

**Jean-Marie Klinkenberg**

## **La relation texte-image Essai de grammaire générale<sup>1</sup>**

### **0. Introduction**

#### **0.1. Un champ vierge**

Que l'image et le texte entretiennent une relation privilégiée, c'est l'évidence même : titres de tableaux, bande dessinée, publicité, journaux, livres pour enfants, pièces de monnaie, monuments, recettes de cuisine, plans de montage, cartes de géographie, articles scientifiques, monuments aux morts, rébus, écrans de téléphone portable, pages Web, sont là pour le prouver. Depuis l'invention de l'écriture, il est rare que l'image aille sans le texte et, aujourd'hui, de plus en plus rare que le texte aille sans l'image. Sans compter que l'écriture — base du texte — provient elle-même historiquement de l'image, et que l'on voit de plus en plus souvent des « mots dans la peinture », selon l'expression de Michel Butor.

Pourtant, il est étonnant de constater que ce problème sémiotique de la relation texte-image n'a jamais été pris à bras le corps. Du moins en tant que problème de portée générale.

Car bien sûr, les études de cas abondent. Il me semble avoir entendu, dans les congrès de sémiotique visuelle (et il doit en être de même dans les colloques d'histoire de l'art...), des dizaines, sinon des centaines, de communications de jeunes chercheurs portant sur « Le texte et l'image dans l'affiche républicaine espagnole » ou « Le texte et l'image dans un corpus de publicités automobiles en Italie ». Et bien sûr, on se souviendra que certains travaux pionniers en sémiotique portaient sur des énoncés où ce rapport texte-image joue à plein. Que l'on pense par exemple à l'étude fondatrice de Roland Barthes sur la « Rhétorique de l'image » (1964), où l'auteur examinait le rapport entre le texte d'une publicité, avec sa connotation

---

<sup>1</sup> Le présent article constitue le développement d'une conférence donnée au GEMCA le 28 janvier 2009 sous le titre *Texte + image : vers une grammaire générale*. Il ne comporte pas les illustrations alors montrées, qui sont ici décrites avec suffisamment de détails. Une version plus complète sera publiée dans le *Bulletin de la classe des lettres et des sciences morales et politiques* de l'Académie royale de Belgique. Le texte fait usage des rectifications de l'orthographe de 1990, proposées par tous les organismes francophones compétents, dont l'Académie française.

d'italianité (ici fournie par le nom des pâtes Panzani), et les motifs iconiques et plastiques du message, avec leurs significations parmi lesquelles on retrouvait également l'italianité.

Mais toutes ces contributions n'ont le plus souvent donné naissance qu'à des concepts *ad hoc*, dont la généralité et le caractère transposable n'ont pas été fréquemment interrogés, à des taxinomies empiriques, comme celle que proposait Abraham Moles (1978) ou encore à des notions générales mais restant vagues, comme celles d'ancrage et de relais, lancées par Barthes (1964) : désignées par des métaphores, elles n'ont jamais été vraiment définies.

Par ailleurs, un certain nombre de propositions qui ont été faites à l'époque de Barthes souffrent d'une perspective glossocentriste, qui caractérisait les débuts de la sémiotique visuelle, et à vrai dire toute la sémiotique du temps : en gros, c'est la thèse qui consiste à dire que le sens du visuel réside dans ce qu'on peut en dire verbalement, ce qui aboutit à subordonner tout le sémiotique au linguistique, voire à l'y résorber<sup>2</sup>. Dans cette dernière version, assurément maximaliste, de la description du rapport linguistique-sémiotique, il n'y a évidemment plus rien à dire sur ce qui n'est pas la langue. Si on s'en tient à la version *light* de cette description, on obtient la hiérarchie suivante : à la langue reviendrait la tâche d'exprimer les significations stables et univoques, à l'icône celle de dire les significations ouvertes, « flottantes », la première ayant pour vocation « d'ancrer » ces dernières<sup>3</sup>. Dans la typologie des rapports texte-image de Moles, le langagier est implicitement donné comme premier, l'image venant comme complément, supplément, contrepoint.

Les progrès opérés depuis lors par la pragmatique, la rhétorique et la sémiotique visuelle permettent de dépasser cette opposition caricaturale : on ne peut plus aujourd'hui sous-estimer la part de non-univocité dans la langue, et on ne peut non plus refuser que l'attribution de sens aux stimuli visuels obéisse à des règles plus strictes que ne l'a cru Barthes. Comme on va le voir, l'image « ancre » le texte, autant que le texte « ancre » l'image.

---

<sup>2</sup> Ce glossocentrisme a encore de belles heures devant lui. On connaît les thèses d'un Bruno Latour, pour qui la science entière se ramène à des énoncés linguistiques, toute déclaration sur le réel étant dès lors impossible. La science ainsi conçue, dans un parfait immanentisme, s'assimile finalement à un débat quasiment littéraire.

<sup>3</sup> Cette conception est évidemment cohérente avec une pensée où la langue est appelée à faire « régner l'ordre ». Citons icile verdict de Hénault (1983), pour qui la sémiologie barthésienne (et on peut en dire autant des suggestions de Louis Marin) « se contentait de nommer terme à terme les figures du contenu ; une telle lexicalisation transformait le texte visuel en texte linguistique et la sémiologie de l'image n'était qu'une variante de la sémiologie linguistique ». Cfr aussi Béguin, 2006 : 30.

## La relation texte-image

Mon objectif sera de constituer ici l'esquisse d'une grammaire de l'interaction texte-image, obéissant à l'exigence de généralité que je viens de pointer.

Une restriction toutefois : le mot « image », que je viens d'utiliser alors que je tends à le bannir de mes travaux, a, c'est connu, un sens très large, que l'on peut gloser comme suit : « *forme significative* perceptible dans l'instant minimum de vision » (Bertin 1967 : 142). Ici, il sera pris au sens restreint de « icône »<sup>4</sup>. Il y a évidemment d'autres « images » : par exemple les courbes, schémas et diagrammes (cfr Badir, 2005 et 2007, Desnoyers, 2005a et b). Dans ces autres énoncés visuels, se nouent des relations spécifiques entre écriture et graphique. Elles méritent assurément un examen particulier, que nous menons ailleurs (Klinkenberg, à paraître a et b), et nous ne ferons donc ici qu'effleurer le sujet. Par ailleurs, il n'est pas davantage question de prendre ici en charge le signe plastique<sup>5</sup> : cela m'eût obligé d'aborder les vastes et pertinentes questions de la typographie, de la mise en page, voire de la qualité texturale du support, au risque d'affaiblir le propos. Ces thèmes seront eux aussi traités ailleurs.

Deux autres notions devront également être présentées en introduction, car elles seront abondamment mobilisées dans ce qui suit : celle de pluricodie, et celle de tabularité.

### **0.2. La pluricodie**

LE rapport texte-image est un cas particulier de pluricodie.

On a parfois appelé « sémiotiques syncrétiques » des manifestations discursives associant des éléments provenant de plusieurs sémiotiques différentes. Cette appellation est doublement critiquable. D'abord parce que le mot « sémiotique » n'y est pas adéquat : ce ne sont pas les sémiotiques elles-mêmes qui sont syncrétiques, mais les énoncés observés. Il vaudrait donc mieux parler de « discours syncrétique ». Ensuite, l'adjectif « syncrétique » renvoie généralement à une combinaison d'éléments peu cohérente. Comme je fais précisément l'hypothèse que la combinaison de l'icône et du texte obéit à des règles strictes, il est particulièrement mal venu. Je préférerai donc les termes « discours pluricode » à l'expression « discours syncrétique ».

---

<sup>4</sup> Un autre concept qui a fait l'objet de controverses sur lesquelles je ne puis m'étendre ici. Sur les raisons de conserver ce concept dans l'appareillage conceptuel de la sémiotique, je renvoie à Groupe  $\mu$ , 1992, et Klinkenberg, 2002b.

<sup>5</sup> Sur la distinction iconique vs plastique, cfr Groupe  $\mu$ , 1992.

Par discours pluricode, j'entendrai donc toute famille d'énoncés considérée comme sociologiquement homogène par une culture donnée, alors même que l'on peut régulièrement isoler dans ses énoncés plusieurs sous-énoncés relevant chacun de codes réputés distincts par le savoir développé dans ladite culture (cfr Klinkenberg, 2000a).

Le plus bel exemple de discours pluricode est peut-être fourni par le cinéma. Cet art a indéniablement son identité et sa stabilité à nos yeux. Et pourtant, il mobilise simultanément le récit, l'image (plastique autant qu'iconique), la kinésique et la proxémique, la langue, la musique, l'iconisme sonore, l'écriture... On trouvera aisément d'autres exemples : le théâtre, qui comme le cinéma manipule le verbe et le corps et plusieurs structurations de l'espace ; la conversation, avec les gestes, les attitudes et le langage; et encore : l'opéra, la publicité, l'architecture, les réalisations informatiques multimédias, les pratiques artistiques où peinture et vidéo se combinent...

Dans ce qui suit, j'utiliserai donc l'expression *énoncé scripto-iconique* pour désigner un énoncé pluricode considéré comme sociologiquement stable, et mobilisant à la fois l'icône et l'écriture : c'est le cas de l'affiche publicitaire, de la page ou de la case de B.D., de la page de revue ou de journal, de la page de l'album pour enfants, de l'image pieuse, des encyclopédies, etc.

Par convention, je désignerai ci-après les sous-énoncés écrits par la lettre L (pour linguistique) et les sous-énoncés iconiques par la lettre V (pour visuel).

### **0.3. Tabularité de l'espace scripto-iconique**

**S**i l'on parle ici d'une interaction du texte et de l'image, c'est par définition d'un texte vu qu'il s'agit, et vu parce qu'il est écrit.

Je convoque ici l'écriture pour en souligner un trait essentiel. En dehors de Harris, 1995, Christin, Cárdenas, 2001, et Perri 2001, les théoriciens de l'écriture sont peu nombreux à avoir pointé cette caractéristique capitale de l'écriture : que cette dernière fait passer la parole du canal oral sur le canal visuel. Ainsi, en passant par la modalité sensorielle qu'est la vision — au moment où ils se coulent dans l'écriture—, les énoncés linguistiques conservent-ils leur caractère linéaire, tout en acquérant un caractère spatial

## La relation texte-image

qu'ils reçoivent des spécificités de ce canal<sup>6</sup>. Grâce à la spatialisation, l'énoncé écrit cesse donc de se déployer dans un espace à une dimension (induisant des relations exclusivement linéaires) mais investit un espace à deux dimensions au moins. Cet espace à deux ou à trois dimensions permet une aperception simultanée. Mais le champ de cette aperception est toutefois orienté (et animé) par des relations linéaires.

La particularité de l'espace écrit réside en définitive dans cette dialectique : il est le champ où se déploient à la fois des relations linéaires et des relations multidimensionnelles. L'énoncé langagier devient image, au sens de Bertin. Tout ceci a en effet pour conséquence que l'écriture assume fatalement des fonctions non-linguistiques. D'où sa relative autonomie par rapport à la langue, et son ambivalence : elle est à la fois une sémiotique linguistique et une sémiotique de l'espace (au même titre que celle qui est à l'œuvre dans le dessin, le plan, l'organigramme, le diagramme, la carte, la maquette, la ville, le jardin ou l'édifice).

---

<sup>6</sup> Les sémiotiques qui transitent par l'un et l'autre de ces canaux sont en effet profondément affectées par les propriétés de ceux-ci. Ceci— soit dit en passant — justifie une prise en considération des canaux, qui était jusqu'à nos jours largement refusée par la sémiotique (voir Klinkenberg, 2000a). Il n'est évidemment pas question de ramener toute la question du sens à la nature du canal (comme le veut une position polémique que résume la formule bien connue de Mc Luhan « Le médium, c'est le message »). Mais de prendre au sérieux l'idée que les aspects matériels de la communication font de plein droit partie du travail de signification et de référence. Je précise que par canal, j'entends d'une première part l'ensemble des stimuli des signes — la notion de stimulus sera expliquée plus loin —, qui dépendent du support matériel permettant leur transmission, et de seconde et troisième parts les caractéristiques des appareils émetteurs et récepteurs, qui ont chacun leurs potentialités propres. C'est que la configuration des énoncés dépend de ces instances de transmission. Les appareils visés ne sont en effet sensibles qu'à une gamme spécifique de phénomènes physiques. Ainsi, le canal auditif ne permet pas de traiter à la fois autant d'informations que le canal visuel, beaucoup plus puissant. Ce dernier permet en effet de faire transiter  $10^7$  fois plus d'informations que le premier, dans le même laps de temps (cfr Groupe  $\mu$ , 1992). Cette différence de puissance a des répercussions non négligeables sur les modes d'organisation des unités sémiotiques dans un syntagme. On peut aisément constater que les sémiotiques s'appuyant principalement sur l'ouïe privilégient les chronosyntaxes, ou syntaxes linéaires (où les informations se présentent pour être traitées les unes après les autres), alors que celles qui exploitent la vision font un usage de toposyntaxes, ou syntaxes tabulaires (un terme qui va être glosé ci-après), où l'on traite simultanément un certain nombre d'informations, cette simultanéité étant autorisée par la puissance du canal. Lorsqu'une même sémiotique mobilise des signes fondés sur des stimulus différents — autrement dit lorsqu'elle exploite des canaux différents —, ceci entraîne une conséquence importante : on ne peut avoir une équivalence parfaite entre la variante du code transitant par le canal *a* et celle qui transite par le canal *b*, puisqu'un canal particulier fait peser des contraintes spécifiques sur l'organisation de ces signes. On doit donc s'attendre à ce qu'en passant d'un canal à un autre, les messages performés à partir d'un code donné perdent une partie de leurs caractéristiques mais gagnent en échange d'autres caractéristiques. Règle que vérifie l'écriture.

À la fois linéaire et spatiale: cette double caractéristique produit une tabularité, notion que le Groupe  $\mu$  a introduite dans *Rhétorique de la poésie* (1977), et qui caractérise la lecture de certains énoncés. Dans la lecture tabulaire, un élément de l'énoncé prend son sens grâce au rapport qu'il entretient avec d'autres éléments, sur au moins deux axes distincts.

## 1. L'homogénéité de l'énoncé : l'indexicalité

**D**ire d'un énoncé qu'il est pluricodique est une affirmation lourde de présupposés et de conséquences : à quelles conditions peut-on affirmer, devant une manifestation sémiotique complexe donnée, qu'elle constitue un énoncé et un seul, et non deux énoncés distincts, sans relations entre eux ? Étudier ces relations, sur le plan syntaxique et sémantique, est la tâche que je me suis assignée dans ce qui suit, mais sa réalisation est, on le voit, suspendue à une question préjudicielle : quand peut-on dire qu'il y a un énoncé pluricodique et un seul ?<sup>7</sup>

Ce qui est posé est donc la double question de l'homogénéité de l'énoncé scripto-iconique et de son unicité : pourquoi décide-t-on de considérer un ensemble de faits comme cohérent et de le distinguer de ce qui l'entoure ?

Deux facteurs interviennent dans cette prise de décision : des facteurs perceptifs et des facteurs que nous nommerons indexicaux.

Partons d'un exemple concret : une œuvre picturale quelconque dans un musée, et dans laquelle il n'y a pas de texte. On y remarquera deux séries de faits, les uns perceptuels, les autres sémiotiques, qui ont tous pour effet d'isoler une portion de l'espace en l'opposant à une autre. Et c'est pourquoi nous donnons un statut à une de ces portions, et que nous l'appelons « tableau », le reste étant un « fond » à quoi nous ne donnons généralement pas d'importance particulière.

Sur le premier plan, perceptuel, un premier examen permet de noter une opposition forte de caractéristiques formelles entre deux séries de propriétés : d'une part un espace encadrant très homogène (même dominante chromatique claire, pas de variation de forme) : l'espace du mur ; de l'autre, un espace hétérogène (modèles, variation chromatique) : celui du tableau. Ces deux espaces sont nettement séparés par une limite rectiligne. On peut donc parier sur l'existence de deux champs distincts, ayant chacun

---

<sup>7</sup> En effet, par définition, l'énoncé global verbo-iconique est pluricodique, ce qui signifie, en d'autres mots, qu'il laisse subsister une hétérogénéité.

## La relation texte-image

leur personnalité, et donc sur leur homogénéité sémantique : ici l'énoncé d'une part — le tableau — et de l'autre le fond sur lequel on le distingue.

Mais ce facteur perceptuel n'est peut-être pas, en droit, suffisant pour nous amener à dire qu'il y a là un couple fond + tableau, et non pas un seul énoncé global. En effet, qu'est-ce qui empêcherait de dire que dire que l'ensemble fond + tableau est un énoncé pluricodique ?

C'est que le facteur perceptif, condition nécessaire, ne constitue pas une condition suffisante. Il se voit obligatoirement relayé par un signe particulier, faisant partie de la famille des index : le cadre, ici, par hypothèse, un cadre doré et mouluré (cfr Groupe  $\mu$ , 1992, chapitre XI).

Cette notion d'index mérite elle aussi un temps d'arrêt.

Pour rappel, l'index — qu'on se gardera de confondre avec l'indice — est un dispositif sémiotique ayant une double fonction : (i) focaliser l'attention sur une portion déterminée d'espace (et spécialement y ségréguer un objet), et (ii) donner un certain statut à cette portion d'espace. Exemple canonique : le doigt pointé vers un objet (mais on pourra aussi penser à la flèche, à l'étiquette, etc).

Il faut noter que ce type de signe ne fonctionne qu'en présence de l'objet désigné, de sorte que sa définition mobilise l'idée de contigüité<sup>8</sup>. En définitive, l'index mobilise toujours trois éléments : (i) la manifestation indexicale proprement dite (/doigt pointé/, /étiquette/, /badge/, /couverture/...), que nous appellerons *indexant*, (ii) la portion d'espace désignée, ou *indexé* (objet, porte de sortie, personne, livre) et (iii) la relation précise que le premier institue avec le second ou *indexation* (« désigner, « donner un statut », etc.). L'index présuppose donc une sémiotique de l'espace : pour que l'indexant puisse désigner une portion d'espace, il faut au préalable que cet espace indexé soit perçu comme une unité homogène et distincte de son entourage : par exemple un bâtiment, distinct de son environnement, un objet, distinct de son contexte. Quant à la relation, elle a, comme on l'a dit, deux fonctions possibles, la seconde étant facultative : la première consiste à focaliser l'attention du partenaire de l'acte indexical sur une portion déterminée d'espace, qui devient donc l'indexé, la seconde à donner un statut sémiotique à cet indexé (« œuvre d'art », « fonction sociale », etc.). Si la nature des objets appelés à devenir des indexants peut être très variable (comme on va le voir, ils peuvent très bien être de nature

---

<sup>8</sup> C'est ce fait qui explique que les index soient fréquemment confondus avec les indices (cfr Klinkenberg, 2000a). Mais la contigüité spatiale qui joue ici n'est pas la même que dans l'indice, où elle est d'ordre causal. La confusion est aussi due à la quasi-homonymie des mots français indice et index (dans maintes autres langues, l'homonymie est même totale).

linguistique : /ce type-là/), ils exploitent donc toujours l'espace : comme l'écriture, les index relèvent donc d'une sémiotique de l'espace, et l'on ne s'étonnera donc pas que l'écriture joue fréquemment un rôle indexical<sup>9</sup>.

Ce sont de telles indications indexicales qui fournissent l'unité du dispositif global. Dans notre exemple, isoler le portrait de son fond et en faire un énoncé homogène est d'autant plus évident que ce portrait est séparé du reste par le cadre. Comme tous les index de ce type, le cadre a une double fonction : produire la clôture, et affirmer l'homogénéité de l'espace ainsi clôturé.

C'est donc à l'indexation que revient la fonction d'établir l'homogénéité des énoncés scripto-iconiques : la planche de BD, la case de BD, la page de l'album, l'affiche, etc. Nous savons ce que c'est qu'une page, à quoi elle sert, comment elle se manipule, et sommes dressés à considérer tout ce qui figure dans l'espace délimité par sa bordure comme constituant une unité de forme et de sens. Nous savons aussi ce qu'est un livre, et y postulons de la même manière que tout ce qui est contenu dans les limites matérielles de son volume, ou dans celles de sa subdivision qu'est la page, constitue une unité de forme et de sens. C'est l'indexation qui lui confère cette qualité.

Ce dispositif est culturel et arbitraire : le type de renvoi qu'il suscite est purement conventionnel. L'index a donc une consistance variable, sa fonction d'indexation peut évoluer dans le temps et se diversifier dans l'espace, il peut avoir un effet plus ou moins puissant, être plus ou moins vite identifiable.

Il faut noter que plusieurs index peuvent également intervenir simultanément. Et dans ce cas, ils peuvent avoir leur syntaxe et leur sémantique propres, sur lesquels je ne puis m'étendre ici longuement.

Leur syntaxe propre : les indexants peuvent par exemple être juxtaposés ou emboîtés. Dans notre exemple du tableau, par hypothèse, les baguettes concentriques se répètent. Si matériellement il n'y a qu'un cadre, sémiotiquement, c'est bien d'un emboîtement de cadres qu'il s'agit ; ce qui correspond sémantiquement à une redondance.

Dans un autre exemple — *La clef des songes*, tableau bien connu de Magritte —, on observe un autre type d'emboîtement. D'un côté, on a un cadre global, qui nous fait percevoir l'ensemble des six cases du tableau comme constituant un énoncé homogène (hypothèse renforcée par

---

<sup>9</sup> Cfr Klinkenberg, 2005. Les indexants peuvent également avoir une substance temporelle : « le roi de France *actuel* ».



## La relation texte-image

l'intervention d'un autre indexant, linguistique celui-là : le titre de l'œuvre ; cfr Bosredon, 1997) ; de l'autre côté, chacune des cases a son propre cadre, qui nous la fait voir comme un sous-énoncé, homogène lui aussi. Le dispositif global nous incitera donc, à six reprises, à nous interroger sur une interaction verbo-iconique particulière : celle de l'/œuf/ visuel (a) et de /l'acacia/ langagier (b), du /soulier de femme/ (a) et de /la lune/(b), du /chapeau melon/ (a) et de /la neige/(b), et ainsi de suite. Mais la répétition de ce type d'interaction, dans un énoncé global, institue ces interactions en règle générale. Dans la bande dessinée, on aura un schéma analogue : l'unité case, avec ses interactions verbo-iconiques propres, l'unité planche, avec ses intégrations de cases, et l'unité album, intégrant les précédentes<sup>10</sup>.

Toute cette longue mise en place — pluricodie, tabularité, indexicalité — était nécessaire pour aborder la question des règles de l'interaction verbo-iconiques.

Comme dans toutes les grammaires, ces règles sont de cinq types, que l'on étudie généralement dans cinq chapitres distincts : les règles qui régissent l'engendrement des constituants de base (en langue, ce seront les sons et les phonèmes), celles qui président à la constitution des unités complexes du plan de l'expression (morphologie), celles qui décrivent les relations entre ces unités (syntaxe), celles du sens (c'est la sémantique, phrastique autant que lexicale) ainsi que celles qui régissent l'usage des énoncés en situation : c'est la pragmatique. Dans la pratique, il est évidemment toujours difficile de maintenir séparés ces différents ordres, et ce sera particulièrement le cas ici. Dans ce qui suit, nous présenterons rapidement les relations morphologiques, et privilégierons les interactions syntaxiques et sémantiques, en faisant dans les deux cas ressortir systématiquement les mécanismes pragmatiques inévitablement à l'œuvre<sup>11</sup>. Nous n'avons par contre pas à nous pencher sur les composants de base : ce sont en effet ceux qu'étudient la sémiotique de l'écriture d'une part et la sémiotique visuelle d'autre part, et ils ne jouent aucun rôle spécifique dans le problème qui nous intéresse ici.

---

<sup>10</sup> Dans les index multiples, chaque index peut aussi avoir sa propre sémantique. Exemple de l'œuvre dans un musée : le /musée/ lui-même est un indexant, mais aussi la /salle/, puis le /cadre/. Chacun de ces indexants a un sémantisme commun (« statut d'œuvre »), mais a aussi un sémantisme propre : (/musée/ renvoie à « statut potentiel d'œuvre » ; /salle/ renvoie à « œuvre appartenant à la série x » ; /cadre/ renvoie à « statut avéré d'œuvre » + « délimitation de l'espace de l'œuvre »...

<sup>11</sup> « La coprésence et l'articulation du verbal et du visuel dans les formes contemporaines d'expression et de transmission de l'expérience contraignent l'analyse des effets de sens à développer, à côté de la pragmatique du verbe, une pragmatique de l'image » (Desgoutte, 2003 : 21).

## 2. Relations morphologiques

Les relations morphologiques entre texte et image mobilisent les fonctions iconiques de l'écriture. Dans ces fonctions, le tracé du signe graphique ou d'un bloc de signes graphiques renvoie iconiquement à un objet ou à un autre signe. Il y a là une interpénétration des signes de deux sémiotiques, que rend possible leur caractère spatial partagé.

Pour décrire correctement ces fonctions iconiques, il faut distinguer, sur le plan de l'expression, les notions de signifiant et de stimulus, cette dernière étant sans doute moins familière au lecteur que la première. Nous entendons par *stimulus* la manifestation concrète et contingente du signe ; autrement dit, ce qui dans la communication le rend transmissible par un canal particulier, en direction d'une sensorialité particulière ; le signifiant, quant à lui, est un modèle. Le rapport entre stimulus et signifiant est donc, généralisé à tout signe, celui qui unit son et phonème dans le signe linguistique. J'appelle objets graphiques les stimuli des systèmes graphiques (Klinkenberg, 2000) : grosso modo, ce sont les constituants du répertoire d'une écriture, mais tels qu'on peut les repérer empiriquement, avant toute assignation d'une fonction particulière à ce constituant dans une écriture donnée (traits constitutifs des lettres, ou des signes diacritiques, espaces, etc.).

L'interpénétration entre le signe scriptural et le signe iconique se laisse donc décrire de la manière suivante : un même stimulus (empirique) renvoie simultanément à deux signifiants, l'un de nature iconique et l'autre de nature scripturale. En termes de rhétorique visuelle, ces interpénétrations produisent des figures du type « chafetière » (cfr Groupe  $\mu$ , 1992), avec cette différence que les deux ensembles en intersection ne sont pas tous deux constitués de stimuli correspondant à des déterminants iconiques : c'est le cas de l'un, l'autre étant constitués d'objets graphiques.

Ce type d'énoncé permet de voir spectaculairement à l'œuvre le caractère ambivalent de l'espace scripto-iconique, organisé par des relations linéaires aussi bien que tabulaires, par des chronosyntaxes autant que par des iconosyntaxes (Edeline 1974, 2003).

Le meilleur exemple de la manifestation de cette interpénétration est sans doute à chercher du côté de ce qu'il est convenu d'appeler les « calligrammes ». Ce terme a certes l'usage pour lui, mais on devrait, en rigueur de termes, lui préférer celui d'iconogramme. Par ailleurs, il tend malheureusement à ne désigner que des énoncés caractérisés par ce qu'on appellera ci-dessous dominante scripturale, à l'exclusion des énoncés à

## La relation texte-image

dominante iconique, où l'on observe pourtant la même interpénétration entre signe scriptural et signe iconique<sup>12</sup>.

Il s'agit d'un phénomène que l'on trouve dans toutes les cultures et dont la tradition remonte à la plus haute antiquité. On peut ainsi la voir fonctionner en ancien égyptien, dans ces cas où l'icône et l'idéogramme — deux entités distinctes en droit — se superposent presque exactement. C'est ce que montre telle gravure pariétale, qui met en scène le rite de l'union des deux pays : un même végétal peut y être lu tantôt comme icône des fleurs caractérisant la Haute- et la Basse-Égypte, tantôt (ou simultanément) comme hiéroglyphe à valeur phonétique (*mhw* et *swaw*, « Haute-Égypte » et « Basse-Égypte»). De nos jours, cette fonction se voit aussi fréquemment utilisée en publicité.

Notons que, comme on vient de le suggérer, l'iconogramme est possible avec divers types d'écriture : dans les écritures phonographiques comme dans les idéographiques.

Une étude exhaustive des règles d'interpénétration du signe graphique et du signe iconique, qu'une sémiotique de l'écriture devrait énumérer et à laquelle nous ne nous astreindrons pas ici, montrerait que celles-ci sont nombreuses.

On verrait ainsi que les énoncés iconographiques peuvent être à dominante iconique ou à dominante scripturale. Les facteurs assurant la prévalence de l'un ou de l'autre type sont de nature quantitative autant que qualitative. Du premier côté, on soulignera le rôle des résidus. Si l'énoncé comporte un résidu iconique ne pouvant pas faire l'objet d'une interprétation scripturale, alors la dominante est iconique. C'est le cas dans l'œuvre de Magritte intitulée *L'art de la conversation*, qui montre un paysage de ruines dans lequel certaines pierres ont la forme de lettres, lettres qui, mises ensemble, permettant de lire le mot « rêve ». Un certain nombre de déterminants du type iconique « ruines » n'y constituent pas des lettres, de sorte que le type « ruines » est dominant. On parlera donc d'iconogramme à dominante iconique

Le contraire s'observe dans le poème de Guillaume Apollinaire *La colombe poignardée et le jet d'eau*. Il s'agit d'un iconogramme, permettant de voir iconiquement l'oiseau, une pièce d'eau, un jet d'eau. Mais le titre échappe à une lecture iconique : il y a donc ici dominante scripturale.

---

<sup>12</sup> On peut voir dans cette préférence un trait de glossocentrisme qui vient s'ajouter à ceux que nous avons déjà énumérés en introduction.

Les cas d'iconogrammes sans résidus, c'est-à-dire où la totalité des constituants sont simultanément lisibles sur les deux modes, sont relativement moins fréquents : c'est vers quoi tendaient certaines parties de l'énoncé égyptien commenté ci-dessus.

Sur le plan qualitatif, notons que l'interpénétration iconographique peut se produire à tous les niveaux de complexité de l'énoncé. Si l'on prend les niveaux d'articulation de la partie scripturale de l'énoncé pour échelle, l'icône peut être produite au haut comme au bas de celle-ci. Dans le premier cas, l'iconogramme est constitué par la totalité (ou la quasi-totalité) de l'énoncé scriptural<sup>13</sup>. Dans le second cas, il l'est par les constituants scripturaux minimaux de cet énoncé : la lettre (comme dans les nombreux cas d'abécédaires où les lettres sont également des icônes d'êtres humains), voire un trait de la lettre, comme dans l'exemple où le point sur l'i du mot « roi » serait manifesté par une couronne<sup>14</sup>. Si l'on prend les niveaux d'articulation de la partie iconique pour échelle, la reconnaissance globale des formes du stimulus peuvent être assurées par l'ensemble des masses graphiques, comme dans le cas de la plupart des « calligrammes ». A un niveau inférieur, les objets graphiques n'icônisent que quelques déterminants du signe iconique. En descendant encore l'échelle, les objets graphiques en arrivent même à n'icôniser qu'un simple formème<sup>15</sup> du signe iconique, comme l'orientation ou des variations d'orientation dans un mouvement : cas de lignes de caractères icônisant le mouvement de la chute que subit un personnage par exemple.

### **3. Relations syntaxiques**

Les instruments essentiels sont ici les déictiques et les connecteurs, auxquels viendront s'ajouter certains autres mécanismes non langagiers. Tous ces instruments ont la particularité d'être des indexants.

#### **3.1. Déictiques**

---

<sup>13</sup> En termes plus rigoureux, on dira que le stimulus de l'icône est constitué de faits spatiaux qui constituent par ailleurs les stimuli de signes glossiques phonémographiques, ou signes renvoyant à la transcription de phonèmes ou groupes de phonèmes.

<sup>14</sup> Toujours en termes plus rigoureux, on dira que le stimulus de l'icône est ici constitué de faits spatiaux qui constituent par ailleurs une partie du stimulus d'un signe phonémographique.

<sup>15</sup> Le formème est le déterminant minimal du paramètre « forme » du signe visuel. Par exemple, position, dimension, orientation, sont des formèmes (cfr Groupe  $\mu$ , 1992).

## La relation texte-image

**T**raditionnellement, on appelle déictiques des signes linguistiques dont le référent n'est identifiable qu'en tenant compte des conditions particulières de la situation de communication : le lieu, le moment, les acteurs. Les déictiques sont donc des indexants linguistiques. En effet, ils ne fonctionnent que par contiguïté, et induisent une suggestion de co-présence physique, que l'espace tabulaire de l'image pourra donc exploiter.

Nous sommes ici conduits à distinguer deux sortes de déictiques : les embrayeurs et les démonstratifs.

### 3.1.1. Démonstratifs et embrayeurs

**E**n français, les démonstratifs sont des pronoms et des adjectifs : cela, cette, celui, voici, etc. Ces termes désignent indexicalement des référents à partir de la position occupée par le locuteur, en tant que celle-ci est identifiable. Prenons le célèbre exemple de *La trahison des images*, de Magritte, où L énonce /Ceci n'est pas une pipe/ En vertu du principe d'homogénéité-clôture, l'objet de la monstration est identifié à partir de la position de l'énoncé. Opérant à partir de l'espace de l'énoncé scripto-ictonique, la monstration est par hypothèse tout entière contenue dans cet énoncé (et n'est pas référée au locuteur du monde). Le spectateur est donc fondé à chercher le lien sémantique entre cet énoncé linguistique et le seul énoncé iconique V qu'il ait à sa disposition dans cet espace clos. Il aura donc à gérer l'opposition apparente entre un énoncé linguistique qui affirme « Ceci n'est pas une pipe » et un énoncé iconique qui semble dire « Ceci est une pipe ». Mais le résultat de cette gestion est une autre histoire...<sup>16</sup>

Le référent des embrayeurs varie avec trois séries de variables : 1) le couple énonciateur/énonciataire, 2) le moment de l'énonciation, 3) le lieu de l'énonciation. En 1), nous aurons les pronoms personnels des deux premières personnes (« je », « toi », « nous », etc.) et les adjectifs et pronoms possessifs relatifs aux mêmes personnes (« mon », « votre », « mien », etc.) ; les formes verbales à ces deux personnes jouent donc aussi un rôle d'embrayage<sup>17</sup>. En 2), les adverbes et locutions adverbiales désignant le temps par rapport à celui de l'énonciateur — « maintenant », « tout à l'heure », « hier », « la semaine prochaine », « ce mois-ci » — ainsi que certains adjectifs : « actuel ». En 3), les adverbes et locutions adverbiales référant à un lieu, situé par rapport à celui de l'énonciateur : « ici », « là », etc.

---

<sup>16</sup> Racontée dans Groupe  $\mu$ , 1999.

<sup>17</sup> Même lorsqu'elles n'entraînent pas l'apparition d'un pronom : ce qui en français est le cas avec les impératifs.

Les interactions scripto-iconiques reposant sur de tels embrayages sont nombreuses : dans la bande dessinée, l'homogénéité thématique du binôme scripto-iconique est assurée par eux.

Les termes d'adresse sont étroitement parents des embrayeurs. Prenons l'exemple d'une publicité pour une plaque de cuisine vitrocéramique. Un personnage féminin s'y adresse à la plaque, luisante comme un miroir : «Rosières, Rosières, dis-moi ce que nous allons lui mijoter ce soir ». Le terme L — /Rosières/ — y est équivalent à un /tu/ : il donne son statut à l'objet, qui est en l'occurrence l'énonciataire interne à l'énoncé V : la plaque-miroir est une Rosières. L'adresse relaye d'ailleurs ici des embrayeurs personnels.

La mobilisation des déictiques pose deux problèmes, dont la solution est d'ordre pragmatique. La première difficulté consiste à identifier correctement l'indexé du déictique à l'intérieur de l'énoncé iconique ; la seconde est de décider s'il faut chercher cet indexé à l'intérieur de l'énoncé iconique ou s'il faut au contraire sortir de ce dernier pour obtenir une interprétation satisfaisante de la relation indexicale. Abordons ces deux problèmes.

### *3.1.2. Aspects pragmatiques : (1) quelle référence interne pour la deixis ?*

**D**ans *La trahison des images*, Magritte ne nous laissait pas le choix : puisque son énoncé ne comportait que la seule unité iconique /pipe/, c'est elle qu'il fallait identifier comme l'indexé de l'indexant qu'est le /ceci/. Mais les choses sont rarement aussi simples. Dans la mesure où plusieurs unités iconiques V peuvent être présentes dans l'énoncé global, cette identification du référent du déictique L n'est pas donnée d'emblée : plusieurs hypothèses entrent en concurrence. L'identification fait alors l'objet d'un calcul, mobilisant le principe de pertinence : on privilégiera l'interprétation laissant le moins de restes (gain de sens maximal), pour le coût cognitif le plus bas (économie de moyens maximale).

Prenons l'exemple d'une publicité pour la sécurité au travail, qui montre en V une main dressée verticalement, et où manque le majeur, remplacé par une végétal ; ce message énonce en L /ça ne repousse pas/ (figure 13).

En droit, le /ça/ pourrait se référer à plusieurs portions iconiques V de l'énoncé : un des doigts présents, le doigt manquant, la racine du végétal, son feuillage, etc. Devant cette pluralité d'hypothèses, le spectateur doit donc se livrer à un calcul sémantique, guidé par la règle de pertinence. Ce

## La relation texte-image

calcul amène ici à choisir le « doigt coupé » comme objet le plus pertinent pour le déictique à l'exclusion du végétal et des doigts intacts : en effet un végétal repousse, un doigt coupé non ; et un doigt intact ne justifierait pas le préfixe re-.

Par ailleurs, la recherche du référent de la deixis implique parfois des calculs inférentiels sophistiqués. Ceux-ci permettent de passer d'un type iconique manifesté V, mais peu pertinent par rapport au déictique, à un autre type non manifesté mais impliqué par le premier. Exemple : une campagne publicitaire contre l'alcool au volant.

Soit L /celui qui n'a pas bu, c'est celui qui conduit/, et V une voiture, identifiable comme un corbillard. L'énoncé V permet une inférence, ici un présupposé de type  $\Pi$ <sup>18</sup> : qui dit « voiture » dit « moteur », « carburant », mais aussi « conducteur » et éventuellement « passager ». L'énoncé L, avec ses verbes « conduire » et « boire », se référera plus économiquement à ces acteurs humains non manifestés qu'au type « voiture », pourtant manifesté. Et on en déduira la proposition « celui qui a bu, c'est celui qui ne conduit pas » (et c'est donc le passager, occupant du cercueil).

### 3.1.3. Aspects pragmatiques : (2) référence interne ou externe de la deixis ?

Jusqu'à présent, nous n'avons envisagé que des relations internes à l'énoncé scripto-iconique. Mais les déictiques, et spécialement les embrayeurs, se caractérisent par leur ambivalence : en droit, ils peuvent se référer tantôt à un objet interne à l'énoncé, tantôt à un objet extérieur à cet énoncé. Par exemple, le couple énonciateur/énonciataire peut être celui que forment des partenaires manifestés iconiquement, mais ce peut aussi être le couple formé par l'énonciateur et l'énonciataire de l'énoncé global : le producteur et le récepteur de cet énoncé. Ici encore, il y a donc plusieurs référents possibles au déictique.

Pour quelles raisons opte-t-on pour l'une ou l'autre des solutions possibles ? De nouveau, on se livre à un calcul de pertinence. Or il arrive qu'on ne puisse déboucher sur une lecture économique en restant à l'intérieur de l'énoncé, parce qu'aucun élément iconique V ne peut entrer de manière pertinente dans la relation déictique. Il faut à ce moment postuler une deixis reliant l'énoncé L au spectateur extérieur (énonciataire) de l'énoncé global.

---

<sup>18</sup> Sur ce concept, voir 4.1.3.

Si un énoncé L comporte plusieurs déictiques, le calcul peut dès lors donner des résultats différents pour chacun d'eux. Prenons l'exemple de l'affiche célèbre /I Want you for US Army/.

« I » (L) y est l'énonciateur de l'énoncé linguistique, et cet énonciateur est présent iconiquement dans V : la deixis reliant « I » et « Uncle Sam » (V) est donc mise en scène. Mais aucun énonciataire potentiel de « You » (L) n'est manifesté iconiquement<sup>19</sup>. On est donc fondé à estimer que la deixis mobilise un énonciataire extérieur à l'énoncé.

Nous parlerons donc dorénavant d'*énonciateurs* ou d'*énonciataires externes*, pour désigner les énonciateurs et énonciataires de l'énoncé scripto-iconique global, et d'*énonciateurs* ou d'*énonciataires internes*, pour désigner les énonciateurs et énonciataires mis en scène iconiquement (V) dans cet énoncé global.

La deixis peut fonctionner dans les deux sens : relier l'énonciateur interne à l'énonciataire externe (c'est le cas avec le « you » de « I want you ») ou un énonciateur externe à un énonciataire interne. Autrement dit, deux mouvements sont possibles, l'un centrifuge, l'autre centripète.

Prenons l'exemple des deux images pieuses où l'on observe la présence, sur le mode iconique (V), du même personnage — le Christ —, et les deux textes L font usage du même embrayeur personnel « vous ». Si l'on observe que ce « vous » est mobilisé dans deux énoncés différents comme /Venez à moi, et je vous soulagerai/ et /par vos souffrances, Seigneur, délivrez les âmes du purgatoire /, on se persuade aisément que l'indexé du /vous/ n'est pas le même dans les deux cas. Dans le premier cas, le fidèle est l'énonciataire externe, le Christ (représenté iconiquement) étant l'énonciateur : deixis centrifuge. Dans l'autre, l'énonciataire indexé est le Christ, toujours représenté iconiquement, et l'énonciateur est le fidèle : deixis centripète.

La deixis peut également relier l'énonciateur externe à l'énonciataire externe. C'est le cas dans une publicité pour un abonnement au mensuel *Fluide glacial* disant en L /Abonnez-vous ou nous tuons ce personnage/ et montrant en V /un poing tendu vers un personnage, poing muni d'un pistolet/. Parmi les nombreux déictiques qu'on peut y repérer, et dont certains sont internes à l'énoncé (/Ce personnage.../), on distingue un couple constitué d'un énonciateur externe (la rédaction ou les services financiers du

---

<sup>19</sup> À moins d'opter pour l'hypothèse couteuse selon laquelle le protagoniste se parlerait à lui-même.



## La relation texte-image

magazine) et d'un énonciataire externe (l'abonné potentiel) désigné par /vous/.

Concluons : le déictique peut mettre en rapport soit

- 1.— deux éléments de l'énoncé iconique V : on parlera de *deixis mise en scène* dans le message iconique
2. — un élément de l'énoncé iconique V et un partenaire de l'énonciation de l'énoncé global (son énonciataire ou son énonciateur). Dans cette hypothèse, on peut avoir deux cas distincts : (2a) *un énonciateur interne et un énonciataire externe* (deixis centrifuge) ou (2b) *un énonciateur externe et un énonciataire interne* (deixis centripète).
3. — deux éléments extérieurs à l'énoncé iconique V : *énonciateur et énonciataires externes*.

### 3.1.4. Aspects pragmatiques : (3) facteurs d'identification de la référence de la deixis

**P**our en terminer avec ces opérations pragmatiques, il convient de synthétiser les facteurs qui interviennent dans la prise de décision du spectateur face aux choix interprétatifs que l'on vient de décrire. Dans son évaluation de la pertinence de ces options, ce spectateur est orienté par plusieurs variables, parmi lesquelles on énumérera : 1) sa connaissance de la sémantique des codes iconiquement représentés, 2) ses connaissances encyclopédiques et 3) les indications génériques.

Pour illustrer le rôle des codes iconiquement représentés, on peut donner l'exemple de la sémantique gestuelle : le geste d'Uncle Sam (figure 15) a un signifié centrifuge, ici mobilisé le long de l'axe frontal. C'est cette valeur qui, rendant couteuse l'hypothèse selon laquelle le protagoniste s'adresserait à lui-même, prédispose en contrepartie à l'identification d'un énonciataire externe : en effet, celui-ci est idéalement situé dans un axe perpendiculaire au plan de l'énoncé.

Pour illustrer le rôle de la seconde variable — les connaissances encyclopédiques —, on peut se tourner du côté des images pieuses : c'est la connaissance de la personne et du rôle du Christ ou des saints qui permet l'interprétation correcte du "vous", et son assignation à un référent tantôt interne (quand le fidèle s'adresse au Christ ou au saint) tantôt externe (quand c'est le Christ ou le saint qui s'adresse au fidèle). Dans la proposition d'abonnement à *Fluide glacial*, c'est la connaissance des rôles dévolus aux acteurs du processus d'abonnement qui permet de les identifier l'un comme « abonné », l'autre comme abonné.

Les contraintes de genre — troisième variable — jouent également un rôle important. On peut ainsi estimer que le « vous » de la publicité a *a priori* de grandes chances d'être centrifuge, s'adressant à un spectateur qui est un acheteur potentiel, et que celui de l'image pieuse sera préférentiellement centripète, s'adressant souvent à Dieu ou à un saint.

### 3.1.5. Rhétorique de la deixis

Tout ceci montre que bien que des hésitations peuvent survenir quant à la manipulation correcte des instruments syntaxiques que sont les déictiques, et que des ambiguïtés ou des superpositions sont possibles. Les énoncés pluricodiques peuvent évidemment jouer sur ces propriétés, ce qui produit une véritable rhétorique de la deixis.

Ainsi, le « vous » de telle publicité montrant une simple silhouette humaine et interrogeant /Que serait la mode sans vous ?/ est-il interne, se référant alors à la personne représentée (qui peut évidemment représenter la communauté des énonciateurs de l'énoncé global) ? ou est-il externe, se référant alors directement à l'énonciateur dudit énoncé global ? Et l'efficacité de cette autre publicité pour une cinémathèque — un dos de femme, avec des bretelles de soutien-gorge figurées par des pellicules de films et la mention /découvrez-moi / — ne vient-elle pas du fait qu'elle superpose deux référents possibles du « moi », correspondant à deux sens du mot /découvrir/ ? Un « moi » qui est celui de l'énonciatrice interne, laquelle invite son partenaire à la déshabiller, et un « moi » externe, faisant de la cinémathèque l'énonciatrice ?

Quant à la réclame pour *Fluide glacial*, riche en déictiques, le « nous » y est simultanément externe (il se réfère alors à « l'abonné », comme on l'a dit) et interne : la présentation sous forme iconique de l'arme de poing susceptible de tuer le personnage présuppose bien, en effet, un tueur interne.

## 3.2. Connecteurs

Le mot connecteur, qui a suscité bien des controverses, désigne globalement des outils linguistiques exprimant des liens syntaxiques, logiques ou sémantiques entre deux portions d'énoncés<sup>20</sup>. Sont ainsi des connecteurs : 1) les termes temporels, qui organisent tantôt des

---

<sup>20</sup> « Un connecteur pragmatique est un mot grammatical (conjonction, adverbe, locution) dont la fonction est d'une part de relier des segments de discours (les énoncés), et d'autre part de contribuer à la constitution d'unités discursives complexes à partir d'unités discursives simples" (Moeschler & Reboul, 1994 : 465).

## La relation texte-image

simultanités d'évènements (« pendant ce temps-là »), tantôt des successions d'évènements (« d'abord », « alors », « ensuite », etc.) ; cette succession pouvant être orientée vers l'antériorité (« l'année précédente ») ou la postériorité (« le lendemain ») ; 2) les termes indiquant un rapport logique, par exemple de causalité ou de conséquence (« donc », « ainsi »), de contradiction (« mais », « toutefois ») ou de sommation (« bien plus », « en outre »), etc.

Une des propriétés linguistiques des connecteurs est que certains d'entre eux peuvent tantôt relier des portions d'énoncé (nous parlerons alors de *connexion interne*), tantôt un énoncé et son énonciation (et on parlera alors de *connexion externe*). Ainsi, un connecteur temporel peut exprimer un lien chronologique entre deux épisodes dans l'histoire racontée — deux unités narratives —, mais il peut aussi connecter deux phases du récit racontant. Il fournit alors des indications sur la dynamique du discours narratif lui-même. Par exemple, « enfin », qui est susceptible d'avoir une première valeur chronologique dans un récit (« Enfin Malherbe vint... »), peut aussi renvoyer au déroulement du récit racontant lui-même : « enfin » voulant alors dire « pour terminer ce récit ».

Ainsi, outre qu'ils donnent une unité à l'énoncé, les connecteurs participent également à la structuration, à l'organisation du texte (comme le font les connecteurs logiques du genre de « d'une part, d'autre part »). Mobilisé dans un énoncé scripto-iconique, un connecteur aura donc un impact sur la partie iconique V de l'énoncé en lui donnant un statut argumentatif. Par exemple en en faisant la première partie d'une opposition (dont la seconde serait signifiée par le contenu linguistique L), ou en en faisant l'antécédent d'un conséquent (signifié par le contenu linguistique).

Nous ne pouvons examiner ici le fonctionnement de tous les connecteurs possibles dans le discours scripto-iconique. Nous n'en examinerons que deux cas, à titre d'exemples : celui du connecteur logique « ou bien » et celui du connecteur dialogique.

### 3.2.1. Connecteurs logiques

**S**oit une publicité pour les voitures Volkswagen : un dessin au trait stylisé montrant un être humain au faciès contracté, et portant à la tempe le pistolet d'une pompe à essence. Le texte énonce : /Ou bien achetez une Volkswagen

Conformément à la règle qui vient d'être établie, le connecteur logique « ou bien » introduit le sous-énoncé L « achetez une Volkswagen », mais a aussi, en retour, un impact sur l'interprétation du sous-énoncé iconique.

Ce dernier, s'il était isolé, ferait l'objet d'interprétations très ouvertes, que divers calculs inférentiels complexes, décrits par ailleurs (des sous-entendus, des présupposés de type  $\pi$ , des tropes visuels de type « chafetière » autant que de type « rencontre »<sup>21</sup>), permettraient de formuler de la manière suivante : « cet énoncé iconique représente un automobiliste désespéré par le prix de l'essence, et désespéré au point d'envisager le suicide ». Mais l'articulation que le connecteur opère en connectant à ce contenu flou la proposition L « achetez une Volkswagen » permet de projeter les données de ce second énoncé sur le premier V. Il permet par exemple d'en particulariser le référent du « vous » (puisqu'il fait de ce personnage l'énonciataire du message), et de modaliser le contenu de l'énoncé en « suivez-vous car l'essence est chère » (ou en « il ne vous reste qu'à vous suicider »).

### 3.2.2 Connecteurs dialogiques

Les connecteurs dialogiques sont fondés sur les attentes qu'induisent les différents types de prise de parole, et qui sont des effets de la force illocutoire dégagée par l'échange.

Ainsi, toute question implique une réponse (l'absence de réponse étant encore une manière de répondre...), tout ordre implique une réaction, etc.

Mobilisé dans un énoncé scripto-iconique, un connecteur dialogique L donnera donc à la portion iconique V de l'énoncé une valeur tantôt de réponse, tantôt de réaction, etc.

Soit une publicité montrant en V deux /claviers d'un ordinateur portant une gaufre et dans la position des deux éléments d'un gaufrier/, métaphore iconique du type chafetière. A la question L /Que faire de votre PC si vous n'avez pas Net Banking ?/, l'image V apporte la réponse : « si vous ne l'utilisez pas pour la connexion à Net Banking, il ne vous reste plus qu'à l'utiliser pour faire des gaufres ». Si l'on généralise la réponse iconique en « votre PC n'est susceptible d'être utilisé que pour faire des gaufres ou pour en faire d'autres usages pour lesquels un tel appareil n'est pas particulièrement indiqué », on en déduit ceci : « l'usage idéal d'un PC est de permettre l'accès à Net Banking. » Soit une affiche pacifiste, énonçant en L

---

<sup>21</sup> Sur tout ceci, cfr Klinkenberg, 1993 et 2002a, Groupe  $\mu$ , 1992, Kerbrat-Orecchioni, 1986.

/Repos !/ et montrant en V un revolver au canon amolli et pendouillant. Le trait de « flaccidité» qui est rhétoriquement attribué au canon de l'arme de poing peut être considéré comme une réaction, iconiquement représentée, à l'ordre L /repos !/.

### **3.3. Index non-linguistiques explicites**

L'examen des articulations scripto-iconiques par déictique ou par connecteur suggère que la connexion est toujours à charge du composant linguistique L de l'énoncé global. Et effectivement, lorsque la connexion semble à charge du composant iconique V, c'est que celui-ci a iconisé le processus de production d'un signe indexical (par exemple un personnage ayant le doigt tendu, le doigt constituant l'indexant).

La connexion peut en effet également se faire par ce que j'appellerai des *index explicites*. Les indexants sont ici des signes non linguistiques, manifestés à la surface de l'énoncé, qui permettent d'articuler explicitement la portion linguistique L de l'énoncé scripto-iconique avec sa portion iconique V.

Parmi ces indexants explicites, je citerai :

- les /traits/ indiquant l'équivalence de la portion linguistique L de l'énoncé et de sa portion iconique V (c'est le phénomène connu sous le nom de co-référence). Exemple fréquent dans les encyclopédies ou dans les infographies
- la /flèche/, qui a cette même fonction, en y ajoutant la connotation de « précision ».
- la /queue de phylactère/, fréquente dans la bande dessinée (voir les figures 11a, b et c), qui joue ce rôle indexical, avec cette fois la valeur précise de « embrayeur d'énonciation ».

Dans ces index non-linguistiques, des éléments apparemment linguistiques peuvent intervenir : par exemple des lettres. Mais c'est pourtant bien ici qu'il faut en traiter, car ces lettres n'interviennent pas à titre de composants d'un énoncé linguistique, mais à titre de composant spatial de l'indexant : des astérisques, ou des chiffres pourraient d'ailleurs parfaitement les remplacer.

La particularité est qu'il y a ici une double indexation. Première indexation : un point dénommé x (indexant) désigne un élément iconique de l'énoncé (indexé). Deuxième indexation : un autre point également dénommé x (indexant) désigne le point spatial qu'est le début d'un texte. L'indexation globale repose sur l'équivalence des deux x, de sorte que la

relation indexicale est transitive : ( x ↔ icône) + ( x ↔ texte)    texte ↔ icône.

Car on le voit, ce qui compte, ce n'est pas la valeur linguistique (a, A) de l'indexant, mais sa double position dans l'espace ; et de ce point de vue, répétons-le, un astérisque aurait pu faire l'affaire. Évidemment, le caractère littéral de l'indexant ajoute une valeur supplémentaire, facultative toutefois. Cette valeur, c'est la séquentialité qui provient de la convention alphabétique : elle introduit un certain ordre dans l'ensemble des rapports indexicaux (successivité, hiérarchie, etc.) ; mais ici encore, des chiffres (1, I) auraient pu faire l'affaire<sup>22</sup>.

Il n'est pas question de faire ici un relevé complet des procédures d'indexation explicite.

Notons toutefois que, selon les genres discursifs historiques, les énoncés peuvent faire un usage très sophistiqué de ces index explicites. Une des sources de cette sophistication provient du fait que l'indexation est une relation transitive. Elle peut donc être exploitée de manière récurrente, démultipliée, et cela de manière en principe infinie. On peut par exemple avoir un premier couple d'indexants littéraux conforme à la description qui vient d'être donnée ; mais, pour obtenir plus de précision, le point sur lequel porte la première indexation peut-être lui même précisé par un indexant de la famille du trait et de la flèche.

### **3.4. Index implicites**

L'indexation peut avoir lieu même lorsqu'il n'y a pas dans l'énoncé de signe spécialisé pour jouer le rôle d'indexant. Cette fonction indexicale peut en effet être assumée par la simple disposition spatiale des éléments du message scripto-iconique

Si un énoncé L /pipe/ apparaît, comme dans *La trahison des images* de Magritte, en dessous d'une icône V de pipe (ou de toute autre icône, d'ailleurs), ou si une mention L /température/ est situé le long d'un vecteur ou d'une courbe dans un graphique, ces dispositions indiquent à suffisance que l'énoncé graphique voisin doit être lu comme ayant la signification « pipe » ou « température ». Nous parlons en ce cas d'*index implicite*.

---

<sup>22</sup> C'est d'ailleurs la règle générale qui est d'application avec la note infrapaginale, qui, bien que ne mobilisant pas d'icône, relève du type d'indexation que nous étudions ci.

## La relation texte-image

Ceci semble bien d'abord être une simple application de la loi de proximité de Gogel (1978), lequel a démontré que deux manifestations visuelles s'influencent d'autant plus qu'elles sont proches. Mais l'interaction, en l'occurrence, n'est pas qu'une donnée physiologique : l'index étant un signe culturel, la sémiotisation particulière de cette proximité physique est elle aussi culturelle.

Dans l'index implicite, le plan de l'expression est ainsi constitué par un mode particulier et codé d'occupation de l'espace. L'énoncé graphique est alors un signe dont on doit savoir, grâce à une règle conventionnelle, qu'il renvoie d'une certaine manière à un objet donné, contigu à ce signe. Prenons l'exemple des mentions sur les édifices : on sait, parce qu'on l'a appris, qu'une mention comme /gendarmerie/ ou /boulangerie/ désigne la qualité (« gendarmerie », « boulangerie ») attribuée à la totalité du volume spatial clos situé en arrière du plan où figure l'expression.

Ce sont ces index implicites que l'on trouve sur les vitrines, les immeubles, les titres d'ouvrages ou d'œuvres picturales, les salles de musée, de classe ou de congrès, les badges de personnel ou des participants à ces congrès, les noms des présentateurs de télévision apparaissant au bas de l'écran, les étiquettes de boîtes à conserve ou de bouteilles, les noms de défunts sur les tombes ou de vivants sur les portes...

Dans les énoncés scripto-iconiques, le sous-énoncé graphique L a avec le sous-énoncé iconique V une valeur de co-référence, comme dans les premiers cas envisagés en 2.3.

Il peut s'agir d'une co-référence relative à des objets (comme dans le cas de la légende en dessous d'une planche d'encyclopédie) mais aussi à des procès (comme avec l'onomatopée accompagnant une explosion dans les bandes dessinées).

### **3.5. Syntaxe, séquentialité et spatialité**

#### *3.5.1. Relations séquentielles entre éléments L et V*

Un certain nombre des dispositifs syntaxiques étudiés établissent une relation entre éléments V et L qui peut être qualifiée de simultanée. C'est par exemple le cas de la relation entre énonciateur et énonciataire (cfr § 3.1.). On peut parfaitement comprendre la source de cette simultanéité : énonciateur et énonciataire sont des instances qui se postulent mutuellement.

D'autres dispositifs syntaxiques établissent une relation de séquentialité entre éléments V et L. C'est particulièrement le cas avec les connecteurs. Dans ce type de configuration, le binôme verbo-iconique peut être décrit comme constituant une séquence, où les éléments entretiennent une relation de type n<sup>23</sup>. Ce phénomène n'a rien pour étonner quand on se souvient du caractère ambivalent de l'espace dans lequel se déploie l'écriture : à la fois spatial et linéaire (cfr § 0.3.). Linéaire, et donc susceptible d'accueillir l'expression de relations vectorielles (cfr Klinkenberg, 2005).

La présence ou l'absence de cette séquentialité, comme sa nature exacte, dépendent du type de relation sémantique mobilisée dans la connexion. Une telle séquentialité peut par exemple être observée dans les paires Ordre-Exécution, Question-Réponse, Circonstance-Action, Antécédent-Conséquent, Cause-Effet, etc.

En principe, chacune des deux variables entretenant le lien séquentiel est susceptible de s'actualiser indifféremment dans un contenu verbal ou dans un contenu iconique. On rencontre donc aisément des exemples illustrant la combinaison question iconique + réponse verbale autant que la combinaison question verbale + réponse iconique

Il est possible que toutes les combinaisons théoriquement possibles ne se réalisent pas dans les faits. C'est ce que devrait établir un examen exhaustif des connexions verbo-éidétiques.

### *3.5.2. Actualisation spatiale des relations syntaxiques*

Il convient de distinguer soigneusement la linéarité qui vient d'être décrite, et qui est une propriété de la relation syntaxique, avec la manifestation de cette relation dans l'espace. Par exemple, dans les derniers cas invoqués, on peut parler d'une linéarité allant de l'ordre à l'exécution, mais cette linéarité n'implique nullement une disposition spatiale particulière des éléments V et L prenant en charge ces deux éléments.

La disposition spatiale des éléments mis en relation syntaxique n'est toutefois pas totalement contingente. On observera en effet un certain nombre de constantes dans ces dispositions. Pour les expliquer, deux facteurs sont à mobiliser

---

<sup>23</sup> Concept qui sera expliqué ci-après en 4.1.3. Sur son application à la temporalité, cfr Groupe  $\mu$  1997.



## La relation texte-image

Le premier est, à nouveau, le caractère ambivalent de l'espace scripto-ictonique. Quoique se déployant dans un espace à deux dimensions (au moins), l'écriture induit des relations linéaires dans cet espace. Dans la mesure où une langue donnée s'écrit par de droite à gauche, on peut s'attendre à ce que cette direction soit privilégiée dans les énoncés scripto-ictoniques qu'elle permet. Un privilège qui sera d'autant plus spectaculaire que la part de l'écriture sera grande dans l'énoncé global.

On peut s'attendre aussi à ce que les relations syntaxiques linguistiques soient iconisées dans l'espace de l'énoncé. Ainsi, la version observée de l'énoncé « Ou bien achetez une Volkswagen », où L se situe en dessous de V, apparaît comme plus grammaticale qu'une version du même message publicitaire où V serait en dessous de L, qui serait ainsi en tête de l'énoncé global. Dans la première version, le connecteur « ou bien » est disposé spatialement entre le sous-énoncé V et le sous-énoncé L, ce qui peut apparaître comme un équivalent de sa position normale dans la phrase et qui est visuellement conforme à son rôle logique. Au lieu que dans la deuxième version, son emplacement en position supérieure gêne la lecture, bien que celle-ci reste parfaitement possible.

Le deuxième facteur à prendre en considération est d'ordre pragmatique et son rôle a déjà été évoqué à plus d'une reprise : ce sont les conventions génériques. Par exemple, une disposition du texte L en dessous de V est plus conforme aux usages des manuels scolaires, des encyclopédies, du discours scientifique et de bien d'autres types d'ouvrages.

### 4. Relations sémantiques

**N**ous n'avons pu étudier les relations syntaxiques entre texte et image sans faire référence à des phénomènes sémantiques, comme la redondance. Il est temps à présent d'aborder résolument cet aspect, en nous focalisant sur le rapport entre les contenus des deux sous-énoncés de l'énoncé global scripto-ictonique<sup>24</sup>.

Si l'on compte bien, théoriquement, il n'y a guère que trois configurations possibles :

1. soit il y a redondance entre les deux contenus (L = V).

---

<sup>24</sup> La question du rapport entre le contenu de l'icône lorsque celle-ci a pour référent l'énonciateur du message linguistique et celui de l'énoncé linguistique est d'une autre nature.

2. soit  $L \neq V$ , et il y a alors rupture de l'isotopie — ou homogénéité sémantique de l'énoncé —, rupture nommée allotopie<sup>25</sup>. Cette allotopie doit être gérée, toujours selon les principes de coopération et d'économie rappelés plus haut. Ce qui débouche donc sur une adaptation d'un des deux contenus. C'est ce que nous appellerons l'accommodation.

3. soit il y a opposition ( $L \longleftrightarrow V$ ). Ce cas peut apparaître comme une variante du précédent.

Un quatrième cas, un peu différent des trois premiers, peut survenir. C'est celui où (a) le sens du sous-énoncé L ou celui du sous-énoncé V, ou les deux, ne fourniraient pas une lecture débouchant sur une interprétation satisfaisante<sup>26</sup> dans l'hypothèse où chacun de ces sous-énoncés resterait isolé, mais où (b) la conjonction des deux énoncés ( $L \times V$ ) débouche sur une lecture satisfaisante.

Chacun de ces types peut évidemment se moduler selon le sémantisme particulier des éléments linguistiques présents. Par exemple, une prédication (faisant par exemple usage du verbe « être ») n'actualisera pas le modèle « opposition » de la même manière qu'un verbe transitif.

#### **4.1. Redondance**

**L**a redondance est le corrélat sémantique du phénomène syntaxique de la co-référence. C'est vers elle que tendent certains discours didactiques, les anciennes publicités ou certains livres pour enfants .

Si l'intervention des index confère une homogénéité aux énoncés scripto-iconiques, on peut dire que cette homogénéité a une source externe. Mais cette homogénéité est aussi interne : et c'est la redondance qui l'assure ; la notion d'isotopie est indissociable de celle de redondance.

##### *4.1.1. Surcroît linguistique et surcroît iconique*

**I**l faut toutefois immédiatement préciser que la redondance ne peut par définition être que partielle.

En effet, les sémiotiques linguistique d'une part et iconique d'autre part n'ont pas, par définition, les mêmes potentialités. Par exemple, pour les raisons qu'on a exposées au paragraphe 0.3., la seconde permet la

---

<sup>25</sup> Sur la notion d'allotopie, cfr Groupe  $\mu$ , 1977,

<sup>26</sup> Sur la notion de lecture satisfaisante, cfr Sperber et Wilson, 1986 et Klinkenberg, 2002b.

## La relation texte-image

présentation simultanée des objets, ce que ne permet pas la première ; en revanche, celle-ci autorise par exemple l'expression de modalités ou de quantifications inaccessible à la seconde. Chacun des deux sous-énoncés V ou L offre donc nécessairement un *surcroit* d'information par rapport à l'autre, surcroit qui n'est le privilège ni de l'image — comme le voulait Barthes — ni du texte.

Soit un dessin, dans un livre pour enfant, énonçant, tant en V qu'en L, /une allumette grillée /. Ce que le sous-énoncé L relatif à l'allumette ne nous dit pas, c'était de quel type d'allumette il s'agit : son référent est en effet une allumette générique, qui aurait aussi bien pu être plate, être en carton ou en cire ; et c'est à V qu'il revient de préciser que l'allumette grillée est une allumette en bois à section carrée. Dans telle page d'un album pour enfant, du à Neal Layton, le spectateur est prêt à croire L, qui lui affirme que V lui montre des mammouths laineux (/Oscar était un mammouth laineux et Arabela aussi/). Mais l'énoncé L nous livre le nom de ces mammouths : Oscar et Arabella. Le linguistique apporte donc un surcroit d'information. A l'inverse, le sous-énoncé V nous assure qu'Oscar porte un bonnet de laine, ou nous parle du paysage de neige et de montagne qui est le biotope des bebêtes, toutes choses sur lesquelles L reste muet. Telle séquence de la tapisserie de Bayeux, comme énoncé global, nous montre un combat (« prelium »), sens exprimé de façon redondante par L et V. Mais les deux sous-énoncés déclinent cette donnée de base redondante chacun à leur manière : à V d'exprimer la présence de chevaux, de fournir le détail des équipements, d'attirer l'attention sur certaines conséquences particulières du combat comme la décapitation, à L d'identifier les combattants comme étant des « Angli » et des « Franci ».

### 4.1.2. Réorganisation de l'espace et réorganisation des contenus

Outre qu'elle met en évidence les surcroits iconique et linguistique, la redondance organise l'espace scripto-iconique commun. Ceci est un corrolaire de ce qui précède : la redondance n'associe en effet qu'une portion de l'espace iconique et une portion de l'espace langagier seulement. Aucun des deux messages n'est donc réductible à l'autre, mais on peut par contre dire qu'ils sont faits d'une série d'éléments dont certains peuvent être communs : ce sont des ensembles logiques en intersection et en exclusion réciproque.

Chacun des sous-énoncés segmente donc les portions d'espace relevant de l'autre. On peut bien faire voir cette manœuvre de segmentation dans l'exemple presque caricatural de l'*Album primo-avrilesque* d'Alphonse Allais : chacune des planches V de cet album nous donne à voir une surface

uniforme, un monochrome tantôt rouge, tantôt noir. Il faut l'intervention presque désespérée du titre L « Récolte de la tomate par des cardinaux apoplectiques au bord de la mer rouge (effet d'aurore boréale) » ou « Combat de nègres dans une cave, pendant la nuit » pour accepter docilement de discrétiser le continuum de V, et admettre que ce /rouge/ renvoie iconiquement à des entités distinctes comme « cardinaux », « tomates », etc., et que ce /noir/ renvoie à des entités qui se déclinent en « nègre », « cave », etc. Dans ce cas-ci, il est indéniable que l'ancrage barthésien joue un rôle important.

Cette sélection, pur phénomène d'indexation, a un corrélat sémantique : la pertinentisation, ou processus d'attribution de sens à des formes discriminées par rapport au fond sur lequel elles se détachent. C'est un processus au cours duquel le récepteur opère un calcul du coût cognitif de chacune des hypothèses envisageables.

En vertu de sa constitution sémantique propre, le sous-énoncé verbal L sélectionne certains éléments dans le sous-énoncé iconique V. Mais l'opération inverse est vraie également. Prenons pour exemple une affiche ancienne pour la compagnie de navigation « Red Star Line », montrant des personnages sur une rive, regardant passer un transatlantique. L'énoncé verbal L, relatif à la navigation en haute mer entre Anvers et New-York, rend très pertinente l'icône V « bateau », et moins pertinents les éléments iconiques V ( le « bac à sable » où joue un enfant, le « panier » qu'une des spectatrices a au bras). Mais inversement, la nature de l'icône (un paysage scaldien) rend, dans L, « Anvers » plus pertinent que « New York ». On peut donc parler d'une mise en évidence, ou de double mouvement de pertinentisation.

#### 4.1.3. Réorganisation sur le mode $\Pi$ et sur le mode $\Sigma$

Chacun des sous-énoncés L et V réorganise les portions d'espace relevant de l'autre aspect, V ou L, établissons-nous au paragraphe précédent. Et nous avons vu que cette manœuvre avait un impact sémantique : la pertinentisation des zones sélectionnées.

Il faut à présent raffiner cette description globale, et faire apparaître que la réorganisation peut procéder tantôt sur le mode  $\Pi$ , tantôt sur le mode  $\Sigma$ , ce qui n'est pas sans conséquence sur l'impact sémantique de l'opération. Un mot d'explication sur ces deux points de sémiotique générale est sans doute nécessaire ici<sup>27</sup>.

---

<sup>27</sup> Sur ceci, voir entre autres Palmer 1975 et 1977, Groupe  $\mu$ , 1992, Klinkenberg, 2000a.

## La relation texte-image

Pour rappel, la relation  $\Pi$  est une relation sémantique fondée sur la relation du tout à ses parties (exemple : « arbre et « tronc »), chaque partie subordonnée au tout entretenant des relations de coordination avec une ou des contre-parties situées au même niveau de subordination (exemple : « tronc » et « racines », deux contre-parties, également subordonnées à « arbre » et coordonnées entre elles). La relation est ainsi nommée parce que les parties coordonnées entretiennent des relations de produit logique (si  $x$  est un arbre, alors  $x$  a un tronc ET des racines, etc.). Ces relations sémantiques  $\Pi$  se distinguent des relations  $\Sigma$ , que la sémantique linguistique traditionnelle a seules accepté de prendre en considération : les relations mobilisées ici sont celle de l'hyponymie et de l'hyperonymie (du genre et de l'espèce). La relation est dénommée  $\Sigma$  parce que les éléments dégagés par l'articulation entretiennent des relations de somme logique (si  $x$  est un arbre, alors  $x$  est un bouleau OU un chêne OU un hêtre, etc.)

Comme l'image relève d'une sémiotique de l'espace, on pourrait croire que seules les relations de type  $\Pi$  y sont pertinentes, et que par ailleurs la sémantique linguistique ne connaît que des relations  $\Sigma$ . Ce n'est pas le cas. On a d'une part démontré qu'une sémantique linguistique complète devait prendre en considération le facteur  $\Pi$  (cfr Groupe  $\mu$ , 1970, Edeline 1972), ce que la logique a d'ailleurs fait, avec la méréologie. D'autre part, contrairement à ce que l'on entend souvent affirmer, l'image peut exprimer une certaine généralité<sup>28</sup>, et est donc susceptible d'exprimer des rapports de type  $\Sigma$ , surtout avec l'aide de sous-énoncés langagiers, comme on va le voir.

---

<sup>28</sup> En affirmant ceci, nous nous opposons à Anna Wierzbicka (1985 : 83, citée par Kleiber, 1990 : 67) pour qui « l'image ne saurait avoir qu'un rôle d'illustration et non pas un rôle théorique dans l'analyse d'un prototype. On ne peut tracer l'image d'une tasse prototypique », argumente-t-elle, « parce qu'une image n'a qu'une forme et ne peut saisir le continu de la série des formes compatibles avec le concept de tasse ». La prémisse du raisonnement est exacte (une image n'a qu'une forme et ne peut saisir le continu de la série des formes compatibles avec le concept), mais non sa conclusion (selon laquelle on ne pourrait tracer une image prototypique). Il est en effet exact que l'icône se conduit différemment du mot quant à son extension : le mot /bœuf/ permet d'atteindre la totalité des instances visées par le signifié « bœuf », et non seulement les occurrences particulières, dotées d'une existence réelle, mais même les instances virtuelles. Comparée au mot, l'icône de bœuf a une extension nécessairement inférieure (on a pu, plaisamment, faire observer qu'en tant qu'icônes, les bovins des panneaux du code de la route ne renvoient pas au boeuf dans toute son extension, mais bien à un « charolais »). Toutefois, on ne peut ainsi raisonner par tout ou par rien, et refuser le rôle d'indicateur de prototypicalité à l'image, sous prétexte d'une différence, bien réelle, avec la langue, qui permet, elle, la saisie de ce « continu de la série ». Comme le prototype étudié des linguistes et des psychologues, le type iconique est une catégorie ouverte, non contingente. Une catégorie (par exemple « bovidé ») peut être représentée par une sous-catégorie, également non contingente (par exemple « charolais »). Si on peut produire l'icône d'une catégorie, c'est parce que les membres de la catégorie sont perçus comme partageant le même faisceau de traits catégoriels discriminatoires. Ceci ne signifie pas qu'ils possèdent du coup une forme

Mais reste que le caractère spatial de l'image rend particulièrement rentable l'exploitation des configurations où c'est le sous-énoncé L qui focalise l'attention sur une portion de l'énoncé I segmenté sur le mode  $\Pi$ . Tous les exemples mobilisés jusqu'à présent dans ce paragraphe relèvent bien de ce mécanisme, comme aussi celui du tableau de Brueghel bien connu intitulé *Paysage avec la chute d'Icare* : la première partie de l'énoncé linguistique (« paysage ») fait que celui-ci est certes redondant avec l'ensemble V, mais sa seconde partie focalise l'attention sur un fragment particulier de l'ensemble V « paysage », au détriment des autres, un fragment qui aurait d'ailleurs pu passer inaperçu du spectateur si le sous-énoncé L n'avait pas joué son rôle : la manifestation iconique de la chute d'Icare.

Le texte L peut ainsi aller jusqu'à discrétiser un ensemble V qui se présente comme continu à la vision. C'est l'exemple maximaliste de l'album primo-avrilisque commenté ci-dessus.

Tous les exemples allégués peuvent laisser croire que le travail de focalisation ne fonctionne que dans un seul sens : autrement dit que seul le texte L a le pouvoir de sélectionner sur le mode  $\Pi$  une partie de l'énoncé visuel. Mais cette hypothèse serait une nouvelle trace du glossocentrisme qui n'a cessé d'être dénoncé.

Comme on l'a vu à propos de la tapisserie de Bayeux, le mécanisme de focalisation agit en fait dans les deux sens, en produisant une sélection réciproque. En principe donc, le sous-énoncé V focalise l'attention sur une portion de l'énoncé L, qu'il extrait de son sous-ensemble, dans le même temps que le sous-énoncé L focalise l'attention sur une portion de l'énoncé V, qu'il extrait de son sous-ensemble. Mais bien sûr, dans cette question des rôles respectifs de L et de V, il y a des usages pragmatiques, déterminant des règles génériques : par exemple, le titre de l'œuvre est réputé donner un sens au visuel et être prioritaire.

Mais il faut à présent démontrer que les mécanismes de focalisation réciproques peuvent aussi opérer sur le mode  $\Sigma$ .

---

moyenne : le paradigme d'un trait du signifiant doit être nécessairement actualisé par une manifestation particulière sur le plan du stimulus (telle forme, telle texture, telle couleur). Mais on peut parfaitement produire le signifiant d'un « bovidé » en général, sans que ceci renvoie ipso facto à un exemplaire contingent et particulier de la catégorie. Pour que le signifiant renvoie à une catégorie générale, il suffit qu'elle représente les traits typiques de la catégorie représentée, traits jouant le rôle d'indicateurs fonctionnels de la classe : bec, queue et aile pour l'oiseau, tablette horizontale surélevée et pieds pour la chaise, queue, oreilles triangulaires et moustaches pour le chat, cornes pour le bovidé, etc.

## La relation texte-image

Dans l'ensemble icono-verbal, la composante L assume ainsi fréquemment la fonction consistant à fournir une indication sur le niveau de généralité ou de particularité à conférer à l'icône V.<sup>29</sup>

Prenons un exemple : soit la photo bien connue de l'écrivain Jacques Prévert par Robert Doisneau. Dépourvue de titre, son niveau de généralité  $\Sigma$  est indifférent aux yeux du spectateur impuissant à identifier l'individu qui a servi de modèle, et cela en vertu des spécificités de l'icône, qui la distinguent sur ce point des signes linguistiques : le type représenté peut aussi bien être « homme » (au sens général de « être humain ») que « homme » (au sens de « être humain de sexe masculin »), « homme âgé », « homme âgé assis » ou « homme âgé assis occidental et devant avoir une identité personnelle ». Mais l'indexant que constitue le titre *Jacques Prévert* pointe un être humain masculin particulier, et stabilise donc le référent à un certain niveau de généralité  $\Sigma$ . Autre exemple : tel tableau de Philippe de Champaigne est connu sous plusieurs titres, qui vont du général au particulier : « Portrait d'homme », « Portrait d'un homme », « Charles Coiffier ».

Prenons attention qu'en particularisant, le sous-énoncé L peut aussi attribuer des traits particulier au type établi à ce niveau de généralité. La photographie de Prévert par Doisneau a été réutilisée dans le cadre d'une campagne pour l'emploi, ou elle est recadrée, pour mieux faire apparaître l'affaissement du personnage, et accompagnée de la mention L « Méfiance, cet homme a l'air d'être au chômage depuis longtemps ». Ce sous-énoncé L particularise certes le référent au niveau « homme âgé assis », mais en assignant cette fois au type identifié la valeur supplémentaire « au chômage ». Mais c'eût pu être « ivrogne », « parisien », « philatéliste », etc. Le tableau de de Champaigne est connu sous un quatrième titre, qui joue un même rôle spécifiant : « Portrait de gentilhomme ».

Résumons-nous : dans la question de la focalisation, le texte L joue un double rôle indicateur :

- (i) il constitue un indicateur de généralité ou de particularité (sur une échelle  $\Sigma$ ) ;
- (ii) dans la mesure où il particularise, il particularise dans une direction sémantique précise : il introduit en effet des traits sémantiques particuliers, spécifiant le type identifié au niveau de généralité obtenu. Il fournit ainsi

---

<sup>29</sup> Exemples que je dois à Odile Le Guern, que je remercie ici pour ses observations. Voir Le Guern, 1989 et à paraître.

l'isotopie sur laquelle le contenu V de l'image doit être indexé. Nous retrouverons donc ici la question du surcroît linguistique, traité en 4.1.1.

Dernier exemple : les didascalies inscrites sur les murs des décors, dans une mise en scène très contemporaine de *La puce à l'oreille*, de Feydeau<sup>30</sup>, où le texte même de la pièce est transcrit sur les parois visibles sur la scène. Loin de structurer l'espace sur le mode  $\Pi$ , comme le font certaines autres didascalies grâce à des indications comme /porte/ ou /fenêtre/ inscrites sur des panneaux nus, les mentions ici présentes le structurent sur le mode  $\Sigma$  : étant donné que l'espace vu — constituant le sous-énoncé V — a un statut très indifférencié (ce pourrait être un hangar, un gymnase, une galerie d'art), sa valeur sémantique se voit stabilisée à un certain niveau de généralité sur l'axe  $\Sigma$ , conformément à la règle (i) ci-dessus : l'espace du sous-énoncé V est particularisé comme étant celui d'une scène de théâtre (et qui plus est, grâce au renvoi à Feydeau, d'une scène de théâtre boulevardier : illustration de la règle (ii) ci-dessus).

Encore une fois, les exemples invoqués ici illustrent un mécanisme où l'orientation sémantique est assumée par la composante L, qui gère la composante V. Mais gardons-nous une fois de plus du glossocentrisme, et évitons de croire qu'il s'agit là d'une règle absolue. Le mouvement inverse est également possible : l'icône V peut parfaitement particulariser un texte général et potentiellement applicable à de nombreux référents. De sorte que l'on peut à bon droit, si l'on entend ne pas abandonner la métaphore barthésienne, parler de « sens linguistique flottant ». Dans ce cas, si ancrage il y a, c'est bien l'icône V qui l'assure. Prenons l'exemple d'un énoncé L comme : /elle est bonne/, assurément elliptique en l'absence de contexte. Comme le démontre l'exemple qui suit, produit d'une rapide enquête menée grâce au moteur de recherche Google, on voit bien les icônes V jouer ce rôle contextuel, et indiquer dans chaque cas l'isotopie sur laquelle il convient d'indexer les sens de « bonne » et de « elle ».

---

<sup>30</sup> Je dois cet exemple à Eleni Mouratidou, que je remercie ici (cfr Mouratidou, 2009).



## La relation texte-image



### 4.2. Allotopie et accommodation sémantique

**La** redondance totale n'est donc qu'un modèle théorique, jamais réalisé. Dans le concret, on n'a que des contenus partiellement convergents.

Il arrive même parfois que le rapport entre les deux portions d'énoncés soit très ténu. Autrement dit, l'énoncé scripto-iconique, bien que désigné comme homogène par la procédure externe de l'indexation (cfr § 1), manque d'homogénéité sémantique interne : il est allotope. Dans ce type de cas, une interprétation serait très couteuse si l'on ne faisait pas intervenir une opération destinée à rendre les contenus des deux portions d'énoncé compatibles — isotopes —, où du moins à justifier et pertinentiser dans l'interprétation globale le maximum d'informations provenant de l'un et de l'autre des sous-énoncés. C'est cette opération que nous nommerons ci-après *accommodation*.

L'accommodation comporte deux modalités.

Dans la première, les données de V comme de L sont acceptées telles quelles. Simplement, elles apparaissent comment insuffisantes pour l'élaboration d'une interprétation globale satisfaisante ; il convient dès lors de procéder, à partir de ces données, à des inférences qui viennent les compléter, mais sans les mettre en cause. Nous parlerons alors d'*accommodation discursive*. Dans la seconde, les données soit de V soit de L sont considérées comme impertinentes et elles sont dès lors remises en cause. Elles font l'objet d'une transformation qui n'est autre que celle que nous trouvons dans la figure de rhétorique<sup>31</sup>. Nous parlerons alors d'*accommodation rhétorique*.

#### 4.2.1. *L'accommodation discursive*

**D**ans l'hypothèse où les contenus faiblement pertinents sont conservés, mais replacés dans un contexte plus large qui les justifie, l'accommodation peut prendre une forme discursive: celle d'un récit ou une argumentation.

Prenons une autre publicité pour une voiture Volkswagen. Elle nous montre en V /un siège de voiture, sur lequel se trouve un chien à l'air ahuri, dont l'arrière-train est enfoncé dans l'espace situé entre le siège et le dossier/. Le sous-énoncé L mentionne /La Golf TDI. Vraiment puissante/

L'intersection entre les deux sous-énoncés L et V est très faible : elle se réduit au contenu « automobile » (iconiquement, nous sommes en présence de sièges d'auto, dont le statut est clairement révélé par un de leurs composants  $\pi$  : les ceintures de sécurité). Pour justifier un maximum d'éléments de V — la position curieuse du chien, qui n'est pas aspiré par un trou noir ou qui n'est pas victime de sa curiosité, ainsi que les marques de griffes que l'on observe sur le siège —, nous devons faire intervenir la chaîne d'inférences suivante, qui mobilise à la fois des relations narratives (antériorité et postériorité) et des relations logiques (antécédence et conséquence) : la puissance d'une voiture lui permet des démarrages brutaux ; lors de ces démarrages, la force d'inertie fait que les objets déposés sur les sièges sont brutalement repoussés vers l'arrière ; cela a dû être le cas de ce chien-là, pris par surprise (surprise qui confirme la brutalité du démarrage, et donc la puissance de la voiture), et qui a vainement tenté de se retenir à quelque chose (d'où les traces de griffes).

#### 4.2.2. *L'accommodation rhétorique*

---

<sup>31</sup> cfr Groupe  $\mu$ , 1992, Klinkenberg, 1993 et 2000a.

L'accommodation des contenus peut aussi prendre la forme d'une transformation rhétorique.

Soit telle publicité pour la bière Chimay, représentée dans un verre accompagné de la mention L /verre de contact/. Le terme « verre » semble être isotope, c'est-à-dire redondant : la bière de V est bien présentée dans un « verre ». Mais, prise globalement, l'expression L « verre de contact » est allotopie : elle signifie en effet d'habitude « lentille correctrice posée sur le globe oculaire », signification peu pertinente en regard du message iconique. On ne peut donc rétablir l'isotopie qu'en donnant un autre sens à « verre de contact ». Dans ce sens nouveau, le mot « verre » retrouverait une indépendance qui le rendrait isotope avec le verre de bière iconique ; du coup, cette indépendance rend également sa liberté au mot « contact », qui est susceptible de trouver une pertinence dans le contexte V. Ce sens pourrait dès lors être « verre qui permet le contact (humain) ».

Dans le cas illustré par ce dernier exemple, la mutation opère à partir de V, qui est considérée comme l'isotopie dominante. Mais cette mutation rhétorique peut parfaitement fonctionner dans les deux sens : de V vers L et de L vers V. Dans telle autre publicité pour la crème Diadermine, montrant une dame tenant un pot de crème entre les mains, le sous-énoncé L énonce /la beauté entre vos mains/. Or, la portion « entre vos mains » de ce sous-énoncé L est redondante avec une autre portion du sous-énoncé iconique V : il y a bien quelque chose « entre les mains » de la personne représentée : un pot de Diadermine. La mutation peut donc affecter et « beauté » L (la beauté, c'est Diadermine) et « Diadermine » V (Diadermine, c'est la beauté).

### **4.3. Contradiction et changement de grille**

La différence entre V et L peut être poussée fort loin : jusqu'à la contradiction.

Dans *La trahison des images*, de Magritte, cette contradiction est d'autant plus visible que la forme grammaticale utilisée est celle de la prédication : l'incompatibilité est totale entre le contenu de l'énoncé linguistique L /Ceci n'est pas une pipe/ et celui du seul énoncé iconique V que le spectateur ait à sa disposition, et qui semble être, traduit en termes langagiers, « Ceci est une pipe ».

Dans le cas particulier d'une allotopie aussi radicale, l'accommodation semble vouée à l'échec. La seule sortie possible consiste alors à changer radicalement de grille d'analyse. Ici, il s'agira de cesser de traiter l'icône

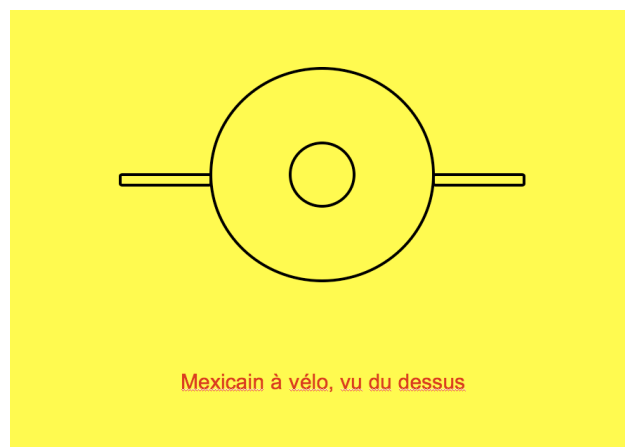
comme un équivalent de l'objet, comme on le fait fréquemment dans les usages quotidiens de l'image, mais de mettre en évidence son statut sémiotique (cfr Groupe  $\mu$ , 1999). Et c'est bien ce que la critique d'art fait dans ce cas célèbre, en soulignant que Magritte est fondé à affirmer « Ceci n'est pas une pipe » puisqu'il ne s'agit pas d'une pipe réelle mais bien d'une icône de pipe.

#### **4.4. La nécessaire collaboration des sous-énoncés**

L'exemple de l'*Album primo-avrilésque* permet de constater qu'un énoncé visuel peut être sémantiquement très opaque (au point de ne pouvoir être désigné comme iconique avec certitude), et que c'est l'intervention d'un énoncé linguistique, au sein d'un énoncé global, qui seule permet de relever le niveau de redondance jusqu'au niveau nécessaire pour obtenir une signification satisfaisante. Mais l'énoncé L peut lui-même être insatisfaisant : elliptique ou allotope.

Dans nombre de cas, c'est l'impact mutuel de L sur V et de V sur L qui permet d'atteindre un niveau de signification acceptable.

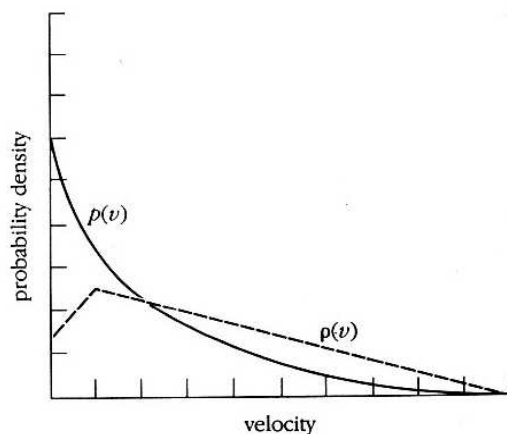
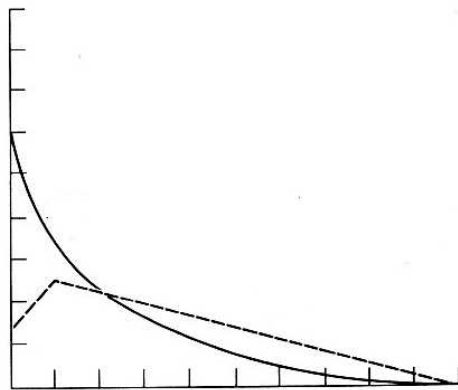
C'est cette règle qu'illustrent les devinettes du genre de celle qui suit. Dans ce cas, c'est le point de vue — qui n'est rien qu'un type particulier de cadrage — qui a abaissé le niveau de redondance. De sorte qu'il faut l'énoncé L « un Mexicain à vélo vu d'en haut » pour obtenir l'identification iconique nécessaire.



Amusement sans grande portée ? Non : ce type de collaboration est celui que l'on trouve dans la plupart des diagrammes scientifiques (lesquels ne sont que très lointainement iconiques). Le diagramme n'a pas en soi besoin

## La relation texte-image

de contenu, puisqu'il énonce des relations de relations. Mais lorsqu'il est présenté seul, comme ci-après dans l'exemple de gauche, il le fait sans se référer à des phénomènes particuliers, et sans fournir d'indications sur la mesure de ces phénomènes. Or il va de soi que nous les manipulons le plus souvent les diagrammes pour les appliquer à des contenus : des phénomènes. Ces phénomènes sont désignés par la composante L des diagrammes (les « données littérales »). Mais à elle seule, cette composante est impuissante à exprimer l'idée de règle et de relation. Il faut la collaboration des formants L et V, que l'on observe dans l'exemple de droite, pour que le diagramme puisse réconcilier le particulier et le général, le syntagme et le paradigme (cfr Klinkenberg, à paraître b).



The average distribution of velocities in an image.

## 5. Conclusion

L'esquisse de grammaire générale qui précède n'épuise évidemment pas tous les cas d'articulation entre image et texte. Comme on l'a dit, elle n'a pas donné le détail de tous les cas possibles de connecteurs, de tous les cas possibles d'embrayage ; elle n'a pas envisagé tous les cas de redondance modulés par le sémantisme de L ; elle n'a qu'effleuré la question importante des schémas et diagrammes ; elle n'a que très partiellement touché à la question du rapport entre l'énoncé et l'énonciation ; elle a laissé de côté la vaste question de la mise en page et des stratégies de lecture (cfr Béguin, 2006, Van der Linden, 2006) ; elle n'a enfin envisagé la question du rapport verbo-iconique qu'en termes de langage écrit (alors que le même type d'interaction existe évidemment avec la langue parlée : depuis le fait de prononcer le nom /Pinocchio/ en faisant le geste d'étirer son nez ou d'iconiser des « guillemets » avec ses doigts jusqu'aux spectacles de lanterne magique).

C'est évidemment à la question de la variation sociale qu'il faut associer celle des genres, dont le rôle n'a été signalé que rapidement. Pour des raisons de clarté, j'ai privilégié des énoncés relevant du genre de la publicité où la relation scripto-iconique fonctionne de manière très spectaculaire. Considérer d'autres corpus — certains aux codes très rigides, comme l'image pieuse, la bande dessinée classique européenne, les encyclopédies, d'autres plus souples comme « les mots dans la peinture » contemporains —, ou encore des corpus anciens, ou extra-européens, permettrait sans nul doute de mettre en évidence des spécificités dans la manipulation des outils généraux ici mis au point.

Mais l'esquisse ici fournie indique à suffisance que ce problème général peut trouver des solutions dans le cadre d'une sémiotique générale, dont le postulat est que le sens est un. Dans le cas présent, cette perspective conduit à ne pas surfaire le rôle de la langue dans un type de communication mixte que toutes les cultures ont exploité, et qu'elles seront amenées à exploiter davantage au fur et à mesure que les techniques multimédias se diffuseront.

## Références

- Badir, Sémir, 2005, À quoi servent les graphiques, *Communication et langages*, 143 : 49-60.
- Badir, Sémir, 2007, Rhétorique des graphiques dans les médias contemporains, *Actes du VIIIe congrès de l'IASV-IASV « Cultures du visible »*, Istamboul, Istanbul Kültür Üniversitesi Yayınları : 279-299.
- Barthes, Roland, 1964, Rhétorique de l'image, *Communications*, 4 : 40-51.
- Beguïn-Verbrugge, Annette, 2006, *Images en texte, image du texte. Dispositifs graphiques et communication écrite*, Presses universitaires du Septentrion (= Communication).
- Bertin, Jacques, 1967, *Sémiologie graphique*, 2e édition, Gauthier-Villars, Paris 1973.
- Bosredon, Bernard, 1997, *Les titres de tableaux. Une pragmatique de l'énonciation*, Paris, PUF (= Linguistique nouvelle).
- Cárdenas, Viviana (2001), Lingüística y escritura : la zona visuográfica, La dimensión plástica de la escritura, n° spécial de *Tópicos del seminario*, 6 : 93-141.
- Christin, Anne-Marie, 2001, *L'image écrite ou la déraison graphique*, Paris, Flammarion (= Champs).
- Desgoutte, Jean-Paul, 2003, *Le verbe et l'image. Essai de sémiotique audiovisuelle*, Paris, L'Harmattan.
- Desnoyers, Luc, 2005a, Les images de la communication scientifique, *Communication et langages*, 146 : 93-113.
- Desnoyers, Luc, 2005b, *La communication en congrès : repères ergonomiques*, Presses de l'Université du Québec.
- Edeline, Francis, 1972, Contribution de la rhétorique à la sémantique générale, *VS*, 3 : 69-78.
- Gogel, Walter C., 1978, Le principe de proximité dans la perception visuelle, *Pour la Science*, 9, 1 : 49-57.
- Groupe  $\mu$ , 1970, *Rhétorique générale*, Paris, Larousse (= Langue et Langage) ; rééd. : Paris, Le Seuil, 1982 (= Points, 146).
- Groupe  $\mu$ , 1977, *Rhétorique de la poésie. Lecture linéaire, lecture tabulaire*, Bruxelles, Complexe.
- Groupe  $\mu$ , 1992, *Traité du signe visuel. Pour une rhétorique de l'image*, Paris, Le Seuil.
- Groupe  $\mu$ , 1997, L'effet de temporalité dans les images fixes, *Texte*, 21-22 : 41-69.
- Groupe  $\mu$ , 1999, Les pièges de l'iconisme, dans *Magritte au risque de la sémiotique* (sous la dir. de N. Everaert-Desmedt), Bruxelles, Publications des Facultés universitaires Saint-Louis : 87-109.
- Harris, Roy, 1995, *Signs of Writing*, Londres, New York, Routledge.
- Hénault, Anne, 1983, *Les enjeux de la sémiotique. Tome 2. Narratologie, sémiotique générale*, Paris, PUF.

- Kerbrat-Orecchioni, Catherine, 1986, *L'implicite*, Paris, Armand Colin, 1986 (= Linguistique).
- Kleiber, Georges, 1990, *La sémantique du prototype. Catégories et sens lexical*, Paris, P.U.F (= Linguistique nouvelle).
- Klinkenberg, Jean-Marie, 1993, Métaphores de la métaphore : sur l'application du concept de figure à la communication visuelle, *Verbum*, n° 1-2-3, 1993 : 265-293 [numéro spécial *Rhétorique et sciences du langage*].
- Klinkenberg, Jean-Marie, 2000a, *Précis de sémiotique générale*, Paris, L Seuil (= Points).
- Klinkenberg, Jean-Marie, 2000b, L'argumentation dans la figure, *Cahiers de praxématique*, n° 35, 2000 [n° *Sens figuré et figuration du monde*] : 59-86.
- Klinkenberg, Jean-Marie, 2002a, La originalidad del sentido retórico, Helena Beristáin (ed.), *El abismo del lenguaje*, México, Universidad nacional autónoma de México, 2002 : 137-150.
- Klinkenberg, Jean-Marie, 2002b. Le cul-de-sac de l'iconisme et la sortie de secours cognitive, *Eco in Fabula. Umberto Eco in the Humanities, Umberto Eco dans les sciences humaines, Umberto Eco nelle scienze umane*, Leuven University Press, Franco Cesati Editore : 79-91.
- Klinkenberg, Jean-Marie, 2005, Vers une typologie générale des fonctions de l'écriture. De la linéarité à la spatialité, *Académie royale de Belgique. Bulletin de la classe des lettres et des sciences morales et politiques*, t. XVI, n° 1-6 : 157-196.
- Klinkenberg, Jean-Marie, à paraître a, L'interaction éido-graphique dans le discours scientifique, Actes de la journée d'études internationale *L'image dans le discours scientifique*, Liège, mars 2007.
- Klinkenberg, Jean-Marie, à paraître b, À quoi servent les schémas ? Tabularité et vectorialité, Actes des journées d'études pluridisciplinaires *Les images scientifiques, de leur production à leur diffusion*, Strasbourg, janvier 2008, à paraître dans *Protée*.
- Le Guern-Forel, Odile, 1989, Images et bases de données, *Bulletin des Bibliothèques de France*, 34, 5 : 422-435.
- Le Guern-Forel, Odile, à paraître, Image de... Entre individu et catégorie, actes du colloque international Le Groupe  $\mu$ . Quarante ans de recherche collective, à paraître dans *Protée*.
- Moles, Abraham, 1978, L'image et le texte, *Communication et langage*, 38.
- Mouratidou, Eleni, 2009 (à paraître), Discours didascalique et praxis critique. Quand le théâtre parle du théâtre et du monde", Actes du colloque international *Signes particuliers. Langue, discours, société*, Université de Bacau, novembre 2008.
- Norman, D. A. et Rumelhart, D. E. (dirs), 1975, *Explorations in Cognition*, San Francisco, W.H. Freeman.



## La relation texte-image

- Palmer, Stephen E., 1975, Visual Perception and Knowledge : Notes on a Model of Sensory cognitive Interaction, in Norman et Rumelhart (dirs) : 279-307
- Palmer, Stephen E., 1977, Hierarchical Structure in Perceptual Representation, *Cognitive Psychology*, 9 : 441-474.
- Perri, Antonio, 2001, Writing, *Key Terms in Language and Culture*, Alessandro Duranti (dir.), Blackwell Publishing.
- Sperber, Dan et Wilson, Deirdre, 1986, *Relevance. Communication and Cognition*, Londres, Basil Blackwell.
- Van der Linden, Sophie, 2006, *Lire l'album*, Le Puy-en-Velay, L'Atelier du poisson soluble.
- Wierzbicka, Anna, 1985, *Lexicography and conceptual Analysis*, Ann Arbor, Karoma Publishers.