

Un environnement multimédia privilégiant l'expression en langage naturel, pour des activités éducatives de type piagétien.

Élisabeth GODBERT

Université de la Méditerranée et CNRS, France

Résumé : *Nous nous intéressons dans cet article au développement de logiciels éducatifs proposant, dans un environnement multimédia, multimodal et réactif privilégiant l'expression en langage naturel, des exercices de type piagétien pour l'aide à l'apprentissage du langage. Dans ce type de système, chaque exercice met en jeu un micro-monde d'objets graphiques sur lesquels l'enfant utilisateur peut agir (modification, déplacement, suppression d'objets, etc.), en formulant des requêtes en langage naturel auxquelles le système répond.*

Nous présentons un logiciel éducatif de cette famille, que nous avons conçu et développé ces dernières années. Nous décrivons en premier lieu les objectifs pédagogiques et les fonctionnalités de notre logiciel, puis le système que nous y utilisons pour le traitement automatique du langage.

Nous décrivons ensuite l'ensemble des composants qui structure les connaissances qu'il a été nécessaire de modéliser, articuler et gérer pour la réalisation de ce système.

- [1. Introduction](#)
- [2. Connexions avec les domaines STI et IHM](#)
- [3. Objectifs pédagogiques du système EREL](#)
- [4. Le système ILLICO pour le traitement du langage écrit](#)
- [5. Description fonctionnelle du système EREL](#)
- [6. Les composants du système EREL](#)
- [7. Validation du système](#)
- [8. Conclusion et perspectives](#)
- [Références](#)



1. Introduction

Nous nous intéressons dans cet article au développement de logiciels éducatifs, proposant, dans un environnement multimédia, multimodal et réactif, des exercices simples de type piagétien pour l'aide à l'apprentissage du langage.

Les travaux de Piaget sur le développement des fonctions cognitives de l'intelligence et de la pensée ont fortement marqué le domaine de la psychologie de l'enfant, ils ont montré que le développement de

l'enfant passe par l'acquisition, par étapes, de "structures de connaissance" (Piaget & Inhelder, 1966). Pour effectuer ces études, Piaget a utilisé un ensemble d'exercices qui visent à enseigner des abstractions par le biais de contextualisations : ces exercices mettent en jeu des contextes simples (ensemble d'objets de différentes formes, de différentes tailles, en nombre variable...) et proposent des problèmes logico-mathématiques de catégorisation d'objets, de sériation, etc. Si les théories piagésiennes ont été quelque peu critiquées par d'autres écoles de psychogénèse, il reste que les exercices de type piagésien sont universellement utilisés par les psychologues, éducateurs et rééducateurs, car ils permettent aux psychologues d'évaluer certains stades du développement cognitif de l'enfant, et aux éducateurs d'aider l'enfant dans l'acquisition de ces fonctions cognitives.

Les expériences piagésiennes, et de nombreuses autres, ont par ailleurs montré que le langage décuple le pouvoir de la pensée en étendue et en rapidité et que la coordination de l'expression en langage naturel et l'expression graphique ou gestuelle est un atout majeur pour la construction de liens entre les aspects langage-dessin-logique des objets perçus (voir par exemple les travaux de Bruner sur l'action, le langage et la pensée dans (Bruner, 1983)).

Par ailleurs, les recherches dans le domaine du traitement automatique du langage naturel (dans la suite, TALN) ont abouti ces dernières années au développement de nombreuses techniques et outils sophistiqués permettant l'analyse et la synthèse de phrases. S'il est clair que l'élaboration d'un système de compréhension et de traitement du langage dans toute sa généralité est une utopie, la réalisation d'un système complet pour l'analyse, la synthèse, et la compréhension de phrases est possible dans le cas d'une couverture linguistique restreinte dans le cadre d'une application particulière, car l'ensemble des connaissances à prendre en compte (sur le langage lui-même, sa sémantique, le monde du discours, etc.) est alors de taille raisonnable.



Les exercices piagésiens mettant chacun en jeu un contexte simple, le développement d'un module de TALN pour la réalisation informatique d'exercices de ce type est envisageable. Il nous a donc semblé intéressant et réaliste de chercher à concevoir des logiciels pour l'aide à l'éducation et/ou la rééducation de fonctions cognitives et langagières, qui reprennent des exercices de type piagésien et utilisent des outils robustes pour le TALN. Plus précisément, nous voulons concevoir des logiciels éducatifs dans lesquels chaque exercice met en place un micro-monde d'objets aux propriétés précises, organisés en une scène que l'on veut pouvoir décrire de façon simple en langage ordinaire et sur laquelle l'utilisateur doit pouvoir agir par l'intermédiaire de requêtes en langage naturel. Nous voulons par ailleurs intégrer ces exercices dans un environnement.

- Dans lequel on coordonne pour le dialogue homme-machine le langage naturel et les possibilités graphiques et, plus généralement, des aspects multimédias et multimodaux (MM) de la communication avec l'ordinateur. On parle alors de dialogue LN+MM et de système multimédia intelligent (Wahlster et al., 1993).
- Que l'on peut qualifier d'environnement graphique réactif car il répond et/ou obéit aux requêtes que formule l'utilisateur pour manipuler (ajouter, supprimer, modifier...) des objets représentés graphiquement à l'écran.

La conception de tels systèmes est complexe, elle dépend des domaines des Systèmes Tuteurs Intelligents (STI), de l'Enseignement du Langage Assisté par Ordinateur, du TALN et de l'Interaction Homme-Machine (IHM) multimédia, et doit donc tirer parti des recherches et techniques développées dans ces domaines.

La partie 2 de cet article propose un bref tour d'horizon sur ces différents domaines connectés à notre problématique.

Les parties suivantes décrivent le système éducatif EREL, que nous avons conçu et développé ces dernières années, qui propose un ensemble d'activités de type piagésien dans un environnement multimodal réactif. Nous décrivons en premier lieu les objectifs pédagogiques d'EREL et les exercices qui y ont été développés. Nous décrivons ensuite le système générique ILLICO qui apporte les outils pour le TALN et qui a déjà été utilisé pour le développement d'applications (Pasero & Sabatier, 1994 et 95). Pour finir, nous décrivons l'architecture d'EREL, en la comparant avec l'architecture classique des systèmes de type STI.

2. Connexions avec les domaines STI et IHM

En ce qui concerne la conception et le développement de logiciels de type STI, les travaux effectués depuis une quinzaine d'années montrent que même lorsque l'objectif pédagogique reste modeste, ces systèmes doivent prendre en compte des connaissances de types très divers, l'architecture complète d'un STI comprenant classiquement (Matthews, 1992) :

- un module expert, où sont rassemblées les connaissances sur le domaine à enseigner et qui sait résoudre un problème ;
- un module de l'apprenant, où se trouvent les connaissances sur l'élève (ses acquis, erreurs, etc.), et qui sait faire un diagnostic et une mise à jour du profil utilisateur ;
- un module pédagogique, qui réunit les contenus didactiques (choix et reformulation des connaissances à enseigner en fonction de stratégies d'apprentissage) ;
- une interface utilisateur, qui recueille et distribue ces connaissances et gère le dialogue homme-machine.

De plus, ces logiciels doivent répondre à la fois :

- au souci de définir des logiciels effectivement utilisables par le public visé, c'est-à-dire concevoir des systèmes dans un contexte utilisateur (Clancey, 1993) ;
- aux critères mis en évidence dans le domaine des STI dans les grilles d'évaluation des produits : critères ergonomiques, critères de génie logiciel (correction, fiabilité, etc.) et critères correspondant aux spécifications des pédagogues.



Il est reconnu cependant que les systèmes développés jusqu'à maintenant dans ce domaine ne répondent pas à cet idéal car il serait extrêmement difficile de respecter tous ces critères, et très peu de ces systèmes sont réellement employés dans des environnements pour l'éducation ou l'apprentissage (voir par exemple (Khuwaja et al., 1996) ou (Dowsing & Long, 1999)).

En ce qui concerne plus particulièrement l'apprentissage des langues à l'aide de systèmes basés sur le graphisme, on peut mentionner des systèmes comme LingWorlds (Douglas, 1995) ou LINGO (Murray, 1995) qui ont été développés pour l'apprentissage d'une langue en contexte. Les auteurs soulignent la complexité de la conception et du développement de tels systèmes, dans lesquels il faut d'une part tenter de répondre aux exigences du domaine des STI, et dans lesquels d'autre part il est nécessaire d'avoir une interface intelligente pour gérer tous les aspects graphiques du système, ainsi qu'un module de TALN qui permette une expression qui ne soit pas trop contrainte en langage naturel pour le dialogue entre l'utilisateur et la machine.

Des modèles pour la gestion de dialogues ont été définis pour le dialogue homme-machine ou le dialogue homme-homme. Certains ont été intégrés dans des logiciels éducatifs, et dans ce domaine on peut mentionner les travaux de M. Joab (Joab, 1990) qui, après un état de l'art sur le dialogue homme-machine et l'interaction dans les systèmes STI, décrit l'intégration des modules d'une architecture STI pour la réalisation informatique d'un dialogue pédagogique en langage naturel.

Dans le domaine de l'interaction homme-machine, depuis une quinzaine d'années, le dialogue en langage naturel a souvent été intégré dans des systèmes utilisant plusieurs modes d'interaction, gérés par des "interfaces utilisateur intelligentes" (voir par exemple (Wahlster et al., 1993 ; Maybury, 1998 ; Cohen et al. 1998)). Ce domaine a vu la formalisation et la réalisation de systèmes utilisant le multimédia et mettant en oeuvre une interaction multimodale dans laquelle les différentes techniques de communication (parole, vision par ordinateur, textes, images, vidéo, etc.) sont combinées et coordonnées (Nigay & Coutaz, 1996) - un média étant défini comme un dispositif servant de support à l'information, et une modalité comme une technique d'interaction -.

Ce bref tour d'horizon sur les domaines connectés à notre problématique nous permet d'avoir une vue globale de l'ensemble des connaissances qui doivent être utilisées pour la conception et la réalisation des systèmes qui nous intéressent. Elles proviennent de deux sources :

1. les connaissances de type pédagogique, qui viennent des spécifications des éducateurs et pédagogues, sont les connaissances à enseigner et les critères didactiques ;

2. les connaissances de type informatique sont les théories, méthodes, techniques, outils informatiques issus des domaines du multimédia, de l'interaction multimodale, du TALN, et du développement de logiciels.



3. Objectifs pédagogiques du système EREL

Le projet EREL (Education, Rééducation, Evaluation du Langage), conçu en 1996 et développé ces dernières années a pour objectif le développement d'un système pour l'éducation et/ou la rééducation du langage et de la cognition chez des enfants (de 5 à 10 ans) présentant des troubles du développement. Ce projet repose sur l'idée suivante : utiliser des techniques d'intelligence artificielle pour développer un système privilégiant l'expression en langage naturel dans un environnement multimédia, ludique, interactif et modulaire, qui :

- propose un ensemble d'exercices conçus pour stimuler, encourager, et aider l'utilisateur à employer le langage naturel pour s'exprimer autour de différents thèmes et interagir avec la machine ;
- utilise conjointement plusieurs supports pour assister l'utilisateur dans son expression ;
- ne demande à l'utilisateur que des manipulations simples ;
- se présente comme un système multi-utilisateurs, qui permette de choisir le niveau de difficulté du travail demandé, en fonction des compétences de l'utilisateur.

Comment définir ou choisir ces exercices ?

Pour le choix des exercices, notre point de départ réside dans les études piagétienne et les théories constructivistes de la psychologie cognitive selon lesquelles l'enfant acquiert des structures de connaissance à partir de ses expériences sur son environnement, la découverte de ces structures découlant non seulement de la perception des objets mais surtout d'activités portant sur ces objets. Il faut donc fournir à l'enfant du matériel expérimental et l'encourager à l'action. Notre objectif est de définir dans EREL des activités qui demandent à l'enfant de raisonner, agir et s'exprimer sur des objets dont les propriétés évoluent dynamiquement.

Les différents stades de développement définis par Piaget comprennent en particulier : la classification d'objets, la sériation, les nombres entiers (exercices de logique quantitative), les déplacements dans l'espace avec la mesure, les structures géométriques, les étapes de la géométrie passant par les relations intrafigurales puis interfigurales (utilisant les coordonnées cartésiennes) et l'algébrisation de la géométrie (groupe des déplacements par exemple).

Sur ces bases, nous avons conçu une suite d'exercices portant sur la désignation et le positionnement les uns par rapport aux autres d'objets de forme, couleur et taille variables :

- en dimension 1, on raisonne sur un alignement d'objets orientés ou non, par exemple un train formé de plusieurs wagons et comportant ou non une locomotive pour l'orienter ;
- en dimension 2, on travaille sur les relations interfigurales et les coordonnées cartésiennes de pions posés sur un damier, que l'on peut déplacer, supprimer, etc.
- en dimension 3, on s'intéresse à une image en trois dimensions, dont on peut décrire les éléments, leurs propriétés et interrelations.

Notons que la description d'une scène peut se faire suivant différents points de vue ou systèmes de repérage (intrinsèque, extrinsèque, déictique) que l'enfant acquiert en plusieurs étapes au cours de son développement cognitif (voir par exemple (Humblot, 1998)).



Pour chacun des jeux proposés par EREL, l'enfant doit utiliser et apprendre à maîtriser le langage associé, qui lui permet d'exprimer ses requêtes. C'est un sous-langage simple du langage naturel, qui contient les expressions linguistiques intuitives nécessaires à la description des propriétés des éléments du contexte, des situations logico-mathématiques et des relations spatiales qui ont une occurrence dans le jeu considéré.

Par ailleurs, EREL a été conçu comme un système multi-utilisateurs et propose donc des exercices de différents niveaux de difficulté, tant pour ce qui concerne l'expression en langage naturel que pour la complexité cognitive des activités.

Ces jeux ont été définis, plus particulièrement, pour aider l'utilisateur à comprendre, appréhender et exprimer en langage naturel différents concepts concernant :

- la catégorisation des objets présents à l'écran, à partir de leur forme, leur couleur et leur taille;
- la construction d'objets à partir de propriétés de ce type ;
- les relations spatiales entre ces objets ;
- la localisation, le placement et le mouvement de ces objets ;
- la sériation d'objets ;
- des éléments de logique quantitative sur l'ensemble des objets de la scène ;
- la désignation d'un objet à partir de ses propriétés spécifiques.

Nous pensons que le dernier point est particulièrement intéressant dans un contexte éducatif : la désignation d'un objet requiert la formulation d'une description référentielle pertinente, qui doit permettre à un auditeur d'identifier un objet particulier parmi l'ensemble des objets du contexte courant (Donellan, 1966). Nous avons mis en oeuvre dans EREL un mode de désignation LN+MM qui combine et fait coopérer, de plusieurs façons, le langage écrit et la sélection graphique à l'écran (voir [cette partie](#)).

4. Le système ILLICO pour le traitement du langage écrit

Dans le système EREL, l'ingénierie linguistique est assurée par le système générique ILLICO (Pasero & Sabatier, 1994 et 1995). Ce système est constitué d'un ensemble d'outils pour le traitement du langage naturel écrit, qui permettent le développement d'applications du TALN, comme des interfaces intelligentes pour bases de données, des systèmes d'aide à la communication, etc.

ILLICO fournit en premier lieu des formalismes simples pour définir le sous-langage approprié à chaque application, par l'écriture de règles lexicales, syntaxiques, sémantiques, conceptuelles et contextuelles. Ensuite, les algorithmes du coeur d'ILLICO permettent d'effectuer un traitement complet des phrases appartenant à ce sous-langage : analyse ou synthèse d'énoncés dont on produit une représentation sémantique de type logique, dans laquelle les expressions référentielles sont résolues. C'est dire que le système "comprend exactement" les textes analysés et est donc à même de faire ensuite des inférences utilisant leur contenu.



Les algorithmes d'ILLICO mettent en oeuvre un traitement interactif du langage en effectuant l'analyse/synthèse partielle des phrases en cours de construction et en en contrôlant à chaque instant la correction. L'analyse/synthèse partielle de phrases permet en particulier de proposer à l'utilisateur le mode de "composition assistée" dans lequel le système aide l'utilisateur à formuler son texte, en synthétisant et en affichant à chaque instant les mots ou expressions qu'il est possible d'utiliser pour continuer la phrase en cours de construction (Pasero & Sabatier, 1994 ; Milhaud 1994). Cette assistance peut être graduée : assistance au niveau lexical seul, ou lexical+syntaxique, etc., selon que l'on connecte ou non les contrôles linguistiques aux différents niveaux.

Pour le développement d'un logiciel d'éducation ou rééducation du langage, la composition en mode assisté est un trait particulièrement intéressant, car elle permet de guider l'utilisateur pas à pas lors de la formulation de ses phrases. Elle permet donc de définir des exercices pour des enfants très jeunes, en tout début de phase d'apprentissage du langage écrit.

5. Description fonctionnelle du système EREL

5.1. Description générale

Chaque activité d'EREL est organisée pour que l'utilisateur s'exprime par des phrases simples, appelées requêtes, que le système l'aide à composer autour d'une scène. Un langage approprié est associé à chaque

activité pour la composition de ces phrases. La scène évolue au fur et à mesure que l'activité se déroule, en fonction des requêtes que formule l'utilisateur et auxquelles le système réagit. Une session d'EREL, schématiquement, est donc une suite de requêtes de l'utilisateur auxquelles le système obéit si elles sont correctes. Si une requête n'est pas correcte, la communication passe par un sous-dialogue pour assister l'utilisateur dans la formulation de son texte.

La correction globale d'une requête requiert sa correction lexicale, syntaxique et conceptuelle, sa correction contextuelle au niveau linguistique (les expressions référentielles peuvent être résolues), sa correction contextuelle au niveau situationnel (l'action est possible ou non) et sa correction par rapport aux règles du jeu (l'action est permise ou non).

Si une requête de l'utilisateur est correcte et complète, le système y réagit, selon le type de la requête :

- si c'est une phrase impérative, il exécute l'ordre donné (modification des éléments du jeu) ;
- si c'est une phrase interrogative, il donne la réponse à la question posée ;
- si c'est une phrase déclarative, il en donne la valeur de vérité.

Dans tous les cas, le système produit la synthèse vocale du texte composant la requête.

Les niveaux de difficulté des activités proposées par EREL sont définis à partir des éléments suivants, dont on laisse le choix à l'utilisateur (l'éducateur ou l'enfant).

- La complexité cognitive intrinsèque des jeux, fonction pour chacun des règles du jeu et du nombre d'objets graphiques à gérer.
- Les sous-langages utilisés, dont le module lexical et le module syntaxique sont plus ou moins étendus selon le niveau choisi. Nous avons défini trois niveaux de complexité grammaticale : le niveau 1 permet la composition de phrases simples sans proposition relative ; le niveau 2 permet la composition de phrases avec propositions relatives simples ; le niveau 3 permet la composition de phrases avec propositions relatives qui peuvent être imbriquées les unes dans les autres.
- Le niveau d'assistance choisi, qui peut être l'un des trois suivants : lexical, ou lexical+syntaxique, ou lexical+syntaxique+conceptuel.



5.2. Les activités proposées

Nous décrivons ici les activités qui ont été développées et qui constituent la version actuelle du système. Pour chacune d'elles, les différents modes opératoires définis précédemment sont disponibles.

Un jeu en deux dimensions : le jeu du damier

Le jeu du damier est un jeu de positionnement de pions de diverses formes et couleurs sur un damier. L'utilisateur peut composer des phrases de type impératif, interrogatif ou déclaratif sur les pions illustrés à l'écran : il écrit des commandes pour agir sur les pions (par exemple les effacer, les déplacer, les permuter ou en créer de nouveaux), ou formule des questions sur la situation des divers éléments du jeu, et le système réagit à chaque phrase composée par l'utilisateur. Les Figures 1 et 2 donnent un exemple de l'une des versions de ce jeu en mode assisté (aide à la composition de phrases au niveau lexical+syntaxique+conceptuel).

Dans la [Figure 1](#), l'utilisateur ayant commencé sa requête par *Ajoute un carré rouge*, le système indique que les seuls mots pertinents pour continuer la phrase sont les locutions locatives *à droite*, *à gauche*, etc. Dans la [Figure 2](#), l'utilisateur ayant composé la phrase complète *Ajoute un carré rouge à droite de l'étoile.*, le système exécute l'action demandée.

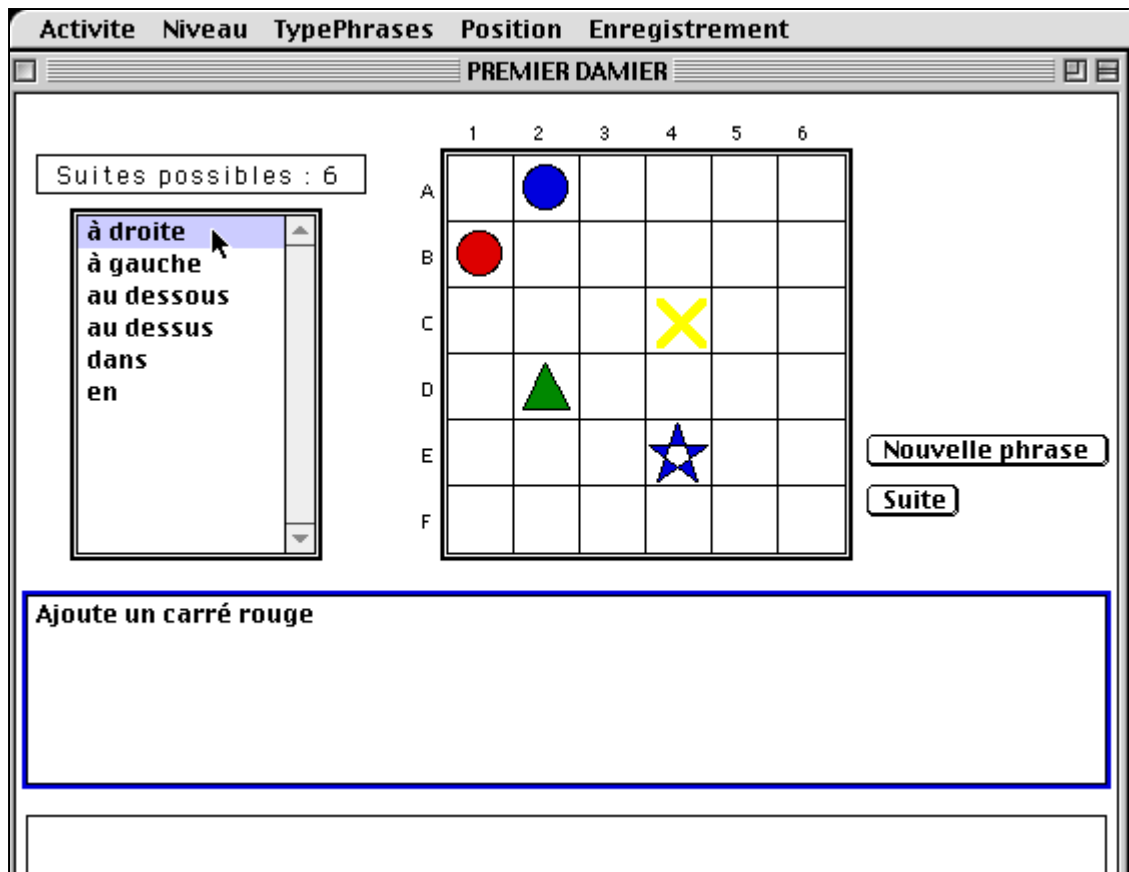


Figure 1 : un jeu en deux dimensions (1).



Voici quelques exemples de phrases (correctes à tous les niveaux) de différents types que peut composer l'utilisateur :

- *Ajoute un carré rouge à droite de l'étoile.*
- *Ajoute un carré gris dans la case F 3.*
- *Permute la croix avec le pion qui se trouve à droite de l'étoile.*
- *Place le rond rouge au-dessus du pion qui se trouve à gauche de la case D 3.*
- *Quel pion se trouve en B 5 ?*
- *Où est le rond bleu ?*

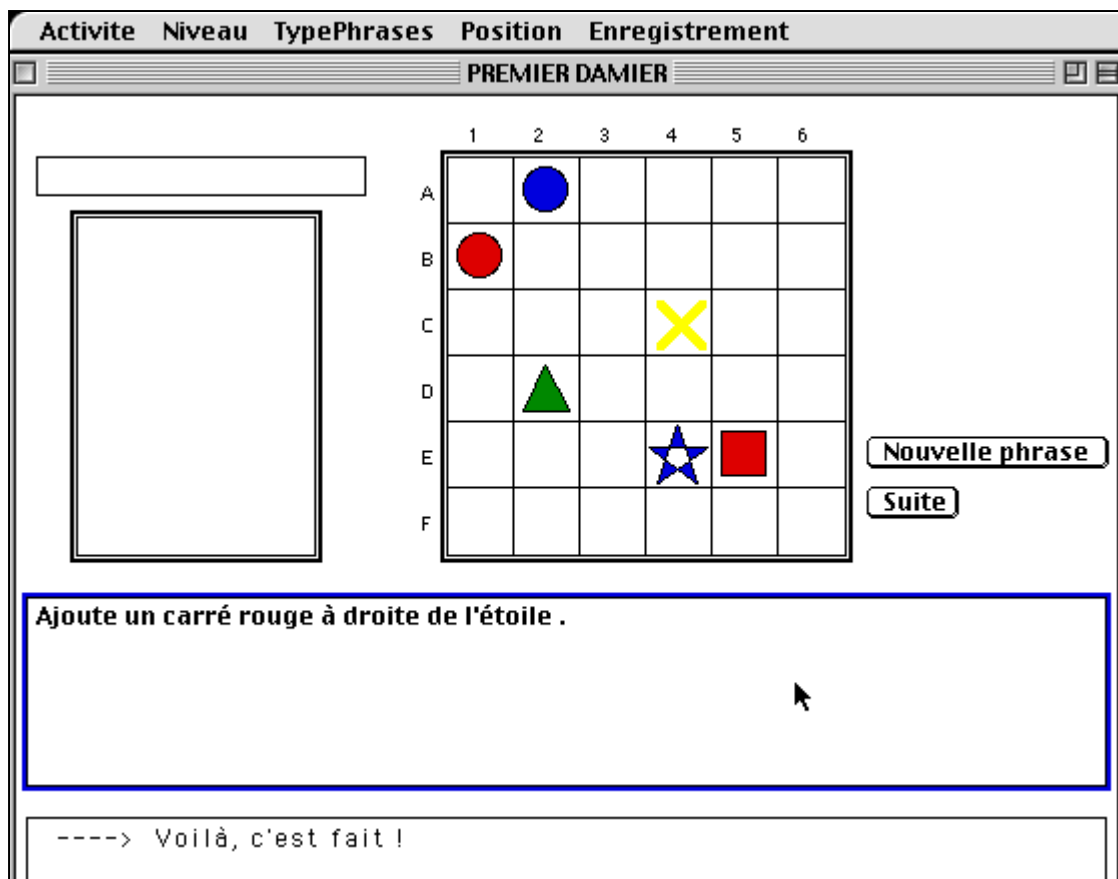


Figure 2 : un jeu en deux dimensions (2).

La résolution des expressions référentielles est effectuée pendant le traitement linguistique des phrases au niveau contextuel. Par exemple, si l'on s'en réfère à la situation illustrée par la Figure 2, les expressions définies *le rond rouge*, *l'étoile*, *le triangle qui est en D2* sont des expressions jugées contextuellement correctes, alors que *le pion bleu*, *le rond* ou *le triangle gris* ne le sont pas, les deux premières ayant plusieurs référents possibles, et la troisième aucun.



Le problème de la désignation d'un objet trouve ici une illustration abondante, il s'agit pour l'enfant de trouver une expression correcte et non-ambiguë pour désigner chaque objet sur lequel il veut agir. Il peut parler *du rond rouge*, *du rond qui se trouve en B1*, *du rond qui se trouve au dessous de la case A1*, *du pion qui est dans la colonne 1*, etc. (le mot *pion* étant dans ce jeu un mot générique) : ces expressions sont toutes correctes et désignent le même objet, que le système sait trouver. Les positions des objets sont exprimées soit de façon absolue (*en B1*), soit relative (*à droite de l'étoile*).

Plusieurs versions de ce jeu ont été développées, dont une version élémentaire destinée à un tout premier apprentissage de la lecture, dans laquelle on ne travaille que sur un ensemble prédéfini d'objets et dans laquelle le mode de composition assistée propose des groupes nominaux non décomposés. La Figure 3 donne un exemple de la façon dont se présente une session de cette version.

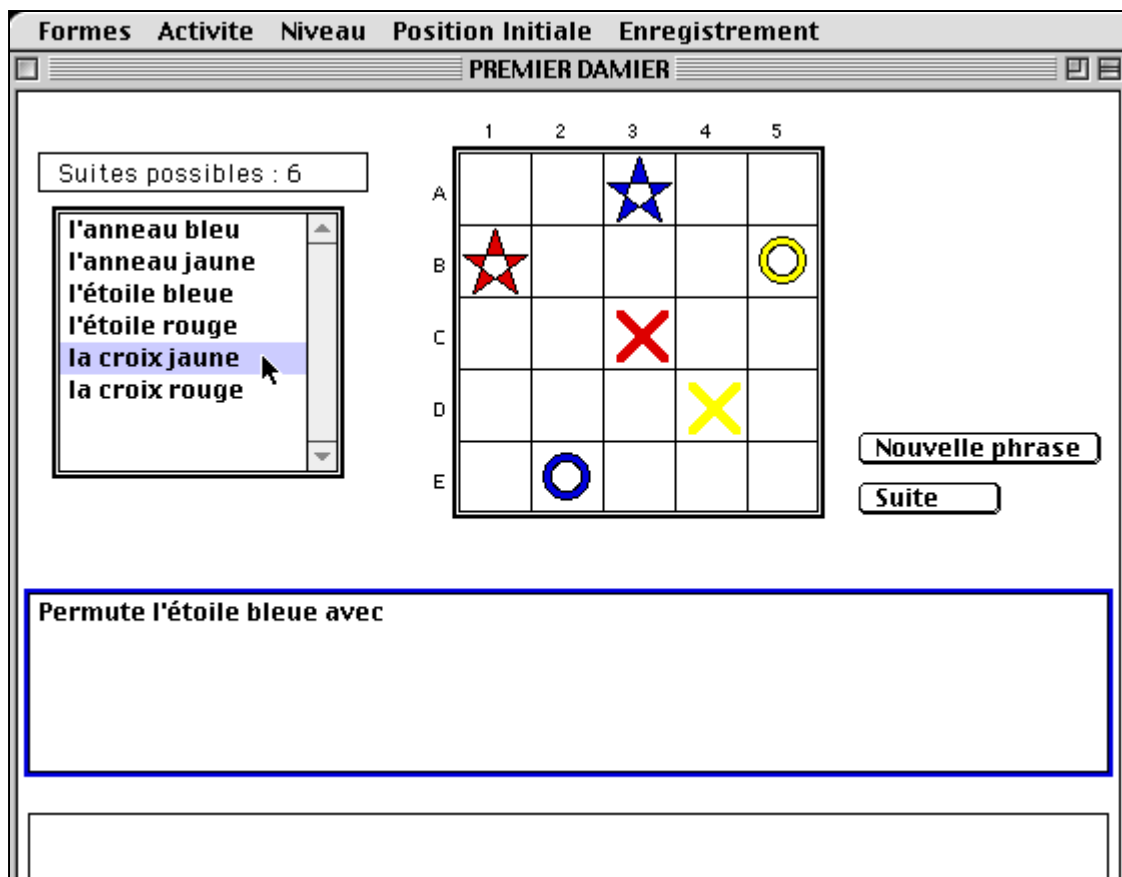


Figure 3 : la version élémentaire du jeu du damier.

Un jeu en trois dimensions : description d'une image

La deuxième activité proposée par EREL est un exercice autour d'un scénario illustré par une image : "A la pêche". L'objectif de ce jeu est d'inciter l'enfant à construire des phrases simples autour de cette scène ; ici les objets dont on parle sont figés.

L'interface se présente telle que dans [la Figure 4](#).

L'utilisateur dispose de deux modes opératoires :

- soit il compose des phrases déclaratives et le système répond si elles sont vraies ou fausses ;
- soit il pose des questions (*quelle est la couleur du foulard ?*, *où se trouve la rivière ? est-ce que le pêcheur fume ?* etc.) auxquelles le logiciel répond.



Pour chaque question, le système calcule et affiche l'ensemble des réponses possibles. Par exemple, à la question *où se trouve le panier ?* correspondent plusieurs réponses dont : *à côté du balluchon, à droite du pêcheur, au bord de la rivière, dans le pré, etc.*

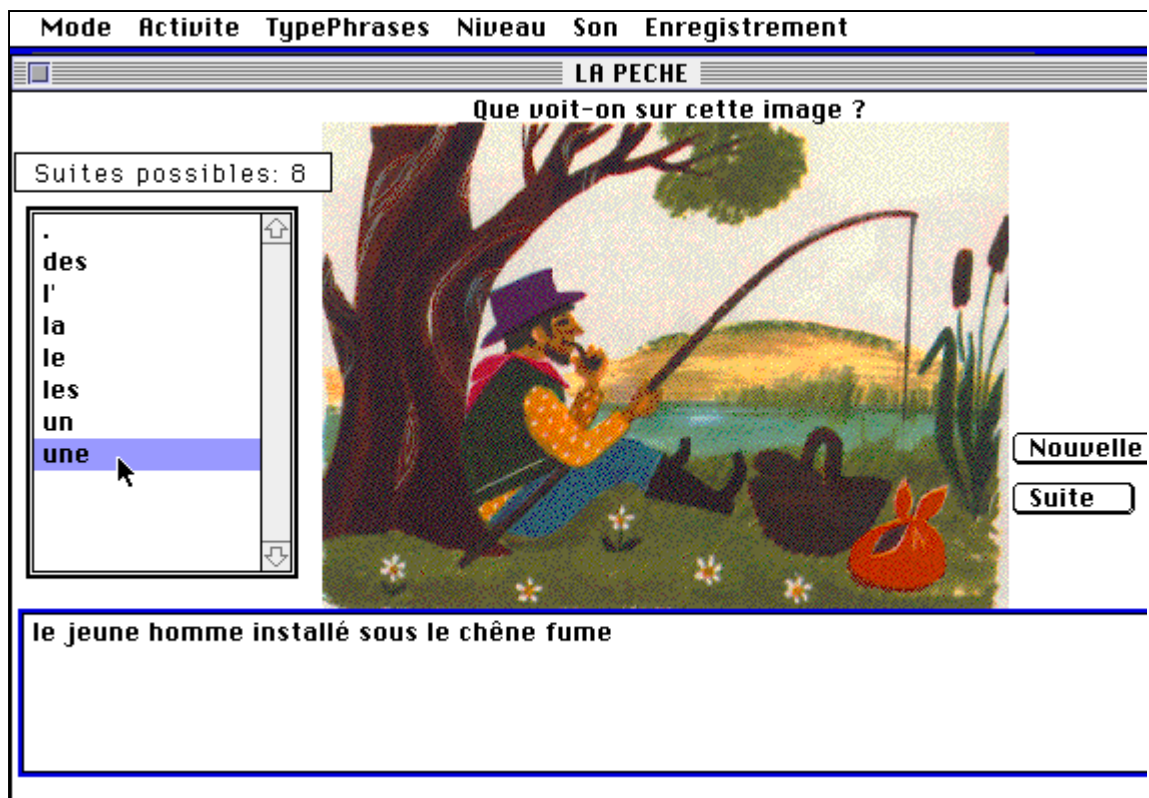


Figure 4 : description d'une image en trois dimensions.

Un jeu en une dimension : un petit train

Il s'agit ici de faire travailler l'enfant sur la composition linéaire d'un ensemble d'éléments, en l'occurrence des wagons et des locomotives (ayant toutes la même orientation).

L'interface se présente comme le montre [la Figure 5](#).

Comme dans le jeu du damier, l'utilisateur compose des requêtes qui lui permettent de supprimer, insérer, permuter, etc. des objets. Il peut également faire reculer ou avancer le train, ou changer la couleur d'un élément. Par exemple il pourra composer la suite de requêtes :

- *Ajoute une locomotive jaune devant le wagon bleu.*
- *Permute la locomotive jaune et le wagon vert.*
- *Enlève le wagon bleu.*
- *Fais reculer le petit train*
- *Supprime la locomotive noire.*
- *Quel élément se trouve devant le wagon jaune ?*
- *Enlève le 3e wagon derrière la locomotive.*



Ici encore, le problème de la désignation d'un objet est un point essentiel. Comme, dans cette activité, il n'existe pas de repère spatial absolu, les objets ne peuvent être localisés que par leurs positions relatives (*devant le wagon jaune, entre le... et la..., à gauche de..., etc.*). Les repères utilisés sont i) celui qui est donné par l'orientation des locomotives pour les expressions comme *devant* ou *derrière*, et ii) celui qui est lié à l'utilisateur pour les expressions comme *à droite*, *à gauche*. Par ailleurs, la position du train par rapport au cadre n'a un rôle que lorsque l'enfant utilise les expressions *faire avancer* ou *faire reculer* le train.

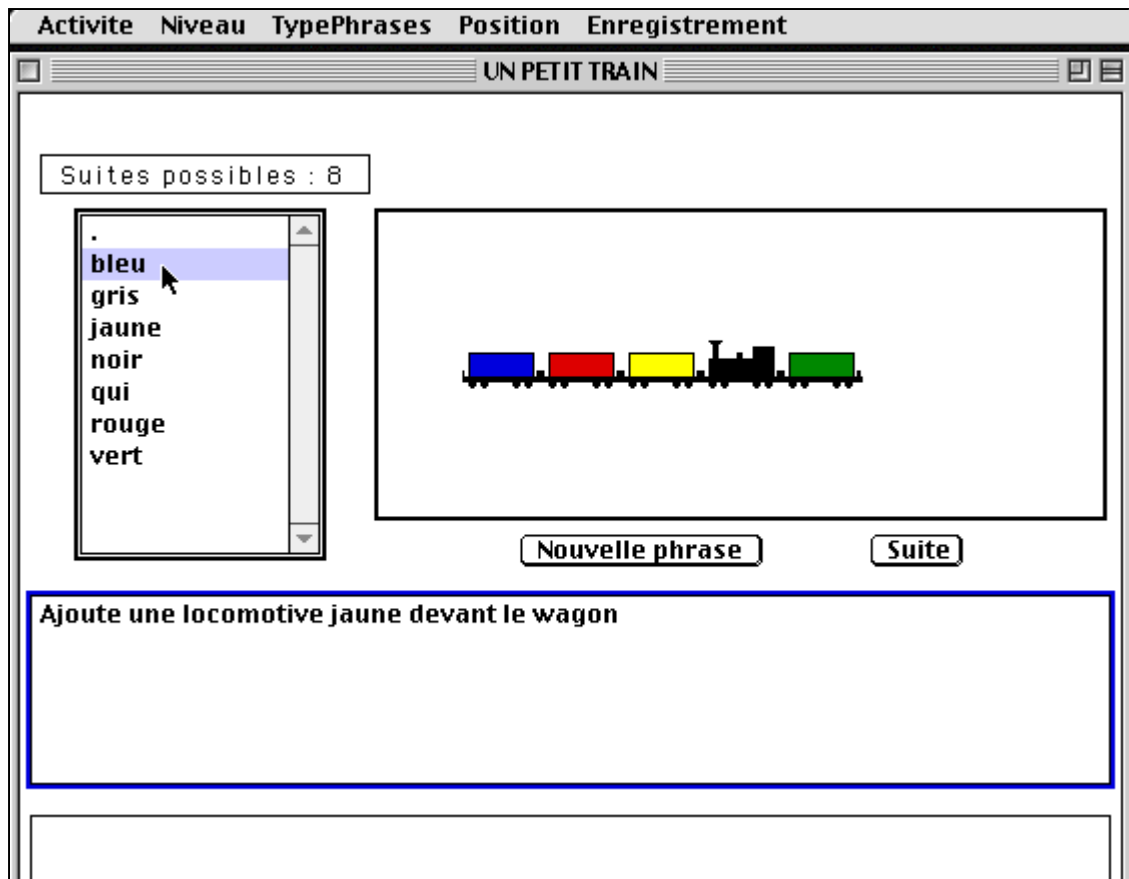


Figure 5 : un jeu en une dimension.

6. Les composants du système EREL

6.1. L'ensemble des connaissances à modéliser et gérer

Pour la réalisation logicielle de chaque exercice d'EREL, et, plus généralement, pour la réalisation de systèmes éducatifs multimédias réactifs privilégiant l'expression en langage naturel, on doit définir et représenter de façon interne :

- les objets du micro-monde considéré et leurs propriétés ;
- le sous-langage (lexique, grammaire, sémantique) qui sera utilisé dans le dialogue entre l'utilisateur et la machine ;
- les règles du jeu ;
- les modalités d'interaction ;
- les spécifications des pédagogues sur les modes d'apprentissage ;
- les algorithmes nécessaires à l'utilisation et à la gestion de tous ces éléments, pour le bon déroulement de l'activité.



Nous décrivons ci-dessous comment dans EREL nous avons classé l'ensemble de ces connaissances à l'intérieur de différents composants, dont chacun apporte à l'ensemble du système des données d'un certain domaine (données statiques, algorithmes, outils, etc.), et nous étudions certains aspects de l'ingénierie des connaissances qu'il est nécessaire d'y mettre en oeuvre. De cette classification, découle l'architecture logicielle d'EREL. Pour chaque composant nous mentionnons les modules de l'architecture classique des STI auxquels il participe.

6.2. Les composants

Le composant Définition des exercices

Il contient les spécifications relatives à chaque exercice : l'ensemble des objets du micro-monde considéré et leurs propriétés, les règles du jeu (règles légales), le langage de requêtes associé, les paramètres qui indiquent le niveau de difficulté de l'exercice, etc.

Le contenu de ce composant provient des spécifications des pédagogues. Dans l'architecture classique des STI, ces connaissances interviennent dans le module pédagogique et le module interface.

Le composant Modèle de l'apprenant

Comme dans l'architecture classique des STI, il permet de définir un profil utilisateur grâce à des informations sur l'élève et sur l'historique de son travail. Dans la version actuelle d'EREL, il est assez rudimentaire : il contient un historique des requêtes composées par l'utilisateur, et nous n'avons encore défini aucun processus pour l'exploitation de ces données.

Le composant Base de données

Il contient :

- pour chaque exercice, la représentation interne du micro-monde d'objets sur lequel on travaille et de leurs propriétés, ces éléments constituant une petite base de données ;
- les processus de raisonnement (règles d'inférence et de calcul), pour la recherche d'informations sur le contenu de cette base et en particulier pour la production de réponses aux requêtes de l'utilisateur ;
- les processus de mise à jour de la base : ajout, modification, suppression de données.

Dans l'architecture classique des STI, ces connaissances interviennent dans le module expert.

Il a donc été nécessaire de définir une représentation interne des objets composant chaque scénario. Dans cette petite base de données, chaque objet est décrit par sa nature conceptuelle, sa couleur et sa place.



La modélisation de la sémantique de l'espace a été faite différemment suivant les jeux. Pour le damier, la localisation d'un objet se fait par ses coordonnées, et les positions relatives des objets sont exprimées à partir des relations cardinales : *à droite, à gauche, au-dessus, au-dessous*. Dans la description de l'image "A la pêche", on a défini un ensemble de "zones" en trois dimensions qui structure et recouvre l'espace décrit par l'image, suivant une approche qui s'inspire des travaux sur la sémantique de l'espace décrits par (Vieu, 1991). Les relations spatiales entre les objets en découlent, et, pour répondre à une question sur la localisation d'un objet, le système est capable de donner dynamiquement toutes les relations possibles entre la zone concernée et les autres zones. Pour le jeu du petit train, on a modélisé la suite des éléments (wagons ou locomotives) qui composent le train, ainsi que la position de l'élément de tête par rapport au cadre, qui est la seule position modélisée de façon absolue. Les positions des autres éléments en découlent, et les actions sur ces éléments s'effectuent par ajout, insertion, ou suppression d'élément, ou par déplacement de l'ensemble dans un sens ou dans l'autre.

Le composant TALN

Pour le traitement des phrases écrites, on utilise le système ILLICO que nous avons présenté dans la partie 4. Il contient i) les formalismes pour la définition des sous-langages du langage naturel associés aux diverses activités, ii) les algorithmes pour la compréhension et la génération de textes en langage naturel.

Par ailleurs, le composant TALN contient également un module simple de synthèse de la parole que nous avons développé pour EREL et qui prend en entrée les phrases complètes et correctes composées par l'utilisateur. Ce module permet de produire une réalisation de la phrase dotée d'un contour prosodique naturel, à partir d'une phonétisation lexicale, de l'analyse lexicale et syntaxique de la phrase écrite effectuée par ILLICO, et d'un ensemble de règles prosodiques (Guizol & Godbert, 1999). La phrase est tout d'abord traduite en la liste des différents phonèmes qui la constitue, avec leur durée et la valeur du pitch ; ensuite, cette liste est traitée par le synthétiseur MBROLA (Dutoit et al., 96) qui dispose des structures formantiques associées à chaque phonème pour produire finalement la phrase réalisée.

Dans l'architecture classique des STI, les connaissances relatives au TALN interviennent en général dans

le module expert et le module interface.

Le composant Graphique

Il contient :

- les outils classiques et méthodes utilisés dans une IHM graphique, pour les entrées-sorties : listes et menus déroulants, boutons, rectangles d'édition, etc.
- une bibliothèque de fonctions graphiques pour l'illustration du micro-monde à l'écran : dessins divers, représentation des déplacements d'objets, modification de leur couleur, etc.

Dans l'architecture classique des STI, ces connaissances interviennent dans le module interface.

Le composant Multimédia et Multimodalité

Ce composant contient les fonctions nécessaires à la définition et la gestion d'une interaction multimédia et multimodale entre l'utilisateur et la machine. Dans l'architecture classique des STI, ces connaissances interviennent dans le module interface.



Si l'on veut permettre à l'utilisateur de formuler ses requêtes en utilisant à chaque instant le canal de communication qui lui convient le mieux, il faut par exemple pouvoir utiliser en parallèle, pour les entrées, les couples <clavier, langage naturel> et <souris, éléments graphiques> (et on pourrait envisager d'ajouter le couple <micro, parole>). De façon générale, les systèmes multimédias et multimodaux intelligents ont à sélectionner des expressions appropriées et de différents types (textes, graphique, etc.), à les coordonner et à les fusionner pour obtenir une représentation explicite du contenu des expressions multimodales.

En ce qui concerne les entrées, l'interface d'EREL est définie de telle manière que les différents médias soient organisés de façon interactive, et non séquentielle, pour qu'à tout instant l'utilisateur puisse s'exprimer par le canal de son choix : le clavier (rectangles d'édition) ou le canal graphique (sélection de mots, d'objets ou de points de l'écran à l'aide de la souris).

C'est dire que le système sait reconnaître une même entité à partir de descriptions différentes, et donc calculer si deux caractérisations sont référentiellement équivalentes.

Par exemple, dans le jeu du damier, l'utilisateur peut utiliser les modalités <clavier, langage naturel> et <souris, éléments graphiques>, ou mixer les deux, pour formuler ses phrases ; pour désigner une case, il pourra par exemple écrire *la case C3* ou bien la sélectionner par un clic, le système se chargeant à partir du clic de générer l'expression *la case C3*.

Pour effectuer la fusion des informations issues de plusieurs modes d'entrée, le système recherche dans les composants TALN et Base de données les données pertinentes (sémantique des mots du lexique, modélisation interne des objets, etc.) qui lui permettent de calculer l'information résultante.

En ce qui concerne les sorties, c'est-à-dire la production de réponses à l'utilisateur, différents modes sont coordonnés : l'activité est illustrée graphiquement à l'écran, les phrases y sont écrites et elles sont synthétisées vocalement. Par ailleurs, dans le jeu du damier, la sémantique de chaque phrase produite est également (si l'utilisateur le désire) représentée de façon schématique. Ajoutons que nous avons également quelques documents sonores (bruits de la nature dans le jeu "A la pêche", etc.) qui complètent l'illustration des activités. De façon générale, il est nécessaire que la présentation résultante des différentes sorties soit cohérente. Par exemple la réponse à une requête exprimant le désir de déplacer un objet sera constituée de la production coordonnée d'un message écrit et d'un message oral, d'actions graphiques (mise à jour de la représentation graphique) et éventuellement de l'audition d'un document sonore. On peut parler dans ce cas de fusion de la réaction du système, de la même façon que l'on parle de fusion dans la fonction de restitution d'informations dans une interface multimédia. Bien sûr, le système effectue également, simultanément, des actions internes à la base de données pour la mise à jour des données, totalement transparentes pour l'utilisateur.



Le composant Dialogue

Il contient les règles cognitives et pragmatiques et les techniques qui permettent de gérer l'enchaînement des actes de communication. Ce composant est en particulier chargé de la production des réponses aux requêtes et de l'aide à l'utilisateur. Dans l'architecture classique des STI, ces connaissances interviennent dans les modules pédagogique, expert et interface.

EREL est un système centré-utilisateur, comme beaucoup d'environnements de STI : il a été conçu et se présente comme un système ludique "au service de l'enfant utilisateur" dans lequel le système obéit aux requêtes que l'utilisateur formule, et lui apporte s'il le désire une assistance pour formuler ses requêtes.

L'initiative est donc laissée en général à l'utilisateur. Cependant, au cours du dialogue, elle est prise par le système pour tout ce qui touche à l'assistance à l'utilisateur : lorsque l'utilisateur demande une aide, le système d'aide contextuelle (aide calculée en temps réel à partir de l'état courant du système) "prend en charge" l'utilisateur, ce qui induit un renversement du paradigme de communication comme le décrivent (Palanque et al., 1994).

La médiation d'autrui est un facteur important dans les processus d'apprentissage (Bruner, 1983), et le traitement des erreurs et l'assistance à l'utilisateur sont des points essentiels dans un système éducatif. Dans un système comme EREL, ils requièrent de mettre en oeuvre une ingénierie qui utilise des connaissances de tous les composants du système.

En ce qui concerne la formulation des requêtes en langage naturel, le système aide l'utilisateur, s'il le désire, à produire des phrases correctes à tous points de vue (bien formées aux niveaux lexical, syntaxique, sémantique, contextuel, pragmatique). Cela correspond à la composition en mode assisté, mise en oeuvre grâce aux outils d'ILLICO décrits dans la partie 4. Si par contre l'utilisateur compose son texte en mode libre, le système n'analyse la phrase que lorsqu'elle est terminée et que l'utilisateur en demande une réponse. Le système détecte alors les éventuelles erreurs que contient le texte, les signale à l'utilisateur et propose son aide (le mode assisté).

Lorsqu'une requête est correcte aux niveaux lexical, syntaxique et sémantique, le système en vérifie la correction contextuelle et, dans le cas favorable, donne sa réponse. Dans l'autre cas, il donne un message d'explication sur l'erreur que contient la phrase de l'utilisateur : au niveau contextuel, une requête peut être incorrecte soit parce qu'elle contient une expression référentielle qui ne peut pas être résolue (elle n'admet pas un référent unique), soit parce qu'elle demande l'exécution d'une action qu'il n'est pas possible de réaliser car elle viole les règles du jeu (par exemple tout simplement le débordement des limites du cadre dans le jeu du train ou du damier, comme *Pose le triangle au-dessus du rond bleu* dans la situation de la [Figure 1](#)).

Le composant Superviseur

Il assure le contrôle et l'orchestration de l'ensemble des composants précédents et de leurs interactions. Dans l'architecture classique des STI, ces connaissances interviennent dans le module interface.

L'ingénierie des connaissances mise en oeuvre dans EREL peut être qualifiée de "multi-ingénierie", chacun des composants fournissant le traitement des connaissances dont il est spécialiste, et le superviseur coordonnant le tout : l'organisation générale du système est telle que les interactions et les processus de coopération entre les composants transitent par le superviseur, qui envoie des données ou des tâches aux différents composants, recueille leurs réponses, et prend des décisions pour la suite du déroulement du jeu.



7. Validation du système

Pour pouvoir continuer notre travail autour d'EREL de façon réaliste, le retour des utilisateurs nous est bien sûr indispensable. EREL est utilisé depuis environ deux ans sur trois sites de la région de Marseille : un cabinet médical, un hôpital de jour et un centre de psychothérapie infantile. Quelques médecins ou

éducateurs spécialisés utilisent EREL avec des enfants psychotiques ou autistes, ou avec des enfants en situation d'échec scolaire. Ils pensent que ce logiciel est un bon outil pédagogique, mais d'une part il est difficile d'évaluer les éventuels progrès des enfants qu'ils font travailler avec EREL, et d'autre part nous n'avons défini aucune méthode précise pour l'évaluation et la validation du système. En fait EREL est utilisé comme un outil, parmi d'autres, que les éducateurs ont à leur disposition pour faire travailler des enfants en âge d'apprentissage du langage. Leur retour nous est précieux, en particulier pour tout ce qui concerne l'amélioration de l'interaction, ils nous ont en particulier suggéré d'ajouter quelques fonctionnalités à la première version du logiciel que nous leur avons donnée.

L'un de ces éducateurs, qui utilise EREL avec des groupes de deux ou trois enfants en situation d'échec scolaire, a lui même ajouté un nouveau média à EREL, en l'occurrence le support papier, qui permet de représenter sous une forme encore différente et sur un support externe ce qui est affiché graphiquement à l'écran. Il s'est surtout intéressé au jeu du damier et aux possibilités qu'il offre au point de vue cognitif et fait travailler les enfants sur l'ensemble des possibilités de création de pions à partir des formes et des couleurs proposées par le jeu.

Il faut cependant reconnaître que les retours des éducateurs sont assez peu nombreux, et nous avons pour l'année qui vient un projet pour élargir le cercle des utilisateurs du système, en collaboration avec un groupe d'instituteurs intéressés par l'utilisation de logiciels éducatifs. Les études et expériences piagétienne sur le développement des fonctions cognitives et langagières chez l'enfant étant une source abondante d'inspiration, nous envisageons l'intégration dans EREL de nouvelles activités, en particulier des exercices de dénombrement. Tous les prolongements du projet seront fonction de nos discussions avec les utilisateurs.

Nous avons également le projet de développer EREL dans d'autres langues, ce qui peut être fait à un coût limité grâce à la modularité de la modélisation des connaissances linguistiques dans ILLICO. La plus grande partie des composants du système EREL seront identiques d'une langue à l'autre, et les seuls modules qui doivent être réécrits dans chaque langue sont le lexique, la grammaire, et le module qui permet de passer de la forme syntaxique d'une unité linguistique à sa représentation sémantique. Actuellement une version simplifiée du jeu "Premier damier" existe en anglais.



8. Conclusion et perspectives

Nous nous sommes intéressés dans cet article à la réalisation logicielle d'exercices de type piagétien conçus pour faciliter chez l'enfant l'apprentissage d'un langage simple associé à des concepts et opérations de type logico-mathématique, dans un environnement multimédia réactif qui privilégie l'expression en langage naturel écrit.

Le logiciel éducatif EREL que nous avons développé dans cette optique ces dernières années est une modeste réalisation d'un tel système. Son développement a montré, entre autres choses, que le système générique ILLICO apportait un ensemble d'outils très appréciables pour le TALN, et EREL tire un grand profit des possibilités offertes par ILLICO.

Nous avons montré aussi que le développement de tels systèmes est loin d'être simple -même si chaque activité proposée ne travaille que sur un micro-monde et un sous-langage très restreints- car il est nécessaire d'y prendre en considération un grand nombre de connaissances de type divers et de mettre en oeuvre une multi-ingénierie pour le traitement de ces connaissances, en particulier si l'on veut parvenir à un dialogue LN+MM aussi naturel que possible entre l'enfant et la machine.

Nos recherches actuelles s'intéressent :

- A la spécification formelle du dialogue multimodal et de l'aide à l'utilisateur dans un système de ce type. Les processus d'aide contextuelle doivent prendre en considération l'ensemble des données constituant l'état du système à chaque instant. Nous pensons utiliser, pour la formalisation du dialogue et de l'aide contextuelle à l'utilisateur, un modèle dérivé des réseaux de Petri, formalisme qui convient bien et qui a déjà été utilisé à de nombreuses reprises dans des systèmes comportant une interface utilisateur complexe (voir par exemple (Bastide et al., 1999)) ou dans des systèmes multi-agents.
- Au contenu des messages d'explication qu'il est nécessaire de donner à l'utilisateur lorsqu'il a

commis une erreur, c'est-à-dire à l'aide que l'on peut apporter à l'utilisateur au niveau cognitif. Dans la version actuelle d'EREL, le contenu des messages d'erreurs est assez succinct : on donne une explication simple sur le type de l'erreur (*la phrase est ambiguë*, ou *l'action demandée est impossible*, ou *interdite*, etc.). Nous voudrions améliorer cette aide en proposant des messages explicatifs détaillés si l'utilisateur le désire (avec un paramétrage du système qui permet de choisir entre différents niveaux d'explication).

La notion d'explication a fait l'objet de nombreuses études dans le domaine des systèmes experts ou systèmes à base de connaissances (voir par exemple (Clancey, 1983)). La production de messages explicatifs en langage naturel est un problème difficile, comme le souligne par exemple (Charnay, 1996), car il requiert de mettre en oeuvre des stratégies pour gérer le dialogue, l'explication, l'argumentation, et pour produire des énoncés en langage naturel (Maybury, 1998 ; Cawsey, 1998). Pour un système éducatif, il faut y ajouter l'exigence de pertinence pédagogique : "Il faut aider l'utilisateur, oui, mais pas trop !" nous disent les éducateurs..., l'idéal étant d'aider naturellement l'enfant à trouver lui-même son erreur.



Remerciements

Le projet EREL a été partiellement financé par le Conseil Général des Bouches-du-Rhône. Nous sommes reconnaissants à H. Garreta, J. Guizol, R. Pasero et M. Rolbert pour l'aide qu'ils nous ont apportée pour le développement d'EREL.

Références

Bastide, R., Palanque, P., Sy O., Le, D.H & Navarre, D. (1999). "Petri-Net Based Behavioural Specification of CORBA Systems". *Proc. of the 20th International Conference on Applications and Theory of Petri-Nets, ATPN'99*. Williamsburg, VA, EU-A.

Bruner, J. (1983). *Le développement de l'enfant, Savoir faire savoir dire*. Paris : PUF.

Cawsey, A. (1998). "Planning interactive explanations". In *Intelligent User Interfaces*, Maybury, M.T. & Wahlster, W. (dirs). San Francisco : Morgan Kaufmann Publishers. pp 404-419.

Charnay, L. (1996). "Dialogue et explication dans les systèmes à base de connaissances : un modèle pour l'acte de dialogue explicatif". *Actes de la conférence Rencontres des Jeunes Chercheurs en Intelligence Artificielle (RJCIA'96)*, Toulouse. p 259.

Clancey, W.J. (1983). "The epistemology of a ruled-based expert system, a framework for explanation". *Artificial Intelligence*. pp 215-253.

Clancey, W.J. (1993). "Guidon-Manage Revisited : A Socio-Technical Systems Approach". *International Journal of Artificial Intelligence in Education*. vol. 4, 1. pp 5-34.

Cohen P.R., Johnson, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., & Clow, J. (1998). "Multimodal Interaction for Distributed Interactive Simulation". In *Intelligent User Interfaces*, Maybury, M.T. & Wahlster, W. (dirs). San Francisco : Morgan Kaufmann Publishers. pp 562-569.

Donellan, K. (1966). "Reference and definite description". *Philosophical Review*, 75. pp 281-304.

Douglas Sarah, A. (1995). "LingWorlds: An Intelligent Object-Oriented Environment for Second-Language Tutoring". In *Intelligent Language Tutors*, Holland, V.M., Kaplan, J.D. & Sams, M.R. (dirs). Mahwah : LEA Publishers. pp 201-220.

Dowsing, R.D. & Long, S. (1999). "An Evaluation of the Impact of AI Techniques on Computerised Assessment of Word Processing Skills". *Artificial Intelligence in Education*, Lajoie, S. & Vivet, M. (dirs). Amsterdam : IOS Press Ohmsha. pp 379-386.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F. & Van Der Vrecken, O., (1996). "The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes ". *Proc. of International Conference of Spoken Language Processing (ICSLP'96)*, Philadelphie, vol. 3. pp 1393-1396.

Guizol, J. & Godbert, E. (1999). "Intégration de la synthèse vocale dans un logiciel éducatif multimédia". *Journées Internationales de Linguistique Appliquée, JILA '99*, Nice. pp 121-125.

Humblot, L. (1998). "Acquisition des marqueurs des systèmes de référence spatiale". *Actes de la Conférence de l'Association pour la Recherche Cognitive (ARC 98)*, Paris. pp 28-35.

Joab, M. (1990). *Modélisation d'un dialogue pédagogique en langage naturel*. Thèse de doctorat en informatique. Paris : Université Paris 6.

Khuwaja, R., Desmarais, M. & Cheng, R. (1996). "Intelligent Guide : Combining User Knowledge Assessment with Pedagogical Guidance". *Intelligent Tutoring Systems, Proc. of ITS'96*. Lecture notes in computer science, Springer-Verlag. pp 225-233.

Matthews, C., (1992). "Going AI. Foundations of ICALL". *Computer-Assisted Language Learning (CALL)*, vol. 5, 1-2. pp 13-31.

Maybury, M.T. (1998). "Planning Multimedia Explanations Using Communicative Acts". In *Intelligent User Interfaces*, Maybury, M.T. & Wahlster, W. (dirs). San Francisco : Morgan Kaufmann Publishers. pp 99-106.

Milhaud, G. (1994). *Un environnement pour la composition de phrases assistée*. Thèse de doctorat en informatique. Marseille : Université Aix-Marseille II.



Murray, J.H. (1995). "Lessons Learned from the Athena Learning Project". In *Intelligent Language Tutors*, Holland, V.M., Kaplan, J.D. & Sams M.R. (dirs). Mahwah : LEA Publishers. pp 243-256.

Nigay, L. & Coutaz, J. (1996). "Espaces conceptuels pour l'interaction multimédia et multimodale". *Technique et science informatiques (TSI)*, vol. 15, 9. pp 1195-1225.

Palanque, P., Salbe, D. & Bastide, R. (1994). "Gestion automatique de l'aide contextuelle multimédia d'une IHM par exécution de sa spécification formelle". *Actes de la Conférence Interaction Homme-Machine (IHM'94)*.

Pasero, R. & Sabatier, P. (1994). "ILLICO for Natural Language Interface". *Language Engineering Convention (LEC)*.

Pasero, R. & Sabatier, P. (1995). "Guided Sentences Composition : Some problems, solutions, and applications". *5th International Conference on Natural Language Understanding and Logic Programming (NLULP)*, Lisbonne. pp 33-44.

Piaget, J. & Inhelder, B. (1966). *La psychologie de l'enfant*. Paris : PUF.

Vieu, L. (1991). *Sémantique des relations spatiales et inférences spatio-temporelles : Une contribution à l'étude des structures formelles de l'espace en Langage Naturel*. Thèse de doctorat en informatique. Toulouse : Université Paul Sabatier.

Wahlster, W., André, E., Finkler, W., Profitlich, H.J. & Rist, T. (1993). "Plan-based integration of Natural Language and Graphics Generation". *Artificial Intelligence*, 63. pp 387-427.

À propos de l'auteur

Élisabeth GODBERT, après des études doctorales au GRTC (Marseille), dans le domaine de la Représentation des connaissances, a rejoint l'équipe "Traitement automatique du langage naturel" du LIM en 1991. Depuis quelques années, elle s'intéresse plus particulièrement à la conception et au développement de systèmes éducatifs pour l'apprentissage du langage. Elle est actuellement Maître de conférences en informatique à la Faculté des Sciences de Luminy, à Marseille.

Courriel : godbert@lim.univ-mrs.fr

Adresse: Laboratoire d'Informatique de Marseille (LIM), Université de la Méditerranée et CNRS, 163 avenue de Luminy, Case 901 13288 Marseille Cedex 9 France.



[ALSIC](#) | [Sommaire](#) | [Consignes aux auteurs](#) | [Comité de rédaction](#) | [Inscription](#)

© *Apprentissage des Langues et Systèmes d'Information et de Communication*, 15 décembre 2000