

ESSAI SUR L'ANALYSE DE RÉPONSES AUX QUESTIONS OUVERTES DANS UN PROGRAMME MULTIMÉDIA

Jeannine GERBAULT, Lise BROSSARD*, Henri PORTINE

Université de Bordeaux 3
gerbault@montaigne.u-bordeaux.fr
brossard@montaigne.u-bordeaux.fr
portine@montaigne.u-bordeaux.fr

PRÉSENTATION DE LA PROBLÉMATIQUE

Tout programme multimédia pédagogique pour l'apprentissage-acquisition des langues étrangères suppose que l'on teste la compréhension de l'utilisateur. Le travail à réaliser peut être de plusieurs types : énigme logique, tâche à accomplir (choisir des objets dans un ensemble selon une certaine finalité, par exemple), résumé, question, etc. Dans le cas d'une question, il peut s'agir d'une question fermée, ou QCM, ou d'une question ouverte.

La plupart des programmes multimédias proposent des questions sous la forme de QCM. L'analyse de réponses en est grandement facilitée. Mais, d'une part, les QCM induisent en partie la réponse. D'autre part, ils ne fournissent à l'utilisateur aucune possibilité de s'exprimer par lui-même. Par conséquent, ils ne sont pas propices à une évaluation formative.

Nous nous sommes donc intéressés au problème de l'analyse des réponses à des questions ouvertes. Une bonne réponse est : syntaxiquement correcte, morphologiquement correcte et correctement orthographiée ou prononcée, sémantiquement appropriée, pragmatiquement adaptée.

Chacun de ces critères est indépendant des trois autres. Nous ne reprendrons pas ici les critères de correction syntaxique et morphologique. Une réponse est *sémantiquement appropriée* (*semantically appropriate*) - et non plus correcte - lorsqu'elle répond effectivement à la question posée. La notion d'appropriation est très différente de celle de correction. La correction est en général binaire. En revanche, l'*appropriation sémantique* permet bien souvent un plus grand nombre de variantes (cf. la problématique de la paraphrase en linguistique).

Une réponse est *pragmatiquement adaptée* (*pragmatically relevant*) lorsqu'elle répond effectivement à la question posée en respectant la situation dans

* DEFLE. Université de Bordeaux 3

laquelle elle s'inscrit. On a un bon indice de la différence entre appropriation sémantique et adaptation pragmatique dans la réaction professorale traditionnelle « Faites une phrase », qui montre que l'adaptation pragmatique est parfois jugée peu importante en situation d'apprentissage.

Considérons une question ouverte portant sur un document ou un extrait de document préalable, noté dorénavant D. On peut distinguer cinq types de réponses selon la forme de la question. Des exemples de chacun de ces types sont présentés Tableau 1.

La réponse peut consister en une identification de la situation dans laquelle D a été produit ou d'un élément de cette situation (Type 1). Elle peut consister en la reprise pure et simple d'une formule de D, avec ajout d'un contexte rendant la réponse syntaxiquement correcte et pragmatiquement adaptée (Type 2). Elle peut aussi être la reformulation d'une formule de D, avec ajout d'un contexte rendant la réponse syntaxiquement correcte et pragmatiquement adaptée (Type 3). Ce type et les deux suivants sont bien évidemment les plus complexes, mais aussi les plus intéressants puisqu'ils débouchent sur une simulation du comportement humain. En effet, un locuteur, langue maternelle, ne reprend que rarement une formule telle qu'elle a été produite. La réponse peut aussi consister en une synthèse élaborée à partir de plusieurs formules de D, avec ajout d'un contexte rendant la réponse syntaxiquement correcte et pragmatiquement adaptée (Type 4). Enfin, la réponse de Type 5 consiste en la formulation d'un raisonnement effectué dans D, c'est-à-dire de liens logiques non exprimés explicitement dans D.

Plus on avance dans la liste des types des réponses attendues, plus on se rapproche de la problématique de l'analyse des textes. Or un texte est à la fois une succession de phrases¹ et un réseau sémantique. Si l'analyse syntaxique conserve alors son importance, l'analyse sémantique et pragmatique doit la compléter. C'est dans ce contexte que s'inscrit notre projet. Considérer le texte comme un réseau sémantique revient à définir les contraintes pesant sur un tel réseau (considéré au niveau informatique comme un graphe) et la technique de construction de ce réseau (quelles opérations ? appliquées comment ?). Or l'exploration d'un texte et sa compréhension supposent toujours un guidage contrôlé par un schème conceptuel lié à des attentes (*conceptually driven processing*) car on n'est jamais vierge devant un texte. Il convient donc de définir le type de schème conceptuel pouvant être à l'œuvre dans telle situation².

DESCRIPTION DU PROTOCOLE D'EXPÉRIMENTATION

Notre objectif est de construire et d'analyser un corpus de réponses à des questions ouvertes afin de déterminer des stratégies d'analyse des réponses. On sait qu'il est vain pour l'instant d'espérer mettre sur pied une analyse automatique brute.

1 À l'écrit, bien évidemment. À l'oral, les choses sont un peu plus compliquées.

2 Nous dirons que le texte à construire (en réponse à une question) ou à reconstruire (en compréhension de la partie du document D sur laquelle porte la question posée) doit être considéré comme « un réseau de concepts ».

C'est pourquoi la tendance actuelle est à cadrer les analyses au moyen d'études sur les attentes et les types de réponses plausibles dans un contexte donné. Les stratégies d'analyse de réponses devront donc tenir compte de ces attentes (à la fois des attentes du concepteur du programme et des attentes de l'utilisateur³) et offrir des investigations cadrées des points de vue sémantique et pragmatique.

Pour ce faire, nous avons monté un protocole d'expérimentation assez lourd dont nous donnons la description. Chaque passation de l'expérimentation se fait individuellement, puisqu'il n'est pas nécessaire que chaque sujet fasse la même chose en même temps. Le sujet subissant l'expérimentation est installé à un poste de travail qui comporte une télévision avec magnétoscope, un ordinateur avec traitement de texte, un magnétophone et un caméscope. Afin de bien faire visualiser le poste de travail, nous le représentons dans la Figure 1.

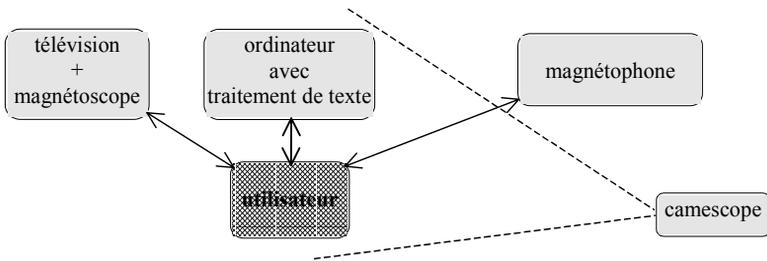


Figure 1.

Dans la Figure 1, les flèches bi-orientées indiquent que l'utilisateur est maître — au moins à certains moments de l'expérimentation — du fonctionnement des appareils auxquels il est relié (à une exception près : le couple télévision + magnétophone est au début manipulé par l'expérimentateur). La position du caméscope montre son rôle : nous donner une image nette du comportement de l'utilisateur.

Nous avons choisi un document vidéo. Pour des raisons évidentes (éviter de fabriquer un matériel *ad hoc*), nous ne voulions pas le fabriquer. Après concertation, nous avons choisi le dialogue « Chez le médecin » de *Je vous ai compris, I*^{®4}. Ce choix a été guidé essentiellement par des considérations de niveau d'apprentissage. Le dialogue en question a été utilisé comme document déclencheur suivi de questions ouvertes fabriquées par nos soins. Le titre du dialogue, trop explicite pour notre expérimentation, a été retiré du document visionné par les sujets.

Le déroulement s'opère, pour chaque utilisateur, en trois phases. Les deux premières se font avec l'expérimentateur qui n'intervient que pour la manipulation des appareils. Pour la troisième phase, l'expérimentateur est en retrait.

Phase 1. L'expérimentateur fait visionner deux fois consécutives le document vidéo par l'utilisateur. Puis il éteint la télévision.

Phase 2. L'expérimentateur demande à l'utilisateur de répondre aux questions écrites qui figurent sur l'écran du traitement de texte. Pour chaque question, on

3 Ces attentes définissent « un profil d'apprenant ».

4 Cédérom d'apprentissage du français langue étrangère, co-production Université de Lille 3 - Neuroconcept.

demande à l'utilisateur d'enregistrer d'abord sa réponse orale sur le magnétophone, puis de frapper au clavier de l'ordinateur sa réponse écrite.

Phase 3. Lorsqu'il a répondu à toutes les questions, on demande à l'utilisateur de visionner à nouveau tout ou partie du document vidéo, pour répondre une deuxième fois aux mêmes questions, mais cette fois seulement à l'écrit. Il peut manipuler le magnétoscope à sa guise, y compris pendant qu'il répond aux questions.

La différence entre les phases 2 et 3 est importante. Elle permet de distinguer les réponses fondées sur la mémoire des dialogues et les réponses avec consultation du dialogue. Toute une série de questions en découle : y a-t-il recherche de reproduction fidèle lors de la phase 3 ? la mémoire constitue-t-elle un grand obstacle dans le déroulement de la phase 2 ? le revisionnement du dialogue dans le cadre de la phase 3 nuit-il à la faculté de synthèse de l'utilisateur ? etc.

ANALYSE DES DONNÉES

Rappelons que nos données sont de trois types :

- document vidéo, chaque sujet a été filmé pendant son travail (Phases 1, 2, 3 ; voir Figure 2),
- document audio, enregistrement des réponses orales aux questions ouvertes (Phase 2),
- document écrit, les réponses aux questions ouvertes ont été tapées sur le clavier de l'ordinateur en deux étapes de l'expérimentation (Phases 2 et 3).

Notre objectif n'est *pas* de résoudre un problème d'informatique, tâche qui incombe aux spécialistes de cette discipline. Mais il est de contribuer à l'identification des paramètres qui sont susceptibles de guider l'élaboration de programmes informatiques pédagogiques dans lesquels il sera possible d'évaluer les réponses à des questions ouvertes des utilisateurs (les apprenants).

Dans cette perspective, la recherche a plusieurs objectifs :

1. Les réponses aux questions ouvertes ne pouvant être analysées automatiquement indépendamment de leur contexte, participer à l'élucidation des directives d'analyse permettant de les analyser automatiquement.
2. Élaborer des contenus de formation pour les tuteurs des centres de ressources, leur permettant de servir d'interface entre le programme multimédia en machine (ou plutôt l'analyse de réponse qu'il fournit) et l'utilisateur du programme (l'apprenant).
3. Étudier le comportement de recherche d'informations par un apprenant en vue de la construction de la réponse à une question ouverte dans un programme multimédia.

Nous pouvons associer des items obtenus lors de notre première phase d'expérimentation (rappelons que d'autres phases sont en cours ou à venir) aux types de réponses (cf. Tableau 1).

Tableau 1. Types et exemples de réponses.

Type de réponse portant sur D, discours contenu dans le document utilisé	Exemple de réponse totalement adéquate	Exemple de réponse fautive de bon type	Erreur de type
1. Identification de la situation	Où se trouve-t-on ? <i>On se trouve chez le médecin.</i>	Où se trouve-t-on ? <i>Ils se trouvent chez le médecin</i>	Que dit-elle de son travail ? [type 3] <i>Elle travaille dans une empresa de travaille temporaire.</i>
2. Reprise d'une formule de D avec ajout d'un contexte rendant la réponse synt. correcte et pragm. adaptée	Quelle est la situation de famille de cette femme ? <i>Elle est mariée et a deux enfants</i>	Quelle est la situation de famille de cette femme ? <i>il a marié / elle a deux enfants</i>	Comment le médecin trouve-t-il ce qui ne va pas ? [type 5] <i>Il le demande Vous travaillez ?</i>
3. Reformulation de D	Quels conseils donne-t-il ? <i>Il lui dit de se reposer et de moins travailler</i>	Quels conseils donne-t-il ? <i>Le conseil de docteur est travailler pour le longtemps</i>	Que font les personnes qui parlent ? [type 1] <i>Ils parlent del someil ; la fille ne peux pas dormir bien</i>
4. Synthèse à partir de plusieurs formules de D	Que dit-elle de son travail ? <i>Elle a beaucoup de dossiers à traiter</i>	Que dit-elle de son travail ? <i>Elle dit que elle travail beaucoup</i>	
5. Reformulation d'un raisonnement de D	Comment le médecin trouve-t-il ce qui ne va pas ? <i>Il lui pose des questions sur son travail et sur sa vie familiale</i>	Comment le médecin trouve-t-il ce qui ne va pas ? <i>Le docteur lui écoute et comme ça elle trouve la solution</i>	Néant [il faudrait une réponse qui reproduise un processus de raisonnement à une question qui n'en appelle pas]

À l'issue de cette première phase d'expérimentation, nous avons été amenés à distinguer parmi les éléments du corpus différents types de fonctions communicatives, en négligeant les réponses inaudibles :

- métacommentaires (non seulement à l'oral mais aussi à l'écrit) ;
- demandes d'informations ;
- réponses effectives.

Il serait important de pouvoir reconnaître automatiquement les métacommentaires et les demandes d'information. En effet, même s'ils apparaissent comme étant à négliger (c'est-à-dire n'appelant pas de traitement spécifique), les métacommentaires permettent de poser un diagnostic sur la démarche de l'apprenant. Quant aux demandes d'information, il faut bien évidemment y répondre, afin de permettre à l'apprenant d'ajuster sa stratégie discursive.

Le travail étant encore en cours, nous nous limiterons ici à un échantillon de notre analyse. Le cinquième type de questions, reformulation de raisonnement, est à évaluer en fonction de la présence ou de l'absence d'un processus. Ainsi, une réponse du type « Le médecin dit qu'elle travaille trop » (exemple : « *Il dit elle travaille trop* ») ne saurait convenir car elle fournit une assertion et non le déroulement d'un processus. De même, une réponse du type « Le médecin demande à la patiente si elle travaille » (exemple : « *Il le demande Vous travaillez ?* ») ne conviendra pas plus pour les mêmes raisons. En revanche, une réponse comme « *Parce que, elle n'a pas les problèmes dans sa famille, elle ne pas boit beaucoup de thé ou café, c'est pour sa que, il a décidé qu'elle a un problème de travail* » — bien que défailante aux niveaux orthographique et syntaxique — sera considérée comme du bon type : elle présente le déroulement d'un raisonnement (ordre « constats, conclusion déduite », marque de déduction « *c'est pour sa que* »), la démarche discursive consistant à reproduire un raisonnement étant de type « constats ou hypothèses, conclusion déduite ». L'évaluation de l'appropriation sémantique et de l'adaptation pragmatique sera donc fondée principalement sur l'adéquation de la démarche discursive à l'attente du locuteur posant la question et sur la présence de marques grammaticales ou lexicales corroborant la démarche discursive en question. Une fois cette évaluation effectuée, on peut passer à l'analyse de la conformité morphologique et syntaxique.

Le problème le plus délicat est le fait que l'appropriation sémantique et l'adaptation pragmatique ne sont pas binaires (*oui-non*) mais scalaires (il y a des degrés) alors que la décision (d'accepter la réponse ou de la refuser) sera, elle, binaire. Dans le cas de la correction morphologique et syntaxique, la binarité est plus fréquente ; la binarité de la décision sera donc plus évidente à mettre en œuvre. On aura remarqué que nous appliquons les critères (pragmatique, sémantique, syntaxique, morphologique) dans l'ordre inverse à la fois de la présentation habituelle et de la présentation que nous en avons donné plus haut. Cette démarche (apparemment inversée) est en fait celle de la production (dans les modèles *top-down*) et de l'interprétation.

D'autres phases d'expérimentation sont en cours et à venir. Nous poursuivons notre travail afin d'identifier des paramètres d'analyse qui soient appropriés pour les différents types de réponses.