



**HAL**  
open science

## Présentation multimodale de l'information et apprentissage

Michel Dubois, Federico Tajariol

► **To cite this version:**

Michel Dubois, Federico Tajariol. Présentation multimodale de l'information et apprentissage. Cinquième colloque Hypermédias et apprentissages, Apr 2001, Grenoble, France. pp.197-209. edutice-00000461

**HAL Id: edutice-00000461**

**<https://edutice.hal.science/edutice-00000461>**

Submitted on 8 Jun 2004

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

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

## PRÉSENTATION MULTIMODALE DE L'INFORMATION ET APPRENTISSAGE

### Le cas d'une leçon d'aéronautique

**Michel DUBOIS\* et Federico TAJARIOL\*\***

\*Laboratoire de Psychologie Sociale - ERIHST - SHS  
Université Pierre-Mendès-France  
BP 47 - 38 040 Grenoble Cedex

\*\*Équipe ARCADE - CLIPS-IMAG - Université Joseph Fourier  
BP53 - 38041 Grenoble Cedex 9.

Michel.Dubois@upmf-grenoble.fr

**Résumé :** *Le multimédia a ouvert de nouvelles perspectives d'apprentissage par sa possibilité de combiner différentes sources d'informations (texte, son, image). Cependant beaucoup d'études ont montré la complexité des traitements cognitifs face à du matériel multimodal. En effet, le traitement du matériel verbal et du matériel imagé n'induit pas les mêmes formes de représentation. Notre étude expérimentale compare six protocoles de présentation d'une même leçon traitant d'aéronautique. Le contenu était structurellement et sémantiquement équivalent. Deux variables indépendantes étaient manipulées : le matériel imagé (image dynamique vs image séquentielle) et le matériel verbal (textuel vs audio vs audio textuel). Les résultats montrent un effet de la variable présentation du matériel verbal avec une supériorité de la présentation audio significativement la plus performante pour mémoriser et comprendre le contenu. Aucun effet d'interaction n'a été trouvé. Ces résultats montrent comment la présentation multimodale d'un contenu affecte le traitement cognitif de l'information durant la lecture et la compréhension et peuvent aider la conception des interfaces.*

**Mots-clés :** *multimédia, multimodalité.*

**Abstract :** *Multimedia has introduced new perspectives of learning by its advantages of combining various forms of information presentation (text, sound, picture). However, many studies emphasized the complexity of the cognitive treatments related to the multimodal material.*

*Visual and verbal processing (illustrations, diagrams, figures, and so forth.) call on different modes of representation. Our experimental study presents one lesson in aeronautics in which six modalities of presentation are compared. The content is structurally and semantically equivalent. Two independent variables are manipulated : visual material (dynamic picture vs sequential picture) and verbal material (textual vs audio vs audio-textual). No interaction effect is found. Results show how the multimodal presentation of a contents affects the cognitive treatment*

*of the information during the reading and the understanding, and how can it helps interfaces designing.*

**Keywords :** *multimedia, multimodality.*

## **INFORMATIONS MULTIMODALES ET APPRENTISSAGE**

### **Les problèmes posés par la rédaction de messages multimodaux**

En matière d'apprentissage, l'exploitation de données multimédia est censée permettre une communication et une compréhension plus efficace (Peeck, 1987) et générerait une plus grande excitation pour le maintien de l'attention. Des auteurs (Feldman & Schoenwald, 1989 ; Fletcher, 1990) comparant l'apprentissage sur multimédia avec l'apprentissage traditionnel ont montré que plus les attributs d'un média étaient utilisés, plus il y avait apprentissage. Le multimédia encouragerait à une exploration active de l'information plus qu'à une réception passive et fournirait une source de matériel riche pour l'esprit qui préfère des stimuli dynamiques aux stimuli statiques (Marmolin, 1992).

Cependant, selon une perspective cognitive, la complexité des combinaisons entre les différents médias (effet de multimodalité) engendre le risque que la production de supports soit inadéquate avec les modèles mentaux des utilisateurs. En effet, on sait dans le domaine de la psychologie cognitive que l'apport de données multimodales entraîne des traitements spécifiques. Il ne suffit pas de penser intuitivement que telle ou telle image (ou tel ou tel son) aura telle répercussion cognitive pour que cela soit le cas. Dans cette perspective, l'étude de la combinaison des médias présentés (texte, image, son, vidéo) devient cruciale. Cela nécessite, au niveau théorique, de comprendre comment l'homme réagit face à tel ou tel support en fonction d'une tâche donnée et d'étudier très précisément les modalités de présentation susceptibles d'augmenter les performances d'apprentissage et d'éviter des coûts cognitifs inutiles à l'utilisateur. Cela nécessite aussi, dans la pratique, d'aider les concepteurs, rédacteurs et formateurs à comprendre comment la conjugaison de différents systèmes de symboles peut être efficace et sous quelles conditions.

### **Effet du matériel verbal et imagé dans le traitement et l'acquisition des connaissances**

#### ***Traitement du matériel verbal***

Les propriétés multiples du code verbal nécessitent des traitements différents. Le codage du matériel verbal s'établirait à partir de deux systèmes de traitements spécialisés : le système visuel et verbal. Ce phénomène a été étudié par Paivio (1986) qui l'explique par la théorie du double codage. Lorsque l'on voit un chien, l'image du chien est activée en mémoire ainsi que le mot correspondant et inversement lorsqu'on lit le mot « chien ». De Vito & Olson (1973) montrent ainsi que le rappel concernant des épreuves utilisant des adjectifs à valeur d'imagerie élevée sont supérieures à celles utilisant des adjectifs à faible valeur d'imagerie.

Les modalités (textuelle ou auditive) de présentation du matériel verbal ont aussi une influence directe sur le traitement de l'information. Ainsi Brooks (1970) a

présenté à des sujets des phrases décrivant des relations spatiales sous forme écrite ou orale. Les sujets reconstituent mieux les relations spatiales lorsque les phrases sont entendues. De même Levin et Divine-Hawkins (1974) ont testé les modalités de présentation auditive ou textuelle d'un matériel verbal (visualisation du contenu de différents épisodes d'un récit). Les rappels sont plus importants lorsque les personnes ont entendu le récit que lorsqu'elles l'ont lu. La représentation mentale visuelle du texte requiert la mise en œuvre du même système de traitement visuel que celui sollicité par la lecture. La prise en charge simultanée de ces deux activités restreindrait les capacités d'encodage et de traitement de l'information.

### ***Traitement du matériel imagé***

De nombreuses recherches ont discuté aussi de l'influence du matériel illustré et de ses modalités de présentation sur le rappel en fonction de différentes tâches (Vernon, 1953 ; Rusted & Coltheart, 1979 ; Mandl & Levin, 1989 ; Rieber, 1990 ; Winn, 1991 ; Betrancourt, 1995 ; etc.). Dwyer (1968) a montré qu'une image accompagnant un texte n'améliore pas forcément le rappel et la compréhension. Le processus d'identification de l'objet est facilité par une présentation réaliste (photographie) alors que la compréhension est plus efficace avec des schémas. Les images se prêtent plus aussi à la représentation mentale de transformations d'éléments ou de processus technique (Bieger & Glock, 1985). Denis & De Pouqueville (1976) introduisent des modalités dynamiques de présentation de l'image. Ils montrent à des sujets sur un écran informatique une liste d'actions présentées soit sous forme de phrases, d'un dessin, d'une photo seule, d'un film ou de trois photos extraites du film et représentant trois étapes de l'action. Toutes les représentations imagées sont mieux mémorisées que les phrases. Les dispositifs permettant de représenter visuellement les étapes de l'action (film ou plusieurs photos) améliorent encore plus le rappel. La présentation d'une image dynamique permettrait de traiter en parallèle deux types d'informations (statiques et dynamiques) ce qui est susceptible d'enrichir l'encodage. Pour Proffitt & Kaiser (1986), l'ajout d'informations dynamiques permet de faciliter l'apprentissage car le mouvement a des caractéristiques permettant d'exciter le niveau perceptif (processus attentionnels de bas niveau). Le traitement de ces informations assure un codage supplémentaire de haut niveau sur la structure de l'objet comme le suivi du mouvement ou le continuum de l'action (cf. aussi Bruce & Green, 1993). Cependant, la bonne perception des informations dynamiques dépend aussi de son degré de ressemblance avec la réalité (Kaiser, Proffitt & Anderson, 1985). Les présentations de l'objet ou du mouvement doivent être prototypiques par rapport à l'objet ou au processus représenté (Rieber, 1990).

À ces éléments différentiels sur le traitement du format verbal et imagé s'ajoutent la complexité de traitement résultant des interactions entre ces formats (co-référentialité).

### **Effet de l'interaction texte, son, image dans le traitement et l'acquisition des connaissances**

Sur un plan général, Mayer & Sims (1994) ont proposé à la suite de Paivio (1986) et de Clark & Paivio (1991), un compte rendu des différents processus en œuvre lors d'une consultation de données multimodales par un sujet. Il y aurait lors

d'une double source de présentation (verbale et illustration), un encodage successif qui permettrait d'établir différentes connections signifiantes. 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 favorisent l'élaboration de connections, plus l'apprentissage devient important. Ainsi chaque média doit contribuer à une meilleure compréhension et structuration de l'autre et permettre d'opérer des traitements cognitifs profonds en co-référence (Dubois & Vial, 1998). Ce principe est en accord avec la notion de charge cognitive de Sweller *et al.* (1990) qui postule que la séparation de deux sources d'information provoque un partage cognitif de l'attention qui requiert alors un coût cognitif élevé.

Selon Mousavi, Low et Sweller (1995), le fait de présenter les informations sous deux modalités différentes (visuelle et auditive) permet de diminuer l'effet d'attention partagée et donc la charge cognitive. La spécialisation de la mémoire de travail entre informations verbales et imagées, permettrait lorsque l'on présente des informations verbales auditives avec des informations imagées un double codage de l'information car les deux mémoires de travail traiteraient l'information en même temps. La quantité de ressources attentionnelles serait doublée (Mayer & Moreno, 1998).

Ces différentes études montrent la complexité des traitements afférents à du matériel multimodal. Selon le format de présentation du matériel verbal (auditif ou textuel) et imagé (statique ou dynamique), la représentation mentale qui en découle est différente. Or la conception de données multimodales (son, texte, image statique ou dynamique) est un exercice de plus en plus courant dans le milieu de la formation (initiale et continue) et pose dès lors de nombreuses questions pratiques. Comment chaque médium agit-il par rapport à l'autre ? Certains médias permettent-ils une meilleure distinctivité de l'information et en quoi sont-ils plus adéquats face aux caractéristiques de tâches spécifiques (rappel, compréhension, résolution de problème) ? Compte-tenu de ces éléments, nous étudions par une expérience ces différents aspects.

## **PRÉSENTATION DE L'EXPÉRIENCE**

### **Procédure et matériel**

#### ***Objectifs et hypothèses***

Cette expérience propose de tester différents modes de présentation et de combinaison de l'information. Trois hypothèses guident ce travail.

H1 : La présentation utilisant des modalités complémentaires (image visuelle et texte sonore) sera plus efficace que celle présentant le matériel sous une modalité exclusivement visuelle (image + texte écrit).

H2 : La présentation d'images dynamiques permet un encodage plus efficace que les images statiques.

H3 : La redondance synchrone de l'information verbale (son + texte) augmente la charge cognitive et détériore l'apprentissage.

### ***Le matériel utilisé***

Le contenu des leçons présentées abordait différentes thématiques d'aéronautique extraites d'un livre<sup>1</sup> de formation présentant la description de divers éléments comme l'aile, la gouverne de profondeur, l'horizon artificiel, les ailerons, etc. et l'explicitation de principes généraux de vol comme la résultante aérodynamique, l'angle d'incidence, l'assiette de l'avion, etc.

L'ensemble de ces connaissances est nécessaire pour l'interprétation de situations typiques de vol, pour la compréhension de vols spécifiques d'un avion et pour son pilotage. Elles se composent pour reprendre la classification classique (Richard, 1995) à la fois de connaissances déclaratives (identification des éléments clés de la situation et compréhension du principe général de fonctionnement) et de connaissances procédurales (connaître les commandes pertinentes face à telle ou telle situation typique et en déduire les procédures à mettre en œuvre).

La leçon se compose d'environ 17 écrans informatiques qui s'enchaînent automatiquement (le sujet ne peut pas contrôler le déroulement du dispositif). Elle a toujours le même contenu mais celui-ci est présenté selon des modalités différentes (groupes indépendants). Chaque condition est équivalente en terme de quantité d'informations (même nombre d'items) et elles ont toutes été pré-testées (un groupe de 6 personnes par condition) pour vérifier le temps de lecture moyen et pour supprimer de manière générique tous les détails pouvant perturber le sujet (adaptation de la taille des schémas, du sous-titrage des légendes, du choix des mots, etc.). Suite à ce pré-test, un certain nombre de modifications a été apporté de manière similaire pour toutes les conditions dans le but de clarifier la leçon (fléchage automatique pour identifier clairement l'élément décrit, masquage de texte ou de parties non pertinentes pour la compréhension, etc.) et pour améliorer la coordination sur chaque écran des différentes informations multimodales (texte, son, image).

### ***Variables expérimentales***

Six protocoles de présentation des informations ont été conçus à partir de deux variables indépendantes intersujets.

La VI1 concerne les images présentées. Ces images représentent des schémas d'ailes d'avion, des représentations simplifiées d'avions en situation de vol et de ses différents mouvements, des instruments de bord comme l'horizon artificiel, etc. Les images sont présentées selon deux modalités :

- images dynamiques : la présentation des mouvements à comprendre se déroule sous forme d'une animation imagée continue sur un même écran ;
- images séquentielles : la présentation du mouvement s'effectue par séquence de différentes images fixes.

La VI2 concerne le support verbal accompagnant les leçons avec trois modalités :

---

<sup>1</sup> Les leçons présentées reprennent les thèmes de formation « Généralités » et « Les instruments de bord » pour l'obtention du Brevet de Pilote Privé et du Brevet de base en aéronautique, extraits du fascicule *Manuel du Pilote d'avion : Vol à vue*, Service Formation Aéronautique et du Contrôle Technique (SFACT), Cepadué Editions, Toulouse, 5<sup>ème</sup> édition, 1992.

- *textuelle* : les informations verbales sont présentées sous forme textuelle et les phrases apparaissent en bas de l'écran (sous la figure ou le schéma). Ces informations sont synchronisées avec l'apparition des différentes images (dynamiques ou séquentielles). Il y a environ une centaine de mots par écran, soit en moyenne 800 caractères. Le texte global n'apparaît cependant que partiellement et en relation étroite avec les images. En effet lorsque l'information verbale n'est plus adéquate avec l'information visuelle de l'écran, un système de masque cache les autres éléments du texte. Il n'y a donc de manière permanente qu'une phrase à l'écran ;
- *auditive* : Les informations verbales accompagnant les images sont présentées de manière auditive et synchrone avec la présentation visuelle correspondant ;
- *textuelle et auditive* : cette condition réunit les deux types de présentation textuelle et auditive (effet de redondance). Un principe de surlignement continu met en évidence le déroulement visuel de la phrase en relation directe avec sa diction sonore (effet « karaoké »).

Une variable intrasujet permet de mesurer les acquisitions de connaissances (pré-test/post-test). Les variables dépendantes correspondent à la somme des bonnes réponses au pré-test et post-test et à leur articulation (mesures répétées).

Les variables contrôlées sont le niveau d'étude des sujets (2<sup>ème</sup> année de Psychologie), le niveau de connaissance en aéronautique (novice tant au niveau théorique que pratique c'est-à-dire n'avoir jamais piloté d'avion, de planeur, de delta plane ou effectué de vol à voile, être ignorant de tout principe de vol, etc.) et le type de Bac (tous bacs hors discipline scientifique).

### ***Participants de l'expérience et modalités de passation***

L'expérience durait environ une heure. Les participants étaient répartis en six groupes correspondant aux six conditions expérimentales. Soixante six sujets ont participé à cette expérience. Le plan d'expérience est donc le suivant : S11 <Image2 \* Verbal3> \* Test 2.

À leur arrivée les sujets étaient avertis du but de l'expérience. Une première phase consistait à mesurer l'empan mnésique des sujets pour tester l'équivalence des groupes. Trois types de mesure étaient effectués.

Une *mesure de l'empan verbal écrit*. Cette épreuve est réalisée à l'aide de paires de suites de mots sans lien sémantique entre eux (cf. figure 1). Ce sont des mots bi-syllabiques de cinq lettres écrits en lettres majuscules noires. La taille de caractère correspond à du times point 28. Les mots sont disposés selon la longueur de la suite sur une ou deux lignes, au centre d'une feuille cartonnée A4. Nous avons réalisé deux listes de mots qui utilisent des mots identiques mais disposés différemment. Il n'y a pas de redondance d'un même mot dans une liste. Le nombre de mots dans chaque suite présentée va en s'accroissant (de 3 à 11 mots toujours disposés sur une seule ligne). Les deux listes sont présentées alternativement à tous les sujets avec un contre-balancement d'ordre. On commence par présenter au sein des deux listes les suites les plus courtes. Les mots sont présentés à la cadence régulière d'un mot par seconde (i.e. pour  $n = 3$  mots la durée de présentation est de 3

secondes). On demande au sujet de reproduire par écrit la suite de mots immédiatement après sa présentation.

Une *mesure de l'empan imagé* réalisée à partir de suites de dessins (cf. figure 1) réalisés à l'encre noire, composés de six figures géométriques de base (rond, carré, losange vertical, losange horizontal, triangle avec sommet vers le haut ou vers le bas) combinées à l'intérieur de 8 éléments typiques (2 traits pleins parallèles horizontaux ou verticaux, 2 traits pointillés parallèles horizontaux ou verticaux, deux gros points noirs disposés horizontalement ou verticalement, deux traits pleins formant le signe + et deux traits pleins formant le signe X). Lors d'un pré-test nous avons retenu 24 figures typiques parmi les 48 combinatoires possibles. La taille des figures est identique (3 cm de diamètre). Nous avons là aussi réalisé deux listes, présentées alternativement au sujet, composées de suites de figures différentes où n augmente progressivement de 3 à 11. Les suites sont agencées au centre d'une feuille cartonnée (format A4). Après chaque présentation le sujet doit reproduire la suite qu'il a visualisée. Pour ce faire, il dispose de tas de figures (format 5 \* 4,5 cm) représentant les 24 figures possibles qu'il doit sélectionner pour reconstruire la suite. L'organisation de la disposition des figures a fait l'objet d'un pré-test pour permettre au sujet de retrouver facilement les figures (les figures de base identiques sont disposées sur une même ligne et les éléments intérieurs identiques sont disposés sur une même colonne). Les suites sont présentées à la cadence de un dessin pour 3 secondes.

Une *mesure de l'empan auditif* : cette épreuve suit le même principe pour la construction du matériel et la passation que la mesure de l'empan mnésique écrit. Nous avons seulement utilisé 63 mots nouveaux de même composition (5 lettres, bi syllabiques). Les temps inter-mots sont réduits au maximum afin d'empêcher une répétition mentale de chaque mot entendu.

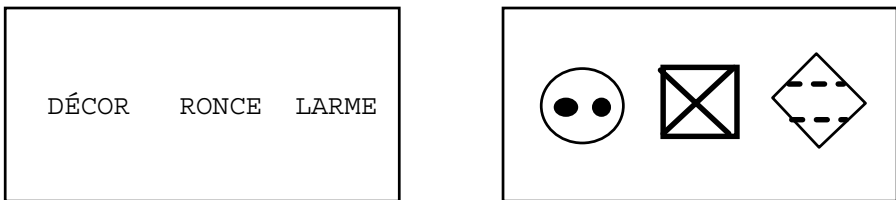


Figure 1. Exemples d'items présentés lors de la mesure de l'empan verbal écrit et imagé.

Lorsque le sujet commet deux erreurs consécutives, l'épreuve est terminée. L'empan correspond à la valeur n de la suite la plus longue que le sujet a été capable de rappeler dans l'ordre.

Suite à cette première phase, les participants étaient soumis à un pré-test pour évaluer leur niveau de connaissances dans le domaine de l'aéronautique. Les sujets effectuaient ensuite une tâche de barrage qui consistait à lire et à mémoriser le nombre de mots féminins parmi une liste de 106 mots. Puis ils étaient conduits dans un box pour assister à la leçon multimédia (durée 6 minutes). Dix minutes après la leçon, ils étaient invités à répondre au post-test (temps libre).



Le pré-test et le post-test sont composés de 26 questions<sup>2</sup> à choix multiples. Chaque question a quatre réponses possibles dont une est « je ne sais pas ». À chaque question il y a une seule bonne réponse cotée 1 point. Les mêmes questions sont utilisées pour le pré-test et le post-test, mais leur ordre chronologique ainsi que les différentes propositions à l'intérieur de chaque question ont été changées aléatoirement pour éviter tout transfert du pré-test au post-test. Ces deux questionnaires ont été prétestés pour s'assurer de la compréhension lexicale de tous les items. La structuration des questionnaires est réalisée de telle manière que le contenu d'une question ne peut pas aider le sujet à trouver la solution à une autre question par effet de report. Les 26 questions se répartissent de la manière suivante :

- 12 sont relatives à des définitions d'éléments d'avion utiles au vol (i.e. « gouverne de profondeur », « ailerons », etc.) ou à des principes explicatifs du vol (i.e. notions « d'incidence », « d'assiette », « de formes de l'aile », etc.). Ce sont des *questions de rappel de terminologie* ;
- 8 font référence à des situations générales de vol et concernent l'interaction entre au moins deux paramètres de pilotage. Elles évoquent la *compréhension de règles de base de pilotage* (i.e. « À incidence constante, une augmentation de la vitesse aura pour effet de ... » ; « Pour maintenir la portance constante en augmentant la vitesse, vous devez agir sur ... ») ;
- 6 sont relatives à la lecture d'instruments de bord présentés dans différentes situations de vol. On fournit ainsi le schéma d'un horizon artificiel et le sujet doit dire si l'avion est en train de piquer ou de cabrer ou on montre une position d'avion et 4 différentes vues d'horizon artificiel et le sujet doit dire à quelle situation de l'instrument de bord correspond l'avion, etc. Ces questions correspondent à des hauts niveaux de *compréhension de situations spécifiques de vol*.

Nous avons aussi élaboré une épreuve de 4 questions absentes lors du pré-test que nous avons nommées *résolution de problèmes pratiques*. Ces questions abordent par l'intermédiaire d'une description d'un exemple de vol les notions développées dans la leçon. Les sujets doivent ainsi décrire l'évolution des indicateurs des instruments de bord et les opérations à effectuer à partir de manœuvres d'avions (i.e. on indique une position de départ de l'avion en vol sur un horizon artificiel et une position à atteindre sur un second et on demande au sujet d'expliciter les manœuvres à effectuer, on leur demande aussi de construire les données de l'horizon artificiel à partir de la vue extérieure de la cabine de pilotage, etc.). Ces questions très complexes nécessitent de construire des inférences à partir de certains indices pour résoudre le problème opérationnel posé. Elles élaborent une synthèse générale des notions abordées durant la leçon.

## Résultats

Les résultats sous forme de tests d'analyse de variance ont été obtenus à l'aide du logiciel statistique SPSS. Nous n'avons pas obtenu de différences significatives entre les différents groupes expérimentaux sur l'ensemble des mesures d'empan

---

2 Le contenu de ces questions reprend ou est adapté des épreuves théoriques pour l'obtention du brevet de pilote.

mnésique, nous centrerons prioritairement notre analyse sur l'influence des différentes modalités de présentation de l'information sur l'apprentissage. Afin de tenir compte des caractéristiques intergroupes et de l'effet d'apprentissage entre le pré-test et le test final (facteur intra-groupe), les analyses effectuées sur les épreuves du questionnaire (pré-test & post-test) seront réalisées à partir de plans mixtes (i.e. épreuve globale et épreuves spécifiques). L'épreuve résolution de problèmes n'étant pas présente lors du pré-test, sera analysée à partir d'un plan inter-sujets.

### **Score global aux épreuves de rétention et de compréhension**

Pour les épreuves du questionnaire présentes à la fois lors du pré-test et du post-test (26 questions), toutes les conditions améliorent significativement en plan intra-groupe les scores de connaissances (au seuil .0000). Quelle que soit la modalité de présentation de l'information, le support multimédia permet un apprentissage satisfaisant. Ce résultat n'est cependant pas surprenant, compte-tenu de la faiblesse du niveau de connaissance de départ des sujets.

Conditions expérimentales	Score au pré-test (sur 26)	Score au post-test (sur 26)
G1 : Dynamique + son	1,36	16,27
G2 : Dynamique + son + texte	0,91	10,91
G3 : Dynamique + texte	2,18	12,45
G4 : Séquentiel + son	3,18	16,18
G5 : Séquentiel + son + texte	2,91	12,27
G6 : Séquentiel + texte	1,55	13,09
Moyenne	2,02	13,53

Tableau 1. Scores des différents groupes expérimentaux au pré-test et post-test.

### **Effet des présentations imagées sur l'apprentissage**

Les groupes avec la présentation imagée dynamique ne diffèrent pas significativement des groupes avec la présentation imagée séquentielle.

### **Effet des modalités du support verbal sur l'apprentissage**

Nous obtenons un effet principal significatif pour le score global ( $F(2,60) = 16,28$  ;  $p < .000$ ). La modalité « auditive » obtient de meilleures performances (M audio = 16,23 ; M textuelle = 12,77 et M audio textuelle = 11,59) (cf. tableau 2).

Pour l'épreuve « terminologie », on constate là aussi un effet principal significatif ( $F(2,60) = 3,64$  ;  $p < .03$ ). La modalité « auditive » obtient de meilleurs scores (8,18) que les modalités « textuelle » (6,82) et « audio textuelle » (6,5).

Pour l'épreuve « bases théoriques de vol », nous n'obtenons pas de différences significatives entre les conditions. Seul un effet tendanciel apparaît ( $F(2,60) = 3,11$  ;  $p < .08$ ) entre la modalité « auditive » (3) et la modalité « audio textuelle » (2,32).

Pour l'épreuve « compréhension », l'analyse révèle un effet principal significatif ( $F(2,60) = 10,06$  ;  $p < .000$ ). L'analyse des contrastes intergroupes montre une différence de performance significative entre le groupe « audio » et le groupe

« audio textuel » ( $F(1,42) = 17,34$  ;  $p < .000$  ;  $M = 5,05$  vs  $2,82$ ) et entre le groupe « audio » et le groupe « textuel » ( $F(1,42) = 18,21$  ;  $p < .005$  ;  $M = 5,05$  vs  $3,23$ ).

Modalités de la présentation verbale	Épreuve globale (sur 26)		Terminologie (sur 12)		Bases théoriques de vol (sur 8)		Compréhension spécifique (sur 6)		Résolution de problèmes (sur 7)	
	Moy.	Écart type	Moy.	Écart type	Moy.	Écart type	Moy.	Écart type	Moy.	Écart type
Audio	16,23	7,61	8,18	2,53	3	2,19	5,05	1,18	3,23	1,99
Audio + textuelle	11,59	9,96	6,5	3,4	2,32	2,41	2,82	4,91	2,14	2,31
Textuelle	12,77	9,42	6,82	2,63	2,73	3,06	3,23	2,18	1,77	3,61
Moyenne	13,53	12,65	7,17	3,31	2,68	2,68	3,7	3,63	2,38	2,94

Tableau 2. Scores moyens et variances des différentes modalités de présentation verbale de l'information.

Pour la résolution de problèmes, on constate là aussi un effet principal ( $F(2,60) = 4,8$  ;  $p < .01$ ). La décomposition de cet effet montre que la présentation « audio » se différencie significativement de la présentation « audio textuelle » ( $F(2,60) = 4,8$  ;  $p < .02$  ;  $M = 3,23$  vs  $2,14$ ) et de la condition « textuelle » ( $F(2,60) = 4,8$  ;  $p < .004$  ;  $M = 3,23$  vs  $1,77$ ).

### Effet de l'interaction entre les modalités imagées et les modalités de support verbal

Pour les scores de l'épreuve globale, l'analyse ne révèle aucun effet d'interaction (effet seulement tendanciel  $< .1$ ) (cf. tableau 3). Les comparaisons par paires révèlent cependant des différences significatives à l'intérieur de la modalité dynamique : G1 (audio) est significativement supérieur à G2 (audio-textuel) ( $F(1,20) = 21,22$  ;  $p < .000$  ;  $M = 16,2$  vs  $10,91$ ) et à G3 (textuel) ( $F(1,20) = 21,22$  ;  $p < .000$  ;  $M = 16,2$  vs  $12,45$ ). Au sein de la modalité séquentielle l'analyse révèle aussi des différences significatives : G4 (audio) est significativement supérieur à G5 (audio-textuel) ( $F(1,20) = 13,65$  ;  $p < .001$  ;  $M = 16,18$  vs  $12,27$ ) et G6 (textuel) obtient aussi des scores significativement plus performants que G5 (audio-textuel) ( $F(1,20) = 6,28$  ;  $p < .02$  ;  $M = 13,09$  vs  $12,27$ ).

Pour les épreuves de « terminologie » et celle de « bases théoriques de vol », on ne constate aucun effet général d'interaction entre les deux variables et les comparaisons intergroupes ne révèlent aucun effet significatif.

Pour l'épreuve de compréhension de situations de vol spécifiques, aucun effet général d'interaction n'est obtenu. L'analyse des comparaisons par paires fait cependant apparaître des différences significatives pour les modalités dynamiques entre : G1 (audio) et G2 (audio-textuel) ( $F(1,20) = 14,76$  ;  $p < .001$  ;  $M = 5,09$  vs  $2,55$ ) et entre G1 et G3 (textuel) ( $F(1,20) = 22,52$  ;  $p < .000$  ;  $M = 5,09$  vs  $2,73$ ). Pour les modalités séquentielles on constate aussi une différence significative entre G4 (audio) et G5 (audio-textuel) ( $F(1,20) = 4,66$  ;  $p < .04$  ;  $M = 5,06$  vs  $3,09$ ).

On constate aussi une différence significative entre G6 (séquentiel/textuel) et G3 (dynamique/textuel) ( $F(1,20) = 5,89$  ;  $p < .02$  ;  $M = 3,73$  vs  $2,73$ ).

Pour l'épreuve résolution de problèmes on ne constate aucun effet d'interaction. Les comparaisons par paires font apparaître des effets significatifs

uniquement au sein des modalités séquentielles. Ainsi G4 (audio) ( $M = 3,27$ ) obtient des scores plus performants ( $p < .000$ ) que G5 (audio-textuel ;  $M = 2,36$ ) et G5 ( $M = 2,36$ ) est plus efficace que G6 (textuel ;  $M = 1,18$  ;  $p < .04$ ). On constate aussi une supériorité ( $p < .02$ ) de la présentation textuelle dynamique (G3 ;  $M = 2,36$ ) sur la présentation textuelle statique (G6 ;  $M = 1,18$ )

Modalités de la présentation verbale	Épreuve globale (sur 26)		Terminologie (sur 12)		Bases théoriques de vol (sur 8)		Compréhension spécifique (sur 6)		Résolution de problèmes (sur 7)	
	Moy.	Écart type	Moy.	Écart type	Moy.	Écart type	Moy.	Écart type	Moy.	Écart type
G1 : Dynamique + son	16,27	6,81	8	3,2	3,18	1,36	5,09	1,09	3,18	3,16
G2 : Dynamique + son + texte	10,91	12,09	6,36	3,85	2,09	2,49	2,55	4,27	1,91	3,29
G3 : Dynamique + texte	12,45	7,27	7,09	2,69	2,64	2,05	2,73	2,41	2,36	5,05
G4 : Séquentiel + son	16,18	9,16	8,36	2,05	2,82	3,16	5,06	1,4	3,27	1,01
G5 : Séquentiel + son + texte	12,27	7,81	6,64	3,25	2,55	2,47	3,09	5,089	2,36	1,45
G6 : Séquentiel + texte	13,09	12,29	6,55	2,67	2,82	4,36	3,73	1,61	1,18	1,76
Moyenne	13,53	12,65	7,17	3,31	2,68	2,55	3,7	3,63	2,38	2,94

Tableau 3. Scores moyens et variances des différentes modalités de présentation de l'information en co-référence aux différentes épreuves.

## DISCUSSION

Quelles conclusions générales peut-on tirer de ces différents résultats ?

En réponse à nos hypothèses, H2 est infirmée. En effet, nous n'observons pas d'effets bénéfiques significatifs des images dynamiques (animation en mouvement continu) par rapport aux présentations séquentielles. En comparant deux types d'animation différents (en continue ou en séquences), cette recherche montre que la présentation de l'évolution des informations imagées par séquences suffit à mettre l'accent sur les changements d'états les plus importants pour permettre une bonne compréhension des phénomènes présentés (cf. Rieber, 1990).

Par contre, H1 et H3 sont confirmées. En effet, globalement sur l'ensemble des épreuves, les présentations du matériel verbal sous forme sonore s'avèrent plus performantes que les présentations verbales textuelles (H1). Lorsque l'information est présentée sous forme auditive, les sujets produisent de meilleures performances que dans la situation où la présentation est textuelle. On retrouve les principaux résultats de Mousavi, Low et Sweller (1995) et de Mayer et Moreno (1998). Ce résultat s'explique si l'on considère que le matériel écrit implique une présentation visuelle (graphique) qui apparemment interfère avec la visualisation nécessaire des images pour mémoriser et comprendre la leçon (interférence entre deux traitements visuels qui perturberait l'encodage).

La dégradation des performances lors de la condition de présentation du matériel verbal sous la modalité « textuelles et auditive » (H3) pourrait s'expliquer d'après le modèle de la mémoire de Baddeley (1992) dans lequel les mots lus sont rapportés dans la mémoire de travail visuel (calepin visuo-spatial) puis ensuite transformer en sons dans la mémoire de travail auditive (boucle phonologique). Une présentation simultanée de sons et de textes créerait ainsi un phénomène de double interférence. Le matériel textuel impliquerait un traitement visuel qui interférerait avec la visualisation de l'image mais il impliquerait aussi une transformation du message écrit en sons qui interférerait avec l'information sonore. Cette redondance surchargerait la mémoire de travail et dégraderait l'encodage de l'information.

Suite à cette étude, on retiendra principalement que quel que soit le type de tests, les modalités de présentation les plus efficaces sont celles qui intègrent une explication sonore avec la visualisation des schémas (dynamiques ou séquentiels). Cependant, les écarts types différents entre les groupes montrent que des approches plus centrées qualitativement sur les performances individuelles sont aussi nécessaires. En effet, malgré les tests d'empan réalisés pour tester les équivalences des groupes, l'apprentissage n'est pas homogène au sein des différentes conditions. Ces aspects méritent d'être approfondis pour mieux mettre en évidence les processus individuels de traitement des différents média selon les caractéristiques des tâches.

## BIBLIOGRAPHIE

- Baddeley A. (1992). *La mémoire humaine théorie et pratique*, Grenoble : Presse Universitaire de Grenoble.
- Bétrancourt M. (1996). *Facteurs spatiaux et temporels dans le traitement cognitif des complexes texte-figure*, Institut National Polytechnique de Grenoble, Thèse de Doctorat.
- Bieger G. R. & Glock M. (1985). « Comprehending spatial and contextual information in picture-text instructions », *Journal of Experimental Education*, vol. 54, n° 4, p. 181-188.
- Brooks L. R. (1970). « An extension of the conflict between visualisation and reading », *Quarterly Journal of Experimental Psychology*, vol. 22, n° 2, p. 91-96.
- Bruce V. & Green P. (1993). *La perception visuelle : physiologie psychologie et écologie*, Grenoble : Presse Universitaire de Grenoble.
- Clark J. M. & Paivio A. (1991). « Dual coding theory and education », *Educational Psychology Review*, n° 3, p. 149-210.
- De Vito C. & Olson A. (1973). « More imagery and the recall of adjectives and nouns from meaningful prose », *Bulletin of the Psychonomic Society*, vol. 1, n° 6-A, p. 397-398.
- Denis M. & De Pouqueville P. (1976). « Realism of figuration and memory for concrete actions », *Bulletin de Psychologie*, n° 30, p. 543-550.
- Dubois M. & Vial I. (1998). « Co-référentialité des informations dans une tâche d'apprentissage d'un vocabulaire étranger », in *Actes du VII<sup>e</sup> colloque de l'Association pour la Recherche Cognitive*, Université de Paris 8, 11, 12 décembre, Paris, p. 199-204.
- Dwyer F. M. (1989). « The effectiveness of visual illustrations used to complement programmed instruction », *Journal of Psychology*, n° 70, p. 158-161.

- Feldman A. & Schoenwald R. (1989). « Development of a computer-based instructional system in pharmaco-kinetics: efficacy in clinical pharmacology teaching for senior medical students », *Journal of Clinical Pharmacology*, vol. 29, p.158-161.
- Fletcher J. (1990). *Effectiveness and cost of interactive videodisc instruction in defense training and education*, Washington : Institute for Defense Analysis.
- Kaiser M. K., Proffitt D. R. & Anderson K. (1985). « Judgements of natural and anomalous trajectories in the presence and absence of motion », *Journal of Experimental Psychology: Learning Memory Cognition*, vol. 11, p. 795-803.
- Levin J. R. & Divine-Hawkins P. (1974). « Visual imagery as a prose-learning process », *Journal of Reading Behavior*, vol. 6, p. 23-30.
- Mandl H. & Levin J. R. (éds) (1989). *Knowledge acquisition from text and pictures*, Amsterdam : North Holland.
- Marmolin H. (1992). « Multimedia from perspective of psychology », in Lars Kjell Dahl (éd.), *Multimedia principles systems and applications*, Springer-Verlag.
- 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, p. 89-401.
- Mayer. E. M. & Moreno R. (1998). « A split attention effect in multimedia learning: evidence for dual processing systems in working memory », *Journal of Experimental Psychology*, vol. 90, n° 2, p. 312-320.
- Mousavi S. Y., Low R. & Sweller J. (1995). « Reducing cognitive load by mixing auditory and visual presentation modes », *Journal of Educational Psychology*, vol. 87, n° 2, p. 319-334.
- Paivio A. (1986). *Mental representations: a dual coding approach*, Oxford : University Press.
- Peeck J. (1987). *The role of illustration in processing and remembering illustrated text*, vol. 1, chap. 4, Willows & Houghton.
- Proffitt D. R. & Kaiser M. K. (1986). « Software techniques the use of computer graphics in motion perception research », *Behavior Research Methods Instruments & Computers*, vol. 18, n° 6, p. 487-492.
- Richard J.-F. (1995). *Les activités mentales : comprendre, raisonner, trouver des solutions*, Paris : Armand Colin.
- Rieber L. P. (1990). « Animation in computer-based instruction », *Educational Technology Research and Development*, n° 38, p. 77-86.
- Rusted. R. & Coltheart V. (1979). « The effect of pictures on the retention of novels words and prose passage », *Journal of Experimental Child Psychology*, vol. 28, p. 516-524.
- 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, p. 176-192.
- Vernon M. D. (1953). « The value of pictorial illustration », *British Journal of Educational Psychology*, n° 23, p. 180-187.
- Winn W. (1991). « Learning from maps and diagrams », *Educational Psychology Review*, n° 3, p. 211-247.