

Advene, une plateforme pour l'édition et la visualisation d'hypervidéos

Olivier Aubert, Pierre-Antoine Champin, Yannick Prié

► **To cite this version:**

Olivier Aubert, Pierre-Antoine Champin, Yannick Prié. Advene, une plateforme pour l'édition et la visualisation d'hypervidéos. Journées Modèles documentaires de l'audiovisuel, Jun 2004, La Rochelle, France. sic_00001270

HAL Id: sic_00001270

https://archivesic.ccsd.cnrs.fr/sic_00001270

Submitted on 13 Dec 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Advene : une plate-forme pour l'édition et la visualisation d'hypervidéos*

Olivier Aubert, Pierre-Antoine Champin, and Yannick Prié

LIRIS - FRE 2672 CNRS - Université Lyon 1
prenom.nom@liris.univ-lyon1.fr

Résumé

Le projet Advene (Annotations de DVD Échangées sur le NEt)¹ vise à faciliter le travail des communautés échangeant des discours (analyses, études) sur des documents audiovisuels au format DVD. Tirant parti des possibilités induites par le caractère numérique du support, des fonctionnalités hypertexte peuvent être ajoutées au contenu audiovisuel, par le biais d'annotations fournissant une structure explicite. Exploitant d'une part le contenu audiovisuel et d'autre part la structure d'annotations, des documents hypervidéos peuvent être produits.

La plate-forme Advene fournit un modèle et des outils pour créer et utiliser des schémas d'annotation ; pour annoter des flux vidéo selon ces schémas ; pour générer et afficher les visualisations vidéo-centrées ou texte-centrées à partir des annotations. Des schémas (définissant des types d'annotations et de relations), des annotations, des relations, des requêtes et des vues peuvent être regroupées et partagées en des unités appelées *recueils*. Les documents hypervidéo sont générés à la demande à partir des recueils (qui contiennent notamment les annotations et la description des vues) et des DVD (flux audiovisuels).

Cet article vise à introduire les différents éléments du modèle Advene, et à en illustrer la mise en œuvre à travers un scénario d'utilisation. Il présente également le prototype Advene, comprenant notamment un lecteur DVD instrumenté (permettant de produire les vues vidéo-centrées) et un serveur web intégré (permettant d'accéder, par le biais d'un navigateur web standard, aux vues texte-centrées).

Mots-clés : Annotation, modèle de document, DVD, partage, vidéo, modes de visualisation, Advene

1 Introduction

La diffusion de l'audiovisuel (AV) numérique est en train de se développer, portée par les connexions internet haut-débit, la télévision numérique et le format DVD. Ceci est valable tant au niveau du grand public, essentiellement pour les loisirs (fictions, documentaires, etc.) que dans une utilisation plus professionnelle : vidéoconférence, surveillance, enseignement, recherche en Sciences de l'Homme et de la Société. Ces derniers domaines font notamment appel à la *lecture active* des documents audiovisuels, pratique dans laquelle il s'agit d'*annoter* les vidéos dans une perspective analytique : enregistrements sportifs, corpus linguistiques ou de situations d'interaction (dans une classe par exemple).

Le contexte de notre travail est celui de la lecture active de documents audiovisuels sur support DVD et du partage des analyses au sein de communautés : analyse filmique pour les cinéphiles ou les chercheurs, utilisation en classe pour enseignants et élèves (étude de vidéos, enseignement en classe de langue ou de littérature). Notre travail a donc pour objectif de prendre en compte des utilisations de vidéos qui soient différentes de la simple visualisation/consommation. Ces utilisations doivent se baser sur des interfaces de visualisation d'*hypervidéos*² dans lesquelles il s'agit d'interagir avec le flux audiovisuel de façon non triviale.

*<http://liris.cnrs.fr/advene/>

1. Projet mené au sein du LIRIS: Laboratoire d'InfoRmatique en Images et Systemes d'information, FRE 2672 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/Ecole Centrale de Lyon

2. Ce terme provient du projet Hypercafé (Sawhney *et al.*, 1996), mais nous l'utiliserons dans un sens étendu : il dénotera aussi bien les vidéos avec des possibilités hypertextes que les documents hypertextes dans lesquels la vidéo joue un rôle non anecdotique.

Les utilisateurs doivent avoir la possibilité d'annoter les flux et de construire leurs propres hypervidéos.

Il est alors nécessaire de disposer de descriptions structurées de documents audiovisuels (mises en place au cours d'un processus d'annotation), et d'outils permettant leur édition, ainsi que leur lecture/visualisation. Pour cela, il faut construire des modèles documentaires permettant tout d'abord de localiser des fragments audiovisuels pertinents pour certaines analyses, ensuite de les qualifier (de leur attribuer une sémantique) et enfin de les mettre en œuvre dans des interfaces graphiques utilisant les structures documentaires en lien avec le flux. Le projet Advène (Annotations de DVds Echangées sur le NEt) propose un tel modèle, ainsi que des outils permettant la lecture active et le travail sur des documents audiovisuels au format DVD³.

Nous avons identifié plusieurs fonctionnalités que notre système doit offrir (et que notre modèle doit supporter) :

- *l'annotation de documents audiovisuels* comme édition d'annotations de type variés, liées au flux AV, formant une structure qui sera utilisée pour construire et interagir avec les hypervidéos. Cette structure obéira à un *schéma d'annotation* ;
- *la construction de vues* des documents audiovisuels et des annotations ;
- *la visualisation de documents hypervidéos* construits à partir du flux et des annotations en mobilisant des vues plus ou moins spécifiques ;
- *la gestion et le partage* des annotations, des schémas d'annotation, et des vues à l'aide de *recueils*.

Dans la suite de cet article, nous présenterons tout d'abord le modèle Advène, puis nous nous concentrerons sur les processus et les interfaces permettant l'annotation, la construction de vues et la visualisation des hypervidéos.

2 Modèle Advène

L'objectif d'Advène est de fournir un cadre d'applications simple et extensible. Nous nous sommes plus particulièrement attachés à la structure *externe* des annotations, c'est à dire la manière dont elles sont contraintes, mises en relation les unes avec les autres, et présentées (ou rendues) dans des hypervidéos. En effet, les annotations ne sont pas simplement des données attachées à des fragments quelconques du flux audiovisuel. Elles sont structurées selon des *schémas d'annotation*, qui dépendent de la tâche sous-tendant la lecture active. Un *recueil* Advène permet de regrouper et de partager un ensemble cohérent d'annotations, les schémas qui les structurent, ainsi que des *vues* permettant le rendu de ces annotations.

En revanche, le modèle Advène pose aussi peu de contraintes que possible sur la structure *interne* des annotations, c'est-à-dire quel type de données est utilisé pour annoter le flux audiovisuel (texte simple ou mis en forme, XML, son...). Il s'appuie sur des types de données définis par ailleurs, et sur un mécanisme de plugiciel pour leur visualisation et leur édition.

2.1 Vue d'ensemble

La notion centrale du modèle Advène est celle de *recueil*. Un recueil est un ensemble pertinent (du point de vue de son auteur) d'éléments Advène, qui peut être échangé. La figure 1 représente, dans le formalisme UML, les différentes classes d'éléments qui constituent un recueil (classe Package), et leurs associations. Les données annotant le flux audiovisuel sont portées par les Annotations et les Relations entre elles. Ces dernières ne sont pas équivalentes, et sont regroupées en *types* (respectivement AnnotationType ou RelationType) décrits dans un Schéma du même recueil. Des requêtes (Query) permettent de sélectionner un sous-ensemble d'éléments

3. Nous pensons en effet que des usages nouveaux de l'audiovisuel ne peuvent réellement se développer que si une utilisation large d'outils se met en place. Pour cela, nous proposons d'une part de considérer des films DVD qui nous permettent de nous affranchir des problèmes de droits liés à la circulation des contenus audiovisuels, tout en utilisant des documents de haute qualité. A cet effet, seules les descriptions d'annotations et d'hipervidéos circuleront sur le réseau au sein des communautés. Nous visons d'autre part un format ouvert et des outils développés suivant les principes du logiciel libre.

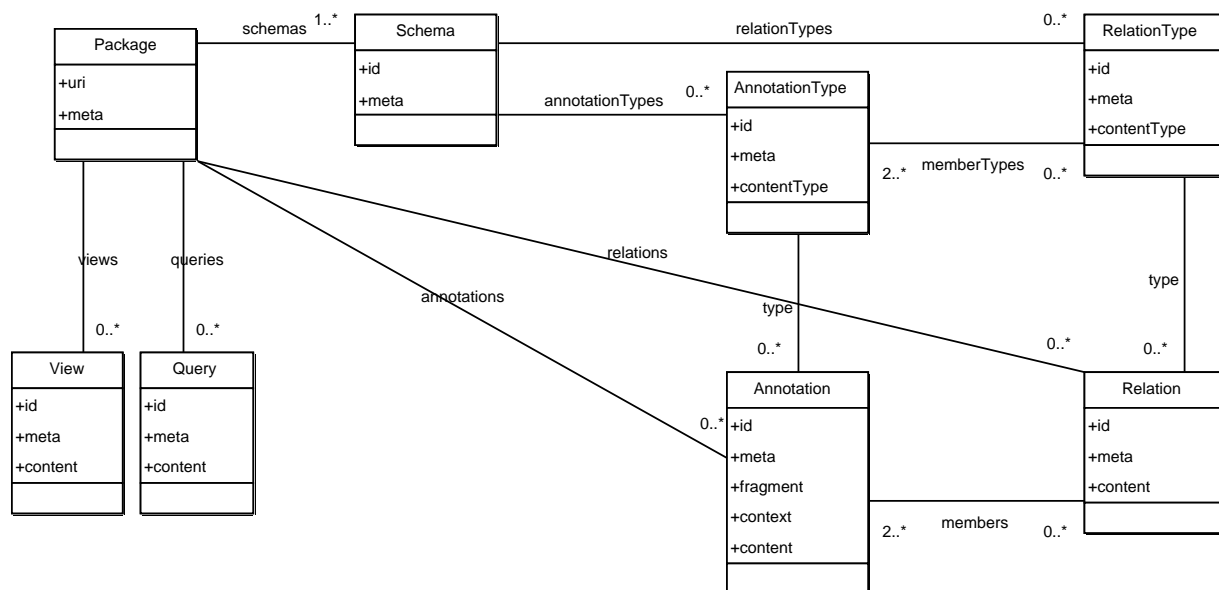


FIG. 1 – Diagramme UML du modèle Advene

du recueil. Les vues (View) permettent quant à elles de produire des hypervidéos en combinant des fragments du flux annoté et les éléments du recueil.

Éléments définis ou importés

Il est important de remarquer que ce diagramme est simplifié, les relations entre certains éléments pouvant revêtir deux formes : la définition et l'import. Chaque élément constituant un recueil peut être *défini* par ce recueil (c'est à dire décrit explicitement), ou *importé* depuis un autre recueil (c'est-à-dire décrit par une simple *référence* à sa définition dans cet autre recueil). De la même façon, un schéma peut définir ou importer les types d'annotation et de relation d'un autre schéma. En fournissant un accès uniforme aux éléments définis et importés, un recueil met à disposition toute l'information nécessaire⁴ à la génération d'hypervidéos à partir des annotations.

2.2 Éléments du modèle

2.2.1 Recueil

Un recueil permet de regrouper dans un unique document tous les éléments Advene (schémas, vues, annotations, etc.) produits ou utilisés par la lecture active d'un ou plusieurs DVD, afin d'en faciliter l'utilisation et l'échange. Il est identifié par un URI. L'auteur et la date de création du recueil sont représentés comme méta-données à l'aide des éléments du Dublin Core.

2.2.2 Annotations, Relations

Les annotations sont des données (appelées le *contenu* de l'annotation) attachées à des portions adressables du flux audiovisuel, que l'on appelle *fragments*. En plus des attributs *id* et *meta*, les annotations ont donc un attribut *fragment* et un attribut *content*.

4. Le flux audiovisuel est lui-même référencé par chaque annotation, quoiqu'il ne soit pas contenu dans le recueil.

Notons que l'identification d'un fragment dans un document audiovisuel numérique dépend de la structure spécifique de ce dernier. Dans une simple vidéo, seule une mesure temporelle (ou éventuellement en nombre d'octets) permet d'adresser un point du flux. Les DVDs, en revanche, sont structurés en titres et en chapitres. La norme MPEG-4 permet même d'adresser des objets (fragments spatio-temporels) dans le flux audiovisuel. Advène utilise pour l'instant des fragments définis comme intervalles entre deux bornes temporelles, mais permet l'extension à d'autres types de fragments.

Par définition, les annotations sont en relation avec le flux audiovisuel, par le biais de leur attribut `fragment`. Il peut également être utile d'exprimer le fait que des annotations sont en relation les unes avec les autres. C'est le rôle de la classe `Relation`: une relation n'est pas reliée directement au flux, mais elle est liée à plusieurs annotations (les *membres* de la relation). Les relations peuvent également contenir des informations additionnelles par le biais de leur attribut `content`.

Un recueil peut définir ses propres annotations et relations, ou les importer depuis un autre recueil. Cependant, toutes les annotations et toutes les relations doivent correspondre à un schéma du recueil (voir ci-après).

2.2.3 Types

Toutes les Annotations et Relations ne sont pas équivalentes entre elles. Les types permettent de contraindre la structure d'annotations ou de relations jouant le même rôle.

Un type d'annotations (classe `AnnotationType`) définit le *type de contenu* que les annotations de ce type peuvent avoir. Dans le prototype Advène, le type de contenu est spécifié par un type MIME (Freed & Borenstein, 1996) (`text/html`, `audio/*`, etc.). Si le type de contenu est `text/xml`, il peut éventuellement être précisé par une DTD, un schéma XML ou Relax NG⁵.

Les types de relations (classe `RelationType`) définissent le nombre de membres, *i.e.* le nombre d'annotations qui doivent participer à une telle relation. Il définissent aussi le type d'annotation auquel chaque membre doit appartenir. Enfin, puisque les relations peuvent également avoir un contenu, les types de relation définissent un type de contenu, de manière similaire aux types d'annotations.

2.2.4 Schémas

Les schémas d'annotation (classe `Schema`) contraignent les annotations et les relations qu'un recueil peut contenir. Plus précisément, ils sont des collections de types d'annotations et de types de relations, regroupés sur la base des tâches de lecture active qu'ils permettent d'accomplir. Par exemple, les types d'annotations et de relations utiles pour l'édition d'un film (comprenant notamment les types d'annotations *Plan*, *Séquence*, et le type de relations *Transition*) seront regroupés dans un schéma *Édition*; un autre schéma *Analyse technique* contiendra les types utiles à l'analyse des techniques cinématographiques (mouvements de caméra, lumière, ...); un troisième contiendra les types utiles à l'exploitation d'un film en cours de langue, etc.

Ce regroupement n'est bien sûr pas exclusif: un type utile à plusieurs tâches, défini dans un premier schéma, pourra ensuite être importé dans d'autres schémas. Par exemple, le type d'annotation *Plan*, appartenant au schéma *Édition*, sera réutilisé dans le schéma *Analyse technique* pour désigner les plans où ont lieu certains mouvements de caméra.

Un schéma et les types qui les composent possèdent un titre et une description. Ces méta-données ont pour rôle d'explicitier la façon d'utiliser les éléments.

2.2.5 Requêtes

Une requête (class `Query`) permet d'accéder à une liste d'éléments Advène (de n'importe quel type décrit dans cette section). Elle peut éventuellement être paramétrable. L'attribut `content` contient une description de la requête dans un langage spécifié par son type MIME. Deux langages sont disponibles dans le prototype actuel: une implémentation immédiate basée sur le langage python, et un langage dédié permettant de filtrer

5. En fait, ce principe peut être étendu à n'importe quel type MIME disposant d'un mécanisme similaire pour contraindre une classe de documents.

des ensembles d'éléments sur la base de conditions. D'autres langages de requête plus simples d'utilisation devront être développés et implémentés à l'avenir.

Alors que les schémas offrent un premier moyen de structurer les annotations, les requêtes permettent quant à elles une structuration plus souple (en fonction du pouvoir d'expression du langage de requête fourni). Cette structuration sert ensuite de base aux hypervidéos générées grâce aux vues (cf. ci après). D'autre part, les requêtes permettent de structurer les annotations *a priori*, c'est à dire au moment de la création des schémas, mais elles peuvent aussi être créées *a posteriori*, pour structurer de manière totalement *ad hoc* un ensemble spécifique d'annotations.

2.2.6 Vues

Les vues (classe View) sont le dernier élément de la chaîne permettant de générer des hypervidéos à partir d'un recueil Advène. Les requêtes permettent de sélectionner un ensemble d'éléments d'un recueil. Les vues permettent quant à elles de *présenter* ces éléments de différentes manières, soit globalement statique, soit globalement dynamique (cf. ci-après). Conformément à la vocation de généralité d'Advène, une vue peut être décrite dans n'importe quel langage de présentation identifié par son type MIME. Dans le prototype actuel, les vues globalement dynamiques sont définies par le biais d'un langage de type événementiel (événement-condition-action), exécuté au sein de l'application qui contrôle notamment le lecteur vidéo. Les vues globalement statiques sont définies par le langage TAL/TALES et permettent de générer dynamiquement des documents HTML accessibles par un navigateur web standard par le biais d'un serveur web intégré à l'application.

Le langage de type événementiel a été développé spécifiquement pour le projet Advène, en utilisant le modèle ECA (Paton, 1999) utilisé couramment dans les bases de données par exemple. Ce langage simple permet de spécifier des actions à entreprendre (affichage d'informations, mise en pause du lecteur vidéo) lors d'occurrences d'événements particuliers (comme le début d'une annotation ou la mise en marche du lecteur vidéo) correspondant à des conditions spécifiées (comme par exemple "le type de l'annotation est *sous-titre*"). Ce mécanisme simple nous permet déjà d'expérimenter des scénarios novateurs d'interaction avec la vidéo.

Le langage TAL/TALES a été développé pour la plateforme applicative Zope⁶. Le souci de modularité et de simplicité d'utilisation ayant présidé à son développement nous a fait adopter ce langage comme langage de description de vues pour Advène. En effet, les expressions TALES, qui permettent de spécifier les données à insérer dans le modèle de document spécifié en TAL, fournissent un accès simple à la structure de données fournie par le modèle Advène. Une différence notable avec XSLT/XPath, autre mécanisme de génération de documents à partir de modèles, est que la source des données est un modèle de données disponibles en mémoire et pas uniquement des données XML, ce qui correspond plus à notre contexte d'utilisation. Par ailleurs, TAL a été conçu pour et permet une intégration transparente avec des outils d'édition WYSIWYG. On peut donc disposer d'éditeurs instrumentés permettant à un public novice de créer des vues simplement, par réutilisation de vues existantes. Enfin, TAL sert principalement à générer des documents HTML (vues globalement statiques), mais peut générer n'importe quel type de document XML. Il permet donc de générer, par exemple, un document SVG ou SMIL (W3C, 2001), et peut donc être utilisé pour décrire des vues globalement dynamiques.

3 Scénarios d'utilisation

Nous proposons ici des exemples d'utilisation d'Advène destinés à mettre en évidence les mécanismes d'exploitation des hypervidéos. Ces exemples se basent sur deux schémas d'annotation :

- le schéma « Découpage en plans » définit le type d'annotation Plan, dont le contenu est une brève description textuelle du plan, et le type de relation « Ordre diégétique » , permettant de lier des plans et d'indiquer leur ordre diégétique ;

6. <http://www.zope.org/Wikis/DevSite/Projects/ZPT/TAL>

- le schéma « Exploitation pédagogique » définit le type d’annotation Référence, dont le contenu est l’URL d’une ressource concernant la séquence annotée, et le type d’annotation Épisode découpant le film en unités pertinentes sur le plan pédagogique. D’autres types, non détaillés ici, permettent la mise en place d’exercices à propos du film.

Dans de nombreux cas, les annotations ont un contenu textuel plus ou moins structuré (texte simple, HTML, XML), à partir duquel il est possible de générer des documents *globalement statiques*. Ces documents peuvent également exploiter des images extraites à la volée du DVD pour illustrer ces documents. Par exemple, les annotations de type Plan peuvent être utilisées pour générer une liste chronologique (temps de la narration) des plans, avec leur première image et leur description. Une telle vue est une *spatialisation* du temps du film. Si on exploite des relations du même schéma, on peut aussi envisager d’ordonner cette liste suivant l’ordre diégétique. Un autre exemple de vue globalement statique, utilisant le schéma pédagogique, est la génération d’une feuille d’exercice et du corrigé à partir des annotations dédiées posées sur le film.

Dans le cas le plus simple de vue *globalement dynamique*, l’hypervidéo produite a une temporalité découlant directement de celle du flux annoté. Par exemple, le film annoté par ses plans peut être visualisé en affichant, en sous-titre, la description de chaque plan au début de celui-ci. Avec le schéma pédagogique, cette temporalité peut être plus ou moins altérée : les annotations de type Épisode peuvent entraîner un arrêt sur image à la fin de chaque épisode, réclamant un redémarrage manuel. Les annotations de type Référence peuvent quant à elles être présentées sous forme d’hyperliens, permettant de passer d’un instant du film à un autre, voire à un tout autre document.

Enfin, comme pour les vues globalement statiques, la structure des hypervidéos produites par les vues globalement dynamiques peut être basée sur la structure induite par les relations entre annotations, « déstructurant » ainsi la temporalité initiale du flux annoté. Dans notre exemple de l’annotation en plans, les relations d’ordre diégétiques peuvent être utilisées pour générer à la volée un re-montage virtuel du film, oubliant l’ordre de la narration au profit de l’ordre des événements.

La figure 2 présente un exemple d’utilisation pédagogique d’Advene, faisant figurer trois modes de visualisation : d’une part, une visualisation globalement statique à l’intérieur d’un navigateur web, présentant une feuille d’exercice utilisant le contenu de certaines annotations (de type Question), enrichi de vignettes tirées du DVD. D’autre part, une autre visualisation globalement statique, utilisée principalement lors de l’édition d’annotations, présentant l’ensemble des annotations sur une ligne de temps. Enfin, la dernière visualisation, globalement dynamique, affiche certaines annotations (selon leur type) comme sous-titres sur le rendu de la vidéo.

4 Travaux proches

Nous présentons rapidement quelques travaux proches, en considérant d’une part ceux qui sont plus axés sur les modèles de description, d’autre part ceux qui s’intéressent plus aux interfaces de visualisation. Il est cependant entendu qu’une description vise souvent à fournir l’architecture d’une interface, tandis qu’une interface de visualisation s’appuie obligatoirement sur un modèle de description des documents audiovisuels.

4.1 Modèles de description

MPEG7 (Sanchez *et al.*, 2002) vise à enrichir les documents AV en ajoutant des annotations de types variés. Cependant, ce format n’intègre pas de notion de visualisation : il se contente, par conception, de fournir un support de stockage des informations, charge aux concepteurs d’applications de les exploiter. Nous postulons avec Advene que la visualisation est liée plus ou moins fortement aux annotations, et que tout du moins, il est important de prendre cet aspect en compte lors d’un processus d’annotation. D’autre part, le format informatique utilisé par MPEG7 étant XML, nous pouvons assurer une interopérabilité avec Advene : une vue Advene permet de générer un fichier MPEG7 contenant les annotations désirées, au moins pour des contenus simples.

Le projet OPALES (Nanard & Nanard, 2001) se place dans le cadre d’un travail collaboratif sur un corpus de vidéo. Il suppose une infrastructure matérielle dédiée, et son modèle largement réflexif est séduisant.

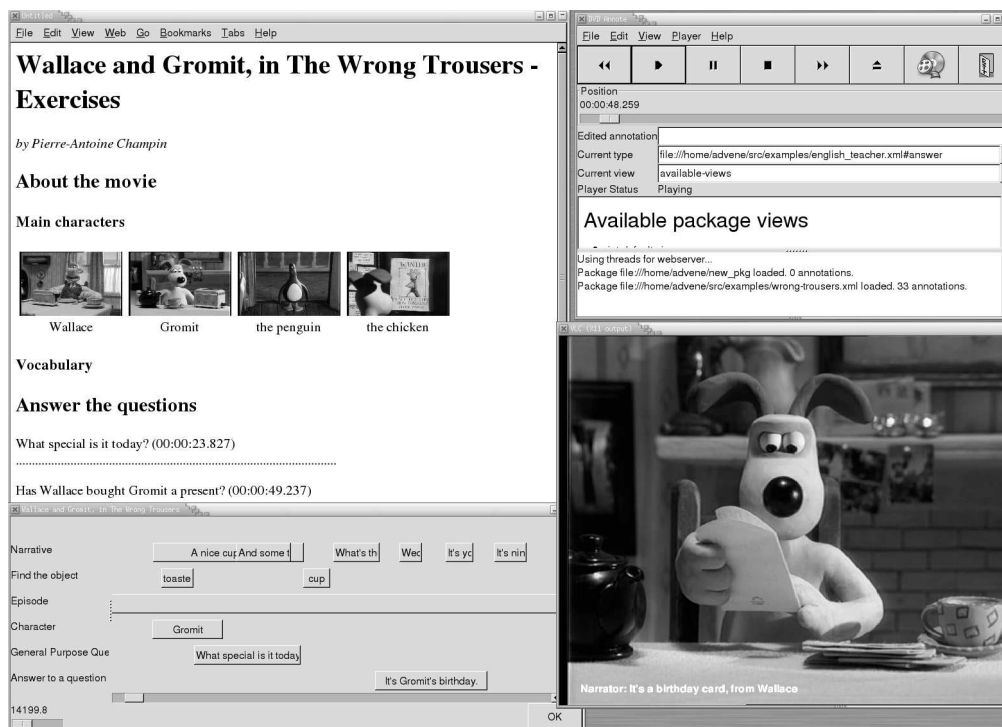


FIG. 2 – Utilisation pédagogique d'Advene

Cependant, il semble difficilement diffusable vers le public visé par Advene: des utilisateurs non-avertis utilisant un matériel standard.

Le projet ATLAS (Laprun *et al.*, 2002) tente de fournir une abstraction couvrant un large spectre d'annotations linguistiques, afin d'assurer une interopérabilité de différents corpus. Il reste contraint au domaine linguistique par le formalisme sous-jacent (le modèle de graphe d'annotations), mais fournit une base intéressante, notamment dans sa définition très flexible des régions (l'équivalent des Fragments d'Advene), qui peuvent adresser différents types de données (flux audio ou vidéo, zone géométriques, etc).

4.2 Interfaces de visualisation

Plusieurs outils existent qui fournissent un sous-ensemble des fonctionnalités proposées par Advene. Le projet le plus proche est Anvil (Kipp, 2001), qui fournit un modèle générique d'annotation, dont on peut enrichir les types, orienté sur l'analyse des dialogues multimodaux. Néanmoins, il reste très orienté comme support d'analyse et n'offre pas de possibilités évoluées de visualisation comme les modes globalement statiques ou globalement dynamiques que nous proposons. D'autre part, il repose techniquement sur la plate-forme Java Media Framework qui ne permet pas à notre connaissance de contrôler un lecteur de DVD, qui est le domaine d'application que nous visons. D'autres projets tels que MagPie (MAGpie, 2003) ou Transana (Thorn, 2002) permettent d'enrichir un document AV par le biais d'annotations, textuelles le plus souvent mais également audio. Cependant, l'unique exploitation envisagée est un rendu de ces annotations au cours de la visualisation du support AV annoté, alors que l'on peut entrevoir d'autres utilisations, comme nous l'avons vu pour Advene, et la contrainte technique (utilisation de JMF) demeure.

En matière de rendu enrichi de vidéo, le projet Hyperfilm (Hyperfilm, 2003) explore des pistes plus originales, dont on peut retrouver l'équivalent dans certaines idées d'Advene. L'outil permet notamment d'inclure des liens actifs dans les vidéos, afin d'offrir à l'utilisateur des pistes alternatives de visualisation d'un média. Cependant, comme les projets précédemment évoqués, il se limite à un mode de rendu des annotations (rendu

lié au déroulement du support AV).

5 Conclusion

Nous avons décrit dans cet article les différents éléments du modèle Advene pour représenter des annotations audiovisuelles, et présenter celles-ci à l'utilisateur sous la forme d'hypervidéos, documents hypermédias générés à partir du document audiovisuel original annoté et des annotations. Les hypervidéos sont des documents multimédias complexes, du fait de la part importante prise par l'élément vidéo et des possibilités d'interaction enrichies qu'elles proposent. Nous avons illustré ce concept par quelques scénarios d'utilisation de l'outil Advene, et des hypervidéos associées.

Les développements du modèle et du prototype Advene se poursuivent. Nous disposons à présent d'une base solide, tant au niveau du modèle que de l'architecture, pour explorer les possibilités des hypervidéos et des expérimentations sont en cours dans le domaine pédagogique pour vérifier la pertinence de nos idées.

Références

- FREED N. & BORENSTEIN N. (1996). *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*. RFC 2046, Internet Engineering Task Force.
- HYPERFILM (2003). Video hyper-media production.
- KIPP M. (2001). Anvil - A Generic Annotation Tool for Multimodal Dialogue. In *Proceedings of Eurospeech 2001*, p. 1367–1370, Aalborg.
- LAPRUN C., FISCUS J. G., GAROFOLO J. & PAJOT S. (2002). A practical introduction to ATLAS. In *Language Resources and Evaluation Conference (LREC)*.
- MAGpie (2003). Media Access Generator (MAGpie). <http://ncam.wgbh.org/webaccess/magpie/>.
- NANARD M. & NANARD J. (2001). Cumulating and sharing end users knowledge to improve video indexing in a video digital library. In *Proceedings of the first ACM/IEEE-CS joint conference on Digital libraries*, p. 282–289, Virginia, USA.
- N. W. PATON, Ed. (1999). *Active Rules in Database Systems*. Springer.
- SANCHEZ J. M. M., KOENEN R. & PEREIRA F. (2002). MPEG-7: The Generic Multimedia Content Description Standard, Part 1. *IEEE Multimedia*, **9**(2), 78–87.
- SAWHNEY N. N., BALCOM D. & SMITH I. E. (1996). HyperCafe: Narrative and Aesthetic Properties of Hypervideo. In *UK Conference on Hypertext*, p. 1–10.
- THORN C. (2002). Creating New Histories of Learning for Math and Science Instruction: Using NVivo and Transana to manage and study large multimedia datasets. In *Conference on Strategies in Qualitative Research*: Institute of Education, University of London.
- W3C (2001). *Synchronized Multimedia Integration Language (SMIL 2.0)*. W3C. <http://www.w3.org/TR/smil20/>.