

## PROCOLE D'ANALYSE DE LA BASE DE DONNÉES *MAGRITTE*. TOUTES LES ŒUVRES, TOUS LES THÈMES

Éric Trudel (Université Laurentienne, Sudbury, Ontario ; [etrudel@laurentienne.ca](mailto:etrudel@laurentienne.ca))  
Louis Hébert (Université du Québec à Rimouski, Québec ; [louis\\_hebert@uqar.qc.ca](mailto:louis_hebert@uqar.qc.ca))

Le présent document rend compte des principes généraux qui ont guidé l'analyse sémantique générale des images de Magritte logée dans la base de données *Magritte. Toutes les œuvres, tous les thèmes*. La plupart des exemples avancés proviennent de la base de données, mais certains ne s'y trouvent pas et sont retenus essentiellement pour leur force illustrative.

### Principe sémiotique général de l'analyse sémantique des images de Magritte

L'analyse sémantique des images magrittiennes repose sur une distinction fondamentale en sémiotique visuelle (Groupe  $\mu$ , 1992), qui admet l'existence de deux sortes de signifiés visuels : (1) les *signifiés iconiques*, associés aux figures (objets) représentées, reconnaissables ou non, lexicalisables ou non; (2) les *signifiés plastiques*, associés aux non-figures, par exemple à la texture de la peinture. La base *Magritte* ne répertorie que les signifiés iconiques présents dans les images<sup>1</sup>, c'est-à-dire en somme les éléments ontologiques qui ont qualité d'êtres, au sens large du terme.

Une autre distinction guide l'analyse des signifiés iconiques, qui se présentent selon deux types : (1) les *figures* (Greimas et Courtés, 1993), contenus qui évoquent directement l'un des cinq sens : par exemple, pour les figures dites iconiques, nous pouvons avoir *ciel* (vue), *chocolat* (goût), et pour les figures dites abstraites, *mort*, *nourriture*; (2) les thèmes (Greimas et Courtés, 1993), contenus qui n'évoquent pas directement l'un des cinq sens, par exemple la *justice*, représentée par des figures comme une balance, un tribunal, etc. La base *Magritte* retient exclusivement les figures<sup>2</sup>.

En termes de nature « ontologique », la base rend compte des entités (*pomme*, *chapeau*, *homme*, *femme*), des actions (*manger*), des activités (*équitation*), des phénomènes (*déformation*, *verticalité*) pris comme tous remarquables dans les œuvres, et note certaines propriétés de ces tous (certains matériaux et certaines parties). Signalons que les couleurs ou les tonalités thématiques ne sont pas recensées, car les reproductions sources sont en noir et blanc.

La *saillance* est le critère fondamental de description et de notation des signifiés iconiques, de leurs propriétés ou de leurs parties. La base décrit ainsi la présence des tous figuratifs et de leurs parties rendus perceptibles et importants de diverses façons : si l'image présente comme sujet principal une femme nue aux yeux fermés, en position couchée, appuyée sur les bras, un coquillage placé sur elle, et dont on voit par profil

---

<sup>1</sup> Pour mieux comprendre la distinction iconique/plastique, nous pouvons nous appuyer sur la distinction pro-pictural/pictural, homologuée elle-même à la distinction faite en sémiotique du cinéma entre le pro-filmique et le filmique : le pro-filmique concerne ce qui est destiné à être filmé, sans « mise en forme » (accessoires, décors, personnes, etc.); il s'agit d'éléments ontologiques, qui ont qualité d'êtres, au sens large du terme. Le filmique concerne la manière de filmer (cadrage, mouvement, effets de pellicule, etc.); il s'agit d'éléments rhétoriques. Ainsi, l'analyse sémantique des images de la base ne décrit que le pro-pictural.

<sup>2</sup> Notons que, contrairement aux œuvres textuelles, il est probablement impossible pour une œuvre picturale de produire des signifiés thématiques sans les corrélés à des signifiés figuratifs.

certaines parties corporelles (dont la reconnaissance est solidement ancrée dans nos codes anthropomorphisants), nous notons sur des lignes analytiques distinctes que la base met à disposition : *femme, nudité, tête, visage, cheveux, yeux* (parce que fermés), *paupière, nez, bras, sein, cuisse, jambe, pied, appui, position couchée, position de profil, coquillage*; ou toute autre partie, essentielle ou accidentelle, ou particularité marquante dans l'œuvre. L'importance d'une entité, de ses propriétés ou de ses parties est toute relative au contexte interne du tableau et au degré de segmentation opérable sur les figures iconiques selon leur traitement sémiotique au sein de l'image. Si le raffinement de l'analyse peut paraître poussé dans certains cas, il peut néanmoins devenir utile pour le chercheur intéressé par exemple à comparer les nez peints par Magritte.

Bref, la *présence remarquable* est le facteur clé de caractérisation des signifiés iconiques et de leurs propriétés. Un autre exemple : pour le portrait en buste d'un homme portant des lunettes, nous notons : *homme, tête, visage, lunettes*. Si les verres de lunettes sont visibles, par des reflets par exemple, nous notons : *verre, reflet*; et ainsi de suite pour certaines autres parties ou éléments que nous pouvons répertorier comme frappants grâce aux interprétants contextuels de l'image.

Les objets peuvent être rendus saillants par des procédés ontologiques (nombre<sup>3</sup>, taille, etc.) : pour un tuba en feu occupant une grande partie de l'espace de l'image, nous consignerons : *tuba, piston, pavillon, tuyau, feu, flamme*. Les procédés rhétoriques, comme la centration ou le gros plan, peuvent être à l'origine de la mise en évidence : pour un mur de briques, les contenus relevés sont : *mur, brique*. La base décrit aussi la présence des entités, parties, matériaux et transformations non prévisibles. Si nous constatons dans l'image une saucisse casquée, nous notons : *saucisse, casque*.

La base ne relève pas les fusions non prévisibles dans la doxa d'un regardant; les éléments fusionnés apparaissent séparément : pour un homme-singe, nous indiquons : *homme, singe*; pour une feuille-arbre : *feuille, arbre*. Mais pour une sirène, nous notons : *sirène, femme, queue de poisson*. Enfin, la base ne décrit pas les métaphores, même entre figures : si un miroir symbolise un tableau, nous notons simplement *miroir*.

### **Dénomination des éléments analytiques**

La dénomination homogène des éléments analytiques est capitale pour la qualité et l'efficacité de la base Magritte : il s'agit de toujours nommer les mêmes entités avec les mêmes termes. Ainsi, nous avons tâché d'éviter autant que possible l'emploi de synonymes (voir ci-dessous).

Pour les choses ou les objets, les substantifs, toujours au singulier (sauf lorsque le singulier n'existe pas en français, par exemple *fiançailles*) sont privilégiés. Si l'emploi d'un substantif n'est pas approprié, dans le cas d'un processus par exemple, ou si nous voulons atteindre une plus grande précision sémantique, un verbe peut exprimer le contenu visé (par exemple, *manger* ne veut pas dire la même chose que *repas*).

Tous les éléments analytiques sont entièrement en minuscules sauf s'il s'agit d'un nom propre (par exemple *Mona Lisa, Napoléon*, etc.) ou s'il s'agit de mots peints par Magritte

---

<sup>3</sup> Même si par exemple une entité est présente en trois occurrences dans une image, elle n'est notée qu'une fois dans l'analyse.

(en ce cas, il importe de respecter scrupuleusement les majuscules, les minuscules et les ponctuations; voir ci-dessous).

L'élément analytique est généralement décrit par un seul mot (substantif ou verbe), c'est-à-dire une lexie simple, par exemple *orange*, *melon*, *bouche*, *feuille*, *marcher*, etc.). Si cela est impossible, ou si nous voulons rendre une plus grande précision selon les besoins de dénomination, nous employons plusieurs mots, c'est-à-dire une lexie composée ou complexe, par exemple *pomme\_de\_terre*, *haut\_de\_forme*, *chapeau\_melon*, *avoir\_un\_objet\_à\_la\_main*, etc. Dans la base, il est indispensable d'unir les mots d'une lexie composée ou complexe par des espaces soulignées pour éviter la séparation des unités dans le lexique de l'analyse. Si *pomme de terre* est inscrit sans espaces soulignées, la base répertorie trois éléments analytiques distincts (*pomme* (1), *de* (1), *terre* (1)) au lieu d'un seul. De plus, il faut, pour les mêmes raisons techniques, substituer les traits d'union par des espaces soulignées.

Les expressions ou groupes de mots ne sont pas utilisés pour caractériser les entités vues, par des adjectifs ou des compléments du nom par exemple. Voici ce qui a été évité : *grand\_oiseau*, *homme\_avec\_chapeau*, etc. Par exemple, pour un tuba en flammes, nous ne pouvons écrire *tuba\_en\_flammes*, parce que *tuba\_en\_flammes* n'est pas une expression autorisable. Nous décrivons alors la chose en la décomposant en deux expressions ou mots autorisés (*feu*, *flamme* et *tuba*). La base ne décrit pas les propriétés sous les appellations du type *homme\_fort*, *tour\_de\_cristal*, *mur\_de\_briques*, etc. Si les propriétés sont elles-mêmes des entités saillantes, elles sont notées séparément sur des lignes analytiques distinctes (entité 1 : *tour*; entité 2 : *cristal*).

Il arrive que, faute de trouver mieux comme dénomination, un mot soit employé dans un sens inexact ou légèrement différent de celui que lui prête un dictionnaire de langue. C'est le cas de *calvitie*, qui signifie en langue « perte, partielle ou totale, involontaire et biologique des cheveux », mais qui est utilisé pour les personnes « chauves » chez Magritte. Pour marquer ces « néologismes de sens », nous leur avons ajouté à la fin le tilde (~) : *calvitie~*.

### **Le traitement des synonymes**

La base Magritte n'utilise pas de synonymes d'une œuvre à l'autre (sauf erreur d'homogénéité) et au sein d'une même œuvre. Ainsi, *montgolfière* sera toujours préféré à *ballon* (celui-ci étant d'ailleurs trop général quand l'entité peut être identifiée avec certitude comme une montgolfière). Ces synonymes faciliteraient le repérage par l'utilisateur, en autant qu'ils soient utilisés de manière conjuguée chaque fois et non en alternance; mais ils fausseraient le décompte en créant deux entités au lieu d'une seule.

### **Le traitement des homonymes et des polysèmes**

L'homonymie (lorsqu'il y a au moins deux unités lexicales différentes ayant chacune un sens distinct) et la polysémie (lorsqu'il y a une seule unité lexicale à laquelle correspondent différents sens) sont avérées ou potentielles entre éléments analytiques : *bouche* [*de métro* / *humaine*], *agenouillement* [*état* / *processus*], *baleine* [*animal* / *de parapluie*] etc.

Pour la désambiguïsation des homonymes et des polysèmes, nous prenons la première des stratégies désambiguïsantes suivantes qui fonctionne :

- Passage à la lexie complexe (en supprimant les mots-outils : « de » dans l'exemple) : *feuille\_(papier)* versus *feuille\_(végétal)*. Nous évitons les descriptions non lexies : *\*bouche\_(humaine)* (nous ne mettons pas *humaine* pour les parties de corps humain) versus *bouche\_(métro)*.
- Indication d'une classe : *melon\_(chapeau)* versus *melon\_(fruit)*. Nous privilégions la classe la plus efficace (petite ou grande) : *melon\_(chapeau)* et non *melon\_(vêtement)* ou *melon\_(culturel)*, mais *loup\_(animal)* et non *loup\_(canidé)*.
- Indication d'un synonyme : *bulle\_(phylactère)*. Nous choisissons le mot le plus commun, ici *bulle*, comme terme désambiguïsé par le terme le moins commun.

Les lexies désambiguïsantes procèdent ainsi du rapport élément-classe (*loup\_(animal)*, *bassin\_(anatomie)*), du rapport partie-tout (*aile\_(oiseau)*, *aile\_(avion)*), du rapport tout-matière (*feuille\_(papier)*) et d'autres rapports possibles<sup>4</sup>. La forme désambiguïsante entre parenthèses correspond souvent à une classe (et en l'absence de mots-outils rien ne l'en distingue), mais dans *aile\_(oiseau)*, par exemple, *oiseau* peut être considéré par l'utilisateur soit comme une classe, soit comme un tout (dont l'aile est une partie).

La désambiguïsation par élément-classe ou par partie-tout pourrait se faire à différents niveaux : *aile\_(oiseau)*, *aile\_(rapace)*, *aile\_(rapace\_nocturne)*, *aile\_(hibou)* ou *aile\_(grand\_duc)*. Pour éviter la multiplication des entrées (*aile\_(hibou)*, *aile\_(aigle)*, etc.), nous nous sommes placés à un degré de généralité élevé : *aile\_(oiseau)* / *aile\_(avion)*. *oiseau* et *avion* peuvent également être considérés comme des parties de lexies : *aile\_d'oiseau* / *d'avion*.

Enfin, si deux homonymes ont le même poids, chaque élément est désambiguïsé : *feuille\_(papier)* / *feuille\_(végétal)*. Si un des homonymes est plus prégnant, en principe seul l'homonyme marqué est désambiguïsé : *loup* / *loup\_(masque)*.

### Catégorisation et lexicalisation des éléments analytiques

Selon Hébert (2007 : 193), il faut distinguer « le problème de la catégorisation de ce qui est représenté dans une image de celui de sa lexicalisation. La catégorisation est l'opération intégrant une unité dans une classe ontologique, une classe d'êtres. [...] La lexicalisation est l'opération consistant à nommer une catégorie ou un représentant de cette catégorie. »

Les catégorisations d'éléments analytiques de la base *Magritte* sont généralement communes aux regardants, sociolectales. Mais certaines catégorisations peuvent être idiolectales, propres à Magritte. Le peintre appelle *grelot* une forme très stylisée de grelot qui pourrait n'être considéré que comme une simple *sphère*. Bref, la catégorie (*grelot*) existe sociolectalement, mais le type visuel que lui fait correspondre Magritte est propre à lui. Nous pouvons aussi considérer que Magritte dénomme une nouvelle classe d'objets parallèle par un terme déjà existant ou qu'il déforme la classe sociolectale en l'élargissant. D'autres catégories sont propres à Magritte mais non lexicalisées par l'artiste ou par les magrittiens (sauf erreur) : *fond\_noir* (par exemple dans *La lunette*

<sup>4</sup> Il ne serait pas utile d'énumérer tous les rapports utilisés dans les stratégies désambiguïsantes : l'important est de comprendre le principe général.

*d'approche*), *fond\_blanc*, *surdimension*, *sous\_dimension*, *forme biomorphe*, etc. Ces importantes catégories non lexicalisées sont bien attestées dans le corpus magrittien et doivent être considérées, donc lexicalisées de manière *ad hoc*.

L'analyse de la base de données aurait tort de ne retenir que les contenus iconiques parfaitement lexicalisables. Certains sont parfaitement lexicalisables, par des lexies courantes (*orange*, *pipe*, *chapeau\_melon*, etc.) ou par des lexies rares (*haut\_de\_guêtre*, *muleta*, etc.). Certaines figures sont difficilement identifiables et/ou lexicalisables : nous les avons désignées par le descripteur *innommable*.

Certaines figures existent et sont dénommées, mais l'analyste ignore leur nom ou leur nom précis : par exemple, dans *La clef des songes*, l'analyste peut indiquer *marteau* alors qu'il faudrait noter *marteau\_à\_main* (petit marteau pour forger). Ici le problème de la lexicalisation rejoint celui de la catégorisation : si nous ne reconnaissons pas la catégorie, nous ne pouvons la dénommer correctement. Par ailleurs, *marteau\_à\_main* dépasse sans doute le degré de précision attendu.

Dans la base *Magritte*, nous avons voulu néanmoins atteindre la plus grande précision attendue (et non possible) dans la catégorisation de l'élément. Voici quelques exemples de précision attendue : *chapeau\_melon* au lieu de *chapeau*, *rose\_(fleur)* au lieu de *fleur*, et de précision parfois possible mais non attendue : *chapeau\_melon\_des\_années\_30*, *rose\_anglaise*. Parfois, le meilleur de degré de précision attendu est rendu dans l'œuvre (*chêne* au lieu de *feuillu*, parce que l'entité est reconnaissable et aisément catégorisée), parfois non (*feuillu* au lieu de *chêne*). Parfois, il est rendu sur un mode incertain (*chêne?*, parce que l'image ne peut nous en donner la certitude, mais que cet arbre ressemble plutôt à un chêne).

### **L'emploi du point d'interrogation**

Nous avons placé un point d'interrogation à la fin de certains éléments analytiques afin d'indiquer : (1) un doute quant à la véritable nature, à la catégorisation d'une entité (mode dubitatif pour traduire une difficulté d'identification due à la taille, à un traitement plastique flou, à la polysémie visuelle, etc.) : *fleur?*, *homme?*; (2) un élément correspondant partiellement à un élément mondain : *avion?* pour les étranges aéronefs de l'œuvre intitulée *Le drapeau noir*. À l'origine, nous avons utilisé le modalisateur *pseudo\_* pour ce type d'élément (*pseudo\_avion* au lieu de *avion?*).

### **Image dans une image**

Si une image est incluse dans un tableau, nous avons essayé de donner non seulement la description du contenu (*montagne*, *arbre*, *ciel*, etc.) mais aussi le genre (*paysage*, *portrait*, etc.). Par contre, nous ne nommons pas le genre de l'œuvre analysée.

### **Mots peints dans les images**

Occasionnellement, les œuvres contiennent des mots isolés et des groupes de mots (expressions, syntagmes et phrases). Sauf la signature de Magritte, ils sont considérés comme des entités et sont notés comme tels dans le lexique analytique des images. Un mot isolé est d'abord placé sur une ligne analytique entre guillemets anglais doubles ("*montagne*") et, sur une ligne différente, le mot est repris entre paire de guillemets doubles anglais (""*montagne*"""). Un groupe de mots est noté entre guillemets doubles

anglais ("*Ceci\_n'est\_pas\_une\_pomme*") et ses mots sont obligatoirement unis par des espaces soulignées pour éviter la scission du groupe de mots dans le lexique; les mots composant ces groupes sont notés entre paires de guillemets doubles anglais ("*Ceci*", "*pas*", "*pomme*", etc.). La ponctuation, les majuscules et les minuscules doivent être respectées. Une apostrophe doit être « collée » au mot qui la suit, pour éviter la coupure du groupe de mots dans le lexique.

L'implantation d'un module de reconnaissance de ce type particulier d'objets inscrits dans les images magrittiennes permettrait de générer un lexique distinct des mots et des phrases contenus dans les images.

Mars 2011

### **Références**

HÉBERT, Louis (2007). *Dispositifs pour l'analyse des textes et des images. Introduction à la sémiotique appliquée*, coll. « Nouveaux actes sémiotiques », Limoges, Presses universitaires de Limoges.

GREIMAS, Algirdas Julien et Joseph COURTÉS (1993). *Sémiotique. Dictionnaire raisonné de la théorie du langage*, Paris, Hachette Supérieur.

GROUPE  $\mu$  (1992). *Traité du signe visuel. Pour une rhétorique de l'image*, coll. « La couleur des idées », Paris, Seuil.