

Chapitre 4

Sémantique multimodale — Intersémiotique

4.1 Introduction : Sémantique des systèmes de signes multimodaux

La question de la communication multimodale intéresse vivement les concepteurs d'interfaces pour ordinateurs, qui y voient, peut-être à juste titre, une nouvelle marche qualitative pour rendre plus aisée et plus efficace¹ la commande de machines par leurs utilisateurs.

Ce sujet soulève quelques problèmes théoriques, dont l'une des questions-clés, transcrite dans la problématique CHM, pourrait se résumer ainsi : lors du processus « ascendant » de reconnaissance et d'interprétation des messages multimodaux, à quel niveau d'analyse convient-il de considérer des éléments qui ne sont plus liés à une modalité particulière ? En des termes en vogue dans la communauté de recherche en Communication Homme-Machine, quand doit s'effectuer la « fusion » des données reçues en événements multimodaux ?

Cette question se trouve au centre de développements théoriques et pratiques autour du concept de « dialogue multimodal », dont on a tenté dans une note antérieure de donner un aperçu général ([Vaillant, 1993]).

La multimodalité n'est toutefois pas, loin de là, un objet exclusif de la communication homme-machine. Bien naturellement, si cette discipline pratique s'y est intéressée, c'est qu'elle trouvait dans ce concept un moyen de simuler l'une des caractéristiques naturelles des systèmes sémiotiques humains. Bien des formes artistiques, bien des genres textuels, bien des codes de communication répandus dans les sociétés

¹L'efficacité est la motivation centrale, même si elle ne se traduit qu'en économies de temps de formation. Les termes « naturel » ou « convivial » reflètent à notre sens une vision un peu déformée de cette motivation : on peut apprendre en quelques semaines à se servir vite et bien d'un clavier de machine à écrire, procédé moins « naturel » que l'écriture manuscrite puisqu'on apprend cette dernière bien plus tôt, à l'école. Le clavier comme outil de saisie de texte ne sera donc pas troqué contre l'ardoise électronique, toute « naturelle » que celle-ci soit, tant que la fiabilité et le temps de reconnaissance de l'écriture manuscrite ne rendront pas les performances de l'ardoise supérieures à celle du clavier (ce qui n'est pas pour 1997)

humaines sont, en somme, multimodaux sans le savoir. L'imbrication des modalités, sur le plan de l'expression, est observable dans des systèmes sémiotiques aussi différents que la conversation gesticulée et la bande dessinée, que le cinéma et la notice technique, que la valse et le théâtre de marionnettes. La description de ces modes d'imbrication sur le plan de l'expression est abordée dans le chapitre 2. C'est la description de l'imbrication *sur le plan du contenu* qui sera l'objet des paragraphes qui suivent.

4.2 Ordre référentiel

Le prolongement de la métaphore informatique, sur le niveau auquel apparaissent des représentations amodales ou multimodales, nous amène à considérer en premier lieu la question des représentations cognitives associées aux textes de systèmes de signes multimodaux. Ces représentations sont-elles modales, amodales, multimodales? À quel « niveau de profondeur » d'encodage?

Les vastes débats théoriques autour de ces différentes questions, et les justifications argumentaires ou expérimentales auxquelles ils ont donné lieu, sont extensivement exposés dans la thèse de Fortis ([1995]), en particulier, pour ce qui nous intéresse, dans le chapitre 5 (« Modalités et Amodalités »); et il ne saurait être question de résumer en quelques lignes l'ensemble de ces questions. Mentionnons simplement quelques points utiles à la clarification de notre propos.

La première distinction qui s'impose, avant d'aborder ce sujet, est celle qu'il faut garder à l'esprit entre modalité *mentale* (d'une représentation) et modalité *manifestée* (sur le plan de l'expression). Les deux sont bien entendu abordables d'un point de vue sémantique : on peut choisir d'étudier les impressions référentielles produites par telle modalité signifiante, quelle qu'en soit la nature, comme on peut s'interroger sur la multiplicité des modalités de représentations induites par un texte, quel que soit son support. Il est par ailleurs incontestable qu'aucune modalité n'a l'exclusivité d'un type particulier de contenu ; la paraphrase d'images en est le témoignage le plus évident.

Faut-il pour autant déclarer nul le lien entre ces deux « modalités »? Personne ne le prétend. Tout d'abord parce que la modalité mentale n'est somme toute jamais manifestée (quand on ne met pas en question sa réalité) que par des supports signifiants d'une certaine modalité. Cela seul prouverait que l'image mentale, par exemple, a un rapport privilégié, *au moins dans l'esprit de ceux qui en utilisent le concept*, avec l'image manifestée. Mais aussi, et ce second argument est moins sophistique, parce qu'il semble que certains types de représentation — et c'est pour cela même que l'on postule leur existence — aient un lien privilégié avec certaines formes d'expression, ou avec l'accomplissement de certaines tâches.

4.2.1 Existence de représentations cognitives modales

Il reste que l'existence même de représentations internes modalisées a fait l'objet d'un vigoureux débat, sur le terrain de la psychologie cognitive, entre les tenants

de l'existence « d'images mentales » et les partisans de l'hypothèse d'un codage propositionnel universel des représentations cognitives. Ce débat est exposé dans son historique par Denis ([1989], pp. 39–61). En voici très schématiquement les principales thèses.

À la sortie de la « préhistoire » — l'ère de la psychologie comportementale, qui ne se posait pas de questions sur la nature des représentations mentales —, et à la suite en particulier des travaux de Piaget ([1966]) sur le développement de l'intelligence symbolique chez l'enfant, on assiste à un regain d'intérêt pour l'image en tant qu'objet cognitif.

Paivio, appuyé sur les résultats d'expériences mettant en œuvre des tâches d'imagerie, conçoit et développe à cette époque le modèle du double codage. Celui-ci suppose que les informations liées à la perception de l'espace sont codées sous forme d'images mentales, invoquées tout particulièrement dans les activités d'imagerie, et coexistant avec des informations plus abstraites codées dans un format verbal (fig. 4.1).

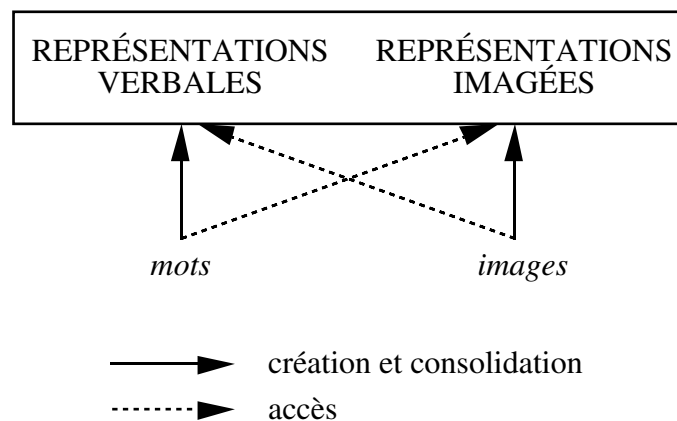


FIG. 4.1: double codage

Pylyshyn critique bientôt ce modèle au nom de l'école fonctionnaliste, pour qui le système cognitif fonctionne en manipulant les symboles amodaux abstraits d'un *langage mental*, dont les fondamentaux sont des concepts préexistant à toute perception. Selon lui, les résultats d'expérience interprétés en faveur de l'existence de représentations imagées peuvent s'expliquer autrement, sans remettre en cause la théorie du langage mental (pour schématiser cette théorie principalement exposée par Fodor [1975], voir fig. 4.2).

Enfin, plus récemment, des psychologues (notamment Kosslyn) élaborent et mettent à l'épreuve des modèles tendant à concilier l'existence de représentations imagées avec un codage symbolique abstrait ; les premières surgiraient ainsi à un niveau opératoire, lors de l'exécution de tâches précises pour lesquelles elles sont utiles, tandis que le second serait à l'œuvre à un niveau plus profond, celui de l'encodage à long terme de nos connaissances. On pourrait représenter ce type de modèles à deux niveaux comme dans la figure 4.3.

Ces derniers types de modèles, on le voit, ont accepté l'idée d'un système de représentation des connaissances propositionnel et amodal, laissant tout au plus

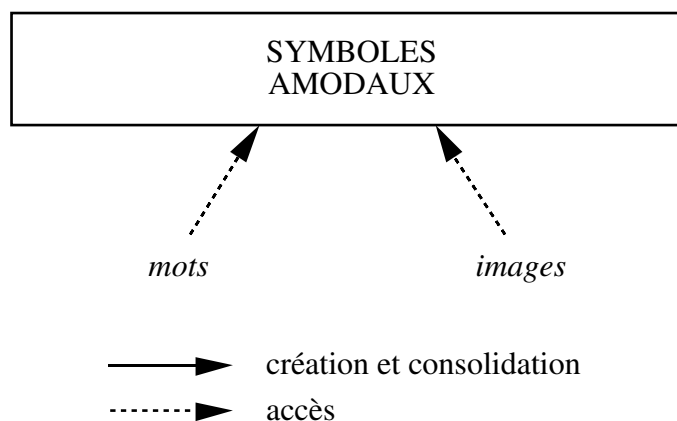


FIG. 4.2: langage mental. Dans cette théorie innéiste, les « concepts » ne sont pas créés par le contact avec la réalité sensible, ils lui préexistent — d'où l'absence de flèches pleines

à l'image un rôle d'outil, dans l'empire du courant fonctionnaliste dominant les sciences cognitives de leur époque. Ce dernier n'a toutefois jamais réussi à apporter de preuve décisive à l'appui de son credo (voir à ce sujet Fortis [1995], p. 127).

Accompagnant le regain d'intérêt, avec l'apparition et l'audience croissante des modèles connexionnistes, pour le caractère fondamentalement *associationniste* du cerveau humain, Rastier ([1991], pp. 207–212) met de son côté en avant la pertinence de l'hypothèse de représentations mentales essentiellement multimodales. On peut, avance-t-il, interpréter tous les résultats de la psychologie imagiste en faveur de cette hypothèse, et celle-ci est en outre corroborée par les acquis de la neuropsychologie (armée de ses nouveaux moyens d'observation du cerveau).

Selon lui, de telles représentations, baptisées *simulacres multimodaux*, comportent un amalgame de souvenirs perceptifs de différentes modalités, et sont convocables à partir de signifiants relevant de ces différentes modalités.

Le schéma d'accès au système des représentations mentales proposé par Rastier (op. cit., p. 210) prend donc une nouvelle forme (fig. 4.4).

C'est-à-dire, pour reprendre la forme des schémas précédents, qu'il peut être représenté, par comparaison avec les fig. 4.1, 4.2 et 4.3, comme sur la fig. 4.5.

Celle-ci, pour se vouloir homogène avec les figures précédentes, est plus réductrice que la fig. 4.4, où les flèches à double sens expriment la détermination du système sémantique psychologique sur les formes de la perception — autrement dit l'existence de *processus descendants* lors d'une *perception sémantique*.

La conception de Rastier s'inscrit dans une tradition de pensée qui est celle de Cassirer, lorsque celui-ci décrit justement le signe linguistique² en tant qu'élément organisateur de la conscience, qui se crée et se nourrit certes des données de la sensation, mais qui en outre organise cette sensation, lui donne forme et vie, la rend manipulable par l'esprit :

²Il s'agit là d'un signe-pensée — nous dirions en termes cognitivistes d'une « représentation mentale » — et non d'un signifiant.

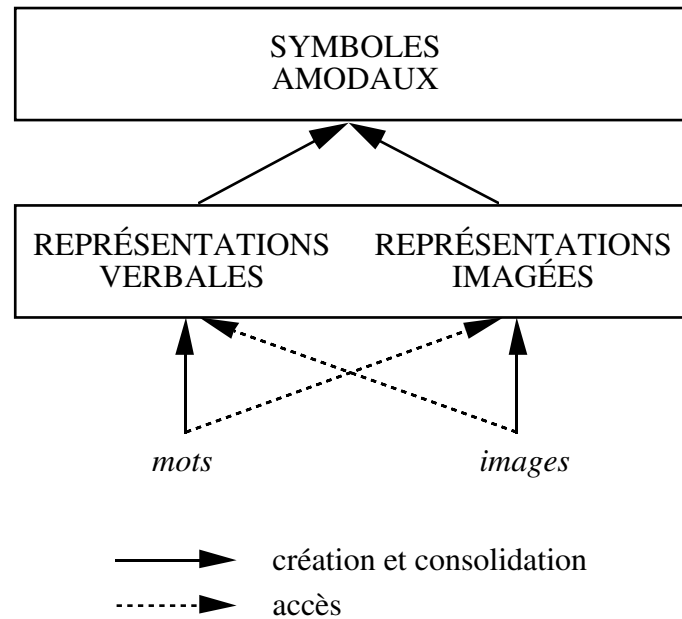


FIG. 4.3: deux niveaux de représentation

Dans tous [les] domaines [de la création spirituelle], de fait, le véhicule du développement qui leur est immanent se trouve dans la constitution d'un monde libre qui leur est propre, un monde de l'image à côté et au-dessus du monde de la perception : un monde qui, de par sa nature immédiate, garde encore toutes les couleurs du sensible, mais qui représente une sensibilité déjà informée et ainsi maîtrisée par l'esprit ([Cassirer, 1972], pp. 28–29).

Ces « images perceptives » actives qui entrent dans la formation des simulacres (Cassirer emploie aussi ce mot) ne sont en outre appréhensibles que comme parties du tout ; le concept de l'objet n'est pas la simple réunion de primitives d'origine sensorielle indépendantes, mais l'*intégration* d'un ensemble de propriétés n'étant conçues précisément que comme propriétés. Celles-ci, pour employer une métaphore mathématique, sont au concept ce que les éléments différentiels sont à l'intégrale : les premiers n'existent qu'en tant que parties de la seconde, et celle-ci n'existe qu'en tant qu'elle est constituée des premiers (idem, pp. 44–49).

Ce modèle rend compte des liens indissociables entre les modalités sans exiger l'acte de foi en un langage mental amodal dont l'existence est indémontrable.

4.2.2 Les représentations modales sont-elles liées à un mode particulier du signifiant ?

Cette question est à la racine de la distinction plus ou moins nette faite entre modalité « mentale » et modalité « manifestée », mentionnée plus haut comme l'un des premiers sujets de confusion et de polémique dans ce débat. Il importe maintenant de ne pas plonger aveuglément dans une autre confusion, celle qui pourrait surgir entre modalité de présentation aux sens de l'objet dénoté par un signe, et moda-

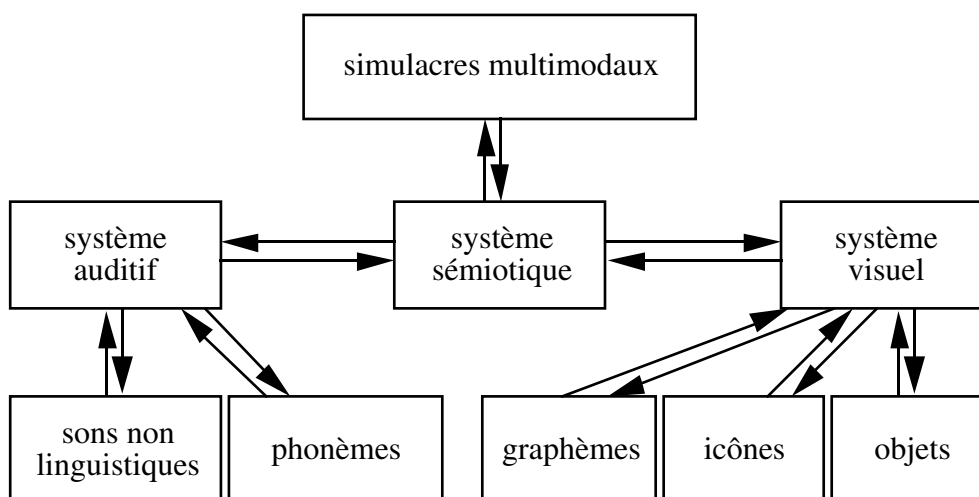


FIG. 4.4: le système sémiotique multimodal (Rastier)

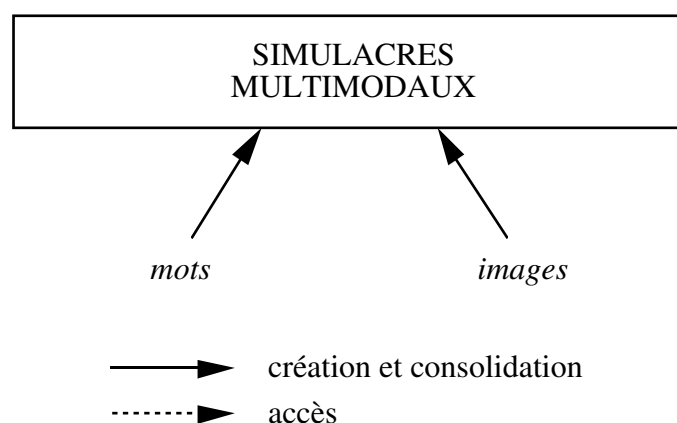


FIG. 4.5: représentations multimodales

lité de présentation du signifiant de ce signe dans un système de signes particulier. Autrement dit, entre modalités perceptives et modalités sémiotiques.

Note : *Y a-t-il des modalités perceptives au même titre que des modalités sémiotiques, ou le monde de la perception n'est-il peuplé que de sensations brutes, au maximum groupées selon leur canal sensoriel d'entrée? Si l'on admet, ce que tout nous porte à croire, que la perception n'est pas purement passive, mais construit ses présentations — par exemple, pour la vision d'une scène, par des stratégies d'exploration visuelle —, alors on a tout lieu de penser que des expériences de nature différente peuvent induire différentes façons de construire leur (re-)présentation. C'est ce que l'on a trouvé exprimé chez Cassirer ([1972], pp. 38–40), dans son introduction à la problématique des formes symboliques. Il est donc légitime de parler de modalité pour la perception et la reconnaissance d'un objet du monde « naturel », tout autant que pour la perception et la reconnaissance, par exemple, d'un mot de la langue ou de tout autre système de signes : en effet, dans le premier cas comme dans le second, il s'agit d'une stratégie de lecture et d'assemblage d'indices, et cette stratégie peut varier sui-*

vant le type d'objet. C'est en ce sens que Greimas et Courtès ([1979], article « Iconicité », p. 118), par exemple, peuvent parler de sémiotiques construites et de sémiotiques naturelles, au lieu d'opposer monde sémiotique et monde naturel³.

Examinons donc les arguments qui tendent à montrer une liaison entre représentations modales et modalités perceptives, en gardant à l'esprit la distinction que nous venons d'évoquer, et en cherchant à voir dans quelle mesure ils s'appliquent aux modalités proprement sémiotiques de présentation des signifiants dans des systèmes de signes particuliers.

1. Fortis ([1995], pp. 173–180), partant de la définition préliminaire suivante : « Selon la conception la plus simple, une représentation modale est une représentation déterminée par la modalité de l'*input* perceptif », expose les modèles expérimentaux, en psychologie, de systèmes à *entrée modale*, et de systèmes à *informations modales* (on retrouve la distinction modalité manifestée vs. modalité interne). Les données qui ont fondé le postulat de ces modèles de représentation « modale » sont pour l'essentiel des études de cas pathologiques d'aphasie ou d'agnosie. Les premiers de ces cas militent pour l'hypothèse d'un accès aux représentations mentales qui serait déterminé par la modalité des signifiants présentés ; d'autres cas portent sur des troubles d'accès à des catégories bien spécifiques, susceptibles de liens sémantiques forts avec la modalité visuelle, et ce quelle que soit la forme de la présentation.

Selon lui, rien dans l'expérience n'infirmerait l'hypothèse que ces deux modèles sont équivalents. Cette hypothèse reviendrait en fait à postuler un codage des représentations dépendant de la modalité des présentations perçues :

Les deux modèles peuvent être rendus équivalents si l'on pose que les informations visuelles, par exemple, conservent la modalité de l'entrée. Selon cette théorie, les informations composant une représentation complexe sont associées sous leur forme modale d'acquisition. Nous retrouvons ici l'idée de représentation plurimodale (Fortis, op. cit., p. 180)⁴.

La neuropsychologie et l'étude des aphasies ou agnosies révèle donc dans son approche spécifique les liens existant entre modalités de représentations mentales et modalités perceptives.

2. Denis, en introduisant son chapitre « Propriétés des images visuelles » (op. cit., pp. 63–102), consacré aux rapports entre image mentale et perception, rappelle dès l'abord ...

³En réalité, notre propos évolue dans une conception de l'univers sémiotique comme univers médiateur entre la conscience et le monde physique ; dans cette conception, la perception est précisément un phénomène sémantique (cf. [Rastier, 1991a])

⁴Fortis trouve d'ailleurs une formulation de cette thèse sous la plume de neuropsychologues : « L'importance relative ou les valeurs "pondérées" de canaux sensori-moteurs différents pourrait éventuellement former la base de la spécificité catégorielle dans le cerveau. Par 'valeur pondérée', nous voulons dire que l'information provenant de chaque canal contribuera quantitativement de manière différente au processus computationnel global impliqué dans l'acquisition initiale et, par suite, la compréhension de stimuli pourvus de sens (Warrington & McCarthy 1987, cité dans [Fortis, 1995], p.180) ». On y voit une esquisse de la structure interne des « simulacres multimodaux » de Rastier.

[...] la parenté « de surface » entre les événements psychologiques que produit l'imagerie et ceux que produit l'activité perceptive (idem, p.64).

Il souligne en outre l'ancrage de l'image dans les expériences perceptives qui en constituent forcément le « matériau de construction » :

[...] À cet égard, il n'est pas abusif de parler, plus que de dépendance ou de parenté, d'une *filiation* de l'imagerie à l'égard de la perception (idem, p. 65).

Puis il s'attache à démontrer, dans les pages qui suivent, l'existence de similitudes fonctionnelles et de similitudes structurales entre l'image comme représentation mentale opératoire et les processus perceptifs. Ces similitudes sont étayées par les résultats de nombreuses expériences de psychologie.

Pour un certain nombre de tâches cognitives, par exemple (mémorisation de mots isolés hors-contexte, mémorisation de structures spatiales, comparaison de taille ou de forme ...), on montre que les résultats obtenus par imagerie mentale, en l'absence d'images perceptives réelles, sont à peu près identiques à ceux que permettent d'obtenir ces images perceptives réelles. En outre, certaines expériences mettent en évidence une véritable relation d'*analogie* entre processus de manipulation d'images mentales et processus d'exploration de perceptions visuelles directes (l'esprit effectue par introspection des « parcours » mentaux sur des « cartes » mentales, ou des « rotations » mentales sur des objets mentaux ...)

Certaines classes de représentations cognitives présentent donc, à la lumière de la psychologie expérimentale, des relations frappantes avec certaines modalités perceptives.

Peut-on transposer ces considérations, qui concernent les modalités de perception directe des objets, aux modalités sémiotiques (sémiotiques *construites* pour être explicite) ? Autrement dit, les représentations modales peuvent-elles avoir un lien privilégié avec une certaine forme de sémiose mise en œuvre dans telle ou telle modalité, et pas avec une autre ? Dans une certaine mesure, oui.

1. Les systèmes de signes sont un mode supplémentaire d'adjonction de signification au monde de l'expérience ; par des règles fixées socialement, ils confèrent des modes d'interprétation à des objets produits pour signifier. Ces objets restent néanmoins, comme ceux perçus et interprétés « spontanément », des objets du monde réel. Les systèmes de signes, par leur plan de l'expression, restent partie du monde réel. *Les modalités sémiotiques et les modalités perceptives ne sont donc pas dans deux univers séparés, mais sont, dans le même univers, des modes de perception et de construction des représentations différents.*

À ce titre, déjà, il n'est pas étonnant de trouver des parallélismes entre le mode de fonctionnement de telle modalité sémiotique et de telle modalité « perceptive », la première se trouvant parfois imiter des processus mis en œuvre naturellement dans la seconde. C'est typiquement le cas du signe visuel iconique, qui fonctionne en tissant entre lui et son objet un type de *reconnaissance de*

formes analogue en certains points à celui qui se met en fonctionnement lors de la perception de l'objet réel (à ce sujet, cf. Eco [1968], pp. 111–112). Cette question est plus particulièrement traitée dans le chapitre 1 (« La nature de l'iconicité »).

2. Les expériences de manipulation d'images mentales, à ce sujet, mettent justement en évidence que certaines modalités sémiotiques contribuent à l'imagination, alors que certaines autres au contraire interfèrent avec elle. Ainsi la modalité iconique (présentation de dessins) semble-t-elle remplacer l'image mentale (elle fait aboutir aux mêmes scores sur un certain nombre de tâches de manipulation ou de rappel), alors que la modalité *linguistique écrite* a un effet notable de *perturbation* de ces mêmes tâches (cf. [Denis, 1989], p. 66, p. 107, p. 120, p. 186), en occupant le même canal sensoriel pour des processus de construction de sens différents (la modalité linguistique orale n'a pas l'effet perturbateur de l'écrit sans avoir l'effet facilitateur de l'icône).
3. L'un des exemples les plus notables est celui des *images motrices*. La participation d'un contenu moteur aux images mentales a été avancée par Piaget et Inhelder [1966]. L'image mentale, montrent-ils, est un outil cognitif qui se développe au cours des premières années de l'enfance. Ce n'est pas une simple trace de la perception, c'est la construction d'une représentation de l'objet à partir d'explorations de ses propriétés. À cet égard, l'image contient du figuré « perceptif » autant que de l'ébauche motrice. Pour toute la partie de la connaissance sur l'objet qui est relative aux façons dont le sujet pensant peut entrer en rapport avec lui, c'est cette représentation de type moteur qui peut être évoquée.

Or on constate en effet un lien fort et évident entre l'évocation d'activités motrices du sujet — et l'on sait comme ces activités tiennent un place importante dans la pensée et la conversation! —, et la modalité de communication gestuelle. Rimé et Schiaratura ([1991], pp. 265–271) décrivent comment les gestes coverbaux surgissent inconsciemment lors de récits d'événements ou d'actions mettant en jeu des schémas moteurs. Denis ([1989], p. 67) expose une expérience d'Engelkamp et Krumnacker (1980), qui montre que pour des tâches de rappel d'actions comme « jouer du piano » ou « lancer une balle », le pantomime, comme mnémotechnique, surpasse encore en efficacité (taux de rappel) l'image visuelle . . .

4. Inutile enfin d'entrer dans le jeu des exemples pour montrer l'ineffabilité de certaines formes sémiotiques dans d'autres systèmes de signes ; les cas de l'intraductibilité linguistique de signifiés musicaux ou picturaux en ont été abondamment servis.

Il n'échappe de toute façon pas à l'intuition qu'au-delà même des préférences individuelles, certains contenus sont plus facilement ou plus efficacement communiqués par un diagramme que par un paragraphe, par une formule mathématique ou logique que par une courbe, parfois par une poignée de mains ou un regard que par des phrases.

4.2.3 Les représentations modales sont-elles liées entre elles ?

Cette question prend un sens lorsque l'on a décidé, comme nous l'avons fait dans les paragraphes qui précèdent (§4.2.1 et §4.2.2 ci-dessus), de raisonner dans l'hypothèse de l'existence de représentations cognitives modales, et de liens privilégiés entre celles-ci et les modalités du signifiant.

En effet, premièrement, si tout signifié était directement un code propositionnel amodal, la question des modalités ne se poserait que sur le plan du signifiant ; il y aurait un plan du contenu amodal et homogène, et un plan de l'expression dont la disposition modale ou multimodale importerait finalement de façon indifférente à la construction du sens.

Deuxièmement, si la corrélation était nulle entre la modalité du signifiant et celle de la représentation évoquée, il serait en fin de compte dépourvu de sens de s'interroger sur les liens entre modalités mentales, puisque celles-ci ne seraient pas accessibles à partir des textes multimodaux ; les modalités du signifiant dans ceux-ci seraient dans un tel cas indifféremment traductibles dans n'importe quelles autres, ou ramenées à une seule, sans que l'on puisse rien dire sur l'univers signifié.

Si en revanche l'imbrication des modalités du signifiant peut avoir un prolongement dans l'ordre référentiel, alors la question devient pertinente.

Les résultats expérimentaux dont on dispose témoignent alors massivement, lors de différentes modalités mentales sont mises en œuvre, de l'associabilité de ces modalités.

1. Le cas de l'image mentale a déjà été abordé ; l'expérience montre que l'image est spontanément évoquée lors des processus de compréhension de textes linguistiques écrits ou oraux, lorsque le contenu de ces textes s'y prêtent. Cette évocation se fait à un plus ou moins grand degré selon les sujets (il existe des sujets spontanément « imageants », alors que d'autres ne s'y livrent que lorsque cela leur est explicitement demandé), et selon la nature d'une éventuelle tâche attendue, mais il apparaît dans tous les cas que la modalité linguistique peut donner lieu à l'évocation d'une représentation mentale imagée. Denis (op. cit.) propose de comprendre ce mécanisme comme la possibilité, pour chaque sujet en situation de compréhension d'énoncé, d'actualiser des *traits figuratifs* des concepts évoqués.

D'une manière plus générale, en situation de réception de messages dans telle ou telle modalité, la convocation de représentations relevant d'une autre modalité est toujours possible : si l'on admet les correspondances entre modalités mentales et modalités sémiotiques, ceci se trouve confirmé à chaque fois que l'on procède à la *traduction* d'une modalité dans une autre (paraphrase d'image, pantomime d'un récit, etc.)

2. En situation d'émission de messages, la convocation simultanée de plusieurs modalités mentales est également attestée par la gesticulation coverbale. Il faut en effet souligner deux points : (a) cette situation de communication a une valeur exemplaire car elle se fait « en temps réel » ; l'auteur du message ne se donne donc pas le temps d'effectuer des opérations de traduction — coûteuses — de modalités dans d'autres : en somme, il n'a pas le temps de

chercher ses mots jusqu'à un résultat qui le satisfasse entièrement de ce point de vue, et (b) le geste est la manifestation spontanée des modalités mentales imagées et motrices : il tient son caractère naturel de l'action, dont il est issu et dont il reste une ébauche, à travers le mécanisme de pantomime (Rimé & Schiaratura [1991], pp. 265–271). En considérant ces deux points, on se convainc que les représentations psychologiques d'ordre référentiel se situant en amont du processus de génération de messages sont multimodales.

David McNeill a consacré un article ([1985]) à la démonstration du caractère indissociable du geste et de la parole lors de la gesticulation co-verbale. Sans adhérer à toutes ses thèses, fondamentalement imprégnées du paradigme cognitiviste (sa formulation exacte est que « les gestes suivent un processus de calcul commun avec la parole », *gestures and speech share a computational stage*), on peut relever l'ensemble des éléments mis en avant pour prouver cette imbrication inextricable des deux modalités :

Les actes de parole sont souvent accompagnés dans notre culture de mouvements des bras et des mains appelés gestes. À une classe d'exceptions près — que l'on décrira plus loin —, ce sont des mouvements qui surviennent seulement pendant la parole, sont synchronisés avec des unités linguistiques, sont parallèles quant à leur fonction sémantique et pragmatique à ces mêmes unités linguistiques, remplissent des fonctions textuelles comme la parole, se dissolvent comme la parole en cas d'aphasie, et se développent conjointement à la parole chez les enfants ([McNeill, 1985], p. 351).

Les différents types de gestes survenant pendant la conversation sont plus particulièrement associés à différents types de contenu : les gestes « iconiques » figurent directement une action (gestes pantomimiques) ou dessinent un objet dans l'espace (gestes pictographiques) ; les gestes « métaphoriques » dessinent dans l'espace la représentation d'un concept abstrait — comme un enchaînement logique — ; les gestes rythmiques (*beats*), les seuls qui peuvent être faits avec la main non-dominante du sujet, se contentent d'accompagner le tempo de la phrase, en marquant les interruptions, les pauses, les corrections ... Or ces différents types de gestes sont affectés différemment par différents types d'aphasie : ainsi les gestes rythmiques sont-ils particulièrement affectés par les aphasies de Broca, où ils disparaissent presque complètement ; les gestes iconiques, en revanche, s'ils subsistent lors des aphasies de Broca, disparaissent dans les cas d'aphasie de Wernicke (aphasie sémantique), où ce sont les gestes rythmiques, au contraire, qui subsistent (*idem*, pp. 361–362).

Une autre constatation frappante est celle de l'émergence simultanée du geste et de la parole chez l'enfant : les gestes déictiques⁵ et pantomimiques se développent en premier lieu, en même temps que les premiers mots isolés à fonction référentielle, alors que les gestes rythmiques apparaissent au dernier stade de l'évolution de la parole, celui de la cohérence textuelle, aux alentours de cinq ans.

⁵Le geste déictique étant ontogénétiquement une pantomime : une évolution du geste consistant à essayer d'atteindre quelque chose avec la main.

Dans le même sens, on citera ce résultat d'une expérience de Rimé : lorsque des sujets sont forcés de parler sans pouvoir faire de gestes (avec les mains et les pieds liés aux montants du fauteuil), (a) on note une suractivité importante des muscles faciaux (au niveau des mâchoires, des joues, des sourcils ...), et (b) le contenu du discours produit est moins imagé que lorsque le sujet peut gesticuler⁶ ([Rimé & Schiaratura, 1991]).

Les différents faits mis en avant semblent conduire à l'idée que le geste co-verbal est d'une part spontané, voire compulsif⁷, et qu'il est d'autre part inséparable de sa contrepartie verbale — qu'à un certain niveau en quelque sorte, les phénomènes psychologiques accompagnant la conversation gesticulée sont homogènes, et que l'on ne peut en reproduire de façon isolée la « partie verbale » ou la « partie gestuelle » (c'est le "*shared computational stage*" de McNeill).

3. Entre deux modalités linguistiques, l'écrit et l'oral, il semble encore que des représentations se forment à un tel niveau d'imbrication qu'il est impossible d'évoquer les unes sans les autres. On voit apparaître ceci par des effets indirects, qui témoignent de l'intervention de représentations de mots écrits lors de tâches de compréhension purement orales.

Ainsi des expériences ont-elles été faites au Portugal sur la répartition des types d'erreur de reconnaissance de mots enregistrés (Morais et al. [1991]). Cette répartition est notablement différente dans un groupe de sujets analphabètes et dans un groupe de sujets formés à la lecture et à l'écriture. Les seconds font proportionnellement moins d'erreurs « segmentales » (mauvaise reconnaissance d'un phonème isolé) que les premiers, mais plus d'erreurs « globales ». Autre différence du même type, les sujets analphabètes réussissent en moyenne beaucoup moins bien des tâches de commutation de phonème que les sujets « lecteurs ».

Enfin, le phénomène de « fusion phonétique » a une incidence différente chez les sujets lettrés et chez les sujets illettrés. Ce phénomène consiste à « entendre » au début d'un mot une consonne initiale qui n'y est en réalité pas, lorsque (a) l'ajout de cette consonne au mot entendu produit un autre mot signifiant, et (b) le *remplacement* de la vraie consonne initiale du mot entendu par cette consonne produit également un autre mot signifiant. On peut par exemple entendre « blanc » alors que le mot prononcé est en réalité « lent », par fusion avec la forme disponible en mémoire « banc ». Or l'expérience a permis de constater chez les sujets formés à la lecture un effet inhibiteur de ce phénomène lorsque l'*orthographe* du fusionnat ne ressemble plus à celle du fusionnant et à celle du fusionné — même si cela ne fait pas de différence à la prononciation. Ainsi ces sujets font-ils nettement moins les fusions *fiz (liz) → feliz*, ou *par (lar) → pelar*, que des fusions comme *cara (lara) → clara*, ou *pena (lena) → plena* (le "e" des premiers exemples ne s'entend plus dans la prononciation du portugais d'Europe) ; alors que chez les sujets illettrés, le taux d'erreur est le même dans tous ces cas.

⁶ « *Samples of the dialogues were submitted to a computerized technique of content analysis, and a significant decrease was recorded for vividness of imagery during movement restriction, as compared with baseline and recovery period* » ([Rimé & Schiaratura, 1991], p. 241).

⁷ On peut pour s'en convaincre observer les mouvements de celui qui est engagé dans une conversation animée au téléphone.

Il semble donc que la disponibilité à l'esprit de représentations graphiques des mots, rendue possible par l'alphabétisation, influe même sur des phénomènes de reconnaissance de mots oraux, qui sont quasiment de l'ordre de la perception.

Ces diverses études reconduisent en somme à l'idée que les modalités mentales sont imbriquées à un degré tel qu'il est plus juste de parler de *représentations multimodales* que de double codage. À l'appui de cette conception, on ne voit en outre pas de quel droit il conviendrait de maintenir dans l'ordre référentiel des systèmes de signes une stricte démarcation entre modalités que l'on n'a par ailleurs aucun moyen de mettre en évidence au niveau sémantique (ordres syntagmatique et paradigmatic) : nous renvoyons le lecteur, pour une discussion sur ce sujet, à Rastier [1991], pp. 214–216.

4.3 Ordres syntagmatique et paradigmatic

La mise en évidence d'une multimodalité « foisonnante » lors des phénomènes d'ordre référentiel accompagnant la composition ou l'interprétation de textes ne doit en effet pas nous conduire à la conception naïve d'une sémantique hétérogène, où chaque modalité de l'expression serait spécialisée dans la signification d'un certain type de contenu.

Greimas observe que « les signifiants de nature sensorielle différente peuvent recouvrir un signifié identique ou, du moins, équivalent : ainsi la langue orale et la langue écrite » ([Greimas, 1986], p. 11), et que :

[...] Quel que soit le statut du signifiant, aucune classification de signifiés n'est possible à partir des signifiants. La signification, par conséquent, est indépendante de la nature du signifiant grâce auquel elle se manifeste. Dire, par exemple, comme cela se fait assez couramment, que la peinture comporte une signification picturale ou que la musique possède une signification musicale n'a pas de sens. La définition de la peinture ou de la musique est de l'ordre du signifiant et non du signifié. Les significations qui y sont éventuellement contenues sont simplement humaines. Tout au plus peut-on dire que le signifiant, pris dans son ensemble, comporte le sens global « peinture » ou « musique » (ibid.)

Cette position peut-elle être vue comme contradictoire avec le fait d'admettre des « liens privilégiés » entre certaines représentations et certaines modalités signifiantes, comme nous le faisons plus haut ?

- Non, si l'on prend garde à ne pas confondre l'ordre référentiel et l'ordre paradigmatic. Greimas se donne pour cadre de travail, comme le titre même de son ouvrage l'indique, une sémantique structurale, et non une sémantique référentielle. L'affirmation de l'indépendance du signifié à l'égard du signifiant ne présume donc en rien de la nature des phénomènes mentaux qui se déroulent lors de la composition ou de l'interprétation des textes.

- Nous avons d’ailleurs souligné plus haut le fait qu’aucune modalité n’a l’exclusivité d’un type de contenu ; la langue en particulier peut dans une certaine mesure suppléer par des mots à l’absence d’un autre mode d’expression (par exemple du dessin).
- Nous proposons donc de comprendre la prise de position de Greimas, un peu extrémiste certes, au moins dans sa formulation, comme une extension à la sémantique multimodale de la notion saussurienne de *valeur* : flottante et arbitrairement liée au signifiant.

Il serait absurde de déclarer nulle l’importance de la modalité choisie pour être porteuse de tel ou tel sens. À l’évidence, on ne choisit pas indifféremment son mode d’expression ; celui-ci est le résultat de la nécessité parfois, du hasard jamais. Admettons néanmoins avec Greimas que le contenu d’un texte relevant d’une unique modalité peut s’étendre sans restriction autre que celle imposée éventuellement par la combinatoire du système de signes employé. Qu’en est-il dans le cas d’un texte multimodal ?

- (a) Les textes monomodaux imbriqués ont des contenus différents.

C’est évidemment le cas, dans la mesure où cela justifie l’existence de plusieurs modalités.

Une sémantique référentielle pourrait avancer comme objection le contre-exemple des « dictionnaires » de traduction d’une modalité dans une autre, comme par exemple les trombinoscopes, les légendes des cartes ou des schémas techniques . . . C’est dans ce cas naturellement la fonction pragmatique de *matériau d’apprentissage* qui justifie la redondance⁸ entre les systèmes de signes.

- (b) Le sens de chaque modalité est influencé par celui des autres.

Il se produit dans les textes multimodaux, *entre* modalités différentes, des phénomènes de sens contextuel comme ceux qu’on trouve en langue. Un exemple canonique en bande dessinée est fourni par le langage des « schtroumpfs » du dessinateur Peyo : comme le fait remarquer Rey, « pour schtroumpfer avec un Schtroumpf, il suffit de savoir de quoi qu’on cause : que le langage en soit responsable (j’en ai ras le schtroumpf, schtroumpfé soit qui mal y schtroumpfe) ou bien l’image, [ce jargon] manifeste comiquement le va-et-vient du sens dans la bande » ([Rey, 1978], p. 171).

Dans le cadre de description monomodal qu’on utiliserait pour analyser isolément l’une des sous-modalités, l’intervention des autres sous-modalités dans le sens de la première pourrait s’expliquer par la notion d’*interprétant* : elle relèverait ainsi de la pragmatique, c’est-à-dire de l’étude des conditions d’interprétation *externes* au texte.

⁸Redondance parfaite, justement, dans une théorie sémantique référentielle : le *sens* du symbole pour « pont » dans une légende de carte géographique s’identifie alors avec la classe d’objets « pont », et chaque occurrence de ce symbole sur la carte aura pour *denotatum* un élément de cette classe ; dans le trombinoscope, le *sens* de la photographie du visage et celui du nom propre ne se distinguent pas : ils s’identifient tous deux à la classe à un seul élément qui constitue l’*essence* de l’individu dénoté. Il va de soi qu’une sémantique interprétative donne une vision non pas contradictoire, mais totalement différente, du sens, dans laquelle cette « redondance » référentielle n’est pas l’horizon de toute description sémantique. Ainsi cette sémantique permettrait-elle par exemple de comprendre pourquoi l’on peut dire d’un visage fin, avec de petites lunettes à monture dorée, qu’il « colle bien » avec un prénom composé et un nom à particule . . .

Mais dans le cas de genres essentiellement multimodaux, il est possible de tenter une description unifiée du texte multimodal, et de ramener ainsi ce phénomène à un phénomène de lecture, c'est-à-dire d'élaboration du sens, qui serait *interne* au texte. Cette approche paraît d'autant plus pertinente qu'elle s'applique à des genres ayant eu le temps de développer leurs codes, et donc certains mécanismes de lecture des *syntagmes* multimodaux : bande dessinée, cinéma, etc.

En bref, et pour aller directement à la synthèse de la conception de la sémantique multimodale qui structure cet exposé, *les modalités se partagent le contenu*.

Des exemples vont nous permettre d'illustrer cette conception. Il s'agit de textes humoristiques appartenant à deux genres multimodaux : le dessin d'humour et la bande dessinée.

Nous rapprocherons volontiers tout d'abord ces textes humoristiques multimodaux des textes humoristiques linguistiques, pour montrer comment des fonctions identiques sont pris en charge différemment dans les deux situations. Nous nous référons pour cela à une approche des « mécanismes » sémantiques de l'histoire drôle introduite par Greimas ([1986], pp. 70–71) et approfondie par Morin [1981]. Selon ces analyses, le « plaisir » de l'histoire drôle naît de la découverte d'un second degré de lecture d'un certain récit, de l'identification d'une seconde isotopie : Violette Morin l'appelle le « récit parasite ». Reprenons pour fixer l'esprit sur un exemple précis l'histoire drôle de Greimas :

« *C'est une brillante soirée mondaine, très chic, avec des invités triés sur le volet. À un moment, deux convives vont prendre un peu l'air sur la terrasse :*

— *Ah ! fait l'un d'un ton satisfait, belle soirée, hein ? Repas magnifique . . . et puis jolies toilettes, hein ?*

— *Ça, dit l'autre, je n'en sais rien.*

— *Comment ça ?*

— *Non, je n'y suis pas allé ! »*

L'histoire drôle est donc structurée autour soit d'une homonymie (jeu de mots — c'est le cas ci-dessus), soit du récit d'un fait qui peut pragmatiquement recevoir deux interprétations opposées. Le jeu tactique pour faire éclater la double lecture en une situation comique consiste (a) en la mise en place d'un contexte initial, fixant la lecture « première » (la *fonction de normalisation* de Morin), et (b) en la mise en évidence, l'activation, de la seconde isotopie dans la dernière partie du récit (la *fonction de disjonction* chez Morin, qu'on appelle en langage courant la « chute » — parce qu'elle est *brutale* comme une chute).

Cette structure mise systématiquement en place pour créer les conditions de la double lecture est conditionnée par les caractéristiques de la modalité du système de signes employé (la langue). La linéarité de celle-ci oblige en effet à une disposition tactique pour faire apparaître le contraste qui est le noyau de l'histoire drôle. Non que l'on ne puisse dire que les deux lectures sont présentes simultanément dans le texte linguistique ; simplement, l'interprétation est capable d'activer des sens et d'en inhiber d'autres, et c'est tout le jeu justement de la lecture que de canaliser tout

d'abord vers une isotopie, créant ainsi un sens « premier » qui est lui aussi, en soi, contextuel, pour obliger d'un coup à l'invalider par un aiguillage vers l'isotopie opposée.

1. Dans le dessin d'humour, les caractéristiques de la modalité créent des conditions toutes différentes. Tout d'abord, le dessin d'humour, que nous distinguons ici bien de la bande dessinée, n'a pas de développement syntagmatique au niveau de la macro-modalité : il n'est pas décomposable en figures (segments homogènes s'assemblant dans une dimension linéaire déterminée), comme l'est par exemple la bande dessinée avec ses *cases*. Il constitue donc à lui seul un texte-signe, qui est en revanche décomposable en *caractères* de sous-modalités différentes : le texte (linguistique) et le dessin.

Or la coprésence de ces deux sous-modalités hétérogènes dans le même message induit une stratégie de lecture et d'interprétation de nature différente de celle des textes linguistiques. Ici, les deux sous-modalités peuvent par exemple porter *simultanément* des messages contradictoires ; la divergence des deux lectures est donc parallèle puisqu'elle ne peut être en série. C'est le cas du dessin de Sergueï (fig. 4.6), publié dans *Le Monde* du 28 septembre 1996 : l'un des acteurs (les États-Unis) proclame par le texte une inquiétude sous-entendue, celle de l'existence d'une *menace importante* (expression “*faire un malheur*”) que feraient peser les Arabes palestiniens sur Israël ; or cette menace importante est contredite par le dessin, qui figure un Palestinien armé d'un lance-pierres faisant face à un Israélien armé d'une mitrailleuse automatique. Le contraste est accentué par une variation de taille, dans le dessin, entre les deux acteurs du conflit⁹.

2. Autre exemple : une variation du mécanisme comique de « récit parasite » réside dans le principe de *fausse justification*. Ainsi Morin décrit-elle une classe d'histoires drôles « à disjonction référentielle » (ce n'est pas un mot homonyme qui fait diverger la lecture, mais une interprétation pragmatique que donnent les acteurs du récit), où ...

[...] l'interlocution confirme la locution par une preuve qui l'infirmé ou inversement l'infirmé par une preuve qui la confirme

⁹Ce dessin appartient à un genre spécifique, celui du dessin d'humour commentant des situations politiques. C'est un genre qui pratique une transposition systématique de domaine, rendue obligatoire par la nature souvent abstraite ou collective des concepts évoqués : ainsi l'affrontement d'un état et d'une population civile est-elle représentée par le combat singulier de deux individus, qui est, lui, *dessinable*. Les transpositions touchent donc aussi bien les acteurs (entités abstraites représentées par des individus : « Oncle Sam » pour les États-Unis d'Amérique) que les processus (processus complexes représentés par des actions concrètes : la paralysation diplomatique par des mains saisissant physiquement des bras pour les empêcher de bouger), et sont observables dans les deux sous-modalités (en langage diplomatique officiel, on ne dit pas : « Je vous retiens avant que vous ne fassiez un malheur », c'est une expression du langage courant réservée aux disputes de cours de récréation). L'interprétation métaphorique est guidée par un ensemble de symboles appartenant à une iconographie spécifique (Oncle Sam, le lance-pierre contre le fusil, etc.), qui, catachrétisées, en constituent des « îlots de confiance ».

Certains même parfois en abusent (ainsi un dessinateur du même quotidien dont les dessins ne sont plus que des assemblages mécaniques de Mariannes-éplorées, de colombes-de-la-peace-transpercées-en-plein-vol ou de fascistes-à-chemise-brune-avec-un-brassard-noir, qui pourraient être engendrées par ordinateur à partir d'une structure de type frame avec une bibliothèque d'icônes adéquate).

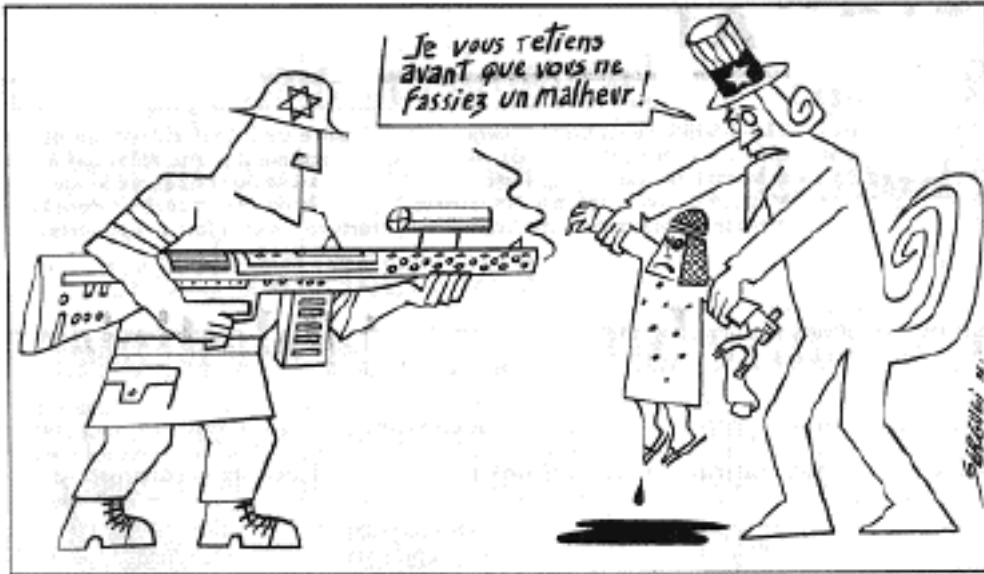


FIG. 4.6: dessin de Sergueï

([Morin, 1981], p. 114).

Exemple de ce ressort :

L'élève doit recopier 100 fois : je ne sais pas compter.

L'instituteur : Pourquoi l'avez-vous recopié seulement 30 fois ?

L'élève : je ne sais pas compter, Monsieur (idem, p. 113).

L'attribution de liens de causalité fantaisistes est à la base de ce type d'histoires ; elle peut être soit volontaire (histoire du charlatan qui vend des têtes de hareng pour rendre intelligent, et qui déclare au client insatisfait venu l'accuser d'escroquerie : « vous voyez que cela commence à marcher »), soit involontaire (histoire du fou lançant des peaux de banane de la fenêtre du train Paris-Orléans pour éloigner les tigres, qui démontre l'efficacité de sa méthode par l'absence de tigres entre Paris et Orléans).

Cette même attribution de liens de causalité fantaisistes sert de ressort au dessin de Jean-Vincent Sénac (fig. 4.7). Le « dégonflement » du raisonnement apparemment sérieux, cependant, au lieu d'être simplement successif à celui-ci dans un texte homogène, met en jeu deux composantes parallèles ; c'est par un aller-retour entre ces deux composantes modales que l'on prend conscience de son caractère absurde.

La crédibilité du conseil donné dans la première partie du texte linguistique s'effondre en effet, on le voit, par le double examen (a) de la justification qui en est fournie par un autre syntagme de la même modalité, et (b) de la personne du conseiller, que l'on observe dans le dessin. C'est en effet un lapin, donc un animal qui a par la nature de son espèce de grands pieds, de grandes oreilles et de grandes dents, qui explique au lecteur que la consommation de carottes est bénéfique pour ces trois organes. Si ce n'était pas un lapin, nous pourrions prendre le texte au sérieux : qui nous empêche de croire en effet, *a priori*, sans connaissances profondes en biologie, que les carottes contiennent



FIG. 4.7: *Régime carotte*, par Jean-Vincent Sénac

des substances bénéfiques pour les oreilles, les dents, les pieds ? D'autre part, sans la justification contenue dans la seconde phrase du texte, le message de l'ensemble ne nous paraîtrait pas *aberrant* comme il le fait là : il nous semblerait presque « naturel » de voir un lapin vanter les carottes, puisque nous savons que les lapins aiment les carottes ; il ne resterait plus de « pas sérieux »¹⁰ dans ce dessin qu'une figure rhétorique un peu fantaisiste consistant à faire parler un animal (une prosopopée). Ce qui est drôle, donc, au bout du compte, c'est que le lapin *justifie* son goût pour les carottes et son prosélytisme par l'attribution à leur consommation de la taille remarquable de ses oreilles, de ses dents et de ses pieds, alors que nous savons que les lapins naissent avec de grandes oreilles, de grandes dents et de grands pieds — comme nous savons qu'il n'y a pas de tigres entre Paris et Orléans.

Nous ne pouvons donc bâtir une lecture de chacune des sous-modalités qu'en faisant intervenir des interprétants issus de l'autre sous-modalité. Proposons donc plutôt une grille de lecture unifiée du texte multimodal, où nous représenterons les interactions de sens qu'elles soient intra- ou inter-modales (fig. 4.8) :

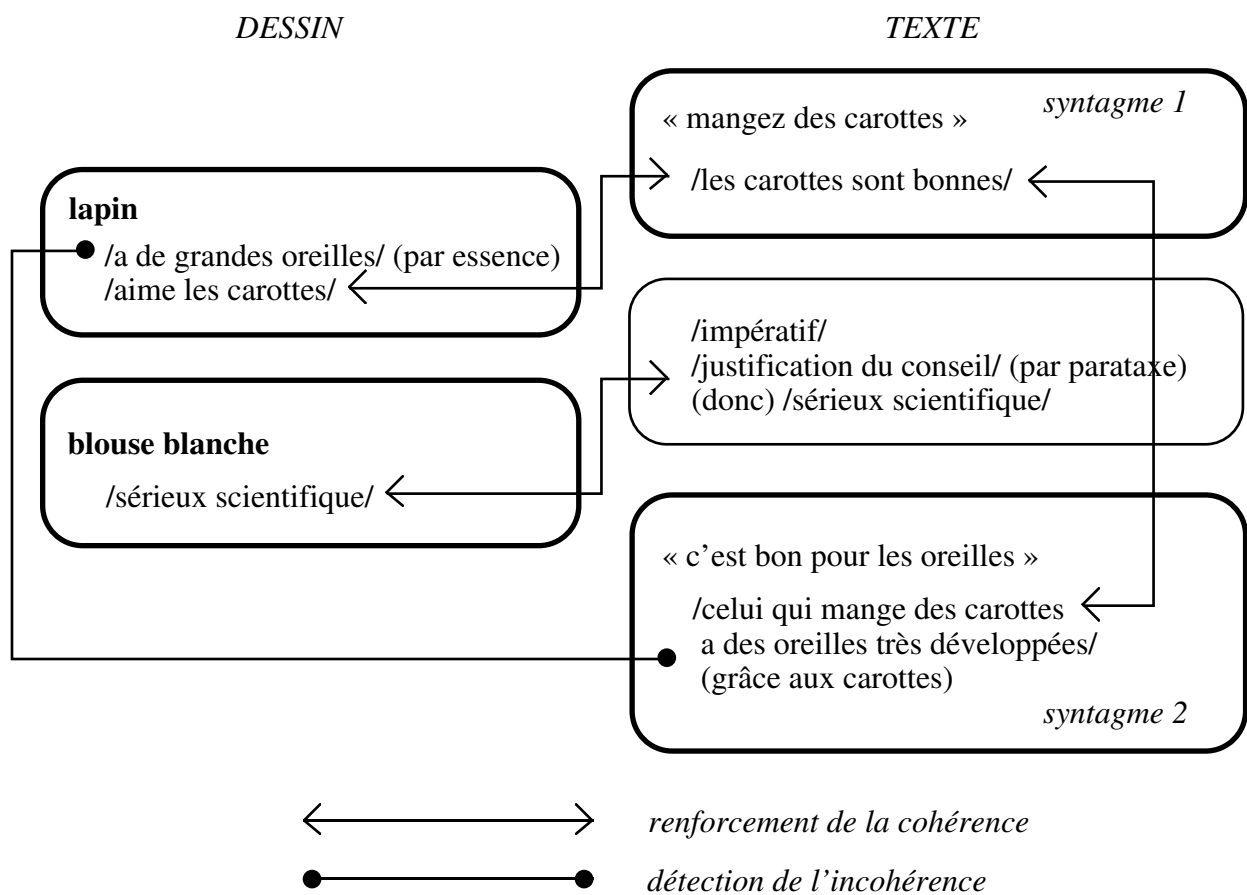


FIG. 4.8: cohérence et incohérence dans le dessin de Sénac

L'équivalent purement linguistique de ce dessin d'humour pourrait se cons-

¹⁰Mais il existe des messages publicitaires télévisés où des lapins conseillent de manger des carottes, et qui sont réalisés dans des buts tout à fait sérieux.

truire ; il faudrait y disposer, dans des syntagmes linguistiques, et de manière linéaire, les informations qui sont ici contenues dans le dessin. Cela pourrait être par exemple : « C'est un lapin en blouse blanche qui dit : "Mangez des carottes! C'est bon pour les oreilles ..." ». Nous ne nous attarderons pas sur les difficultés qui surgiraient si l'on voulait alors renforcer la saillance des grandes oreilles chez le lapin, chose qui se fait aisément dans le dessin en les allongeant un peu plus, par exemple.

3. Enfin, examinons une petite bande dessinée qui illustre bien l'exploitation que peuvent faire de la multimodalité les genres autorisant un développement syntagmatique (fig. 4.9).



FIG. 4.9: *Tracer Bullet*, extrait de *The Calvin and Hobbes Tenth Anniversary Book*, par Bill Waterson (Londres : Warner Books, 1995).

On comprend immédiatement en lisant cette courte bande dessinée que l'auteur joue sur un contraste entre la vie quotidienne d'un enfant de cinq ans et la vie imaginaire suscitée et entretenue en celui-ci par la lecture de récits d'aventures (ici policières). Ce qui invite à se souvenir du commentaire de Greimas sur les contrastes privilégiés des histoires drôles :

[...] un premier classement superficiel montre la préférence des bons mots pour les oppositions du type :

adultes	vs	enfants
majoritaires	vs	minoritaires
normaux	vs	fous
humains	vs	animaux ou choses, etc.

([Greimas, 1986], p. 71)

C'est la même opposition anthropologique fondamentale entre univers adulte et univers enfantin qui est mise en scène dans la bande dessinée *Tracer Bullet*. L'univers familial de la salle de séjour d'une maison ordinaire, où une mère excédée découvre une bêtise et demande — de façon purement rhétorique, car Calvin est enfant unique — qui en est le coupable, bascule dans l'univers des romans policiers de série noire, à la Peter Cheyney, où un détective cynique, cupide et alcoolique se charge d'enquêtes particulièrement périlleuses pour le compte de fragiles et séduisantes créatures.

Nous avons ici cependant une structure signifiante complexe, où les traits caractéristiques /adulte/ (détective dur à cuire) et /enfant/ (petit garçon fantaisiste) apparaissent dans presque tout le récit parallèlement. Alors que dans l'histoire drôle telle que l'analyse Morin, la découverte, au cours d'une disjonction brutale, du terme masqué jusque là par son terme opposé marque la fin de l'histoire, cette bande dessinée, elle, laisse traîner de la troisième case à la fin un hiatus comique entre la situation réelle et la situation imaginée.

Le détective « Tracer Bullet » fait en effet irruption à la troisième case, par une technique de changement de décor qui s'apparente aux techniques cinématographiques (*flash back* et retour au narrateur), mais en gardant la taille et la figure du petit garçon Calvin. À la dernière case enfin, malgré le second changement de décor qui ramène à la confrontation du petit garçon et de la mère, le texte continue dans le style du roman policier.

Le premier changement de rails du récit est introduit par le « Who did this ? » de la mère, interprété comme s'il s'agissait véritablement de découvrir la solution d'une énigme : c'est l'élément disjoncteur de l'histoire. Le détour par l'autobiographie du détective Tracer Bullet construit une image de dur qui a connu beaucoup d'épreuves et ne recule devant rien. Enfin, l'acceptation de l'autorité maternelle par le petit garçon (« I took the case ») est justifiée bouffonnement, à la manière des romans policiers, par la cupidité des détectives privés et leur sensibilité au charme féminin (« a tall brunette »), à laquelle est également implicitement attribuée la frayeur de l'irruption d'une mère en colère (« my heart did a few calisthenics »).

Une partie du *plaisir de lecture* de cette bande dessinée provient en fait de la reconnaissance par le lecteur d'une série de signes connotant l'univers de la série noire, que l'auteur a pris la peine de faire nombreux et détaillés (chapeau, imperméable, jeu de cartes, brasseur d'air, cigarettes, bouteille d'alcool, pistolet, téléphone style *rétro*, . . . jusqu'à l'éclairage en contre-jour à travers les jalousies !) Il ne s'agit plus seulement de faire émerger un récit parasite qui balaye brutalement le récit « premier » et met fin à l'histoire, mais de laisser se développer une double lecture comique, par la disproportion des domaines transposés. Un récit purement linguistique fondé sur les mêmes ressorts — il est naturellement possible d'en écrire — ne serait plus une histoire courte mais un récit assez long, devant rendre par des moyens linguistiques ce qui est ici véhiculé par le dessin.

Il y a en somme une grammaire dans les systèmes de signes multimodaux, c'est-à-dire non seulement une façon d'associer les unités signifiantes, mais aussi, sur le plan du contenu, des modes de polarisation et de circulation du sens analogues à

ceux que l'on peut mettre en évidence dans des textes linguistiques.

Il serait sans doute abusif de parler de *syntaxe* multimodale, dans la mesure où la syntaxe est la façon dont la disposition des éléments conditionne le sens. S'il y a donc, très vraisemblablement, une syntaxe dans les systèmes de signes multimodaux, ce n'est pas pour autant une *syntaxe multimodale* dans la mesure où elle règle la disposition d'unités homogènes, les signes de la macro-classe, et non pas celle d'unités hétérogènes.

En ce sens, on ne pourrait véritablement parler de syntaxe que dans des cas comme celui de la langue japonaise écrite, où les *kana* s'intercalent avec les *kanji* dans la même dimension linéaire, et où leur disposition mutuelle conditionne le sens. Nous savons bien entendu que la syntaxe bien définie de ce système de signes multimodal est héritée de celle de la langue japonaise parlée, dont il est historiquement un système hybride de notation. Nous nous sommes interdit de déclarer *a priori* qu'une langue écrite est forcément le calque d'une langue orale (cf chap. 2) ; il faut néanmoins bien reconnaître ici un cas un peu particulier, qui fait de ce système de signes multimodal un système à syntaxe parce qu'il est transposable dans un système monomodal donné.

La plupart des autres systèmes pratiquent, nous l'avons vu, l'entaxe, c'est-à-dire qu'ils disposent des unités signifiantes des sous-modalités à l'intérieur des figures de la modalité super-ordonnée. Les syntagmes des sous-modalités sont des caractères des figures de la macro-modalité. Une fois ce seuil franchi, en revanche, les sous-modalités ne partageant pas le même espace de déploiement, la question de leur disposition mutuelle ne se pose pas. Ainsi, dans une bande dessinée, s'il est important pour le sens qu'une légende soit rattachée à une case ou à la suivante, il est au contraire indifférent la plupart du temps qu'elle se situe du côté gauche ou du côté droit de la case. De même dans une conversation gesticulée, s'il est important de rattacher un geste à un syntagme linguistique donné, on a par contre une latitude sur le moment de début de ce geste et sur sa durée, tant qu'il reste dans les limites du syntagme.

Mais refuser la formule facile de « syntaxe multimodale » n'empêche pas de constater la richesse du déploiement *syntagmatique* du sens dans des systèmes multimodaux. Comme nous l'ont montré les quelques exemples ci-dessus, l'apport de sens d'une sous-modalité vers l'autre, à l'intérieur d'une même figure de la macro-modalité comme à travers les limites de cette figure, sont permanents.

Ces différents exemples confirment la validité d'une théorie de lecture interprétative des textes multimodaux, dont le mécanisme fondamental serait comme en langue l'afférence.

4.4 Ordre herméneutique

On retrouve en effet dans les systèmes de signes multimodaux les phénomènes d'activation de traits sémantiques à longue portée qui sont à la base de la constitution d'un « contexte global » fournissant le fond sur lequel s'interprètent les unités sémantiques locales. On retrouve sans étonnement en somme le cercle herméneu-

tique, par lequel le sens des unités permet de construire le sens du texte, qui lui-même détermine le sens des unités. Il n'y a pas là de « cercle vicieux » mais ni plus ni moins qu'un phénomène de recherche d'équilibre dans un système complexe, pouvant donner lieu à des réitérations — ici des réinterprétations (ce thème a été abondamment développé par [Rastier, 1987], en particulier chapitre IX).

Comme nous l'avons écrit plus haut, l'interprétant d'un texte relevant d'une modalité peut parfois lui-même relever d'une autre modalité. Dans un système de signes fondamentalement multimodal, on peut avec profit décrire de façon unifiée les syntagmes de ce système, en évitant de renvoyer les sous-modalités considérées comme « non-principales » au rang d'interprétants, mais en les faisant au contraire concourir directement au parcours interprétatif. Cette démarche est légitime, puisqu'on en fait après tout autant dans les textes linguistiques écrits, en faisant concourir simultanément signes alphabétiques et signes de ponctuation à l'élaboration du sens (il y a déjà là une multimodalité, comme le souligne Rastier [1994], p. 212).

Il reste cependant des interprétants extra-textuels, et là encore ceux-ci peuvent être d'une autre modalité que le texte lui-même. En sortant du texte, on passe alors de la *multimodalité* à l'*intermodalité*.

Les interprétants extra-textuels peuvent être par exemple des facteurs pragmatiques liées à la situation d'énonciation. Il en va ainsi de l'énoncé « Il fait froid », qui peut être interprété comme « Fermez la fenêtre » si l'énonciateur est en situation de supériorité sociale par rapport au destinataire : l'interprétant relève alors d'une sémiotique ethnologique des groupes sociaux ([Rastier, 1987], p.230).

Mais on peut étendre cette importance des interprétants extra-textuels à tout un ensemble de connaissances encyclopédiques qui s'imposent irréductiblement à tout membre d'une société humaine, quelle que soit son époque. Il y a autour de chaque texte, quelle que soit sa modalité, un vaste entour intersémiotique qui conditionne la compréhension et l'interprétation ; c'est l'absence de cet entour qui fait qu'un néophyte européen d'aujourd'hui ne « comprend pas », n'arrive pas à attacher de sens à, une peinture chinoise antique (nous en parlions au chapitre 1, au sujet de la « perte de ressemblance » des signes iconiques avec leurs modèles).

Sebeok, dans une étude réalisée en 1984 pour le *Department of Energy* [Sebeok, 1984], a dû essayer de répondre à la question : comment concevoir un langage capable de porter à nos lointains descendants un signal de danger absolu à proximité des sites où sont enterrés aujourd'hui des déchets nucléaires à vie longue ? S'appuyant sur l'exemple de la survivance du sens des mythes grecs, portés tant bien que mal jusqu'à nous par une continuité culturelle, et sur le contre-exemple de l'enfouissement du secret des hiéroglyphes, son rapport conclut qu'il n'y a pas d'iconicité qui résiste aux siècles, et que les seules images interprétables le sont dans le cadre d'une tradition culturelle et mythologique¹¹.

Une illustration plus légère du même principe nous est donnée par une catégorie de messages satiriques consistant à transformer le sens d'une œuvre connue en en

¹¹Ce qui le conduit d'ailleurs à la conclusion effrayante que la seule solution pour transmettre le message du danger mortel que représentent nos poubelles à d'éventuelles sociétés futures moins civilisées consisterait à établir dès aujourd'hui une caste de « prêtres de l'atome » chargés de perpétuer un tabou de nature quasi-religieuse de génération en génération !

proposant un interprétant totalement différent de celui usuellement reçu. Ainsi, selon le « Guide bleu Languedoc-Roussillon »¹², quelques vieux Nîmois se plaisent à expliquer l'attitude de leur célèbre compatriote l'empereur Antonin (représenté en statue, non loin de la Maison Carrée, dans la position du salut romain), en disant qu'il « surveille s'il pleut » (« *Espincho sé plou* »). Autre exemple, la description donnée par Ambrose Bierce du Laocoon, célèbre statue antique représentant le prêtre troyen et ses fils en train de se faire étouffer par des serpents¹³ : « célèbre sculpture antique représentant un prêtre de ce nom et ses deux fils pris dans les anneaux de deux énormes serpents. L'adresse et la diligence avec lesquelles le vieil homme et les garçons soutiennent les serpents et les maintiennent à bonne hauteur ont été considérées à juste titre comme l'une des plus nobles illustrations artistiques de la supériorité de l'intelligence humaine sur l'inertie brute. »¹⁴

Plus généralement, toute production culturelle fait un usage plus ou moins affirmé de conventions et de symboles qui la rattachent à une tradition, à une histoire antérieure. Dans l'univers de la peinture ou du dessin, les cristallisations graphiques de certains de ces symboles constituent ce qu'on appelle les conventions iconographiques.

Pour comprendre un dessin de presse concernant des sujets politiques, par exemple, nous montrions ci-dessus que l'on doit reconnaître un certain nombre de symboles qui représentent conventionnellement des entités, processus ou caractères abstraits. Ainsi la colombe revient-elle souvent comme symbole de paix et d'espoir retrouvé, et le serpent comme symbole de mensonge et de duperie.

Il est intéressant de reconnaître dans l'histoire de l'établissement et de la consolidation de tels symboles visuels des phénomènes locaux, dont on peut marquer les jalons dans une histoire culturelle. Ainsi, pour reprendre les deux derniers exemples, la colombe et le serpent sont-ils tous les deux des symboles bibliques : on sait que la colombe, portant un rameau d'olivier à Noé, lui apporta la nouvelle de la fin de la colère divine (*Genèse* 8, 11), et que le serpent, en mentant à Ève sur les conséquences de son geste, l'incita à manger du fruit défendu (*Genèse* 3, 1-7)¹⁵. Ces symboles ne sont pas nécessairement universels : le serpent, dans la mythologie grecque, était le symbole d'Asclépios (Esculape), dieu de la médecine ; et même, plus généralement, « le serpent était un signe de la divinité » ([Lessing, 1964], p. 16). La cosmogonie aztèque avait fait du *Quetzalcoatl*, le serpent à plumes, le fondateur de l'univers.

Comme elles sont la marque de leur aire culturelle, les conventions iconographiques sont également la marque de leur époque. Ainsi Émile Mâle [1984] a-t-il découvert dans un catalogue d'images allégoriques de la fin du XVI^{ème} siècle la grande source d'inspiration des représentations d'entités abstraites de tout l'art religieux italien du XVII^{ème}. L'ouvrage, *Iconologia ovvero descrizione dell'Imagini universali cavate dall'antiquità et da altri luoghi*, par Cesare Ripa, publié pour la première fois en 1593, que l'auteur décrit lui-même sur la page de titre comme « nécessaire aux

¹²Paris : Hachette, 1988.

¹³Scène décrite par Virgile, *Énéide*, livre II.

¹⁴Ambrose Bierce, *The Devil's Dictionary*. New York (U.S.A.), Dell Publishing, 1991 (première édition 1911).

¹⁵Récit qui d'ailleurs a contribué à forger une image négative de la femme au moins autant que du serpent.

poètes, peintres, sculpteurs et autres, pour représenter les vertus, les vices, les émotions et les passions humaines » ([Białostocki, 1968]), a constitué un dictionnaire d'allégories qui s'est imposé jusqu'au XVIII^{ème} siècle en Italie et en France, puis qui a sombré dans l'oubli avec le XIX^{ème} siècle. Les iconographies évoluent, au fur et à mesure que les peintres inventent de nouvelles représentations.

On peut étendre ces réflexions sur les conventions iconographiques à d'autres modalités. Les stéréotypes existent dans les langues, où ils représentent la doxa d'une époque : la parole en est tissée, comme l'a illustré Flaubert dans son exercice de style célèbre *Bouvard et Pécuchet*. Les gestes qui accompagnent la conversation, et que nous qualifions de *spontanés* par opposition par exemple aux gestes plus rigoureusement codifiés de la langue des signes utilisée par les sourds, sont conventionnels et culturels à notre insu : le geste de refus consistant à balancer l'index comme un métronome est par exemple typiquement français. Ils évoluent aussi, comme le bras d'honneur qui a progressivement été remplacé par le majeur dressé depuis les années 60, et qui a quasiment disparu aujourd'hui de la pratique courante [Calbris & Montredon, 1986].

4.5 Conclusion

Nous ne découvrons bien entendu pas l'existence de ces conventions culturelles permettant la lecture et l'interprétation. Ce qui nous intéresse particulièrement ici est le fait que le « réseau herméneutique » qui nourrit et est nourri de sens par chaque texte individuel, est multimodal. De même que les tableaux d'art chrétien se nourrissent de textes bibliques, de même que les expressions faciales des personnages de bande dessinée (cf. à ce sujet [Eco, 1968], II.9, p. 120) comme de ceux des romans de Stendhal ([Rastier, 1968]) se nourrissent de sémiotiques « naturelles » de l'attitude humaine, de même nous rencontrerons à tout instant, dans notre étude des systèmes de signes iconiques dont l'usage est répandu de nos jours, des rappels et des conventions culturelles dont nous retrouverons l'origine dans des sémiotiques littéraires, mythiques, graphiques, gestuelles ... propre à la civilisation qui les a engendrés.

