

Chapitre 13

Vers des environnements d'apprentissage adaptatifs intégrant hypermédias et agents pédagogiques

13.1. Introduction

Dans l'enseignement traditionnel, les interactions entre apprenants et enseignants (ou entre apprenants travaillant en groupe) sont multi-supports (dessins au tableau, photocopiés, documents audio-visuel), multimodales (parole, gestes, postures, regard), coopératives (échanges ayant a priori un but commun comme la compréhension d'une partie de cours) et adaptatives (par exemple, l'enseignant modifie le déroulement de son cours pour répondre aux questions). Une des difficultés pour l'enseignant consiste aussi à savoir adapter sa communication à des apprenants ayant des niveaux très différents. Vis à vis de cette multimodalité et du caractère coopératif et adaptatif de ces interactions humaines, les outils actuels des Environnements Interactifs pour l'Apprentissage Humain (EIAH) semblent limités.

Quels sont les impacts de la multimodalité, de l'adaptabilité et de la coopérativité sur la pédagogie (qualité de l'enseignement, attention des apprenants, interactivité du cours, adaptation à différents niveaux et filières d'apprenants) ? Le projet MICAME (<http://www.iut.univ-paris8.fr/micame/>) – Modélisation Informatique de la Coopération, de l'Adaptabilité et de la Multimodalité dans

l'Enseignement – a pour objectif de mieux modéliser les mécanismes qui sous-tendent ces interactions multimodales et coopératives afin d'envisager la spécification d'Interfaces Homme-Machine multimodales (clarification, empathie et encouragement), plus conviviales, plus pédagogiques, plus efficaces mais aussi permettant d'adapter non seulement le contenu d'un cours à différents niveaux d'apprenants, mais surtout la communication de ce cours.

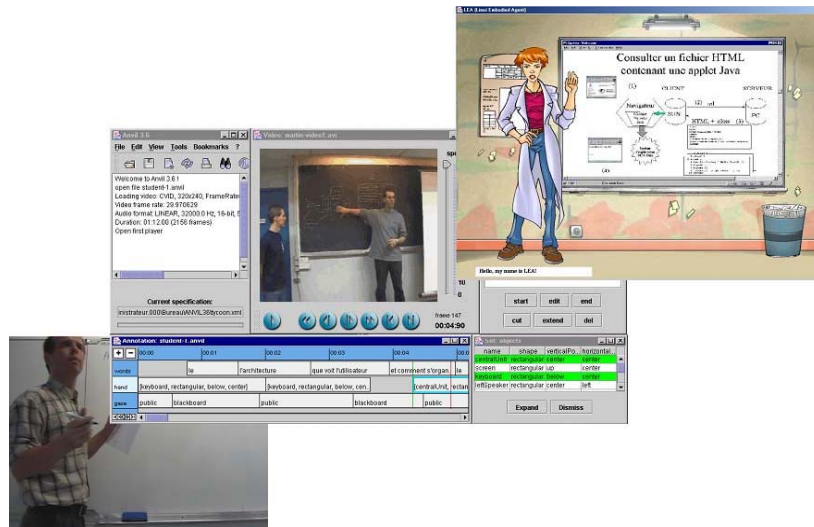


Figure 13.1. Le projet MICAME

Le projet que nous avons défini se décompose en plusieurs étapes illustrées par la figure 13.1 : 1) construction d'un corpus vidéo pédagogique [MAR 02], 2) élaboration d'un modèle de communication multimodale pédagogique, 3) spécification et implémentation d'un système intégrant un hypermédia adaptatif et un agent conversationnel [ABR 02], 4) évaluation du système par des apprenants. Dans cet article, nous décrivons le point 3) qui nécessite de pouvoir faire coopérer un agent pédagogique (personnage animé placé à gauche dans la figure 13.2, fournissant des indications verbales et non-verbales) et un hypermédia adaptatif. Chacun de ces deux composants doit intégrer une adaptabilité par rapport au niveau scolaire de l'apprenant et à l'historique de ses consultations du cours. Le problème consiste alors à coordonner ces deux composants adaptatifs de manière transparente et intuitive pour l'apprenant. Nous nous focalisons sur les problèmes posés par la coopération à mettre en œuvre entre ces composants. Les sections 2 et 3 présentent les domaines des hypermédia adaptatifs éducatifs et des agents pédagogiques. La section 4 décrit comment notre système intègre ces deux composants. La section 5 est consacrée à la question de l'évaluation.

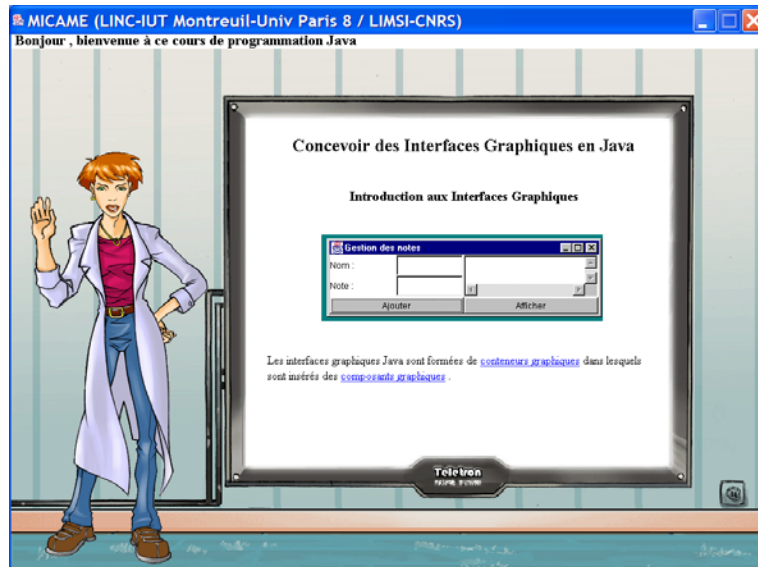


Figure 13.2. Exemple de coopération entre un agent pédagogique et un hypermédia.

13.2. Hypermédias adaptatifs éducatifs

L'enseignement constitue un domaine privilégié d'application des hypermédias adaptatifs [BRU 01]. De nombreuses recherches sont menées dans cette direction depuis le milieu des années 90. Citons le cours hypermédia adaptatif "2L690: Hypermedia structures and systems" (<http://www.wis.win.tue.nl/2L690/>) basé sur le système AHA [DEB 02], le système TANGOW [CAR 99] ou bien le système METADYNE [DEL 00].

L'objectif dans ce contexte est de présenter des documents pédagogiques multimédias d'une manière individualisée, adaptée à l'apprenant. L'architecture des systèmes hypermédias adaptatifs est d'une manière générale basée sur trois composants : a) un modèle de l'apprenant qui représente les informations connues par le système à propos de l'apprenant, b) un modèle du domaine qui structure les éléments pédagogiques auxquels l'utilisateur peut accéder, c) un modèle de l'hypermédia qui détermine en fonction du modèle du domaine et du modèle de l'apprenant quels éléments pédagogiques et quelles possibilités de navigation sont proposés à l'utilisateur à un moment donné et sous quelle forme. Ces trois composants sont étroitement liés. Le modèle de l'hypermédia dépend directement du modèle de l'apprenant et du modèle du domaine. Le modèle du domaine est structuré en fonction du type d'informations pris en compte par le modèle de

l'apprenant. Certains auteurs ajoutent un modèle des fonctionnalités qui offre une représentation conceptuelle des différentes tâches que peut effectuer l'utilisateur [VIL 02].

13.2.1. *Le modèle de l'apprenant*

L'adaptation peut s'effectuer en fonction d'une évaluation des connaissances de l'apprenant sur le domaine d'apprentissage, en fonction d'une évaluation de connaissances indépendantes du domaine d'apprentissage (par exemple : la familiarité avec l'outil informatique) et/ou en fonction de modèles cognitifs (capacités cognitives, personnalité de l'apprenant,...). Des évaluations pourront être effectuées par le système de manière dynamique pendant le processus d'apprentissage, par exemple par validation de tests. Certains éléments du modèle de l'apprenant pourront aussi être informés préalablement au processus d'apprentissage, soit par l'apprenant lui-même, soit par une équipe pédagogique. Cette dernière possibilité peut être particulièrement pertinente lorsque le système est utilisé en complément d'une formation classique.

D'autres paramètres, moins spécifiques au cadre de l'EIAH, pourront bien sûr être utilisés tels des préférences de langue, la prise en compte de caractéristiques techniques du terminal informatique utilisé, ou bien des caractéristiques physiques de l'apprenant (handicap visuel ou auditif, par exemple).

13.2.2. *Le modèle du domaine*

Il est le plus souvent structuré selon une hiérarchie de concepts abstraits qui représentent des entités pédagogiques que l'apprenant doit valider, soit simplement en parcourant des documents pédagogiques, soit en réussissant des tests. L'ensemble de ces concepts est muni d'un ensemble de relations (relation de prérequis, relation entre un concept et un ensemble de sous-concepts, etc.), formant ainsi un graphe conceptuel qui définit l'organisation du domaine d'apprentissage.

Dans la plupart des systèmes actuels, les documents pédagogiques présentés par l'hypermédia sont sélectionnés à partir d'un ensemble de documents multimédia écrits dans un format qui n'est pas propre au système (il peut s'agir de formats HTML, Microsoft Word ou Excel, etc.). Le système associe à un concept un document, un ensemble de documents, ou bien un document complexe composé à partir de plusieurs documents source. Ces documents multimédias peuvent être associés aux nœuds du modèle du domaine de différentes manières [BRU 96] : à un concept peut correspondre un index de documents ou de fragments de documents, ou bien un document unique. Le système peut aussi faire appel à une base de

données de documents multimédias et associer de manière dynamique des documents pédagogiques aux concepts en fonction du modèle de l'utilisateur. On parle alors d'hypermédias adaptatifs dynamiques [VAS 95], [DEL 98].

D'autres travaux s'attachent à définir la structure interne des documents pédagogiques eux-mêmes. Ces travaux utilisent généralement le langage XML. Ainsi, la DTD de [BON 00] comprend des éléments comme `<EXERCISE>...</EXERCISE>`, `<EXAMPLE>...</EXAMPLE>` ou bien `<EXPLANATION> ...</EXPLANATION>`. Le modèle du domaine peut être alors représenté à l'intérieur des documents : [BON 00] définit un élément `<CONCEPT>...</CONCEPT>` qui permet d'associer un nom de concept à un document, et permet de spécifier du contenu ou des liens conditionnels avec des éléments du type `<CONDITION user_level= "2 "> ...</CONDITION>`.

13.2.3. Le modèle de l'hypermédia

L'adaptation des documents pédagogiques peut affecter le contenu fourni à l'apprenant (contenu adaptatif), la forme et la mise en page des documents multimédias (présentation adaptative), ainsi que les possibilités de navigation offertes à l'apprenant (support de navigation adaptatif). De nombreuses techniques ont été utilisées comme par exemple dans le cas de l'adaptation du support de navigation des tris de liens avec présentation par ordre décroissant de pertinence ou bien des annotations de liens par 'clés' visuelles (colorisation de l'ancre du lien avec définition d'une sémantique des couleurs). Nous ne détaillerons pas ici l'ensemble de ces techniques et renvoyons le lecteur intéressé à [BRU 01].

13.3. Agents Pédagogiques

L'agent pédagogique est un paradigme relativement récent pour les Environnements Interactifs d'Apprentissage Humain. Il se situe à la croisée de deux directions de recherche : les agents conversationnels et les environnements d'apprentissage à base de connaissance.

Les agents conversationnels sont des interfaces homme-machine dotées de capacités conversationnelles inspirées de la communication humaine. On s'intéresse ici principalement aux agents conversationnels animés (*Embodied Conversational Agents*) qui ont une apparence humaine (tête ou corps entier) qui vise à leur permettre de générer des sorties verbales et non-verbales, de gérer des fonctions conversationnelles comme les tours de parole (regard, expression faciale), le retour d'information sans interrompre l'utilisateur, la capacité à donner des signaux qui indiquent l'état de la conversation comme les hésitations ou les

corrections [CAS 00]. La réalisation multimodale du comportement de l'agent peut être atténuée ou au contraire exagérée suivant le contexte.

Plusieurs équipes développent des agents pédagogiques (IntelliMedia, Carte, Affective Computing Research Group du MIT Media Lab, PALS Research Lab). Un état de l'art des différents projets d'agents pédagogiques est décrit dans [JOH 00].

Nous avons étudié plusieurs agents pédagogiques comme Steve [RIC 99], Cosmo [LES 00], Jacob [EVR 00], Herman the bug [LES 99], Adele [SHA 99], Autotutor [GRA 03], Stella2 [FEN 02]. Cette étude fait ressortir les dimensions caractéristiques suivantes :

- les modalités disponibles en entrée (selon le système, l'utilisateur peut juste naviguer avec la souris, cliquer sur des boutons, taper des phrases dans une boîte de dialogue ou parler) ;
- les modalités disponibles en sortie (selon le système l'agent peut parler, afficher une phase, comporter une tête ou un corps entier, être en 2D ou 3D, être réaliste ou plutôt de type cartoon) ;
- les différents rôles possibles de l'agent pédagogique liés aux aspects psychologiques de l'apprentissage [GRA 02], [DIL 93], [BAY 03] : expert, tuteur, mentor ;
- les applications pertinentes des agents pédagogiques (environnement virtuel, diagnostic, ...).

Les actes communicatifs pédagogiques et comportements multimodaux suivants ressortent de cette synthèse :

- agir physiquement dans l'environnement éducatif : montrer comment exécuter une action physique / démonstration interactive (e.g. réparation), exécuter les actions spécifiées par l'étudiant ;
- aider à la navigation (dans un environnement 3D, pour conseiller un autre cours) ;
- guider les actions de l'étudiant (interface et tâche) ;
- guider l'attention (regard, geste, posture, locomotion) et éviter les ambiguïtés ;
- fournir différents degrés de retour sur les actions de l'étudiant (e.g. féliciter, approuver, désapprouver, s'étonner...) ;
- signaux de conversation (tours de parole) ;
- enseigner via un dialogue : produire un acte causal amenant l'étudiant à se poser une question (gesticuler, se gratter la tête), fournir / demander des explications rationnelles, indications et conseils sur la manière de résoudre les problèmes, présenter des informations, initier un dialogue avec l'étudiant, répondre à des questions, permettre à l'étudiant d'apprendre via un apprentissage basé sur la tâche

(*task-based learning*) : l'étudiant doit poser les bonnes questions et obtenir certaines réponses de l'agent.

Les agents pédagogiques actuels ont leurs limites. Le comportement de l'agent et l'application pédagogique sont intégrés au niveau logiciel et ne bénéficient pas des avancées plus récentes en terme de spécifications XML à différents niveaux (coopérations entre modalités, émotions) plus facilement réutilisables, par exemple pour générer différents cours. De plus, l'agent adapte rarement son comportement multimodal pour répondre à des besoins d'étudiants de différents niveaux, ou pour réagir dynamiquement aux actions de l'étudiant. Enfin, le comportement multimodal de ces agents pédagogiques est fondé sur des règles générales issues de la littérature en socio-linguistique et rarement sur une analyse appliquée des stratégies communicatives multimodales d'enseignants telles que l'on peut les observer et les annoter sur des vidéos.

13.4. Le système MICAME

Notre objectif est de tirer parti des avancées dans les domaines des hypermédias adaptatifs éducatifs et des agents pédagogiques afin de proposer un système d'EIAO qui intègre :

- la présentation de documents pédagogiques sous forme d'hypermédia ;
- un agent pédagogique dont les actions de communication sont finement synchronisées avec la présentation hypermédia ;
- une adaptation des contenus pédagogiques et de leur présentation, de la navigation parmi les documents pédagogiques, et du comportement communicatif de l'agent en fonction des caractéristiques et des actions de l'apprenant.

L'architecture du système MICAME est représentée par la figure 13.3. Le système est composé d'un modèle du domaine formé de documents LMML (voir ci-dessous), d'un modèle de l'apprenant et d'un contrôleur dont le rôle est de répartir les tâches communicationnelles entre la présentation hypermédia et l'agent pédagogique.

Nous considérons dans un premier temps un modèle de l'apprenant simplifié qui prend en compte une évaluation du niveau général de l'apprenant concernant le domaine (cette évaluation est introduite initialement dans le système ; elle peut être par exemple le résultat d'évaluations préalables) ; un ensemble de concepts à valeurs booléennes (non validé, validé), ainsi que l'historique des documents pédagogiques consultés par l'apprenant.

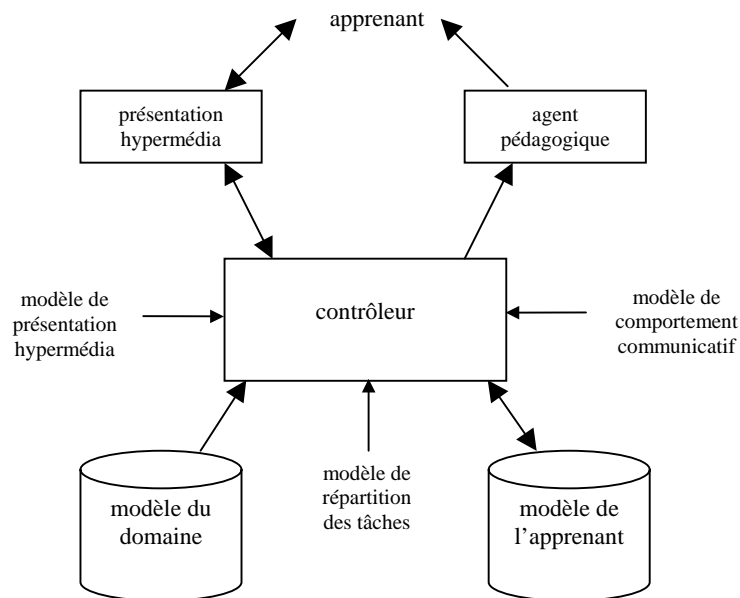


Figure 13.3. Architecture du système MICAME

13.4.1. Représentation des documents pédagogiques

La plupart des recherches menées depuis une dizaine d'années dans le domaine de l'enseignement assisté par ordinateur ont abordé la question de la production et de la gestion d'environnements d'apprentissage à grande échelle, avec le plus souvent le Web comme support. Ces travaux ont pour principal objet l'organisation et la réutilisation de briques multimédia – qui peuvent être préexistantes – caractérisées à l'aide de méta-données, et trouvent actuellement leur continuation dans le processus de normalisation des technologies de l'information pour l'éducation (<http://comelec.afnor.fr/afnor/cn36>). Notre approche au contraire s'intéresse en priorité à l'organisation interne des briques composant un cours, à un niveau de granularité très fin. Notre objectif étant de présenter de manière simultanée et coopérative des documents pédagogiques par un agent conversationnel d'une part, et sous la forme d'un hypermédia d'autre part, le comportement de l'agent ne doit pas seulement se baser sur un modèle de l'apprenant mais aussi sur des informations de nature structurelle et sémantique concernant les documents pédagogiques. L'agent pédagogique pourra être amené à intervenir auprès de l'apprenant, par exemple pour le motiver ou peut-être le blâmer. Nous ne souhaitons cependant pas limiter son rôle à un rôle de supplément au matériel pédagogique. Nous souhaitons que l'agent puisse intervenir dans la présentation elle-même,

certaines explications pouvant être par exemple délivrées oralement par l'agent plutôt que présentées de manière textuelle. Ainsi, le système devra pouvoir décider de la présentation d'un élément pédagogique :

- soit sous forme hypermédia uniquement ;
- soit de manière coopérative : partiellement par l'agent et partiellement sous forme hypermédia ;
- soit avec introduction de redondance : par l'agent et sous forme hypermédia ;
- soit par l'agent uniquement.

Nous retenons l'idée de stricte séparation entre contenu et forme mise en avant par les travaux en ingénierie documentaire et qui a présidé à la naissance de SGML puis de XML. Les documents pédagogiques mis en jeu par le système MICAME n'intègrent pas de détails concernant la forme de présentation par l'un ou l'autre de ces deux médias de sortie que sont l'agent et la présentation hypermédia. Chaque document doit par contre faire apparaître sa structure pédagogique et sa structure énonciative. Le système utilisera ces informations structurelles pour décider, en fonction du modèle de l'utilisateur, de la répartition des tâches communicatives entre l'agent et la présentation hypermédia.

Le modèle de documents pédagogiques LMML (Learning Material Markup Language : <http://www.lmml.de>) propose un tel cadre. Il s'agit d'un langage XML qui définit un ensemble d'éléments « sémantiques » pour la représentation de matériel pédagogique ainsi qu'un ensemble d'attributs permettant d'associer des méta-données. Ce langage peut être aisément étendu, par exemple avec la définition de relations entre documents. La présentation d'un document LMML peut s'effectuer sur un navigateur HTML par l'intermédiaire d'une mise en forme assurée par une feuille de style XSL.

Un document LMML est composé d'éléments structurels (section, paragraphe, liste, table, etc.), d'éléments de contenu pédagogique (exemple, exercice, illustration, motivation, définition, etc.), d'éléments multimédia (texte, son, image, animation) et de liens hypertextuels. Les éléments structurels et les éléments de contenu pédagogique peuvent prendre des attributs permettant de leur associer des méta-informations comme un niveau de difficulté ou une évaluation qualitative du caractère théorique ou formel. Les documents pédagogiques présentés par le système MICAME seront décrits en LMML étendu, dans un premier temps, avec des notions de concept et de relation de prérequis permettant la définition d'un modèle du domaine simplifié. Ce langage sera étendu à nouveau dans le futur de manière à définir des modèles du domaine plus complexes (voir notre discussion à ce sujet à la section 2). La figure 13.4. présente un exemple de document pédagogique utilisé par le système MICAME (il s'agit ici d'un élément de cours de programmation en langage Java).

```
<section difficulty="low">
  <definition>
    <text>
      Un bouton est un composant graphique sur lequel
      l'utilisateur peut cliquer.
    </text>
  </definition>
  <example>
    <image uri="images/frameAvecBouton.jpg" />
    <code>
      JFrame f=new JFrame("Exemple de bouton"); // création de la fenêtre
      JButton b=new JButton("Go !"); // création du bouton
      f.getContentPane().add(b); // ajout du bouton dans la fenêtre
      f.pack();f.setLocation(100,100);f.setVisible(true); // affichage de la fenêtre
    </code>
  </example>
  <remark>
    Il faudra ajouter un écouteur afin d'exécuter une action lorsqu'on
    clique sur le bouton.
  </remark>
</section>
```

Figure 13.4. Exemple de document pédagogique

13.4.2. L'agent conversationnel LEA

Nous utilisons dans le cadre du système MICAME l'agent conversationnel LEA. Il s'agit d'un agent animé 2D utilisant un catalogue d'images des différentes parties du corps (pupille, sourcils, tête, bras, corps). Ainsi pour les bras 46 images ont été dessinées afin de représenter différentes configurations (un ou deux bras, positions et configurations de la main) qui peuvent être utilisées pour différents types de gestes (icôniques, emblèmes et déictiques par exemple). Les combinaisons de ces différentes images ainsi que du message à synthétiser vocalement peuvent être spécifiées à un bas niveau en XML. Trois apparences d'agents ont été conçues avec la même technologie et ont fait l'objet d'études expérimentales pour différentes stratégies de comportement multimodal de la part de l'agent : redondant, complémentaire, monomodal [BUI 03]. Un langage XML permet de spécifier à un niveau plus abstrait le degré de redondance ou au contraire de complémentarité que doit montrer l'agent dans son comportement lorsqu'il fait référence à des objets ainsi qu'une certaine variabilité aléatoire afin d'éviter un comportement répétitif de l'agent [ABR 03].

13.4.3. Adaptabilité et répartition des tâches

L'adaptabilité du système peut intervenir selon trois dimensions, en fonction du modèle de l'apprenant (tableau 13.1.) : les modalités utilisées, le détail des informations communiquées et le type de parcours. Ces dimensions regroupent celles étudiées classiquement dans les interfaces multimodales et dans les hypermédias.

Modalités	Hypermédia seul
	Agent avec Hypermédia
Niveau de détail des informations communiquées	Hypermédia : simple / détaillé
	Agent : simple / détaillé (nombre et longueur des interventions, activité non-verbale)
Parcours	Très guidé par le système
	Liberté de parcours importante laissée à l'apprenant

Tableau 13.1. Dimensions de l'adaptabilité

L'adaptabilité est gérée à plusieurs niveaux : dans la sélection et dans l'ordonnancement des étapes communicatives, mais aussi dans la réalisation interne de chacune de ces étapes. A la première exécution du système, l'algorithme déroule les étapes communicatives suivantes : message de bienvenue, introduction au cours, choix de la 1ère section (par le système ou par l'apprenant). Ensuite, pour chaque section sélectionnée, les étapes sont les suivantes : introduction à cette section, présentation du contenu de la section, présentation des liens possibles vers les autres sections.

Dans la figure 13.2., l'apprenant avait un niveau faible et consultait le cours pour la première fois. Le système génère alors un parcours fortement guidé afin que l'apprenant ne se perde pas, des pages peu denses dans la partie hypermédia en se limitant aux éléments pour lesquels l'apprenant a validé les pré-requis et une contribution importante verbale et non-verbale de l'agent pédagogique. Dans le deuxième exemple (figure 13.5.), l'apprenant a un bon niveau et a déjà consulté des parties du cours. Le système génère alors une partie hypermédia plus dense complétée par une contribution essentiellement non-verbale de la part de l'agent pédagogique.

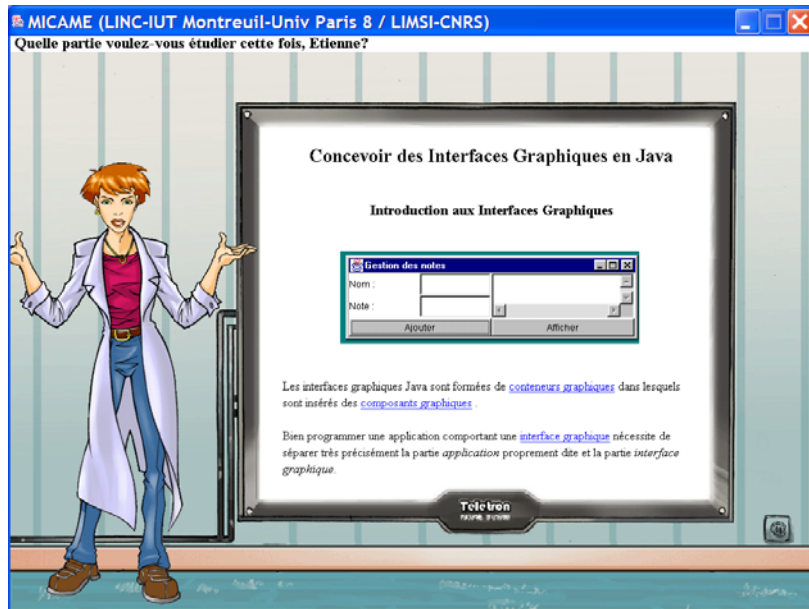


Figure 13.5. Exemple d'écran d'accueil

13.5. Evaluation

Déterminer les caractéristiques de l'agent qui interagira avec des utilisateurs nécessite de se poser de nombreuses questions. Tout d'abord, est-ce qu'un tel agent est la 'meilleure' interface, où meilleur doit être défini dans le contexte de l'interface en question : la plus efficace, la plus naturelle, le moyen le plus facile d'interagir avec le système, etc. Ensuite, si l'interface profite vraiment de l'apport d'un agent, quelles doivent en être les spécificités ? Doit-il être en 2D ou en 3D ? Son modèle doit-il être réaliste ou bien caricatural ? Doit-il ressembler à un jeune homme, à une femme mûre ? Doit-il avoir un aspect humanoïde ou au contraire doit-il avoir l'aspect d'un animal de compagnie, ou encore d'un personnage de dessin animé populaire ? L'agent doit-il plutôt ressembler à un français ou bien à un japonais ? Une fois que les informations sur la géométrie de l'agent ont été déterminées pour une application donnée et un public ciblé, il faut se demander non seulement comment l'agent doit communiquer des informations, mais aussi comment l'agent doit interagir avec les utilisateurs ? Il faut donc décider si l'agent doit être introverti ou extraverti, emphatique ou peu expressif ? Quelle modalité l'agent doit-il utiliser pour appuyer son discours : son visage plutôt que son regard, ou bien encore des gestes ? Est-ce que l'agent doit être redondant dans son comportement, et s'exprimer

avec tout son corps ? En quoi cette redondance des comportements affecte-t-elle le style de l'agent ? Cette redondance agit-elle sur la qualité de l'interaction, et si oui, selon quelle dimension (apprentissage, maintien de l'attention, attirance de l'attention, etc.) ? Quel rôle l'agent doit-il avoir dans l'interaction (preneur de décision, informateur, conseiller, etc.) ? quels doivent être ses buts ? doit-il persuader l'utilisateur, ou tout simplement fournir de l'information, ou bien doit-il se montrer emphatique ? A partir de cette longue liste de questions, on peut entrevoir la diversité des choix qui doivent être établis lors de la création d'un agent. L'évaluation de ces choix et leurs effets sur l'interface est donc une nécessité.

La métaphore de la conversation humain - humain est très souvent appliquée [CAS 00] à la création d'agents. Un premier type d'évaluation possible est donc de comparer le comportement de l'agent avec celui d'un être humain. La comparaison peut se faire en utilisant une évaluation objective : par exemple, une comparaison entre des données réelles de paramètres labiaux et des valeurs calculées de ces paramètres permet d'évaluer un modèle algorithmique des mouvements des lèvres durant la parole. Au contraire, une évaluation d'intelligibilité est obtenue en faisant passer des tests de perception à un ensemble de sujets. Certaines études [BES 03] ont montré que les résultats obtenus par les deux types d'évaluation, évaluation objective et évaluation d'intelligibilité, ne sont pas nécessairement corrélés. En effet, en prenant de nouveau l'exemple de l'algorithme de calcul du mouvement des lèvres durant la parole, les modèles utilisant une hyper-articulation donnent de bien meilleurs résultats lors de tests de perception, même si les mouvements obtenus par ce type de modèles ne sont pas très naturels.

L'évaluation en soi doit se faire à différents niveaux : au niveau atomique où un aspect particulier de l'agent est observé ; au niveau de l'utilisateur où les réactions de celui-ci sont évaluées ; au niveau de l'application dans laquelle l'agent est impliqué.

13.5.1. *Evaluation au niveau atomique*

Ces évaluations s'intéressent à un seul aspect de l'agent à la fois. Par exemple, elles évaluent le comportement de l'agent en fonction d'une modalité particulière (visage, regard, corps, etc.). [KRH 04] met en évidence que le mouvement des sourcils par rapport au pitch de la voix n'a pas le même rôle pour le hollandais ou pour l'italien. Le regard a été beaucoup étudié [KRA 04], [GAR 03]. Ce comportement a un impact important sur la façon dont un utilisateur perçoit un agent. Buisine et Martin [BUI 04] ont étudié si les stratégies qui utilisent la multimodalité (i.e. communiquer la même information par le regard et/ou la parole et/ou le geste...) ont un impact non seulement sur la façon dont l'agent est perçu, mais aussi sur le niveau de rétention des informations. Ces auteurs ont mené une étude dans laquelle trois agents d'aspects physiques différents utilisent ou bien une

seule modalité (par exemple, le regard ou le geste), ou bien un comportement redondant (tel que le regard et le geste) pour communiquer une information donnée. D'autre part, Cassell et Thòrison [CAS 99] ont montré que les utilisateurs préfèrent nettement dialoguer avec des agents qui communiquent de manière multimodale. Les utilisateurs trouvent les agents plus naturels s'ils communiquent par le regard, les expressions du visage et les mouvements du corps ; de plus la multimodalité a une influence sur le caractère persuasif de l'agent.

13.5.2. *Evaluation au niveau de l'utilisateur*

Ces évaluations s'intéressent plus particulièrement au point de vue de l'utilisateur. Les recherches de Höök [HÖÖ 00], [HÖÖ 03], [HÖÖ 04] en sont un bon exemple. Höök s'interroge sur la contribution qu'une caractéristique donnée d'un agent peut avoir sur l'interaction que les utilisateurs ont avec celui-ci. A travers plusieurs applications, elle examine comment le rôle, le comportement, la parole et les émotions de l'agent agissent sur l'état émotionnel des utilisateurs. Elle argumente que la création d'un agent doit se faire dans une boucle mélangeant développement et évaluation. Ainsi, un agent ne peut être défini qu'à travers des études d'évaluation qui examinent les effets que celui-ci peut avoir dans l'interaction et l'expérience que les utilisateurs ont avec le système. Nass et ses collègues ont établi d'importants résultats sur le design d'un agent en basant leur recherche sur des études psychologiques. Ils ont entre autres étudié le rôle que la culture [LEE 98], la personnalité [ISB 04] ou l'apparence physique [NAS 04] peuvent exercer sur la perception de l'agent par les utilisateurs.

13.5.3. *Evaluation au niveau de l'application*

Alors que plusieurs recherches ont confirmé la capacité des agents à motiver les utilisateurs [CAS 00], [LES 97], [MOR 01], peu de travaux ont étudié l'apport qu'un agent peut avoir pour une tâche précise en terme de performance de l'utilisateur [DEH 00]. Plusieurs expériences ont examiné l'effet que l'aspect vestimentaire des agents [MCB 04], l'intonation de la voix [DAR 04] ou le sexe [BUI 04] peuvent exercer sur l'effet de confiance transmis par un agent.

13.5.4. *Evaluation d'agents pédagogiques*

Stone et Lester [STO 96] ont établi trois propriétés essentielles qu'un agent pédagogique doit posséder : il doit discuter, présenter des informations, donner des explications aux bons moments et avec le niveau de détail nécessaire ; il doit tenir compte du contexte (tel que le niveau de l'utilisateur) lors de son intervention ; il doit maintenir une continuité en ayant un comportement cohérent, pédagogique et crédible. En plus de l'évaluation de l'agent en tant que tel, un agent pédagogique doit donc être examiné selon ces trois dimensions.

S'appuyant sur expérimentation, Baylor [BAY 03] a soulevé la question de savoir s'il était plus efficace d'avoir un seul agent pédagogique (appelé Mentor) qui combinerait une expertise intellectuelle avec un support de motivation, ou bien deux agents qui se répartiraient les tâches – apport de connaissance (l'expert) et apport de motivation (le motivateur). Dans le cadre d'un système non adaptatif, Baylor a montré qu'utiliser deux agents pédagogiques se partageant de tels rôles a un impact positif sur l'apprentissage ainsi que sur la qualité perçue des agents.

Une autre étude récente a montré que lorsque des étudiants interagissent avec des agents qui parlent pour communiquer des informations plutôt qu'ils n'écrivent, l'intérêt que les étudiants portent à l'agent, mais aussi le niveau de rétention d'information et la performance à des tests, étaient augmentés [MOR 01].

Darves et Oviatt [DAR 04] ont exploré le rôle de la qualité de la voix dans le design d'un agent. Les auteurs ont basé leur étude sur des expériences montrant que si un professeur parle avec une voix dynamique et énergique, voix associée à l'extraversion, [ISB 04], [SCH 79] les étudiants restent plus concentrés et plus attentifs. Elles ont observé que les agents ayant une telle voix (voix extravertie) étaient plus efficaces pour stimuler et motiver les jeunes utilisateurs.

13.6. Perspectives et conclusions

Nous envisageons d'autres scénarios de coopération entre l'hypermédia et l'agent pédagogique, notamment pour des animations ou exercices plus interactifs. Pour la sélection des actes communicatifs et pédagogiques de l'agent, nous tiendrons compte des évaluations où les agents pédagogiques se sont révélés le plus efficace.

Le comportement de l'agent pédagogique peut s'adapter en plusieurs points de contrôle : type d'acte communicatif à mettre en oeuvre, variation de la réalisation de cet acte (par exemple, réaliser un *request* par « vous devez revoir » ou « veuillez revoir »), moment de la réalisation d'un acte communicatif, choix du type de dialogue, réalisation multimodale (plus ou moins de redondance, plus ou moins d'émotion) ou bien sûr par la sélection du contenu verbal.

Outre les actes communicatifs répertoriés dans notre étude bibliographique, l'agent pédagogique peut intervenir dans les actes suivants pour lesquels la communication non-verbale est bien adaptée :

- attirer l'attention / insister sur ;
- encourager (en général ou sur un point particulier) ;
- actes sociaux (e.g. dialogue de bienvenue) ;

- montrer comment exécuter une action ;
- aider à la navigation : interactivité (« voulez-vous plus d'informations sur X ou Y ? »), branchements entre les sections (« Je vous conseille de suivre maintenant le cours sur ... »), lien vers une autre partie du cours, éclaircir les transitions et les liens entre les chapitres, conseiller de réviser un point particulier ;
- lister / énumérer les étapes d'un traitement séquentiel ou expliquer les alternatives possibles ;
- avertir de ce qu'il ne faut pas faire ;
- informer sur des points difficiles / importants / sérieux / ennuyeux (définition, théorie, important à retenir, problème à résoudre) ;
- montrer des éléments amusants, motivants, positifs (analogie, exemple, nouveau point abordé, solution, avantage d'une solution, astuce) ;
- répéter une information déjà consultée.

Parmi les réalisations multimodales de ces actes communicatifs, les coopérations et relations temporelles entre expressions faciales et déictiques (pointer avec la main et encourager avec l'expression faciale) nous semblent particulièrement riches.

Nous avons décrit l'intégration d'un agent pédagogique et d'un hypermédia adaptatif éducatif. Le système que nous proposons repose sur un modèle du domaine spécifié en XML dans lequel les documents sont représentés en fonction d'un modèle de structure pédagogique. Un module de contrôle est chargé de répartir les tâches de présentation à l'apprenant entre l'hypermédia et l'agent. Ce travail ouvre de nombreuses perspectives de recherche.

La structure pédagogique simplifiée de document que nous avons prise en compte pour l'instant devra être complétée. D'autre part, le prototype actuel ne permettant l'action de l'apprenant sur le système que par l'intermédiaire des liens dans l'hypermédia, il serait intéressant d'étudier les possibilités d'une action en entrée sur l'agent. La question de l'influence du degré de coopération/redondance entre l'agent et l'hypermédia dans les informations présentées à l'apprenant reste ouverte. Certaines informations peuvent être présentées par l'hypermédia et reprises par l'agent, d'autres informations peuvent être présentées par l'un des deux seulement. Nous envisageons d'utiliser le prototype actuel pour réaliser une étude d'évaluation de différents scénarios de répartition des tâches pour laquelle nous ferons appel à un public d'apprenants avec un protocole s'inspirant par exemple de [BUI 02]. En situation d'enseignement classique, les interactions entre enseignants et apprenants font intervenir plusieurs modalités communicatives (parole, gestes, regard, posture, expressions faciales, dessin sur un tableau, transparents). Bien que l'utilisation de vidéos éducatives se développe dans le domaine des EIAH, celles-ci sont rarement utilisées comme des ressources permettant d'annoter et de mieux

comprendre la multimodalité de la communication dans ce contexte pédagogique. Comme nous l'avons vu plus haut, même dans le domaine des agents pédagogiques, le comportement multimodal de l'agent est souvent limité en expressivité et n'est pas fondé sur une analyse fine de comportements multimodaux d'enseignants, mais plutôt sur des règles générales issues par exemple de la littérature en sociolinguistique. Un travail futur consistera à étudier l'apport que peut représenter l'utilisation de corpus vidéo pour la détermination de scénarios pédagogiques [VER 01], la spécification du comportement multimodal de l'agent, et plus généralement les conséquences pour la production de contenus pédagogiques interactifs.

Remerciements : MICAME a été soutenu par le BQR de l'Univ. Paris 8. L'agent LEA a été développé au LIMSI-CNRS par S. Abrilian, S. Buisine, C. Rendu et J.-C. Martin dans le projet NICE (www.niceproject.com). Le système MICAME est développé au LINC par J.-H. Réty, P. Dubrulle et J.-C. Martin.

Bibliographie

- [ABR 02] ABRILIAN S., BUSINE S., RENDU C., MARTIN J.-C., "Specifying Cooperation between Modalities in Lifelike Animated Agents", *Proc. of International Workshop on Lifelike Animated Agents: Tools, Functions, and Applications, 7th Pacific Rim International Conference on Artificial Intelligence (PRICAI'02)*, 2002.
- [ABR 03] ABRILIAN S., MARTIN J.-C., BUISINE S., "Algorithms for controlling cooperation between output modalities in 2D Embodied Conversational Agents", *Fifth International Conference on Multimodal Interfaces (ICMI'2003)*, ACM Press, p. 293-296, 2003.
- [BAY 03] BAYLOR A., "The impact of three pedagogical agent roles", AAMAS2003, Melbourne, Australia, 2003.
- [BES 03] BESKOW J., *Talking Heads: Models and Applications for Multimodal Speech Synthesis*, PhD thesis, Centre for Speech Technology, KTH, Stockholm, Sweden, 2003.
- [BON 00] BONFIGLI M. E., CASADEI G., SALOMONI P., "Adaptive Intelligent Hypermedia using XML", *Proc. of SAC 2000 - ACM Symposium on Applied Computing*, Como, Italy, 2000.
- [BRU 96] BRUSILOVSKY P., "Adaptive Hypermedia: an Attempt to Analyze and Generalize", *Multimedia, Hypermedia and Virtual Reality*, Springer-Verlag LNCS n.1077, p. 288-304, 1996.
- [BRU 01] BRUSILOVSKY P., "Adaptive Hypermedia", *User Modeling and User-Adapted Interaction*, vol. 11, p. 87-110, 2001.
- [BUI 02] BUISINE S., ABRILIAN S., RENDU C., MARTIN J.-C., "Towards Experimental Specification and Evaluation of Lifelike Multimodal Behavior", *Workshop on Embodied conversational agents - let's specify and evaluate them!*, First International Joint Conference on Autonomous Agents & Multi-Agent Systems, Bologna, Italy, 2002.

- [BUI 03] BUISINE S., ABRILIAN S., MARTIN J.-C., "Evaluation of individual multimodal behavior of 2D embodied agents in presentation tasks", *Workshop Embodied conversational characters as individuals*, 2nd International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS'03), Melbourne, Australia, 2003.
- [BUI 04] BUISINE S., ABRILIAN S., MARTIN J.-C., "Evaluation of multimodal behaviour of embodied agents", In Z. RUTTKAY and C. PELACHAUD, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [CAR 99] CARRO R. M., PULIDO E., RODRÍGUEZ P., "Task-based Adaptive learner Guidance On the WWW: the TANGOW System", *8th International Conference on Human-Computer Interaction (HCI'99)*, 1999.
- [CAS 99] CASSELL J., THORISSON K., "The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents", *Applied Artificial Intelligence*, 13(3), 1999.
- [CAS 00] CASSELL J., SULLIVAN J., PREVOST S., CHURCHILL E., *Embodied Conversational Agents*, MIT Press, 2000.
- [DAR 04] DARVES C., OVIATT S., "Talking to digital fish", In Z. RUTTKAY and C. PELACHAUD, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [DEB 02] DE BRA P., STASH N., "AHA! Adaptive Hypermedia for All", *SANE 2002*, Maastricht, 2002.
- [DEH 00] DEHN D., MULKEN S., "The impact of animated interface agents: a review of empirical research.", *International Journal of Human-Computer Studies*, vol. 52, p. 1-22, 2000.
- [DEL 98] DELESTRE N., PECUCHET JP., GREBOVAL-BARRY C., "L'architecture d'un hypermédia adaptatif dynamique pour l'enseignement", *NTICF'98 (Nouvelles Technologies de l'Information et de la Communication dans les formations d'Ingénieurs et dans l'industrie)*, Rouen, p. 383-391, novembre 1998.
- [DEL 00] DELESTRE N., METADYNE : un hypermedia adaptatif dynamique pour l'enseignement, Thèse de doctorat, Université de Rouen, 20 janvier 2000.
- [DIL 93] DILLENBOURG P., HILARIO M., MENDELSON P., SCHNEIDER D., BORCIC, B., "Intelligent Learning Environments", Report from the project *Les systèmes explorateurs intelligents*, NFP23 Program; Project N° 4023-2701, 1993.
<http://tecfa.unige.ch/tecfa/research/memolab/report93.book.html>
- [EVE 00] EVERS M., NIJHOLT A., "Jacob - An Animated Instruction Agent in Virtual Reality", *ICMI 2000*, LNCS 1948, Springer-Verlag, p. 526-533, 2000.
- [FEN 02] FENTON-KERR T., "Some roles and designs for speech-enabled interface agents in language learning", *International Workshop on Lifelike Animated Agents: Tools, Functions, and Applications*, held in conjunction with the 7th Pacific Rim International Conference on Artificial Intelligence (PRICAI'02), 2002.
- [GAR 03] GARAU M., SLATER M., VINAYAGAMOORTHY V., BROGNI A., STEED A., SASSE M.A., "The impact of avatar realism and eye gaze control on perceived quality of

communication in a shared immersive virtual environment", *In Proceedings of the SIG-CHI conference on Human factors in computing systems*, Fort Lauderdale, FL, USA, April 5-10, 2003.

- [GRA 02] GRAESSER A. C., LEON J. A., OTERO J. C., "Introduction to the psychology of science text comprehension", In *The psychology of science text comprehension*. J. A. L. J. OTERO, & A. C. GRAESSER, Mahwah, NJ: Erlbaum, p. 1-15, 2002.
- [GRA 03] GRAESSER A. C., MORENO K., MARINEAU J., ADCOCK A., OLNEY A., PERSON N., "AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head?", *Proceedings of Artificial Intelligence in Education*, p.47-.54, 2003.
- [HÖÖ 00] HÖÖK K., PERSSON P., SJÖLINDER M., "Evaluating users' experience of a character-enhanced information space", *Journal of AI Communications*, 13(3), p. 195-212, 2000.
- [HÖÖ 03] HÖÖK K., BULLOCK A., PAIVA A., VALA M., CHAVES R., PRADA R., "FantasyA and SenToy", *In Proc. of the conference on Human factors in computing systems*, p. 804--805, 2003.
- [HÖÖ 04] HÖÖK K., "User-centred design and evaluation of affective interfaces", In Z. RUTKAY and C. PELACHAUD, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [ISB 04] ISBISTER K., NASS C., "Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics", *International Journal of Human-Computer Studies*, To appear.
- [JOH 00] JOHNSON W. L., RICKEL J. W., LESTER J. C., "Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments", *International Journal of Artificial Intelligence in Education*, vol. 11, p. 47-78, 2000.
- [KRA 04] KRAEMER N.C., TIETZ B., BENTE G., "Effects of embodied interface agents and their gestural activity", In Z. RUTKAY and C. PELACHAUD, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [KRH 04] KRAHMER E., SWERTS M., "More about brows", In Z. Ruttkay and C. Pelachaud, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [LEE 98] LEE E.-J., NASS C., "Does the ethnicity of a computer agent matter? An experimental comparison of human-computer interaction and computer-mediated communication", *In WECC'98, The First Workshop on Embodied Conversational Characters*, October 1998.
- [LES 97] LESTER J., CONVERSE S., KAHLER S., BARLOW T., STONE B., BHOGAL R., "The Persona Effect: Affective Impact of Animated Pedagogical Agents", *CHI '97*, Atlanta, 1997.
- [LES 99] LESTER J. C., STONE B. A., STELLING, G. D., "Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments", *User Modeling and User-Adapted Interaction* 9(1-2), p. 1-44, 1999.
- [LES 00] LESTER, J. C., TOWNS, S.G., CALLAWAY, C.B., VOERMAN, J.L., FITZGERALD P. "Deictic and emotive communication in animated pedagogical agents", in *Embodied*

Conversational Agents, J. Cassell, Sullivan, J., Prevost, S., Churchill, E. (Eds.), The MIT Press, p. 123-154, 2000.

- [MAR 02] MARTIN J.-C., RÉTY J.-H., BENSIMON N., "Multimodal and Adaptive Pedagogical Resources", *Electronic proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC'2002)*, Las Palmas, Spain, 2002.
- [MCB 04] McBREEN H., MORTON H., JACK M., "Experimental evaluation of the use of ECAs in eCommerce applications", In Z. RUTTKAY and C. PELACHAUD, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [MOR 01] MORENO R., MAYER R., SPIRES H. A., LESTER J. C., "The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents?", *Cognition and Instruction*, vol. 19, p. 177-213, 2001.
- [NAS 04] NASS C., ROBLES E., WANG Q., "Apparent attention", In Z. RUTTKAY and C. PELACHAUD, editors, *From Brows till Trust: Evaluating Embodied Conversational Agents*, Kluwer, To appear.
- [RIC 99] RICKEL J., JOHNSON W. L., "Animated agents for procedural training in virtual reality: Perception, cognition, and motor control", *Applied Artificial Intelligence*, vol.13, p. 343-382, 1999.
- [SCH 79] SCHERER K.R., "Personality markers in speech", In SCHERER K.R. and GILES H. editors, *Social Markers in Speech*, Cambridge Univ. Press, Cambridge, UK, p. 147-209, 1979.
- [SHA 99] SHAW E., GANESHAN R., JOHNSON W. L., MILLAR D., "Building a case for agent-assisted learning as a catalyst for curriculum reform in medical education", *Proceedings of the Ninth International Conference on Artificial Intelligence in Education*, 1999.
- [STO 96] STONE B., LESTER J., "Dynamically sequencing an animated pedagogical agent", *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, p. 424-431, 1996.
- [VAS 95] VASSILEVA J., "Dynamic Courseware Generation: at the Cross Point of CAL, ITS and Authoring", *International Conference on Computers in Education*, p. 290-297, 1995.
- [VER 01] VERCLYTTÉ L., USEILLE P., LABOUR M., "Scénario Interactif/scénario pédagogique interactif : convergence et divergence d'une écriture", *H2PTM'01*, 2001.
- [VIL 02] VILLANOVA-OLIVER M., *Adaptabilité dans les systèmes d'information sur le Web : modélisation et mise en œuvre de l'accès progressif*, Thèse de doctorat, Institut National Polytechnique de Grenoble, 18 décembre 2002.