

CONCEPTION MULTIMODALE : QUELQUES ORIENTATIONS POSSIBLES (APPLICATION À UN VOCABULAIRE ÉTRANGER)¹

Michel DUBOIS, Isolde VIAL, Thierry BOLLON

ERIHST, Laboratoire de Psychologie Sociale,
Bâtiment Sciences Humaines et Mathématiques,
Université Pierre Mendès France,
BP 47, 38040 Grenoble Cedex 9
midubois@upmf-grenoble.fr

Résumé : À la suite de travaux de psychologie cognitive concernant l'augmentation de la capacité d'imagerie mentale de mots, cette recherche vise à préciser l'effet sur l'apprentissage d'un vocabulaire étranger de différentes modalités de présentation multimédia. Pour ce faire, nous avons mené une expérience pour comparer l'efficacité de quatre modalités de présentation du matériel d'apprentissage de mots étrangers sur écran d'ordinateur : présentation « conventionnelle » (mot français et étranger avec sa prononciation sonore), présentation conventionnelle « illustrée » (idem avec une illustration de l'objet correspondant au mot), présentation « imagée » sonore (qui permet de relier sémantiquement grâce à la technique du mot clé - cf. Atkinson & Raugh (1975)- les deux mots grâce à une liaison imagée et sonore) et une présentation « imagée textuelle » (idem avec une liaison imagée et textuelle). Différents modes de rappel étaient proposés (audio, visuel et audiovisuel) pour évaluer l'impact des modalités d'encodage en fonction du rappel. Les résultats montrent un avantage significatif en faveur des présentations « imagées » ce qui confirme l'hypothèse de la nécessité d'une intégration complémentaire et en corréférence des différents média. Mais au delà se pose aussi le problème du choix et de la combinatoire des différents média pour aider à la construction d'une représentation adaptée à l'objectif de l'apprentissage (type de rappel et caractéristiques de la tâche).

Mots clés : apprentissage, charge cognitive, double codage, informations multimodales, multimédia, psychologie cognitive.

Abstract : This article reports findings on the effect of diverse modalities of multimedia presentation on foreign vocabulary learning. In the experience related here, we have compared the efficiency of four modalities of presentation of computer learning equipment of foreign words : a conventional presentation (french and foreign words with sound-pronunciation), an illustrated conventional presentation

1 Remerciements : ce travail est soutenu par le Centre National des Études sur les Télécommunications (CNET-PAA/TSA/UST, 38-40 Av. Général De Gaulle, 92131 Issy les Moulineaux, France) dans le cadre d'une convention de recherche.

(the same adding an illustration of the object corresponding to the word), an auditive presentation with an image (which creates a semantic link thanks to the linkword technique -cf. ATKINSON & RAUGH- using sound and imagery), a textual presentation with an image (the same adding text and imagery). Different recall modalities were used (auditory, visual and audiovisual) to study the effect of encoding modalities according to the recall. Results show a significant advantage of the « image » presentation, which confirms the hypothesis of the necessity of integration in co-reference different media. The issue of choice and combination of different media should be further studied to help building a representation adapted to the purpose of learning.

Key words : *cognitive load, cognitive psychology, dual coding, learning, multi-media, multimodale informations.*

INTRODUCTION

Les théories classiques de la psychologie permettent de prédire les performances du sujet lorsqu'il est face à un matériel verbal ou imagé (Ericsson et Kintsch, 1995 ; Anderson, 1993 ; Simon, 1990 ; Kosslyn et al., 1984). En revanche, peu de modèles permettent de décrire le comportement du sujet lorsque ce dernier doit intégrer plusieurs types d'informations simultanément (Kosslyn et al., 1990). Ce type de situation est le propre de l'utilisation des outils multimédia qui permettent de rassembler sur un même support des médias relevant de différents formats informationnels (image-son-écriture). Cette potentialité du multimédia permet de créer différentes sortes de présentation d'une même information (aspects multimodaux avec combinaison de médias visuels et auditifs, icôniques et verbaux...). Le multimédia ne peut, toutefois, se contenter d'être un « mélange » des supports traditionnels, d'en imiter un seul, ni de les reproduire tous, stricto sensu (Brouste & Cotte 1993). Il peut constituer un nouvel outil pédagogique du fait de ses potentialités de combinaison des différents médias pour améliorer l'apprentissage² (Agnola 1996).

Si logiquement, le format imagé et/ou l'utilisation de sons, doivent induire une meilleure performance, cette multiplicité des matériels introduit une complexité des traitements. L'image, par exemple, engendre potentiellement une pluralité de significations (en fonction des sujets, des moments, des contextes...). De plus, les combinaisons entre des données de différentes natures (son-image-texte) en référence mutuelle n'engendrent pas toujours l'efficacité attendue. D'autres études (Dumas de Rauly, 1987) affirment que globalement les différentes sources d'information ont des effets massifs sur la mémorisation.

L'utilisation d'un matériel multimédia soulève dès lors deux questions principales du point de vue du fonctionnement cognitif (Seel & Strittmatter, 1989) :

- comment les connaissances acquises par médias textuel, graphique ou sonore sont-elles traitées et représentées ?

² Les secteurs liés à l'apprentissage enregistrent aujourd'hui la croissance la plus importante des ventes. Près de 40% des applications multimédia concerneraient les secteurs de la formation.

- quels sont les effets des combinaisons texte-figure-son sur l'apprentissage et quels sont les critères pour agencer cette combinaison de manière efficace en regard de telle ou telle activité ?

Certains auteurs affirment que plus de sens sont mobilisés, plus la possibilité de développer le concept chez l'apprenant est importante. Wilmet (1994) montre des différences de rétention de l'information en fonction du mode de communication (oral, imagé et simultané) avec une meilleure performance lors de la combinaison sonore-visuelle. Nous proposons ainsi l'hypothèse générale, en nous appuyant également sur les travaux de Atkinson et Raugh (1975) et de Lieury et al. (1988) que l'induction d'un format figuratif devrait induire de meilleures performances dans une tâche d'apprentissage verbal. Nous développons cette hypothèse générale en intégrant dans notre expérience différents types de format figuratif : dessin illustratif, incitation à l'utilisation d'images mentales.

À partir de ces résultats, nous tenterons de fournir certains éléments utiles à la conception d'un support multimédia dans une tâche d'apprentissage d'un vocabulaire étranger³.

CADRE THÉORIQUE

Image mentale et traitement de l'information

L'image mentale semble faciliter le traitement de l'information verbale (Paivio, 1969). Dans une expérience d'apprentissage de paires de mots associés, Bower (1970) a montré que les sujets adultes qui devaient créer une image mentale entre les référents de deux mots, apprenaient les associations plus rapidement que dans la situation de contrôle. Il apparaît toutefois que ce ne sont pas les images en elles-mêmes qui seraient bénéfiques pour le rappel ultérieur, mais les liens associatifs qu'elles créent. Epstein, Rock & Zuckerman (1960) ont utilisé des images dites « en interaction spatiale » (ex. « l'illustration d'un rocher brisant une bouteille ») pour faciliter le rappel des objets comparativement aux deux mêmes éléments mis côte à côte. Bower, Karlin & Dueck (1975), à partir de matériel imagé incohérent, révèlent qu'un commentaire établissant un lien sémantique entre les éléments de l'illustration favorise son rappel en augmentant la profondeur du traitement. On retrouve des résultats similaires pour la mémorisation des mots (Craik, 1973) ou des visages (Bower & Karlin, 1974). Les effets de rappel et de reconnaissance sont supérieurs dès lors que le protocole présenté permet de comprendre l'image, d'identifier ce qui la conceptualise ou de mieux structurer mentalement l'objet qu'elle présente, c'est à dire de tenir compte du message (contenu) caché derrière le moyen (médium).

L'image devient surtout utile lorsqu'elle complète ou lie les différents éléments du texte (description de caractères « physiques » ou de relations spatiales). Pour Bransford & Jonhson (1972), les images peuvent donner des informations

3 Pour plus de précisions cf. Dubois M., Kouabenan R.D., Mallein P., Toussaint Y., *Étude socio-cognitive des usages du multimédia*, Rapport final, CNET Issy les Moulineaux, Janvier 1998.

contextuelles qui facilitent la compréhension du texte associé. Inversement, Barthes (1964) a mis en évidence que l'écriture lorsqu'elle est intégrée dans une image (légende, commentaire...) permet de réduire la polysémie de l'image. Barthes attribue au verbal le rôle d'ancrage et de relais. Dans le même sens, Burton, Bruce & Johnston (1990) dans leur modélisation de la mémoire montrent que l'association par une relation sémantique renforce les liens entre les objets présentés et favorise leur rappel, ce que ne permet pas une simple relation de contiguïté.

Interaction des formats verbaux et figuratifs

L'efficacité de l'image est attestée sur différents types de tâches et du fait de processus multiples. Jonassen (1982), attribue aux illustrations trois fonctions dans un texte didactique :

- une fonction attentionnelle car elle éveille l'attention et augmente l'intérêt du texte ;
- une fonction explicative en aidant à la compréhension ;
- une fonction rétentionnelle (amélioration la rétention en mémoire).

Si, les images semblent appropriées à la description d'informations spatiales, en revanche, les aspects spécifiques des relations spatiales peuvent être décrits verbalement de façon plus détaillée (Bieger & Glock, 1985). Le langage verbal (écriture) paraît plus adapté aux tâches de précision, son sens étant relativement univoque.

Il ressort de ces différents travaux que la combinaison de texte et de figure n'est réellement efficace que lorsque les deux informations contenues dans l'un et l'autre sont complémentaires et adaptées à chaque présentation. Par exemple, l'ajout d'une illustration à côté d'un mot à mémoriser (Lieury & al., 1988) ne favorise pas son apprentissage.

D'un point de vue cognitif, les traitements figuratifs (illustration, schéma, figure...) et verbaux renvoient à des modes de représentation différenciés (pour les traitements auditifs en co-référence, les travaux sont relativement rares). Les partisans des modèles propositionnels (Anderson & Bower, 1973 ; Anderson, 1983) soutiennent qu'il n'existe qu'une seule représentation des connaissances en mémoire (abstraite et amodale) dont l'unité est la proposition (une relation liant deux arguments). En mémoire, les stimuli figuratifs et verbaux activeraient le même code propositionnel. Selon cette hypothèse, les images sont mémorisées sous forme propositionnelle comme le matériel verbal.

Shepard (1967) a montré expérimentalement que la reconnaissance d'images était supérieure à la reconnaissance de mots et de phrases. De plus, la rétention d'un matériel concret, en général des images ou des dessins, est meilleure que celle d'un matériel verbal (Ducharme & Fraisse, 1965 ; Fraisse & Leveille, 1975). Paivio, Rogers & Smythe (1968), expliquent cette supériorité en faisant l'hypothèse que les images sont stockées sous forme figurative et verbale. Elles bénéficieraient donc d'un double codage, tandis que le matériel verbal serait stocké sous forme uniquement verbale. Paivio (1971a) s'appuie sur le fait que les mots concrets sont mieux mémorisés que les mots abstraits, et que, parmi les mots concrets, ce sont les mots

qui évoquent le plus facilement des images qui sont les mieux retenus. Il postule l'existence de deux sous-systèmes de codage interconnectés, mais fonctionnellement distincts, pouvant intervenir dans l'activité mnémonique et l'apprentissage, mais également dans les activités perceptives et intellectuelles. Selon Betrancourt (1992, 1996), le gain d'apprentissage significatif lorsque la présentation utilise deux médias associés, par rapport aux situations où l'un des deux médias est présenté, semble infirmer l'existence d'un système commun et amodal de traitement de l'information. Il existerait donc d'après ces résultats, plusieurs modes de représentation différenciés, codant chacun des caractéristiques spécifiques du stimulus.

Sur un plan général, Mayer & Sims (1994) ont proposé à la suite de Paivio (1971b, 1986) et de Clark & Paivio (1991), un compte rendu des différents processus en œuvre lors d'une consultation de données multimédias par un sujet. Il y aurait lors d'une double source de présentation verbale/orale et visuelle/illustrée, un encodage successif qui permettrait d'établir différentes connections significantes. Les deux représentations initiales (verbale et visuelle) nécessiteraient d'abord d'élaborer des connections référentielles performantes, en fonction de la spécificité du matériel, qui seraient dans un second temps traitées en co-référence. Plus les dispositifs de présentation vont favoriser l'élaboration de ces trois connections plus l'apprentissage sera important.

Problématique

Dans cette perspective, le multimédia, offre certaines potentialités pour induire l'imagerie mentale (Peeck 1987 ; Marmolin 1992). On devrait donc s'attendre à une efficacité supérieure du multimédia en termes d'apprentissage comparativement aux produits traditionnels (Monet, 1995 ; Boursier & Taufour, 1994).

Notre objectif n'est pas ici de se positionner sur le fait que l'information figurative soit codée de façon amodale ou bimodale mais d'analyser les interactions possibles lors du traitement de l'information par le sujet des différents formats de présentation de l'information. En d'autres termes, il s'agit d'essayer d'identifier les effets de ces différents formats sur l'apprentissage de langues étrangères. Il a été observé que le fait d'inciter le sujet à créer une image mentale dans ce type de situation d'apprentissage donnait de meilleures performances. Deux expériences d'apprentissage de langues étrangères s'avèrent intéressantes de ce point de vue, celle d'Atkinson & Raugh (1975) avec l'utilisation d'une consigne d'une imagerie spontanée de la part du sujet et celle de Pressley (1977) par la création d'une image mentale imposée. Toutes deux confirment l'efficacité de l'utilisation d'un mot-clé pour associer sémantiquement le mot et sa traduction. De même, Lieury & al. (1988) ont montré la pertinence de fournir un dessin qui intègre le mot-clé et sa traduction. Dans cette condition le sujet n'a pas à créer lui-même une représentation mentale de la situation et les résultats sont largement améliorés.

Dans la continuité de ces travaux, nous nous attendons à ce qu'effectivement l'intégration des données en co-référence induise une meilleure performance. Ainsi, les sujets devraient produire un meilleur rappel lorsque l'information verbale aura été accompagnée d'une information figurative autorisant une construction de liens

sémantiques et phonétiques entre les informations. Les modes de présentation intégrant le matériel permettraient de faire opérer des traitements supplémentaires au sujet. La représentation imagée intégrative créerait des correspondances entre les données présentées qui permettraient au sujet de traiter le matériel en profondeur et qui lui ouvriraient des voies de récupération plus performantes. En revanche, la simple présence de multiples sources d'informations devrait selon Sweller & al. (1990) provoquer un partage de l'attention pour l'activité d'intégration des données et pour l'acquisition d'un schéma correspondant. Cette séparation des informations entraînerait une augmentation de la charge cognitive du fait des traitements supplémentaires nécessaires à l'intégration des différents éléments en référence mutuelle. On devrait donc observer une baisse de performance pour les situations où le matériel informationnel nécessite un traitement supplémentaire de l'information pour que le sujet puisse construire une représentation cohérente et récupérable en mémoire. Ainsi, on s'attend à de moins bonnes performances si la présentation, en ajoutant une simple illustration ne permet pas d'intégration des informations entre elles.

Nous avons réalisé une expérimentation visant à intégrer ces différents aspects théoriques. Il s'agit de tester différents modes de présentation de l'information et leur interaction avec différents modes de rappel dans une tâche de mémorisation d'un vocabulaire russe. La méthode mnémotechnique du mot-clé, testée par Atkinson & Raugh (1975), est reprise et adaptée à différents supports multimédias.

PRÉSENTATION DE L'EXPÉRIENCE

Méthode

Les sujets

Les 60 participants étaient des étudiants de 1^{ère} année à bac+6 de l'Université de Grenoble. Aucun des sujets n'avait appris le russe. Les sujets étaient répartis aléatoirement dans chacun des groupes expérimentaux.

Variables expérimentales

Nous avons manipulé deux Variables Indépendantes. La V.I.1 était *le mode de présentation de l'information sur multimédia* et avait 4 modalités. La V.I.2 est *le mode de rappel du vocabulaire* et avait 3 modalités. Le plan de l'expérience était : S5 <P4*R3> (P était le facteur mode de présentation et R était le facteur mode de rappel).

Les 4 modalités de présentation ont été appelées P1, P2, P3 et P4. Dans toutes les modalités de présentation, le mot russe et la traduction étaient donnés aux sujets auditivement et textuellement sur l'écran de l'ordinateur. En condition contrôle P1 (présentation conventionnelle), seuls le mot russe et la traduction étaient donnés. En P2 (présentation conventionnelle illustrée), une illustration représentant le mot à traduire était ajoutée. En P3 et P4, nous avons utilisé une méthode suscitant une imagerie mentale imposée, en utilisant la technique du mot-clé inséré

dans une phrase. Le mot-clé était proche phonétiquement du mot russe et la phrase permettait de créer un lien sémantique entre le mot-clé et la traduction. Par exemple, pour Коцѣр, se traduisant par « feu » et se prononçant « Kastiore », le mot-clé choisi était « castor » et la phrase était « devant le feu, des castors ». Cette phrase était présentée à l'oral en P3 (présentation imagée sonore) et présentée à l'écrit sur l'écran en P4 (présentation imagée textuelle). En plus de cette phrase, une image (utilisant l'illustration de P2), mais recontextualisée dans une scène signifiante représentant le lien entre le mot-clé et le mot russe s'affichait sur l'écran.

Le rappel du vocabulaire consistait à présenter aux sujets un mot russe pour lequel ils devaient trouver et écrire la traduction. En plus des 19 mots russes appris, nous avons ajouté 11 mots russes « distracteurs » non appris, proches alphabétiquement et phonétiquement des mots présentés. Au sein de chaque groupe expérimental, trois sous-groupes homogènes ont été constitués (nombre de sujets = 5). À chaque sous-groupe correspondait un mode de rappel spécifique. Le mot était donné soit seulement à l'écrit sur l'écran (rappel visuel : Rv), soit seulement à l'oral (rappel auditif : Ra) soit à l'écrit et à l'oral (rappel audiovisuel : Rav).

La Variable Dépendante utilisée était la réussite à la tâche du vocabulaire : nombre de traductions correctes.

Procédure et matériel

L'expérience se déroulait en deux phases : une phase d'apprentissage des mots russes et une phase de rappel.

Pour la phase d'apprentissage, le matériel se composait de 19 couples de mots (le mot russe et la traduction) à apprendre (présentation successive de 19 écrans). La liste des mots à apprendre était la même pour tous les groupes. Les 19 écrans étaient présentés deux fois de suite et la durée de présentation de chaque écran était de 15 secondes chacun. La phase de rappel a consisté à présenter une seule fois et successivement 30 mots russes pendant 15 secondes chacun. La tâche du sujet était d'écrire la traduction pour chacun des mots russes qu'il reconnaissait.

PRÉSENTATION DES RÉSULTATS

Nous observons un effet de la variable mode de présentation ($F(3,48) = 7,049$, $p < .0005$). La décomposition de cet effet primaire montre que P1 se différencie de P3 (4,8 vs 8,5, $F(1,48) = 11,13$, $p < .01$). De même, P2 se différencie de P3 (3,7 vs 8,5, $F(1,48) = 18,39$, $p < .0001$) et nous observons le même type de résultat entre P2 et P4 (3,7 vs 6,5, $F(1,48) = 6,26$, $p < .01$) (cf. fig. 1). On peut noter que les présentations P3 et P4 n'engendrent pas de confusion en terme de traduction entre le mot clé et le mot correct à traduire.

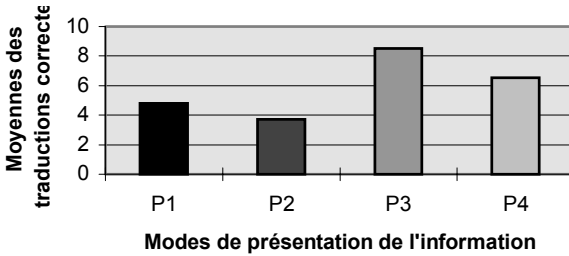


Figure 1. Traductions correctes en fonction du mode de présentation de l'information.

La variable mode rappel produit également un effet significatif ($F(2,48) = 16,07$, $p < .0000$). Les sujets de la modalité Rav produisent de meilleures performances que ceux de la modalité Rv ($7,15$ vs $2,75$, $F(1,48) = 26,61$, $p < .0000$). De même, les sujets de la modalité Ra produisent un meilleur rappel que ceux de Rv ($7,8$ vs $2,75$, $F(1,48) = 27,15$, $p < .0000$).

L'interaction de nos deux variables produit aussi un effet significatif ($F(6,48) = 7,136$, $p < .0000$). (cf. résultats tableau 1)

Modalités de présentation	Modalités de rappel	Moyennes des traductions correctes
P1	Audiovisuel	8,2
	Auditif	3,8
	Visuel	2,4
P2	Audiovisuel	2
	Auditif	4,6
	Visuel	4,6
P3	Audiovisuel	10,4
	Auditif	14,6
	Visuel	0,6
P4	Audiovisuel	8
	Auditif	8,2
	Visuel	3,4

Tableau 1. Nombre de traductions correctes selon l'interaction des modalités de rappel et de présentation.

Nous observons un premier effet simple qui indique un rappel différentiel pour la modalité Rav en fonction des modes de présentation ($F(3,48) = 6,905$, $p < .001$). La méthode des comparaisons par paires au seuil .01, nous permet d'observer des effets significatifs entre P2 et P3 ($2,0$ vs $10,4$), P2 et P1 ($2,0$ vs $8,2$) et entre P2 et P4 ($2,0$ vs $8,0$). Nous obtenons un second effet simple significatif avec la modalité Ra ($F(6,48) = 12,891$, $p < .000$). Les comparaisons par paires (au seuil .01) indiquent des effets significatifs entre P3 et P1 ($14,6$ vs $3,8$), P3 et P2 ($14,6$ vs $4,6$) et P3 et P4 ($14,6$ vs $8,2$). Pour la modalité Rv, l'analyse par contraste n'indique qu'un seul effet significatif ($F(1,48) = 4,259$, $p < .05$) à savoir entre P2 et P3 ($4,6$ vs $0,6$). Pour la modalité P1, il apparaît un effet significatif en fonction des modalités

de rappel ($F(2,48) = 4,877, p < .01$). La méthode des comparaisons par paires (au seuil .01) permet d'observer une différence entre Rav et Rv (8,2 vs 2,4). La modalité de présentation P2 ne produit pas de performance différentielle quelque soit le type de rappel. La modalité P3 donne lieu à un effet significatif en fonction de la variable rappel ($f(2,48) = 27,478, p < .000$). Nous observons (comparaisons par paires au seuil .01) deux effets significatifs entre Rav et Rv (10,4 vs 0,6) et Ra et Rv (14,6 vs 0,6). Un effet simple apparaît également pour la modalité P4 ($F(2,48) = 3,925, p < .026$). L'analyse des contrastes permet d'opposer les groupes Rav et Rv (8,0 vs 3,4, $F(1,48) = 42,261, p < .0000$) ainsi que Ra et Rv (8,2 vs 3,4, $F(1,48) = 6,133, p < .0168$).

DISCUSSION

Pour la variable type de présentation de l'information, il apparaît en premier lieu que les sujets produisent de meilleurs rappels lorsque l'information présentée est intégrée sous forme auditive (P3) que dans le cas où l'on n'a ajouté qu'une image simple au vocabulaire (P2) et dans la situation contrôle (P1). On retrouve les résultats de Lieury et al. (1988) dans lesquels une image intégrant deux concepts permet un meilleur apprentissage qu'une image simple. L'ajout d'une image ne représentant qu'une information n'autorise pas une intégration et n'améliore pas les performances des sujets. La situation où l'information est intégrée sous forme textuelle (P4) produit également de meilleurs résultats par rapport à la situation non-intégrée (P2). L'hypothèse de l'intégration de l'information peut donc être maintenue (Bower, 1970). La quantité d'information à traiter par le sujet n'apparaît pas induire de partage attentionnel limitant les effets d'apprentissage selon l'hypothèse de surcharge cognitive (Sweller et al., 1990) dès lors que les différentes sources d'informations sont intégrées. En effet, les deux présentations à fortes charges informationnelles génèrent les meilleurs apprentissages, l'intégration des informations entre-elles permettant aux sujets de mieux mémoriser les informations qui lui sont présentées : le lien entre le mot russe et la traduction française.

Pour les trois situations de rappel, le résultat le plus intéressant à noter est la moindre performance en situation uniquement visuelle (Rv). En effet, les deux autres situations qui produisent les meilleures performances intègrent un code oral compatible avec la modalité dominante des différentes modalités de présentation de l'information de la phase d'apprentissage.

Lorsque le rappel est audio-visuel (Rav), les résultats en termes d'apprentissage de vocabulaire sont meilleurs en situation de présentation de l'information intégrée auditivement (P3) et plus faible dans la présentation avec image simple (P2). On peut donc conclure que la présence d'une phrase sous forme auditive joue un rôle intégrateur lors d'une présentation à dominante auditive. La présentation simultanée d'informations figurative et verbales auditive en co-référence sémantique améliore l'apprentissage (Mousavi, Low et Sweller, 1995).

Dans la situation de rappel auditif (Ra), on observe d'une manière générale un meilleur apprentissage lorsque les informations sont intégrées sous forme auditive. La condition intégrée sous forme auditive (P3) permet un meilleur

apprentissage que celle intégrée sous forme textuelle (P4). Ces résultats s'accordent de nouveau avec ceux de Mousavi, Low et Sweller (1995) car une présentation dans laquelle les deux éléments intégratifs (image et texte) sont sous forme uniquement visuelle permettent un moins bon apprentissage que s'ils sont sous forme visuelle et auditive. De plus, la phrase écrite engendre un traitement textuel des informations, ce qui semblerait incompatible avec un mode de rappel auditif.

La dispersion des moyennes en rappel visuel (Rv) est moindre que dans les autres conditions de rappel, ce qui signifie que cette modalité de rappel n'induit pas de grosses variations de performances. De plus, dans l'ensemble ces dernières restent relativement faibles par rapport aux autres conditions de rappel. On remarque cependant une différence entre une présentation intégrée sous forme auditive (P3) et la présentation avec une image simple (P2). Cette baisse attendue des performances en P3 s'expliquerait par le type d'intégration induit par la phrase auditive, cette phrase créant un traitement interférent avec la modalité visuelle de rappel. La situation avec une image simple donne des résultats faibles mais meilleurs que ceux de la situation P3 puisque la présentation laisse la possibilité aux sujets de traiter l'information textuelle, sans incompatibilité forte entre phase d'apprentissage et de rappel.

CONCLUSION

On constate l'influence de certaines modalités de présentation de l'information sur la mémorisation. La conception de données suscitant une imagerie mentale (ici l'adaptation de la technique du mot clé associée à des combinaisons sonores et imagées), permet d'encoder l'information de manière plus efficace. Ces différents résultats montrent que l'ajout d'une image à un corpus sonore et textuel peut constituer une aide notable pour les sujets sous certaines conditions. L'illustration doit favoriser un traitement du matériel présenté. On rejoint l'ensemble des résultats révélés dans d'autres contextes.

L'ajout d'un médium supplémentaire ne s'avère pas efficace si l'information apportée est inopérante pour faciliter un traitement d'information en profondeur (P2 n'obtient pas de différence significative dans l'apprentissage par rapport à P1). Par contre, les sujets de P2 (introduction d'une illustration) ont de moins bonnes performances que les sujets des présentations intégrées alors que la redondance de l'information à traiter est introduite là aussi sous différentes modalités informationnelles (texte, son, image). Ce résultat peut s'interpréter dans deux perspectives : un effet de charge cognitive plus importante qui rendrait l'encodage plus difficile (Sweller et al., 90) en partageant l'attention ou un effet de traitement sémantique contre-performant. Dans ce cadre, on peut supposer que l'image engendre potentiellement une pluralité de significations (en fonction des sujets, des moments, des contextes...). Elle instaurerait des significations multiples, voire des confusions chez le sujet.

L'amélioration des multimédias éducatifs, semble donc passer par la recherche d'une conception de médias en référence mutuelle pour éviter ces deux effets. Dans cette perspective on note des recherches encore peu nombreuses consacrées

aux effets de l'utilisation simultanée de différents médias sur un même support. Les interactions inter-médias n'ont pas encore fait l'objet de recherches expérimentales systématiques dans le cadre du multimédia. Pourtant elles constituent une voie d'étude intéressante pour aider les concepteurs à déterminer certains paramètres pertinents dans l'agencement des différents médias en fonction d'une tâche donnée. Si l'écriture multimédia veut être efficace, elle se doit de concevoir des supports en accord avec les spécificités et les limitations cognitives des individus. Cela nécessite de comprendre comment l'homme réagit face à tel ou tel support en fonction d'une tâche donnée. Il est évident que ce qui fonctionne bien pour l'apprentissage de vocabulaire étranger, peut ne pas correspondre à d'autres tâches. On peut cependant affirmer avec de nombreux auteurs, que le taux de rappel sera amélioré uniquement si les images illustrent l'information centrale du texte, si elles représentent un nouveau contenu qui est important pour la tâche et si elles établissent une relation structurelle dont le texte fait mention. La présentation doit établir des références entre médias sur le contenu de la tâche, en permettant une redondance multimodale de l'information clairement explicitée. Les différents supports doivent activer un schéma unique en se complétant et en permettant d'organiser les différents éléments qui composent la tâche. Le sujet est ainsi invité à opérer des traitements en profondeur.

Des affinements semblent aussi nécessaires, en ce qui concerne les modèles théoriques. En effet, la présentation 3 apparaît souvent plus efficace que la présentation 4. Ce fait serait dû non pas aux effets des informations écrites ou auditives sur l'apprentissage mais plutôt à l'effet de la combinaison entre les différents médias par rapport à une tâche donnée et en fonction d'un type de rappel spécifique. Ces résultats suggèrent donc que, dans ce type de tâche, l'information auditive en référence à une image permet un meilleur apprentissage qu'une information textuelle en référence à la même image. Cet aspect peut compléter le modèle de Mayer et Sims (1994). La construction des connections référentielles significatives semblent tributaires de plusieurs facteurs. Tout d'abord, des modalités de présentation et de la construction en co-référence des données multimodales, mais aussi des caractéristiques de la tâche (ici les différentes formes de rappel). Plus les conditions du rappel sont en adéquation avec les modalités d'encodage, plus l'apprentissage semble important. Dans notre expérience, à partir d'un contenu strictement identique, une modalité de rappel audio privilégie les formes d'encodage sous forme audio (et inversement pour un rappel visuel). Ces résultats doivent être approfondis, un des objectifs étant de mettre en évidence les critères déterminant l'efficacité des combinaisons entre différents médias, selon les caractéristiques des tâches.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Agnola M. (1996). *Passeport pour le multimédia*, Paris : CFPJ Éditions.
- Anderson J. R. (1983). *The architecture of cognition*, Cambridge, MA : Harvard University Press.
- Anderson J. R. (1993). *Rules of the Mind*. Hillsdale, New Jersey : LEA.

- Atkinson R., Raugh M. (1975). « An application of mnemonic keyword method to the acquisition of an russian vocabulary », *Journal of Experimental Psychology : Human Learning and memory*, vol 104, n° 2, 126-133.
- Barthes R. (1964). « Rhétorique de l'image », *Communications*, 4, Paris : Le Seuil.
- Betrancourt M. (1992). *Interaction texte/figure, effet de leur disposition relative sur l'apprentissage*, rapport INRIA dans le cadre de l'obtention du DEA, Dir. A. Bisseret, Grenoble.
- Betrancourt M. (1996). *Facteurs spatiaux et temporels dans le traitement cognitif des complexes texte-figure*, Thèse de doctorat, Sciences Cognitives, INPG, Grenoble.
- Bieger G. R., Glock M. (1985). « Comprehending spatial and contextual information in picture-text instructions », *Journal of Experimental Education*, p. 181-188.
- Boursier P., Taufour P.A. (1994). *La technologie multimédia*, Paris : Hermès.
- Bower G. H. (1970). « Imagery as relational organizer in associative learning », *Journal of Verbal Learning and Verbal Behavior*, 9, 529-533.
- Bower G. H., Karlin M. B. (1974). « Depth of processing pictures of faces and recognition memory », *Journal of Experimental Psychology*, 103, 751-757.
- Bower G. H., Karlin M. B., Dueck A. (1975). « Comprehension and memory for pictures », *Memory and Cognition*, 3, 216-220.
- Bransford J. D., Johnson M. K. (1972). « Contextual prerequisites for understanding : some investigations of comprehension and recall », *Journal of Verbal Learning and Verbal Behavior*, 11, 717-726.
- Brouste P., Cotte D. (1993). *Le multimédia, promesses et limites*, Paris : ESF éd.
- Clark J. M., Paivio A. (1991). « Dual coding theory and education », *Educational Psychology Review*, 3, 149-210.
- Craik F. I. M. (1973). « A "Levels of analysis" view of memory », in P. Pliner, L. Krames & T. Alloway (Eds), *Communication and affect : Language and thought*, New York : Academic Press.
- Ducharme R., Fraisse P. (1965). « Etude génétique de la mémorisation des mots et des images », *Canadian Journal of Psychology*, 19, 253-261.
- Dumas de Raully T. (1987). *Choisir et utiliser les supports visuels et audiovisuels*, Paris : Les Éditions d'Organisation.
- Epstein W., Rock I., Zuckerman C. B. (1960). « Meaning and familiarity in associative learning », *Psychological Monographs*, 74, n° 491.
- Ericsson K. A., Kintsch W. (1995). « Long-term working memory », *Psychological review*, 102, 2, 211-245.
- Fraisse P., Leveille M. (1975). « L'influence du codage visuel de phrases sur leur mémorisation à court terme », *L'Année Psychologique*, 75, 409-416.
- Jonassen D. H. (1982). *Technology of text : principles for designing, structuring and displaying texts*, New Jersey : Educational Technologies Publication.
- Kosslyn S. M., Brunn J., Cave K. R., Wallach R. W. (1984). « Individual differences in mental imagery ability : A computational analysis », *Cognition*, 18, 195-243.

- Kosslyn S. M., Flynn R. A., Amsterdam J. B., Wang G. (1990). « Components of high-level vision : A cognitive neuroscience analysis and accounts of neurological syndromes », *Cognition*, 34, 203-277
- Lieury A., Clinet C., Gimonet M., Lefebvre M. (1988). « Représentations imagées et apprentissage d'un vocabulaire étranger », *Bulletin de Psychologie*, Tome XLI, n° 386.
- Marmolin H. (1992). « Multimedia from perspective of psychology », *Multimedia : principles, systems and applications*, Springer-Verlag, Ed. Lars Kjeldahl.
- Mayer R. E., Sims V. K. (1994). « For whom is a picture worth thousand words ? Extensions of a dual-coding theory of multimedia learning », *Journal of Educational Psychology*, vol. 86, n° 3, 389-401.
- Monet D. (1995). *Le multimédia*, Paris : Ed. Flammarion.
- Mousavi S. Y., Low R. & Sweller J. (1995). « Reducing cognitive load by mixing auditory and visual presentation modes », *Journal of Education Psychology*, vol. 87, n° 2, 319-334.
- Paivio A. (1969). « Mental imagery and associative learning in memory », *Psychological Review*, 76, 241-263.
- Paivio A. (1971a). « Short term sequential memory for pictures and words », *Psychonomic Science*, 11, p. 137-138.
- Paivio A. (1971b). *Imagery and verbal processes*, New-York : Holt, Rinehart & Winston.
- Paivio A. (1986). *Mental representations : a dual coding approach*, Oxford, England : Oxford University Press.
- Paivio A., Rogers T. B., Smythe P. C. (1968). « Why are pictures easier to recall than words », *Psychonomic Science*, 11, p. 137-138.
- Peeck J. (1987). *The role of illustration in processing and remembering illustrated text*, Willows & Houghton, vol. 1, ch. 4.
- Pressley M. (1977). « Children's use of the keyword method to learn simple spanish words », *Journal of Educational Psychology*, n° 69, p. 465-472.
- Seel N., Strittmatter P. (1989). « Presentation of information by media and its effects on mental models », in H. Mandl & J.R. Levin (Eds), *Knowledge acquisition from text and pictures*. Advances in psychology, North Holland, 58, p. 37-58.
- Shepard R. N. (1967). « Recognition memory for words, sentences and pictures », *Journal of verbal learning and verbal behavior*, 6, 156-163.
- Simon H. A. (1990). « Invariants of human behavior », *Annual Review of Psychology*, 41, 1-19.
- Sweller J., Chandler P., Tierney P., Cooper M. (1990). « Cognitive load as a factor in the structuring of technical material », *Journal of Experimental Psychology*, vol. 119, n° 2, 176-192.
- Wilmet B. (1994). « Les hypermédias au service de l'enseignement », *Revue de l'EPI*, n° 73, mars 1994, p. 177-187.