

	<p style="text-align: right;">Date : 27/06/2007</p> <p>Audiovision et indexation de films</p> <p>James M Turner École de bibliothéconomie et des sciences de l'information Université de Montréal, Montréal, Québec Canada</p> <p>Suzanne Mathieu Ville de Montréal, Montréal, Québec Canada</p>
Meeting:	157 Classification and Indexing
Simultaneous Interpretation:	No
<p>WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL 19-23 August 2007, Durban, South Africa http://www.ifla.org/iv/ifla73/index.htm</p>	

Résumé

L'accès aux documents audiovisuels devrait être aussi facile que l'accès aux imprimés. Toutefois, une telle réalité nous échappe pour le moment. Les méthodes utiles au traitement documentaire d'imprimés ne s'appliquent pas nécessairement très bien aux documents audiovisuels et multimédias. Dans le cadre de ce projet, nous avons étudié des textes d'audiovision et des descriptions textuelles d'images afin d'évaluer ces derniers en tant que sources de mots-clés pour l'indexation d'images en mouvement. Nous avons trouvé ces sources utiles pour cette fin. Dans un deuxième temps, nous avons étudié la possibilité de traduire automatiquement vers d'autres langues des mots-clés provenant de textes d'audiovision, toujours aux fins d'indexation. Les résultats sont encourageants également.

Introduction

Comme nous le rappellent les *Recommandations relatives aux documents audiovisuels et multimédias à l'usage des bibliothèques et autres institutions* de l'IFLA, l'accès à ces documents devrait être aussi ouvert et aussi libre que celui des documents imprimés. Toutefois, une telle réalité nous échappe pour le moment. Les méthodes utiles pour le traitement documentaire d'imprimés ne s'appliquent pas nécessairement très bien aux documents audiovisuels et multimédias. Par ailleurs, il faudrait mettre davantage d'efforts avant d'arriver à bien offrir aux utilisateurs de ces documents l'accès adéquat.

Nous faisons rapport ici de nos travaux dans le contexte du Réseau de recherche E-Inclusion (<<http://e-inclusion.crim.ca/?q=fr>>). Ce réseau a pour but « de fournir aux créateurs des outils audio-vidéo puissants qui leur permettent de se concentrer sur les aspects de création et ainsi d'améliorer la richesse de l'expérience multimédia pour les personnes ayant une déficience sensorielle ». Le réseau est une initiative du *Centre de recherche informatique de Montréal* (CRIM, <<http://www.crim.ca/>>) et comprend plusieurs universités ainsi que des

partenaires, dont l'Office national du film du Canada (ONF) et l'INCA, anciennement l'Institut national canadien pour les aveugles et AudioVision, une agence de description.

Le projet 3,1 du réseau de recherche a trait à l'identification de types d'informations dont les personnes aveugles ou malvoyantes ont besoin pour bien comprendre des films, des vidéos ou des émissions de télévision. L'objectif de ce projet était de produire des lignes directrices pour les personnes qui font la description d'images en mouvement pour ce public. On appelle ces personnes des *descripteurs* (*descriptrices* au féminin). Afin de mener à bien notre recherche, nous avons visionné plusieurs films avec un public aveugle ou malvoyant. Nous avons demandé à ces participants de commenter la qualité de l'audiovision. Nous cherchions à savoir si la description d'images était adéquate et si non, quelles améliorations pourraient se faire. Un des buts du Réseau de recherche E-Inclusion est d'étudier la possibilité d'offrir la description variable, c'est-à-dire de la description paramétrable par l'utilisateur. Par exemple, l'utilisateur pourrait choisir d'entendre plus ou moins de description ou encore la description de seulement certains types d'information. Ceci est d'intérêt parce qu'il existe une gamme étendue de problèmes de perte de vision. En réalité, peu d'utilisateurs d'audiovision sont complètement aveugles. Par ailleurs, dans le contexte du vieillissement des populations, de plus en plus de gens perdent la vision parce qu'ils vieillissent et vivent plus longtemps qu'avant. De plus, même les gens qui voient parfaitement bien prennent souvent plaisir à visionner des films décrits en audiovision. C'est qu'on y donne parfois des informations qui auraient passé inaperçues sans la description ou qui fournissent un contexte favorisant la compréhension.

Afin d'étudier ces problèmes, il fallait identifier les caractéristiques des films visionnés avec les aveugles et malvoyants. Cela comprend un travail initial d'analyse, lequel est décrit dans la partie suivante. En cours de route, nous avons également étudié le vocabulaire de description. Ce dernier est le sujet du présent rapport.

Contexte

Depuis 1995, nous travaillons sur quelques projets de recherche ayant comme objectif d'étudier les caractéristiques des textes d'audiovision, afin d'évaluer son utilité pour l'indexation d'images en mouvement. Nous nous intéressons particulièrement à la question de l'indexation plan par plan de films, d'émissions de télévision et d'autres documents audiovisuels. Pour ce faire, nous avons développé une méthode pour la transcription des textes d'audiovision et pour faire la correspondance avec l'image (Turner 1998). Plus tard, nous avons raffiné la méthode (Turner et Colinet 2004, Mathieu et Turner 2007a).

Nous avons développé la notion d'« épisodes » d'audiovision, des unités d'information possédant une certaine cohérence logique. Lorsque le sujet de description change, c'est un nouvel épisode qui commence. Or l'audiovision se fait sur une piste audio particulière, dans laquelle on insère l'information dans des espaces sonores disponibles. Ainsi, il n'est pas toujours possible de réciter le texte au même moment où l'image correspondante s'affiche. Nous avons étudié les divers schèmes d'insertion, afin d'identifier si l'épisode d'audiovision s'entend en même temps que l'image s'affiche, ou encore un plan avant ou après l'image, deux plans avant ou après, et ainsi de suite.

On peut offrir l'indexation plan par plan en connectant dans une base de données le texte de description à l'image correspondante, en isolant les mots-clés utiles. Les utilisateurs créateurs de films ou d'émissions de télévision, ceux qui étudient ces produits, ceux dans les salles de nouvelles de télévision et d'autres utilisateurs peuvent bénéficier énormément de ce type d'indexation.

Par ailleurs, les bases de données qui offrent l'indexation plan par plan d'images en mouvement offrent d'habitude une description écrite de l'action de chaque plan également.

L'utilisateur peut donc lire rapidement cette description, sans avoir à visionner le plan, afin de réduire la liste de possibilités aux plus intéressantes, pour ensuite visionner seulement les plus prometteurs. Ainsi, il peut réaliser des économies importantes de temps.

Nous cherchions à savoir si le texte de l'audiovision pouvait être utilisé pour de telles descriptions, c'est-à-dire si assez de mots-clés étaient communs dans les deux types de textes. Si oui, il serait possible de réaliser d'autres économies en évitant la nécessité de formuler des descriptions tout en générant l'indexation à partir du seul texte de l'audiovision. Cependant, nos travaux antérieurs nous faisaient voir que le texte de l'audiovision se réfère souvent à une séquence de plans et donc pas toujours à des plans individuels. Tout dépend du film. Voici quelques exemples : parfois l'action se déroule très rapidement. On monte une séquence de courts plans afin de maintenir la stimulation du public. Dans ce cas, il est inutile de décrire chaque plan. On décrit plutôt la séquence. Dans d'autres films, l'action se déroule lentement et des plans restent longtemps à l'écran. Il y a donc peu de concurrence pour les espaces sonores, surtout en l'absence de dialogues. Dans un tel cas, le descripteur peut prendre tout son temps pour décrire l'image. D'autres films sont très courts, par exemple les films expérimentaux de Norman McLaren que nous avons étudiés dans le cadre du présent projet. Il n'y avait que de la musique comme son. De plus, le film au complet est composé d'un seul plan. Il faut même remettre en question la notion de plan, parfois, car certains films expérimentaux sont gravés ou dessinés à même la pellicule. C'est ainsi que nous ne pouvons formuler des règles pour couvrir toute situation. À cause de cette variété de formes possibles pour les images en mouvement, il faut plutôt identifier des tendances des divers types de produits.

Le réseau de recherche E-Inclusion

Ce projet de recherche s'adresse à trois aspects du thème du programme de la section Classification et indexation de l'IFLA : l'accès par sujet pour les langues minoritaires, la collaboration avec les utilisateurs dans le design d'outils d'accès et l'amélioration de l'accès par sujet aux ressources. Le premier thème est traité par les aspects bilingues (français et anglais) de notre projet. Bien que nous ne traitons pas de l'aspect multilingue dans ce rapport, on peut faire des liens avec des projets antérieurs (par exemple Turner et Hudon 2002) afin d'étendre les résultats vers d'autres langues. La collaboration avec les utilisateurs se fait par nos discussions avec les participants aveugles ou malvoyants, lesquels nous ont fourni beaucoup d'informations utiles. Par l'analyse des sujets discutés avec ces participants, nous avons pu formuler nos lignes directrices. Par ailleurs, leurs commentaires sur les types utiles de description nous guident dans la recherche des termes d'indexation les plus utiles pour ce public. L'amélioration de l'accès par sujet aux ressources se fait en offrant un accès quelconque aux séquences et aux plans, car ce type d'accès n'est pas disponible actuellement, sauf dans des situations particulières, telles que les cinémathèques de plans d'archives et les salles de nouvelles de la télévision.

En analysant le texte de l'audiovision des films étudiés, nous avons cherché également à valider la typologie des types d'informations véhiculées par les films déjà étudiés. En ajoutant à ce petit corpus, nous avons pu consolider nos résultats en les comparant aux résultats déjà obtenus. Nous voulions voir si cette typologie était déjà assez complète ou s'il fallait y introduire de nouvelles catégories. Le peu de travail dans ce domaine fait que chaque nouveau film étudié augmente un peu le corpus de films analysés et aide à confirmer ou à infirmer nos résultats. Voilà ce qui explique également le nombre important de références à notre propre travail dans le texte.

Par rapport aux résultats dans ce rapport, nos objectifs du projet étaient de :

- valider la typologie de types de description trouvés dans des textes d'audiovision

- comparer les mots-clés de l’audiovision avec ceux des descriptions d’utilisateur rédigées pour chaque plan, afin d’évaluer la possibilité de générer automatiquement l’indexation
- comparer les mots-clés en anglais avec ceux du français, afin d’évaluer la possibilité de produire l’indexation automatiquement entre ces deux langues et dans d’autres langues

Nous souhaitons atteindre ces objectifs afin d’acquérir suffisamment de connaissances pour obtenir des réponses adéquates à nos questions de recherche. Dans la partie qui suit, nous traitons de la méthodologie utilisée pour y arriver.

Méthodologie

Dans le cadre du projet nous avons analysé onze films, dont deux longs métrages, deux documentaires, et sept courts films d’animation. Chaque film était découpé en séquences et plans, et l’audiovision en épisodes. De plus, une description en texte libre était rédigée pour chaque plan par un assistant de recherche qui visionnait le film sans le son, afin d’éviter d’être influencé par les dialogues, la narration ou l’audiovision. Les fichiers utilisés pour l’analyse sont le produit de ce travail.

Un logiciel permettait de jumeler les segments de films avec le code temporel, ce qui permettait de les identifier précisément. Le code temporel inclut l’heure, la minute, la seconde, et chaque image. Les films fournis par l’Office national du film du Canada nous parvenaient avec le code temporel déjà brûlé à même l’image.

Afin de répondre à nos objectifs de recherche, nous avons catégorisé les épisodes d’audiovision en utilisant la typologie développée lors d’un projet antérieur (Turner 1998) puis raffiné lors d’un deuxième projet (Turner et Colinet 2004). Nous avons également pris en considération la typologie présentée par Piety (2003) mais nous l’avons jugé trop général pour les besoins de notre analyse. Nous avons consulté également les lignes directrices de l’Independent Television Commission (2000).

Des fichiers Excel ont été créés pour les productions analysées. Ceux-ci comprennent le dépouillement du film en séquences et plans, chacun avec sa description et le texte de l’audiovision (figure 1). Dans le cas des longs métrages, les informations pour chaque chapitre sont stockées dans une feuille. Grâce à cette méthode, nous pouvions voir un survol des données en plus d’automatiser certains calculs. Cela représente une technique raffinée par rapport à celles utilisées dans nos projets antérieurs, dans lesquels l’information était stockée dans des bases de données, lesquelles ne permettent pas de voir les données en survol.

Les fichiers Excel contiennent donc de nombreux types de données et permettent plusieurs types d’analyse. Dans ce rapport, la discussion est focalisée sur les aspects pertinents pour la section Classification et indexation, d’abord sur la relation entre le texte de l’audiovision et celui de la description écrite, et ensuite sur la relation entre les mots-clés dans deux langues, le français et l’anglais.

#	Chapitre	TC in	TC out	Type	AV	Description	Texte de l'audiovison
7	7	00:45:35	00:45:36	Décor?	1	Gros plan si	Photos de l'homme chauve que Nino poursuivait.
8	7	00:45:26	00:46:26	Action	1	Pionnée sur	Recouvre la tête.
9	7	00:47:25	00:47:28	Action	2	Amélie est	Dans son lit en fer forgé, Amélie regarde à nouveau les photos.
10	4	00:47:33	00:47:33	Action	2	Amélie, ass	Les visages se succèdent.
11	5	00:47:36	00:47:36	Person	2	Photos d'un	Un homme barbu.
12	6	00:47:38	00:47:38	Person	2	Photos d'un	Une femme qui éternue.
13	7	00:47:41	00:47:43	Person	2	Photos d'un	Un jeune homme noir à l'air séneux.
14	0	00:47:45	00:47:45	Person	2	Photos d'un	Un bébé
15	9	00:47:46	00:47:47	Person	2	Photo d'un	Un homme qui remet son bonnet.
16	10	00:47:50	00:47:52	Action	2	Amélie est	Amélie est endormie, la télécommande de son téléviseur posée près d'elle.
17	11	00:47:54	00:47:56	Action	2	On voit une	La télévision diffuse une course cycliste.
18	12	00:47:56	00:48:00	Action	2	On voit la	Un cheval s'échappe de son pré en sautant une barrière et se met à galoper au milieu des coureurs.
19	13	00:48:01	00:48:03	Action	2	Amélie se	Amélie ouvre les yeux, enregistre la séquence.
20	14	00:48:12	00:48:14	Action	1	On voit les	Amélie descend les escaliers de son immeuble.
21	15	00:48:16	00:48:19	Action	1	On voit Amé	Elle s'arrête sur un palier et dépose une grosse enveloppe craft sous un paillason.
22	16	00:48:21	00:48:24	Action	1	Amélie se	En se redressant, elle voit un trousseau de clefs subtilé sur la porte d'en face.
23	17	00:48:27	00:48:29	Action	1	Amélie pren	Elle le prend et sort.
24	18	00:48:30	00:48:32	Action	1	Amélie desc	Dehors il y a du brouillard, l'épicer derrière son étal.
25	19	00:48:41	00:48:42	Action	1	Lucien toum	Il se tourne vers son commis.
26	20	00:48:53	00:48:53	Action	1	L'épicer tap	Il le tape.
27	21	00:49:03	00:49:04	Action	1	Lucien baiss	Lucien s'éclipse, Amélie est furieuse.
28	22	00:49:21	00:49:23	Action	1	Amélie man	Elle marche dans la rue d'un pas décidé.
29	23	00:49:24	00:49:27	Action	1	Amélie man	Pousse la porte du condonnier sur laquelle est inscrit: (ci dé-minute.
30	24	00:49:31	00:49:33	Action	1	On voit Amé	Sur le palier, elle remet le trousseau de clefs à sa place.
31	25	00:49:34	00:49:34	Décor?	1	Amélie se	Dans sa poche, une clef apparaît comme par radioscopie.
32	26	00:49:39	00:49:39	Action	1	Amélie redé	Elle respire.
33	27	00:49:40	00:49:41	Action	2	Amélie redé	Georgette est au kiosque à journaux.
34	28	00:50:00	00:50:02	Action	2	La dame du	Elle tend une pile de journaux à Georgette qui semble ne pas comprendre.
35	29	00:50:03	00:50:04	Lieu	1	Gros plan si	Au café, Hipolito est au comptoir.
36	30	00:50:46	00:50:48	Action	1	Gina replac	Elle remet en place le cou d'un client.
37	31	00:50:59	00:51:01	Action	1	On voit Jos	Georgette et Joseph se regardent du coin de l'œil.
38	32	00:51:09	00:51:11	Action	2	On voit Lud	Vouant regarder sa montre, Lucien renverse le cageau.
39	33	00:51:13	00:51:15	Action	2	Amélie lano	Amélie regarde fixement l'épicer.
40	34	00:51:17	00:51:19	Action	1	Amélie déba	Avec sa clef secrète, elle entre chez Collignon.
41	35	00:51:23	00:51:24	Action	1	Amélie man	Elle traverse le couloir.
42	36	00:51:27	00:51:29	Action	1	Se tourne	Elle passe dans le salon où une paire de charentaises attend au pied d'un fauteuil.
43	37	00:51:33	00:51:36	Action	1	Amélie soul	Elle souève les pantoufes pour regarder la peinture.
44	38	00:51:37	00:51:37	Image	1	On voit le	42
45	39	00:51:39	00:51:41	Action	1	Amélie déba	Elle ouvre le placard à chaussures.
46	40	00:51:46	00:51:46	Action	1	Amélie rega	Limite des lacets.
47	41	00:51:58	00:51:00	Action	1	Amélie se	Dans le mille, le hain, elle inhéruet le dentifice avec un tube de crème pour les pieds.

Figure 1. Un fichier Excel avec la liste des plans et le texte audiovison correspondant.

Pour chaque plan, lorsqu'on compare le texte de la description écrite avec celui de l'audiovison, nous pouvons estimer la fréquence à laquelle les mots-clés paraissent dans les deux textes. Cela permet d'évaluer la possibilité d'automatiser le processus d'indexation et d'identifier lequel des deux types de texte est plus utile pour cette fin. Nous avons fait une copie du fichier Excel pour le film *Le fabuleux destin d'Amélie Poulain*, soit l'un des films étudiés sous plusieurs angles par de nombreux chercheurs du projet plus large du réseau E-Inclusion. Cette version du fichier (figure 2) montre le même chapitre 7, réorganisé pour montrer les colonnes A, G et H, soit le numéro de plan, la description écrite et le texte de l'audiovison. Les mots-clés qui figurent dans les deux types de description sont affichés en orange. Comme on peut voir, les mots-clés correspondants ne paraissent pas toujours avec les mêmes plans (par exemple dans la séquence des plans 22 à 25), ceci parce que l'audiovison est inséré dans des espaces sonores disponibles du film. Aux fins de cette analyse, nous avons donc permis d'inclure les mots-clés qui paraissent à quelques plans de celui qu'ils décrivent.

Afin de comparer les mots-clés en français avec ceux en anglais, nous avons employé une méthode utilisée dans plusieurs projets antérieurs. Un dictionnaire bilingue offre une ou plusieurs acceptions dans la langue cible pour chaque mot de la langue source. En comparant deux textes, dont un de chaque langue, si l'une des acceptions de la langue cible correspond au mot-clé équivalent de la langue source, nous considérons qu'il y a équivalence. Aux fins de cette analyse, nous avons utilisé le logiciel Reverso Online (Reverso 2007), en employant la fonctionnalité de recherche bilingue dans le dictionnaire *Collins* (anglais-français).

Resultats et analyse

D'abord, on discute des résultats obtenus en comparant le texte de la description écrite avec celui de l'audiovison. Ensuite, nous discutons de comment les mots-clé utilisés dans le texte anglais de l'audiovison se comparent à la traduction automatique vers le français de ces

mots-clés.

La figure 2 présente les mots-clés (de couleur orange) pour le chapitre 7 de *Le fabuleux destin d'Amélie Poulain*.

#	Description	Texte de l'audiovision
1	Gros plan sur la photo de l'homme chauve.	Photo de l'homme chauve que Nino poursuivait.
2	Plongée sur le tableau que M. Dufayel s'apprête à signer.	Reprendre la toile.
3	Amélie est assise dans son lit, un album à la main.	Dans son lit en fer forgé, Amélie regarde à nouveau les photos.
4	Amélie, assise dans son lit, regarde l'album.	Les visages se succèdent.
5	Photos d'un homme barbu.	Un homme barbu.
6	Photos d'une femme qui étouffe.	Une femme qui étouffe.
7	Photos d'un homme noir.	Un jeune homme noir à l'air sérieux.
8	Photos d'un bébé.	Un bébé.
9	Photo d'un homme avec une main dans son bonnet.	Un homme qui remet son bonnet.
10	Amélie est couchée dans son lit, elle dort, une télécommande près de son bras.	Amélie est endormie, la télécommande de son téléviseur posée près d'elle.
11	On voit une course de vélo dans un écran de télévision.	La télévision diffuse une course cycliste.
12	On voit la course de vélos dans l'écran télé puis Amélie qui se réveille.	Un cheval s'échappe de son gré en sautant une barrière et se met à galoper au milieu.
13	Amélie se réveille et appuie sur sa télécommande.	Amélie ouvre les yeux, enregistre la séquence.
14	On voit les jambes d'Amélie qui descendent les marches d'un escalier.	Amélie descend les escaliers de son immeuble.
15	On voit Amélie qui s'agenouille pour mettre une enveloppe sous un palissonn.	Ele s'arrête sur un palier et dépose une grosse enveloppe kraft sous un palissonn.
16	Amélie se relève de devant le palissonn et se retourne.	En se redressant, elle voit un trousseau de clés oublié sur la porte d'en face.
17	Amélie prend le trousseau de clés.	Ele le prend et sort.
18	Amélie descend les escaliers.	Derrière il y a du brouillard, l'épicerie derrière son étal.
19	Lucien tourne son visage vers l'épicer.	Il se tourne vers son commis.
20	L'épicer tape la tête de Lucien.	Il le tape.
21	Lucien baisse la tête et s'en va de derrière l'étal.	Lucien s'éclipse, Amélie est furtive.
22	Amélie marche furtivement dans la rue.	Ele marche dans la rue d'un pas décidé.
23	Amélie marche furtivement dans la rue.	Pousse la porte du coordonnier sur laquelle est inscrit: 10 dé-minute.
24	On voit Amélie qui remet les clés dans la serrure.	Sur le palier, elle remet le trousseau de clés à sa place.
25	Amélie se détourne de la serrure et on voit une clé sur elle.	Dans sa poche, une clé apparaît comme par radioscopie.
26	Amélie redescend les escaliers.	Ele redescend.
27	Amélie redescend les escaliers.	Georgette est au kiosque à journaux.
28	La dame au kiosque à journaux remet des journaux à Georgette.	Ele tend une pile de journaux à Georgette qui semble ne pas comprendre.
29	Gros plan sur le visage de la dame du kiosque aux journaux.	Au café, Hippolyte est au comptoir.
30	Gina replace le cou d'un vieux homme.	Ele remet en place le cou d'un client.
31	On voit Joseph qui regarde le tabac de Georgette par un miroir.	Georgette et Joseph se regardent du coin de l'œil.
32	On voit Lucien qui échappe le cagou. Amélie qui arrive.	Voulant regarder sa montre, Lucien renverse le cagou.
33	Amélie lance un regard noir à Hippolyte.	Amélie regarde fixement Hippolyte.
34	Amélie débarras la porte puis entre chez Collignon.	Avec sa clé secrète, elle entre chez Collignon.
35	Amélie marche dans le couloir.	Ele traverse le couloir.
36	Se tourne vers une autre pièce et y entre.	Passé dans le salon où une paire de charentaises attend au pied d'un fauteuil.
37	Amélie soulève une paire de pantoufles qu'elle retourne.	Ele soulève les pantoufles pour regarder la peinture.
38	On voit le chiffre 42 sous une pantoufle.	42
39	Amélie détourne la tête.	Ele ouvre le placard à chaussures.
40	Amélie regarde les étagères d'un garde-robe.	Lime des lacets.

Figure 2. Le même fichier Excel, réorganisé pour afficher les mots-clés de la description écrite et ceux du texte de l'audiovision.

Dans une analyse de six chapitres de ce film, on trouve au total 238 plans. De ceux-ci, exactement 50 n'ont aucun mot-clé figurant dans les deux types de texte. Pour les 188 autres plans, 463 mots-clés figurent dans les deux descriptions. Le ratio est donc d'environ 2,5 mots-clés par plan.

Afin de répondre à notre question sur lequel des deux types de description est plus utile que l'autre pour des fins de production de mots-clés utiles pour l'indexation, nous observons d'abord que les mots-clés essentiels figurent dans les deux textes, comme on pouvait s'y attendre. Toutefois, dans chaque type de texte se trouvent d'autres mots-clés qui ne figurent pas dans l'autre texte. Nous pouvons donc constater qu'il serait utile de se servir des deux types de texte pour générer automatiquement des mots-clés correspondant à l'image, pourvu que le processus puisse être automatisé. S'il fallait choisir entre les deux, nous pouvons conclure que l'un ou l'autre pourrait aller, puisque les deux contiennent les mots-clés essentielles pour l'indexation.

Afin de comparer les mots-clés en anglais avec ceux en français, nous avons utilisé un court film intitulé *Home Security / En toute sécurité*, produit par l'Office national du film du Canada, lequel offrait l'audiovision dans les deux langues. Chaque mot-clé provenant de l'anglais (langue originale du film) était filtré par le dictionnaire, afin de voir si la traduction française offerte par le dictionnaire contenait le mot en français de l'audiovision du film. Nous avons créé un fichier Excel pour analyser les données (figure 3). La colonne A donne les mots-clés en anglais, tels qu'ils paraissent dans le scénario pour l'audiovision fourni par l'Office national du film. La colonne B donne les mots-clés en français provenant du scénario

pour l'audiovision, fourni également par l'Office national du film. La colonne C contient les résultats provenant de Reverso Online / dictionnaires *Collins*. La colonne D contient un X lorsqu'il y a équivalence et la colonne E contient un X lorsqu'il n'y en a pas.

	A	B	C	D	E
	Keyword	Mot clé	Reverso	OUT	NON
14	ball	balle	balle	X	
15	bark	aboyer	aboyer	X	
16	barred	recouvertes de barreaux	baricadé		X
17	baseball	baseball	base-ball	X	
18	bat	batte	batte	X	
19	beaten	reçoit un coup	battu		X
20	beats	cogne	battre		X
21	behind	derrière	derrière	X	
22	bells	sonneries	sonnerie	X	
23	biker	motard	motard	X	
24	blue	bleu	bleu	X	
25	board	office	conseil		X
26	bolts	part en courant	filer		X
27	bombs	bombes	bombes	X	
28	boys	garçons	garçons	X	
29	breaking into	entre	s'introduire par effraction		X
30	breaks	se brise	se brise	X	
31	building			0	
32	bulges			0	
33	burglar	voleur	cambricoleur		X
34	butterfly	papillon	papillon	X	
35	button	bouton	bouton	X	
36	cage	cage	cage	X	
37	Canada	Canada	Canada	X	
38	car			0	
39	cat	chat	chat	X	
40	ceiling	chandelier	plafond		X
41	channel	chaine	chaine	X	
42	chase	poursuit	poursuit	X	
43	chattering	claquant	bavarder		X
44	city	ville	ville	X	
45	closes	ferme	fermer		X
46	clothes	vêtements	vêtements	X	
47	commercial	publicité	publicitaire	X	
48	content	content	content	X	
49	continue	continuent	continuant	X	
50	cops	policiers	flics		X
51	crashes	traverse	percuter		X
52	criminal	bandit	criminel		X

Figure 3. Les mots-clés en anglais avec équivalents en français, et les mots-clés générés par le traducteur automatique.

Au cours de ce film de cinq minutes, il y avait un total de 252 mots-clés en anglais dans le scénario de l'audiovision. Parmi ceux-ci, 43 n'avaient pas d'équivalent en français (à cause de constructions linguistiques dans la version française du film). Dans le fichier Excel, ce phénomène est représenté par le chiffre zéro dans les colonnes B et C. Parmi les 209 autres mots-clés, 154 avaient un terme équivalent en français et 55 n'en avaient pas. Environ 74 pour cent des termes avaient donc un terme correspondant qui était directement équivalent. Dans le contexte de production de termes d'indexation, ce résultat est encourageant parce qu'il indique qu'environ trois quarts du temps, un traducteur automatique générera des mots-clés utiles dans une autre langue (du mois entre l'anglais et le français). Cependant, dans notre analyse, nous n'avons pas considéré les synonymes. Si, par exemple, le traducteur était jumelé à un thésaurus ou autre réseau sémantique, ce pourcentage pourrait être amélioré. Nous observons, par exemple, dans l'analyse de la figure 3, que lorsqu'il n'y a pas d'équivalence, les mots-clés des colonnes B et C peuvent être considérés comme synonymes ou termes reliés de façon assez proche.

Discussion

Nos résultats indiquent que du texte créé pour d'autres fins peut servir pour l'indexation de plans ou de séquences de films. Dans le cas de mots-clés provenant de descriptions écrites ou de textes d'audiovision, on y trouve assez de termes utiles à l'indexation pour penser à exploiter ces textes. Lorsqu'on ajoute d'autres sources textuelles, par exemple les sous-titres

pour malentendants, lesquels représentent les dialogues et la narration de films, on peut se servir de ce bassin de textes pour produire de l'indexation assez complète aux plans et aux séquences. On peut appliquer ces résultats de recherche en trouvant des moyens pratiques d'exploiter les fichiers textuels afin d'y faire le forage de mots-clés et de les associer à des plans et à des séquences. On peut imaginer que cela se fera assez facilement, au fur et à mesure que la production de films et de télévision se fait en format numérique. Dans un tel environnement, on peut construire des bases de données modulaires pour la gestion de la production, lesquelles contiendraient déjà plusieurs sortes de textes. À ces modules, on ajoute ceux de la postproduction. Il ne restera qu'à programmer le système pour identifier et générer automatiquement les mots-clés, pour ensuite les attacher aux plans ou séquences correspondants, fournissant ainsi les termes d'indexation. De telles activités sont relativement faciles à programmer. Il y a d'autres possibilités pour l'exploitation de tels textes, par exemple les lier à des outils de gestion de vocabulaire ou encore pour générer automatiquement l'indexation dans d'autres langues.

Les films d'animation étudiés dans le cadre de notre projet sont intéressants parce qu'on y trouve des schèmes qui ne figurent pas dans les films étudiés précédemment. Plusieurs choses l'expliquent. Une partie de ces films sont expérimentaux, comprenant des techniques comme le dessin ou la gravure de l'image directement sur pellicule et donc sans caméra. D'autres films emploient la technique de la pixillation, où des comédiens bougent seulement un peu entre prises fixes pour créer un effet d'animation. C'est un autre cas où la notion de « plan » ne s'applique pas.

En plus de ces questions par rapport aux plans, les films d'animation manquaient souvent de dialogues, ce qui fait qu'il n'y avait pas de problème d'espaces sonores disponibles pour l'insertion d'épisodes d'audiovision. Dans ces films, on entend l'audiovision presque toujours (98 pour cent du temps) en même temps que l'image correspondant s'affiche à l'écran, et non pas avant ou après, comme c'est souvent le cas de films tournés avec des comédiens. On trouve donc dans une partie des films étudiés des schémas non rencontrés dans nos travaux antérieurs.

Une autre question éclairée par cette recherche est la très grande variété de problèmes de perte de vision et de styles cognitifs parmi les usages. Ainsi, il serait hautement désirable de pouvoir offrir aux utilisateurs la possibilité de paramétrer l'audiovision lorsqu'ils regardent un film. Un paramètre serait le niveau de détail dans la description. Lors de la préparation de textes d'audiovision, on pourrait identifier et étiqueter des épisodes comme étant de niveau 1, 2 ou 3, par exemple. Chaque utilisateur pourrait alors choisir un niveau de description pour le visionnement. Un autre paramètre serait les types d'information contenus dans la description. Notre typologie pourrait servir de base pour offrir des possibilités aux utilisateurs. Ainsi, ceux qui voient assez bien pour savoir s'il fait soleil ou s'il pleut, par exemple, pourraient alors choisir de ne pas entendre cette information. C'est ce genre de problème que nous étudierons au cours de la prochaine étape du projet, en 2008 et 2009.

Conclusions

Ce rapport traite de deux aspects du travail accompli dans le contexte du projet du Réseau de recherche E-Inclusion. D'autres rapports traitent d'autres aspects, et un rapport général de recherche est disponible (Mathieu et Turner 2007b). Dans le présent rapport on fait état de l'utilité de « recycler » du texte produit aux fins de description de l'image à l'intention des aveugles et malvoyants afin de générer automatiquement de l'indexation de plans ou de séquences d'images en mouvement. Nous avons trouvé que des mots-clés provenant de ces textes sont souvent utiles comme termes d'indexation.

On évalue par ailleurs la possibilité de générer automatiquement des mots-clés dans

d'autres langues, afin de produire de l'indexation bilingue et éventuellement multilingue. En mesurant la simple correspondance entre termes en anglais et français, sans tenir compte de synonymes ou de quasi-synonymes, nous avons trouvé que trois quarts du temps il y a correspondance, soit assez pour exploiter avantageusement cette technique.

Notre étude contribue aux travaux sur la problématique générale de l'utilisation de textes créés pour d'autres fins comme source pour le forage de termes d'indexation de plans individuels ou de séquences de films ou d'émissions de télévision. De concert avec d'autres types de textes, ceux étudiés dans le cadre du présent projet forment un corpus riche en mots-clés disponibles pour recyclage comme termes d'indexation, pourvu qu'on puisse automatiser les processus nécessaires pour y arriver. Ainsi, on pourrait offrir l'accès aux plans et aux séquences de films et émissions de télévision, ce qui serait autrement impossible parce que trop coûteux à produire par des indexeurs humains.

Remerciements

Nous remercions Patrimoine canadien pour sa contribution au financement du réseau de recherche E-Inclusion. Nous remercions également nos partenaires au Centre de recherche en informatique de Montréal (CRIM), pour leurs précieux conseils et leur direction, et pour l'atmosphère de collaboration amicale dans laquelle nous avons pu faire notre travail. Nous sommes bien reconnaissants à Mme Anne Jarry de l'Institut national canadien pour les aveugles (INCA) pour son aide avec l'organisation de sessions de visionnement avec des participants aveugles et malvoyants, ainsi que pour ses observations et ses conseils toujours astucieux. Nous remercions chaleureusement tous les participants qui ont accepté de nous fournir des données en offrant leurs opinions et leur point de vue sur des façons d'améliorer l'audiovision. Merci à Alain Lapierre, de l'INCA également, pour son aide technique et organisationnelle. Finalement, nous tenons à remercier Linda Rivest, notre assistante de recherche, qui a préparé pour l'analyse une bonne partie des données.

References

- Independent Television Commission [Royaume-Uni]. 2000. ITC Guidance on standards for audio description / Nick Lodge.
<http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/archive/audio_description_stnds/itcguide_sds_audio_desc_word.doc>.
- Mathieu, Suzanne, et James M Turner. 2007a. Audiovision ou comment faire voir l'information par les personnes aveugles et malvoyantes : lignes directrices pour la description d'images en mouvement. Congrès annuel de l'Association canadienne des sciences de l'information, *Actes*. <http://www.cais-acs.ca/proceedings/2007/mathieu_2007.pdf>.
- Mathieu, Suzanne, et James M Turner. 2007b. Réseau de recherche E-Inclusion, thème 3 : Audiovision interactive et adaptable, projet 3.1, rapport final.
<<http://hdl.handle.net/1866/1307>>.
- Piety, Philip J. 2003. Audio description, a visual assistive discourse: an investigation into language used to provide the visually disabled access to information in electronic texts. Master of Arts thesis. Washington: Georgetown University.
- Reverso Online. 2007. <<http://www.reverso.net>>.
- Turner, James M. et Emmanuël Colinet. 2004. Using audio description for indexing moving images. *Knowledge Organization* 31, no. 4, 222-230.
- Turner, James M. et Michèle Hudon. 2002. Multilingual metadata for moving image databases: preliminary results. *Advancing Knowledge: Expanding Horizons for*

- Information Science, Proceedings of the 30th annual conference of the Canadian Association for Information Science*, ed. Lynne C. Howarth, Christopher Cronin, Anna T. Slawek. Toronto: Faculty of Information Studies, 34-45.
- Turner, James M. 1998. Some characteristics of audio description and the corresponding moving image. *Proceedings of the 61st ASIS annual meeting, Pittsburgh, PA, October 24-29, 1998*, ed. Cecilia M Preston. Medford, NJ: Information Today, 108-117.