Date : 07/08/2007

Audio description text und Filmerschließung

James M Turner
École de bibliothéconomie et des sciences de l'information
Université de Montréal, Montréal,
Québec, Canada

Suzanne Mathieu
Ville de Montréal, Montréal,
Québec, Canada

Übersetzung: Michaela Probst,
Deutsche Nationalbibliothek, Leipzig

Meeting: **157 Classification and Indexing**
Simultaneous Interpretation: No

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

Abstract

Der Zugriff auf audiovisuelle Medien sollte ebenso ungehindert und bequem sein wie auf Printpublikationen. Allerdings ist dies bisher nicht der Fall. Denn Erschließungsmethoden, die für konventionelle Medien entwickelt wurden, eignen sich nur bedingt für die Aufbereitung von audiovisuellem und multimedialem Material. In unserem Projekt haben wir untersucht, wie a*udio description text* und Beschreibungstexte zur Erzeugung von Indextermen bei der Erschließung von Filmen genutzt werden können. Diese Quellen waren sehr ergiebig und nützlich. Im zweiten Teil der Untersuchung haben wir uns mit den Möglichkeiten der automatisierten Übersetzung von Stichwörtern aus a*udio description text* in andere Sprachen beschäftigt. Auch hier können wir ermutigende Ergebnisse vorlegen.

Einführung

Die *IFLA Guidelines for audiovisual and multimedia materials in libraries and other institutions* fordern, daß der Zugriff auf audiovisuellen Medien ebenso ungehindert und bequem sein sollte wie auf Printpublikationen. Allerdings ist dies bisher nicht der Fall. Denn Erschließungsmethoden, die für konventionelle Medien entwickelt wurden, eignen sich nur bedingt für die Aufbereitung von audiovisuellen und multimedialen Werken. Es sind noch große Anstrengungen nötig, um Nutzern zukünftig

einen besseren Zugriff auf solches Material zu ermöglichen. Wir möchten über unsere Arbeit im Rahmen des E-Inclusion Research Network (http://e-inclusion.crim.ca/?q=en) berichten. Dieser Forschungsverbund hat das Ziel, „leistungsstarke Ton-Bild-Werkzeuge…zur Verbesserung der Wahrnehmungsmöglichkeiten von Hör- und Sehbehinderten zu entwickeln“. Im Forschungsverbund, der beim CRIM (*Centre de recherche informatique de Montréal or Montréal Centre for Computer Research*) angesiedelt ist, arbeiten verschiedene Universitäten und Partner, wie *National Film Board of Canada* (NFB) und CNIB/INCA (früher *Canadian National Institute for the Blind*) sowie AudioVision, ein Anbieter für akustische Bildbeschreibung, zusammen.

Projekt 3.1 des Forschungsverbundes beschäftigt sich u.a. mit der Identifizierung von Informationen, die Sehbehinderte benötigen, um einen Film erfassen zu können. Die Zielstellung des Projekts war die Erarbeitung von Richtlinien für Personen, die Filmbeschreibungen für diese Zielgruppe erstellen, sogenannte *describer*. Um das Problem besser zu verstehen, haben wir mit Sehbehinderten eine Reihe von Filmen angesehen und sie um eine Einschätzung der Audiodeskription gebeten. Wir wollten feststellen, ob die Beschreibung als ausreichend empfunden oder ob Verbesserungen gewünscht werden. Ein Ziel des *E-Inclusion Research Network* ist die Entwicklung von variablen Beschreibungen, die von den Nutzern selbst gesteuert werden können, z.B. der Umfang der Beschreibung oder die Beschränkung auf bestimmte Arten von Informationen. Dies ist notwendig, weil es eine große Bandbreite von Sehbehinderungen gibt, und darüber hinaus in der alternden Gesellschaft immer mehr Menschen unter altersbedingt eingeschränkter Sehkraft leiden. Außerdem schätzen manchmal sogar Sehende Filme mit akustischer Bildbeschreibung, da diese Zusatzinformationen zum besseren Verständnis des Films beitragen.

Um diese Probleme untersuchen zu können, mussten wir zunächst die Eigenschaften der Filme, die wir mit den Sehbehinderten angeschaut haben, genau bestimmen. Dazu waren umfangreiche Hintergrundarbeiten nötig. Im nächsten Abschnitt berichten wir aus dieser Arbeit. Bei der Analyse der Filme haben wir uns auch mit der Frage beschäftigt, ob das Vokabular, das bei der akustischen Bildbeschreibung verwendet wird, für die Erschließung nutzbar gemacht werden kann – dieses Teilproblem ist der Schwerpunkt des Textes.

Kontext

Seit 1995 arbeiten wir an einer Reihe von Projekten, die a*udio description text* auf eine mögliche Verwendung bei der Erschließung von Filmen untersuchen. Wir waren besonders an der Frage interessiert, wie einzelne Einstellungen in Filmen, Fernsehproduktionen und anderen audiovisuellen Doku-

menten erschlossen werden können. Im Rahmen unserer Untersuchungen haben wir ein Verfahren zur Transkription von a*udio description text* und der Verknüpfung mit den Filmsequenzen entwickelt (Turner 1998), in späteren Projekten wurde diese Methode ausgebaut (Turner und Colinet 2004, Mathieu und Turner 2007). Die Grundidee war, dass einzelne „Episoden“ der akustischen Bildbeschreibung, quasi Informationspartikel, bestimmten Themeneinheiten zugeordnet sind. Wenn sich das Thema der Beschreibung ändert, beginnt auch eine neue „Episode“. Da akustische Bildbeschreibung im Zwei-Kanal-Ton-Verfahren angeboten wird, ist der Beschreibungstext nicht immer dann zu hören, wenn das beschriebene Bild erscheint. Wir haben die Plazierung der Beschreibungen untersucht und erfasst, ob sie in der gleichen Einstellung wie das Bild zu hören sind oder eine Einstellung vor oder nach dem Erscheinen des Bildes, zwei Einstellungen früher oder später und so weiter.
Durch das Verknüpfen des Beschreibungstextes mit dem korrespondierenden Bild in einer Filmdatenbank, die auch verzeichnet, ob Text und Bild zum gleichen Zeitpunkt gesendet werden, und durch das Identifizieren von geeigneten Begriffen ist es möglich, jede einzelne Einstellung eines Films zu erschließen. Dies ist für alle hilfreich, die Filme und Fernsehproduktionen selbst herstellen oder sie untersuchen. Besonders nützlich ist diese Form der Erschließung für die tägliche Arbeit von Fernsehredaktionen, da so jederzeit der schnelle Zugriff auf das Material gewährleistet werden kann.
Filmdatenbanken, die jede einzelne Einstellung erschließen, enthalten eine Beschreibung, die in Textform wiedergibt, was in der jeweiligen Einstellung zu sehen ist. So brauchen Nutzer zunächst nur den Text überfliegen und können dann eine kleinere Anzahl von Filmsequenzen anschauen, um schließlich die endgültige Auswahl zu treffen. Aufgrund dieser Vorgehensweise kann der Recherchevorgang effektiv gestaltet werden, denn es wird nur wenig Zeit für das Anschauen der Filmsequenzen benötigt.
Wir haben uns die Frage gestellt, ob a*udio description text* in ausreichendem Maße redundant ist, um diese Form der Beschreibung zu ersetzen. Wenn dies der Fall wäre, könnte zukünftig auf das Verfassen der Beschreibungen verzichtet werden und die Erschließung ausschließlich mit Hilfe von a*udio description text* erfolgen. Die Ergebnisse unserer vorhergehenden Untersuchung haben allerdings gezeigt, daß a*udio description text* zwar oft einer Sequenz zugeordnet ist, aber nicht unbedingt einer bestimmten Einstellung. Daß dies vom jeweiligen Film abhängt, sollen die folgenden Beispiele erläutern. Einige Filme zeichnen sich durch eine sehr unruhige Handlung aus, und das Interesse des Zuschauers wird durch viele kurze, schnell aufeinanderfolgende Schnitte aufrechterhalten. In diesem Fall ist es nicht unbedingt sinnvoll, jede Einstellung zu beschreiben, sonder eher das Geschehen der gesamten Filmsequenz. Bei anderen Filmen verläuft die Handlung eher ruhig, und die Einstellungen sind relativ lang, so daß die Plazierung von akustischen Beschreibungen keine Probleme bereitet, da häufig nur Hintergrundmusik zu hören ist und es keine Dialoge gibt. In einem solchen Fall ist ausreichend

Zeit für die Beschreibung des Bildinhaltes jeder einzelnen Einstellung. Einige Filme, die wir im Rahmen unseres Projekts untersucht haben, z.B. die experimentellen Animationsfilme von Norman McLaren, sind kurz und haben nur Hintergrundmusik. Man könnte sogar sagen, dass der gesamte Film aus nur einer einzigen Einstellung besteht bzw. nicht einmal aus einer Einstellung im engeren Sinne, wenn das Bild auf das Filmmaterial gezeichnet oder eingeritzt ist. Filme und andere audiovisuelle Medien sind so unterschiedlich strukturiert, dass es nicht möglich ist, allgemeingültige Regeln aufzustellen. Vielmehr ist es nötig, eigene Erschließungsmodelle für die verschiedenen Typen zu erarbeiten.

E-Inclusion Network projekt

Die Untersuchung befaßt sich mit drei Aspekten der Themen der *Classification and Indexing Section*: sachlicher Zugriff in Minderheitssprachen, Zusammenarbeit mit Nutzern bei der Entwicklung von Suchwerkzeugen für den sachlichen Zugriff und Erleichterung und Ausbau des sachlichen Zugriffs auf Bibliotheksbestände. Der sachliche Zugriff in Minderheitssprachen, der nicht das eigentliche Thema dieses Papiers ist, steht in Beziehung zu den bilingualen (Englisch und Französisch) Aspekten unseres Projekts, die an frühere Untersuchungen (z.B. Turner/Hudon 2002) zur Erweiterung des Anwendungsbereichs auf andere Sprachen anknüpfen. Durch die Zusammenarbeit mit Nutzern und insbesondere die Gespräche mit Sehbehinderten haben wir viel darüber erfahren, wie a*udio description text* zum Verständnis von Filmen beiträgt. Die Analyse dieser Gespräche, die im Anschluß an das gemeinsame Anschauen der Filme stattfanden, ist in die Ausarbeitung unserer Richtlinien eingeflossen. Außerdem helfen die Hinweise auf besonders geeignete Formen der Beschreibung dabei, ein auf diese Zielgruppe ausgerichtetes Erschließungsvokabular zu entwickeln. Die Erleichterung und der Ausbau des sachlichen Zugriffs soll durch die Erschließung einzelner Einstellungen, die bisher nur in Filmarchiven und Fernsehredaktionen verbreitet war, erreicht werden.
Die Analyse von a*udio description text* der untersuchten Filme dient der Überprüfung der Typologie von Beschreibungsarten, die wir in früheren Projekten herausgearbeitet haben. Durch die größere Menge der untersuchten Filme sind wir nun in der Lage, unsere Ergebnisse durch Vergleich mit den schon vorliegenden Resultaten abzusichern und zu entscheiden, ob die Einführung neuer Kategorien notwendig ist oder ob die schon vorliegenden für die Beschreibung der untersuchten Filme ausreichen. Bisher gibt es nur wenige Untersuchungen dazu (daher beziehen sich die Verweise zumeist auf unsere eigene Arbeit), aber jede weitere Filmanalyse hilft uns, die Ergebnisse zu verfeinern.
Im Zusammenhang mit den in diesem Papier vorgestellten Ergebnissen hat unser Projekt folgende Aufgaben:

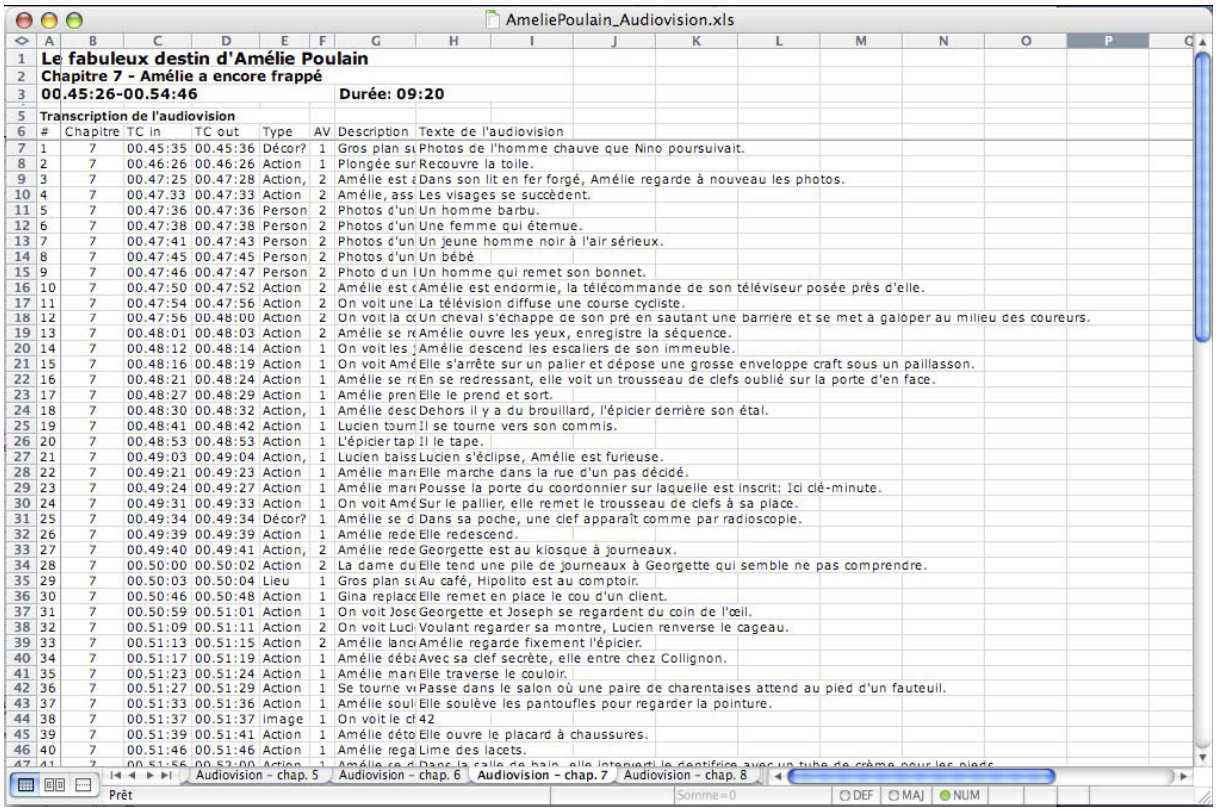
Überprüfen der Typologie von Beschreibungen, die a*udio description text* anbietet.
Vergleich der Stichwörter von a*udio description text* mit den Nutzer-Stichwörtern für die einzelnen Einstellungen, um die Eignung dieser Texte als Indexierungsgrundlage zu bewerten.
Vergleich der englischen und französischen Stichwörter, um die Eignung von automatischen Indexierungsverfahren für diese und andere Sprachen zu überprüfen.
Wenn diese Aufgaben abgeschlossen sind, werden unsere Kenntnisse hoffentlich ausreichen, um die offen Fragen zufriedenstellend beantworten zu können. Im nächsten Abschnitt berichten wir von unserer Vorgehensweise.

Methodik

Im Rahmen unseres Projekts haben wir elf Filme analysiert, zwei in Spielfilmlänge, zwei Dokumentarfilme und sieben kurze Animationsfilme. Für die Analyse wurde jeder Film in Sequenzen und einzelne Einstellungen zerlegt, die akustische Beschreibung in Abschnitte gegliedert sowie jede Einstellung mit einer (Bild-)Beschreibung in Freitext versehen. Die letztgenannte Aufgabe hat ein Forschungsassistent übernommen, der die Filme ohne Ton angeschaut hat, um nicht von Dialogen und Erklärungen sowie der akustischen Bildbeschreibung beeinflusst zu werden. Die Arbeitsergebnisse bildeten die Grundlage der Datenbanken, die wir für die Analyse benutzt haben.
Wir haben Standardsoftware benutzt, um die einzelnen Segmente der Filme durch einen Zeitcode zu erfassen und sie so genau identifizieren zu können. Der Zeitcode gliedert sich in Stunde, Minute, Sekunde und Bild. Filme aus dem Bestand des National Film Board of Canada wurde in einer DVD-Version mit Zeitcode für uns bereitgestellt.
Im Hinblick auf unsere Fragestellung haben wir die Abschnitte der akustischen Bildbeschreibung bestimmten Kategorien zugeordnet, die in einem früheren Projekt (Turner 1998) entwickelt und in einem späteren Projekt (Turner and Colinet 2004) ausgebaut wurden. Zunächst haben wir erwogen, für unsere Analyse die Typologie von Piety (2003) zu nutzen, die aber für unsere Zwecke zu grob strukturiert war. Die *ITC guidance on standards for audio description* (Independent 2000) waren dabei sehr hilfreich.
Zur Erfassung der Analysedaten haben wir für jeden Film eine Excel-Datei (Abb. 1) erstellt. Die Sequenzen und Einstellungen der Filme sind jeweils mit der Beschreibung und dem dazugehörigen *audio description text* verknüpft. Filme in Spielfilmlänge wurden kapitelweise in eigenen Tabellen erfasst, um angesichts der Fülle des Materials nicht den Überblick zu verlieren und die für unsere Untersuchung notwendigen Berechnungen zu ermöglichen. Dies stellt eine Verbesserung dar, da wir in frühe-

ren Projekten Software benutzt haben, die Informationen zu den einzelnen Einstellungen, aber keinen Überblick über das gesamte Material geboten hat.

Die Excel-Datei enthält zahlreiche Dateneinheiten und ermöglicht so verschiedene Auswertungen. In diesem Papier beschränken wir uns auf Aspekte, die mit Fragestellungen zusammenhängen, die in der *Classification and Indexing Section* thematisiert werden, einerseits das Verhältnis von *audio description text* zur Beschreibung in Freitext und andererseits die Verknüpfung von englischen und französischen Stichwörtern.



AmeliePoulain_Audiovision.xls

Le fabuleux destin d'Amélie Poulain
Chapitre 7 - Amélie a encore frappé
00.45:26-00.54:46 **Durée: 09:20**

Transcription de l'audiovision

#	Chapitre	TC in	TC out	Type	AV	Description	Texte de l'audiovision
1	7	00.45:35	00.45:36	Décor?	1	Gros plan su	Photos de l'homme chauve que Nino poursuivait.
2	7	00.46:26	00.46:26	Action	1	Plongée sur	Recouvre la toile.
3	7	00.47:25	00.47:28	Action,	2	Amélie est a	Dans son lit en fer forgé, Amélie regarde à nouveau les photos.
4	7	00.47.33	00.47:33	Action	2	Amélie, ass	Les visages se succèdent.
5	7	00.47:36	00.47:36	Person	2	Photos d'un	Un homme barbu.
6	7	00.47:38	00.47:38	Person	2	Photos d'un	Une femme qui éternue.
7	7	00.47:41	00.47:43	Person	2	Photos d'un	Un jeune homme noir à l'air sérieux.
8	7	00.47:45	00.47:45	Person	2	Photos d'un	Un bébé
9	7	00.47:46	00.47:47	Person	2	Photo d un l	Un homme qui remet son bonnet.
10	7	00.47:50	00.47:52	Action	2	Amélie est c	Amélie est endormie, la télécommande de son téléviseur posée près d'elle.
11	7	00.47:54	00.47:56	Action	2	On voit une	La télévision diffuse une course cycliste.
12	7	00.47:56	00.48:00	Action	2	On voit la co	Un cheval s'échappe de son pré en sautant une barrière et se met à galoper au milieu des coureurs.
13	7	00.48:01	00.48:03	Action	2	Amélie se re	Amélie ouvre les yeux, enregistre la séquence.
14	7	00.48:12	00.48:14	Action	1	On voit les j	Amélie descend les escaliers de son immeuble.
15	7	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 paillasson.
16	7	00.48:21	00.48:24	Action	1	Amélie se re	En se redressant, elle voit un trousseau de clefs oublié sur la porte d'en face.
17	7	00.48:27	00.48:29	Action	1	Amélie pren	Elle le prend et sort.
18	7	00.48:30	00.48:32	Action,	1	Amélie desc	Dehors il y a du brouillard, l'épicier derrière son étal.
19	7	00.48:41	00.48:42	Action	1	Lucien tourn	Il se tourne vers son commis.
20	7	00.48:53	00.48:53	Action	1	L'épicier tap	Il le tape.
21	7	00.49:03	00.49:04	Action,	1	Lucien baiss	Lucien s'éclipse, Amélie est furieuse.
22	7	00.49:21	00.49:23	Action	1	Amélie mar	Elle marche dans la rue d'un pas décidé.
23	7	00.49:24	00.49:27	Action	1	Amélie mar	Pousse la porte du coordonnier sur laquelle est inscrit: Ici clé-minute.
24	7	00.49:31	00.49:33	Action	1	On voit Amé	Sur le pallier, elle remet le trousseau de clefs à sa place.
25	7	00.49:34	00.49:34	Décor?	1	Amélie se d	Dans sa poche, une clef apparaît comme par radioscopie.
26	7	00.49:39	00.49:39	Action	1	Amélie rede	Elle redescend.
27	7	00.49:40	00.49:41	Action,	2	Amélie rede	Georgette est au kiosque à journeaux.
28	7	00.50:00	00.50:02	Action	2	La dame du	Elle tend une pile de journeaux à Georgette qui semble ne pas comprendre.
29	7	00.50:03	00.50:04	Lieu	1	Gros plan su	Au café, Hipolito est au comptoir.
30	7	00.50:46	00.50:48	Action	1	Gina replace	Elle remet en place le cou d'un client.
31	7	00.50:59	00.51:01	Action	1	On voit Jose	Georgette et Joseph se regardent du coin de l'œil.
32	7	00.51:09	00.51:11	Action	2	On voit Luci	Voulant regarder sa montre, Lucien renverse le cageau.
33	7	00.51:13	00.51:15	Action	2	Amélie lanc	Amélie regarde fixement l'épicier.
34	7	00.51:17	00.51:19	Action	1	Amélie déb	Avec sa clef secrète, elle entre chez Collignon.
35	7	00.51:23	00.51:24	Action	1	Amélie mar	Elle traverse le couloir.
36	7	00.51:27	00.51:29	Action	1	Se tourne v	Passe dans le salon où une paire de charentaises attend au pied d'un fauteuil.
37	7	00.51:33	00.51:36	Action	1	Amélie soul	Elle soulève les pantoufles pour regarder la pointure.
38	7	00.51:37	00.51:37	Image	1	On voit le ch	42
39	7	00.51:39	00.51:41	Action	1	Amélie déto	Elle ouvre le placard à chaussures.
40	7	00.51:46	00.51:46	Action	1	Amélie rega	Lime des lacets.
41	7	00.51:56	00.52:00	Action	1	Amélie se d	Dans la salle de bain, elle intervertit le dentifrice avec un tube de crème pour les pieds

Audiovision - chap. 5 | Audiovision - chap. 6 | Audiovision - chap. 7 | Audiovision - chap. 8

Prêt | Somme=0 | DEF | MAJ | NUM

Abb. 1:

Exeltabelle mit *audio description text* nach Einstellungen geordnet

Durch den Vergleich des Beschreibungstextes mit der akustischen Beschreibung können wir für jede einzelne Einstellung berechnen, wie häufig Stichwörter in beiden Texten auftreten. Aus den Ergebnissen ergibt sich, ob eine automatisierte Indexierung sinnvoll und welcher der beiden Texte für diesen Zweck besonders geeignet ist. Wir haben hier die Tabelle für den Film *Le fabuleux destin d'Amélie*

Poulain, der zu den Filmen gehört, die im Rahmen des *E-Inclusion Projects* von einer Reihe von Wissenschaftlern intensiv untersucht wurden. Eine Version der Tabelle (Abb. 2) zeigt die Spalten A, G und H zu Kapitel 7 des Films, in denen die Nummer der Einstellung, die Freitextbeschreibung und *audio description text* zu sehen sind.

AmeliePoulainJamesAnalyse.xls

	A	G	H
1	**Le fabuleux destin d'Amélie Poulain**		
2	**Chapitre 7 - Amélie a encore frappé**		
3	**0**	**Durée: 09:20**	
5	Tr		
6	#	Description	Texte de l'audiovision
7	1	Gros plan sur la photo de l'homme chauve.	Photos de l'homme chauve que Nino poursuivait.
8	2	Plongée sur le tableau que M. Dufayel s'apprête à couvrir.	Recouvre la toile.
9	3	Amélie est assise dans son lit, un album à la main.	Dans son lit en fer forgé, Amélie regarde à nouveau les photos.
10	4	Amélie, assise dans son lit, regarde l'album.	Les visages se succèdent.
11	5	Photos d'un homme barbu.	Un homme barbu.
12	6	Photos d'une femme qui éternue.	Une femme qui éternue.
13	7	Photos d'un homme noir.	Un jeune homme noir à l'air sérieux.
14	8	Photos d'un bébé	Un bébé
15	9	Photo d'un homme avec une main dans son bonnet.	Un homme qui remet son bonnet.
16	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.
17	11	On voit une course de vélo dans un écran de télévison.	La télévision diffuse une course cycliste.
18	12	On voit la course de vélos dans l'écran télé puis Amélie qui se réveille.	Un cheval s'échappe de son pré en sautant une barrière et se met à galoper au milieu
19	13	Amélie se réveille et appuie sur sa télécommande.	Amélie ouvre les yeux, enregistre la séquence.
20	14	On voit les jambes d'Amélie qui descendent les marches d'un escalier.	Amélie descend les escaliers de son immeuble.
21	15	On voit Amélie qui s'agenouille pour mettre une enveloppe sous un paillasson.	Elle s'arrête sur un palier et dépose une grosse enveloppe kraft sous un paillasson.
22	16	Amélie se relève de devant le paillasson et se retourne.	En se redressant, elle voit un trousseau de clefs oublié sur la porte d'en face.
23	17	Amélie prend le trousseau de clefs.	Elle le prend et sort.
24	18	Amélie descend les escaliers.	Dehors il y a du brouillard, l'épicier derrière son étal.
25	19	Lucien tourne son visage vers l'épicier.	Il se tourne vers son commis.
26	20	L'épicier tape la tête de Lucien.	Il le tape.
27	21	Lucien baisse la tête et s'en va de derrière l'étal.	Lucien s'éclipse, Amélie est furieuse.
28	22	Amélie marche furieusement dans la rue.	Elle marche dans la rue d'un pas décidé.
29	23	Amélie marche furieusement dans la rue.	Pousse la porte du coordonnier sur laquelle est inscrit: Ici clé-minute.
30	24	On voit Amélie qui remet les clefs dans la serrure.	Sur le pallier, elle remet le trousseau de clefs à sa place.
31	25	Amélie se détourne de la serrure et on voit une clef sur elle.	Dans sa poche, une clef apparaît comme par radioscopie.
32	26	Amélie redescend les escaliers.	Elle redescend.
33	27	Amélie redescend les escaliers.	Georgette est au kiosque à journeaux.
34	28	La dame du kiosque à journeaux remet des journeaux à Georgette.	Elle tend une pile de journeaux à Georgette qui semble ne pas comprendre.
35	29	Gros plan sur le visage de la dame du kiosque aux journeaux.	Au café, Hipolito est au comptoir.
36	30	Gina replace le cou d'un vieil homme.	Elle remet en place le cou d'un client.
37	31	On voit Joseph qui regarde le tabac de Georgette par un miroir.	Georgette et Joseph se regardent du coin de l'œil.
38	32	On voit Lucien qui échappe le cageau, Amélie qui arrive.	Voulant regarder sa montre, Lucien renverse le cageau.
39	33	Amélie lance un regard noir à l'épicier.	Amélie regarde fixement l'épicier.
40	34	Amélie débarre la porte puis entre chez Collignon.	Avec sa clef secrète, elle entre chez Collignon.
41	35	Amélie marche dans le couloir.	Elle traverse le couloir.
42	36	Se tourne vers une autre pièce et y entre.	Passe dans le salon où une paire de charentaises attend au pied d'un fauteuil.
43	37	Amélie soulève une paire de pantoufles qu'elle retourne.	Elle soulève les pantoufles pour regarder la pointure.
44	38	On voit le chiffre 42 sous une pantoufle.	42
45	39	Amélie détourne la tête	Elle ouvre le placard à chaussures.
46	40	Amélie regarde les étagères d'un garde-robe.	Lime des lacets.

Audiovision - chap. 4 | Audiovision - chap. 5 | Audiovision - chap. 6 | Audiovision - chap. 7 | Au

Prêt | Somme=50 | DEF | MAJ | NUM

Abb. 2:

Ansicht der Exeltabelle mit den Stichwörtern aus dem Beschreibungstextes und *audio description text*

Stichwörter, die sowohl in der Freitextbeschreibung als auch bei *audio description text* auftreten, erscheinen in Orange. Dabei ist zu beachten, daß die zusammengehörigen Stichwörter nicht immer derselben Einstellung zugeordnet sind, z.B. bei den Einstellungen 22-25. Der Grund dafür ist, daß die akustische Beschreibung nur dann eingefügt werden kann, wenn keine für das Verständnis wichtigen Geräusche oder Dialoge zu hören sind. Bei unserer Untersuchung haben wir deshalb – unabhängig von ihrer exakten Zuordnung in der Tabelle – alle Stichwörter berücksichtigt werden, die eine bestimmte Einstellung beschreiben.

Für den Abgleich französischer und englischer Stichwörter haben wir ein Verfahren angewandt, das sich schon in früheren Projekten bewährt hat. Ein zweisprachiges Wörterbuch bietet zu jedem Wort in

der Ausgangssprache eine oder mehrere Bedeutungen der Zielsprache. Beim Abgleich der beiden Texte in unterschiedlichen Sprachen sind Stichwörter dann verknüpft worden, wenn eine der Bedeutungen in der Zielsprache dem Schlüsselwort in der Ausgangssprache des Beschreibungstextes entsprach. Für diesen Abgleich haben wir die Software *Reverso Online* (Reverso 2007) mit hinterlegtem Collins-Wörterbuch (Englisch-Französisch), das zahlreiche ausführliche Begriffsbestimmungen enthält, benutzt.

Ergebnisse und Auswertung

Zunächst betrachten wir die Ergebnisse des Vergleichs von Beschreibungstext und *audio description text*. Im zweiten Teil der Auswertung geht es um die automatische Übersetzung englischsprachiger Stichwörter des Beschreibungstextes ins Französische. Abbildung 2 zeigt die Stichwörter aus Kapitel 7 des Films *Le fabuleux destin d'Amélie Poulain* in Orange.

Es wurden sechs Kapitel des Films untersucht, also insgesamt 238 Einstellungen. Nur bei 50 Einstellungen traten Stichwörter nur im Beschreibungstext und nicht in *audio description text* auf. Bei den übrigen 188 Einstellungen waren 463 Stichwörter in beiden Beschreibungsformen enthalten. Das sind pro Einstellung durchschnittlich etwa 2,5 Stichwörter.

Im Hinblick auf unsere Fragestellung, ob der Beschreibungstext für die Erzeugung von Stichwörtern besser geeignet ist als *audio description text*, haben wir erwartungsgemäß festgestellt, dass die wichtigen Begriffe in beiden Texten vorkommen. Allerdings gibt es in jedem Text eine Reihe weiterer Stichwörter (in schwarzer Schrift), für die es keine Entsprechung im anderen Text gibt. Wir gehen davon aus, dass es sinnvoll ist, beide Texte für Erzeugung von Stichwörtern, die dem Bild adäquat sind, einzusetzen, wenn der dafür nötige Verarbeitungsprozeß automatisiert werden kann. Wenn man sich für eine der beiden Textarten entscheiden müßte, wären wegen der Übereinstimmung bei wichtigen Begriffen beide für die Erzeugung von sinnvollem Erschließungsvokabular geeignet.

Den Abgleich der Stichwörter in Englisch und Französisch haben wir anhand des Kurzfilms *Home Security / En toutre sécurité* vorgenommen, der vom National Film Board of Canada produziert wurde und eine akustische Beschreibung in beiden Sprachen bietet. Wir haben jedes englische Schlüsselwort erfasst und überprüft, ob in der französische Übersetzung des Wörterbuchs der französische Begriff vorkommt, der auch in der akustischen Beschreibung enthalten ist. Zur Analyse der Daten haben wir eine Exceltabelle mit drei Spalten (Abb. 3) erstellt. Spalte A enthält die englischen Stichwörter aus dem Manuskript der akustischen Beschreibung des National Film Board. In Spalte B finden sich die entsprechenden französischen Stichwörter aus dem französischsprachigen Manuskript der akustischen

Beschreibung des National Film Board. Spalte C zeigt die Ergebnisse des Abgleichs mittels Reverso Online / Collins-Wörterbücher. Ein „x“ erscheint in Spalte D, wenn eine Verknüpfung vorliegt und in Spalte E, wenn keine Verknüpfung möglich ist.

AnalyseTraducteur.xls

	A	B	C	D	E
1	Keyword	Mot clé	Reverso	OUI	NON
14	ball	balle	balle	X	
15	bark	aboyer	aboyer	X	
16	barred	recouvertes de barreaux	barricadé		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	0		
32	bulges	0	0		
33	burglar	voleur	cambrioleur		X
34	butterly	papillon	papillon	X	
35	button	bouton	bouton	X	
36	cage	cage	cage	X	
37	Canada	Canada	Canada	X	
38	car	0	0		
39	cat	chat	chat	X	
40	ceiling	chandelier	plafond		X
41	channel	chaîne	chaîne	X	
42	chase	poursuit	poursuit	X	
43	chattering	claquant	bavarder		X
44	city	ville	ville	X	
45	closes	fonce	fermer		X
46	clothes	vêtements	vêtements	X	
47	commercial	publicité	publicitaire	X	
48	content	content	content	X	
49	continue	continuent	continuent	X	
50	cops	policiers	flics		X
51	crashes	traverse	percuter		X
52	criminal	bandit	criminel		X

Feuil1 Feuil2 Fe

Prêt

Abb. 3:

Englische Stichwörter, französische Entsprechungen und automatisch erzeugte Übersetzungen

Das Manuskript der akustischen Beschreibung zu diesem fünfminütigen Film enthält 252 englischsprachige Stichwörter. Aufgrund von spezifischen Sprachmustern des Französischen existiert zu 43 dieser Stichwörter keine Entsprechung in der französischen Version. Diese Fälle sind in den Spalten B und C der Tabelle durch „0“ gekennzeichnet. Von den übrigen 209 Begriffen haben 154 ein französisches Äquivalent und 55 keines. Demnach gibt es für etwa 74% der Begriffe eine direkte Entsprechung. Dieses Ergebnis ist sehr ermutigend für das Vorhaben, die Stichwörter zum Erzeugen von Erschließungsvokabular zu nutzen, da die automatisierte Übersetzung in 74% der Fälle brauchbare Stichwörter in der anderen Sprache liefert (zumindest bei Englisch und Französisch). Allerdings wäre die Trefferquote höher, wenn ein Thesaurus in die Übersetzungssoftware eingebunden wäre und somit auch Synonyme erkannt würden. In Abbildung 3 wird deutlich, daß das Übersetzungsprogramm in vielen Fällen zwar keine Übereinstimmung feststellen konnte, es sich bei den Betreffenden Begriffen aber um Synonyme oder eng verwandte Begriffe handelt.

Diskussion

Die hier vorgestellten Ergebnisse zeigen, dass Texte, die eigentlich anderen Zwecken dienen, einzelnen Filmeinstellungen oder -sequenzen zugeordnet und für die Erschließung verwendet werden können. Die Beschreibungstexte und *audio description text* enthalten eine ausreichend große Anzahl von Stichwörtern für die Erschließung von Filmen. Kombiniert man diese mit den Stichwörtern der Untertitel für Hörgeschädigte, die die Dialoge des Films wiedergeben, könnte eine beinahe vollständige Erschließung der Einstellungen und Sequenzen erzielt werden. Um diese Ergebnisse für die Praxis nutzbar zu machen, müssen Verfahren zur Einbeziehung der Textdaten bei der Erzeugung von Stichwörtern und ihrer Verknüpfung mit den Einstellungen und Sequenzen entwickelt werden. Wir denken, dass sich dieses Problem im Laufe der Zeit lösen wird, da immer mehr Filme in Digitalformat produziert werden und zur Verwaltung der Filmdateien Datenbanken notwendig sind. Anders gesagt, beschreibende Texte werden bereits im Zuge der Filmproduktion beim Erstellen der Datenbank entstehen, so daß die automatische Erzeugung von Stichwörtern und ihre Verknüpfung mit den Einstellungen und Sequenzen schon zu diesem Zeitpunkt erfolgen können. Eine entsprechende Software ist leicht zu programmieren. Man kann sich noch weitere Möglichkeiten zur Nutzung derartiger Texte vorstellen, wie die Bearbeitung durch Sprachprogramme, um einen Thesaurus oder Erschließungsvokabular in anderen Sprachen zu erzeugen.

Die Animationsfilme, die wir im Rahmen dieses Projekts analysiert haben, waren besonders interessant, da sie anders strukturiert sind als die in früheren Projekten untersuchten Filme. Es gibt eine Reihe von Gründen dafür. Einige Filme hatten experimentellen Charakter, auch hinsichtlich der Technik, denn wenn das Bild ohne Benutzung einer Kamera direkt auf das Filmmaterial gezeichnet oder eingeritzt wird, gibt es keine Einstellungen. Bei anderen Filmen wurde Pixilation benutzt, ein Stop-Motion-Verfahren, bei dem die Bewegungen der Darsteller durch Verzögerung effektvoll verfremdet werden. Auch hier gibt es keine Einstellungen im engeren Sinn.
Abgesehen von fehlenden Einstellungen haben Animationsfilme oftmals auch keine Dialoge, so dass die akustische Beschreibung problemlos integriert werden kann. Dies hat zur Folge, dass die Beschreibung fast immer (zu 98,29%) zeitgleich mit dem Erscheinen des Bildes zu hören ist, bei Realfilmen ist sie hingegen häufig vor oder nach dem jeweiligen Bild zu hören. Wir haben bei der Analyse der Filme Strukturen und Gestaltungsprinzipien herausgearbeitet, die sich von den Ergebnissen früherer Untersuchungen unterscheiden.
Weiterhin hat sich im Laufe unserer Arbeit herausgestellt, dass die Gruppe der sehbehinderten Nutzer sowohl hinsichtlich des Sehvermögens als auch der kognitiven Fähigkeiten so heterogen ist, dass den Nutzern unbedingt die Möglichkeit geboten werden sollte, die akustische Beschreibung für jeden Film individuell anzupassen. Eine Möglichkeit ist die Stufung der Beschreibung nach Detailliertheit. Bei der Erstellung von *audio description text* müssten die jeweiligen Niveaus je nach Ausführlichkeit der Beschreibung z.B. als Stufe 1, 2 und 3 markiert werden, die der Nutzer seinen Bedürfnissen entsprechend auswählen kann. Ein weiterer Aspekt betrifft die zu beschreibenden Objekte. Unserer Typologie bietet dem Nutzer die Möglichkeit, das aus der akustischen Beschreibung auszuschließen, was er noch selbst erkennen kann, z.B. sonniges oder wolkiges Wetter. Dieses Beispiel veranschaulicht die Probleme, mit denen wir uns in der nächsten Phase des Projekts beschäftigen möchten.

Zusammenfassung

In diesem Papier berichten wir über zwei Fragestellungen aus unserer Arbeit im Rahmen des *E-Inclusion Research Network project*, weitere Aspekte werden an anderer Stelle behandelt. Hier wollten wir zeigen, wie Stichwörter aus Texten, die der Bildbeschreibung für Sehbehinderte dienen, zur Erschließung von einzelnen Einstellungen und Sequenzen von Filmen genutzt werden können. Wir haben festegestellt, dass solche Stichwörter sehr gut für die Verwendung als Erschließungsvokabular geeignet sind.

Wir haben auch untersucht, wie die Stichwörter automatisch in andere Sprachen übersetzt und so für die bilinguale oder sogar multilinguale Erschließung genutzt werden können. Da durch die automatische Übersetzung in drei Viertel der Beispiele eine Entsprechung ermittelt werden konnte – Synonyme oder eng verwandte Begriffe wurden nicht berücksichtigt – ist ein solches Verfahren praxistauglich. Diese Untersuchung leistet einen Beitrag zur Auseinandersetzung mit der Frage, wie vorhandene Texte, die eigentlich anderen Zwecken dienen, für die Erzeugung von geeignetem Vokabular für die Erschließung einzelner Einstellungen und Sequenzen von Filmen oder Fernsehproduktionen genutzt werden können. Ähnlich wie bei einer Reihe von anderen Textressourcen stellen auch die von uns untersuchten Texte eine große Anzahl von Stichwörtern zur Verfügung, die sich, sofern sie automatisiert weiterverarbeitet werden, sehr gut als Erschließungsvokabular nutzen lassen. Somit kann der Zugriff auf einzelner Einstellungen und Sequenzen deutlich verbessert werden, und zwar zu einem Bruchteil der Kosten, die für eine intellektuelle Erschließung nötig wären.

Danksagung

Wir danken *Canadian Heritage* für die Unterstützung der Gründung *E-Inclusion Research Network.* Wir danken unseren Partnern vom *Centre de recherche en informatique de Montréal (CRIM),* für die Hilfe und Beratung und die kollegiale Arbeitsatmosphäre. Wir sind Anne Jarry vom *Canadian National Institute for the Blind (CNIB)* sehr dankbar für die Hilfe bei der Organisation von Filmveranstaltungen mit Sehbehinderten und ihre unschätzbaren Hinweise und Ratschläge. Wir bedanken uns bei allen Teilnehmern, die durch ihre Ansichten und Meinungen dazu beigetragen haben, die akustische Bildbeschreibung weiter zu verbessern. Dank an Alain Lapierre, auch vom CNIB, für die organisatorische und technische Unterstützung. Zum Abschluß möchten wir Linda Rivest, unserer Forschungsassistentin, danken, die die Daten für unsere Untersuchung aufbereitet hat.

Quellen

Independent Television Commission [UK]. 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 and James M Turner. 2007. Audiovision ou comment faire voir l'information par les personnes aveugles et malvoyantes : lignes directrices pour la description d'images en mouvement. Canadian Association for Information Science, annual conference, *P*roceedings* [in press].

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. and Emmanuël Colinet. 2004. Using audio description for indexing moving images. *Knowledge Organization* 31, no. 4, 222-230.

Turner, James M. and 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.