

Réflexions linguistiques et sémiologiques pour une écriture didactique du multimédia de langues

Anne-Laure FOUCHER

Université Paris 6 , France

Résumé : *Dans une combinaison originale de domaines souvent envisagés de manière fermée que sont la didactique des langues, la linguistique, la sémiologie et l'ALAO (Apprentissage des Langues Assisté par Ordinateur), nous mettons en avant la nécessité de concevoir l'image multimédia comme un objet composite, lieu véritable d'interactions entre modes iconique, auditif et linguistique. De cette option théorique découle, selon nous, une philosophie qui préside à la conception d'un multimédia pour l'apprentissage des langues. Nous en donnons les grandes lignes sous forme de jalons méthodologiques destinés à une écriture du multimédia en langues.*

- [1. Introduction](#)
- [2. Enseignement/apprentissage des Langues et multimédia](#)
- [3. "L'image multimédia" est un lieu d'interactions multimodales](#)
- [4. Implications pour une philosophie de conception du multimédia de langues](#)
- [5. Remarques méthodologiques pour une écriture du multimédia](#)
- [6. Préalables méthodologiques relatifs à l'introduction de l'image dans l'image multimédia](#)
- [7. Quelques jalons méthodologiques pour un cadre d'écriture didactique du multimédia de langues](#)
- [8. Conclusion](#)
- [Références](#)

1. Introduction

Lors d'une table ronde en 1996 autour du thème "Environnements interactifs pour l'apprentissage des langues", était soulignée l'importance, pour qui désirait exploiter pédagogiquement des documents hypermédiés, d'un passage par les trois étapes suivantes : la première étant la recherche de produits susceptibles de plaire aux apprenants ; la seconde consistant en un repérage dans les produits des parties "dans lesquelles le rapport texte/son/image [favoriserait] une compréhension aisée et l'envie de s'exprimer ; ce repérage [supposant] un minimum de connaissances sémiologiques." (Chanier & al, 1996 : p. 252) ; la dernière étant consacrée à la conception de "tâches amenant l'apprenant à manipuler (en compréhension comme en production) des données sonores, visuelles et graphiques" (Idem : p. 252). Parallèlement à ces considérations et plus récemment, T. Lancien suggérait qu'on optimise la multicanalité [1] dans le multimédia [2], multicanalité qu'il définit comme "le fait que coexistent sur un même support différents canaux de communication" et qui "ne prend de sens que selon les choix que fait la personne qui les consulte" (Lancien, 1998 : p. 24), cela afin de favoriser l'apprentissage de la langue. Il ajoutait que la multicanalité d'un multimédia devait être évaluée en considérant la nature, la pertinence et la richesse des liens entre les médias.

Que l'on se place du point de vue de l'utilisateur de produits hypermédias déjà constitués et où il s'agit de repérer les parties d'un produit hypermédia dont l'organisation des modes favoriserait certains types d'activités, ou que l'on se range plutôt du côté du concepteur de produits considérant la multicanalité comme un apport fécond pour l'apprentissage et dont l'objectif est alors d'optimiser cette même multicanalité, mobiliser des connaissances d'ordre à la fois sémiologique et linguistique se révèle être un passage obligé.

Pourtant, le domaine des TIC (Technologies de l'Information et de la Communication) pour les langues étrangères laisse encore aujourd'hui bien peu la place à une réflexion didactique sur la mise en forme des contenus en rapport avec la spécificité du support multimédia[3]. Or s'intéresser à l'approche multimédia ou hypermédia, c'est-à-dire introduire, dans une logique uniquement formatrice, des messages multimodaux tels que son, texte et image pour des séquences d'apprentissage, c'est minimalement s'interroger sur les rapports que peuvent entretenir les modes en question et tout particulièrement ceux qu'entretiennent l'iconique et le linguistique en tant que deux modes de signifier spécifiques : le support multimédia est au service d'un projet didactique, d'objectifs pédagogiques et non l'inverse.

Nous reviendrons rapidement sur la problématique de l'image et du multimédia pour l'enseignement des langues et montrerons combien il est fructueux de considérer l'image multimédia comme un lieu d'interactions entre les différents modes en présence dans un produit. Ces considérations nous amèneront à définir les implications d'une telle optique dans la philosophie d'écriture d'un multimédia pour les langues, implications que nous déclinons en propositions méthodologiques à l'usage des concepteurs.

2. Enseignement/apprentissage des Langues et multimédia

De même que l'on a paré l'audiovisuel (l'image fixe puis la vidéo) de toutes sortes de vertus pour l'apprentissage, on fait aujourd'hui du multimédia une panacée pour l'enseignement/apprentissage des langues. Le multimédia étant encore largement considéré comme une addition de textes, de sons et d'images (au sens large), on lui prête les qualités des modes qu'il met en présence. Pourtant toutes les qualités de la vidéo, fonction motivationnelle (Ginet, 1997 ; Stratfold, 1994), fonction rhétorique (Laurillard, 1993), fonction de relais de l'expérience des apprenants (Stratfold, 1994), etc. combinées avec celles d'autres modes (textes, sons, etc.) ne donnent pas la spécificité véritable du multimédia. Il est évident que le multimédia ne peut être conçu uniquement sur la base d'une pure et simple juxtaposition des qualités attribuées aux modes qui le composent. De même, on ne peut dire seulement que le multimédia, c'est une maniabilité et un relatif contrôle offerts à l'utilisateur.



Quand on dit qu'articuler texte et image semble favoriser l'apprentissage (Mayer & Anderson, 1992), ce ne sont pas les termes "texte" et "image" qui sont importants ici mais bien le terme "articuler". Les problèmes couramment rencontrés par des apprenants utilisant le multimédia nous en fournissent la preuve. Utiliser le multimédia, ce n'est pas utiliser les modes qui le composent de manière isolée. Ainsi L. Plowman, ayant expérimenté un environnement multimédia interactif sur de jeunes enfants, explique que :

Learners are exposed to a relatively unfamiliar combinaison of media, such as moving and still video images, computer-generated text and graphics, and sound. The addition of a computer control to this combinaison of media is initially unfamiliar to many users and can lead to difficulties in navigation through a program. Transferring knowledge of the structural conventions of other media to interactive multimedia may make excessive demands on cognitive processing (Plowman, 1994 : p. 278).

En effet, lorsqu'un apprenant travaille sur un logiciel multimédia intégrant de la vidéo, il est amené à passer de la bande vidéo aux activités pédagogiques, de la bande vidéo aux différentes aides qu'on lui propose qu'elles soient lexicales, textuelles ou sous forme vidéo. Il doit donc effectuer des balancements entre des éléments filmiques (de fiction ou issus d'un document authentique) et des éléments pédagogiques. A chacun de ces sauts effectués, il y a rupture de la construction filmique en termes de causalité, de temporalité : la linéarité filmique est éclatée et il y a remise en cause des liens causaux et temporels. Or ce sont ces liens mêmes, organisés en un tout cohérent, qui guident la compréhension du sens. Il y a également rupture de la progression pédagogique : devoir passer aux séquences filmiques entre les activités pédagogiques ou même au sein d'une activité peut perturber la "saisie" des connexions, des liens faits entre ces activités.

De même, Lin (1994, cité dans Duquette & Renié, à paraître), comme L. Plowman (1996a), souligne que les utilisateurs d'hypermédias éprouvent des difficultés liées au non recours à des stratégies métacognitives adéquates (du fait même de leur perte de repères et du coût cognitif investi pour la compréhension de ces environnements) : "[Learning in hypermedia systems] is likely to be especially difficult for those users who lack skills planning, monitoring, reflecting, and regulating their personal learning strategies" (Lin, 1994 : p. 487) (pour des comptes-rendus d'utilisations de vidéo interactive dans des dispositifs d'anglais d'après des retours d'apprenants, cf. F. Pasquier, 1997 : p. 395 et suivantes).

Parallèlement à cela, et contrairement à la tendance actuelle qui est de donner aux apprenants, dans des dispositifs d'autoformation, des produits à "explorer", sans médiation pédagogique, où l'apprenant butine et chemine à sa guise (Chanier, à paraître), on est amené à constater la très faible utilisation d'un cédérom encyclopédique, par exemple, alors que des produits a priori peu modernes (produit "papier"), des didacticiels avec vidéos et des vidéos pédagogiques ont la faveur des apprenants et totalisent de nombreuses heures d'utilisation (Demaizière, 1996 ; Demaizière & Foucher, à paraître).

Toutes ces considérations nous conduisent à dire que le multimédia doit être pensé (en amont par le concepteur) en termes d'interactions des modes en jeu et non simplement comme une co-présence de ces modes, ce qui implique aussi qu'il est important de considérer que les stratégies mises en place par un apprenant lorsqu'il travaille sur une vidéo par exemple (de type journal télévisé ou film de cinéma) ne sont pas les mêmes que lorsqu'il manipule un multimédia (cf. Chanier & al, à paraître ; pour un panorama sur les recherches concernant multimédia et stratégies d'apprentissage, cf. Duquette & Renié, à paraître). Et face à des inconvénients énoncés comme celui-ci :

on ne peut pas ne pas se demander comme pour l'environnement textuel et hypertextuel des vidéos authentiques, si ces aides écrites (surtout le sous-titrage intégral et l'accès à un lexique) ne risquent pas de favoriser chez beaucoup d'apprenants et par solution de facilité, des pratiques de compréhension mot à mot, au détriment de stratégies de compréhension plus authentiques. (Lancien, 1998 : p. 56),

nous choisissons une optique de conception du multimédia différente et nous parions sur des apprenants qui prennent l'initiative et pensent leur apprentissage autrement qu'en termes de facilité (Demaizière, Foucher, à paraître).



La suite de notre texte se propose de reconsidérer les liens des médias introduits dans le multimédia pour que l'on passe d'une conception du multimédia encore largement discontinuée et fragmentée à une conception intégrative et interactive. Nous prendrons comme point d'entrée un mode spécifique, celui de l'image.

3. "L'image multimédia" est un lieu d'interactions multimodales

Notre réflexion prenant appui sur le mode spécifique qu'est l'image, nous proposons d'adopter d'emblée la terminologie d'"image multimédia", l'expression permettant d'évoquer à la fois l'idée, sous le terme d'image premièrement, de la présence d'images non seulement visuelles (dessins, photos, icônes, etc.) mais aussi sonores et textuelles et l'idée, sous le terme multimédia, que tous les modes sont à envisager de concert.

Ce que nous désirons mettre en avant ici n'est donc pas une typologie des images dans le multimédia mais bien les caractéristiques générales de ce que nous appelons "image multimédia", c'est-à-dire les caractéristiques de l'interaction, de la combinatoire de textes et d'images déployés sur un écran, le son y étant associé de plus en plus. Parler d'image multimédia ou s'intéresser aux images dans le multimédia ne se situe pas, selon nous, au même niveau d'analyse. Parler d'image multimédia, c'est envisager les images visuelles, sonores, textuelles à un niveau plus global d'interaction alors que, pour nous, considérer les images dans le multimédia est réducteur dans la mesure où seules les images visuelles sont prises en compte, dans leurs relations aux autres modes certes (on parlera le plus souvent alors de rapports image/langue), mais toujours dans un sens, celui qui va des images vers les autres modes.

Nous pensons, de plus, qu'une typologie des images dans le multimédia et a fortiori une typologie des

fonctions de l'image dans le multimédia n'est pas "arrêtable", "figeable" en ce sens que premièrement, l'image est de toutes façons associée à des modes divers dans des écrans de logiciels et ne l'est qu'en relation à des activités, à des tâches à effectuer et que deuxièmement, à la moindre manipulation de l'apprenant[4] le statut de l'image est modifié. Il est clair qu'introduire une séquence vidéo dans une image multimédia n'est pas équivalent à l'intégration d'un dessin animé dans une image multimédia mais il est, à notre sens, infiniment plus rentable et productif de dégager des spécificités générales de l'image multimédia car ce sont elles qui feront le socle d'une écriture didactique du multimédia.

Ces caractéristiques sont à prendre en compte dans la conception d'un logiciel de type multimédia ou à ne pas ignorer lors de l'introduction d'un tel multimédia dans une séquence d'enseignement/apprentissage de quelle que forme que ce soit. Pour mettre en lumière ces caractéristiques de l'image multimédia et proposer une définition, nous nous appuyons sur les points de vue qu'ont sur la nature de l'image visuelle et sur le son différentes disciplines telles que la linguistique, la sémiologie, la didactique des langues, la psychologie cognitive et les théories du cinéma, la filmologie et la sémiologie du cinéma[5].



En choisissant de ne pas faire de typologie de l'image dans le multimédia des langues, nous évitons d'une part de mener notre étude dans un seul sens, c'est-à-dire de l'image vers la langue principalement et, d'autre part, nous nous débarrassons de l'encombrant mythe du document authentique qui parasite encore le paysage didactique.

3.1. "Image multimédia"

L'image multimédia est, pour nous, une organisation multimodale, et donc perceptive (sur un écran d'ordinateur) où des images visuelles[6] sont en interaction avec d'autres modes (images sonores, paroles, etc.). Dans un logiciel multimédia d'apprentissage de langue, de type didacticiel ou non, l'apprenant est confronté à des images (vidéo ou non) imbriquées dans des écrans où sont combinés des modes faisant appel au sens de la vue et de l'ouïe. Les modes iconiques qui peuvent être présents sur les écrans sont des scènes visuelles animées et complexes (images vidéo couleur, etc.), des scènes visuelles complexes mais fixes, des graphiques, des dessins, etc., et de la "langue" écrite (textes et "images" de ces textes). Les modes auditifs peuvent être, quant à eux, du discours oral (oral spontané ou texte écrit et oralisé), des bruits et éventuellement de la musique.

En ce qui concerne les modes visuels et auditifs liés à la langue, on distingue le plus souvent trois sources énonciatives différentes : le discours oral des protagonistes de la bande vidéo (ou de la vignette dessinée ou simplement d'un échange sans que les interlocuteurs soient représentés de manière imagée), la transcription écrite et remaniée[7] de ce discours des personnages, le discours didactique qui comprend consignes, commentaires, résumés de contenus, transitions entre les différents points de contenu, etc. Ce discours didactique, dans certains produits, est parfois présenté sous forme orale notamment pour les commentaires. Ces différents modes sont, à l'écran, organisés de façon spatiale pour les modes visuels et de manière temporelle pour les modes auditifs.

L'image multimédia renvoie à l'image au sens de "texte visuel" (quand il n'y a pas de son), ou mieux au sens de texte multimodal, avec du son. Ce texte multimodal possède une construction propre de tous ses éléments constitutifs (bruits, musique, texte(s) écrit(s), "paroles" et images animées, fixes, etc.) avec une "syntaxe" différente de celle d'un texte purement linguistique ou d'un simple montage d'images fixes ou animées ou d'une juxtaposition des deux. Nous insistons ici sur les guillemets pour le mot syntaxe, M. Tardy (1973 : pp. 78-80) nous mettant en garde dans *Le professeur et les images*, sur ce "contresens ontologique" résultant d'un usage "douteux" de la notion de langage et menant certains pédagogues à faire de l'initiation aux messages visuels un apprentissage de la lecture et de l'écriture du visuel.

3.2. Caractéristiques de l'image multimédia

Nous donnons ici les caractéristiques principales de l'image multimédia, celles qui, selon nous, ont une incidence sur la création d'un multimédia de langues.



La multimodalité est la présence, sur un même écran de logiciel et dans le logiciel dans son entier, de différents modes iconiques (séquences vidéos, images visuelles, icônes, animations diverses), linguistiques (textes écrits ou interventions orales), auditifs (musique, sonorisations diverses). Nous utilisons donc l'expression "image multimédia" pour qualifier ce que d'autres appelleront "message" ou "texte pluricodique"[8] pour désigner un document "[...] où le sens des données langagières interfère constamment avec celui de données d'origine différente et travaillées par d'autres codes, telle l'image", ce peut être un spot publicitaire, (Minot, 1993 : p. 23) ou encore "formes pluricodées" ou "langages pluricodiques" (Boutan & al, 1988 : p. 27) . A l'intérieur de cette image multimédia, bien entendu on trouvera des séquences vidéo par exemple où les modes linguistique et visuel sont associés pour faire sens. Mais cette première combinaison de modes sera intégrée à une combinaison d'autres modes encore et donc plus large. On ne trouvera pas, dans l'image multimédia, de mode totalement isolé.

La multimodalité est toujours au moins à deux niveaux : on a, premièrement, une juxtaposition ou combinaison de différents modes et, deuxièmement, ces mêmes modes ont une nature multimodale (une séquence vidéo, par exemple, comprend images animées et son, les deux étant livrés conjointement). Dans l'image multimédia, il n'y aura pas d'image (visuelle) seule mais c'est, de toutes façons, rarement le cas par ailleurs : l'image (visuelle) fixe n'est pratiquement jamais seule, y sont joints courts textes, mentions diverses, légendes ou opérateur de lisibilité (Masselot-Girard, Jouvenot, 1994 : p. 39). Ainsi sous une photographie dans une exposition par exemple, on pourra trouver l'énoncé "sans titre" écrit noir sur blanc. De même, l'image animée n'est que rarement isolée, une bande-son y étant pratiquement systématiquement associée et c'est une loi du genre pour le domaine des langues : au cinéma, on privilégie la parole (on parle de verbo-centrisme parfois) mais il en est de même dans les vidéos destinées à l'apprentissage des langues, ce qui est tout à fait justifié ! On imagine difficilement apprendre une langue sans au moins l'entendre. " Pour retrouver des images données sans paroles, il faut sans doute remonter à des sociétés partiellement analphabètes, c'est-à-dire à une sorte d'état pictographique de l'image [...] " (Barthes, 1964 : p. 43), état qui ne pourrait être qu'antérieur à l'apparition du livre malgré les débats autour de notre entrée dans une soi-disant civilisation de l'image, au détriment de l'écriture, etc. De même, un texte (mode linguistique écrit) a une image globale donnée par la forme du texte dans son ensemble : " Le texte est une image, comme on peut s'en rendre compte en examinant le comportement d'enfants qui ne savent pas encore lire et feuilletent un livre illustré où le texte apparaît comme une image parmi d'autres, une image à déchiffrer à la lueur du contexte iconique." (Escarpit, 1976 : p. 121).

Multimodalité didactisée et à visée pédagogique

Nous distinguons l'image multimédia du texte pluricodique car là où dans un texte pluricodique (au sens donné plus haut), les sens véhiculés par les différents modes se combinent pour produire un sens nouveau, dans l'image multimédia (de produits de formation en langues), de la combinaison des différents modes n'émerge pas de sens nouveau mais de cette combinaison doit émerger un savoir (que l'apprenant s'appropriera par reconstruction). Cette combinaison de modes se fait dans un contexte d'apprentissage pour servir l'apprentissage de contenus à assimiler par le biais d'activités. L'image multimédia est donc un texte pluricodique où l'interaction des modes en jeu est dédiée à l'émergence, à plus ou moins long terme, chez l'apprenant, d'un savoir. A court terme et ponctuellement, pour ce même apprenant, ces interactions modales sont destinées à la mise en oeuvre d'échanges pédagogiques sous forme d'activités, de QCM, etc.



Pour définir un peu plus avant la nature de cette interaction des modes en jeu, nous rapprochons le logiciel multimédia (pour l'apprentissage de la langue) du manuel scolaire papier traditionnel considéré comme objet sémiolangagier ou écrit iconisé (déployé sur une double page). Le logiciel multimédia est, selon nous et par un certain côté, comme le manuel, "un produit discursif d'une énonciation de type analytique et outil de construction des schèmes mentaux qui permettent de traiter l'information et de construire les savoirs" (Masselot-Girard, 1994 : p. 4), énonciation de type analytique car sont communiqués les savoirs préalablement didactisés et outil pour l'apprenant qui reconstruit ces savoirs. Bien sûr, là où le manuel ne déploie que des blocs textuels et des images visuelles, pour le logiciel multimédia doivent être ajoutés deux dimensions supplémentaires : le son et les possibles manipulations ou transformations par l'apprenant de l'organisation de tous ces modes (visuel, linguistique, sonore).

L'interaction des différents modes en jeu dans l'image multimédia et plus largement dans le logiciel multimédia est donc (devrait être) le résultat d'une mise en forme didactique des modes sélectionnés dans un but de facilitation de l'apprentissage, de re-construction des savoirs par l'apprenant.

Multimodalité et intermodalité dynamitables

L'intermodalité est la relation qui s'instaure entre les différents modes à partir du moment où ces modes sont en présence : par exemple, l'apprenant peut se référer à une séquence vidéo (un mode visuel et auditif) pour répondre par écrit à une sollicitation et cela à plusieurs écrans de distance, se référer alors à ce mode pouvant être pertinent ou non pour mener à bien l'activité. Dans l'image multimédia, il y a des interactions de modes préalablement décidées par les concepteurs mais aussi des interactions de modes convoquées par l'apprenant selon ses besoins : en effet, même si celles-ci sont prévues par le concepteur, elles peuvent ne pas être mises en oeuvre. Il y a donc interactions multimodales dans l'image multimédia mais celles-ci sont variables et plus ou moins variées selon les modes proposés.

La caractéristique principale de l'image multimédia selon nous réside dans son organisation à tout moment dynamisable^[9] (Foucher, 1998). L'image multimédia est une organisation mouvante : un apprenant cliquera plus volontiers sur des aides visuelles, un autre convoquera à intervalles réguliers des aides textuelles ou lexicales ; un autre encore refusera de demander de l'aide sous quelque forme que ce soit ; un autre choisira enfin un type de mode en fonction de l'activité qu'il est en train de faire.

4. Implications pour une philosophie de conception du multimédia de langues

La variabilité de l'image multimédia, et il apparaît difficile aujourd'hui de ne pas proposer aux apprenants toutes les potentialités mentionnées ci-dessus, a évidemment des implications dans la philosophie qui préside à sa création, à son écriture, c'est-à-dire à la mise en oeuvre de tous les modes pouvant potentiellement entrer en jeu par une manipulation de l'apprenant. Elle implique également une conception spécifique de la liaison didactique des langues-TIC, de la langue et de ces rapports avec l'image notamment.



4.1. Interroger les rapports entre images et langue

Les caractéristiques que nous avons mises en lumière, à savoir essentiellement la variabilité de l'image multimédia, nous amènent à repenser les rapports que peuvent entretenir les images et la langue dans un logiciel multimédia. D'une part, compte tenu de ces caractéristiques, il est peu pertinent de concevoir ces rapports en termes de stabilité où une image serait associée à une séquence sonore de façon permanente par exemple. D'autre part, les rapports traditionnellement établis en didactique des langues entre image et langue (redondance et illustration principalement) sont bouleversés du fait même de cette "non-permanence" des interactions entre modes.

4.2. Interroger les rapports de l'image et du son

De même, la variabilité de l'image multimédia remet en question le rapport au son dans ses interactions avec les autres modes. On a l'habitude de dire que si le regard est une exploration spatiale et temporelle d'un "donné à voir", limité car maintenu dans le cadre de l'écran, l'écoute au contraire est l'exploration d'un "donné à entendre", qui est plus un "imposé à entendre", beaucoup moins délimité, avec des contours changeants, en tout cas au cinéma. En effet, du fait de la nature physique du son, de l'absence de paupières pour les oreilles, de l'omni-directionnalité de l'écoute et de l'absence d'une culture auditive véritable, on peut difficilement en exclure ou y sélectionner quoi que ce soit. Cette conception du son, valable pour le son au cinéma entre autres, ne l'est plus dans le cas du multimédia où le son peut être à loisir manipulé : baissé voire coupé complètement, sélectionné, et même ralenti dans certains logiciels visant un travail sur la prononciation.

On constate que souvent encore dans le multimédia, la mise en scène du son souvent privilégiée est celle d'un son "in", c'est-à-dire d'un son "visualisé" où l'on entend la voix d'un protagoniste en même temps qu'apparaît son image à l'écran. (Quitte parfois à détourner le dispositif physique de la communication, dans le cas par exemple d'un appel téléphonique où l'on entendra les deux protagonistes et où, sur un écran divisé en deux, chacun des protagonistes apparaît dans sa moitié d'écran avec un combiné à la main). Trois types de sons (nous plaçons musique et bruitages à part) peuvent être généralement isolés : des sons visualisés dans le cas des séquences vidéo ou de vignettes fixes et même animées représentant de personnages, la bande sonore portant le plus souvent les dialogues de ces personnages ; des sons "textualisés" où des textes sont oralisés ; des sons non-visualisés dans le cas, par exemple, de consignes orales émanant de l'instance pédagogique (qui n'est pratiquement jamais représentée, sauf sous forme

d'icône parfois). Ici se pose le problème du dispositif énonciatif mis en oeuvre dans le logiciel : oraliser un texte, c'est lui donner une voix mais qu'en est-il de la prise en charge des énoncés par une voix off, de l'identification par l'apprenant de cette voix off émanant d'une source invisible d'autres lieux, d'un autre temps ?

Nous pensons qu'une construction des rapports entre les différents modes selon des principes de type "visualisation / monstration" n'est plus valable (à supposer qu'elle l'ait été d'un point de vue didactique et pédagogique) pour l'image multimédia d'aujourd'hui : étant donné que la mise en sons peut être dynamitée par l'apprenant qui manipule ces sons (des commandes de type revoir sans le son ou réécouter le dialogue sont courantes), étant donné que les interactions entre les différents modes peuvent être soumises à des changements divers, on ne peut écrire l'image multimédia comme on a écrit les méthodes audiovisuelles, les méthodes de langue papier.



De plus, envisager l'image multimédia comme un lieu d'interactions entre modes permet d'éviter l'écueil d'un réalisme naïf, à un objet physique correspondrait un signe donné ainsi que l'écueil du béhaviorisme qui fait qu'à ce signe donné est associé un comportement.

Voyons maintenant quelles peuvent être les conséquences de ces caractéristiques de l'image multimédia pour la mise en oeuvre d'images multimédia dans un scénario pédagogique pour un logiciel d'apprentissage de langue étrangère.

5. Remarques méthodologiques pour une écriture du multimédia

Nous donnons ci-après quelques remarques méthodologiques qu'il nous semble important de considérer lorsque l'on conçoit un multimédia pour les langues.

5.1. Interaction de modes et convocation variable

Les termes d'image multimédia et de multimodalité définis, nous pensons qu'il est important de tenir compte des relations, des interactions qui s'instaurent inévitablement entre les différents modes à partir du moment où ces modes sont en présence, c'est ce que nous appelons intermodalité, et des interactions qui sont susceptibles d'être construites par l'apprenant, et cela parfois à plusieurs écrans de distance.

Dans l'image multimédia, certaines interactions entre les modes (texte de l'activité pédagogique avec consignes + séquence vidéo ; séquence vidéo + aide lexicale, etc.) sont préalablement décidées par le concepteur mais ces interactions de modes sont convoquées par l'apprenant selon ses besoins : des interactions prévues peuvent ne pas être mises en oeuvre par l'apprenant. C'est ce que nous avons appelé une multimodalité dynamitable (par l'apprenant). Cette variabilité des interactions fait que l'introduction d'un mode, visuel ou autre, ne doit plus être pensé en termes statiques mais en termes de convocation variable. Il est alors nécessaire, dans le scénario pédagogique envisagé, de ne pas faire "comme si" l'apprenant allait forcément traiter le mode en question dans le cadre d'une activité pédagogique mais, au contraire, de penser le mode introduit comme pouvant être appelé à plusieurs reprises dans le cadre de cette même activité ou ne pas être convoqué du tout.

5.2. Importance des repères

Nous avons vu qu'il n'y a pas de mode isolé dans l'image multimédia : on peut avoir à un premier niveau une juxtaposition ou une combinaison de différents modes qui ont par ailleurs une nature multimodale (ce qui constitue le second niveau).

Face aux différents modes présents dans l'image multimédia mais aussi entre ces différentes images multimédia tout au long du didacticiel, il nous semble particulièrement pertinent d'insister sur la présence de repères indiquant à la fois le niveau de profondeur auquel est arrivé l'apprenant dans le logiciel et la zone en "surface" dans laquelle il évolue. L'apprenant doit pouvoir obtenir à tout moment, s'il le désire, ces deux repères. Les liens référentiels, sous forme de redondances entre images par exemple, mais aussi les charnières modales (que nous abordons plus loin) sont des éléments de repérage.



6. Préalables méthodologiques relatifs à l'introduction de l'image dans l'image multimédia

Facilité de lecture de l'image, passivité du sujet qui "lit" cette image, appel à un mode de consommation établi de cette image, décryptage aisé dû à de longues expositions à des documents audiovisuels entre autres, etc. Des recherches à la fois en psychologie cognitive et en sémiologie ont depuis longtemps montré que ce sont là quelques présupposés qui, s'ils circulent encore en didactique des langues, ne sont guère recevables (Bresson, 1981 ; Eco, 1970 ; Thon & al, 1993 par exemple). En effet, reconstruction du réel, l'image visuelle présente trois dimensions : l'image est une réalité matérielle à valeur figurative, qui nous donne toutes les apparences (ou presque) ; une réalité esthétique et une réalité intellectuelle à valeur signifiante. Ne pouvant être définie uniquement par ses éléments visuels, elle est appréhendée de manière trans-sensorielle, l'affectif jouant un rôle important dans cette appréhension. Nous pensons que cette définition a des conséquences directes sur l'introduction de l'image dans une approche multimédia des langues.

6.1. Irréductibilité de l'image

La réduction d'une des trois dimensions de l'image visuelle est impossible : réduire une dimension de l'image ne diminue pas les deux autres mais réorganise les trois dimensions différemment pour donner une autre image aussi "chargée" sémantiquement. Nous estimons donc qu'une intervention didactique ou pédagogique sur le contenu même d'images existantes (comme cela a été fait au temps des M.A.V. sur les images fixes) est vaine en termes de réduction de polysémie, de facilitation de la compréhension, etc. De même, pour la création d'images pédagogiques, conçues à des fins d'enseignement/apprentissage, il apparaît inutile de miser sur un quelconque contrôle des trois réalités de l'image visuelle pour canaliser l'appréhension de cette image.

6.2. Au moins deux niveaux de "lecture" de l'image

L'image étant un système de représentations du monde, elle présente à ce titre au moins deux niveaux de "lecture". Un premier niveau, celui de la dénotation : les objets du monde sont reconnus et nommés ; et un second niveau, celui cette fois de la connotation : une lecture "culturelle", "symbolique" de ces objets du monde et de leurs configurations. Il est nécessaire de considérer ces deux niveaux de manière indissociable lors de toute introduction d'image dans une séquence d'apprentissage de langue.

La lecture de l'image est subjective et affective et se fera en fonction de l'histoire personnelle du lecteur, de ses représentations. Il est par conséquent tout à fait illusoire de vouloir maîtriser ces deux paramètres que sont l'affectivité et la subjectivité : *"L'efficacité iconique tient [...] à sa position médiatrice ou polysensorielle, qui active et gouverne en nous plusieurs registres de sens."* (Bougnoux, 1994 : p. 107)

De plus, il est difficile et surtout douteux en termes de processus cognitifs de comparer un comportement de "lecteur" d'images en situation d'apprentissage et en situation "intime" devant un poste de télévision chez soi. Il est donc hasardeux, si l'on adopte le point de vue du concepteur d'images pédagogiques, de copier simplement la mise en forme de ces images télévisuelles.

Si la "polysémie" de l'image est irréductible et si la subjectivité et l'affectivité présidant à l'appréhension de cette image sont incontrôlables, alors il apparaît que ce sont les modalités d'introduction de l'image dans l'image multimédia et plus largement dans le logiciel d'apprentissage, et les conditions d'appréhension de cette même image, qui sont didactiquement pertinentes et donc à considérer.



6.3. Ancrage et construction d'images pédagogiques

Si nous nous plaçons du côté du concepteur et/ou scénariste d'images pédagogiques à introduire dans un logiciel de langue, il est souhaitable, selon nous, de signaler un travers dans lequel tombent nombre de ces concepteurs d'images pédagogiques. Nous avons déjà dit que nombre d'images pédagogiques étaient

scénarisées et tournées selon les deux principes de visualisation/monstration et de redondance : ainsi, dans les images pédagogiques mettant en scène des personnages, est filmé celui qui parle et est filmé ce dont parlent ces personnages. Etant entendu que si le thème d'une conversation entre deux protagonistes tourne autour d'horaires de train par exemple, les personnages seront souvent mis en scène dans une gare. Plus finement, si un des personnages énonce qu'il cherche la consigne pour les bagages, il est fort probable que la pancarte signalant cette consigne fera l'objet d'un gros plan.

Il est important, selon nous, de dépasser cette correspondance de surface entre signifié linguistique et situationnel et objets optiquement identifiables : on ne parle pas de trains uniquement dans des situations de type gare ou voyages ! Il est souhaitable de travailler les images également au niveau connotatif et non simplement en restant à un niveau dénotatif.

6.4. Des images pédagogiques à la manière de...

Certains produits d'apprentissage de langues s'inspirent de genres télévisuels familiers (sitcoms, débats télévisés, etc.) pour monter leurs images pédagogiques. Ainsi, *Rendez-vous à l'annexe*, une méthode vidéo de français langue étrangère (FLE) parue chez Hachette, met en scène, à l'instar des sitcoms d'AB Production, un groupe de jeunes montant une entreprise de livraisons de pizzas. Nous ne mettons pas en cause ces genres mais copier ce type de modèle répond à des critères pédagogiques discutables : il semble que l'on copie un modèle (le plus souvent télévisuel ou cinématographique) parce qu'intuitivement on a intégré que les images pédagogiques sont ennuyeuses car elles ne ressemblent à rien en termes de "genre" (comme si un film de fiction par exemple ne pouvait pas être ennuyeux). On obtient alors des productions qui mixent avec plus ou moins de bonheur une mise en forme didactique, des impératifs pédagogiques et les formes conventionnelles du genre copié.

Nous estimons, quant à nous, que les images conçues à des fins pédagogiques constituent un genre à part entière et que leur conception, leur scénarisation doit s'affranchir de modèles conçus à des fins autres.

6.5. Ergonomie et dimension de l'image dans l'image multimédia

La dimension de l'image dans l'image multimédia est une donnée importante dans l'appréhension et le traitement de cette image par l'apprenant : la pénibilité de la lecture d'une image en plein écran, lorsqu'on est placé à portée de clavier, ce qui est le cas le plus fréquent puisque l'apprenant pilote le logiciel, est à prendre en compte.

6.6. Le "beau" et l'utile

La motivation ne se décrète pas (Viau, 1994 ; Vallerand & Thill, 1993) et ce ne peut être le média utilisé et tout particulièrement l'image qui engendre et porte entièrement la motivation de l'apprenant. Cependant nous pensons que l'image est tout de même facteur de motivation, ne serait-ce que par son pouvoir d'attraction de l'oeil (qui reste volontiers rivé à l'image). En terme d'enseignement/apprentissage des langues, l'utilisation des images paraît donc porteuse. Mais il y a un paradoxe. Dans une mise en oeuvre didactique, la polysémie et l'ambiguïté des images font peur car elles sont hors de contrôle. Les images, utilisées dans une séquence d'apprentissage, le sont donc avec précaution et doivent le plus souvent être au service du discours didactique qui, lui, se tient à l'écart de toute tentation stylistique, humoristique, etc. Le résultat est que, souvent, il y a platitude des images et manque de rythme dans le scénario, les images pouvant dès lors perdre leur pouvoir de motivation, de captation de l'intérêt et faire surgir des comportements de type "zapping".



Se pose le problème de la qualité de l'image en parallèle avec le discours didactique. A notre avis, on ne peut se passer du "beau" sous prétexte d'"utile". Le rapport entre des images comme support d'apprentissage et le discours didactique doit être envisagé de manière autre que sous l'angle de la primauté absolue du discours didactique (au sens d'images au service de ce discours). Des images pauvres peuvent à coup sûr gêner une mise en forme didactique astucieuse et innovante.

Esthétique des images, beauté des images sont évidemment des notions éminemment subjectives mais il semble tout de même qu'éviter un certain amateurisme en matière de cadrage notamment, de netteté des images, est une règle de base. (On trouve sur le marché des logiciels dont les séquences vidéo présentent des images saccadées, tremblotantes et parfois floues !).

Cette question de pauvreté et de platitude esthétiques nous semble pouvoir être mise en parallèle avec le revirement méthodologique que l'on a connu en didactique du FLE, à savoir la "réintroduction" progressive des textes littéraires dans l'apprentissage du FLE. La primauté a longtemps été à l'oral, aux dialogues de "survie" dans la langue (scène chez le boulanger, comment saluer, comment demander son chemin, etc.). Puis peu à peu, le littéraire a refait son apparition. De la même façon qu'une évolution s'est fait sentir dans le passage de textes non littéraires (en FLE en tous cas) à des textes littéraires, on est passé de l'image dépouillée des méthodes audiovisuelles à des images plus riches de sens et donc d'interprétations (souvent des images importées, télévisuelles le plus souvent). En création, pour les logiciels multimédia, il faut aller vers des images vraiment travaillées esthétiquement et scénarisées.

7. Quelques jalons méthodologiques pour un cadre d'écriture didactique du multimédia de langues

Tous les jalons donnés ci-après sont à l'usage des concepteurs mais peuvent évidemment offrir des critères pertinents pour l'analyse d'un multimédia que l'on aimerait introduire auprès d'apprenants.

7.1. Substituer la notion d'ancrage à la notion de redondance

Langue et image sont deux modes d'abstraction différents donc deux modes de signifier différents, ne fonctionnant pas de la même manière : on ne peut donc obtenir des "calques", les faire coïncider (Foucher, 1996). De même, il n'y a pas de correspondance semblable entre son et image. Nous pensons donc que les rapports langue / son / image ne doivent pas être envisagés en termes de redondance mais pensés en termes d'ancrage.



7.2. Tenir compte de phénomènes d'association entre certains modes

Si l'on admet qu'il n'y a pas de redondance entre son et image et que c'est l'image qui donne au son ses dimensions, son lieu, qui décide où situer ce son. La localisation auditive est, en effet, énormément influencée par la localisation visuelle, par la visualisation. Et quelle que soit la source du son, nous avons tendance, nous spectateurs, à associer ce que nous voyons à l'écran et ce que nous entendons alors que les deux au départ peuvent être totalement dissociés. (M. Chion (1992 : p. 88) parle, à ce propos, de phénomène de "synchrèse"). Le son est donc toujours défini par rapport à des critères de visualisation ou de non-visualisation, donc toujours par rapport au champ de l'écran. A ce titre, il apparaît particulièrement important de soigner les dispositifs énonciatifs mis en place dans le logiciel : l'apparition à l'écran de personnages, par exemple, si elle coïncide avec une intervention pédagogique sonore (commentaires de réponse par exemple), peut amener des confusions sur les instances énonciatives, les énonciateurs en question. (Nous reviendrons sur les dispositifs énonciatifs en jeu dans les logiciels de langues autour des trois catégories suivantes : dilution, polarisation et mouvement des instances énonciatives).

7.3. Travailler les "périphériques modaux" de l'image visuelle dans l'image multimédia

C'est le manque de focalisation assertive (Joly, 1994 : p. 83) plutôt que la polysémie qui définit l'image visuelle. C'est par l'intermédiaire du langage notamment qu'elle affirmera ou dénier quelque chose, c'est par l'introversion du verbal par exemple que l'image sera temporalisée. Cette caractéristique de l'image visuelle a, selon nous, des répercussions didactiques fortes. Ainsi, si l'on veut introduire de l'image visuelle dans un scénario multimédia et donc dans une image multimédia, il apparaît primordial de situer, de contextualiser cette image. Le plus souvent les images introduites le sont de façon abrupte, sans apprêt (nous ne parlons pas ici de justifications de l'utilisation de l'image qui pourraient apparaître en amont). La contextualisation de l'image visuelle dans l'image multimédia doit se faire à plusieurs niveaux. Nous en avons dégagé trois : un niveau diégétique, un niveau pédagogique et un niveau relatif à l'actualité de l'apprenant.

Contextualiser l'image au niveau diégétique

Dans le multimédia, T. Lancien (1998 : p. 56) souligne à juste titre que "[la linéarité des vidéos] semble en quelque sorte entrer en contradiction avec les possibilités d'écoute fragmentée et parcellisée qui sont

proposées". L. Plowman (1996b), M. Stratfold (1994) et D. Laurillard (à paraître) parlent, eux, de cohérence narrative dans les ILEs (Interactive Learning Environments) pour pallier l'impact négatif de l'insertion ponctuelle des modes (plus ou moins fréquente, plus ou moins longue).

Nous proposons donc de contextualiser de manière diégétique les images : cette contextualisation pourra prendre effet dans le cas de séquences vidéo longues et découpées en modules plus courts consultables dans le désordre (nous nous référons ici à des images conçues à des fins pédagogiques, non authentiques en ce sens qu'elles ne s'ancrent pas dans une réalité extérieure). En effet, si les modules vidéo courts ne sont pas autonomes narrativement mais réfèrent à des événements antérieurs au niveau diégétique, il est important selon nous de les situer. Ce sera sans doute moins vrai pour des produits basés sur une narration découpée en situations indépendantes.



Cette contextualisation diégétique pourra prendre la forme d'un court texte (en message furtif par exemple) sous la vignette vidéo ou d'un commentaire en voix off au tout début de la séquence (convocable à loisir).

Contextualiser l'image au niveau pédagogique

Les images, les séquences vidéo qui sont introduites dans le scénario pédagogique, c'est-à-dire le plus souvent en tant que supports d'activités langagières, doivent à notre sens également être contextualisées dans le scénario pédagogique, d'une part lorsque la séquence est présentée à l'apprenant pour la première fois (avant qu'il ne la traite) et, d'autre part, dans le cas où il y a réutilisation d'une même séquence pour une série d'activités différentes.

La contextualisation pédagogique à intégrer lorsque l'image apparaît pour la première fois comme support d'une activité pourra être indépendante de la consigne de traitement de la séquence donnée à l'apprenant comme par exemple dans :

(1) Vous pouvez écouter et regarder la vidéo en utilisant les contrôles vidéo ensuite vous passerez aux activités de compréhension (consigne extraite de *Camille*)

ou intégrée à cette consigne :

(2) La réceptionniste de Tyson Computers Ltd. parle de son travail. Regardez la vidéo et recherchez les informations de manière à compléter la description du poste. Cliquez [...] (consigne extraite de *Business Challenges Interactive*)

(3) A votre avis, au moment où le commandant Mahoudo dit : "Nous allons maintenant lancer le départ des différentes unités d'intervention", le départ des camions est :

- un événement en train de se réaliser.
- un événement à venir.

(consigne extraite des *Pompiers au Secours du Temps*).

Dans le cas d'une séquence vidéo introduite à plusieurs reprises dans le logiciel multimédia mais support d'activités langagières différentes, il s'agira bien entendu de varier cette contextualisation pédagogique.

Contextualiser l'image dans l'actualité de l'apprenant

Cette contextualisation ne se fera que dans le cas d'intégrations au logiciel multimédia de documents vidéo inscrits dans une actualité et notamment de journaux télévisés, de reportages télévisés, etc. L'intérêt de l'introduction de ce type de documents authentiques est avant tout de faire travailler les apprenants sur des thèmes d'actualité liés au pays dont ils apprennent la langue et les documents sont choisis alors selon ces critères. On note cependant que si la contextualisation du document est traditionnellement faite pour les méthodes de langue ou les manuels éducatifs, elle l'est rarement dans les logiciels multimédia.

Pour le logiciel multimédia, cette contextualisation ne sera guère différente de celles que l'on peut trouver

sur papier et elle prendra la forme de "coordonnées énonciatives" : il s'agira d'abord classiquement de dater l'extrait, cet extrait étant ancré dans une réalité à laquelle participe, même de loin, l'apprenant : l'apprenant n'est pas le téléspectateur natif s'informant sur un sujet "chaud" concernant un grand nombre de compatriotes. L'inconvénient de la datation d'un tel document dans un produit de langue multimédia (ou dans un manuel quand il s'agit de presse écrite par exemple) est évident et accélère d'autant la péremption du document ; mais travailler sur l'actualité admet difficilement les grands décalages temporels entre l'actualité du document et celle de l'apprenant. Il s'agira aussi de préciser d'où est extrait le document, de donner à l'apprenant la source médiatique du document.



Les modalités d'affichage de ces coordonnées seront en revanche différentes et pourront, dans le logiciel multimédia, faire l'objet d'une fiche signalétique du document à la disposition de l'apprenant et consultable à loisir.

7.4. Travailler les charnières modales

Nous avons parlé de contextualisation de l'image visuelle dans l'image multimédia et au sein du scénario pédagogique. Une manière de "situer" également les images visuelles, selon nous, consisterait à travailler les liens entre les modes : les charnières modales. Pour le multimédia, on ne parlera évidemment pas de légende au sens courant de "texte qui accompagne une image et l'explique" mais plutôt de "charnières modales". Nous définissons ces charnières modales comme les liens existant entre deux modes : par exemple, un renvoi au mode visuel et auditif que constitue la séquence vidéo peut être signifié à l'apprenant par une intervention orale de l'instance pédagogique. Dans ce cas précis, la charnière modale sera linguistique et orale. Mais les charnières modales peuvent prendre la forme d'un texte écrit, d'un commentaire oral, d'une animation ou d'une autre image.

A l'examen de plusieurs logiciels multimédia de langue, on constate que les charnières modales présentes entre les écrans sont de deux types : elles apparaissent essentiellement sous forme de boutons et d'icônes avec parfois un mot sur le bouton et plus rarement sous forme linguistique (écrite ou même orale). A sein d'un même écran, c'est l'inverse qu'on observe : "l'interface" est peu utilisée pour guider la lecture de l'écran (rares sont, par exemple, les affichages décalés dans le temps de certains modes qui pourraient induire un ordre pour le traitement de ces modes qui s'affichent). Ce sont le plus souvent des charnières modales linguistiques, écrites surtout mais aussi orales, qui indiquent un ordre de circulation dans l'écran : nous trouvons des énoncés du type "Cliquez sur une des images, écoutez le début du dialogue et regardez la mimique du personnage. Ensuite choisissez dans la liste d'expressions celle qui correspond à l'image", cet énoncé permettant à l'apprenant de hiérarchiser tous les modes présents à l'écran (visuel, sonore, puis linguistique).

Mais le plus souvent ces charnières modales sont quasi-inexistantes : l'apprenant a sous les yeux une organisation multimodale complexe qui n'est pas hiérarchisée en termes de traitement : par quoi commencer ? pour quoi faire ? Le schéma pré-construit que nous retrouvons toutefois le plus souvent est celui-ci : c'est une consigne (donc une charnière modale linguistique) qui permet le passage au mode visuel / auditif, puis le passage aux sollicitations se fait directement par l'intermédiaire d'un bouton (ces sollicitations, dans une activité de compréhension par exemple, pourront être orientées sur les contenus du mode visuel / auditif). Nous pouvons également trouver directement le mode visuel / auditif à traiter et un passage direct aux questions de compréhension, sans charnières modales, l'ordre de présentation des modes étant considérée comme suffisante pour hiérarchiser le traitement de ces mêmes modes. Cet enchaînement entre les modes reproduit des pratiques de classe (en présentiel) que connaissent les apprenants.



Il semble donc que la spécificité du multimédia de présenter ensemble différents modes, avec la possibilité de les hiérarchiser, de les associer, etc. n'est pas exploitée et cela, bien que l'apprenant se trouve seul devant le logiciel. Cette situation d'individualisation de l'apprentissage et les potentialités du multimédia devraient permettre de mettre en oeuvre des activités autres que celles proposées en classe pour un groupe-classe. Rechercher dans un document vidéo des réponses à des questions préalablement formulées, en cliquant dès qu'une réponse est trouvée, par exemple, est tout à fait envisageable. Ces charnières modales pourraient proposer à l'apprenant un "ordre de lecture" des différents modes présents dans l'image multimédia et entre les différents écrans. Ce schéma ne serait bien entendu qu'un guide de

lecture de l'image multimédia "à disposition" de l'apprenant, celui-ci restant maître à bord. Le concepteur de logiciel multimédia, pour sa part, veillera à varier les charnières modales qu'il développe, systématisera leur introduction dans le logiciel en associant certains types de charnières à certains types d'activités par exemple.

Ces charnières modales présentent de nombreuses caractéristiques (pouvant être croisées) que nous avons regroupées dans la classification suivante :

Charnières internes ou inter-écrans

Dans un logiciel d'apprentissage des langues, le passage d'un mode à l'autre peut se faire à "l'intérieur" d'un écran ou entre écrans.

Charnières orientées

Les charnières modales sont orientées : elles relient deux modes dans un ordre qui sera celui du traitement à effectuer sur ces modes. L'orientation de traitement est indicative. Ainsi, le traitement d'une consigne (mode linguistique) précède en général le traitement d'une séquence vidéo (mode visuel et auditif) mais un apprenant peut décider de traiter la vidéo sans objectif délimité avant de la revoir avec une consigne plus précise.

Charnières imposées ou proposées

Elles peuvent être imposées ou proposées (c'est-à-dire à convoquer par l'apprenant). Le choix dépendra ici de l'optique didactique privilégiée. Toutefois, à l'heure actuelle, la tendance est davantage à des logiciels laissant l'apprenant décider de son parcours : imposer des charnières modales, c'est en quelque sorte rigidifier le cheminement de l'apprenant, ce qui peut gêner voire irriter certains d'entre eux. Aussi pensons-nous que mettre à la disposition de l'apprenant ces charnières permet à la fois de rassurer certains apprenants et de ne pas bloquer certains autres.

Charnières diégétiques, pédagogiques ou "topographiques"

Elles peuvent être diégétiques, pédagogiques ou "topographiques", destinées respectivement à relier différentes séquences vidéo entre elles, à relier une activité pédagogique et un mode à traiter pour mener à bien l'activité, ou à indiquer les étapes du cheminement de l'apprenant dans le logiciel.



Charnières linguistiques, auditives ou iconiques

Elles peuvent être linguistiques, auditives ou iconiques. Linguistiques, elles pourront être orales ou écrites ; auditives, elles pourront prendre la forme de signaux (à l'instar des livres-disques des enfants où un son de clochette indique qu'il faut tourner la page) ; iconiques, elles pourront se réaliser sous forme de montages d'images, d'affichage différé, d'animation, de changement de couleur, etc.

7.5. Quelques exemples de charnières modales

Ces charnières modales peuvent prendre, par exemple, la forme d'un message furtif associé à chaque mode présent sur un écran, les contenus de ces messages furtifs (donc à disposition de l'apprenant mais pas imposés) pouvant être de différents types. Pour nous, il ne s'agit pas d'assimiler systématiquement ces charnières modales à des consignes : nous redisons que baliser fortement un cheminement en introduisant, à chaque changement de mode, des énoncés du type "Vous venez de faire ceci, nous allons maintenant passer à cela" serait parfaitement insupportable pour qui travaillerait sur ce logiciel (la tendance étant à l'hypermédia et aux interfaces "intuitives"). Mais elles permettraient d'une part d'éviter que les apprenants, par exemple, ne se réfèrent à des modes n'ayant plus rien à voir avec l'activité pédagogique qu'ils mènent pour répondre aux sollicitations. Elles pourraient, d'autre part, permettre de tenir compte de la spécificité de la lecture des textes dans une situation d'interaction sujet-ordinateur (Rouet, 1991), du fait par exemple que des textes écrits sont rarement lus dans leur intégralité en une fois et que parfois "quatre sauts successifs" sont nécessaires pour mener une lecture exhaustive (cf. expérience du CRESLEF sur le didacticiel de littérature Didatext dans Masselot-Girard, 1994).

Nous proposons ici dans des cas particuliers que nous décrivons quelques charnières modales envisageables pour un logiciel multimédia de formation en langues.

Pour une phase de compréhension par exemple, élaborée à partir d'une séquence vidéo longue et morcelée (découpée en séquences), on pourra proposer (ou fixer, selon l'optique du concepteur) que ces séquences extraites d'une vidéo plus longue soient reliées entre elles quant au contenu diégétique. On peut imaginer alors des messages furtifs sous des vignettes vidéo en images arrêtées représentant la situation de chacune des séquences. On pourra choisir de mettre ces messages furtifs sous forme de pavés de textes, mais on pourra également les envisager sous forme d'intervention orale (il s'agira alors de déterminer avec rigueur "qui" interviendra et de permettre à l'apprenant d'identifier ce "qui"). La charnière modale peut aussi prendre la forme d'une intervention orale "visualisée" dans une fenêtre. Le choix de l'instance intervenant à ce moment est évidemment déterminante, sa représentation visuelle également : choisir de visualiser l'instance pédagogique re-situant la séquence dans son contexte ou préférer visualiser un des personnages de la vidéo résumant brièvement la situation devra faire l'objet d'une réflexion en amont. Le statut conféré à une charnière modale ne sera sans doute pas perçu par l'apprenant de la même manière selon qu'il s'agit d'une intervention prise en charge pédagogique "directement" par l'instance pédagogique ou de l'intervention d'un personnage.

Si l'on privilégie dans ce cas des charnières modales iconiques, on peut imaginer plus simplement, pour ce message furtif, une image fixe d'un moment clef de la situation, accompagnée d'une bulle de type bande dessinée (à l'instar des romans-photos) ou d'un carton (comme dans les films muets).



En revanche, dans le cas d'une phase d'activités langagières où il s'agirait de systématiser l'utilisation d'une structure dans une série d'exemples en langue étrangère, il paraît peu judicieux d'imaginer que chacun des exemples de cette liste soit oralisé par une voix différente, la contextualisation serait là de toute façon insuffisante. Associer à chacun des exemples travaillés une mini-séquence vidéo offrirait une contextualisation plus riche mais cela apparaît peu raisonnable en termes de faisabilité et de rentabilité.

Au sein d'une activité langagière par exemple, pour articuler consigne pédagogique, support de l'activité et zone de réponse de l'apprenant, on peut décider de privilégier des charnières iconiques sous forme de couleur : les différents cadres à l'écran (par exemple, une vignette vidéo et un pavé de texte) changeront de couleur pour guider la lecture. Un affichage différé dans le temps est également une solution permettant d'induire un ordre de lecture mais de ne pas l'imposer.

Si l'apprenant convoque une aide visuelle, par exemple pour mener à bien une activité de compréhension, en cas de réponse incorrecte par la suite, on peut lui suggérer que, pour cette activité, une aide textuelle serait peut-être plus adéquate : la charnière entre les deux types d'aides peut alors prendre la forme d'un commentaire méthodologique oral pris en charge par l'instance pédagogique. On peut plus simplement afficher une icône indiquant qu'à ce moment-là, une aide textuelle peut être consultée : cette icône ne ferait pas partie de la barre de commandes présente tout au long du didacticiel mais serait spécifique du type d'activité mise en oeuvre à ce moment.

7.6. Travailler localisation sonore et localisation visuelle : l'exemple de la multimodalité énonciative mise en oeuvre dans les logiciels multimédias

Nous avons défini plus haut la multimodalité et l'intermodalité (cf. 3.2.). Pour montrer combien est complexe la mise en oeuvre de tous les modes dans les logiciels d'apprentissage de langues, nous avons choisi de prendre pour exemple les dispositifs énonciatifs en jeu dans ces logiciels. Un dispositif énonciatif est, pour nous ici, l'organisation de la pluralité des "voix" dans un échange linguistique ou la mise en oeuvre de la "polyphonie" dans des situations particulières (voix de l'instance pédagogique, voix des personnages des séquences vidéo, voix de l'apprenant, voix autre). Les modes en jeu ici sont : le mode auditif (certaines instances énonciatives "s'expriment" sur le mode sonore), le mode textuel (la prise en charge de l'énonciation se fait textuellement) et le mode visuel (où les instances énonciatives sont représentées à l'aide de vignettes dessinées ou apparaissent tout simplement dans la séquence vidéo).

Cette polyphonie ou multimodalité énonciative émane essentiellement, dans les logiciels d'apprentissage des langues, de trois instances, de trois voix : celle du pédagogue, par le biais des consignes, des commentaires, des écrans de résumés de contenus, etc. ; celle de l'apprenant, par ses réponses aux sollicitations ; celles des personnages de la vidéo dans des situations de la vie quotidienne le plus souvent.

Nous avons constaté que dans les logiciels multimédia de langue, dont l'objectif premier est l'apprentissage de la langue, et ce, souvent selon une démarche dite communicative, la multimodalité énonciative est mise en oeuvre selon trois grandes catégories. La première catégorie, que nous appelons "polarisation des instances énonciatives", signale des logiciels dans lesquels on sait toujours qui s'adresse à qui et dans quel but, dans lesquels les différentes instances énonciatives sont clairement identifiables par l'apprenant. La multimodalité énonciative est alors "lisible" par l'apprenant.



La seconde catégorie, que nous appelons "mouvement des instances énonciatives", concerne des logiciels dans lesquels l'apprenant, par exemple, aura une double position d'interlocuteur, interlocuteur de l'instance pédagogique (c'est le cas pour tous les logiciels) et interlocuteur des protagonistes de la séquence vidéo. Ce peut être également un personnage de la vidéo qui endosse le rôle du pédagogue (ou d'un interlocuteur de l'apprenant). La troisième catégorie, que nous appelons "dilution des instances énonciatives", qualifie des logiciels dans lesquels ces instances énonciatives ne sont pas clairement identifiables, dans lesquels se pose le problème pour l'apprenant de savoir qui parle, à qui et pourquoi et quel est le statut que celui-ci doit adopter lors de sa réponse à une éventuelle sollicitation.

Certains dispositifs énonciatifs mis en place dans les écrans présentent des non-sens énonciatifs (et par là même pédagogiques) : l'instance pédagogique est diluée, le statut conféré à l'apprenant (doit-il se mettre dans la peau d'un personnage ? répondre en son nom propre ?) n'est pas arrêté, le statut des personnages et leur rôle de locuteur natif est également flou. Ainsi, on peut rencontrer une image visuelle mettant en place un dispositif énonciatif à deux interlocuteurs alors que le pavé de texte associé à cette image renvoie à trois dispositifs dont un est censé correspondre à celui de l'image. L'échange pédagogique apparaît dans ces cas singulièrement compromis. Le croisement des données iconiques et des données langagières peut donc introduire une confusion dans les dispositifs énonciatifs. Veiller à une grande maîtrise des rapports entre les données iconiques et langagières, minimalement pour qu'elles ne soient pas divergentes ou contradictoires, semble être nécessaire.

Il est nécessaire de bien circonscrire les différents rôles que peuvent endosser les différentes instances énonciatives : une organisation multimodale différente de l'écran pour le rôle de chaque instance énonciative pourrait constituer une réponse. A ce titre, la technique de surimpression d'un personnage incarnant une instance particulière, par exemple, peut contribuer à bien différencier son statut ; de même un personnage sortant d'un cadre, par exemple, indiquera alors un passage d'un monde, celui de la vidéo, à un autre monde, celui du pédagogue.

La place de la séquence vidéo dans l'écran et son environnement (boutons, bandeaux, etc.) a aussi sans aucun doute une influence sur la perception que peuvent en avoir les apprenants : c'est là sans doute rappeler des évidences mais on ne regarde pas de manière identique une vidéo plein écran et une vignette vidéo en quart d'écran, l'apprenant n'y voyant certainement pas de la même manière l'utilité de la séquence dans l'activité pédagogique et le statut de son pseudo-locuteur. Bien "délimiter" les différentes instances énonciatives peut être, comme nous l'indiquons plus haut, pris en charge par des potentialités de l'interface multimédia. Mais prendre soin au contenu des interventions sonores ou écrites lors de la rédaction, si l'on a choisi ce mode-là, semble être un minimum, ces interventions construisant les échanges entre apprenants, pédagogues et personnages de la vidéo (consignes, aides, feed-back, etc.).

8. Conclusion

La spécificité du multimédia pour la formation tient non pas à la juxtaposition de médias sur un même support mais aux liens tissés entre ces différents médias : ce tissage, que nous définissons en termes d'intégration et d'interactions, si on le veut véritablement au service de l'apprenant et de son parcours d'apprentissage, appelle de la part des concepteurs une réflexion allant nécessairement de la linguistique à la sémiologie en passant bien évidemment par la didactique.



En termes méthodologiques, pour écrire le multimédia dans un cadre de formation en langues, il apparaît tout à fait important de :

- fournir une variété de liens clairement définis entre les différentes sections d'un programme,

fournir en corollaire des charnières modales entre les modes pour permettre un maximum de choix de la part de l'apprenant dans son cheminement entre les activités et au sein des activités ;

- fournir des charnières modales "topographiques" pour que l'apprenant se familiarise avec la manière dont vont être présentés les contenus, les modes, etc. ;
- fournir des charnières modales différentes pour les nouvelles informations auxquelles accédera l'apprenant ;
- prêter un grand soin au choix des modes textuels, iconiques, visuels, auditifs pour contextualiser les formes linguistiques ;
- fournir des charnières méthodologiques pour faciliter le choix par l'apprenant d'une stratégie adéquate.

Références

Références bibliographiques

Barthes R. (1964). "Rhétorique de l'image". *Communications* 4, "Recherches sémiologiques". Seuil, pp. 40-51.

Bodenreider O., Le Bouleix F., Kohler F. & Kolher C. (1996). "Du multimédia intégré au multimédia réparti ou comment passer d'une philosophie HyperCard à une philosophie WWW ?". Bruillard E., Baldner J.-M. & Baron G.-L. (eds) (1996). *Actes des troisièmes journées scientifiques "Hypermédias et apprentissages"*. Paris : INRP - EPI.

Bougnoux D. (1994). "Nous sommes sujets aux images". *Revue Esprit*, "Vices et vertus de l'image", pp. 96-109.

Boutan P., Chailley M., Vigne H. (1988). "Il était une fois... la vie. Enseigner la fiction". *Repères* 74, "Images et langages", pp. 27-38.

Bresson F. (1981). "Compétence iconique et compétence linguistique". *Communications* 33, "Apprendre des médias". Seuil, pp. 165-196.

Chanier T. & Pothier M. (eds) (à paraître). "Apprentissage des langues et environnements informatiques hypermédia". *Etudes de Linguistique Appliquée* 111.

Chanier T. (à paraître). "Hypertexte, hypermédia et apprentissage dans des systèmes d'informations et de communication : résultats et agenda de recherches". *Etudes de Linguistique Appliquée* 111, "Apprentissage des langues et environnements informatiques hypermédia".

Chanier T., Duquette L., Laurier M. & Pothier M. (à paraître). "Stratégies d'apprentissage et évaluation dans des environnements multimédia d'aide à l'apprentissage du français". *Actes des Journées scientifiques et techniques du Réseau Francophone de l'Ingénierie de La Langue de L'Aupelf-Uref (JST'97)*, Avignon..

Chion M. (1992, réédition). *Le son au cinéma*, coll. Essais. Paris : Ed. Cahiers du Cinéma.

Demaizière F. & Foucher A.-L. (à paraître). "Individualisation et initiative de l'apprenant dans des environnements (et des dispositifs) d'apprentissage ouverts : une expérience d'autoformation guidée". *Etudes de Linguistique Appliquée*, 111.



Demaizière F. (1996). "Autoformation, nouvelles technologies et didactique. Réflexions et propositions méthodologiques". *Les Sciences de L'Education pour l'Ere Nouvelle*, vol. 29, 1-2, pp. 67-100.

Duquette L. & Renié D. (à paraître). "Stratégies d'apprentissage dans un contexte d'autonomie et environnement hypermédia". *Etudes de Linguistique Appliquée* 111, "Apprentissage des langues et environnements informatiques hypermédias".

Eco U. (1970). "Sémiologie des messages visuels". *Communications* 15, "L'analyse des médias", Seuil, pp. 11-51.

Eco U. (1988). *Le signe* (trad.fr). Bruxelles : Labor.

Escarpit R. (1976). *Théorie générale de l'information et de la communication*. Paris : Hachette.

Foucher A.-L. (1996). "Didactique des langues et Nouvelles Technologies pour la Formation : entre linguistique et psychologie cognitive". *Actes du Deuxième Colloque Jeunes Chercheurs en Sciences Cognitives*, Presqu'île de Giens, pp. 309-312.

Foucher A.-L. (1998). *Didactique des Langues et Nouvelles Technologies pour la Formation : entre linguistique, sémiologie de "l'image multimédia" et enseignement/apprentissage des langues*, thèse pour le doctorat de Didactique des Disciplines, Université Paris 7.

Ginet A. (et al.) (1997). *Du laboratoire de langues à la salle de cours multimédia*. Nathan, Paris.

Joly M. (1994). *Introduction à l'analyse de l'image*. coll. 128, Paris : Nathan Université.

Joly M. (1994). *L'image et les signes*. Paris : Nathan.

Lancien T. (1998). *Le multimédia*. coll. Didactique des Langues Etrangères, Paris : CLE International.

Laurillard D. M. (à paraître dans *Computers and Education*). "Multimedia and the learner's experience of Narrative".

Laurillard D.M. (1993). *Rethinking University teaching: a framework for the effective use of educational technologies*. Longdon : Routledge.

Lin X. (1994). "Metacognition : Implications for Research and Hypermedia-based Learning Environment". *Proceedings of Selected Research and Development Presentations at the 1994 National Convention of The Association for Educational Communications and Technology Sponsored by the Research and Theory Division*, Nashville, pp. 485-502.

Linard M. (1996, édition réactualisée). *Des machines et des hommes*. coll. Savoir et formation, Paris : Ed. L'Harmattan.

Masselot M. (1993). "Eléments pour une didactique des textes pluricodiques". *Repères 7*, "Langage et images", pp. 151-160.

Masselot-Girard M. (1994). "Le manuel, un livre à part". *Cahiers du CRESLEF 37*, "Images et savoirs", pp. 03-15.

Masselot-Girard M., Jouvenot Y. (1994). "Les savoirs des manuels de littérature". *Cahiers du CRESLEF 37*, "Images et savoirs", pp. 15-54.

Mayer R.E. & Anderson R.B. (1992). "The instructive narration : helping students built connection between words and pictures in multimedia learning". *Journal of Educational Psychology*, 84, pp. 484-490.

Metz C. (1964). "Le cinéma : langue ou langage ?". *Communications 4*, pp. 52-90.

Minot F. (1993). "Le récit pris au mot et à l'image - contes de fées et publicités télévisées en classe primaire". *Repères 7*, "Langage et images", pp. 23-40.

Pasquier F. (1997). *La vidéo interactive pour la formation. Création et usages d'un dispositif multicanal de diffusion de vidéo numérique. Le cas de la compréhension de l'anglais oral*, thèse pour le doctorat de Didactique des Disciplines, université Paris 7.



Plowman L. (1994). "The 'Primitive Mode of Representation ' and the evolution of interactive multimedia". *Journal of Educational Multimedia and Hypermedia 3* (3/4), pp. 275-293.

Plowman L. (1996a). "Towards and understanding of interactive multimedia 'literacy' : mapping the space". Paper presented at Literacy and Culture, the 33rd annual conference on Cognitive Ergonomics, pp. 167-172, Granada, Spain.

Plowman, L. (1996b). Narrative, linearity and interactivity: making sense of interactive multimedia. *British Journal of Educational Technology 27* (2), pp. 92-105.

Rouet J.-F. (1991). *Compréhension de textes didactiques par des lecteurs inexpérimentés dans des situations d'interaction sujet-ordinateur*. Thèse pour le doctorat de psychologie, Poitiers.

Stratfold M. (1994). *Investigation into the design of Educational Multimedia : Video, Interactivity and Narrative*. PhD thesis, Open University.

Tardy M. (1973). *Le professeur et les images*. Paris : PUF.

Thon B., Marquié J.-C., Maury P. (1993). "Le texte, l'image et leurs traitements cognitifs". *Actes du Colloque Interdisciplinaire du CNRS, "Images et langages. Multimodalité et modélisation cognitive"*, pp. 29-39.

Vallerand R. J., Thill E. E. (dir) (1993). *Introduction à la psychologie de la motivation*. Laval, Québec : Ed. Études Vivantes.

Viau R. (1994). *La motivation en contexte scolaire*. Paris : PUF.

Références logiciels

Break the Ice (1992). Videologic Limited.

Chanier T., Pothier M. & Lotin P. (1996). *L'acte de vente. Deuxième cédérom de "Camille - Travailler en France"*. Paris : CLE International/Nathan.

Demaizière F., Redon M. (1994). *Les Pompiers au secours du temps*. CNEAO - Equipe Ordi, Universités Paris 6 - Paris 7.

LTV Allemand. Jériko.

Lussan E. (1996). *Opération Teddy Bear*. Paris : Flammarion.

Rendez-vous à l'annexe. Paris : Hachette.

Talk to me. Auralang, Médiaconcept.

Notes

[1] Nous préférons le terme de "multimodalité" qui a l'avantage, selon nous, de prendre en compte la nature et les spécificités des modes en jeu contrairement à "multicanalité" qui renvoie plutôt aux supports véhiculant ces modes.

[2] Dans tout ce texte, le terme "multimédia" renverra à "des informations stockées sur des supports multiples et diffusées par le média électronique des systèmes d'information. Ainsi un système multimédia favorise la communication interactive d'informations dans un format intégrant des ressources non restreintes aux textes, soit des ressources verbales (texte + audio), soit des ressources verbales et non verbales (diagrammes, images fixes ou animées, vidéo). Cette possibilité de jouer sur des canaux de communication variés (visuel pour le texte et l'image ; oral pour l'audio et les sons) a ouvert des perspectives particulièrement intéressantes en apprentissage des langues, en permettant à l'apprenant de coupler des procédures cognitives de traitement basées sur les aspects verbaux et non-verbaux du langage." (Chanier, 1998 : p. 5). Plus précisément, nous parlerons du "multimédia intégré", où toutes les ressources sont disponibles sur un support unique, un cédérom par exemple (par opposition au "multimédia réparti", où les ressources sont dispatchées sur la Toile (WWW). (Bodenreider & al, 1996 : p. 282).



[3] En tout cas en France, on sait qu'ailleurs et notamment en Grande-Bretagne des chercheurs en ALAO ont tissé des liens étroits entre didactique et linguistique. De même, des expertises, dans le domaine des environnements d'apprentissage assisté par ordinateur, plus centrées sur des disciplines scientifiques comme les mathématiques et la physique ont intégré la didactique.

[4] Nous pensons ici à des logiciels comme *Opération Teddy Bear* (Lussan, 1996) où l'apprenant peut cliquer partout sur l'image fixe à l'écran et cela, de manière intuitive puisqu'aucun "indice" n'indique l'endroit où cliquer susceptible de déclencher "quelque chose". En cliquant sur des éléments visuels ou textuels, l'apprenant déclenche des animations sonores, visuelles, fait parler un personnage, etc.

[5] Christian Metz distingue quatre façons d'aborder le film : la critique de cinéma, l'histoire du cinéma, la théorie du

cinéma, réflexion fondamentale menée de l'intérieur du monde cinématographique et enfin, la filmologie, "*étude scientifique menée de l'extérieur par des psychologues, des psychiatres, des sociologues, des pédagogues, des biologistes*" ("*Le cinéma : langue ou langage ?*", *Communications* 4, 1964, p.89).

[6] Dans la classification de Peirce, l'image est une sous-catégorie de l'icône, l'icône étant "*un signe qui renvoie à son objet en vertu d'une ressemblance, du fait que ses propriétés intrinsèques correspondent d'une certaine façon aux propriétés de cet objet [[équivalence]] De sorte que sont des icônes une photocopie, un dessin, un diagramme mais aussi une formule logique et surtout une image mentale.*" (Eco, 1988 : p. 75).

L'icône (orthographié comme en américain, sans accent) comprend donc trois sous-catégories, l'image, le diagramme et la métaphore.

"*L'image, c'est le signe iconique qui met en oeuvre une ressemblance qualitative entre le signifiant et le référent. [Pour ce qui est de l'image visuelle] elle imite, ou reprend, un certain nombre des qualités de l'objet : forme, proportions, couleurs, texture, etc.*" (Joly, 1994 : p. 33).

L'image peut ne pas être visuelle : un bruitage renverra à une image sonore. Un parfum de synthèse dans un yaourt renverra à une image olfactive tandis qu'à un matériau imitant le cuir, par exemple, correspondra une image tactile.

[7] La question de la "visualisation" à l'écran de la transcription d'un oral est un problème. Doit-on adopter un système de transcription spécifique (à l'instar de celui adopté par le GARS (Groupe Aixois de Recherche en Syntaxe) par exemple simplifié) ou maintenir les formes conventionnelles de l'écrit connues des apprenants mais inadéquates pour l'oral ?

[8] terminologie adoptée par le groupe INRP "Sémiotiques" mais aussi par le CRESLEF (Centre de Recherches et d'Études Sémio-Linguistiques pour l'Enseignement du Français) de l'UFR des Sciences de L'Homme, du Langage et de la Société de l'Université de Besançon.

[9] cf. seconde acception du terme "dynamiter" donnée par le Robert : faire éclater les règles traditionnelles d'un système.



A propos de l'auteur

Anne-Laure FOUCHER : ATER à l'université Paris 6. Elle vient de soutenir une thèse intitulée Didactique des Langues et Nouvelles Technologies pour la Formation : entre linguistique, sémiologie de l'image multimédia et enseignement/apprentissage des langues. Ses recherches actuelles s'articulent autour de l'image multimédia, des stratégies d'apprentissage et des dispositifs d'autoformation.

Mél/Courriel : foucher@moka.ccr.jussieu.fr

Adresse : Université Paris 6 - Pierre et Marie Curie, case courrier 7023 4, Place Jussieu, 75252 Paris cedex 05.

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

© *Apprentissage des Langues et Systèmes d'Information et de Communication*, 1998