

Cognitive Modelling of Icon Comprehension

Philippe Dessus* & Daniel Peraya**

* University of Grenoble & IUFM, France
Philippe.Dessus@upmf-grenoble.fr

** University of Geneva, Switzerland
Daniel.Peraya@tecfa.unige.ch

Chapter submitted to E. McKay (Ed.), *Enhancing Learning through HCI*. Hershey: Idea Group Publishing

Introduction

Nowadays, Human-Computer Interaction is mainly carried out through icon-based interfaces. Thus, before verifying whether HCI enhances learning, it would be necessary to verify that, and how, HCI can be actually fostered by icons, when compared with textual commands. This latter question entails to solve issues about the meaning conveyed by icons. For instance, the command expressed by the word “print” has only one meaning, but understanding an icon is more cognitively complex. Since each icon represents an object –in that case, a printer–, this object has often to be understood, by metonymy, as the action made possible by this object – in that case, “to print”. But the correct matching between a printer icon with the command “to print” is not so straightforward whoever the user is. An icon is indeed polysemic; it can refer either to “print”, “switch the printer on”, “feed paper into printer”, and so on, depending on the background knowledge of the user.

How to objectively assess the degree of polysemy conveyed by icons? Do icons actually improve the access to the intentional software function? How better this access is when tooltips are combined with icons? What is the effect of the user’s background knowledge on icon comprehension? The purpose of this chapter is to answer these questions, which are mainly related to how representations of an icon are cognitively computed and referred to the corresponding software function.

The interface of any recent software is composed of numerous icons. These icons are graphic representations of objects that are intended to be close to the user’s mental model, thus easier to recognize. But the matching between these representations and the actual function of the software raises some cognitive problems. The user has to perform a conversion between the object represented and the function of the software, which is not univocal (see the notion of congruency, Duval, 1999). For instance, a “paper sheet with one folded border” icon, though faithfully represented, will not necessarily lead to its actual function (i.e., “New Blank Document” in most softwares): there are numerous alternate candidates like “Clear”, “Page Layout”, or even “Page Break”. Conversely, the function “New Blank Document” may be represented by numerous icons like “a blank window”, “a blank paper sheet without folded border”, and so on (Barsalou, 1999).

The quality of the matching thus depends both on the fidelity of the symbolic representation (i.e., the object-to-image relation, or how close to the represented object is the picture) and on the user’s knowledge about how the software works (i.e., image-to-function relation, or knowledge about the software functionalities, its domain, and so on, see Hammond & Barnard, 1984, for a comprehensive review). In turn, both fidelity of the symbolic representation and user’s knowledge may affect the user’s cognitive processes involved in icon identification, namely, her ability to recognise what the icon means, and her ability to associate its meaning to the adequate software command or function (Richardson, 2001). This two-fold conversion entails some issues. Firstly, we have to tackle how to determine precisely how much icons are polysemic. This process is closely akin to a decision process (i.e., I am certain, 75%, that this icon depicts a house). Secondly, the second part of the conversion process (image description to software function) needs to model both the amount of knowledge about the software as well as general knowledge about the objects represented in the software (e.g., knowledge about the publication process is necessary for manipulating icons of desktop publishing systems).

The aim of this chapter is to provide some computerized modelling of these two processes. We designed two systems that model or assess some cognitive processes involved in icon recognition. The first system, the *Iconometer*, after Leclercq’s (1992) work, is a web-based tool for the collective assessment of the meaning conveyed by an icon (Peraya & Strasser, 1999; Simitsek, 2002). Since the polysemy of an icon is a highly subjective dimension, its

assessment by several raters may assure a better reliability of the scheme. The second system, called *AIR*, for Assisted Icon Recognition (Dessus & Peraya, 2005), models the user's background knowledge necessary to match the icon with its adequate function. Such a system would be useful for designing icons that are coherent the user's knowledge, and thus are easier to understand. The remainder of this chapter is as follows. The first part is a three-fold literature review about the icon comprehension process from HCI viewpoint. In a second empirical part we present our two systems and how they were tested, and then we present some further applications of these systems on learning. Finally comes a discussion about technical and practical issues.

Icon Comprehension: a Literature Review

There are two ways to tackle the issue of how to translate an icon between two semiotic registers (pictorial and functional). The low-level procedure, in which the image of the icon is automatically filtered to get its main features that can in turn be analyzed (Forsythe et al., 2003); and the high-level procedure, which accounts for the human subjectivity involved in this conversion process. We review here higher-level approaches of icon comprehension, by describing three major research trends: the semiotics approach, the study of the role of knowledge, and the effect on icon comprehension of additional artefacts like tooltips.

Semiotics

The study of the meaning of icons, as representations, pertains to semiotics. Goguen (1999) pointed out that "A user interface for a computer system can be seen as a semiotic morphism from (the theory of) the underlying abstract machine (what the system does) to a sign systems for windows, buttons, menus, etc." Such morphisms are metaphorical devices that help users to grasp how to do in order to perform a given task.

Before considering computer icons per se, we will briefly review works on semiotics from Barthes (1985) and Eco (1976) to Darras (1996) and Duval (1995), who answered some important questions like "what is an image?", "what is an analogical representation?", "Are there different categories of icons?"

Since early studies in semiology were involved in the analyses of images as well as analogical languages, the very definitions of the sign and the linguistic code they used were inappropriate to account for visual languages. In photography, for example, "the relation between signified and signifiers is not one of 'transformation' but of 'registration', and *the absence of code* obviously reinforces the myth of photographic 'naturalness' [...] (Barthes, 1985, p. 33, italics added). Then, the Peircean distinction between index, icon and symbol fuelled research about the notion of iconicity, namely the degree of analogy between the representative—the icon—and the represented—the object (e.g., Moles, 1981). But trying to find out a one-to-one correspondence between what is meant and real world objects is unsatisfactory. Difficulties also arise in the semantic processing of images. Scholars like Barthes (1985) argued that images are intrinsically polysemous: a reader can pick up some "signified" and be unaware of others. Thus, two main processes often described in structural theories of communication appear to be unlikely. First, the "monosemantic" processing of images, given the richness of the network of metaphorical and metonymic relations activated by each image "reader" (Meunier & Peraya, 2004). Second, the very process of "decoding", which cannot definitively be viewed as coding in reverse, as showed either by many daily experiences or by scientific and teaching practice—in particular in intercultural situations.

Breaking with this paradigm right when structuralism was mainstream, Eco (1976) suggested that significance would occur in the analogical ratio that is to be found between the sign and the perceptive model of the object rather than between the sign and its referent. This new trend of research waited twenty years to emerge, with the work of the μ Group (1992) on the

perceptive foundations of visual semiotics. The very notion of sign was deeply modified, and composed of the significant, the referent and the type, whose role is to keep equivalence between the two others.

Darras (1996) went even further in integrating outcomes of research on mental models (Denis, 1991). Darras analyzed the effects of the various categories of visual representations on the cognitive mechanisms and the categorisation processes. A difference was drawn between two types of visual representations: *simile* and schemas. *Simile* are reproducing perceptual phenomena while schemas are simplifications carried out by cognitive processes. More precisely, *simile*, e.g., photographic or cinematographic images, are visual representations that “display a high degree of visual resemblance to their referents” (Kindler, 2004, p. 236) and are grounded in perceptive processes. Scales of iconicity have to be used to organize these visual representations according to their realism. Schemas, on the other hand, are progressively elaborated during childhood by frequent exposition to iconotypes, that are social prototypical representations (e.g., ☀ is the iconotype of the sun).

Finally, Duval (1995, 1999) claimed that the semiotic system (e.g., algebra, sentences, graphs) in which the representation is coded has to be taken into account for the categorisation of images. So a given object (e.g., a line) can have several representations (e.g., a drawing or an equation), depending on the semiotic system in which it is coded. Two kinds of semiotic processes can be carried out upon these representations: transformations and conversions. A transformation is a process that produces new meaning within the same system of representations, whereas a conversion is a translation of the representation of the same object between to different semiotic registers or systems. Duval pointed out that this conversion process can be more or less transparent, depending on whether it is performed like a unit-to-unit translation, or involves a more complex one. For instance, in the Clark and Chase’s experiment (1972, cited by Duval, 1995), subjects had to decide what sentence accurately described the image (see figure 1 below). The first sentence was the most rapidly selected because of the congruence of the conversion process (i.e., same word order); the last sentence however, although being true, was less rapidly selected. The congruence—either actual or not—between icons and their verbal description is one of the problems we will tackle here.

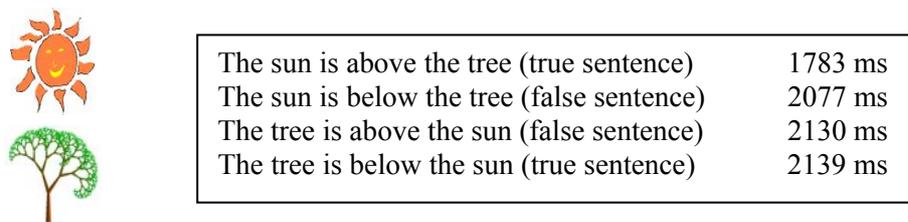


Figure 1 – Material of the Clark and Chase’s experiment (1972, quoted by Duval, 1995, pp. 47-48).

To sum up what semiotics contributed to our understanding of the nature of icons, we will outline, firstly, that some visual representations are related to cognitive summaries and are built from a reorganization of our cognitive resources; secondly, other representations, like imitations, are closer to perceptive events and thus require specific processing. Moreover, it is worth noting that images used in communication (e.g., instructional settings, HCI) have often been considered as *simile*—they effectively share some features with objects of real world—whereas they actually should be considered as schemas—their intended function is to foster learning or understanding.

The Role of Knowledge in Icon Identification

It is a very common fact that user’s background knowledge can have an effect on her performance on icon identification. However, few studies have empirically addressed this

problem, studies about cultural differences notwithstanding (e.g., Evers, Kukulska-Hulme, & Jones, 1999). Hammond and Barnard (1984) designed a typology of the different sources of knowledge required for an adequate human-computer interaction:

- ☞ knowledge of the domain (as represented in real life as well as in the system),
- ☞ knowledge of the computer version of the domain,
- ☞ knowledge of the non-computer “workbase” version,
- ☞ knowledge of the problem that can be solved by the system,
- ☞ knowledge of the system operations,
- ☞ knowledge of natural language,
- ☞ knowledge of the physical interface as well as the dialogue interface.

The multiplicity of these sources of knowledge entails some difficulties for assessing rigorously the effect of each of them on the identification task. Since knowledge about the functioning of the computer system is easily assessed through questionnaires (e.g., Liu, Crum & Dines, 1998), knowledge of natural language as well as domain knowledge have to a large extent been ignored. Likewise, Moyes and Jordan (1993) pointed out that “world knowledge” is necessary to perform particular tasks with icons; which is sometimes cultural (e.g., Evers et al., 1999; Piamonte, Abeysekera & Ohlsson, 2001). Some expertise reversal effects have even been showed: experts in icon design were unexpectedly less able than novices to recognize the function of computer icons (Holloway & Bailey, 1996). The mental model the icon functions for novices matched better with icons than experts mental models. Richardson (2000) showed a counterintuitive effect of expertise on the use of icons: although novices stated that they preferred using icons, they did not actually used them in performing tasks; conversely, experts stated that they used icons the least and obtained the best results in using them.

The Effects of Tooltips

Even though icons are more closely akin to user’s mental models than text-based interfaces, the frequent use of tooltips in recent interfaces paradoxically shows the difficulty of icon identification. Some recent studies assessed the effects of different artefacts like tooltips on icon comprehension, which are short texts temporarily displayed when the mouse is pointed over an icon. It is rather paradoxical to add such textual information to icons for helping their comprehension, since they were originally designed to allow comprehension without any textual reference. The usefulness of such tooltips lies perhaps in the difficulty to give an icon a meaning that is univocal (related to the world of objects) and corresponds to a unique function (related to the functioning of the software). Likewise, as we will show below, tooltips can help users to perform a semiotic conversion between two registers (Duval, 1999). Richardson (2000) carried out an experiment to assess the interaction of user expertise and use of tooltips on icon comprehension. He showed that the more expert are the user, the less the tooltips are used. This result confirms the assumption that icon usage is learned and not directly obvious.

The review of the literature shows some problems to be addressed. Firstly, how the icon identification process does account for users’ background knowledge? Our aim is to simulate some cognitive processes involved in the icon identification task. Two systems were implemented. Firstly, the *Iconometer* is a tool for assessing the degree of polysemy of icons; secondly, the *AIR system* is an AI-based tool that models knowledge necessary to identify the function of icons. Both systems are intended to give a solution to a more general question: how the meaning of an icon (i.e., its function) is conveyed to users, according to their different levels of knowledge. Through the test of these systems, the aim of this chapter is to bring solutions to the following questions:

- ☞ how to assess and if necessary minimize the degree of polysemy of icons?

- ☞ to what domains pertain the knowledge necessary to identify icons?
- ☞ how to translate the imaged information conveyed by icons into a verbal one?
- ☞ how to simulate the cognitive process of icon identification?

The *Iconometer*, A Tool for Assessing the Degree of Polysemy of Icons

Our first tool, the *Iconometer*, is a web-based system (written in PHP and MySQL) that helps test the degree of polysemy of an icon through a questionnaire. This system is freely accessible through Internet, and displays a set of images (see Figure 2 below, as well as Morand, 2005; Ott, 1998; Simitsek, 2002) and for each of them one has to formulate a hypothesis about the represented object, as well as the degree of certainty of this hypothesis (between 1 to 100).

Images à analyser: banque d'images (Tourisme) de Peraya
Description du thème:

<p>Image n°: 1</p> 	<p>Ajouter une hypothèse</p> <p>Hypothèse:</p> <p>C'est un/une <input type="text"/> (nom) .</p> <p>Certitude: (entrer un chiffre entre 1 et 100)</p> <input type="text"/> <p><input type="button" value="Ajouter"/></p>
<p>Image n°: 2</p> 	<p>Ajouter une hypothèse</p> <p>Hypothèse:</p> <p>C'est un/une <input type="text"/> (nom) .</p>

Figure 3 – Screenshot of the *Iconometer* (Peraya & Strasser, 1999). For each image in the left pane, the user has to formulate a hypothesis about what it is represented, as well as a degree of certainty of this formulation.

Then, for each image or hypothesis several values are computed:

- ☞ *the frequency of each hypotheses*, or the number of each hypothesis divided by the overall number of formulated hypotheses;
- ☞ *the certainty weight of each hypothesis*, or the average degree of certainty expressed for each hypothesis, derived in local weight (accounting for the number of participants who formulated the hypothesis) and in overall weight (accounting for the overall number of participants, whatever the hypothesis they formulated);
- ☞ *the potential of evocation of each image*, or the sum of all the certainty indices of the different hypotheses that have been formulated about the image;
- ☞ *the efficacy ratio of each image*, or the overall weight of the correct hypothesis divided by its potential of evocation.

For each image, the graphical representation of the frequency x local weight of each hypothesis provides information about its polysemy (Peraya & Strasser, 1999). This graphical representation is certainly the most useful clue of this application since it offers the easiest information to be interpreted (see Figure 3):

- ☞ only one hypothesis with a high local weight would mean that the image is monosemic (star);
- ☞ a hypothesis “low frequency” with low local weight would mean that this hypothesis, though weak, could be changed (square);

- ☞ several hypotheses with low frequency and high local weight lead to examine its rightness: if right, then it would be not enough prototypical (many alternate candidates), if false, then the participants will change their mind with difficulty (triangle);
- ☞ a high frequency with low local weight hypothesis would either be changed if it is true (weak participants conviction) or subject to misinterpretation if it is false (circle).

Moreover, analyzing this diagram for each icon gives an understanding of what are graphic features that induce false hypotheses, and thus helps the redesign of the icon for reducing its polysemy.

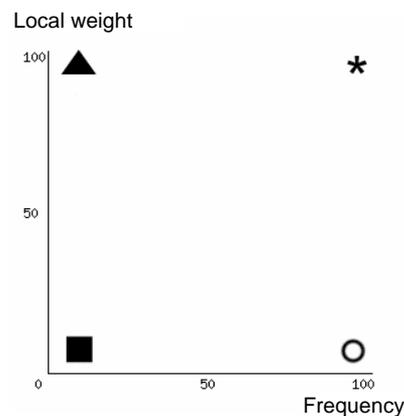


Figure 3 – Screenshot of the Iconometer (Peraya & Strasser, 1999). Frequencies and local weight of different hypothesis.

The AIR system, a Tool that Accounts for Knowledge in Comprehending Icons

Users' knowledge involved in the use of icons is difficult to assess, and thus model. Our second system, AIR (Assisted Icon Recognition), aims at modelling knowledge necessary for identifying an icon. We explore here the effects of a three-parts knowledge corpus composed of a) general background knowledge; b) knowledge of the software used; c) knowledge about semiotics. Moreover, research showed (e.g., Richardson, 2000) that expert users do not need explicit help about the function of icon (they simply point over the right icon and click on it), novices may benefit from the help of additional informations like tooltips, which are short texts temporarily displayed when the mouse is pointed over an icon. Thus the effects of such tooltips has to be tested as well. We carried out an experiment (fully reported in Dessus & Peraya, 2005) in order to test our system, and to compare the result of this test with human data.

Icon Similarity Modelling with Latent Semantic Analysis

We performed a computer simulation using Latent Semantic Analysis (LSA, Landauer & Dumais, 1997), a model for the representation of the meaning of words and sentences. LSA takes a huge corpus as input and yields a high-dimensional vector representation for each word. It is based on a singular value decomposition of a word x paragraph occurrence matrix, which implements the idea that words are given similar representations if they occur in similar contexts. The advantage of such vector representation is that it is straightforward to compute the similarity between vectors, usually by means of the cosine function. AIR relies on this LSA measure of semantic similarity between the verbal description of an icon and the description of the candidate functions as they appear in the software help. Such a matching resembles a human one when a user successively performs a matching between the perceived icon and the candidate software functions. Such human icon recognition uses background knowledge represented by large text corpora described below.

LSA is an appropriate tool for processing this kind of conversion, because they can be viewed as the translation of a document from a language (textual) to another (pictorial

translated into text). Dumais et al. (1997) used LSA in a cross-language retrieval task. A bilingual corpus containing the same information in both languages was firstly processed. Then queries about words (or documents) in one language are performed to retrieve corresponding words (or documents) in the other language. Westerveld et al. (1999) successfully carried out this procedure for an image to text translation.

Corpora

The corpus used is composed of three parts:

- a 24 million word corpus from a French daily newspaper (*Le Monde*), representing a general common knowledge base, which is here used as the “novice user knowledge base”;
- a .3 million word corpus from the *Microsoft Word Help*, representing knowledge about the software (i.e., the entire text of the *Microsoft Word 97 Help*);
- a .03 million word corpus from a master course about semiotics held by the second author at Geneva university, representing expert knowledge about icons.

Experimental Task

We asked 74 participants (high school pupils for novices and master students for experts) to match each of the 23 icons with the adequate function (selected among 5). This task was completed through a questionnaire (see Table 1) for comparing purposes with computerized identifications. In order to take into account expertise effect, two groups of participants were composed. After we ensured that no one of them had used *Microsoft Word*, we composed a novice group, with 35 secondary pupils (average age: 13), and an expert group, with 39 master students (average age: 24.5) engaged in an educational software design course. Participants were given the same MCQ questionnaire, composed of 23 questions, one per icon, with 5 alternative choices.

Table 1 – *Questionnaire Excerpt (one question among 23). Items have been translated from french. The right answer is item 3.*

	<input type="checkbox"/>	Creates a new, blank spreadsheet or a Microsoft graphic file based on the default template. Click where you wish insert a new incorporated object for a spreadsheet file or a graphic. In the Insertion menu, select Object, then on the New Object tab.
	<input type="checkbox"/>	Creates a table. Use tables for structuring informations and design attractive documents, incorporating, for example, texts with columns and images nearby. The Insert Table button allows to rapidly create a simple table containing an equal number of rows and columns.
	<input type="checkbox"/>	Creates a document. Click on New (File menu). In order to create a new blank document, click on the General tab, then double click on the Blank document icon. In order to create a document based on a template or an assistant, click on the corresponding tab, then double click on the appropriate name (template or assistant).
	<input type="checkbox"/>	Pastes (Edition menu). Insert the text in the clipboard at the insertion point and replaces the selected text. This command is only available when some text (or object, cell) was cut or copied.
	<input type="checkbox"/>	Creates a hypertext link. Users can enrich Web pages and Word documents by inserting hypertext links to other documents. The hypertext link allows the reference to either another part of the same document, another Word document or web page, or a file created using another software.

Hypotheses

The four following hypotheses were tested.

1. *The performance of the system is affected by the corpora that are used.* We expect to find that, the more the corpora contain specialized (i.e., expert-like) knowledge, the higher the identification performances. Thus, the “help” corpus would allow better performances than the “semiotic” corpus, and the latter would be better at identifying icons than the “newspaper” corpus. Likewise, the addition of the three corpora would entail better performances than when separate corpora are used.

2. *The use of tooltips enhance the performance of the system.* When information from tooltip is added (e.g., “print” concerning the “printer” icon), LSA identifies more icons than without this information.
3. *Simile vs. figurative icons are differently identified by humans and machine.* The *simile* icons (i.e., object-like) will be less recognized than the figurative (i.e., abstract) ones by the machine, though better recognized by human participants. For humans, *simile* icons are directly identified, without any translation from another register. Conversely, the *AIR system* will have better performance in identifying figurative icons, fostered by a more extensive description.
4. *The AIR system mimics the overall results in identification of human participants.* The pattern of results obtained by humans vs. machine will be similar. More precisely, the performance with the help of tooltips will be better than without this help, for both humans and machine. We also predict an interaction between this help and the participant’s expertise: the effect of tooltip will be higher for novices than for experts.

Materials

The *Microsoft Word 97* standard icon bar (23 icons) was used in this study (see figure 4). We chose that icon bar because experts generally have been exposed to it-conversely, we ensured that it was not the case for novices.



Figure 4 – Screenshot of the Microsoft Word 97 standard icon bar. A number was attributed to each icon for identification purposes.

Conversion Scheme from an Iconic to a Verbal Form

As previously outlined, one of the most important problems about icon identification is how (and if) a given icon is “translated” into another format. For the purpose of this study, such translation was necessary, because LSA do not directly takes into account iconic descriptions (i.e., no use of picture recognition processes). Thus, each icon of the bar had to be translated into verbal propositions. In order to prevent variability of such coding and to have a high congruency between icon and their verbal description (Duval, 1999), an a priori conversion scheme was elaborated, as follows. This conversion scheme was applied independently by the two authors to each icon. Possible cases of disagreement between them were subsequently discussed and resolved.

- ☞ *No translation concerning the function of icons.* No link between the icon and its functional use within the software had to be held. For instance, an icon representing a printer had to be converted into “printer” and not to “print”.
- ☞ *No translation concerning colour.* Though all icons are coloured, we preferred not to code this feature, because it has no correspondence with other corpora (i.e., a blue printer behave strictly like a red one).
- ☞ *Code globally.* When parts of the icon can be identified, we preferred to code them globally than analytically. For instance, the icon number 15 (see Figure 4) had to be coded as a “table”, not as “several cells”; the magnifying glass was coded as such, and not as “a circle with a perpendicular bar”).
- ☞ *Order of reading.* The possible elements of an icon had to be coded as they were read (from left to right and from up to down).

☞ *Each icon had to be described coherently to each other.* For instance, if the icon number 0 is described as “a paper sheet with a folded corner”, this description had to appear verbatim for the icon number 4.

Simile vs. Figurative Icons

Once we have an overall coding scheme concerning the format in which the icons are translated, we need now a way to determine whether an icon is a *simile* or figurative. However, this discrimination is sometimes difficult to perform. Peraya (1998) presented several criteria for categorizing icons. He distinguished simple icons (one iconeme, e.g., the icons number 2 or 10), from “complex” ones, composed of several “iconemes”, or sub-elements (e.g., the icons number 12 or 18). More finely, Peraya analyzed the semiotic register involved in the “reading” of each icon. For instance, an icon can be exclusively composed of analogical elements (i.e., *simile*), or, as symbols, more arbitrary (i.e., figurative). Finally, some icons are mixed, integrating both *simile* and figurative elements. We categorized (*see Table*) each icon of the material upon the nature of the semiotic register used to identify it. Some icons were put apart, because of their shared nature (e.g., the icon number 10, “left arrow”, can be either *simile* (representing a back movement) or figurative (representing the function previously used).

Table 2 – A categorization of icons between simile, schemas, and indetermined.

<i>Simile</i> icons	0, 2, 3, 4, 6, 9, 14, 15, 17
Figurative icons	1, 5, 12, 13, 16, 18, 19, 20, 21, 22
Indetermined icons	7, 8, 10, 11

Processing

The description of each of the 23 icons was compared with its description in the help, using different combinations of knowledge bases as multidimensional spaces: each of them separately, and the combination of the three corpora (overall corpus) stand for expert knowledge base. We then compared the results obtained with those of the human experiment. For each comparison, LSA processed the semantic similarity between the verbal description of the icon and each candidate functionality (i.e., each functionality of the 22 remaining icons), by comparing their vectors in the multidimensional space. For instance, the verbal description of the icon 0 (New Document) was successively compared with each icon functionality. The highest similarity, above an arbitrary threshold was considered as the choice by LSA.

But all the results were not so straightforward (e.g., sometimes there were several values above the given threshold). Thus we categorized the result given by LSA as follows (partially after Goonetilleke et al., 2001):

☞ *no identification*, when no functionality was sufficiently close (related to an arbitrary threshold) to the icon description.

If it was the case, two other cases arose:

☞ *right identification*, when the functionality which was sufficiently close to the icon description was the right one;

☞ *confusion*, when the functionality sufficiently close to the icon description was not the right one.

Depending on the number of icon that are “identified” above the arbitrary threshold, two other categories arose:

☞ *ambiguous identification*, when several functionalities were above this threshold;

☞ *unique identification*, when only one functionality was sufficiently close to the icon description (either adequate or not).

Results

We present here the main results of this experiment, detailed by hypothesis:

Hypothesis 1. This hypothesis was not confirmed. The performances of the *AIR system* are very low when the sole description of icons is compared to the different functionalities. The highest rate is 22% identified icons. Results also show that the humans outperformed the *AIR system* on the task of icon recognition (about 22% for the overall corpus vs. 77% for novices and 83% for experts).

Hypothesis 2. This hypothesis is confirmed. We show that adding the tooltip information to the icon description interestingly improves the system performance (from 22% to 48%), while human performance with the tooltip assistance raises to 81% for novices and to 91% for experts. So, human performance varies in the same way as the *AIR system* does.

Hypothesis 3. This third hypothesis is partly confirmed. Unexpectedly, figurative icons were badly identified by the system; though most of the unidentified icons by humans were *simile*, for novices, and figurative, for experts. This result, as for the previous hypothesis, shows the same pattern of variations, when humans were compared with machine.

Hypothesis 4. This fourth hypothesis is confirmed, since the pattern of performances novices vs. experts is as expected: experts identified slightly more icons than novices (90.8 vs. 81 with tooltips; and 82.5% vs. 77% without the help of tool tips), and they were significantly favoured by the help of tooltips (90.8% vs. 82.5%, $t(20)= 2.7$; $p<.05$). This result confirm those of Richardson (2000) showing that only experts benefit from the use of tooltips.

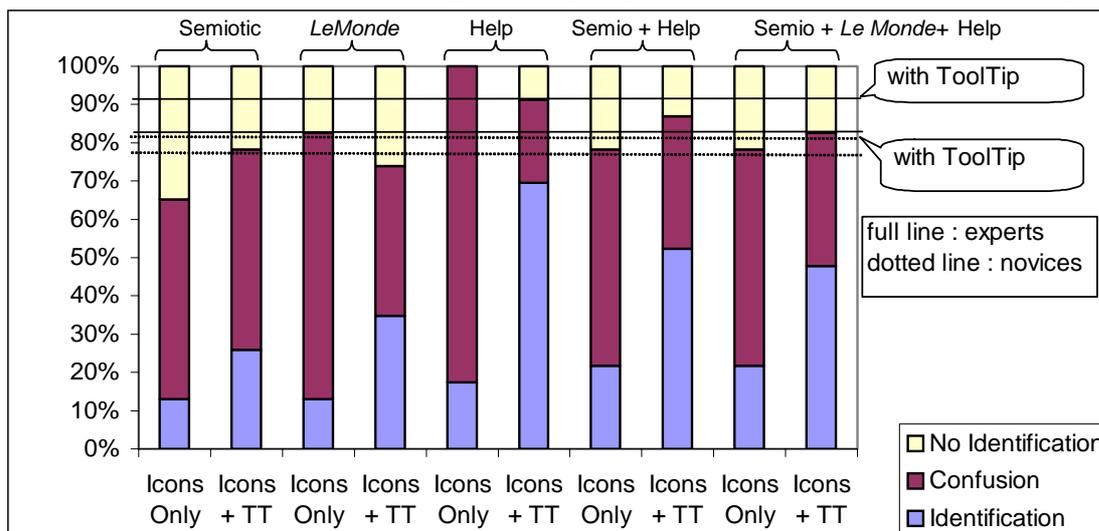


Figure 5 – Icons correctly identified, no identified, confused (% per corpus and with/out help of tooltips).
Caption: TT, Tooltip.

Discussion

The problem we tackled here is how to perform conversions between semiotic registers in order to identify software icons. We presented here some empirical observations showing that a) a tool can be appropriately used for assessing the polysemy of icons (allowing to perform the object-to-image relation); b) overall human performance in identifying icons can be simulated by a software (allowing to perform the image-to-function relation). However, some issues remain to be addressed. Firstly, the poor performances of the semiotic corpus, when added with the other corpora show that it have to be improved by selecting more appropriate texts. Secondly, the conversion scheme we elaborated, though efficient, has to be tested with more icons and raters, in order to assess its reliability. Thirdly, we plan to use eye-tracking techniques in order to understand more precisely how users successively grasp elements of an icon.

Cognitive Modelling of Icon Comprehension

A model of some cognitive processes engaged in the identification of icons can then be proposed. This model is mainly deduced from the results of the second experiment. In order to identify an icon, users may have to perform the following processes.

1. *Describe verbally the icon.* In order to perform this task, users need specialized knowledge on semiotics (e.g., knowing that “web” can be represented by “earth”, see icons 12 and 13).
2. *Retrieve from knowledge the corresponding function.* This retrieval can be fostered by tooltips only if the user has sufficient knowledge of the software. Thus, when users lack this kind of knowledge, the tool tips do not have any effect on identification.

Some Applications on Learning

We can now consider the application of these two tools for enhancing learning with computers. Firstly, the *Iconometer* can be used to control the validity of the interpretation of an icon (and of an image as well). This system can be used either in primary school contexts (for training pupils to read the adequate interpretation of an image), in university courses (for training students in semiotics to grasp difficulties in interpreting images), or by software designers who want to test the design of their icons. More generally, such a system could be used to help teachers analyze the different representations of their students about a phenomenon. Secondly, *AIR* can be used for automatically adapting the icons to the level of knowledge of its intended software users. In so doing, the software learning curve would be steeper than without such an adaptation. Thirdly, a coupling of both tools could tackle the main drawback of the *Iconometer*. For example, since two synonyms (i.e., house and home) are represented as two different hypotheses, *AIR* could be used to detect such a relation and to represent both words as a unique hypothesis.

The Symbol Grounding Problem

More generally, the different problems tackled here have to do with the symbol grounding problem, i.e., how meaning is grounded into perceptual experience. A debate opposed the proponents of a grounded approach of meaning to those of a computational one. We review the main arguments of this debate.

Barsalou (1999, see also Glenberg & Kaschak, 2003) claimed that simple conjunctions of symbols do not, by themselves, generate meaning: they have to be grounded in another “thing” (e.g., perceptual cues). This claim refers to the famous Searle’s (1980) Chinese room argument, which states that any intentionality (i.e., meaning) can emerge from a computer program manipulating abstract symbol systems. Some opponents to this view (e.g., Landauer, 2002) stated that a machine analyzing regularities within a corpus containing two different kinds of percepts (e.g., icon descriptions and software functionalities) can link them without any additional clue (i.e., rule). From this point of view, Shaw’s (2003) distinction between “grounding” and “situation” is helpful: “Grounding requires rules of reference that specify what the symbol denotes (refers to), whereas situating requires rules of usage that specify what the symbol connotes (means in context).” (id., p. 45). In addition, since one can situate a symbol which is not grounded, one cannot ground a symbol without situating it; thus, a symbol can be denoted—but not connoted—by a second hand experience.

Our second experiment showed results confirming the latter position, and icons can be described and computed with LSA (i.e., denoted) without any direct perceptual information. More generally, this study showed that one can *situate* an icon description (i.e., understand its context, or what it connotes) without having to *ground* it (i.e., knowing what are the objects it refers to, or what it denotes). Thus, situating the meaning of words can be carried out from second-hand, while grounding such meaning needs first-hand experience (e.g., shared attention, ostensive contexts, see Tomasello, 2003).

Acknowledgements

The second system was designed during a sabbatical of the first author at TECFA, university of Geneva, Switzerland. We would to thank: Benoit Lemaire and Cyril Rebetez for providing thoughtful comments on an earlier version of this chapter; the teachers, as well as David Delavier, Michèle Delavier, Stéphane Lattion and Stéphane Morand for their help in passing the tests of the second experiment; Bérénice Jaccaz for helping collect data; Catherine Frété for helping translate parts of this chapter, as well as Vera Britto for checking the english.

References

- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Barthes, R. (1985). *The responsibility of forms*. Berkeley: University of California Press.
- Clark, H. H. & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3, 472-517.
- Darras, B. (1996). *Au commencement était l'image [Let There be Image]*. Paris : ESF.
- Denis, M. (1991). *Image and cognition*. New York: Harvester Wheatsheaf.
- Dessus, P., & Peraya, D. (2005). Le rôle des connaissances dans l'identification automatique d'icônes: une comparaison avec des humains [The effect of knowledge in the automatic recognition of icons: a comparison with humans]. *Revue d'Intelligence Artificielle*, 19, 195-214.
- Dumais, S., Letsche, T. A., Littman, M. L., & Landauer, T. K. (1997). Automatic cross-language retrieval using Latent Semantic Indexing. *Paper presented at the AAAI Symposium on Cross-Language Text and Speech Retrieval*.
- Duval, R. (1995). *Sémiosis et pensée humaine [Semiosis and Human Thinking]*. Bern: Lang.
- Duval, R. (1999). Representation, vision and visualization : Cognitive functions in mathematical thinking. Basic issues for learning. *Proc. Annual Meeting of PME Society*. Morelos (Mexico).
- Eco, U. (1976). *Theory of Semiotics*. Bloomington: Indiana University Press.
- Evers, V., Kukulska-Hulme, A., & Jones, A. (1999). *Cross-cultural understanding of interface desing: A cross-cultural analysis of icon recognition*. Paper presented at the Proc. Int. Workshop on Internationalisation of Products and Systems, Rochester.
- 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.
- Goguen, J. (1999). An introduction to algebraic semiotics, with application to user interface design. In C. Nehaniv (Ed.), *Computation for Metaphor, Analogy, and Agents* (pp. 242-291). Berlin: Springer.
- 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 μ (1992). *Traité du signe visuel. Pour une rhétorique de l'image [Treatise of the visual sign. Towards a rhetoric of the image]*. Paris: Seuil.
- 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.
- Holloway, J. B., & Bailey, J. H. (1996). Dont' use a product's developers for icon testing. *Proc. Conf. CHI'96*. Vancouver.
- Kindler, A. M. (2004). Researching impossible? Models of artistic development reconsidered. In W. E. Eisner & M. D. Day (Eds.), *Handbook of research and policy in art education* (pp. 233-252). Mahwah: Erlbaum.
- 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.
- Leclercq, D. (1992). *Audio-visuel et apprentissage [Audiovisual and learning]*. Liège : Liège University, Educational Technologies Service, Unpublished report.
- Liu, K., Crum, G., & Dines, K. (1998). Design issues in a semiotic description of user responses to three interfaces. *International Journal of Behaviour & Information Technology*, 17(3), 175-184.
- Meunier, J.-P. & Peraya, D. (2004). *Introduction aux théories de la communication. Analyse sémio-pragmatique de la communication médiatique [An Introduction to Communication Theories. Semio-pragmatics Analyse of the Mass Media-Based Communication]*(1st ed., 1993). Bruxelles : De Boeck.
- Moles, A. A. (1981). *L'image communication fonctionnelle [Image and Functional Communication]*. Tournai : Casterman.

- Morand, S. (2005). *Iconometer Site Web* (version 2). Retrieved November 9, 2005 from <http://tecfa.unige.ch/sip/index.php?module=pnIconometer&func=main>
- 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.
- Ott, D. (1998). *Iconometer Site Web* (1st version). Retrieved November 9, 2005 from <http://tecfa.unige.ch/campus/infospace/index.php?display=rooms&id=130601>
- Peraya, D. (1998). Structures et fonctionnement sémiotiques des icônes de logiciels et d'environnements informatiques standardisés (ILEIS) [Structures and semiotic functions of software icons and ILEs]. *Recherches en Communication*, 10, 101-140.
- Peraya, D., & Strasser, D. (1999). L'iconomètre : un outil de formation et de recherche pour mesurer le degré de polysémie des représentations visuelles [The Iconometer: a tool for training and research for measuring how polysemic are visual representations]. *Paper presented at the Fifth Colloque Européen sur l'Autoformation [European Self-Education Conference]*, Barcelona (Spain).
- Piamonte, D. P. T., Abeysekera, J. D. A. & Ohlsson, K. (2001). Understanding small graphical symbols: a cross-cultural study. *International Journal of Industrial Ergonomics*, 27, 399-404.
- 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.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457.
- Shaw, R. (2003). The agent-environment interface: Simon's indirect or Gibson's direct coupling? *Ecological Psychology*, 15(1), 37-106.
- Simitsek Inau (2002). *Développement d'un iconomètre [Implementation of an Iconometer]*. Geneva: TECFA, Université de Genève, unpublished Master thesis in psychology.
- Tomasello, M. (2003). *Constructing a Language, a Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Westerveld, T., Hiemstra, D., & de Jong, F. (1999). Extracting bimodal representations for language-based image retrieval. *Paper presented at the Multimedia'99 Conference*, Vienna (Austria).