

**« LES ICÔNES, C'EST DU CHINOIS ! »**

**SEMIOTIQUE ET COGNITION DANS L'IDENTIFICATION D'ICÔNES**

Philippe Dessus\* & Daniel Peraya\*\*

\* Université de Grenoble & IUFM

Philippe.Dessus@upmf-grenoble.fr

\*\* Université de Genève, Suisse

Daniel.Peraya@tecfa.unige.ch

TITRE COURANT : Les connaissances dans l'identification d'icônes

**Introduction**

Les interfaces homme-machine font actuellement un grand usage d'icônes. Ce sont des représentations imagées d'objets censées être plus proches des modèles mentaux des utilisateurs que les interfaces à ligne de commande, donc plus aisés à reconnaître. Cependant certaines inférences sont nécessaires pour que l'utilisateur arrive à identifier la fonction du logiciel représentée par l'icône à partir, notamment, du sens descriptif de celle-ci. L'utilisateur doit d'abord donner un sens à la représentation, autrement dit lui attribuer un sens descriptif. Ensuite seulement, il pourra, par un processus d'inférence, lui attribuer la fonctionnalité du logiciel qu'elle est censée désigner. En effet, dans la plupart des cas, cette fonctionnalité n'est pas immédiatement identifiable par l'utilisateur. Pour donner un exemple, l'icône correspondant à la fonction « imprimer » n'est compréhensible qu'à partir d'un processus assimilable à un trope, ou à l'usage d'un mot dans un sens figuré, déplacement sémantique qui correspond, dans cet exemple, à un processus de type métonymique. Le problème apparaît donc complexe : il s'agit, d'une part, d'« élucider » la fonction désignative de la représentation et, d'autre part, d'inférer, à partir de là, la fonctionnalité visée par le concepteur du logiciel. Ce second processus pourrait s'apparenter à celui décrit pour la langue par la pragmatique en terme d'acte de discours indirect ou parfois de trope illocutoire (e.g., « pouvez-vous me passer le sel ? »).

La métonymie est une figure rhétorique générale qui, on le sait, consiste en un transfert de dénomination entre deux notions, l'une étant liée à l'autre par une relation obligée (par exemple cause et effet, contenu/contenant, partie pour le tout, etc.). Il s'agit, dans notre exemple, de désigner l'action (imprimer) par la représentation de l'objet technique qui permet cette action (une imprimante). Bien plus que la métaphore, la métonymie semble être un

procédé largement utilisé dans les systèmes d'icônes de logiciels et d'environnements informatiques standardisés (ILEIS, voir Peraya, 1998) : la gomme (l'objet) renvoie à la fonction « effacer » (ce que permet l'objet), la disquette est utilisée pour représenter tous les supports de stockage et le support de stockage pour désigner la fonction « sauvegarder », etc. On peut donc penser que ce procédé s'est progressivement stabilisé et standardisé à la manière des actes de discours indirect conventionnel comme « Est-ce que vous avez l'heure ? », qui constitue une façon habituelle de demander l'heure qu'il est. Mais toutes les inférences ne se basent pas sur une telle régularité, tout au contraire.

L'identification et la compréhension des icônes, comme celles de toute représentation iconique, sont au centre des questions relatives au rapport entre les représentations matérielles ou sémiotiques et les représentations mentales. Mais en même temps, c'est le rapport entre les registres de représentation différents – linguistique et iconique – qui se trouve questionné, le sens d'une image étant souvent référé à l'interprétation linguistique de celle-ci (Barthes, 1964 ; Eco, 1970). C'est d'ailleurs pour faciliter la compréhension des icônes et l'identification de leur fonctionnalité correspondante que les concepteurs ont conçu les info-bulles, ces courts textes d'aide qui apparaissent lorsque la souris reste un court délai au-dessus d'une icône. L'info-bulle assure ici le rôle essentiel d'ancrage (Barthes, 1964) du sens et de l'interprétation, réduisant grandement la polysémie propre à l'image. Toutefois, dans le cadre de la sémiotique classique, le processus de passage du registre de représentation iconique ou registre de représentation linguistique ne fait l'objet d'aucune analyse. La notion de conversion entre des registres de représentations telle que l'a développée Duval (1995, 1999 et 2006) constitue aujourd'hui une voie féconde pour aborder ces problèmes du point de vue théorique, mais aussi de celui de leur application, par exemple, aux rapports entre les info-bulles et les icônes correspondantes. Reconnaisant ainsi l'importance du langage verbal et des info-bulles dans les processus de compréhension et d'identification de la fonction des icônes, nous serons donc amenés, dans l'étude que nous présentons ici, à analyser ces processus aussi sur une base linguistique, et en conséquence sur celle de la description linguistique des icônes. C'est dans ce contexte que la notion de conversion nous sera particulièrement utile.

Ce chapitre a pour objet les processus d'identification à l'œuvre dans l'utilisation d'icônes de logiciels, et nous apporterons quelques réponses à des questions qui nous paraissent importantes. L'utilisation d'icônes dans les interfaces améliore-t-elle vraiment l'accès à leur fonctionnalité ? En particulier, les info-bulles sont-elles réellement un moyen supplémentaire

d'en faciliter l'accès ? Quel est l'effet des connaissances des utilisateurs sur la compréhension d'icônes ? Est-il possible de simuler l'utilisation de ces connaissances dans ce processus ? Nous passerons en revue, dans ce chapitre, deux thèmes centraux liés à la compréhension des icônes : qu'apporte l'approche sémiotique à la formulation et à la compréhension des ces problèmes ? Quel est le rôle de la connaissance et quels sont les effets d'artifices additionnels tels que les info-bulles dans le processus d'identification des icônes ?

### **De l'image à la fonctionnalité, comment identifie-t-on une icône ?**

Détaillons maintenant les deux moments de l'identification. Pour commencer, l'élucidation de la fonction désignative de la représentation révèle déjà un certain nombre de difficultés. Les premières relèvent du type de représentations visuelles auquel s'apparentent les icônes. Celles-ci ne sont ni des photographies, ni des dessins réalistes ; ce sont au contraire des dessins stylisés, schématisés qui reprennent en conséquence certains traits caractéristiques de l'objet. On peut donc imaginer un ensemble d'icônes « synonymes », chacune correspondant à la représentation personnelle de chacun des graphistes l'ayant dessinée. La variété des icônes représentant une imprimante pour la même fonctionnalité « Imprimer le document en cours », présente dans la quasi-totalité des logiciels, en constitue une démonstration exemplaire (*voir figure 1*).



Figure 1 – Quelques icônes de la fonctionnalité « Imprimer » (source <http://www.fresherimage.com/Icons/print/>)

Mais, à ce stade, rien ne garantit encore que l'utilisateur inférera de cette icône la fonctionnalité appropriée : toute image, de nature analogique, est polysémique et suscite donc de multiples interprétations. Pour prendre un autre exemple simple, l'icône représentant « une feuille de papier avec un coin plié », même très fidèlement représentée et reconnue comme telle par l'utilisateur, ne le conduira pas nécessairement à sa fonctionnalité véritable

(« Nouveau document »). En effet, cette dernière pourrait tout aussi bien être : « Effacer tout le document », « Mise en page », « Nouvelle page ». La monosémie des représentations visuelles se fonde sur l'hypothèse selon laquelle la compréhension consisterait en une symétrie des opérations de codage et de décodage. Or, celle-ci a été invalidée par l'expérience quotidienne et, depuis fort longtemps, par la pratique pédagogique – notamment en situation interculturelle – autant que par la littérature scientifique (par exemple Almassy, 1974 ; Barthes, 1964 ; Coste, 1975 ; Erni, 1986 ; Lefebvre-Mignot, 1979 ; Moles, 1981).

La compréhension du sens se construit, certes à partir d'indices repérables dans la représentation, mais principalement sur la base de processus inférentiels. C'est donc en s'appuyant sur ceux-ci que l'utilisateur pourra découvrir la fonctionnalité à partir de la représentation de l'objet. Dans la section suivante, nous allons décrire comment la recherche sur la sémiotique s'est progressivement intéressée à caractériser ces processus inférentiels.

### **Identification d'icônes et approche sémiotique : de la perception à la cognition**

Les icônes, parce qu'elles sont des représentations imagées, relèvent de la sémiotique, et plus particulièrement de la sémiotique visuelle et de l'étude des langages iconiques. L'hypothèse générale qui en sous-tend leur usage, notamment dans les interfaces informatiques, est qu'un marquage visuel constitue une aide efficace à l'identification, la mémorisation et à la recherche de l'information. Retraçons brièvement les quelques étapes de l'évolution de la sémiotique visuelle qui ont profondément transformé la problématique de la compréhension des images, et donc des icônes.

#### **Le signe iconique et l'iconicité**

L'évolution de la sémiotique visuelle pourrait donc être vue comme l'évolution d'une certaine conception du rapport entre le signe iconique et le réel ou plus généralement entre *semiosis* et pensée humaine, donc entre *semiosis* et *noesis* (Duval, 1995). La sémiologie s'intéresse très tôt à l'« image » et aux langages analogiques (Barthes, 1964 ; Metz, 1964), mais les modèles classiques du signe et du code linguistiques ne peuvent rendre compte des langages visuels. En effet, les opérations de décodage sur lesquelles s'est construite à l'époque l'approche structurale du sens et de la signification, se heurtent à l'analogie de l'image et à l'impossibilité de la décomposer en unités discrètes. La photographie par exemple, « en raison de sa nature absolument analogique, semble bien constituer un langage sans code. » (Barthes, 1964, p. 46). On privilégiera donc l'interprétation linguistique de l'image pour lui donner sens (*cf.*, ci-dessus, la fonction d'ancrage).

Dans leur volonté de trouver une solution théorique au problème de l'analogie, les chercheurs se tournent alors vers la distinction de Peirce (1931/1938) entre indice, icône et symbole. Pour cet auteur, un signe analogique (l'icône) possède certaines propriétés de l'objet représenté, ou possède une certaine ressemblance avec l'objet auquel il se réfère. Eco (1970) a longuement analysé la définition de l'icône chez cet auteur et en a proposé une formulation nouvelle, basée tant sur l'expérience perceptive de l'objet que sur son expérience réelle : les signes iconiques « ne possèdent pas les propriétés de l'objet représenté, mais reproduisent quelques conditions de la perception commune, sur la base de codes perceptifs normaux et en sélectionnant des stimuli qui – d'autres stimuli étant éliminés – peuvent me permettre de construire une structure perceptive qui possède – par rapport aux codes de l'expérience acquise – la même signification que l'expérience réelle dénotée par le signe iconique. » (Eco, 1970, p. 14). Enfin, l'iconicité est pour Eco une question de degrés, ce qui suppose donc une analyse des différents types de représentations visuelles et l'impossibilité des les traiter toutes de la même façon.

Progressivement, le degré d'« iconicité », ou la « quantité de réalisme [...] contenu dans la représentation » selon Moles (1981, p. 102) – entre le représentant (l'icône) et le représenté (dans notre cas, la fonctionnalité) deviendra le principal critère de classification des représentations, dont la plus achevée est l'échelle d'iconicité en 12 degrés proposée par Moles (1981). Cependant, au-delà de son indéniable valeur heuristique, cette définition ne paraît guère satisfaisante du point de vue théorique, car elle fonde clairement l'iconicité dans le rapport entre la représentation et l'objet lui-même. Elle s'inscrit donc dans le courant de la sémantique extensionnelle « qui postule une correspondance entre les signifiés et les éléments du réel » (Berendonner, 1991, p. 583), et réintroduit inévitablement le réel dans le processus de sémiologie, dont il devait par définition être exclu.

### **Une approche cognitive du signe analogique**

L'enjeu du débat autour de l'iconicité renvoie à la nature du signe sémiotique en général et, plus particulièrement, aux statuts du référent d'une part, du signifié d'autre part, dans le modèle de triangle sémiotique<sup>1</sup>. La première question porte sur les rapports entre le langage et la

---

<sup>1</sup> Ce modèle a connu de nombreuses variantes selon les auteurs : par exemple, symbole, référence, référent (Ogden & Richard, 1923) ; signe, *significatum/designatum* et *denotatum* (Morris, 1946) ; mot, sens et chose (Ullman, 1957) ; signifiant, signifié/concept, référent (Saussure, 1972). Ces différents modèles ont été longuement commentés par Eco (1988) et synthétisés par Hébert (2006).

réalité et nous venons de voir combien cette question se pose avec acuité à propos du signe analogique. La seconde, qui divise elle aussi les linguistes, concerne la distinction entre concept et signifié : peut-on les assimiler jusqu'à les confondre en une seule catégorie ? Doit-on les distinguer ?

Dans le domaine de la sémiotique visuelle, il faudra attendre les années 1990 pour voir se développer, de notre point de vue, une reformulation du triangle sémiotique compatible avec l'exigence de l'extériorité du référent et des travaux qui se développent alors sur les images et de modèles mentaux (Denis, 1979, 1989 ; Johnson-Laird, 1983, 1988 ; Langacker, 1987). Les chercheurs du Groupe  $\mu$  (1992) s'attachent à établir les fondements perceptifs d'une sémiotique visuelle, autrement dit à analyser le processus sensation *versus* perception *versus* cognition. Si la position « perceptiviste » de Eco (1970) constitue une étape importante dans cette évolution, nous pensons que, dans le contexte structuraliste largement dominant à l'époque, cette ouverture vers les modèles perceptifs et la modélisation de l'expérience acquise de l'objet ne pouvait être entendue.

Pour les membres du Groupe  $\mu$ , la relation analogique ne peut être modifiée si la conception du signe demeure inchangée. Ils abandonnent donc la représentation bipartite saussurienne et proposent une reformulation du triangle sémiotique composé du signifiant, du référent et du type (voir figure 2). Le référent est « un *designatum* (et non le *denotatum*<sup>2</sup> par définition extérieur à la *sémiosis*), mais un *designatum* actualisé ». En d'autres termes, c'est l'objet déjà entendu comme membre d'une classe validée par le type, qui est, quant à lui, « un modèle intériorisé et stabilisé qui, confronté avec le produit de la perception, est à la base du processus cognitif » (*id.*, p. 137). Le signifiant enfin, est un « ensemble modélisé de stimuli visuels correspondant à un type stable, identifié grâce à des traits de ce signifiant, et qui peut être associé à un référent reconnu [...] Il entretient avec ce référent des relations de transformations. » (*ibid.*). La fonction du type, dont la définition est proche de celle du modèle mental (Denis, 1989 ; Denis & de Vega, 1992), est alors de garantir l'équivalence (ou « identité transformée ») du référent et du signifiant visuel.

---

<sup>2</sup> Les auteurs font référence au *denotatum* dans le triangle sémiotique classique, tel que celui de Morris.

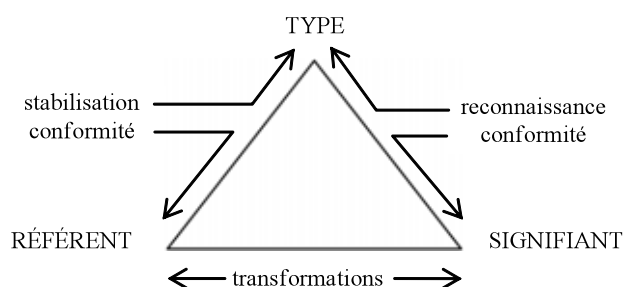


Figure 2 – Le signe iconique (Groupe  $\mu$ , 1992, p. 136).

Si l'on cherche à appliquer ce modèle à l'exemple initial, celui de l'icône de l'imprimante évoqué ci-dessus, on conviendra que la figure 1 rassemble différentes représentations visuelles – autrement dit, des « signifiants » différents – d'un « type unique », c'est-à-dire de l'image mentale de l'imprimante. Cette image mentale, s'est construite progressivement au cours de nos interactions avec l'objet réel – l'objet « imprimante » – qui progressivement s'est constitué en un référent, un *designatum*, partie intégrante du processus de sémiotisation.

### **Similis et schémas chez Darras**

Pour le Groupe  $\mu$ , les mécanismes cognitifs sont donc constitutifs du signe iconique et indissociables de sa définition. Darras (1996) prolonge cette approche et propose de considérer les mécanismes cognitifs et les processus de catégorisation conceptuelle comme la base de distinction entre les différentes représentations visuelles. L'auteur, qui s'est longtemps intéressé aux dessins d'enfants, à l'éducation aux médias et à la formation des enseignants dans ce domaine, découvre chez les adultes une façon de dessiner qui rappelle à s'y méprendre – ce que confirment d'ailleurs des épreuves de classement effectuées par des juges externes – celle des dessins d'enfants. Il appellera « imagerie initiale » cette façon de dessiner que partage la majorité des adultes et des enfants. Selon lui, l'imagerie initiale, qui comme la photographie traite du visible, répond pourtant à une logique toute différente :

Toutefois, l'immense différence provient du fait que les figures de l'imagerie initiale ne sont pas des projections optiques mais avant tout des projections de l'esprit. [...] l'acte photographique privilégie les aspects du vu et du visible alors que cette préoccupation est très marginale en imagerie initiale. (Darras, 1996, p. 36).

C'est donc bien la place et l'importance relative des processus cognitifs par opposition aux processus perceptifs qui fonderait la différence entre ces deux types de représentations visuelles, mais aussi de façon plus générale entre les différentes sortes d'imagerie. L'auteur y reconnaîtra deux grandes familles principales : les *similis* (issus de la pensée visuelle) et les

schémas (issus de la pensée figurative). Les premiers ont pour caractéristique principale de simuler les phénomènes (pouvant être indexés par gradient de réalisme) tandis que les seconds opèrent par simplification et abstraction. Les *similis* sont des représentations visuelles qui ressemblent à ce qu'elles représentent : elles s'enracinent dans le réalisme et dans les codes qui permettent de singer, de mimer l'objet (*e.g.*, photographies). Les *schémas*, au contraire, obéissent à une économie non plus perceptive, mais cognitive basée sur les notions de prototype et de typicalité empruntées respectivement à Cordier (1993) et à Rosch (1978). Selon Darras, les schémas sont produits à partir d'une reconfiguration de notre matériel cognitif :

La procédure graphique fonctionnerait par convocation, prélèvement, transfert, conversion et assemblage des traits et des informations figuratives de l'entité cognitive qui constitue le résumé cognitif. (*op. cit.*, p. 105).

Enfin, dans les caractéristiques propres du résumé cognitif, Darras mentionnera la consensualité, « propriété statistique et pragmatique des catégories naturelles » (p. 114), trait essentiel puisqu'il implique que les icônes constituent des représentations partagées par les utilisateurs, autrement dit, à propos desquelles existe un consensus parmi ces derniers.

### **Duval et les processus de transformation et de conversion**

Enfin, nous évoquerons plus succinctement un des aspects des travaux de Duval (1995, 1999, 2006), qui identifie ensuite deux points importants dans le traitement des représentations, qu'elles soient visuelles ou non. Le premier est la différence entre deux processus principaux : la transformation et la conversion. La *transformation* est un processus interne au registre sémiotique considéré qui permet de produire de nouvelles significations, de nouveaux messages, dans le même système (*e.g.*, calculer). D'autre part, la *conversion* est une transformation entre deux registres sémiotiques. Duval explicite clairement cette différence à propos des différents registres de représentations mathématiques.

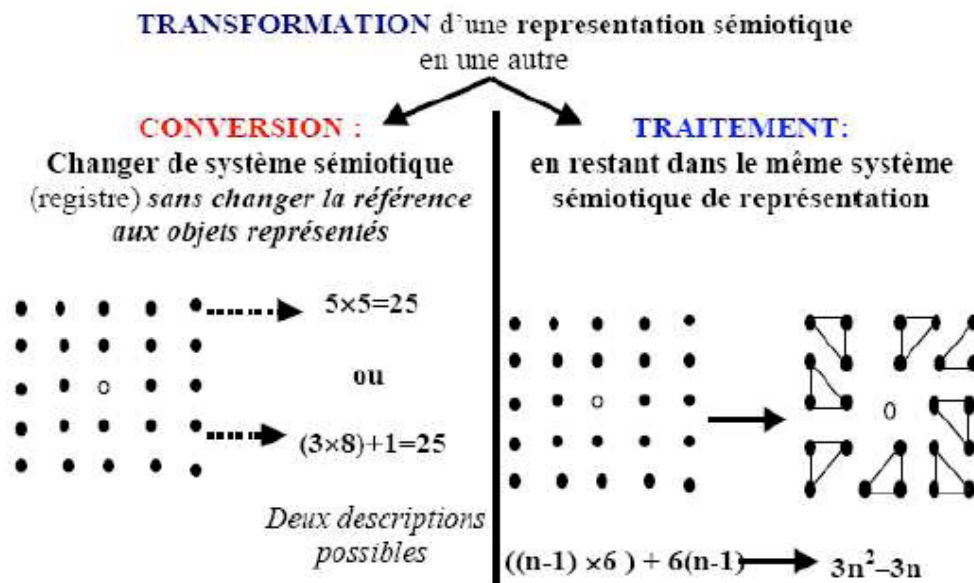


Figure 3 – Différence entre transformation et conversion de représentations « cognitivement hétérogènes », selon Duval (2006, p. 12).

## Discussion

Nous retiendrons de l'ensemble de ces cadres de référence brièvement évoqués trois points principaux. Tout d'abord, que certaines représentations visuelles, les *schémas*, s'apparentent à des résumés cognitifs et sont construites à partir d'une réorganisation de notre matériel cognitif. D'autres, comme les *similis*, en revanche, sont plus proches d'événements perceptifs et demandent donc des processus de traitement différents. Qu'en est-il pour les icônes d'ordinateur ? Ensuite, le processus de conversion, ou passage d'un registre sémiotique à un autre, paraît central : identifier une icône, ce n'est pas seulement la reconnaître, c'est arriver à la convertir dans un système sémiotique différent, car la description du fonctionnement du logiciel fait nécessairement appel à un autre système sémiotique que celui des icônes. Enfin, nous garderons la notion de programme figuratif attaché aux résumés cognitifs et la nécessité d'un consensus, de la part du groupe ou de la communauté, pour valider et le résumé cognitif et sa représentation, point de vue par ailleurs assez conforme à l'interactionnisme social (Bronckart, 1996) et aux courants sociocognitivistes.

## Les effets du contexte (connaissance) dans l'identification des icônes

Le contexte dans lequel le message est émis influe sur la construction, par le lecteur, de la signification du message (*i.e.*, de l'icône). Travailler dans un logiciel de mise en page ou un logiciel de CAO induit un certain cadre d'interprétation (Goffman, 1991) des icônes, lié à l'environnement et aux tâches que l'on y mène. De plus, les connaissances mêmes de

l'utilisateur peuvent être considérées comme faisant partie du contexte (Scalisi, 2001) ou des encyclopédies d'interprétation (Klinkenberg, 1996). Nous nous centrerons, dans cette section, tout d'abord sur les formes et la nature des catégories de ces connaissances. Ensuite, nous décrirons une manière d'influer sur le contexte dans les récentes interfaces d'ordinateurs, concevoir des info-bulles.

### **Les types de connaissances impliquées dans l'identification d'icônes**

L'utilisation de logiciels nécessite de nombreuses catégories de connaissances différentes, qu'Hammond et Barnard (1984) ont classées dans les huit catégories suivantes :

- connaissances de la langue naturelle,
- connaissances du domaine, ainsi que du domaine tel que représenté dans le système,
- connaissances des opérations du système,
- connaissances du domaine tel dans un environnement de bureau, sans informatique,
- connaissances du problème traité,
- connaissances de l'interface,
- connaissances des autres machines et des procédures,
- connaissances de l'interface de dialogue.

Nous avons souligné, dans un autre article (Dessus & Peraya, 2005), le principal problème lié à l'utilisation des connaissances pour l'identification d'icônes. Les connaissances directement liées au système sont bien plus aisées à prendre en compte que celles liées à la connaissance de la langue naturelle, voire à des connaissances encyclopédiques (Moyes et Jordan, 1993 ; Klinkenberg, 1996), pourtant essentielles : les multiples variantes de l'icône correspondant à la fonctionnalité « imprimer » (*voir la figure 1*) s'expliquent en partie par le fait que des utilisateurs différents peuvent avoir une représentation différente de l'objet « imprimante ».

### **Le rôle des info-bulles dans l'identification des icônes**

L'identification des icônes graphiques des interfaces *WIMP* (*Windows, icon, menu, pointing device*, ou fenêtres, icônes, menus et souris) ne doit pas constituer un processus à ce point évident, puisque leurs concepteurs ont conçu les info-bulles, courts textes d'aide apparaissant lorsque la souris reste un court délai au-dessus d'une icône. Les études sur cette nouvelle aide sont encore rares, et Richardson (2000) a montré, sans surprise, que la sollicitation de ces info-bulles était liée au niveau d'expertise du logiciel, mais aussi au degré de ressemblance de

l'icône avec la fonctionnalité. Toutefois, des travaux supplémentaires restent à réaliser, car les connaissances encyclopédiques n'ont pas été manipulées dans cette étude. Le rôle des info-bulles ne peut être envisagé sans référence aux difficultés du passage d'un registre sémiotique – l'iconique – à un autre registre, le verbal, même si globalement l'info-bulle assure une fonction d'ancrage du sens (Barthes, 1964). Nous pouvons penser que ces info-bulles, en présentant une information verbale uniquement centrée sur le fonctionnement de l'ordinateur – et donc jamais sur l'icône elle-même –, facilitent justement le processus de conversion icône-fonction.

\*

La revue de la littérature ci-dessus montre quelques problèmes importants. Comment les icônes et leurs fonctionnalités correspondantes sont représentées mentalement et traitées pour être appariées ? Quel est le rôle des info-bulles dans ce processus ? Quels processus de conversion permettent de passer d'un registre sémiotique à l'autre (*e.g.*, de l'icône au texte de l'info-bulle, cette dernière étant la description linguistique de la fonctionnalité) ? Quel est le rôle des connaissances de l'utilisateur ? Peut-on simuler de tels processus ? Les recherches respectives en intelligence artificielle et en sémiotique ont suivi des chemins très différents, malgré leurs motivations communes, et leurs modèles respectifs sont également très différents : les premiers sont plus formels et plus définis, les seconds plus intuitifs et généraux (Gudwin, 1999).

Nous avons, quant à nous, conçu un système qui tente d'articuler les deux démarches précédentes et de répondre à la question importante suivante : comment le « sens » d'une icône (*i.e.*, sa fonctionnalité) est décodé par l'utilisateur, selon son niveau de connaissances ? Pour cela, nous partons d'une méthode d'analyse de grands corpus textuels, censés représenter les connaissances encyclopédiques et techniques des utilisateurs. Nous faisons de plus le postulat, maintenant largement admis (McNamara, 1999), que les représentations imagées (« de surface ») sont vraisemblablement réinterprétées et recodées en un format propositionnel, primitif et « profond », avant qu'une correspondance avec la fonctionnalité du logiciel puisse être trouvée (Kosslyn, 1981). Nous présenterons plus bas (*cf.* section « Méthode de conversion image vers texte ») une manière systématique de convertir le sens de l'image des icônes dans un format textuel, ce qui, ensuite, permet de pouvoir le comparer avec des fonctionnalités possibles (*e.g.*, issues des info-bulles).

**Un outil modélisant les connaissances mises en œuvre dans l'identification d'icônes**

## Description du dispositif informatique

Nous avons simulé, en utilisant un logiciel d'analyse statistique de grands corpus textuels (analyse de la sémantique latente, Landauer & Dumais, 1997, voir Lemaire & Dessus, 2003, pour une présentation en français), le processus d'appariement entre l'objet représenté par l'icône et la fonction de cette dernière (voir Dessus & Peraya, 2005, pour une présentation complète de cette étude). Le mode de fonctionnement de LSA suppose que le sens d'un paragraphe (ou, dans notre cas, la description de l'icône) est défini à partir de l'ensemble des mots (*i.e.*, des éléments de l'icône) qui le composent. Inversement, le sens d'un mot (*i.e.*, un élément de l'icône) est défini à partir de l'ensemble des paragraphes (*i.e.*, description de l'icône) dans lesquels il apparaît. Cette double récursion est résolue par un traitement statistique proche d'une analyse factorielle. Pour modéliser les connaissances humaines, nous avons utilisé trois types de corpus, tout d'abord séparément, puis par combinaisons :

1. année 1999 du quotidien *Le Monde* (149 Mo, représentant des connaissances générales) ;
2. aide en ligne de *Word* (intégralité de l'aide en ligne du logiciel, 1,5 Mo, représentant des connaissances particulières sur le logiciel) ;
3. TP sémiotique (analyses sémiotiques d'icônes par des étudiants, 200 ko, représentant des connaissances en sémiotique).

## Hypothèses

Cinq hypothèses ont été testées dans cette étude. Les quatre premières sont reprises de l'étude *princeps* décrite en détail dans Dessus et Peraya (2005), la dernière est originale et réalisée à des fins de contrôle.

*Hypothèse 1 : Effet des corpus de connaissances sur la performance de LSA en identification d'icônes.* Plus les corpus font mention de connaissances spécialisées, plus les performances de LSA seront élevées et nous devrions observer de plus un effet cumulatif de ces connaissances (l'ajout de tous les corpus devrait donner des performances maximales).

*Hypothèse 2 : Effet des indications des info-bulles sur la performance de LSA en identification d'icônes.* Ajouter une information textuelle (*i.e.*, le contenu de l'info-bulle) devrait systématiquement augmenter le taux d'appariement icône-fonction, et ce quel que soit le corpus de connaissances.

*Hypothèse 3 : Effet de la nature des icônes (schémas vs simili) sur la performance de LSA en identification d'icônes.* La distinction entre schémas et *simili* (Darras, 1996) qui, transposée

aux corpus des icônes, permet de discriminer les icônes figuratives et celles qui, strictement analogiques, sont assimilables à des *simili*. Nous posons tout d'abord une différence entre les taux d'identification obtenus par LSA et ceux des participants humains. En effet, LSA traite des corpus linguistiques plus ou moins riches. On peut donc penser que son processus d'identification s'appuiera sur la richesse et l'homogénéité des bases de connaissances, par un double processus cumulatif (via la capacité de LSA à inférer du sens à partir de l'accumulation de corpus, d'une part, et de la sommation de vecteurs, d'autre part). L'hypothèse que nous faisons est que les icônes figuratives seront plus aisément identifiées à leur fonction par LSA, dans la mesure où elles demandent une description linguistique plus détaillée, alors que les *simili* devraient l'être mieux par les humains, car directement perçues sans traitement cognitif complexe sur la base d'une représentation analogique fondée en perception.

*Hypothèse 4 : Similarité des performances humaines avec celles de LSA.* Les configurations de résultats obtenus avec la méthode de simulation seront identiques à celles obtenues avec des participants humains, chargés d'apparier chaque icône à la fonction correspondante, par un questionnaire. Notre hypothèse est que les résultats avec info-bulle seront meilleurs que sans, et que les performances des participants seront fonction de leurs expérience.

*Hypothèse 5 : Effet du sens de l'identification.* Schématiquement, dans un contexte de travail sur ordinateur, deux processus d'identification peuvent se présenter. Soit l'utilisateur se demande à quelle fonctionnalité du logiciel correspond une icône donnée, qu'il ne connaît pas, ou bien dont il pressent la fonctionnalité (condition icône vers fonctionnalité), soit l'utilisateur veut mettre en œuvre une fonctionnalité donnée dans le logiciel, et il se demande quelle icône permet son déclenchement (condition fonctionnalité vers icône). Nous faisons l'hypothèse que les performances des participants humains aux questions fonctionnalité vers icône seront meilleures que celles icône vers fonctionnalité, et ce d'autant plus que leur niveau en informatique est élevé. En effet, la deuxième situation est bien plus fréquemment rencontrée que la première, car les utilisateurs d'un logiciel passent bien plus de temps à se demander comment réaliser telle action qu'à se demander à quelle fonctionnalité correspond telle ou telle icône. Il est donc probable qu'ils réalisent des performances plus élevées dans la deuxième condition, et que cela dépendra également de leur niveau en informatique.

### **Description des participants et de leur tâche**

Nous avons comparé ces résultats aux performances de 86 participants humains. Dans la condition « icône vers fonctionnalité », nous avons présenté à 74 participants humains (collégiens novices et experts, étudiants de second cycle en informatique éducative) face à

une tâche similaire, c'est-à-dire en leur faisant passer un questionnaire à choix multiple (QCM) présentant chaque icône (*voir figure 4*), leur tâche étant de choisir la fonctionnalité adéquate de chaque icône parmi cinq. Pour tester la condition inverse, « fonctionnalité vers icône », nous avons demandé à 12 participants (6 étudiants utilisateurs novices de *Word*, et 6 étudiants de second cycle en informatique éducative) de répondre à un QCM présentant chacune des 23 fonctionnalités-cibles de *Word*, chacune assorties à un choix de 5 icônes.



Figure 4 – Copie d'écran de la barre d'icônes utilisée dans l'étude.

### Classement des icônes

Nous avons montré plus haut l'intérêt de l'opposition entre *similis* et schémas. Pourtant si toutes les icônes relèvent de la catégorie des schémas, certaines plus que d'autres présentent un caractère analogique marqué : ce sont en général des icônes d'objets qui, une fois identifiées, permettraient de manière relativement conventionnelle d'identifier la fonction correspondante. On peut penser que ces icônes demandent d'asseoir la compréhension de la fonction sur l'identification de l'objet. Pour les besoins de cette étude, certaines icônes de la barre ont été considérées comme des représentations mimétiques, comme des *similis*, dans la mesure où le degré d'analogie perceptive nous semblait dominer leur caractère schématisant (la distinction est parfois ténue et une épreuve interjuges serait une bonne façon de trancher).

Tableau 1 – Classement des icônes selon les catégories *similis*, figuratives et indéterminées.

Icônes <i>similis</i>	0, 2, 3, 4, 6, 9, 14, 15, 17
Icônes figuratives	1, 5, 12, 13, 16, 18, 19, 20, 21, 22
Icônes indéterminées	7, 8, 10, 11

### La méthode de conversion image vers texte

Il est également fondamental de disposer d'un moyen de faire correspondre, pour l'identification des icônes, leur description avec leur fonctionnalité. Pour cela, il faut que le format de leur codage soit identique. Cette contrainte, en réalité, ne semble pas incompatible avec la recherche sur l'imagerie (Kosslyn, 1981), qui pose qu'un traitement profond de l'image est réalisé. Il est important de noter que, comme il n'est pas possible de connaître la nature exacte du langage formel utilisé pour ce codage, nous utiliserons ici arbitrairement un codage verbal.

Il existe deux manières de régler le problème de la conversion entre deux registres sémiotiques (imagé et fonctionnel). La *procédure de bas niveau*, qui consiste à filtrer les informations de l'image de manière à ne conserver que ses traits principaux, qui peuvent ensuite être analysés (Forsythe *et al.*, 2003). La *procédure de haut niveau*, plus subjective, qui consiste à faire décrire l'icône par des juges avant toute analyse de ses fonctionnalités. LSA étant un système traitant uniquement des informations textuelles, il a été nécessaire de coder de manière textuelle les différentes images. Il a fallu réaliser un codage textuel le plus valide et fidèle possible, de manière à conserver une congruence maximale entre les images et leur codage textuel (Duval, 1995). Nous avons élaboré *a priori* une série de consignes permettant de coder textuellement chaque icône de la manière la plus fidèle possible.

- Pas d'interprétation fonctionnelle des icônes (*i.e.*, pas de liens avec ce que l'icône représente d'un point de vue informatique). Par exemple, une icône d'imprimante sera nommée « imprimante », plutôt que « imprimer ».
- Pas de codage de la couleur.
- Lorsque des éléments de l'icône sont identifiables, privilégier leur dénomination globale, plutôt qu'analytique (*i.e.*, une icône représentant une loupe sera décrite en tant que telle, plutôt que « un cercle avec une barre en bas à droite »).
- Description des éléments des icônes de gauche à droite et de haut en bas.
- Conserver la cohérence entre les différentes descriptions d'un même objet imagé : « la feuille à coin plié » est nommée à l'identique, qu'elle soit dans l'icône « Nouveau document » ou dans l'icône « Aperçu avant impression ».

Par exemple, l'icône 0 sera décrite « feuille de papier à coin plié », l'icône 1 « dossier ouvert », l'icône 2 « disquette », l'icône 3 « feuille qui sort de l'imprimante ». Il nous a paru essentiel de rationaliser cette étape de conversion qui, comme indiqué en préliminaires de la description de cette étude, est censée correspondre à un traitement propositionnel, profond, des images, que nous posons proche de celui réalisé à la lecture de la « surface » de ces dernières (Kosslyn, 1981). Ainsi, LSA simulerait le traitement de la représentation profonde de l'image, et c'est à partir de cette représentation qu'une recherche de la fonctionnalité la plus proche serait réalisée. Ce qui paraît, à première vue, comme un artefact seulement destiné à faciliter une comparaison par informatique serait donc assez proche de la manière dont les humains réalisent effectivement ce type d'identification.

Pour des raisons de place, nous ne présenterons ici que des résultats généraux, sans tests statistiques. Nous laissons le lecteur se reporter à la description complète (Dessus & Peraya, 2005).


### **Traitement des données**

Nous avons tout d'abord compilé individuellement, avec LSA, les différents corpus ci-dessus, puis, dans un deuxième temps, certaines combinaisons (*voir tableau 3*). Par exemple, *Le Monde* devrait simuler les connaissances des participants novices, alors que le corpus composé de tous les autres (TP+aide+*Le Monde*) devrait simuler celui des participants experts. Ensuite, nous avons comparé systématiquement chaque description d'icône, selon les deux modalités (*i.e.*, icône seule, icône+info-bulle) aux descriptions des fonctionnalités correspondantes présentées dans les QCM des participants humains, récupérées dans l'aide en ligne. LSA a calculé, pour chaque comparaison, la proximité sémantique entre la description textuelle de l'icône et les cinq fonctionnalités candidates, en comparant leurs vecteurs, sans toutefois intégrer ces descriptions à l'espace sémantique. Ces différentes valeurs nous permettent déjà de calculer une valeur de seuil (arbitrairement la moyenne de toutes les proximités calculées, augmentée d'un écart type). Il est donc à noter que les résultats présentés plus bas diffèrent de ceux présentés dans l'étude *princeps*, puisque réalisés à partir des QCM et non par comparaison avec les 23 fonctionnalités possibles.

À titre d'exemple, la description textuelle de l'icône 0 (colonne 3 du tableau 2) a été successivement comparée par LSA à chacune des cinq fonctionnalités du QCM – dont celle présente à la colonne 4 du tableau 2). Il a été utile de réaliser cette comparaison exhaustive (différente de la tâche donnée aux participants humains), afin de mieux comprendre comment se comporte le système, et surtout pour détecter des régularités dans les non-identifications d'icônes. La fonctionnalité ayant la valeur maximale est considérée comme la fonction appariée à l'icône par LSA. Si cette valeur de proximité est inférieure au seuil, cela signifie qu'aucune fonction n'est suffisamment proche de la description de l'icône, et que nous sommes en présence d'une « *non identification* ». Dans le cas où la valeur est supérieure ou égale au seuil, deux cas sont à envisager : soit l'icône correspond bien à la fonction, ce qui veut dire que l'identification est correcte. Le résultat est donc « *Identification correcte* » ; soit l'icône correspond à une autre fonction, ce qui est une « *confusion* ». Selon le nombre d'icônes au-dessus du seuil, nous avons constitué deux catégories : les icônes « *ambiguës* » (*i.e.*, plusieurs fonctions sont suffisamment proches), et les icônes « *uniques* » (*i.e.*, une seule fonction, adéquate ou non, est au-dessus du seuil fixé). Cette distinction, également faite par

Goonetilleke *et al.* (2001), est utile, car l'ambiguïté des icônes amène à des durées de décision plus importantes, alors que l'unicité autorise toujours une récupération de l'éventuelle erreur (*i.e.*, cliquer sur une icône non appropriée).

Tableau 2 — Extrait du corpus traité par LSA.

Icône	Info- Bulle	Description textuelle	Fonctionnalité <i>Microsoft Word 97</i> (aide en ligne)
	Nouveau document	Feuille de papier à coin plié.	Création d'un document. Dans le menu Fichier, cliquez sur Nouveau. Pour créer un document vide, cliquez sur l'onglet Général, puis double-cliquez sur l'icône Document vide. Pour créer un document fondé sur un modèle ou un assistant, cliquez sur l'onglet correspondant au type de document que vous souhaitez créer, puis double-cliquez sur le nom du modèle ou de l'assistant à utiliser. Conseil : Pour créer rapidement un nouveau document fondé sur le modèle par défaut, cliquez sur Nouveau. Remarque Si vous ne voyez pas le modèle ou l'assistant souhaité dans la boîte de dialogue Nouveau, vous pouvez afficher la liste complète des modèles et des assistants Microsoft Word et installer ceux qui vous intéressent. Pour plus d'informations, cliquez sur...

## Résultats

*L'hypothèse 1 n'a pas été confirmée* : d'une part, les performances de LSA ne vont pas croissantes lorsqu'on spécialise les corpus (vraisemblablement parce que le corpus sémiotique est trop peu important en taille), d'autre part, ces performances ne sont jamais très supérieures au hasard (jusqu'à 26 %), du moins tant que seule l'information issue de l'icône est traitée.

En revanche, *l'hypothèse 2 est confirmée* : l'ajout de l'information contenue dans les info-bulles permet d'améliorer notablement l'identification (*voir le tableau 3*), et cette dernière est croissante lorsqu'on spécialise les connaissances utilisées, avec le même plafonnement lors de l'ajout du corpus sémiotique.

*L'hypothèse 3 n'est pas vérifiée* : LSA échoue à identifier surtout les icônes schémas, qui auraient dû être mieux identifiées. En revanche, des résultats intéressants apparaissent en étudiant les résultats des humains (*voir le tableau 4*). Les novices identifient moins bien les icônes *similis*, sans profiter des informations de l'info-bulle : l'ambiguïté des traits des icônes, l'absence de connaissances sémiotiques et le fait qu'ils aient une mauvaise connaissance des fonctions les empêchent de reconnaître certaines icônes *simili*. En revanche, les experts identifient moins bien les icônes schémas (connaissances sémiotiques) sans info-bulle, mais les identifient toutes avec cette information : ils parviennent à résoudre des ambiguïtés, car ils ont une bonne connaissance des fonctionnalités du logiciel. Si nous ne sommes pas parvenus à reproduire les performances en identification d'icônes des participants humains, nous

montrons qu'ajouter l'information des info-bulles à celle des icônes améliore les performances de la simulation.

À propos de *l'hypothèse 4*, les performances humaines vont dans le sens de la simulation : les participants experts identifient correctement plus d'icônes que les novices, et les deux groupes en identifient correctement plus avec l'aide de l'info-bulle que sans.

Pour finir, *l'hypothèse 5* est en partie vérifiée : les participants novices ont de meilleures performances en identification dans le sens icône vers fonctionnalité que dans le sens inverse. En revanche, et contrairement à ce qui était attendu, les experts ont des performances similaires dans les deux sens de l'identification. Il est donc plus difficile, pour une personne ne connaissant pas un logiciel, d'identifier une icône à partir d'une fonctionnalité que l'inverse : le processus d'identification icône-objet assurant plus facilement une identification ultérieure objet-fonctionnalité.

Tableau 3 – Principaux résultats concernant le taux d'identification des icônes par réponse au QCM. Le seuil du hasard est de 20 %.

		LSA			Humains	
		<i>Monde</i>	<i>Monde+Aide</i>	<i>Monde+Aide+Sémio</i>	Novices	Experts
Icône vers fonctionnalité	sans info-bulles	9 %	22 %	22 %	77 %	83 %
	avec info-bulles	39 %	43 %	43 %	81 %	91 %
Fonctionnalité vers icône	sans info-bulles	22 %	26 %	26 %	26 %	70 %
	avec info-bulles	39 %	57 %	43 %	74 %	100 %

Ces résultats rendent plausible le processus suivant. Les participants, pour reconnaître une icône, disposent de connaissances sémiotiques et, le cas échéant, de l'information de l'info-bulle. Ils réaliseraient une première conversion, verbale, avec l'aide de leurs connaissances sémiotiques (*i.e.*, répondre à la question « à quoi ressemble cette icône ? »). Ensuite, une deuxième conversion, fondée sur les connaissances techniques du logiciel, est réalisée, menant à la fonctionnalité du logiciel, et pouvant être aidée par l'info-bulle. Les participants novices, ne disposant pas de connaissances sémiotiques suffisantes, ne peuvent correctement convertir les icônes proposées et, par conséquent, profitent moins des informations de l'info-

bulle. De plus, ils ne disposent pas non plus des connaissances techniques suffisantes pour passer de la fonctionnalité à l'icône. Les experts, qui disposent, eux, de ces connaissances, en profitent.

Tableau 4 – Catégories (*simili* ou schémas) des icônes non correctement identifiées (condition icône vers fonctionnalité).

Corpus	Modalité	Nombre d'icônes non correctement identifiées	
		Schémas	<i>Simili</i>
<i>Le Monde</i>	Icône	5	1
	Icône + Info-bulle	2	1
Le Monde + Aide + TP sémiotique	Icône	4	0
	Icône + Info-bulle	1	1
Humains novices	Icône	1	3
	Icône + Info-bulle	1	4
Humains experts	Icône	4	1
	Icône + Info-bulle	1	1

### **Quel contexte pour l'identification d'icônes ? Le problème du *symbol grounding***

Le rôle des connaissances et du contexte dans la compréhension a été abordé de multiples façons par différentes disciplines. Nous avons déjà évoqué les rôles des encyclopédies d'interprétation et des inférences, largement utilisés dans le cadre des analyses sémiotiques ou sémiocognitives. D'un point de vue linguistique, l'importance du contexte n'a cessé de croître avec la linguistique de l'énoncé et les différents courants pragmatiques (Ducrot, 1972, 1980 ; Recanati, 1980, 1981). De simples conjonctions de symboles ne pourraient générer du sens, ces derniers doivent être fondés dans autre chose (*e.g.*, des percepts issus de l'action). Pour reprendre une célèbre expérience de pensée, une personne anglophone placée dans une chambre chinoise (Searle, 1980), sans connexions avec le monde extérieur, et disposant d'un dictionnaire anglais-chinois, ne peut acquérir une véritable compréhension de la langue (Harré & Wang, 1999). Nous pouvons replacer cette expérience, terme à terme, dans le cadre de notre problème : un système informatique peut-il identifier des icônes, sans autres connaissances du monde, et surtout de la langue, que celles correspondant à un corpus de textes décrivant des connaissances encyclopédiques, le fonctionnement du logiciel et des inférences expertes à propos de sémiotique ? Un triple parallèle entre ces deux situations peut être fait. Le premier concerne la capacité même de conversion (ou, au sens large, traduction), c'est-à-dire de disposer d'une méthode valide et fiable pour convertir des objets codés dans un format (*e.g.*, une langue) dans un autre. Le dictionnaire proposé par Searle fait office de

méthode, mais nous pouvons très bien imaginer des cas (comme le nôtre), où ce dictionnaire n'existe pas, et qu'il faille déterminer un moyen de pallier son absence. Le second parallèle concerne l'exposition à un corpus comme un moyen d'inférer des connaissances à partir de ce dernier et la capacité à faire se correspondre des descriptions d'icônes à leur fonctionnalité, pour peu que cette capacité repose sur une analyse des co-occurrences. De la même manière que les humains infèrent une signification à des événements ou des mots en fonction de leurs co-occurrences dans divers contextes. À ces deux titres, le choix de LSA est justifié, des recherches ayant à la fois montré ses capacités à inférer le sens de mots à partir de dictionnaires bilingues (Westerveld *et al.*, 1999) ; et aussi, (Landauer & Dumais, 1997), à inférer le sens de mots à partir de l'analyse de co-occurrences inter- et intraparagraphes (*i.e.*, contextes). Le troisième parallèle permet de valider le choix du codage verbal pour coder le contenu des icônes. Initialement, ce choix était dû aux contraintes de traitement informatique (LSA ne pouvant traiter que du texte), mais les résultats montrent, de plus, que même si le codage verbal des icônes engendre nécessairement une perte importante d'informations, il en reste suffisamment pour réaliser l'appariement.

Plus largement, ces considérations réfèrent au problème du *symbol grounding*, qui a récemment fait l'objet d'un débat entre les tenants d'une approche computationnelle du sens (*e.g.*, Fodor, 1986) avec ceux d'une approche « fondée » (*grounded*). Rappelons-en brièvement les principaux arguments. Barsalou (1999, voir aussi Glenberg & Kaschak, 2003) a montré que les sciences cognitives, en étant largement influencées par une approche logique et computationnelle de la représentation, en sont venues à proposer des langages représentationnels fondamentalement amodaux (*frames*, réseaux sémantiques, systèmes de productions, etc.). Les objets réels décrits par ces systèmes ont donc fait l'objet de conversions, mais le problème principal est que leur représentation n'est reliée à la réalité que de manière arbitraire, différente de celle utilisant notre perception. Cela pose, selon Barsalou, certains problèmes importants : d'une part, celui de la réalité neuropsychologique de telles représentations, d'autre part, celui de la « transduction » : nos états perceptuels sont traduits dans un système de représentation complètement différent, qui décrivent ces états de manière amodale, et fait ainsi perdre la relation perception/représentation. Barsalou note que cette perte empêche ces symboles amodaux d'être à nouveau, et fidèlement, reproduits en symboles perceptuels. Quelques opposants à cette vue (*e.g.*, Goldstone & Rogosky, 2002 ; Landauer, 2002) montrent qu'une approche computationnelle analysant les *patterns* de mots au sein de leurs contextes d'apparition rend compte de manière satisfaisante de la connaissance inférée

de la « perception » de ces mots. Landauer (2002) montre ainsi que des approches visant à détecter de telles régularités (co-occurrences d'éléments) au sein d'environnements (corpus) ont donné des résultats satisfaisants, que ce soit en étudiant le langage, la construction de connaissances à partir de textes, ou même la perception.

Dans l'étude présentée ici, et même si toutes nos hypothèses n'ont pas été confirmées, nous avons montré qu'une machine analysant les régularités dans un corpus contenant *un même* système de symboles codant *deux types* différents de signifiés (*e.g.*, descriptions d'icônes et fonctionnalités de logiciels) peut les mettre en rapport sans indice supplémentaire (*i.e.*, règle d'identification icône-fonctionnalité). Il nous reste maintenant à expliquer pourquoi cela fonctionne. LSA, placé dans l'équivalent d'une chambre chinoise, sans dictionnaire image-fonction explicite, parvient, par l'analyse des régularités d'un corpus, à relier une partie des icônes à leur fonction dans le logiciel.

- Un humain, ou une machine, peut *situer* un mot (*i.e.*, comprendre son contexte, son sens descriptif) sans le *fonder* (*ground*, *i.e.*, savoir quels sont les objets, contextes, usages, etc. auxquels il renvoie). L'inverse n'est pas possible (Shaw, 2003) ;
- Situer des mots peut se faire *de seconde main*, alors que les fonder nécessite une expérience de première main, donc perceptuelle (voir les phénomènes d'attention partagée, d'ostension, décrits dans Tomasello, 2003) ;
- *LSA situerait donc des mots sans les fonder*. L'analyse de corpus de mots à propos d'icônes et de fonctionnalités de logiciels permet donc d'activer une *situation* des icônes, préalable à leur compréhension optimale.

## **Discussion**

La recherche que nous avons présentée plus haut permet de mieux comprendre le processus d'identification d'icônes, lorsqu'il est mis en relation avec son contexte ou avec les connaissances de l'utilisateur du logiciel. Nous avons montré que le processus d'identification icône-fonctionnalité utilise différents sous-processus de conversion et qu'il peut en partie être automatisé à l'aide d'une méthode statistique fondée sur l'analyse de régularités. À plus long terme, il serait possible d'envisager un logiciel qui adapte automatiquement ses icônes en fonction des connaissances de ses utilisateurs. Une autre application possible pourrait être un outil d'évaluation de la complexité des icônes, toujours en tenant compte des connaissances de l'utilisateur et du domaine. Un tel logiciel pourrait intégrer un système de recueil des durées de réponse, ce qui permettrait une analyse plus fine du processus d'identification.

Les icônes d'ordinateur sont un objet d'étude intéressant pour comprendre les processus cognitifs engagés dans la compréhension sémiotique, notamment pour approfondir les questions suivantes :

- l'analyse de la compréhension des icônes pourrait aider à la compréhension du rapport entre images matérielles et images mentales, ainsi qu'à l'explicitation des processus de genèse de l'icône ou du programme figuratif attaché aux résumés cognitifs ?
- si, par définition, l'image d'une icône est similaire à l'objet qu'elle représente, pourquoi faut-il nécessairement une phase d'apprentissage des logiciels (Scalisi, 2001), ainsi que des info-bulles ?
- est-il possible de construire des systèmes artificiels capables de sémosis ? (Gudwin, 1999).

La réponse à ces questions nécessitera de nouvelles études expérimentales permettant de mieux comprendre les complexes processus de conversion entre systèmes sémiotiques.

### **Note des auteurs**

Cette étude a été menée pendant que le premier auteur était chercheur invité à TECFA, université de Genève, à l'occasion d'un congé de recherches du ministère de l'éducation. Nous remercions Benoît Lemaire et Cyril Rebetez pour leurs commentaires d'une précédente version de ce chapitre, Silvia Cadillo pour sa collaboration dans la passation des questionnaires « fonction vers icône », ainsi que David et Michèle Delavier, Stéphane Lattion et Stéphane Morand pour leur aide pendant la passation des questionnaires, et enfin Bérénice Jaccaz pour son aide dans la saisie des données.

### **Références**

Almassy, P. (1974). Le choix et la lecture de l'image d'information. *Communication et Langages*, 22, 57-69.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.

Barthes, R. (1964). Rhétorique de l'image. *Communications*, 4, 91-134.

Berrendonner, A. (1991). Référent. In R. Doron & F. Parot (Eds.), *Dictionnaire de Psychologie* (p. 583). Paris : PUF.

Bronckart, J.-P. (1996). *Activité langagière, textes et discours. Pour un interactionisme sociodiscursif*. Berne : Peter Lang.

- Cordier, F. (1993). *Les représentations privilégiées. Typicalité et niveau de base*. Lille : Presses Universitaires de Lille.
- Coste, D. (1975). Les piétinements de l'image. *Études de Linguistique Appliquée*, 17, 5-29.
- Darras, B. (1996). *Au commencement était l'image*. Paris : ESF.
- Denis, M. (1979). *Les images mentales*. Paris : Presses Universitaires de France.
- Denis, M. (1989). *Image et cognition*. Paris : Presses Universitaires de France.
- Denis, M. & De Vega, M. (1992). Modèles mentaux et imagerie mentale. In M.-F. Ehrlich, H. Tardieu & M. Cavazza (Eds.), *Les modèles mentaux, approche cognitive des représentations* (pp. 79-100). Paris : Masson.
- Dessus, P., & Peraya, D. (2005). Le rôle des connaissances dans l'identification automatique d'icônes: une comparaison avec des humains. *Revue d'Intelligence Artificielle*, 19, 195-214.
- Ducrot, O. (1972). *Dire et ne pas dire*. Paris : Hermann.
- Ducrot, O. (1980). *Les mots du discours*. Paris : Minuit.
- Duval, R. (1995). *Sémiosis et pensée humaine*. Berne : Lang.
- Duval, R. (1999). *Conversion et articulation des représentations analogiques, Séminaire de recherche 1*, Direction de la recherche et du Développement, IUFM Nord Pas de Calais.
- Duval, R. (2006). Transformations des représentations sémiotiques et démarches de pensée en mathématiques. In *Acte du colloque de la COPIRELEM* (pp. 67-89). Strasbourg : IREM de Strasbourg.
- Eco, U. (1970). Sémiologie des messages visuels. *Communications*, 15, 11-51.
- Eco, U. (1988). *Sémiotique et philosophie du langage*. Paris : PUF (édition italienne 1984).
- Erni, S. (1986). Réalité et perception de la réalité – la communication visuelle dans la coopération au niveau du développement. *Éducation des Adultes et Développement*, 27, 157-174.
- Fodor, J. A. (1986). *La modularité de l'esprit*. Paris: Éditions de Minuit.
- Forsythe, A., Sheehy, N., & Sawey, M. (2003). Measuring icon complexity: An automated analysis. *Behavior Research Methods, Instruments, & Computers*, 35(2), 334-342.
- Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. *The Psychology of Learning and Motivation*, 43, 93-126.

- Goffman, E. (1991). *Les cadres de l'expérience*. Paris : Minuit.
- Goldstone, R. L., & Rogosky, B. J. (2002). Using relations within conceptual systems to translate across conceptual systems. *Cognition*, 84, 295-320.
- Goonetilleke, R. S., Shih, H. M., On, H. K., & Fritsch, J. (2001). Effects of training and representational characteristics in icon design. *International Journal of Human-Computer Studies*, 55, 741-760.
- Groupe  $\mu$  (1992). *Traité du signe visuel. Pour une rhétorique de l'image*. Paris : Seuil.
- Gudwin, R. R. (1999). From semiotics to computational semiotics. *Proc. 9<sup>th</sup> Int. Congress of the German Society for Semiotic Studies (IASS)*. Dresde.
- Hammond, N., & Barnard, P. (1984). Dialogue design: Characteristics of user knowledge. In A. Monk (Ed.), *Fundamentals of Human-Computer Interaction* (pp. 127-164). Londres: Academic Press.
- Harré, R., & Wang, H.-T. (1999). Setting up a real 'Chinese Room': an empirical replication of a famous thought experiment. *Journal of Experimental and Theoretical Artificial Intelligence*, 11, 153-154.
- Hébert, L. (2006), « Les structures du signe. Le signe selon Klinkenberg ». In L. Hébert (dir.), *Signo* [en ligne], Rimouski (Québec), document disponible à <http://www.signosemio.com>.
- Johnson-Laird, P. (1983). *Mental Models: Towards a cognitive Science of Language, Inference and Consciousness*. Cambridge : Cambridge University Press.
- Johnson-Laird, P. (1988). Modèles mentaux en sciences cognitives, *Bulletin de Psychologie*, XLI (383), 60-88.
- Klinkenberg, J.-M. (1996). *Précis de sémiotique générale*. Bruxelles : De Boeck.
- Kosslyn, S. M. (1981). The medium and the message in mental imagery: a theory. *Psychological Review*, 88(1), 46-66.
- Landauer, T. K. (2002). On the computational basis of learning and cognition: Arguments from LSA. *The Psychology of Learning and Motivation*, 41, 43-84.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104(2), 211-240.

- Langacker, R. (1987). *Foundations of Cognitive Grammar, I, Theoretical Prerequisites*. Stanford : Stanford University Press.
- Lefebvre-Mignot, Y. (Ed.). (1979) Audio-visuel et développement. *Tiers Monde*, XX(79).
- Lemaire, B. & Dessus, P. (2003). Modèles cognitifs issus de l'Analyse de la sémantique latente. *Cahiers Romains de Sciences Cognitives*, 1(1), 55-74.
- McNamara, T. P. (1999). Single-code versus multiple-code theories in cognition. In R. J. Sternberg (Ed.), *The Nature of Cognition* (pp. 113-136). Cambridge: MIT Press.
- Metz, C. (1964). Le cinéma, langue ou langage, *Communications*, 4, 52-90.
- Moles, A. A. (1981). *L'image communication fonctionnelle*. Tournai : Casterman.
- Morris, C (1946). *Signs, Language and Behavior*. New York : Prentice Hall.
- Moyes, J., & Jordan, P. W. (1993). Icon design and its effect on guessability, learnability, and experienced user performance. In J. L. Alty, D. Diaper & S. Guest (Eds.), *People and Computers VIII (HCI'93)* (pp. 49-59). Cambridge: Cambridge University Press.
- Ogden, C. & Richard, I. (1923). *The Meaning of Meaning*. Londres : Routledge and Kegan Paul.
- Peirce, C. S. (1931/1938). *Collected Papers* (Vol. II, *Elements of Logic*). Cambridge, Mass. : Harvard University Press.
- Peraya, D. (1998). Structures et fonctionnement sémiotiques des icônes de logiciels et d'environnements informatiques standardisés (ILEIS). *Recherches en Communication*, 10, 101-140.
- Recanati, F. (1980). Qu'est-ce qu'un acte illocutionnaire ? *Communications*, 30, 190-215.
- Recanati, F. (1981). *Les énoncés performatifs*. Paris : Minuit.
- Richardson, M. (2000). Do people use icons? An investigation into use of the text-prompt associated with icons. Unpublished Msc. in Information processing, University of York, York.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, N.J.: Erlbaum Associates.
- Scalisi, R. (2001) A semiotic communication model for interface design. *Proc. Conf. Cosign 2001 (Computational Semiotics)*. Amsterdam.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457.

Saussure F. de (1972). *Cours de linguistique générale* (Edition critique préparée par T. de Mauro). Paris : Payot.

Shaw, R. (2003). The agent-environment interface: Simon's indirect or Gibson's direct coupling? *Ecological Psychology*, 15(1), 37-106.

Tomasello, M. (2003). *Constructing a Language, a Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Ullman, S. (1957). *The principles of semantics* (2<sup>e</sup> éd.). Glasgow : Glasgow University Publications.

Westerveld, T., Hiemstra, D., & de Jong, F. (1999). Extracting bimodal representations for language-based image retrieval. *Paper presented at the Multimedia'99 Conference*. Vienne (Autriche).