A Collection of Scholarly Book Reviews from the Platforms of Electronic Sources in Humanities and Social Sciences OpenEdition.org

Chahinez Benkoussas[*,**]**, Hussam Hamdan**[*,**,***]**, Patrice Bellot**[*,**]**, Frédéric Béchet**[***]**,
Elodie Faath**[**]**, Marin Dacos**[**]

*Aix-Marseille Université, CNRS, LSIS UMR 7296, 13397, Marseille, France
**Aix-Marseille Université, CNRS, CLEO OpenEdition UMS 3287, 13451, Marseille, France
***Aix-Marseille Université, CNRS, LIF UMR 7279, 13288, Marseille, France
Email: {chahinez.benkoussas,hussam.hamdan,patrice.bellot,frederic.bechet}@univ-amu.fr
{elodie.faath,marin.dacos}@openedition.org

Abstract

In this paper, we present our contribution for the automatic construction of the Scholarly Book Reviews corpora from two different sources, the OpenEdition platform which is dedicated to electronic resources in the humanities and social sciences, and the Web. The main target is the collect of reviews in order to provide automatic links between each review and its potential book in the future. For these purposes, we propose different document representations and we apply some supervised approaches for binary genre classification before evaluating their impact.

Keywords: Genre classification, Digital Library, Review, OpenEdition.

1. Introduction

Automatic recommendation of books is a hot topic in natural language processing and information retrieval. It can be performed by combining content analysis and collaborative filtering, for example by taking into account the behavior of the users (browsing, reading or shopping) or by analyzing the user comments and the short reviews on online web sites (bookshops or social networks like LibraryThing[1]) as it has been proposed for the Social Book Search track of INEX evaluation (Kazai, 2011).

Unlike previous studies on book recommendation, we are dealing here exclusively with scientific content in the social science domain. We are interested in collecting two kinds of reviews:

- long reviews of scientific books written by expert reviewers in scientific journals;
- short reviews that can be found on the Web or in social media.

These two kinds of reviews express two complementary points of view on a given book: deep expert analysis on the one hand, short reaction representing an opinion on the other hand.

In our knowledge, this will be the first corpus of this kind to be collected and that is freely available (via OAI-PMH access on the OpenEdition.org platform).

This paper is structured as follows: section 1 describes the OpenEdition platform that is the source of the books and the long reviews considered in this study. Section 2 describes the bootstrap corpus that has been manually collected from this platform for supervised learning of genre classification models. Section 3 and 4 present the two processes that have been designed for enriching this first corpus with a supervised approach. Lastly, Section 5 presents some perspectives.

2. Description of the OpenEdition Platforms of Books, Journal Papers and Scientific Blogs

OpenEdition[2] is the umbrella portal for OpenEdition Books, Revues.org, Hypotheses and Calenda, four platforms dedicated to electronic resources in the humanities and social sciences (books, journals, research blogs, and academic announcements respectively). Founded in 1999, Revues.org today hosts over 300 online journals, i.e. over 106,000 articles, proceedings and editorials, 99,900 are in full text in Open Access. Founded in 2009, Hypotheses today hosts 600 blogs fed by a community of 1200 bloggers from around the world. All content is in Open Access. Hypotheses hosts various types of blog: research, fieldwork, seminar, and monitoring blogs etc. The platform itself is organized into linguistic subdivisions which correspond to French-speaking, German-speaking, and Spanish-speaking bloggers.

OpenEdition Books[3], the newest OpenEdition platform, distributes reference books from publishers in the humanities and social sciences. Its aim is to build an international library in the digital humanities, while encouraging publishers to adopt Open Access in the long term. The platform promotes all cultural domains, through

---

1 http://www.librarything.com

2 http://www.openedition.org

3 The online publication is made through the conversion of articles into XML TEI and then into XHTML and allows the viewing of the full text in web browsers. The specific technical quality needed for the publishing of scientific texts is provided by many functions: metadata management, multiple indexes, management of endnotes, automatic table of contents, numbering of paragraphs and attribution of DOI.

all historical periods in most scientific languages. By spring 2014, 1,271 books are available on the OpenEdition Books platform.

On this platform, some journals and blogs are only devoted to book reviews and most of them publishes reviews regularly. If in journals, reviews are pre-classified and can be easily detected according to metadata, this is not the case for scientific blogs. Another difficulty in our building collection task is that neither the journals nor the blogs have hyperlinks between the reviews and the books they talk about. Therefore these links have to be automatically estimated in order to be able to add them to our book/review corpus.

3. A Collection of Scientific Book Reviews

This section describes the learning corpus we collected from the OpenEdition platform.

In contrast with other collections of reviews like movie reviews from popular Web platforms, book reviews written by and for scientists are much more complex to process because they are very often themselves long scientific papers. If it has been shown that movie reviews (content and quantity) are influenced by movie sales (Duan & Whinston, 2008), reviewing scientific books is one of the activities of the researchers. Hence, the reviews are a very valuable resource for scholarly work in a digital library context. Here is an extract of such a review:

```
<title level="a" type="main">
<bibl> <author> <forename> Peter </forename> <surname>
Bing </surname> </author>,
    <hi rend="italic">   <title   xml:id="title1"
    level="m">The Scroll and the Marble. Studies in
    Reading    and    Reception    in    Hellenistic
Poetry</title> /    <title         xml:id="title2"
level="m">The Well-Read Muse. Present  and  Past  in
Callimachus and the Hellenistic   Poets</title></hi>
</bibl></title>
<author>
    <persName>   <forename>   Benjamin   </forename>
<surname> Acosta-Hughes </surname> </persName>
</author>
[...]
<body> The Scroll and the Marble. Studies in Reading
and Reception in Hellenistic Poetry. [...] It is not
standard for the articles of a not so senior scholar in
my  field  to  be  published  together,  indeed this is
usually something that would happen at the conclusion
of a career, but this particular case is different.
Bing is the author of some of the most seminal articles
in  Hellenistic  poetry  scholarship,  particularly  on
epigram,  and  in  the  recent  "boom"  in  epigram
scholarship,   occasioned   in   large   part   by   the
publication   of   the   new   epigrams   attributed   to
Posidippus,  scholars  have  turned  in  particular  to
Bing's  work  as  a  hermeneutic  guide  to  the  genre of
literary epigram. All of the articles in this volume
are excellent, some are seminal work.
[...]
```

This first data collection is subdivided in two parts: a set of reviews from OpenEdition that are publicly available online and a set of Web URLs that correspond to Web reviews of OpenEdition books. It can be employed for further research in opinion analysis, for recommendation or for book search systems.

In order to train classifiers for automatically classifying documents as reviews/non reviews, we manually extracted a French corpus from the Revues.org platform.

3.1. Training and Test Corpora

We built a training corpus of documents in French, pre-classified into two categories, the first one contains a set of 498 reviews of books and are extracted from the journals of the OpenEdition's Revues.org platform and the second, a set of 280 documents that are not reviews but other scientific papers which can be scientific articles, editorials, call for papers or full-text books.

We constructed a test corpus from the Revues.org platform. The purpose of testing on this corpus is to know how much the classifiers are efficient for identifying new RVs. It is composed of 200 RVs and 100 s.

3.2. Automatic Identification of Reviews

3.2.1. Basic Textual Features

A major challenge for the textual documents classification problem is the representation of a document. We used the bag-of-words representation, the simplest but most widely used approach. In this case, a document is represented by a vector of word counts. For the basic configuration, the order of appearance of the words in the documents is ignored, all XML TEI tags are removed and the individual words constitute the classification features. In our case, the documents more than 160,000 different words. We used different classification approaches to tackle this binary genre text classification problem.

3.2.2. Naive Bayes (NB) and Naive Bayes Multinomial (NBM)

Naive Bayesian method is one of the popular techniques for text classification (McCallum & Nigam, 1998; Lewis & Ringuette, 1994). Given a set of training documents D, each document is considered an ordered list of words. We use $w_k(d_i)$ to denote the word in position k of document d_i, where each word is from the vocabulary $V = < w_1, w_2, …, w_{|v|} >$. The vocabulary is the set of all words we consider for classification. Let $C = \{c_1, c_2, …, c_{|C|}\}$ be the classes (in our case, $C = \{RV, NonRV\}$ and the classification is binary). In order to perform classification, we need to estimate the posterior probability $P[c_j \mid d_i]$. Assuming that the probabilities of the words are independent:

$$P[c_j|d_i] = \frac{P[c_j]\prod_{k=1}^{|d_i|} P[w_k(d_i)|c_j]}{\sum_{r=1}^{|C|} P[c_r]\prod_{k=1}^{|d_i|} P[w_k(d_i)|c_r]} \quad (1)$$

by employing Laplacian smoothing:

$$P[w_t|c_j] = \frac{1+\sum_{i=1}^{|D|} N(w_t,d_i)P(c_j|d_i)}{|V|+\sum_{s=1}^{|V|}\sum_{i=1}^{|D|} N(w_s,d_i)P(c_j|d_i)} \quad (2)$$

where $N(w_t, d_i)$ is the frequency of the words w_t in document d_i. The class with the highest $P[c_j|d_i]$ is assigned as the class of the document. (Liu et al., 2002)

3.2.3. Support Vector Machines (SVM)

SVM designates a learning approach introduced by Vapnik in 1995 for solving two-class pattern recognition

problem (Vladimir et al., 1995). The SVM method is based on the *Structural Risk Minimization* principle (Vladimir et al., 1995) from computational learning theory. In their basic form, SVMs learn linear threshold function. Nevertheless, by a simple "plug-in" of an appropriate kernel function, they can be used to learn linear classifiers, radial basic function (RBF) networks, and three-layer sigmoid neural nets. (Joachims et al., 1998)

3.2.4. Decision Trees (DT)

A decision tree is a tree in which each branch node represents a choice between a number of alternatives, and each leaf node represents a decision. for each attribute from the training set the training algorithm finds the normalized information gain ratio (difference in entropy) and creates a decision node that most effectively splits the data , then it repeats the same operation recursively on the remaining samples obtained from the previous splitting (Quinlan, 1993).

3.3. Expanding the collection of long reviews by means of automatic classification

Based on the structured documents (XML-TEI), different techniques can be applied for automatic classification. One of the most popular machine learning method tackling this problem is the Multinomial Naive Bayes model. Faced with the great variety of document styles present on the two platforms and the multitude of languages, we have tested several approaches. The hypothesis is that classifying texts as reviews can be accomplished by taking into account the lexicon (a common lexicon for the majority of the reviews) of the book. This differs from classical customer review analysis (Hu & Bing, 2004) in the sense that no clear "product feature" can be given as an input to detect and to analyse reviews. In the following section, we summarize the experiments and the results we obtained.

3.4. Classification Results on OpenEdition.org

The objective of this study is to establish a methodology to well identify RV of book in the Revues.org platform. After removing all XML TEI tags present in the documents   and indexing the collection using the "*bag-of-words*" approach.

For the construction of the Naive Bayes Multinomial (NBM), SVM and DT models, we used Weka[4] and report hereafter the results we obtained with the default parameters. The automatic classification result is evaluated using precision, recall and F_1-measure for each class (RV, NonRV).

	RV			NonRV		
	Recall	*Precision*	*F_1-measure*	*Recall*	*Precision*	*F_1-measure*
NBM	91%	96%	93,43%	81%	64%	71,5%
NB	93%	85,3%	89%	68%	82,9%	74,7%
SVM (RBF)	**98,5%**	94,3%	96,3%	88%	**96,7%**	92,1%
SVM (Linear Kernel)	98%	**97%**	**97,5%**	**94%**	95,9%	**94,9%**
DT	94,5%	93,1%	93,8%	86%	88,7%	87,3%

Table 1. Evaluation of classification performance for RV and NonRV classes using Revues.org collection test

We compared the performance of the different learning methods commonly used for text classification. Each method represents a different machine learning approach: density estimation using a NB/NBM classifier, the SVM method using RBF and linear kernels and the Decision Tree/rule learner.

The first line of Table 1 shows the results of the 10-fold cross validation on the training set using NBM algorithm and the other lines show the results of the NB, SVM (RBF/linear kernels) and DT classifiers on the Revues.org test corpus.

We observe that on the Revues.org test collection the SVM classifiers perform best among the other classifiers. The SVM with RBF and linear kernels classifiers have identified respectively 197,198/200 RV and 88,94/100 NonRV, the DT classifier was able to detect 189/200 RV and 86/100 NonRV, whereas the NB classifier, 186/200 RV and 71/100 NonRV were correctly predicted.

When analysing the documents wrongly classified, we noticed that the classifiers fail in the case of documents having some citations in other languages than French.

This experiments have allowed to find out what method is promising for well identifying RV of books. However, we choose the SVM (with the linear kernel) method to detect the RV in all OpenEdition platforms because it tackles many problems like the high dimensional input space, the multitude of irrelevant features existing in the bag of words and the case when the document classification problem is linearly separable.

4. Searching for Book Reviews on the Web

Our second purpose is to build a collection of reviews from the Web and not only from the OpenEdition platforms by finding on-line reviews of each OpenEdition book. The main problem holds in the absence of obvious features which make us capable to determine whether a Web page is a review or not.

For this purpose, we chose about 127 books classified in 3 categories: 49 books about environment, 63 about sociology and 15 about information science. Then, we queried Google Web Search by concatenating the title of the book (exact phrase search) and the name of all its authors, we added the keyword "*review*" for the English books but not for the French ones because we remarked that it could increase in the most of cases the possibility of delaying the appearance of the review in the ranking result. We then downloaded the first 20 Web pages for each book and performed a manual annotation. This annotation was performed by one expert in this domain, we didn't ask another experts to give their evaluation because of the simplicity of determining if a web page is review or not and this corpus is still a test corpus for evaluating the difficulty of this task in such context.

2000 Web pages were manually evaluated and classified into several classes such as: review, advertisement,

4 http://www.cs.waikato.ac.nz/ml/weka/

interview, bibliography or even 'access denied'. We filtered out all advertising Web pages which never contain any good review of book. In the 600 remaining pages we used for classification learning, 97 are reviews of books.

4.1. Feature selection

For each retrieved Web page, we kept its URL and its HTML content. Three feature sets have been considered:

- A. We deleted all HTML and Java Scripts tags in the Web pages.
- B. We employed *Boilerpipe* (Kohlschütter et al., 2010) for extracting the actual content of the Web pages[5].
- C. We used only the words in the URLs as features.

4.2. Results

We tested four different classification approaches (Naive Bayes, Naive Bayes Multinomial, SVM, Decision Tree). Table 2 presents the precision, recall and F-measure for the two classes Review (RV) and Not-Review (NotRV).

We remark that SVM classifier gives the best result with the feature set B. The F-measure of non reviews is always better than F-measure of reviews. For the purpose of building a collection of reviews automatically, the scores of the RV class is the more important.

	Fea ture Set	RV			NonRV		
		Recall	Precisio n	F$_1$-meas ure	Recall	Precisi on	F$_1$-meas ure
NBM	A	0.546	0.390	**0.455**	0.835	0.905	**0.869**
	B	0.268	0.356	0.306	0.907	0.865	0.885
	C	0.534	0.362	0.431	0.830	0.908	0.867
NB	A	0.588	0.442	0.504	0.857	0.915	0.885
	B	0.608	0.454	**0.520**	0.859	0.919	**0.888**
	C	0.625	0.444	**0.519**	0.859	0.927	**0.891**
SVM (Linear)	A	0.485	0.618	0.543	0.942	0.905	0.923
	B	**0.629**	**0.663**	**0.646**	**0.938**	**0.929**	**0.934**
	C	0.455	0.597	0.516	0.945	0.906	0.925
DT	A	0.515	0.641	**0.571**	0.944	0.910	**0.927**
	B	0.495	0.632	0.555	0.944	0.906	0.925
	C	0.420	0.552	0.477	0.939	0.900	0.919

Table 2. Evaluation of classification performance for RV and classes in web collection using cross-validation and different approaches: Naive-Bayes Multinomial (NBM), Naive-B

5. Perspectives: Linking Reviews to Books with BILBO

Once a document has been classified as a long scientific review, we have to link it with an identifier of the book reviewed. This process is not straightforward as a lot of ambiguities can occur.

Our reference parsing software BILBO (Kim et al.,2011) is employed in order to detect bibliographic references in the reviews and to obtain the DOI of the corresponding book via CrossRef[6] API if such an identifier exists. BILBO is constructed with our own annotated corpora from Digital Humanities articles from the OpenEdition Revues.org platform. The robustness of BILBO (Kim et al., 2012a), based on linear-chain conditional random fields, allows largely a language independent tagging. Our preliminary results seem to show that the quality of linking references to books via BILBO and CrossRef is good enough to be deployed (> 80% of the links are correct).

6. Conclusion

We have presented two collections of book reviews: the first one was extracted from OpenEdition Platform in XML/TEI format, these reviews, along with metadata, can be freely downloaded (OAI-PMH) and can be used for further analysis (opinion mining, recommendation...). The second collection was collected from the Web using Google and a generated query. These reviews, along with metadata, can be freely downloaded (please contact the authors) and can be used for further analysis (opinion mining, recommendation...). We presented the classification results we obtained by means of classical machine learning approaches for text classification that allowed to extend the collections of reviews.

7. Acknowledgements

8. References

Duan W., Gu B., Whinston A.B. (2008). Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems*, vol. 45, no. 4, p. 1007-1016.

Hu M., Bing L. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining,* ACM.

Joachims T. (1998). Text Categorization with Support Vector Machines: Learning with Many Relevant Features. *in Proceedings of European Conference on Machine Learning (ECML-98)',* pp. 137-142.

Kazai G. (2011). Overview of the INEX 2010 book track: Scaling up the evaluation using crowdsourcing. *Comparative Evaluation of Focused Retrieval*. Springer Berlin Heidelberg, pp. 98-117.

Kim Y.M., Bellot P., Faath E., Dacos M. (2011). Automatic Annotation of Bibliographical References in Digital Humanities Books, Articles and Blogs. *BooksOnline 2011 at CIKM 2011,* Microsoft Research Ed., ACM Press.

---

5 The number of real pages decreased to 576 because of some empty documents generated by Boilerpipe (88 reviews and 488 not-reviews).

---

6 http://www.crossref.org

Kim Y.M., Bellot P., Tavernier J., Faath E., Dacos M. (2012a). Evaluation of BILBO Reference Parsing in Digital Humanities via a Comparison of Different Tools. *ACM 12th Symposium on Document Engineering DocEng'12*.

Kim Y.M., Bellot P., Faath E., Dacos M. (2012b). Annotated Bibliographical Reference Corpora in Digital Humanities. *8th international conference on Language Resources and Evaluation (LREC)*, full paper.

Kohlschütter C., Fankhauser P., Nejdl W. (2010). Boilerplate detection using shallow text features. *Proceedings of the third ACM international conference on Web search and data mining*, New York, USA, pp. 441-450.

Liu B., Lee W.S., Yu, P.S., Li X. (2002), Partially Supervised Classification of Text Documents, *in ICML'02: Proceedings of the Nineteenth International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA , pp. 387-394.

Lewis D.D., Ringuette M. (1994), Comparision of two learning algorithms for text categorization. *Symposium on Document Analysis and Information Retrieval*.

McCallum A., Nigam K. (1998), A comparison of event models for Naive Bayes text classification.

Quinlan R. (1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers , San Mateo, CA .

Vapnik V.N. (1995), *The Nature of Statistical Learning Theory*, Springer .

Appendices

Example of French Review of Book from Revues.org Platform:

[…]

```
<body>
    <p>L'ouvrage de A.F. Robertson peut être perçu
```
comme un avatar de ces démarches régénératrices d'une anthropologie prétendue moribonde et pour laquelle la cure oscille entre la recherche de nouveaux champs d'étude et la "revisitation" d'objets anciens. L'auteur opte explicitement pour cette première approche puisqu'il se propose d'étudier le développement planifié dans une perspective anthropologique "comme un système symbolique C..), un ensemble de catégories visant à manipuler les gens et les ressources dans l'espace et le temps" (p.87). L'ouvrage débute par un long chapitre dans lequel est exposée l'histoire des idées et des politiques de développement.</p>

    <p>Dans le chapitre consacré aux organisations il présente les éléments d'une problématique qui appréhende les bureaucraties (instruments de l'application des politiques de développement) et les communautés locales (champs d'application des politiques) comme des arènes politiques, des lieux de compétition entre acteurs sociaux. L'apport le plus

original de l'ouvrage est sûrement cette idée de la confrontation entre deux formes de planification : celle de l'État et celle du "peuple", comme nœud de la contradiction inhérente à l'implantation de projets. Il est regrettable que l'auteur n'investigue pas davantage les formes locales de planification notamment dans l'étude du cas malais qu'il présente. Malgré la faiblesse de cet aspect de la démonstration et le flou qui entoure parfois les conditions d'application de l'anthropologie au contexte de développement, le texte de A. F. Robertson est fondateur en ce qu'il explore de manière exhaustive et pertinente un domaine délaissé par l'académisme anthropologique.</p>
    </body>

[…]
```

Example of French Review of Book from the Web:

[…]

1   Cet ouvrage préfacé par la géographe et sociologue, Jacqueline Coutras, et dirigé par Sylvette Denèfle, sociologue à l'université François-Rabelais, est la restitution du colloque interdisciplinaire « Femmes et villes » qui s'est déroulé au Palais des Congrès de Tours les 8 et 9 mars 2002, en regroupant une soixantaine de participant(es). « Les femmes dans la ville » (pp. 25-179), « Rien n'est jamais gagné (pp. 180-356) et « Rien n'est jamais perdu » (pp. 357-539), sont les trois grandes parties qui structurent le livre et regroupent trente-cinq contributions. Le cœur de la réflexion est axé sur des lectures croisées d'études urbaines et de travaux concernant le genre, et sur les usages de la ville et leurs conséquences par rapport aux évolutions sociales.

2   Femmes et villes offre donc une multiplicité d'approches théoriques, méthodologiques et thématiques, fort intéressantes, mais du coup impossibles à résumer. Ceci posé, les pratiques féminines et masculines - qui semblent « normales et naturelles » (Denèfle, p. 15) aux usagers d'un espace urbain - confrontées aux pratiques urbaines pouvant modifier, conforter ou transformer les normes sociales de sexe, sont un des dénominateurs communs de l'ouvrage et dévoilent « l'extrême complexité de l'intrication des systèmes normatifs » (ibid., p. 16). De même, la plupart des contributions mettent en perspective l'intérêt à prendre en compte la dimension du genre dans les comportements urbains, notamment du côté de l'action politique pour repenser l'aménagement du territoire en fonction des usages sexués de la ville. Car ce qui est illustré ici, par des études de cas ou des enquêtes, est un large panorama des marques de l'assignation féminine dans les villes selon différents espaces : espace domestique (Bauhardt, pp. 41-49 ; Corbeau, pp. 167-179 ; Diaz, Plat, Pochet pp. 135-153 ; Dussuet, pp. 359-378 ; Robin, pp. 65-75), espace du partage des pouvoirs (Bigoteau, Roux, pp. 325337 ; Lepinard, pp. 339-356 ; Naceur, pp. 239-247 ; Nigaud, Ripoll, pp. 489-499 ;  Saïdi-Sharouz,  pp. 443-452),  espace

professionnel (Dupont-Kerlan, Fontaine, pp. 123-133 ; Gubin, Piette, pp. 379-398 ; Le Sager Diouf, pp. 25-39 ; Monqid, pp. 471-487), espace d'expression de lutte et d'indépendance (Bastiat, pp. 501511 ; Bourdarias, pp. 51 -63 ; Dermen-jian, Loiseau, pp. 101-112 ; Flahault, pp. 399-408 ; Hancock, pp. 409-417 ; Sniter pp. 529-539), et surtout espace de contraintes (Bahi, pp. 183-193 ; Camus, pp. 225-238 ; Diaz, Plat, Pochet, pp. 135153 ; Dris, pp. 249-264 ; Maillochon, pp. 207-223 ; Pennec, pp. 87-100 ; Thalineau, pp. 113-121 ; Vella, pp. 265-282 ; Zeilinger pp. 196-205).

3    Parmi d'autres aspects, c'est ce dernier qui interroge, car il semblerait bien que les villes d'Europe ou d'ailleurs (Afrique subsaharienne, Algérie, Inde du Sud, Iran,

4    Mali, Maroc, Mexique, Tunisie) demeurent des lieux difficiles pour les femmes, car ils sont sources de contraintes récurrentes et liés au fait que les femmes sont tout simplement des femmes. Globalement, ces contraintes sont d'ordre matériel et symbolique et s'entremêlent. Dans les deux cas de figure, les auteurs relèvent les modes de déplacement, les horaires (de travail et/ ou de sorties domestiques ou autres), le choix de l'habitat, les lieux fréquentés, la mobilité, les prises de responsabilités (professionnelles, politiques, militantes), en lien avec la charge familiale, la précarité du statut de la femme (âge, niveau d'études, choix de sa sexualité, droits coutumiers et religieux, ségrégation du genre), les risques encourus (agressions, viols) à l'extérieur (rue, quartiers, transports en commun, bars, etc.), et selon des temporalités particulières (soirée, nuit). Pour exemple, l'enquête coordonnée par l'Institut de démographie de l'université Paris 1 (Maryse Jaspard, dir, Les violences envers les femmes. Une enquête nationale, rapport coordonné par l'Institut de démographie de l'université Paris 1, Paris, Éd. La Documentation française, 2003), effectuée au début de l'année 2000 sur le territoire français, est résumée ici par Florence Maillochon (pp. 207-223). À partir d'un échantillon national de 6 970 femmes âgées de 20 à 59 ans et résidant en métropole, l'étude tend à démontrer que « les grandes zones urbanisées sont particulièrement touchées par toutes les formes de violence [...], que les femmes qui habitent dans des immeubles collectifs en ville sont plus fréquemment victimes que celles qui habitent une maison » (Maillochon : 218), et chiffre à 0,1 % les femmes victimes d'agressions sexuelles (attouchements sexuels, viols ou tentative de viol) et à 1,7 %, les agressions physiques (vols, coups, menaces avec arme). Cette présentation chiffrée de la violence exercée contre les femmes par des hommes propose une lecture nouvelle de ces violences, en ce sens que même si les actes ne sont pas minimisés en tant que tels, ils vont à contre-courant de ce qui est régulièrement déversé dans certains espaces associatifs militants (e.g. Ni Putes

Ni Soumises) et relayé par quelques médias à l'image de TF 1 et Canal+. Ce que dénoncent ces discours d'experts (Camus, pp. 225238 ; Maillochon, pp. 207-223) qui, outre le fait qu'ils rappellent que « les médias désignent habituellement les jeunes issus des banlieues comme principaux coupables de [la] montée du sentiment d'insécurité » (Camus : 225), auprès des femmes surtout, tentent d'expliquer pourquoi l'espace urbain favorise ces stéréotypes. Mais peut-être faut-il aussi reconsidérer toutes ces analyses au prisme d'éléments non comptabilisés dans les études et les enquêtes des chercheurs(ses). Car certains sévices commis ne sont ni dits ni connus, et ne peuvent donc être pris en compte dans les statistiques officielles.

5    Dès lors, si dans leurs pratiques, les femmes excluent certains espaces ou en sont exclues, elles s'en approprient d'autres par défaut. Ce sont ces parcours au féminin que propose cet ouvrage, auxquels le lecteur est invité à découvrir les méandres comme les squares, les jardins publics, les bas des immeubles (Dussuet, pp. 359-378), les places publiques, les commerces de proximité, les galeries marchandes, les marchés, les cimetières, les librairies, les universités du temps libre (Pennec, pp. 87-100 ; voir aussi Diaz, Plat, Pochet, pp. 135-153), les lieux associatifs (Demerjian, Loiseau, pp. 101-112), et enfin les lieux de mémoire (musées, mémoriaux dont les monuments aux morts et les statues : voir à ce sujet, Snitter, pp. 529- 539). Des parcours qui font dire, à juste titre, à Sylvette Denèfle (p. 20), qu'il serait temps pour ceux qui en ont la charge (élus, architectes, etc.) de prendre en compte les comportements des femmes et des hommes dans les villes afin de « servir la qualité de la vie de tout un chacun et [...] d'adoucir les injustices » relatives aux modèles identitaires de sexe. Car ainsi que le mettent en lumière la plupart des contributions de Femmes et villes, les femmes ont un rôle essentiel dans les espaces qui forgent la ville, parce qu'elles ont des responsabilités étagées (domestiques, professionnelles, militantes, culturelles), dont il n'est plus possible d'ignorer l'importance, et surtout les inégalités sociospatiales.