Image Retrieval:
A Comparative Study on the
Influence of Indexing Vocabularies

Elaine Ménard

Assistant Professor, McGill University, School of Information Studies,
3459 McTavish, MS72C, Montréal Quebec Canada H3A1Y1 <elaine.menard@mcgill.ca>

Elaine Ménard is Assistant Professor at the School of Information Studies, McGill University (Montreal, Canada). Professor Ménard has completed her Ph.D. at the École de bibliothéconomie et des sciences de l'information, Université de Montréal. She holds a Master in information studies, a Master in translation and a Bachelor in translation from the same university. Her main research interests deal with indexing, cataloguing, classification, information retrieval, cross-language information retrieval systems, and multimedia documents.

ABSTRACT: This paper reports on a research project that compared two different approaches for the indexing of ordinary images representing common objects: traditional indexing with controlled vocabulary and free indexing with uncontrolled vocabulary. We also compared image retrieval within two contexts: a monolingual context where the language of the query is the same as the indexing language and, secondly, a multilingual context where the language of the query is different from the indexing language. As a means of comparison in evaluating the performance of each indexing form, a simulation of the retrieval process involving 30 images was performed with 60 participants. A questionnaire was also submitted to participants in order to gather information with regard to the retrieval process and performance. The results of the retrieval simulation confirm that the retrieval is more effective and more satisfactory for the searcher when the images are indexed with the approach combining the controlled and uncontrolled vocabularies. The results also indicate that the indexing approach with controlled vocabulary is more efficient (queries needed to retrieve an image) than the uncontrolled vocabulary indexing approach. However, no significant differences in terms of temporal efficiency (time required to retrieve an image) was observed. Finally, the comparison of the two linguistic contexts reveal that the retrieval is more effective and more efficient (queries needed to retrieve an image) in the monolingual context rather than the multilingual context. Furthermore, image searchers are more satisfied when the retrieval is done in a monolingual context rather than a multilingual context.

1.0 Introduction

For the past several years, the increasing development of databases and collections consisting of various types of text or multimedia documents has posed new challenges for the retrieval process. Searching for images is one of the most popular categories of searches on the Internet (Tjondronegoro and Spink 2008, 340–41). Although in recent years it has become easier to identify images with, among others, the use of search engines specifically designed for this purpose, image retrieval still presents a significant degree of difficulty. "Images are notoriously difficult to retrieve with accuracy, as is evident to anyone who has searched for im-

ages on the World Wide Web" (Harpring 2002, 20). Among the many types of images available on the Internet, the ordinary image (i.e., non-artistic) occupies an important place in the user's search of the Web. By ordinary images, we mean images of common objects used daily, such as a television, a coffee mug, a tennis racket, etc. The growth of the Internet has highlighted the pressing need to develop tools for the description of images in order to facilitate their retrieval, as we find them in most resources: personal Web pages, museum collections, digital libraries, commercial products and services catalogues, government information, and so on. In general, we consider the image as *language independent* (Gonzalo et al.

2006). However, image indexing gives the image a language status similar to any text document, which may affect retrieval. Unfortunately, language barriers still prevent users from accessing information (Marlow et al. 2007), including visual information.

Since the beginning of the 19th century, cataloguing, classification and indexing processes have been primarily devoted to textual documents. However, obtaining access to multimedia material presently raises a lot interest. According to Guinchat and Menou (1990), the indexing process of visual material poses particular problems because of its own nature and mode of consultation. Clearly, the choice of suitable indexing terms for image description is an enduring issue whether it is in a monolingual or multilingual retrieval context. This research project examines what occurs at the retrieval stage when an image is indexed according to each of the two following approaches: traditional image indexing, which recommends the use of controlled vocabularies, that is, the indexing terms are chosen from an artificial language whose main function is to generate formal document representation; or free image indexing, which uses uncontrolled vocabulary, that is, the indexing terms being used to describe the image do not follow any pre-established rules. While these two indexing approaches show common characteristics, there are also differences that may influence image retrieval. This research aims to establish whether one of these indexing approaches surpasses the other in terms of effectiveness, efficiency and satisfaction of the searchers for image retrieval in a multilingual context.

2.0 Related studies—an overview

2.1 Meaning of images

Over the years, image indexing has been the focus of several key studies. Panofsky (1955) identified three levels of meaning in works of art: pre-iconography for the primary or natural subject matter, iconography for the secondary subject, and iconology for the intrinsic content of the work. A few years later, Markey (1983; 1988) applied these levels for the identification of themes or concepts illustrated in images. Shatford (1986) defined three groups of attributes, which are "specific of," "generic of" and "about," corresponding to the three classes in Panofsky's (1955) classification. In the same manner, Krause (1988) divided the information contained in an image into "hard indexing" (what can be observed in a picture) and "soft indexing" (subjective meaning

and the personal response that it evokes). Layne (1994) suggested that an image could be *of* and *about* something. While *ofness* is mainly concrete and objective, *aboutness* is more abstract and subjective. The various types of attributes of images and the terms used to describe them have been the subject of several studies. In order to determine what the appropriate access points are, image indexing and retrieval have been studied extensively over the years (Turner 1993; Turner 1994; Ornager 1996; Armitage and Enser 1997; Jörgensen 1998; Chen 2001; Goodrum and Spink 2001; Choi and Rasmussen 2002; Choi and Rasmussen 2003; Greisdorf and O'Connor 2008). When examining the different approaches to indexing an image, it is clear that the majority of the images are indexed on a minimum and often offer a single point of access (Jörgensen 1998, 162). Several studies (Besser and Snow 1990; Roddy 1991) show that most of the indexing approaches are not suitable for picture researchers, while other studies (Ohlgren 1980; Krause 1988; Turner 1993) emphasize the fact that the main problem concerning image retrieval is the approach chosen for the indexing process.

2.2 Controlled vocabulary indexing

The controlled vocabulary is a language with its own terms, syntax and semantics (Wellisch 1995, 214). The main advantage of vocabulary control is to promote consistency in indexing and to increase, through a system of referrals, the probability of matching the words chosen by the indexer and those of the researcher (Jörgensen 2003, 105; Arsenault 2006, 141). The use of controlled vocabularies offers many advantages for retrieval, browsing and interoperability while also aiming to facilitate the indexing process. However, controlled vocabularies also present some weaknesses, the main one being that they represent the concepts in an artificial way (Macgregor and McCulloch 2006, 294). The indexing terms offered by controlled vocabularies often have very few relevant connections with the terms used by individuals in the formulation of their queries (Furnas et al. 1987). Furthermore, controlled vocabularies have a tendency to become quickly outdated, therefore constituting another disadvantage since neologisms take considerable time before they are integrated in the various controlled vocabularies (Lancaster 2003, 255). Moreover, the use of these vocabularies remains a complex task for the majority of indexers (Goodrum 2000, 66). Finally, most controlled vocabularies suggested by metadata schemas and commonly used for image indexing only exist in

English. Consequently, an indexer having little knowledge of the English language and wishing to use these vocabularies will face a major linguistic problem unless an effective translation mechanism is offered to facilitate their use.

2.3 Uncontrolled vocabulary indexing

Unlike indexing with controlled vocabulary, indexing with uncontrolled vocabulary is not governed by any predetermined rules regarding the choice of words used. The indexing terms are extracted from the natural language (Cleveland and Cleveland 2001, 35; Chu 2003, 47). In recent years, we observed the advent of a new phenomenon perfectly illustrating indexing with uncontrolled vocabulary. It is the collaborative indexing whose main characteristic is to use uncontrolled vocabulary for indexing documents. Collaborative indexing, also known as "social tagging," has emerged as a means of organizing information resources on the Web and is contradictory to the traditional philosophy of controlled vocabularies (Macgregor and McCulloch 2006, 292). The primary objective of collaborative indexing is to enable an individual to share the indexing terms for a set of resources (Macgregor and McCulloch 2006, 294). The main advantage of this indexing approach, according to Marlow et al. (2006), is to solve the problem of vocabulary, as perceived by Furnas et al. (1987) when different people use different terms to describe similar concepts. In other words, this system of social tagging has the potential to improve the conventional indexing with controlled vocabulary which generates many problems during the retrieval process (Marlow et al. 2006). Collaborative indexing offers the benefit of providing additional access points that are often very different from the conventional ones (Kipp 2007). The tags produced with collaborative indexing may take any form, according to the user's inspiration. In addition, the tags can be constructed from a single language or may combine several languages (Marlow et al. 2006). The popularity of social indexing increased quickly and is no longer for textual documents only. For example, online image sharing systems also use this type of indexing (Le Deuff 2006, 67; Marlow et al. 2006; Mathes 2006; Angus et al. 2008; Rorissa 2008). These systems allow users to index and share their own images. People can share images on a theme and create an exchange community based on their labels (tags) or keywords. The system Flickr® offers users the possibility to index their own images and make them public, that is, the

images can be viewed by all or only a group of individuals chosen by the user of the system.

Given its recent nature, little is known about collaborative indexing. Golder and Huberman (2005) analyzed the structure of sharing systems using this type of indexing. The results of this study indicate a great variability in the keywords assigned, ranging from the very general (e.g., dog) to the most specific (e.g., Fido sleeping after playing all afternoon with Mary and John). Some individuals use many keywords, others only a few to describe an image according to how they perceive it. However, in cases where several words are used, it seems that this procedure is for personal purposes rather than public. Indeed, the meaning of some keywords often remains ambiguous for the larger number of users (e.g., mylife, strangeday_14nov07). The information added by many users of the system is useful only to the extent that all users understand the content in the same way and if there is an overlap in their choice of categories of keywords. Nevertheless, the collaborative indexing, even if it largely depends on the ability of individual indexers, can be of great value, provided it is used across the Web and not just by a few individuals (Golder and Huberman 2005).

2.4 Dilemma

Several elements emerge from this review of literature. First, the ordinary image, contrary to the artistic image, is not really described in the literature. Second, this review highlights that we know very little about the influence of the vocabulary used for indexing images on its retrieval in a multilingual context. Third, analysis of the literature reveals that there are two approaches to indexing images, one using the more traditional controlled vocabulary and the other advocating the use of uncontrolled vocabulary. The latter approach is attracting increased interest with the phenomenon of collaborative indexing.

The use of either controlled or uncontrolled vocabulary raises a number of difficulties for the indexing process that impact image retrieval. The choice between the two approaches is extensively discussed in the literature surrounding this study where scholars explore which is the best vocabulary to be used for the indexing process. Some results indicate that uncontrolled vocabularies offer a better retrieval performance than the controlled vocabularies (Savoy 2005), a point of view shared by Rao Muddamalle (1998). Other scholars (Markey et al. 1980; Kamps 2004) consider that the controlled vocabularies im-

Knowl. Org. 36(2009)No.4
E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

203

prove the precision of results at the time of retrieval. It is clear, nevertheless, that many searchers recognize the advantages of both indexing approaches according to the circumstances (Arsenault 2006). However, it seems that some authors have increasingly recognized the usefulness of combining the two approaches when indexing images (Matusiak 2006; Macgregor and McCulloch 2006; Enser et al. 2007).

As noted in the work of several researchers (Goodrum and Spink 2001; Goodrum et al. 2003; Jörgensen and Jörgensen 2005; Tjondronegoro and Spink 2008), we see an evolution in how queries are formulated for image retrieval. For example, queries consisting of a single term are less frequent than before. This transformation point forces us to reconsider the way in which the image must be indexed and whether the controlled vocabulary traditionally employed for image indexing is well suited to this particular type of document. The study of the Web image searcher's behaviour is therefore necessary in order to improve retrieval systems. In the same way, it is essential to carry on the study of retrieval strategies of image searchers in order to establish better indexing and, more specifically, a vocabulary more adapted to the needs and behaviours of the image searchers.

With this in mind, several questions are raised about image retrieval in a multilingual context. We can consider whether the choice of the vocabulary used for image indexing will influence the retrieval results in a multilingual context and, if such is the case, what this influence is. Moreover, we can investigate which image indexing approach is preferable to implement to facilitate the retrieval of images in a multilingual context. As few studies address these questions, finding the image indexing process that allows for effective and efficient retrieval with a high degree of searcher satisfaction in a multilingual context is a vital step requiring more attention.

3.0 Methods

3.1 Image Database

As a first step of this exploratory study, a database containing ordinary images extracted from a commercial online catalogue representing common objects was constructed. A total of 3,950 images were selected and then indexed according to four indexing forms: French and English controlled vocabularies, and French and English uncontrolled vocabularies. For reasons of availability and economy, indexers involved in the project had no professional experience

in indexing. The indexing with uncontrolled vocabularies was performed by two different indexers (one native French speaker and one native English speaker) while the controlled vocabulary indexing (French and English) was produced by an indexer fluent in both languages.

For the indexing process with controlled vocabulary, the indexer was asked to use the *Nouveau dictionnaire visuel multilingue* (Corbeil and Archambault 2003). Even if this dictionary cannot be formally considered an authentic controlled vocabulary, it was chosen for three reasons. First, this dictionary contains words and images of common objects similar to the type of images included in the database. Second, it provides a form of term standardization allowing maximum control on regionalisms, archaisms and anglicisms. Third, this dictionary offers both French and English terms (in addition to Spanish, German and Italian terms), which facilitated the indexing process while maintaining a form of inter-linguistic consistency. The process of indexing with controlled vocabularies proceeded as follows: the indexer was shown the image to be indexed and then located the best corresponding image in the dictionary. The French and English terms associated with the image of the dictionary were then used as indexing terms. Between one and five controlled indexing terms could be assigned to each image.

For the indexing process with uncontrolled vocabularies (French and English), the indexers saw the images and then used their own words as indexing terms. No limit was imposed on the number of indexing terms assigned to one image. At the end of the indexing process, all images were associated to four categories of indexing terms as illustrated in Figure 1.

3.2 Retrieval System

An image retrieval system specifically designed for this experiment and linked to the image database was used for the simulation. This system, called El@ine, was programmed with the language ASP.NET 1.1 and used a MySQL database. The search interface of El@ine included two sections. In the upper part of the display, the images to be retrieved were shown to the participants who could enlarge the images by clicking on them. In the lower portion of the display, there was a search box where the participants typed their French queries when trying to retrieve the images presented to them. No terminological suggestions were made to the participants who had to

204

Knowl. Org. 36(2009)No.4

E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

Examples of images extracted from the database	Controlled Vocabulary		Uncontrolled Vocabulary	
	French Terms	English Terms	French Terms	English Terms
	brodequin de tra-vail	heavy-duty boot	bottes de travail Ko-diak	Kodiak work-boot
	aspirateur-traîneau	cylinder vacuum cleaner	aspirateur Shark	Shark portable vacuum
	chaussure de sport	running shoe	espadrilles Wilson	women's sneaker

Figure 1. *Images and the associated indexing terms*

Image	Original query	Translated query
	bottes	boots
	bottines	boots
	souliers	shoes
	souliers sport	shoes sport
	souliers bruns	shoes brown
	bottes de construction	boots of construction
	aspirateur	vacuum cleaner
	poussiere	dust
	balayeuse	sweeper
	aspirateur portable	vacuum cleaner portable
	aspirateur central	vacuum cleaner exchange
	balayeuse electrique	sweeper electric
	soulier	shoe
	chaussures	shoes
	course	race
	basket	shoe
	souliers course	shoes race
	espadrille	espadrille

Figure 2. *Examples of queries with the translated equivalent*

transpose in their own words the image shown. Once the mapping between the query terms and the index-ing terms was done, the display interface showed again the image to be retrieved to the participant in the upper part of the screen with the resulting images shown in the lower part. If needed, participants could then make another query. For our study, it was de-cided that no text would be displayed with the im-ages. This approach stands in contrast to the conclu-sions of some studies advocating the simultaneous

2005; Clough et al. 2006; Karlgren and Olsson 2006), and also differ from our analysis of several key search engines such as Google Images, Yahoo! Search, Alltheweb and AltaVista revealing that these engines often offer to their users the possibility to visualize keywords with the retrieved images. This decision was taken so that the participants would not be influ-enced by a term they did not think of themselves for the formulation of their queries.

This study compared image retrieval within two contexts: a monolingual context, that is, where the language of the query (French) was the same as the indexing language (French); and a multilingual context, that is, where the language of the query (French) was different from the indexing language (English). For this second context, the French queries of the participants were first translated into English by Babel Fish (Yahoo! 2008), a machine translation system integrated in the retrieval system, and then mapped to one of the English indexing forms. For the purpose of this study, the machine translation of the queries was better suited because the queries for images generally contained only a few words (Chen 2001; Fukumoto 2004; Spink and Jansen 2004; Pu 2005). Hence, the translation could be done more quickly and at a low cost. No verification was made whether the translations obtained were correct or not. If the machine translation system returned no result, the French query terms were retained. The query translation process was done without the knowledge of participants who did not know to which indexing form they were linked. Figure 2 illustrates a few query examples and the equivalent translated queries:

3.3 Participants

For this research, a nonprobability sample was used where the elements of the population are chosen because of the correlation between their characteristics and goals of the research (Fortin 1996, 362). With this type of sample, it is possible to both increase the usefulness of the information and limit the number of subjects (Contandriopoulos et al. 1990, 62). Moreover, it was a voluntary sample since each participant had to take an appointment to participate in the experiment.

In order to strengthen the validity of the results, the simulation of the retrieval was made with 60 participants randomly divided into six independent groups assigned to one of the six indexing forms (the four forms mentioned, in addition to the combination of French controlled and uncontrolled vocabularies and the combination of English controlled and uncontrolled vocabularies). For ethical considerations, our participants were aged 18 and older. In addition, to ensure the homogeneity of the group of participants (Fortin 1996, 34), three selection criteria were defined: participants whose mother tongue is French, undergraduate students of the Université de Montréal and, given the nature of the tasks to perform during the experiment, the participants should have no professional experience in a field involving image indexing and retrieval. Each participant had to meet these minimal inclusion criteria to be selected to perform the retrieval simulation. These criteria were used to control the bias which may come from heterogeneous participants. However, we were aware that the sample size and too much homogeneity could limit the generalization of the statistical results to the single category of participants selected for our research (Fortin et al. 2006, 180).

The recruitment of participants was done using posters displayed in strategic locations at Université de Montréal (display boards, libraries, etc.). A monetary compensation of $20 was allocated to each respondent suitable for the experiment. The retrieval simulation was conducted in a relatively short period, from September 27 to November 13, 2007, to prevent the effect of data contamination.

3.4 Data Collection

During this experiment, performed with the retrieval system El@ine, the participants were shown each of the 30 images, in the same order of presentation. Each participant was randomly associated with one of the six conditions (one indexing form) established for this study. Once the participants were satisfied with the result of the retrieval task, they recorded the retrieved image directly into the retrieval system. Before searching for the next image, the participants were asked to evaluate their degree of satisfaction regarding the retrieval results they obtained after each image retrieval task. The variables used in our study were recorded directly by the retrieval system. Each transaction of the participants was retained in the database, with the date and time (to the nearest second) at which it took place. For each image retrieval, the following variables were recorded:

– The time of the beginning of each retrieval task
– The indexing form
– The queries used by the participant to retrieve an image
– The translation returned by the system if applicable
– The total number of results for each query
– The number of queries used for each image
– The time of the end of each retrieval task
– The image selected by the participant for each retrieval task
– The evaluation of participant satisfaction after each retrieval task

3.5 Data analysis

The recorded data at the time of the retrieval simulation of tracking, as well as responses to the satisfaction questions, have been transcribed into an Excel file. To test the research hypotheses, descriptive statistics were computed and the two-factor analysis of variance was used. Also known under the name of two-way ANOVA, this statistical test is regularly used to determine whether or not the differences observed between the means measured with specific populations are significant. In addition, analysis of variance is used to make comparisons of means between independent groups (Ouellet 1994, 251). Regarding our research, this statistical test was feasible since it included a multiple independent variable (six indexing forms), a single dependent variable (retrieval performance) and a ratio level of measurement of the dependent variable (Sproull 1995, 257). As part of our research, analysis of variance was used to compare the averages obtained for each indexing form and to determine whether there were significant differences. In short, the analysis of variance allowed for the verifying of the statistical significance of differences between averages obtained at the time of the simulation, for effectiveness, temporal and human efficiency, and satisfaction of the image searcher. In addition, to complete the two-way ANOVA test, the Tukey HSD (Honestly Significant Difference) Test was used to compare averages for each pair of indexing approaches. This test allows for the classification of the average processed into subsets and for control over the Type I errors (Howell 2001, 426–27).

3.6 Measures

The quantification of the image retrieval performance of each indexing approach was based on the usability measures recommended by the ISO 9241-11 standard, that is, effectiveness, efficiency and user satisfaction (Association française de normalisation 1998). In general, effectiveness refers to the ability to achieve a given goal, while efficiency rather refers to the ability to perform a given task with minimum time and effort (Brangier and Bracenilla 2003, 47). Traditionally, several indicators can be considered for these two measures. For our research, the measures were defined as such:

– *Effectiveness of image retrieval*: measured by the success rate of retrieval calculated using the retrieved number of images divided by the total number of images to be retrieved.

– *Temporal efficiency of image retrieval*: measured by the average time (in seconds) for each retrieved image.
– *Human efficiency of image retrieval*: measured by the average number of queries used for each retrieved image.
– *Satisfaction of the image searcher*: measured by a scale of participant evaluation with the results obtained for a specific retrieval task. Satisfaction was assessed individually after each retrieval task.

4.0 Findings

The six following indexing forms were compared: French controlled vocabulary (FCV), French uncontrolled vocabulary (FUV), combination of French controlled and uncontrolled vocabularies (FCUV), English controlled vocabulary (ECV), English uncontrolled vocabulary (EUV) and combination of English controlled and uncontrolled vocabularies (ECUV).

4.1 Retrieval effectiveness

For this study, we considered the success rate, that is, the ability to achieve the objective (shown image retrieval), as the main indicator of effectiveness. Figure 3 shows the average proportions of retrieved images and highlights the observed differences of each indexing form.
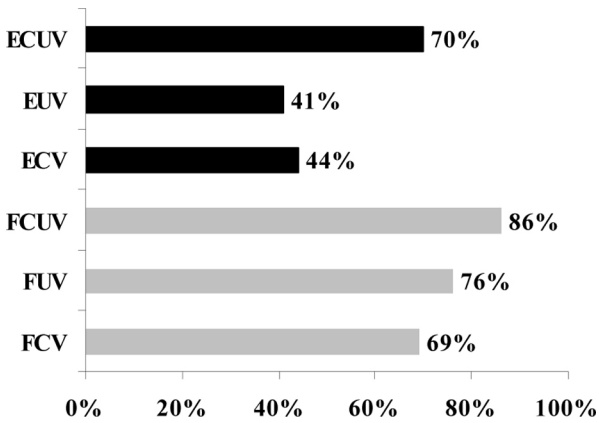

Figure 3. *Effectiveness of image retrieval (% of retrieved images)*

4.1.1 Monolingual retrieval context

We observed that a greater proportion of images were retrieved when they were indexed with the combination of controlled and uncontrolled vocabularies, compared respectively to the controlled vocabulary and the uncontrolled vocabulary approaches. A two-

factor ANOVA test was carried out on these findings to ascertain whether the observed differences were significant. The factors taken into account were the indexing language (two levels: French and English) and the indexing approach (three levels: controlled vocabulary, uncontrolled vocabulary and the combination of controlled and uncontrolled vocabularies). The results indicated that there was no significant interaction between the language and the indexing approach and it was therefore possible to independently examine the effects of the language and the indexing approach on the proportion of retrieved images. Since the two-factor ANOVA test revealed a difference between the average proportions of retrieved images in terms of the indexing approach, the Tukey HSD Test was used to compare averages for each pair of indexing approaches. The differences between the proportions of retrieved images were significant between the combination of controlled and uncontrolled vocabularies approach and the controlled vocabulary approach, and also between the combined approach and the uncontrolled vocabulary approach. However, no significant difference was observed between the controlled and the uncontrolled vocabulary indexing approaches.

4.1.2 Multilingual retrieval context

Our analysis revealed a considerable influence from the indexing language on the effectiveness of retrieval since the difference between the proportions of retrieved images indexed using a French vocabulary (monolingual retrieval context) and images indexed using an English vocabulary (multilingual retrieval context) is significant. Additionally, the combination of French controlled and uncontrolled vo-

cabularies (FCUV) was the most effective indexing form, in terms of retrieved images while the English uncontrolled vocabulary (EUV) was the least effective form.

4.2 Temporal efficiency of retrieval

Regarding efficiency, we distinguished between two forms: temporal efficiency and human efficiency (Brangier and Bracenilla 2003, 50; Ménard 2007). First, the temporal efficiency is measured by the time in seconds, on average, used to retrieve an image. Figure 4 shows the mean average time required to retrieve an image.

4.2.1 Monolingual retrieval context

The results of the two-factor ANOVA test showed that there was no significant effect of the language on the average time required to retrieve an image as the difference between the monolingual and multilingual contexts is not significant. Then we observe that there is an effect of the indexing approach on the average time needed to retrieve an image. However, the Tukey test revealed that the observed differences were not significant between each pair of indexing approaches.

4.2.2 Multilingual retrieval context

The French controlled vocabulary (FCV) was the most efficient indexing form, while the combination of English controlled and uncontrolled vocabularies (ECUV) was the least efficient form. However, the results of our analysis showed that the observed difference in terms of temporal efficiency (time re-
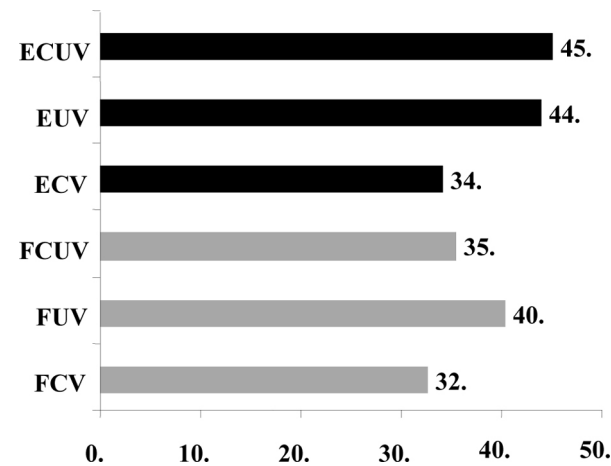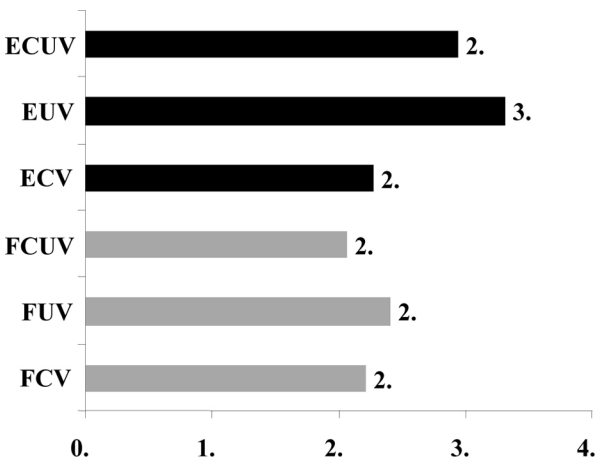

Figure 4. *Temporal efficiency (in seconds)*

Figure 5. *Human efficiency (in queries)*

208

Knowl. Org. 36(2009)No.4

E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

quired to retrieve an image) between the two linguistic contexts was not significant.

4.3 Human efficiency of retrieval

Human efficiency is measured by the average number of queries used to retrieve an image. Figure 5 shows the average number of queries required to retrieve an image and highlights the differences observed.

4.3.1 Monolingual retrieval context

The results of the two-factor ANOVA test on the average number of queries required to retrieve an image revealed that there were no significant interaction between the indexing language and the indexing approach. Once again, the influence of the indexing language and the indexing approach could therefore be studied independently. As indicated by the Tukey Test, the difference between the average number of queries required to retrieve an image was only significant between the controlled and uncontrolled indexing approaches. The indexing with controlled vocabulary was the most efficient in terms of queries required compared to the indexing with uncontrolled vocabulary. No significant difference, however, was observed between the combined and the controlled indexing approaches, or between the combined and the uncontrolled indexing approaches.

4.3.2 Multilingual retrieval context

In addition, the combination of both French controlled and uncontrolled vocabularies (FCUV) stood out as the most efficient indexing form with an average of 2.1 queries per retrieved image while the English uncontrolled vocabulary form (EUV) was the least efficient with an average of 3.3 queries per retrieved image. Our findings revealed a significant difference between the indexing forms using a French vocabulary compared to the indexing forms using an English vocabulary. In the context of our study, image retrieval in the monolingual context required fewer queries, on average, to retrieve an image than in the multilingual context.

4.4 Satisfaction of the image searcher

For this study, the satisfaction of the image searcher corresponded to a scale of arbitrary values ranging from 1 to 5. The value "1" represented the minimum satisfaction, while "5" represented the maximum sat-

isfaction. The average levels of satisfaction for all 30 images to be retrieved have been calculated and are shown in Figure 6.
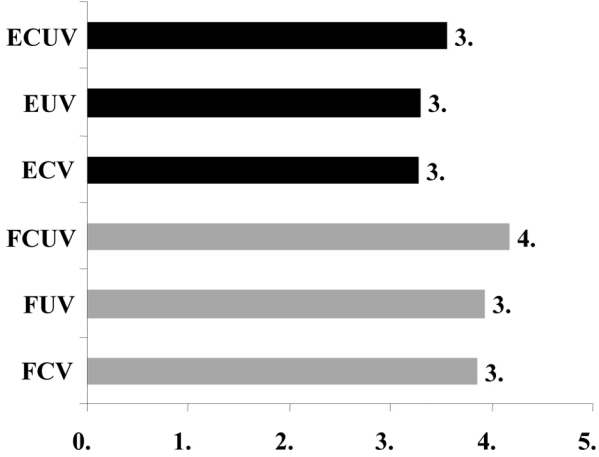

Figure 6. *Satisfaction of the image searcher*

4.4.1 Monolingual retrieval context

The two-factor ANOVA test on the mean satisfaction rate recorded after each one of the 30 image retrieval tasks revealed that there was no significant interaction between the indexing language and the indexing approach. Therefore we independently examined the influence of the indexing approach and the indexing language on image searcher satisfaction. The results of the Tukey test indicated that the difference between the satisfaction rates was significant between the combined and the controlled vocabulary approaches. No significant difference was noticeable between the combined and the uncontrolled approaches, or between the controlled and uncontrolled indexing approaches.

4.4.2 Multilingual retrieval context

We observed that the participants seemed much less satisfied when images were indexed with the English controlled vocabulary form (ECV) and English uncontrolled vocabulary form (EUV). Our analysis also revealed that that the indexing language played an influencing role since the observed difference in the average satisfaction rate between the monolingual and multilingual context is significant, the participants being more satisfied when the image retrieval occurs in the monolingual context rather than the multilingual context.

Knowl. Org. 36(2009)No.4
E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

209

5.0 Discussion and conclusions

The results of our research have led to the conclusion that the retrieval of ordinary images representing common objects, when initiated with a shown image, is more effective (proportions of retrieved images) when the images are indexed with the combination of controlled and uncontrolled vocabularies rather than the controlled vocabulary or the uncontrolled vocabulary approaches respectively. Since the combined indexing approach holds the advantages of both indexing approaches, the effectiveness of the retrieval is maximized. The results of our study also indicate that differences in terms of temporal efficiency (average time required to retrieve an image) are not significant. For human efficiency, the results indicate that the controlled vocabulary indexing approach is more efficient than the uncontrolled approach in terms of the number of queries required to retrieve an image. In the context of this study, the use of the controlled vocabulary offers the maximum consistency for the representation of concepts contained in the image and has an influence on the human efficiency of image retrieval. However, the combination of the controlled and uncontrolled vocabularies, while improving effectiveness by providing more mapping possibilities between the query terms and the indexing terms, results in an increased number of images displayed and hence a loss in temporal efficiency. With regard to the satisfaction of the image searcher, the results of the analysis confirm that searchers consider themselves more satisfied when retrieving images indexed with the combined approach rather than the controlled approach. Regardless of its subjectivity, the assessment of image searcher satisfaction reinforces the results obtained for the effectiveness and efficiency of image retrieval from the perspective of the image searcher.

In regard to the two studies of linguistic retrieval contexts (monolingual and multilingual), the results of this research, conducted with the system El@ine, have shown that the retrieval in the monolingual context is more efficient (queries used to retrieve an image) and more satisfactory for the searcher. Moreover, the retrieval in the monolingual context is more effective than the retrieval in the multilingual context, with a greater proportion (25%) of retrieved images. This result differs slightly from those generally reported by the studies on multilingual image retrieval. For example, the retrieval performance of the multilingual system *Eurovision* proposed by Clough and Sanderson (2006, 706) showed a performance for the monolingual retrieval superior by about 11% com-

pared to the multilingual context, under conditions similar to our study in some respects (image retrieval initiated by a shown image, use of a machine translation system, queries of the participants translated into English). Thus, the analysis of the retrieval effectiveness performance obtained by our study clearly demonstrates that it is less effective when the indexing language is different from the query language, all indexing approaches combined. Since participants did not know to which form of indexing their queries were associated, we can reject the possibility that they changed the way they formulated their queries in relation to a particular indexing language. Consequently, the observed differences between the proportions of retrieved images obtained in multilingual and monolingual context, regardless of the indexing approach, are likely related to the processing of the queries, that is, the translation mechanism used in the retrieval system.

The observed differences when comparing the effectiveness of the two linguistic retrieval contexts can perhaps be explained by the translation of queries to be carried out when the language of the query is different from the indexing language. Several linguistic resources can be used in cross-language information retrieval (CLIR) systems: bilingual or multilingual dictionaries (Pirkola et al. 2001; Hedlund et al. 2004), machine translation (MT) systems (Chen and Gey 2004; Zhang and Vines 2004) and parallel or comparable corpora (Braschler and Schäuble 2000; Xu and Weischedel 2005). The review of the literature conducted on the various linguistic resources that can be used in CLIR systems demonstrated their advantages and limitations. The methodological choice made when designing the retrieval system El@ine has focused on the participants' queries being translated by an incorporated machine translation system (Fluhr 2006, 237). At first, it was believed that the queries used for image retrieval could be properly translated by a machine translation system, since they generally contain only a few words, that is, 3.7 words per query on average, as noted by Goodrum and Spink (2001). Studies by Chen (2001), Fukumoto (2004), Spink and Jansen (2004) and Pu (2005) also confirmed that, in general, people tend to formulate short queries to retrieve images. For this research, analysis of queries shows that the participants made queries containing an average of 1.3 words. This average is slightly less than what is usually found in the literature and is probably a consequence of the image type used for this study.

We could think a priori that short queries generated by the retrieval tasks performed for this study would be most appropriately translated by the machine translation system. However, a partial analysis of the translations obtained automatically indicates that the queries are sometimes too short, that is, they do not provide sufficient context (Oard and Diekema 1998, 231; Kishida 2005, 435; Gey et al. 2005, 427) to be interpreted and, therefore, are incorrectly translated (see Figure 2). Indeed, the translation is often faced with many problems of semantic and syntactic ambiguities (Braschler 2004, 189). In the context of this research, the semantic ambiguity of the request for translation, mainly due to a lack of context, may result in a misinterpretation. This is indeed one of the main problems of CLIR systems using a machine translation mechanism (Kishida 2005, 439). As a result, retrieved and displayed images do not always correspond to the desired image, or queries whose translation is wrong do not generate any results. This largely explains the differences between the proportions of images retrieved in monolingual and multilingual context.

The failure of the machine translation system was mentioned by Fluhr (2006, 237) who emphasized that the quality of translation systems [automatic] cannot guarantee that the translation is good. In the case of a bad translation of the concepts of the query, the search may not succeed. For example, French queries of "vélo" or "bicyclette" would be, in principle, easily translated into the English term "bicycle." However, we noted that the indexing forms with ECV or EUV often used the word "bike" to describe this object. However, the word "bike" translated back into French, by the same machine translation system, is the word "vélo" and not "bicyclette." In other words, if the participant used the word "vélo" in the query, the image of this object could be retrieved, but not if using the word "bicyclette," which is incorrect since the words "vélo" and "bicyclette" are considered synonyms.

It is worth mentioning that these kinds of terminological difficulties are not confined to the retrieval in a multilingual context. The retrieval in the monolingual context also has its share of problems. For example, a participant who uses the word "chandail" to identify an image indexed with the word "maillot" or "pullover" meets the same type of obstacle. Thus, we can suppose that a retrieval device that would include a system of references such as a thesaurus could improve effectiveness both in monolingual and multilingual retrieval contexts. However, as high-lighted by several studies (Markkula and Sormunen 1998; Baca 2003; Jörgensen 2003), there are many obstacles for the use of these controlled vocabularies, including the cultural bias of the term included and inconsistencies chiefly due to their intrinsic complexity (Jörgensen 2003, 98–99).

The present research studied the retrieval of ordinary images representing common objects within two different linguistic contexts. By doing so, this study fills an important gap in the literature about ordinary images. These observed differences between both linguistic contexts may be explained by the semantic ambiguity occurring at the time of the query translation. However, other research seems necessary to further the knowledge obtained by this study, in order to enrich the field of information organization and representation. For example, the linguistic context suggested by this study concerned only two languages from the same family of Indo-European languages. The results of this research could be complemented by the study of other languages to not only identify constants in indexing, but also in the retrieval performance in terms of effectiveness, efficiency and the satisfaction of the image searcher. Comparisons between indexing vocabularies extracted from other linguistic groups (Semitic, Chinese, Niger-Congolese, Austronesian, etc.) could also be considered. In the same manner, the definition of the sample used for this research predicted that the respondents had to be undergraduate students whose mother tongue is French. A comparison with other populations speaking different mother tongues could also enrich the knowledge of image retrieval performance in a multilingual context.

Finally, the results of this research are closely related to the image type used. Thus, one can assume that the study of indexing and retrieval of other image categories could adequately complement the results of this study. The analysis of indexing terms of documentary images, such as images related to a specific field (sports news, medical imaging, etc.) or artistic images (museum objects, famous works, etc.) is a crucial research avenue to obtain a better understanding of the best indexing approach to adopt to optimize image retrieval. Moreover, there are ever-increasing multimedia documents available in information systems on the Internet. The methodology proposed by this study could be applied to other types of documents, such as audio files or videos for example, in order to study how to index and retrieve these multimedia documents and make the necessary suggestions to improve organization and, consequently, retrieval.

Knowl. Org. 36(2009)No.4

211

E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

References

Angus, Emma, Thelwall, Mike, and Stuart, David. 2008. General patterns of tag usage among university groups in Flickr. *Online information review* 32: 89-101.

Armitage, Linda H., and Enser, Peter G. B. 1997. Analysis of user need in image archives. *Journal of information science* 23: 287-99.

Arsenault, Clément. 2006. L'utilisation des langages documentaires pour la recherche d'information. *Documentation et bibliothèques* 52: 139-48.

Association française de normalisation. 1998. *Exigences ergonomiques pour travail de bureau avec terminaux à visualisation (TEV). Partie 11, Lignes directrices relatives à l'utilisabilité*. Genève: Organisation internationale de normalisation.

Baca, Murtha. 2003. Practical issues in applying metadata schemas and controlled vocabularies to cultural heritage information. *Cataloging & classification quarterly* 36 no.3: 47-55.

Besser, Howard, and Snow, Maryly. 1990. Access to diverse collections in university settings: the Berkeley Dilemma. In Petersen,Tony, and Molholt, Pat, eds., *Beyond the book: extending MARC for subject access.* Boston: G. K. Hall, pp. 203-24.

Brangier, Éric, and Barcenilla, Javier. 2003. *Concevoir un produit facile à utiliser*. Paris: Éditions d'Organisation.

Braschler, Martin. 2004. Combination approaches for multilingual text retrieval. *Information retrieval* 7: 183-204.

Braschler, Martin, and Schäuble, Peter. 2000. Using corpus-based approaches in a system for multilingual information retrieval. *Information retrieval* 3: 272-84.

Chen, Aitao, and Gey, Frederic C. 2004. Multilingual information retrieval using machine translation, relevance feedback and decompounding. *Information retrieval* 7: 149-82.

Chen, Hsin-Liang. 2001. An analysis of image queries in the field of art history. *Journal of the American Society for Information Science and Technology* 52: 260-73.

Choi, Youngok, and Rasmussen, Edie M. 2002. Users' relevance criteria in image retrieval in American history. *Information processing & management* 38: 695-726.

Choi, Youngok, and Rasmussen, Edie M. 2003. Searching for images: the analysis of users' queries image retrieval in American History. *Journal of the American Society for Information Science and Technology* 54: 498-511.

Chu, Heting. 2003. *Information representation and retrieval in the digital age*. Medford, NJ: Information Today.

Cleveland, Donald B., and Cleveland, Ana D. 2001. *Introduction to indexing and abstracting*, 3rd ed. Englewood, Colo: Libraries Unlimited.

Clough, Paul, and Sanderson, Mark. 2006. User experiments with the Eurovision cross-language image retrieval system. *Journal of the American Society for Information Science and Technology* 57: 697-708.

Clough, Paul, et al. 2005. *The CLEF 2005 cross-language image retrieval track*. Available http://www.clef-campaign.org/2005/working_notes/workingnotes2005/hersh05.pdf.

Contandriopoulos, André-Pierre, et al. 1990. *Savoir préparer une recherche: la définir, la structurer, la financer*. Montréal: Presses de l'Université de Montréal.

Corbeil, Jean-Claude, and Archambault, Ariane. 2003. *Le Nouveau dictionnaire visuel multilingue*. Montréal: Québec Amérique.

Enser, Peter G. B. et al. 2007. Facing the reality of semantic image retrieval. *Journal of documentation* 63: 465-81.

Fluhr, Christian. 2006. Le rôle de la terminologie dans les systèmes d'informations multilingues. In Mustafa El Hadi, Widad, ed., *Terminologie et accès à l'information.* Paris: Hermès Science, pp. 235-55.

Fortin, Marie-Fabienne.1996. *Le processus de la recherche: de la conception à la réalisation.* Montréal: Décarie.

Fortin, Marie-Fabienne, Côté, José, and Filion, Françoise. 2006. *Fondements et étapes du processus de recherche.* Montréal: Chenelière Éducation.

Fukumoto, Toru. 2004. *An analysis of image retrieval behavior for metadata type and Google image database.* Available http://www.nier.go.jp/fukumoto/fukumoto_ICCE2004.pdf.

Furnas, G. W., et al. 1987. The vocabulary problem in human-system communication. *Communications of the ACM* 30: 964-71.

Gey, Frederic C., Kando, Noriko, and Peters, Carol. 2005. Cross-language information retrieval: the way ahead. *Information processing & management* 41: 415-31.

Golder, Scott A., and Huberman, Bernardo A. 2005. *The structure of collaborative tagging systems*. Available: http://tinyurl.com/h584l.

Gonzalo, Julio, Karlgren, Jussi, and Clough, Paul. 2006. *iCLEF 2006 overview: searching the Flickr*

212

Knowl. Org. 36(2009)No.4
E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

WWW photo-sharing repository. Available http://tinyurl.com/ykoj8x.

Goodrum, Abby A. 2000. Image information retrieval: an overview of current research, *Informing science* 3. Available http://inform.nu/Articles/Vol3/v3n2p63-66.pdf.

Goodrum, Abby A., and Spink, Amanda. 2001. Image searching on the Excite Web search engine. *Information processing & management* 37: 295-311.

Goodrum, Abby A., Bejune, Matthew M., and Siochi Antonio C. 2003. *A state transition analysis of image search patterns on the web*. Available http://www.springerlink.com/media/6p8qd4b65j3vpne80ecn/contributions/0/e/w/7/0ew7r6lelaafcqv8.pdf.

Greisdorf, Howard F., and O'Connor, Brian C. 2008. *Structures of images collections: from Chauvet-Pont d'Arc to Flickr*. Westport, Conn.: Unlimited Libraries.

Guinchat, Claire, and Menou, Michel. Sciences et techniques de l'information et de la documentation, 2nd ed. rev. and augm. Paris: Unesco.

Harpring, Patricia. 2002. The language of images: enhancing access to images by applying metadata schemas and structured vocabularies. In Baca, Murtha, ed., *Introduction to art image access issues, tools, standards, strategies.* Los Angeles: Getty Research Institute, pp. 20-39.

Hedlund, Turid, et al. 2004. Dictionary-based cross-language information retrieval: learning experiences from CLEF 2000-2002. *Information retrieval* 7: 99-119.

Howell, David C. 2001. *Méthodes statistiques en sciences humaines*. Paris: De Boeck.

Jörgensen, Corinne. 1998. Attributes of images in describing tasks. *Information processing & management* 34: 161-74.

Jörgensen, Corinne. 2003. *Image retrieval: theory and research*. Lanham MD: Scarecrow Press.

Jörgensen, Corinne, and Jörgensen, Peter. 2005. Image querying by image professionals. *Journal of the American Society for Information Science and Technology* 56: 1346-59.

Kamps, Jaap. 2004. Improving retrieval effectiveness by reranking documents based on controlled vocabulary. In McDonald, Sharon, and Tait, John Berlin, eds., *Advances in Information Retrieval: 26th European Conference on IR research, ECIR 200.* Berlin: Springer, pp. 283-95.

Karlgren, Jussi, and Olsson, Fredrik. 2006. *Trusting the results in crosslingual keyword-based image retrieval*. Available http://clef.isti.cnr.it/2006/working_notes/workingnotes2006/karlgren CLEF2006.pdf.

Kipp, Margaret E.I. 2007. Tagging practices on research oriented social bookmarking sites. In Arsenault, Clément, and Dalkir, Kimiz, eds., *ACSI/CAIS 2007*, 35th Congrès annuel de l'Association canadienne des sciences de l'information, Université McGill, Montréal, Québec, 10-12 mai 2007. Available http://www.cais-acsi.ca/proceedings/2007/kipp_2007.pdf.

Kishida, Kazuaki. 2005. Technical issues of cross-language information retrieval: a review. *Information processing & management* 41: 433-55.

Krause, Michael G. 1988. Intellectual problems of indexing picture collections. *Audiovisual librarian* 14: 73-81.

Lancaster, F. W. 2003. *Indexing and abstracting in theory and practice*. Champaign, Ill.: Graduate School of Library and Information Science, University of Illinois.

Layne, S.S. 1994. Some issues in the indexing of images. *Journal of the American Society for Information Science* 4: 583-88.

Le Deuff, Olivier. 2006. Folksonomies - les usagers indexent le web. *Bulletin des Bibliothèques de France* 51: 66-70.

Macgregor, George, and McCulloch, Emma. 2006. Collaborative tagging as a knowledge organisation and resource discovery tool. *Library review* 5: 291-300.

Markey, Karen. 1983. Computer-assisted construction of a thematic catalog of primary and secondary subject matter. *Visual resources* 3: 16-49.

Markey, Karen. 1988. Access to iconographical research collections. *Library trends* 37 no. 2: 154-74.

Markey, Karen, Atherton, Pauline, and Newton, Claudia. 1980. An analysis of controlled vocabulary and free text search statements in online searches. *Online Review* 4: 225-36.

Markkula, Marjo, and Sormunen, Eero. 2000. End-user searching challenges indexing practices in the digital newspaper photo archive. *Information Retrieval* 1: 259-85.

Marlow, Cameron, et al. 2006. *Position paper, tagging, taxonomy, flickr, article, toRead*. Available: http://www.danah.org/papers/WWW2006.pdf.

Marlow, J., et al. 2007. Multilingual Needs Of Cultural Heritage Web Site Visitors: A Case Study Of Tate Online. In . Trant, J., and Bearman, D., eds, *International Cultural Heritage Informatics Meeting (ICHIM07): Proceedings.* Toronto: Archives & Museum Informatics. Available http://www.

Knowl. Org. 36(2009)No.4

E. Ménard. Image Retrieval: A Comparative Study on the Influence of Indexing Vocabularies

213

archimuse.com/ichim07/papers/marlow/marlow.html.

Mathes, Adam. 2006. *Folksonomies - cooperative classification and communication through shared metadata*. Available http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html.

Matusiak, Krystyna K. 2006. Towards user-centered indexing in digital image collections. *OCLC systems & services* 22: 283-98.

Ménard, Elaine. 2007. Étude sur l'influence du vocabulaire utilisé pour l'indexation de l'image numérique sur le repérage en contexte multilingue. In Arsenault, Clément, and Dalkir, Kimiz, eds., *ACSI/CAIS 2007*, 35th Congrès annuel de l'Association canadienne des sciences de l'information, Université McGill, Montréal, Québec, 10-12 mai 2007. Available http://www.cais-acsi.ca/proceedings/2007/menard_2007.pdf.

Oard, Douglas W., and Diekema, Anne R.1998. Cross-language information retrieval. In Williams, Martha E., ed., *Annual review of information science and technology* 33. Medford, NJ: Information Today, pp.223-56.

Ohlgren, T. 1980. Subject indexing of visual resources: a survey. *Visual Resources* 1: 67-73.

Ørnager, Suzanne. 1996. The newspaper image database: empirical supported analysis of users' typology and word association clusters. In *SIGIR 95*, Seattle, http://tinyurl.com/fzumv.

Ouellet, André. 1994. *Processus de recherche : une introduction à la méthodologie de la recherche*. Sainte-Foy: Presses de l'Université du Québec.

Panofsky, Erwin. 1955. *Meaning in the visual arts: papers in and on art history*. Garden City, NY: Doubleday.

Pirkola, Ari, et al. 2001. Dictionary-based cross-language information retrieval: problems, methods, and research findings. *Information Retrieval* 4: 209-30.

Pu, Hsiao-Tieh. 2005. A comparative analysis of web image and textual queries. *Online information review* 29: 457-67.

Rao Muddamalle, Manikya. 1998. Natural language versus controlled vocabulary in information retrieval: A case study in soil mechanics. *Journal of the American Society for Information Science and Technology* 49: 881-87.

Roddy, K. 1991. Subject access to visual resources: what the 90s might portend. *Library hi tech* 9: 45-49.

Rorissa, Abebe. 2008. User-generated descriptions of individual images versus labels of groups of images: a comparison using basic level theory. *Information processing & management* 44: 1741-53.

Savoy, Jacques. 2005. Bibliographic database access using free-text and controlled vocabulary: an evaluation. *Information processing and management* 41: 873-90.

Shatford, Sara. 1986. Analyzing the subject of a picture: a theoretical approach. *Cataloging & classification quarterly* 6 no. 3: 39-61.

Spink, Amanda. 2002. A user-centered approach to evaluating human interaction with Web search engines: an exploratory study. *Information processing & management* 38: 401-26.

Spink, Amanda, and Jansen, Bernard J. 2004. *Web search: public searching of the Web*. Boston: Kluwer Academic.

Spink, Amanda, et al. 2001. Searching the Web: the public and their queries. *Journal of the American Society for Information Science and Technology* 52: 226-34.

Sproull, Natalie L. 1995. *Handbook of research methods: a guide for practitioners and students in the social sciences*, 2nd ed. Metuchen, NJ: Scarecrow Press.

Tjondronegoro, Dian, and Spink, Amanda. 2008. Web search engine multimedia functionality. *Information processing & management* 44: 340-57.

Turner, James M. 1993. Subject access to pictures: considerations in the surrogation and indexing of visual documents for storage and retrieval. *Visual resources* 9: 241-71.

Turner, James M. 1994. *Determining the subject content of still and moving image documents for storage and retrieval: an experimental investigation*. Doctoral Dissertation, University of Toronto, Toronto.

Wellisch, Hans H. 1995. *Indexing from A to Z*. New York: H. W. Wilson.

Xu, Jinxi, and Weischedel, Ralph. 2005. Empirical studies on the impact of lexical resources on CLIR performance. *Information processing & management* 41: 475-87.

Yahoo! 2008. *Yahoo Babel Fish*. Available http://fr.babelfish.yahoo.com/.

Zhang, Ying, and Vines, Phil. 2004. Using the web for automated translation extraction in cross-language information retrieval. In *Proceedings of SIGIR'04*, Sheffield, South Yorkshire, UK. Available http://tinyurl.com/jv9l7.