses, followed by working papers and communications.

(3) Particular features of grey literature: Compared to white publications (articles), the part of non English documents is more important. Two third of the authors of grey documents didn't publish their results in articles or books; a smaller part of grey information was published in journals up to four years later. Nearly half of grey publications are freely available on the web, most of the time in institutional (academic) repositories.

6. Conclusion

The goal of our research was to provide additional empirical evidence on the importance of grey literature in Social Science and Humanities, based on a sample of recent publications. The selected topic, bilingualism, is not a strategic or frontline domain such as biotechnology or genomics. Its

specificity lies in the interdisciplinary approach, at the frontier of social sciences, humanities and medical sciences and in the strong relationship with the societal environment.

Our sample shows a dynamic and developing scientific research activity and production where the grey literature occupies a significant place, complementary to the articles published in academic journals. Its signification – more than 30% of all publications – can be described by three aspects, timeliness, uniqueness, and community.

Timeliness: One part of the grey documents, in particular dissertations, working papers and communications, present research results one or more years before formal and peer reviewed publications.

Uniqueness: Apparently, one part of the research results presented in grey documents has not been published elsewhere so far, and probably will not be. In these cases, grey literature is the only vector of scientific communication.

Community: Whereas articles in scientific journals are mainly published in English and available for the international scientific readership, grey literature is partly published in the community-specific languages Flemish and French and thus more deeply linked to the Flemish and Walloon scientific communities.

What about quality? Articles published in scientific journals – 62% in our sample – have undergone the usual quality control of scientific information, selection by editors, peer review, revision. Nevertheless, at least 60% of the grey documents have also been evaluated before publication: dissertations are evaluated by academic commissions; and communications have been selected by conference program committees and commented by the public (auditors), followed in some cases by revision before publication. The “quality gap” between grey and white appears thus to be rather relative than absolute, grey literature being at least partly published with added value through scientific peer evaluation.

A final comment should be made on open access to grey literature. The OpenDOAR site records actually (November 2007) 14 institutional, two disciplinary and one aggregating repository located in Belgium, most of them with multidisciplinary and multilingual (English, Flemish and French) content. Even if some of them present limited search functionalities, these repositories appear to be the best way to preserve (archive) and make available grey documents, especially in contrast to hard to find personal web pages or specific academic web sites such as lists of working papers without metadata and/or search facilities. In the future, authors of grey documents as well as scientific and academic institutions should be more conscious of this problem and act in favour of the utility and use especially of institutional repositories as a complementary vector of scientific communication⁸, in particular for working papers, master and doctoral theses and conference proceedings, in order to make and keep them available for a larger scientific community.

⁸ See for instance Lynch (2003), Correia & Neto (2002) and Davis & Connolly (2007).

7. Bibliography

Correia A.M.R.; Neto M.D. "The role of eprint archives in the access to, and dissemination of, scientific grey literature: LIZA – a case study by the National Library of Portugal". *Journal of Information Science* 2002, 28, 3, 231-241.

Davis P.M.; Connolly M.J.L. "Institutional Repositories". *D-Lib Magazine* 2007, 13, 3/4.

<http://www.dlib.org/dlib/march07/davis/03davis.html>

Fabbro F. "The bilingual brain: cerebral representation of languages". *Brain and Language* 2001, 79, 2, 211-222.

Farace D.J.; Frantzen J.; Schöpfel J.; Stock C.; Boekhorst A.K. "Access to Grey Content: An Analysis of Grey Literature Based on Citation and Survey Data". In *Seventh International Conference on Grey Literature: Open Access to Grey Resources. GL7 Conference Proceedings, Nancy, France, Dec 5-6, 2005*; Farace D., Frantzen J. Eds.; TextRelease: Amsterdam, 2006.

Hartsuiker R.J.; Pickering M.J.; Veltkamp E. "Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals". *Psychological Science* 2004, 15, 6, 409-414.

Leseaux S.A.; Siegel L.S. "The development of reading in children who speak English as a second language". *Developmental Psychology* 2003, 39, 6, 1005-1019.

Lynch C.A. "Institutional repositories: Essential infrastructure for scholarship in the digital age". *ARL Bimonthly Report* 2003, 226, 1-7.

<http://www.arl.org/newsltr/226/ir.html>

Portes A.; Schauffler R. "Language and the 2nd generation – bilingualism yesterday and today". *International Migration Review* 1994, 28, 4, 640-661.

Schöpfel J.; Stock C.; Farace D.J.; Frantzen J. "Citation Analysis and Grey Literature: Stakeholders in the Grey Circuit". *The Grey Journal* 2005, 1, 1, 31-40.

Schöpfel J.; Farace, D.J. "Grey Literature". In Bates, M.J. & Maack, M.N. (ed): *Encyclopedia of Library and Information Sciences*. 3rd edition. Taylor & Francis 2007 (forthcoming).

Swenden W.; Jans M.T. "Will it stay or will it go? Federalism and the sustainability of Belgium". *West European Politics* 2006, 29, 5, 877-894.