

Chapitre 10¹

Evaluation de documents par oculométrie

10.1. Introduction

L'idée développée dans ce chapitre est d'analyser la trace du regard dans l'activité de lecture d'un document pour en tirer un diagnostic sur la pertinence de ce document, aux niveaux de la structure et du contenu.

Les temps ne sont pas éloignés où l'expression écrite (et son corollaire la lecture) était fondamentalement associée à un support physique unique : le papier. Aujourd'hui un nouveau chantier est apparu, c'est celui de la « dématérialisation » de ces supports papier avec pour objectif d'en extraire et d'en conserver automatiquement l'information pertinente. Cette mise sous forme virtuelle est faite un peu à l'image de ce que fait l'homme, qui par sa capacité à lire et à comprendre, parvient à identifier puis à garder en mémoire une trace de l'information perçue [BON 89].

Comment cette capacité à lire et à comprendre peut-elle être observée et analysée ? Il semble généralement admis [RAY 92] que les mouvements oculaires d'un sujet engagé dans une tâche de lecture, reflètent bien les processus cognitifs à l'œuvre dans cette activité. L'enregistrement des mouvements oculaires fournit une trace spontanée, directe et mesurable de l'attention du sujet en action.

En Sciences Humaines, l'étude des mouvements oculaires est souvent utilisée d'une part comme un outil d'observation des témoins d'une activité mentale ou d'un fonctionnement spatial, et d'autre part comme un moyen de vérification des modèles comportementaux. Dans l'étude de l'activité de lecture, par exemple, l'analyse de l'emplacement, du nombre et de la durée des fixations oculaires, de l'amplitude des saccades et des variations de la direction du regard, sont autant d'indices des opérations mentales de traitement des informations visuelles qui se déroulent, d'une

¹ Ce chapitre a été rédigé par J. CAELEN, V. EGLIN et S. HOLLARD

façon automatique ou contrôlée. Depuis l'"eye-mind hypothesis", émise en 1980 par M. Just et P. Carpenter [JUS 80]², l'observation des mouvements oculaires s'est affinée, et a permis, par exemple, de mettre en lumière une stratégie globale de déplacement du regard vers une position optimale, en fonction d'indices uniquement visuels, et des tactiques locales de réajustement en fonction d'interprétations plus élaborées de données visuelles et linguistiques. C'est la théorie dite "stratégie et tactiques", proposée par le groupe Regard de l'Université Paris V [O'R 90].

Le parcours oculaire se caractérise :

- par un aspect temporel, le temps de fixation sur les différentes parties du texte, le temps de lecture total,
- par un aspect spatial : taille des empan, localisation des points de fixation,
- par un aspect stratégique : nombre des fixations, retours arrière.

Ces caractéristiques dépendent en grande partie du type de texte lu. Si l'on connaît le type de texte proposé, ce qui est généralement le cas dans la lecture de pages Web, alors la trajectoire oculaire apporte des indications sur la façon dont le sujet en construit le sens. Par exemple, un temps de fixation très long³ sur une même zone du texte, une fréquence élevée des retours arrière⁴, ou une réduction de la taille des empan⁵ indiqueront des difficultés de construction du sens. Ces difficultés peuvent provenir 1) du lexique utilisé : termes insuffisamment définis, ou ambigus, etc., 2) de la construction syntaxique, par exemple référents difficilement identifiables, enchâssement de relatives, 3) des marques énonciatives : insuffisance de ponctuation, de connecteurs, etc., ou encore 4) de l'expression elle-même : manque de cohérence dans l'enchaînement des énoncés, information manquante, etc. L'analyse des mouvements oculaires de sujets engagés dans une recherche sur le Web va donc permettre d'évaluer, au delà de leur propre style cognitif dans cette tâche, le degré de difficulté d'interprétation présenté par la tâche elle-même.

Les documents électroniques sont des objets structurés, qui ont un contenu. Celui-ci est souvent multimédia c'est-à-dire textuel et imagé (figures, tableaux, images, croquis, etc.) et la présentation est hiérarchisée (titres, sous-titres, blocs de texte, paragraphes, etc.). Lorsqu'on lit un document on opère à plusieurs niveaux (perception, compréhension, interprétation) en utilisant des stratégies diverses : dans un premier temps on peut effectuer un *survol* pendant lequel la structure du document est perçue, certains titres ou figures étant globalement analysés, puis dans un second temps une *lecture* fine de certaines parties jugées pertinentes est mise en œuvre. Ces deux phases peuvent d'ailleurs interférer de manière itérative ou

² Selon l'"eye-mind" hypothesis, le traitement d'un mot se déroule exactement pendant la fixation de l'œil sur ce mot. La fin du traitement déclenche la fixation suivante.

³ Une fixation dure entre 100 et 500 millisecondes (ms). La durée moyenne est de 250 ms.

⁴ La proportion normale de retours arrière, ou « régressions », n'est que de 10 à 15 % des saccades.

⁵ La taille normale d'un l'empan est d'une vingtaine de lettres.

imbriquée. En retour, dans le cadre de l'évaluation d'un document, on peut être tenté d'utiliser les informations portées par la trajectoire du regard d'une manière diagnostique : repérer les points de focalisation, les zones d'intérêt qui peuvent donner de précieux renseignements sur la structure du document, suivre les points de fixation et leur densité, les points de reprises, etc., qui peuvent permettre de mesurer l'efficacité de l'acte de lecture pour améliorer la qualité de rédaction.

Ce chapitre présentera d'abord quelques données sur l'oculométrie, puis trois recherches illustrant l'intérêt de cette technique : évaluation de la structure de documents électroniques, évaluation de sites web, et construction du sens.

10.1.1. Mouvements oculaires et vision

La motricité du regard se traduit par deux familles de mouvements. Les mouvements oculaires involontaires contribuent à stabiliser sur la rétine l'image perçue par le regard, et à coordonner le regard des deux yeux. Les mouvements oculaires volontaires permettent la capture ou le suivi d'une cible, ils assurent le déplacement du regard dans l'espace. Les mouvements saccadiques, qui déplacent l'œil entre deux fixations, jouent un rôle primordial dans la compréhension de la perception des mots ou des images [RAY 92].

Le champ récepteur oculaire comprend trois zones : fovéale, parafovéale et périphérique. La zone fovéale, limitée aux deux degrés de vision centrale (1 degré de part et d'autre du point de fixation), détient la capacité de percevoir les détails fins grâce à la concentration des cônes récepteurs de la rétine (photorécepteurs privilégiés de la vision diurne). La zone parafovéale englobe environ 5° de part et d'autre du point de fixation. Au delà de ces 5°, on parle de vision périphérique

Une information ne peut être extraite des objets que lorsque ceux-ci sont fixés ou sur le point de l'être. Pour que le déplacement de l'œil soit programmé sur un objet qui n'est pas placé sur la fovéa (i.e. en vision centrale), ce dernier doit présenter un fort signal rétinien. Il est alors jugé important, ce qui permet la programmation de la saccade vers cette nouvelle cible. Au fur et à mesure que la direction du regard rejoint la cible, l'activité se déplace vers l'axe fovéal. La distance qui lui reste à parcourir pour se placer au centre de la fovéa représente l'erreur motrice.

Les objets se trouvant dans des zones extra-fovéales jugées informatives (transmettant des signaux rétiniens forts et nombreux) seront naturellement fixés, que cet objet soit transparent (laisse voir l'arrière plan) ou opaque. Ces zones informatives peuvent regrouper des objets catégorisés en vision extra-fovéale

comme sémantiquement reliés à la recherche en cours ou comme visuellement marquants (fortes variations de luminosité, de contours, etc.). Les travaux de Balota [BAL 95] et de Underwood [UND 90] ont montré qu'une certaine information est extraite de la zone parafovéale : par exemple, des mots s'y trouvant et ne faisant pas l'objet de fixations sont tout de même perçus. Cette information, principalement sémantique, est soit suffisante pour être traitée, soit jugée ambiguë ou incomplète et permet de programmer une fixation.

La fixation initiale suffit à dégager une information globale sur le sens de la scène affichée et à déterminer les zones d'intérêt qui seront parcourues. Dans les instants qui suivent, la taille de la saccade diminue et la durée de la fixation augmente avec la densité d'informations présentées à l'affichage et la difficulté de la tâche impartie.

Le mouvement saccadique, ou saccade, dure de 100ms pour une amplitude de 40 degrés d'arc, à moins de 25ms pour une amplitude de 10 degrés d'arc. La relation durée/amplitude est linéaire. Ces déplacements représentent sensiblement 10% du temps de balayage d'une image. Une saccade de 25ms est suivie d'une fixation, en moyenne de 250ms. Les saccades vers un point donné sont commandées soit par une réaction réflexe due à l'apparition d'un nouveau signal rétinien dans le champ extrafovéal, soit par les nécessités de structuration de l'information. La saccade présente un délai de déclenchement compris entre 150 et 300 ms. Le délai est plus court lorsque le mouvement s'effectue vers le haut, et/ou vers la droite.

Comme dans tout système de traitement, dans le système oculomoteur il existe un délai entre l'arrivée d'un signal et l'action commandée par sa sortie. Biologiquement, ces délais sont compensés au niveau central par un modèle dit d'adaptation de gain. De plus, certaines ambiguïtés de perception visuelle sollicitent les représentations mentales du sujet. Robinson [ROB 75] a proposé un modèle de fonctionnement du système oculomoteur saccadique dont le schéma de principe s'applique aussi bien aux captures fovéales d'un objet qu'aux transformations visuomotrices.

10.1.2. Vision et cognition

Dans la prise d'information par le regard sont convoqués deux processus : l'un périphérique, celui de la perception (incluant la reconnaissance des formes) proprement dite et l'autre central, celui de la cognition (rappel mémoriel, compréhension au plan sémantique et interprétation au niveau pragmatique).

Il existe trois familles de théories concernant le traitement et la reconnaissance visuelle en fonction de la nature du stimulus (objet, scène ou visage), qui prennent

en compte la tâche demandée au sujet (simple catégorisation ou identification) [MAR 00]. Quel que soit le type d'objet, les expériences montrent que le temps d'identification du stimulus s'avère peu sensible :

- au type de présentation de l'objet (dessin, photo, croquis), pourvu que celle-ci conserve les informations relatives à la structure de l'objet [BIE 87, 95] ;

- aux translations et aux dilatations subies par l'objet. En revanche, les rotations 2D ou 3D du stimulus influent fortement sur la rapidité de l'identification.

L'activité cognitive produit des images mentales qui présentent des ressemblances fonctionnelles et structurales avec l'activité perceptive visuelle. Les représentations mentales sont de deux types : forme abstraite assimilée à un codage propositionnel de l'information, et forme analogique de l'objet représenté. Cette dernière est une image mentale qui conserve les propriétés structurales et fonctionnelles de l'objet. Ainsi, l'objet réel perçu, le document visuel ou image matérielle qui le représente, le percept ou image née de la perception (qui peut n'être que partielle), et l'image mentale détiennent des caractéristiques communes qui correspondent à la même structure générale de l'objet.

10.1.3. *Regard et structure du document*

Une image de document (comprise au sens large de texte et d'image), à la différence d'une image quelconque, possède une structure et des niveaux d'information bien hiérarchisés. L'auteur a organisé la structure du document en régions d'intérêt [DOE 96]. La recherche d'information par un lecteur sur ce document n'est donc pas aléatoire, elle résulte de stratégies exploratoires de ces zones d'intérêt bien spécifiques au document. On distingue : la lecture complète (de type lettres, articles, romans, etc.), le survol (correspondant à la lecture dite familièrement en diagonale des journaux, des magazines, des catalogues, etc.) et enfin l'inspection ciblée d'une région particulière en vue d'une recherche intentionnelle d'information (dictionnaire, encyclopédie, manuel, guide, etc.). Le processus de recherche d'information nécessite donc la connaissance de plusieurs facteurs : le type de document traité (document technique, scientifique, publicitaire, manuscrit, etc.), les styles de mise en page et l'organisation logique des entités présentées. Ainsi, sur la base de l'analyse des mécanismes de capture d'informations chez l'homme, il est intéressant de chercher à obtenir une description du document liée aux déplacements successifs de l'œil, mettant l'accent sur les zones informatives et attractives, [EGL 97].

10.2. Évaluation de la structure du document

Pour évaluer l'organisation spatiale du document (ou structure), nous avons choisi d'exploiter la stratégie exploratoire de *survol*. La base des documents de tests a été la plus diversifiée possible. Le système de structuration des données qui a été développé procède par *anticipation du regard* ; il permet d'établir des liens entre les différentes zones d'un document sur lesquelles, à un instant donné, toute l'attention est focalisée. Pour cela, il a fallu adapter pour l'exploration des documents, différents éléments liés à la perception visuelle humaine. En redimensionnant le champ visuel et en l'adaptant à l'espace du document, il a fallu reconsidérer certaines propriétés fondamentales de la perception visuelle au cas particulier de la stratégie de survol. Cette tâche a première vue très simple met en jeu des mécanismes de repérage d'informations très complexes, à partir desquels nous avons retenu une hypothèse fondamentale : il existe chez l'homme une émergence de l'information par zones d'intérêt, matérialisées sur l'exemple de la figure 10.1 par des zones de fixations. Plus précisément encore, on s'aperçoit sur un grand nombre de résultats que les fixations sont localisées pour la plupart dans les zones d'images (grandes illustrations ou icônes), de titres et de sous-titres, c'est-à-dire dans les zones que le rédacteur du document aura particulièrement soignées et mises en évidence (taille des caractères, graisses, espaces inter-lignes, couleur, etc.).

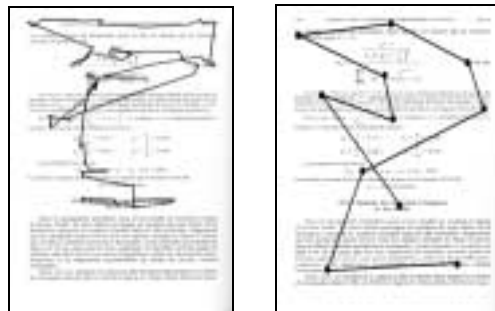


Figure 10.1 : Exemple de mesures oculométriques d'observateurs humains dans un contexte de survol du document (exemple réel à gauche et simulé à droite).

L'exploration est décrite par le trajet oculaire du regard sur le document et la segmentation se traduit par un découpage en zones d'intérêt, découpage d'autant plus fin que les régions explorées sont nombreuses.

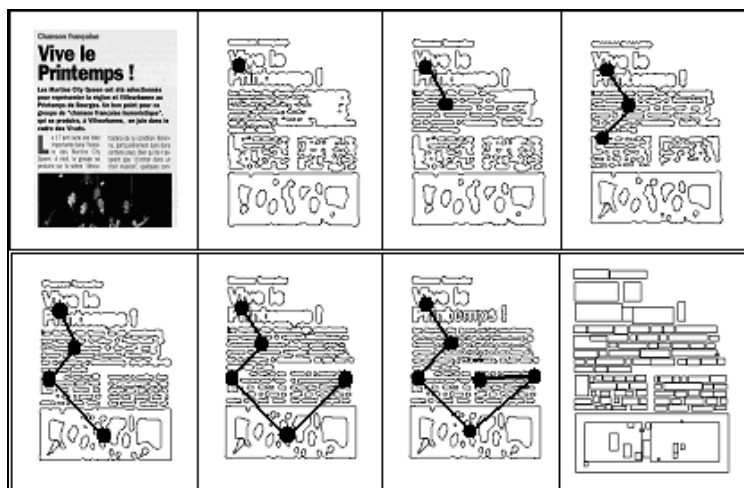


Figure 10.2 : Convergence des fixations vers une représentation unifiée où la succession des fixations aboutit à une reconstruction de la page.

La segmentation du document repose sur une étape fondamentale de prétraitement (où des caractéristiques de bas niveau de l'image du document sont extraites) aboutissant à une esquisse de segmentation physique. Cette première étape concerne la simulation des caractéristiques de la rétine en utilisant des séries de filtrages fréquentiels passe-bas par transformées de Fourier et en particulier les filtrages directionnels de Gabor qui privilégient les directions horizontales et verticales des textes. Le principe consiste à choisir un point de fixation (déterminé à partir de caractéristiques locales de l'image) puis à filtrer l'image en ce point en diminuant progressivement la fréquence de coupure avec l'éloignement au centre de fixation. On parvient ainsi à une représentation de l'image centrée en un point de fixation particulier correspondant à un centre d'intérêt du document avec une décroissance de précision du centre à la périphérie de l'image (figure 10.2.). A partir d'une série de descriptions centrées en différents points de fixation on obtient un parcours visuel sur l'image et une reconstruction possible de l'unité physique du document. A partir du découpage visuel des différentes entités du document, la récupération des différents niveaux informatifs et logiques de la page devient alors possible.

10.2.1. Etiqueter les zones informatives par une analyse typo-dispositionnelle

La recherche d'information sur les documents faite par le lecteur dépend à la fois d'un objectif de recherche (ou d'une consigne) et de l'organisation des données

(étape de segmentation). Pour trouver rapidement une caractérisation visuelle de ces données et mettre en évidence les régions d'intérêt du document, il s'est avéré indispensable en plus de l'étape de segmentation précédemment exposée, de mettre en place une méthode de classification des écritures imprimées. Cette caractérisation va reposer sur une analyse de texture qui renseigne sur l'aspect du tracé des polices, sur leur fréquence d'apparition, et sur leur graisse, [EGL 01]. En considérant le texte comme une texture, on définit le caractère comme son entité élémentaire. Une page de texte peut être considérée comme un ensemble de petits graphismes, les lettres, qui génèrent une impression « macroscopique » de texture. Les caractéristiques visuelles de cette texture dépendent de la disposition des lettres, de leur fréquence d'apparition, de la police, de la graisse, de la présence ou non d'italique, de la langue, [RIC 89]. Nous ne nous intéressons pas au sens contenu dans le texte, mais à sa forme, à l'impression visuelle qui en ressort et nous analysons là encore la page sous sa forme image. Dans ce contexte, nous proposons l'usage de paramètres statistiques, capables de mettre en évidence certaines caractéristiques de la police du texte observé et de faire une distinction aisée entre les éléments dominants tels que les titres, ou les sous-titres, les paragraphes plein texte et les objets discrets mais informatifs tels que les notes ou les lignes isolées. Nous proposons pour chaque page analysée de baser l'étiquetage des zones de texte sur l'interprétation de mesures statistiques de complexité, de densité et de visibilité des polices de caractères. C'est à partir des boîtes englobantes de chaque zone de texte que nous procédons à cette analyse.

La première étape du processus d'étiquetage des textes est directement issue de la phase initiale de segmentation de l'image du document en blocs homogènes présentée précédemment. A ce niveau, on ne s'intéresse qu'aux données textuelles, celles dont l'orientation privilégiée est horizontale. Le principe général de l'analyse des blocs de textes par l'image et la texture est illustré à la figure 10.3.

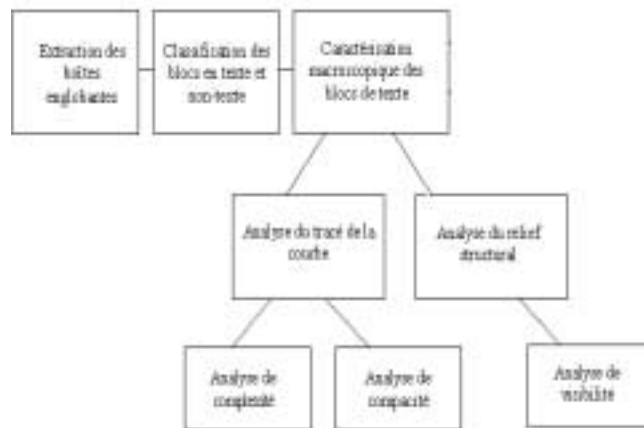


Figure 10.3 Principe des traitements pour l'étiquetage fonctionnel des blocs.

Les quatre attributs de texture que nous avons retenus sont à l'origine du classement des textes en trois familles fonctionnelles (les titres, les paragraphes plein texte et les notes ou lignes isolées). Il s'agit tout d'abord de la complexité, qui est quantifiée par une mesure d'entropie effectuée sur le bloc du texte. On peut dire que plus une courbe est complexe, plus la valeur de l'entropie est élevée. Les compacités horizontale et verticale sont des densités directionnelles correspondant respectivement aux transitions horizontales et verticales mesurées sur le tracé des polices. La dernière primitive correspond à la visibilité des blocs qui peut être vue comme une caractérisation du relief structural des polices (graisse).

Pour chaque bloc, nous procédons d'abord à l'évaluation de la valeur numérique des attributs, puis à l'intégration des valeurs dans un espace où tous les résultats issus de l'ensemble des blocs de la page sont représentés. L'espace nous renseigne sur la distribution relative des blocs les uns par rapport aux autres en fonction de l'attribut considéré. Le label définitif d'un bloc de texte résulte de la mise en commun des résultats des quatre attributs pour ce bloc et permet de donner une description qualitative très complète de sa nature et de sa fonction dans la page.

Pour valider cette approche, notre intérêt s'est porté sur les documents composites tels que les journaux, les articles scientifiques normalisés de conférence internationale, et des premiers travaux sont actuellement réalisés sur des pages Web. Sur une base de tests de 120 documents, cette analyse associée à la simulation de

l'exploration rapide de la page conduit à des résultats de segmentation et d'étiquetage très prometteur. Les stratégies exploratoires très diversifiées devraient nous conduire à considérer de nouveaux paramètres davantage liés à une connaissance a priori des pages traitées.

10.2.2. *Evaluation comparative des sites Web*

Les travaux portant sur la lecture d'image ont mis en évidence trois types de parcours oculaire que l'on retrouve dans le cas des sites Web :

10.2.2.1 Les schémas de lecture

Balayage en « Z » :

Le lecteur lit le document suivant un parcours en Z, quitte à revenir, par la suite, à un point perçu comme éventuellement intéressant. Ce modèle s'explique, chez les occidentaux, par un facteur culturel : l'écriture se déroule de gauche à droite et de haut en bas. Cette théorie conduit à accepter que, pour d'autres langues, l'habitude d'écriture et de lecture influence l'enchaînement des saccades.

Balayage « hiérarchique »

Ce modèle se base sur l'hypothèse que les lecteurs construisent leurs représentations à partir de composantes de trois types : vivantes, mouvantes et stables rattachées aux catégories, (a) êtres vivants, (b) objets inanimés mobiles et (c) objets inanimés statiques. Lors de la lecture d'une image fixe, l'ordre de balayage est prédéterminé par la hiérarchie de ces trois composantes. Le lecteur fixe en premier les êtres humains ou animaux, puis les objets capables de mouvement (voitures, eau, nuages, etc.), et enfin les objets restant statiques (maison, montagne, etc.).

Balayage en « spirale »

En partant du coin supérieur gauche, le lecteur parcourt l'image selon un balayage en spirale, tout en se rapprochant de son centre. Initialement proche des bords, le rayon de la trajectoire diminue au fil des tours jusqu'à ce que toute la surface ait été parcourue. Lors d'enregistrements de parcours du regard, ce modèle est plus rarement observé, hormis lorsque l'image affichée présente des champs circulaires. L'attrait de ce modèle réside dans le rapprochement que l'on peut opérer avec les raisonnements de l'esprit humain : la déduction qui va du général au particulier, du principe à la conséquence, s'opposant à l'induction qui va du particulier au général, des faits aux lois. Dans ce cadre, le lecteur recherche

initialement une compréhension globale de la scène avant d'en déterminer les différentes parties.

10.2.2.2 Organisation des sites commerciaux

D'une façon générale, les pages d'accueil de la plupart des sites Web à vocation commerciale peuvent être regroupées en quatre catégories, par rapport à la façon dont est organisé l'accès à l'information pour le client :

- C1 : il y a un moteur de recherche ou des listes de liens vers les catégories principales d'information. Le concepteur y privilégie le concept de tâche ;

- C2 : il y a un niveau métaphorique (icônes, photos, etc.) pour accéder aux catégories principales d'information. Le concepteur y privilégie le concept de tâche activé par des représentations ;

- C3 : il y a des menus fournissant tout ou partie de l'arborescence des catégories (ex : affichage du premier niveau de profondeur correspondant aux catégories principales, et des pages situées en profondeur sur le même axe). Le concepteur privilégie l'accès structuré aux informations et l'exhaustivité ;

- C4 : il y a des zones hybrides associant deux ou plusieurs modes d'accès précédents (par exemple, menus et métaphores des catégories principales).

La présentation de ces moyens d'accès peut elle aussi être classée en quatre grands types :

- P1 : présentation linéaire (comme un texte illustré),

- P2 : présentation imagée (comme une image intégrant graphisme et texte),

- P3 : présentation grille (avec des zones d'intérêt),

- P4 : présentation escamot (avec des textes volatils).

Ce qui, combiné avec les catégories Ci, constitue une variété de 16 types de pages.

Nous avons sélectionné 10 sites de commerce électronique appartenant au domaine du tourisme, recouvrant une partie de ces 16 types. Ces sites proposent une information de type agence de voyage en langue française. La conception de ces sites est centrée *produit*, *client* ou *mixte*.

- En approche *produit* la page d'accueil présente tous les types de produits et services disponibles (cette approche est aussi dite *centrée institution*),

- En approche *client* le concepteur du site privilégie un certain type de client à travers ses contraintes de choix ou ses centres d'intérêts,

- L'approche *mixte* tente de répondre aux diverses préoccupations et buts de recherche du client tout en cherchant à lui montrer, dès la page d'accueil, le plus grand nombre possible de produits.

10.2.2.3. Expérimentation et résultats

Les paramètres déterminant le but du sujet sont limités à trois : la date, le prix, la destination. Les sujets reçoivent consécutivement les consignes suivantes, la première plus précise que la seconde :

1 Je cherche un vol pour New York à moins de 2500 francs pour la dernière quinzaine de mai

2 Je prépare un voyage exceptionnel pour cet été

Les sujets examinent tous les sites, dans un ordre aléatoire. [LEV 88]. Nous avons considéré les tâches comme variables indépendantes, le nombre et la durée des fixations, la répartition des durées des fixations, et la somme des temps de regard sur les divers éléments comme variables dépendantes. Les saccades et fixations « caractérisent les changements de visée dans l'examen d'une image » Les parcours oculaires sont étudiés en enchaînements de points de mesures puis de fixations de durée $\geq 250\text{ms}$ ⁶.

Il y avait deux grands types de sujets selon leur niveau de pratique d'Internet : les experts et les novices. Indépendamment du but de l'expérimentation, l'enregistrement de la trajectoire oculaire a permis de montrer que les novices sont davantage portés à regarder les zones d'images. Il apparaît également que les stratégies d'exploration sont mises en œuvre par les utilisateurs indépendamment de la tâche à accomplir.

Quel que soit le sujet, les fixations supérieures à 500ms sont rares (moins de 20%) et peuvent être utilisées pour marquer son intérêt ou son incompréhension vis-à-vis d'un mot ou d'un objet.

⁶ Sur le plan méthodologique, le choix d'une durée minimale de fixation à 250ms nous assure que le sujet a le temps de traiter partiellement l'information capturée par son regard avant la fixation suivante, donc de la diriger dans un but exploratoire, même si des études [LOF 81] ont montré que 100ms constitue un seuil assurant une reconnaissance fiable de l'image.



Figure 10.4 : Exemple de représentation en points de mesure

Deux grandes familles de comportements se dégagent des parcours enregistrés :

1 Le sujet cherche davantage à identifier une cible précise, qui correspond au but de sa recherche, qu'à découvrir le sens complet du contenu de la page. Il la parcourt, prélève des échantillons et localise l'information qui l'intéresse. L'étude de la trajectoire du regard sur le document électronique et l'analyse des points de fixation de ce type de lecteur permettent de reconstruire la structure du document et d'obtenir différents niveaux informatifs.

2 Le sujet examine l'ensemble du document et recherche tous les accès à l'information possibles. Les temps de parcours sont longs en général dans ce cas qui vise l'exhaustivité. Ce type de lecteur, plus rare, est intéressant pour étudier par exemple la sémantique des titres de rubriques, les notes, légendes, etc..

Comportements des sujets

L'internaute novice ou occasionnel regarde la plupart ou la totalité des photos. Il les considère souvent comme des liens vers l'information recherchée. L'internaute expert ne regarde que les photos qui peuvent cacher un lien ou une information métaphorique.

Sur 64 parcours observés :

18 peuvent être assimilés à un balayage en Z,

11 peuvent être assimilés à un balayage hiérarchique,

5 peuvent être assimilés à un balayage en spirale,

30 ne peuvent être rapprochés d'aucun modèle connu, ce qui montre que l'exploration de pages Web ne s'apparente pas toujours à celle de documents classiques.

L'analyse de la séquence des 10 premières fixations montre généralement que, lorsque les fixations changent d'objet (icône, titre,...), chacune s'enchaîne dans la zone parafovéale de la précédente. Ceci confirme l'importance du traitement basse résolution dans la définition de la cible suivante.

L'influence de la présence d'un moteur de recherche est importante : dans le cas d'une présentation explicite, le sujet peut s'y attarder puis l'utiliser ou non. Un moteur qui n'apparaît que partiellement à l'écran n'est que très rarement fixé. Il passe inaperçu..

Sites	L'utilisent	Ne l'utilisent pas	Ne s'y attardent pas
X	65%	10%	25%
Y	31%	26%	42%
Z	60%	0%	40%

Tableau 10.1 : Comportement des sujets vis-à-vis d'un moteur de recherche

Les types de catégorisation et d'approche de conception utilisés ne constituent une limite ni aux possibilités ou rubriques présentées au client, ni à l'ergonomie des composants de la page. Ils sont donc équivalents du point de vue de la quantité d'information. Cependant, du point de vue qualitatif, la présentation « grille » est moins bien perçue par les sujets car elle a tendance à être « trop chargée » en informations. La présentation « linéaire » recueille leurs préférences lorsqu'elle est associée à une approche « client ». Le choix d'une structure « imagée » ou « escamot » ne semble pas constituer un paramètre pertinent sur les préférences des utilisateurs. L'abus de photos ou d'images nuit au temps de recherche et n'apporte pas toujours à l'esthétique.

Analyse ergonomique

Nous allons tenter de répondre à quelques questions :

- 1 comment mesurer la complexité informationnelle de la page ?
- 2 comment préconiser un bon agencement des éléments d'information ?
- 3 comment évaluer rapidement la qualité d'un document ?

L'analyse statistique du temps d'observation d'une page ne permet pas d'obtenir un résultat significatif ($F=3,63$; $p=0,073$). De plus, cette variable intègre une erreur due au comportement de certains sujets qui fixent très longtemps le lien choisi avant de l'activer. En revanche, l'analyse du nombre de fixations nécessaire à l'exploration de chaque page d'accueil fournit une différence significative entre les catégories de sujets pour une comparaison de différentes pages ($F=5,45$; $p=0,031$).

En revanche, le temps d'exploration de la page semble être un bon révélateur de la densité des informations ou complexité informationnelle de la page. Le temps d'exploration est corrélé au nombre de fixations lui-même lié à la quantité des rubriques et à la compréhension du moteur de recherche s'il y en a un.

Le nombre de fixations est sensible à la nature de l'information présente : il existe par exemple sur un site, une rubrique dont le titre représente exactement le but de la première consigne (le fait de disposer d'une rubrique « vols »), pour lequel le nombre de fixations est évidemment faible.

Le découpage en zones dont le contenu, tout en correspondant à un découpage fonctionnel de l'espace d'exploration, revêt un caractère sémantique pour le client, facilite le repérage et diminue le temps de recherche (trajectoires du regard plus courtes). Il est bon de pouvoir repérer sans ambiguïté :

- Z1 : l'accroche (sens publicitaire du terme),
- Z2 : le logo/signature,
- Z3 : les produits,
- Z4 : les questions des utilisateurs (FAQ),
- Z5 : les autres zones de la page (promotions, références, etc.).

La perception initiale d'une zone comme informative est favorisée par la présence d'un nombre limité d'éléments au sein de celle-ci, ainsi que par la présence d'un objet clairement identifiable ou prototypique (ex : moteur de recherche). Au sein de l'espace occupé sur l'écran, la distribution spatiale ou densité de l'information dans les zones joue un rôle important.

La distribution des fixations (en %) dans chaque zone, permet de comparer, au sein d'un échantillon, l'intérêt suscité par les différentes zones durant les premiers instants d'exploration puis tout au long de l'exploration. Le début de l'exploration est particulièrement intéressant pour mesurer l'efficacité de la présentation.

Associées à un outil statistique, ces exploitations permettent, lors de la conception ou de l'évaluation de plusieurs versions d'une page :

d'évaluer la rapidité avec laquelle les internautes trouvent les liens,

de comparer l'intérêt accordé par les internautes à différentes zones ; initialement à travers la distribution de la première fixation entre les différentes zones, puis de façon continue à travers l'évolution de ces distributions jusqu'à un seuil de représentativité.

La tâche a une forte influence sur la recherche d'information et en particulier pour la validation des catégories d'informations principales, représentées souvent par des liens. Il y a donc lieu de valider l'analyse ergonomique avec un scénario complet de tâches.

En résumé,

Le tracé en points de mesures livre le parcours précis du regard, et fournit une première estimation des zones auxquelles l'internaute s'intéresse,

Les fixations de 100ms dévoilent les objets de la page susceptibles de présenter un intérêt,

Les fixations de 250ms permettent de savoir quels objets ou quelles informations ont été identifiées comme porteuses d'intérêt par l'internaute. Une information jugée essentielle par le concepteur doit faire l'objet chez le sujet d'une ou de plusieurs fixations; inversement une information non fixée peut-être éliminée,

Le temps d'exploration de la page semble être un bon révélateur de la densité des informations ou de la complexité informationnelle de la page. Pour une tâche donnée, la meilleure page sera celle pour laquelle le temps d'exploration sera le plus faible,

Le nombre de fixations des sujets *ayant un but précis* s'avère sensible à la nature de l'information présente. Il faut viser à la précision sémantique des rubriques pour réduire le nombre de fixations,

L'évolution du nombre de fixations dans le temps permet de déceler une difficulté de choix, ou une gêne, car cela révèle une prise de décision difficile de l'internaute. L'explication de la cause peut être fournie par un questionnaire associé à l'étude du parcours (régressions et durées des fixations),

L'étude de la répartition des fixations initiales permet d'estimer si le niveau de contraste informationnel atteint par la zone considérée ou les objets qui la composent est satisfaisant,

La distribution des fixations dans chaque zone, en fonction du rang de la fixation, indique l'évolution de l'intérêt de la population de lecteurs au cours de l'exploration. Durant les premiers instants d'exploration, elle permet, en fonction du paramètre étudié, d'évaluer l'intérêt suscité par les différentes zones.

10.3. Mouvements oculaires et construction du sens

Une étude plus fine du document peut-être opérée en utilisant maintenant le regard dans l'acte de lecture visant à la compréhension d'énoncés. Nous avons utilisé l'enregistrement des mouvements oculaires dans un travail centré sur l'approche des processus cognitifs à l'œuvre dans la construction du sens, notamment lors de l'interprétation de la polysémie [HOL 99, HOL 00]. Ce travail s'appuie sur la théorie de la construction dynamique du sens [FUC 83, 85], [VIC 96].

L'observation des mouvements oculaires des sujets pendant la lecture constitue pour nous une technique d'observation "écologiques", c'est à dire respectant au maximum les processus à l'œuvre dans la lecture, en évitant notamment toute tâche interférente, et toute interruption ou modification du rythme naturel de lecture.

L'ambiguïté et son traitement

« Inscrite au cœur même de la langue » [FUC 96], l'ambiguïté se manifeste dès qu'à une forme unique correspondent plusieurs sens. On distingue généralement l'**homonymie**, cas où ces sens n'ont aucun lien entre eux⁷, de la **polysémie**, cas où, au contraire, les sens sont apparentés⁸.

Le traitement de l'ambiguïté a été abordé en psychologie cognitive en termes d'**activation** en mémoire du sujet. Selon l'hypothèse de "l'activation multiple", lors de la lecture d'une expression polysémique, toutes ses significations potentielles, stockées en mémoire à long terme, seraient activées dans un premier temps, puis désactivées ensuite. Selon l'hypothèse de "l'activation unique", au contraire, une seule signification serait activée d'emblée, et éventuellement désactivée au profit d'une seconde, par la suite. Il existe également des hypothèses mixtes.

De même, en informatique, le traitement de l'ambiguïté est généralement soit de type "parallèle", correspondant en gros au processus d'activation multiple décrit ci-

⁷ Ainsi « une bonne pêche » peut être cueillie dans un verger, ou présentée sur l'étal du poissonnier, voire, dans un langage plus familier, faire partie d'un récit de bagarre...

⁸ " un **polysème** est un signe dont l'unicité est historiquement attestée, et auquel correspond une diversité de sens dénotatifs possibles qui, en synchronie, sont ressentis comme sémantiquement apparentés : ainsi le *plateau* (support plat servant à transporter des objets, ou étendue de pays plate), ou le *violon* (instrument ou musicien).[FUC 94].

dessus, soit de type " séquentiel ", correspondant au processus d'activation unique. **L'approche linguistique**, quant à elle, propose un renversement de perspective, en considérant la plurivocité et la variabilité comme premières, et comme inscrites de façon contrôlée et réglée dans la sémantique du système de la langue [FUC 88, 94, 96].

La question ne se pose plus alors en termes de processus d'élimination ou de choix forcé dans un ensemble de significations préétablies, mais en termes de **construction** : À toute expression polysémique est associé un **noyau de sens**, sous déterminé, à partir duquel peuvent se déployer une pluralité de significations **construites de manière dynamique**, au cours du processus d'interprétation, en fonction d'indices contextuels. [FUC 83, 85] La modélisation de cette théorie linguistique s'effectue dans le cadre mathématique de la géométrie différentielle, et s'implémente à l'aide de réseaux neuronaux [VIC 96]

Dans le cadre de cette théorie, nous avons cherché à comprendre comment s'effectuait la construction du sens, notamment dans le cas où le lecteur est confronté à des éléments polysémiques, situation où sa faculté d'interprétation est mise à l'épreuve, plus intensément que dans le cas où la signification est évidente. Nous avons observé des sujets confrontés à des éléments polysémiques dans le cas où aucun des éléments constitutifs du texte n'induit une interprétation préférentielle, et dans le cas où, au contraire, le co-texte est fortement inducteur.

Notre but est de mieux connaître les processus cognitifs à l'œuvre dans ce type d'activité, afin d'être à même d'améliorer les échanges en langue naturelle dans le dialogue Homme-Machine.

Une approche expérimentale : trois observations

Notre approche passe donc par l'observation des stratégies, notamment oculaires, de sujets confrontés à un matériel linguistique contenant des éléments polysémiques

. Nous nous sommes demandé :

- à quel moment de la lecture commençait la construction du sens,
- sur quels éléments s'appuyait cette construction,
- et enfin dans quelle mesure la polysémie était perçue.

Nous avons réalisé trois observations :

La première consiste à faire lire à un grand nombre de sujets des textes à **interprétations multiples**, ou textes « polysémiques », c'est à dire contenant des mots ou des expressions porteurs de plusieurs sens, insérés dans un **contexte**

énonciatif non inducteur. Les textes ne comportaient pas de titre. Nous supposons que les sujets, ne disposant d'aucune indication pour guider leur interprétation, élaboreraient chacun très librement leur propre scénario, et que les scénarios recueillis pour un même texte seraient donc très divers. L'analyse de ces scénarios et des réponses des sujets devait nous indiquer les éléments qui avaient servi de base à la construction de chacune des différentes significations attribuées à chaque texte.

La deuxième a observé « on line » le comportement oculaire de sujets lisant des phrases isolées, comportant des expressions polysémiques identiques, insérées dans différents **contextes fortement inducteurs**. Les expressions polysémiques étaient constituées de syntagmes verbaux **syntactiquement identiques**⁹. En revanche, leur **valeur sémantique était complètement différente** dans chaque phrase. Nous supposons que, les sujets commençant à construire une signification dès le début de la lecture, l'expression polysémique serait intégrée dans un scénario déjà partiellement construit, et prendrait donc d'emblée une coloration sémantique différente selon le contexte. Si le parcours oculaire reflète cette valeur sémantique, alors il serait différent à chaque fois, et les sujets n'auraient même pas conscience de la similitude des expressions. Nous voulions donc comparer les parcours oculaires d'un même sujet sur les expressions polysémiques dans les différentes phrases.

La troisième a consisté en une **tâche d'association libre**, proposée sur le Web, à partir des termes polysémiques utilisés dans les textes à interprétations multiples. Nous voulions comparer les significations attribuées au terme polysémique selon que celui-ci était présenté isolément, ou en contexte. Nous supposons que la signification construite en présence d'un contexte, même non inducteur, différencierait de celle évoquée en l'absence de tout contexte, cette dernière étant plus proche du « **noyau de sens** » des termes polysémiques.

Nous donnons ci-dessous :

- a) la consigne générale qui introduisait les textes
- b) deux exemples de textes à interprétations multiples (l'observation en comptait huit.)

⁹ Les expressions cibles sont constituées des syntagmes verbaux « il a encore gagné » et « arrive encore en tête ». Nous avons utilisé deux types de verbe (transitif et intransitif), deux temps (présent et passé composé), et différents types de compléments. Pour faciliter l'exploitation du parcours oculaire, nous avons augmenté la taille de l'empan en ajoutant à l'expression verbale l'adverbe « encore », lui-même polysémique (sens duratif : « La chanson à texte... arrive **toujours** en tête » ; sens répétitif : « le RPR arrive **de nouveau** en tête » ; le sens notionnel : « Ce bruit s'appelle encore de la musique ? » n'est pas illustré ici).

- c) les six phrases contenant les syntagmes verbaux polysémiques « il a encore gagné » et « arrive encore en tête ».

Nous ne présenterons pas ici la totalité du travail réalisé, mais nous en citerons les principaux résultats.

a) Textes : Les consignes

- Lisez les pages dans l'ordre
 - Sur chaque page, vous trouverez un petit texte et une question. Nous vous demandons de lire le texte, puis de répondre immédiatement et spontanément à la question, sans réfléchir.
 - Passez ensuite à la page suivante.
- Sur la dernière page, vous pourrez inscrire vos remarques, et éventuellement, des réponses supplémentaires

b) Textes : Deux exemples

L'homme avait l'air préoccupé.
Son front dégarni était barré d'une ride soucieuse. Derrière d'épaisses lunettes, ses yeux rougis clignaient sans cesse. Il continuait à mâchonner une cigarette éteinte, qu'il ne songeait même pas à rallumer. Il se tourna vers son jeune collègue:
- "Et ce trou" lui dit-il, "tu y as pensé? comment allons nous faire pour le combler?"
L'autre avait l'air plus serein. La question ne semblait pas l'inquiéter et c'est presque en souriant qu'il répondit, d'une voix douce:
- "Ne t'en fais pas, j'ai tout prévu! Nous aurons tout ce qu'il faut demain matin et personne ne saura jamais qu'il y a eu un trou! Toutes mes prévisions vont dans ce sens-là!" et il ajouta, avec une pointe d'ironie: "tu sais que je fais entièrement confiance aux méthodes modernes de prévisions".

QUESTIONS:
A votre avis,
Où se passe cette scène?
.....
Quelle est la profession des personnages?
.....
Pouvez-vous entourer l'endroit du texte qui vous a suggéré votre réponse?

Texte 1 (« trou »)

Maintenant il était seul, dans cette ville inconnue, qu'il ressentait comme vaguement hostile. Le café s'était peu à peu vidé de tous ses clients, et les serveurs commençaient à empiler les chaises dans la salle du fond. Il sentit une sourde inquiétude l'envahir. Il leur avait fait totalement confiance en leur donnant le plan de la pièce. Et s'il avait eu tort? S'ils n'étaient pas honnêtes? Cette pensée l'avait déjà effleuré au début de l'entretien, mais il l'avait repoussée. Maintenant, elle revenait, plus insistante, plus précise, et il avait du mal à s'en défendre. Il sortit du café et se dirigea vers son hôtel. Il appréhendait le moment où il se retrouverait seul dans sa chambre. Il essaya de respirer calmement, en s'efforçant de ralentir son allure, tandis qu'il avançait dans cette nuit moite, qui étouffait le moindre bruit.

QUESTIONS:
A votre avis,
Où se passe cette scène?
.....
Quelle est la profession des personnages?
.....
Pouvez-vous entourer l'endroit du texte qui vous a suggéré votre réponse?

Texte 2 (« plan »)

c) Les Phrases

La chanson à texte a toujours gardé ses adeptes. "les feuilles mortes", par exemple, arrive encore en tête des "dix meilleures chansons françaises" pour la moitié des personnes interrogées.

Depuis qu'il a changé d'entraîneur, ses performances se sont améliorées et il a encore gagné en vitesse et en précision, jusqu'à se retrouver au niveau des plus grands.

Je pense souvent à cette image que je suis seule à voir et dont je n'ai jamais parlé. J'ai essayé d'oublier, je me suis étourdie. Mais, cette pensée m'arrive encore en tête quelquefois et le même regret m'envahit toute entière.

Mon client n'est pas un assisté! Il a travaillé toute sa vie. Même après le licenciement, il a encore gagné sa vie honnêtement, avec beaucoup de courage. Sa famille est là pour en témoigner.

Grâce au maintien du candidat vert dans ce canton, il est à prévoir que le R. P. R. arrive encore en tête, et même avec un meilleur score qu'au premier tour.

Noah joueur de tennis n'avait plus la faveur des médias, Noah chanteur ne l'avait pas vraiment non plus, mais dans son rôle d'entraîneur il a encore gagné le cœur des Français.

10.3.1 Observation 1 : les textes à interprétations multiples

Les textes étaient construits autour de termes ou de syntagmes polysémiques, insérés dans un contexte non inducteur, et présentés sans titre. Le texte d'entraînement, placé en tête, était faiblement polysémique ; Le dernier texte comportait une simple homonymie¹⁰ (avocat). Nous pensons que ce type d'ambiguïté, aisément repérable, aiderait le lecteur à prendre conscience du caractère polysémique des textes qu'il venait de lire, et ce, juste avant qu'il ne commence à répondre au questionnaire.

Les textes ont d'abord été présentés à 150 sujets " tout venant ", 42% de femmes et 58% d'hommes, de 15 à 75 ans, de niveau d'études divers (36% de non bacheliers), et de profession variées, dont 41% de salariés.

¹⁰ Ce type d'ambiguïté, qui ne donne le choix qu'entre deux solutions, est aussi appelé « ambiguïté alternative ».

Chaque texte mettait en scène une situation comprenant un personnage central et un environnement. Le lecteur devait lire le texte, puis répondre " immédiatement et spontanément " à une question, portant en général sur la profession du personnage. Sa réponse révélait, en fait, son interprétation personnelle du texte, la signification accordée à l'expression-cible, et les éléments du texte qu'il avait considérés comme informatifs.

10.3.2. Observation 2 : les phrases

12 sujets, âgés de 23 à 52 ans, 8 féminins, 4 masculins ont participé à l'observation. Les six phrases étaient construites autour de 2 expressions polysémiques¹¹, insérées chacune dans 3 contextes différents, fortement inducteurs. Les phrases étaient présentées sur l'écran d'un ordinateur, séparées par une ligne vide. L'expression polysémique occupait toujours la même place dans la phrase présentée, c'est à dire au début de la deuxième ligne.

Les sujets étaient placés dans le dispositif d'enregistrement des mouvements oculaires. Nous leur demandions de lire simplement les phrases, pour nous permettre de régler l'appareil, en ajoutant qu'une question générale pouvait leur être posée à la fin. Notre but était d'obtenir une lecture la plus naturelle possible, en évitant aussi bien un survol trop rapide, qu'un effort de mémorisation trop important.

10.3.3. Observation 3 : les associations libres

Les mots isolés sont les mots cibles polysémiques utilisés. Dans les textes cités en exemple, ce sont les mots « trou » (texte 1), « plan » et « pièce » (texte 2). Ils ont été présentés sur un site Web¹², grâce à un protocole électronique, pour une tâche d'association libre (observation 3). Les internautes devaient donner le premier terme qui leur venait à l'esprit à la lecture du terme présenté. Les termes cibles faisaient partie d'une liste de 140 mots. A travers les termes associés, nous voulions trouver le sens attribué à chacun des mots cibles présenté sans contexte, pour le comparer à la signification construite pour ce même terme en contexte non inducteur. Nous avons obtenu en moyenne 912 réponses pour chaque mot.

¹¹ Les expressions polysémiques sont constituées des deux syntagmes verbaux: « il a encore gagné » et « arrive encore en tête ».

¹² **WebCompletion**, protocole interactif, mis au point au CLIPS, par D. Vaufreydaz, J.P. Meiyé et J. Rouillard.

10.3.4. Perception de la polysémie

Dans les phrases

Nous avons observé, pour chacun des sujets, les stratégies oculaires sur les expressions cibles (point d'impact, parcours, temps de fixation) : bien que les expressions lues soient identiques, les stratégies de lecture, chez un même sujet, sont différentes pour chaque phrase. La figure ci-dessous présente l'enregistrement des fixations supérieures à 250ms sur les phrases contenant l'expression « il a encore gagné », pour le même sujet.

La taille des points est proportionnelle à la durée totale des fixations sur cette zone. Les chiffres indiquent le rang de la fixation.

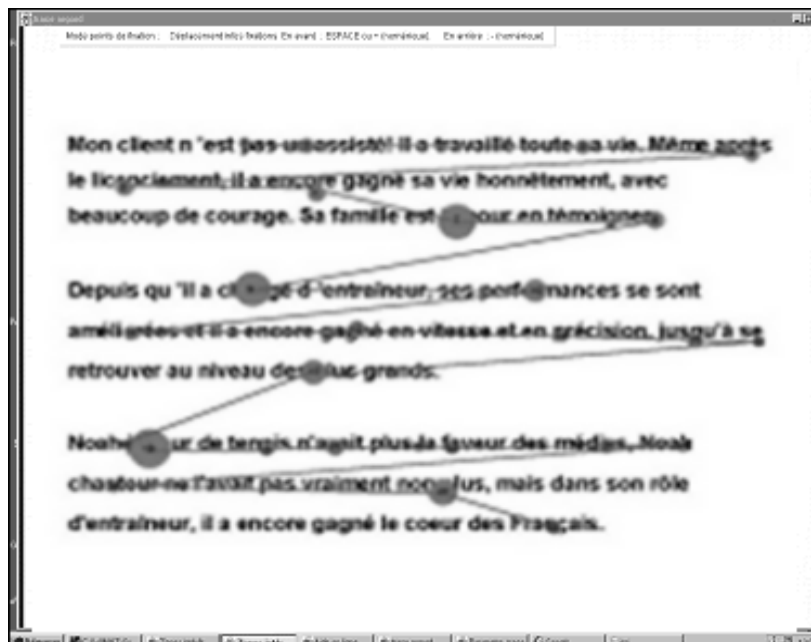


Figure 10.5 « il a encore gagné » Fixations supérieures à 250ms

Aucun des 12 sujets qui ont lu les phrases n'a perçu d'emblée la polysémie, même quand nous leur demandions s'ils avaient remarqué quelque chose dans la composition des phrases. Bien plus, leur surprise était grande quand, après la passation, nous leur montrions la similitude des expressions utilisées. Rappelons que

les six phrases étaient construites autour de deux expressions polysémiques identiques, qui se répétaient 3 fois chacune, qui plus est au même emplacement.

Si la forme morpho-syntaxique des expressions était la même, en revanche, leur signification était différente, puisque propre à chacune des phrases. Le fait que les stratégies oculaires soient également propres à chaque phrase laisse à penser qu'il n'y a pas d'analyse morpho-syntaxique préalable à la construction du sens, mais qu'au contraire cette construction commence avec la lecture même et influence la perception des éléments suivants. L'importance des premiers indices d'interprétation fournis au sujet qui aborde un écrit apparaît donc clairement ici. Elle est confirmée par l'analyse des scénarios construits par les lecteurs des textes à interprétations multiples.

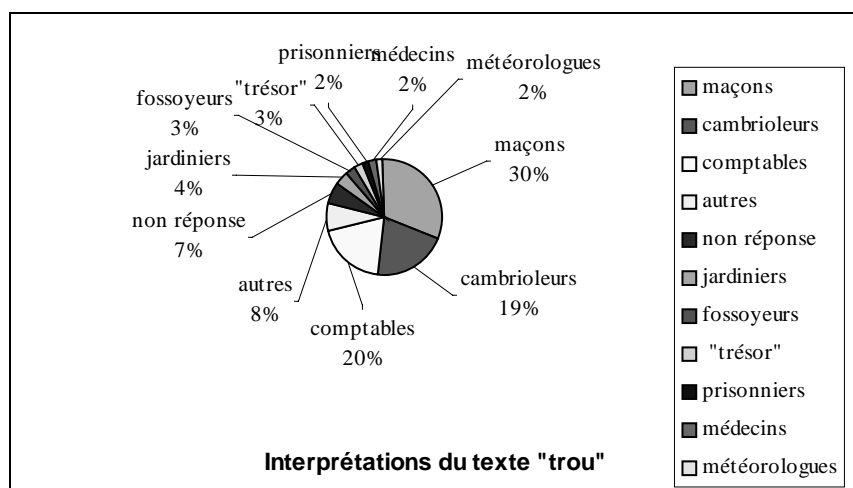
Dans les textes

Lors de la lecture des textes, tous les sujets ont immédiatement construit une interprétation, mais la polysémie n'a été perçue d'emblée que par 7 sujets sur 150. Ces sujets ont donné plusieurs réponses. D'autres ont mentionné dans leurs remarques finales la possibilité d'interprétations différentes. Un sujet dit par exemple : « à chaque lecture, je change d'avis ». 42 sujets ont fait des commentaires après la lecture des huit textes, souvent pour souligner leur polysémie, 35 n'ont fait que des remarques générales.

10.3.4. Les comportements interprétatifs

Au cours de la première observation, nous avons recueilli une très grande variété d'interprétations, que nous avons analysées en termes de scénarios, en regroupant les réponses par classes de scénarios apparentés. Par exemple, pour le texte "trou", cité en exemple, nous avons rassemblé les réponses : *assassins, bandits, braqueurs de banque, cambrioleurs, criminels, escrocs, gangsters, mafiosi, malfrats, truands, tueurs, tueurs professionnels, voleurs*, dans une même classe "cambrioleurs". Le nombre de classes de scénarios s'échelonne de 2 à 14 selon les textes. Le texte 8 (la simple homonymie "avocat") ne suscite que 2 classes de réponses (*homme de loi* et *légume*), tandis que le texte "trou" a donné lieu à 11 classes de réponses différentes, présentées dans la figure ci-dessous. Ce texte s'avère par ailleurs l'un des plus difficiles, avec 10 non-réponses¹³.

¹³ Nous avons établi un premier classement des textes, selon la proportion de sujets qui y ont répondu, le texte le plus facile étant celui auquel tout le monde a répondu.



Nous avons également recueilli les différents éléments ayant servi à établir les diverses interprétations. Certains de ces éléments sont effectivement présents dans le texte, d'autres peuvent en être inférés¹⁴, d'autres encore sont "importés" par le lecteur. Celui-ci semble donc construire, à partir des éléments les plus saillants pour lui, sa propre interprétation d'un texte, le rendant ainsi significatif. Une fois le scénario établi, il est affiné, avec quelquefois une extrême précision. Par exemple, pour le texte cité en exemple, (texte "trou"): "ingénieur et technicien plus jeune", "Employés du service des eaux", "Informaticien comptable en train de truquer un compte, un budget", "Intervenants sur marché financier pour le compte de clients"; "collecteurs de déchets radioactifs", voire "services financiers d'un ministère d'une "démocratie populaire" ou république bananière". Un sujet place les personnages « dans un bistrot, pendant la pause de midi », un autre « dans une pièce sombre et étroite », un autre encore les imagine sur « un forage géologique, réalisant des échantillonnages par carottage en terrain naturel sauvage ». Ce type d'interprétation, que nous appelons « réponses élaborées » est le fait des sujets les plus âgés. Le choix même des professions varie également selon l'âge. Les figures ci-dessous illustrent ces différences.

¹⁴ Le lecteur intéressé par l'exploitation de l'implicite pourra se reporter à l'ouvrage que C. Kerbrat-Orecchioni lui a consacré : « L'implicite » [Ker 86].

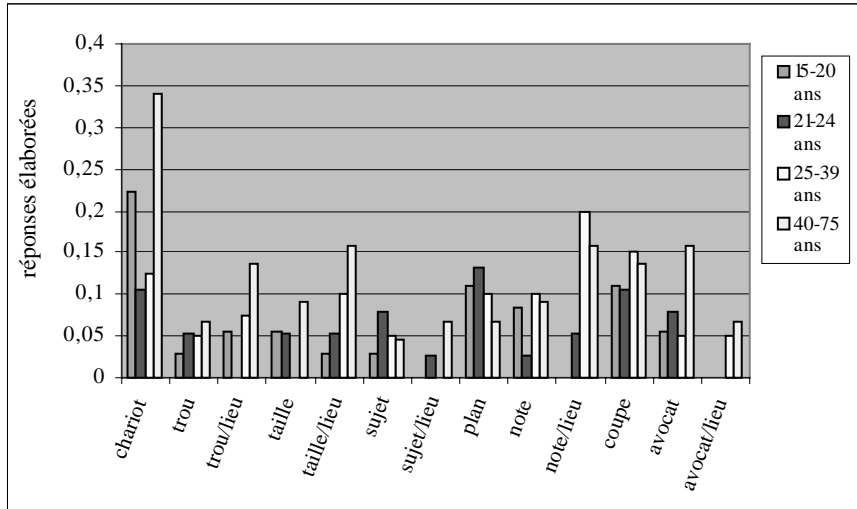


Figure10.6 a) : Taux de réponses élaborées selon l'âge des sujets pour tous les textes

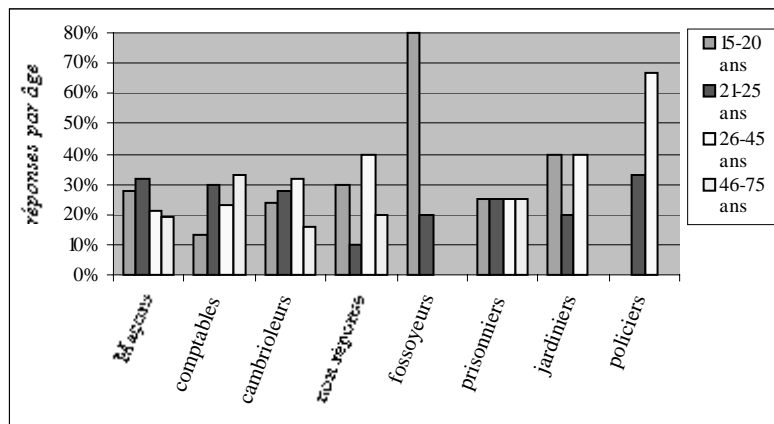


Figure10.6 b) : Les interprétations du texte « trou » selon l'âge des sujets pour les professions les plus fréquentes

La plupart du temps, le scénario, une fois déclenché, ne se modifie plus, quelles que soient les informations apportées ultérieurement. Par exemple la zone « cigarette éteinte » entraîne la réponse « *inspecteurs de police* », de même que les zones « ride soucieuse » et « jeune collègue » sélectionnées ensemble ; La zone « yeux rougis », aiguille la réponse vers « *informaticiens* » ; « ride soucieuse » entraîne la réponse

voire biologistes). Un sujet souligne que son interprétation (*comptables*) a été influencée par la présence du terme « combler » et non pas « boucher » dans le fragment « comment allons-nous faire pour le combler ? ». Ainsi, les deux termes « boucher » et « combler », sémantiquement apparentés et également pertinents dans ce contexte, voient leur interprétation différenciée en fonction de leur degré de concrétude : « boucher » plus concret, plus approprié à des éléments physiques, aurait induit l'interprétation « maçons », tandis que « combler » a entraîné l'interprétation « comptables ».

La présence du terme « collègues » écarte un sujet de l'interprétation « prisonniers » qu'il aurait choisie sans cela, au profit de l'interprétation « maçons ». Ainsi « collègues » ne fait-il pas partie des scénarios associés au métier de maçon pour ce sujet.

Lors de l'observation 3 nous avons recueilli, sur le Web, les termes associés librement à chacun des mots cibles, présentés isolément (sans contexte). Il est à remarquer que quand seul le mot polysémique a été sélectionné comme zone informative, les réponses s'apparentent aux associations libres. Par exemple pour "trou": "*fossoyeur*", "*croque mort*". Ces réponses n'apparaissent pas quand la zone "*prévisions*" ou "*méthodes modernes de prévisions*" est sélectionnée. Il semblerait, d'après ces réponses, que certains sujets, devant un contexte peu inducteur, renoncent à l'utiliser, et en restent à l'interprétation primitivement évoquée par le terme polysémique. En revanche, d'autres sujets n'ont sélectionné aucune zone du texte, mais ont construit un, ou plusieurs scénarios, comme si la seule présence du contexte les y incitait, même s'ils n'y trouvaient aucun élément inducteur.

Les zones d'intérêt indiquées dans cette observation ont été confirmées par les enregistrements des parcours oculaires sur les deux premiers textes. Les temps de fixation sur ces zones sont effectivement plus importants que sur le reste du texte.

10.3.6. Implications pour l'évaluation de textes

Ces observations permettent de confirmer que la perception visuelle doit bien être considérée comme une activité à part entière, et non comme le résultat de la stimulation de récepteurs sensoriels passifs. Cette étude met en évidence la diversité des "motifs" qui se dégagent d'un même "fond", en l'absence d'éléments inducteurs nets. Elle permet de mieux cerner les stratégies mises en œuvre par les lecteurs dans ce cas de figure. Quand aucun élément inducteur n'est donné par le texte, le lecteur s'attarde sur les éléments polysémiques, puis construit sa propre signification, selon ses connaissances et son expérience personnelles. L'interprétation construite reflète l'individualité du lecteur. En revanche, quand la signification est franchement induite par le contexte, alors le lecteur ne s'attarde pas sur les éléments

polysémiques, dont il n'a d'ailleurs aucune conscience. Il semble donc bien, à la lumière de cette étude, que si la lecture d'un terme isolé donne lieu à l'évocation d'un « noyau de sens », en revanche, la lecture de plusieurs termes associés déclenche un processus de construction sémantique, qui va guider la perception même des termes suivants, et intégrer des éléments liés à l'individualité du lecteur, en proportion variable selon le degré de précision du contexte. La connaissance de ce phénomène peut-être utile lors de la mise au point d'énoncés qui vont donner lieu à des inférences de la part d'un utilisateur, notamment dans le dialogue homme-machine.

10.4 Conclusion

Vis-à-vis de l'application des modèles d'exploration existants aux documents Web, il semble inopportun d'assimiler l'exploration d'une page Web à celle d'une image mais plutôt à celle d'un texte illustré. Les résultats obtenus semblent différer sensiblement de ceux obtenus par Yarus [YAR 67] sur les images :

- les schémas de lecture privilégient généralement le centre du document
- le premier regard n'est pas dirigé vers l'icône porteuse du plus grand sens, mais vers la zone ayant le plus fort pouvoir informatif pour l'utilisateur (cette zone dépend bien sûr du but poursuivi par ce dernier).

De même, des différences majeures la séparent d'un document papier :

- grande variété de couleurs pour la police et les fonds de zones,
- nombre important d'images et de photos,
- le concepteur de pages commerciales ne recherche pas forcément une structuration équilibrée du document, mais il veut que l'internaute trouve rapidement le lien intéressant, et/ou ne quitte pas la page sans avoir perçu certains objets.

Il apparaît donc important, lors de la conception d'une page Web, de privilégier les points suivants :

- adapter la présentation de l'information le plus possible à un balayage en Z,
- éviter l'utilisation de photos ou d'images qui ne représentent pas (ou ne recouvrent pas) des liens, et ce, afin de préserver les processus associés aux balayages hiérarchiques,
- conserver les « modèles » de représentations que possèdent les utilisateurs experts (pour cela faire explorer la page conçue par un expert avant de la mettre en ligne ou de la soumettre à une analyse plus poussée).

L'étude de la distribution des fixations initiales sur les différents objets ou zones de l'écran constitue un bon indicateur du pouvoir d'attraction de ceux-ci, et de fait, un bon moyen de vérifier qu'une information jugée importante par le concepteur (qui est souvent le vendeur dans un site à vocation commerciale) sera perçue par l'utilisateur (souvent le client).

Pour un sujet dont on connaît les habitudes de lecture, l'utilisation de la durée de la fixation semble appropriée pour étudier la complexité d'un objet, le manque de clarté sémantique, ou la présence d'une ambiguïté.

Le nombre d'informations présentes sur la page est important, mais il est conditionné par l'effort de catégorisation réalisé lors de la conception de la page. La nature et l'aspect physique de ces éléments d'information conditionnent les premières fixations. L'agencement des informations sur la page reste peu important dans la mesure où il permet des enchaînements de fixations en zone parafovéale. L'agencement des informations dans une zone contribue au contraste obtenu et à la hiérarchisation de cette zone au niveau perceptif.

Le contenu sémantique de ces informations doit tenir compte des processus de construction du sens mis en évidence par les études sur la polysémie, notamment la prégnance des premiers indices significatifs rencontrés par le lecteur, et l'effet déterminant du contexte sur l'interprétation finale. L'importance des caractéristiques individuelles de l'internaute (âge, culture, expertise) qui ressort de ces études appelle également une attention particulière lors de la mise au point des systèmes d'interaction.

Un paramètre de mesure robuste et simple est le temps de parcours visuel pour la recherche d'une information ciblée : plus ce temps est bref plus l'information est présentée de manière pertinente. On peut se servir de ce critère pour comparer plusieurs solutions concurrentes entre elles au moment de la conception.

Remerciements : Ces études ont eu le support de la Région Rhône-Alpes (projet SHIVA de type Emergence). Nous remercions également Brigitte Meillon pour le développement des logiciels de capture et de mesure oculaires, ainsi que Michel Nicolaï [NIC 01] qui a contribué à une partie des analyses et des résultats dans le cadre de son DEA en sciences cognitives.

Pour plus de détails sur les travaux de S. Hollard, B. Meillon et M. Nicolaï, voir le site Web du laboratoire Clips : <http://www-clips.imag.fr/>

10.8. Références

- [BAL 95] BALOTA D.A., POLLATSEK A., & RAYNER, K., «The interaction of contextual constraints and parafoveal visual information in reading», *Cognitive Psychology*, 17, p. 364-390, 1995.
- [BEA 90] BEAUVOIS J.L., ROULIN J.L., TIBERGHIE G., *Manuel d'études pratiques de psychologie, T1 : pratique de la recherche*, Presses Universitaires de France, 1990.
- [BON 89] BONNET C., GHIGLIONE R., RICHARD J.F., *Traité de psychologie cognitive, I, Perception, Action, Langage*, Dunod, 1989.
- [CAE 98] CAELEN J., BISSERET A., MALLEIN P., « La plate-forme MultiCom : vers un laboratoire d'usage et d'utilisabilité », *Congrès NIMES'98, éd. EC2, Nîmes*, p. 103-106, 1998.
- [CAT 95] CATLEDGE L., PITKOW J., «Characterizing Browsing Strategies in the World-Wide Web», *Computers Networks & ISDN Systems*, n° 27, p. 1065-1073, 1995.
- [CHAU 99] CHAUMERLIAC E., VILLARD, M., Perception de la polysémie au travers de textes courts, proposés à des adolescents normolecteurs et mauvais lecteurs , *Mémoire présenté pour l'obtention du certificat de capacité d'orthophoniste*, Université Claude Bernard, Lyon 1, 1999.
- [CHE 98] CHENE D., TIJUS C., POITRENAUD S., « Une méthode de conception des interfaces basée sur la catégorisation : théorie, méthode et outils informatiques », *Actes du colloque ERGO IA '98*, p. 206-215, 1998.
- [DOE 97] DOERMANN, D., ROSENFELD, A., RIVLIN, E., «The function of documents», *Proceedings of the International Conference on Doc. Analysis and Recognition, ICDAR-97*, vol.2, p. 1077-1081, 1997.
- [EGL 01] EGLIN, V., GAGNEUX A., «Functional labeling and printed text featuring», *Proceedings of the International Conference on Doc. Analysis and Recognition, ICDAR-01*, p. 27-45, 2001.
- [EGL 97] EGLIN, V., EMPTOZ, H., «Logarithmic Spiral Grid and Gaze Control for the Development of Strategies of Visual Segmentation on a Document». *Proceedings of the International Conference on Doc. Analysis and Recognition, ICDAR-97*, p.689-692, 1997.
- [EGL 98] EGLIN, V., Contributions à la structuration fonctionnelle des documents imprimés. Exploitation de la dynamique du regard dans le repérage de l'information, *Thèse Institut National de Sciences Appliquées de Lyon*, 249p., 1998.
- [FAN 99] FANCHINI., H., « L'analyse ergonomique des besoins (AEB) », *Actes du colloque Ingénierie des connaissances IC'99*, Polytechnique, Palaiseau, 14-16/06/99.
- [FAV 54] FAVERGE J.-M., *Méthodes statistiques en psychologie appliquée*, Presses Universitaires de France, 1954.
- [FIS 75] FISHBEIN M., AJZEN I., «Belief, attitude, intention and behaviour : An introduction to theory and research», *Reading MA*, Addison-Wesley Publishing Company, 1975.
- [FUC 83, 85] FUCHS, C. ET LE GOFFIC, P., « Ambiguïté, paraphrase et interprétation », *Modèles linguistiques*, V (p. 109-138) et VII (p. 27-51), Lille, Presses universitaires, 1983/1985.

- [FUC 88] FUCHS, C., « Représentation linguistique de la polysémie grammaticale » *T. A. informations*, 1-2, P.7-20, 1988.
- [FUC 94] FUCHS, C., *Paraphrase et énonciation*, OPHRYS, 1994
- [FUC 96] FUCHS, C., *Les ambiguïtés du français*, OPHRYS, 1996.
- [HOL 00] HOLLARD, S., « Interprétation de textes polysémiques : une étude expérimentale », *Troisième Atelier de Sémantique : "Corpus - verbalisation de l'action"*, Toulouse le Mirail, 5-6 octobre 2000, consultable sur <http://www.univ-tlse2.fr/erss/ateliersem.html>
- [HOL 99] HOLLARD, S., « Interpretation of polysemous texts : an experimental study », *Proceedings of The 4th EUROLAN Summer School on Human Language Technology*. IASI, Romania, 1999.
- [JUS 80] JUST M. A., CARPENTER, P. A., « A Theory of Reading, From Eye Fixation to Comprehension », *Psychological Review*, 87, p. 329-354, 1980.
- [KAY 90] DE KAYSER V., « Fiabilité humaine et la gestion du temps dans les systèmes complexes ». In J. Leplat and G. de Terssac (eds). *Les facteurs humains de la fiabilité dans les systèmes complexes*, p. 85-108, Octares, Toulouse.
- [KER 86] C. KERBRAT-ORECCHIONI C., « L'implicite », Armand Colin, 1986.
- [LEE 98] LEEK E.C., « Effects of stimulus orientation on the recognition of common polyoriented objects », *Psychonomic Bulletin & Review*, 5, p. 650-658, 1998.
- [LER 96] LEROUX M., « Méthodologie d'une spécification d'interface guidée par un modèle cognitif. Application à la spécification des futures aides à la décision pour les contrôleurs aériens », *Journées IHM96 de Grenoble*, 1996.
- [LEU 98] LEULIER C., BASTIEN J.M.C., SCAPIN D.L., *Compilation of Ergonomic Guidelines for the Design and Evaluation of Web Sites, Rapport de recherche de l'INRIA, « Commerce et Interaction »*, Février 1998.
- [LEU 99] LEULIER C., BASTIEN J.M.C., SCAPIN D.L., *L'évaluation des sites Web : Apports respectifs des questionnaires de satisfaction et de l'inspection ergonomique*, 1999.
- [LEV88] LEVY-SCHOEN A., « Les mouvements des yeux comme indicateurs des processus cognitifs », in J.P. Caverni, C. Bastien, P.Mendelsohn et G.Tiberghien (eds), *Psychologie cognitive, modèles et méthodes*, P.U.G., p. 329-347, 1988.
- [LOF 81] LOFTUS G., « Tachistoscopic Simulations of Eye Fixations on Pictures », *Journal of Experimental Psychology : Human Learning & Memory*, 1981.
- [MAR 00] MARENDAZ C., ROUSSET S., CHARNALLET A., « Reconnaissance des scènes, des objets et des visages », in A. Delorme et M. Flückiger (Eds), *Perception et Réalité*. Montréal, Gaëtan Morin.
- [MAS 76] MASSE, D., *Le contrôle des mouvements oculaires, Thèse Institut National Polytechnique de Grenoble*, 1976.
- [MEN 93] MENDELSON P., NGUYEN-XUAN A., « Psychologie cognitive et EIAO », 3^o journées EIAO de Cachan, Eyrolles, 1993.

- [MOL 86] MOLES A., *Théorie structurale de la communication et société*, Coll. CNET, ENST, Masson, 1986.
- [NIC 01] Nicolai M., *Suivi du regard et ergonomie cognitive*, Dea de sciences cognitives, Grenoble, 2001
- [NIE 97] NIELSEN J., *Be Succint : How to Write for the Web*, Jakob Nielsen's Alterbox, 15 mars 1997, <http://www.useit.com/alterbox/9703b.html>
- [NIE 99] NIELSEN J., *Differences Between Print Design and Web Design*, Jakob Nielsen's Alterbox, 24 janvier 1999, <http://www.useit.com/alterbox/990124.html>
- [NORn 83] NORMAN D.A., «Some Observations on Mental Models», in Gentner & Stevens eds, *Mental Models*, Hillsdale, N-J, Lawrence Erlbaum Associates, 1983.
- [O'R 90] J.K. O'REGAN, « Eye movements and cognitive processes» in E. Kowler (ed.), *Eye movements and their role in visual and cognitive processes*, North-Holland, Elsevier Science Publishers, 1990.
- [RAS 86] RASMUSSEN J., *Information processing and human-machine interaction : an approach to cognitive engineering*, New York, North Holland, 1986.
- [RAY 92] RAYNER K., POLLATSEK A., «Eye Movements and Scene Perception», *Canadian Journal of Psychology*, 46, p. 342-376, 1992.
- [RIC 89] RICHAUDEAU, F. *Manuel de typographie et de mise en page*, Paris : Retz, 128 p., 1989.
- [UND 90] UNDERWOOD G., CLEWS S., EVERATT J, «How do readers know where to look next? Local information distributions influence eye fixations», *Quarterly Journal of Experimental Psychology*, 42A(1), p. 39-65, 1990.
- [VER 90] VERGNAUD G., « La théorie des champs conceptuels », *Recherches en didactique des mathématiques*, vol. 10, n° 2/3, p. 133-170, 1990.
- [VER 91] VERMERSCH P., « L'entretien d'explicitation », *Les Cahiers de Beaumont*, n°53, p. 63-70, 1991.
- [VER 90] VERMERSCH P., «Questionner l'action : l'entretien d'explicitation», *Psychologie Française*, n° 35-3, p. 227-235, 1990.
- [VIC 96] VICTORRI, B., *La polysémie : construction dynamique du sens*, Paris, Hermès, 1996.
- [YAR 67] YARBUS A.L., *Eye Movement and Vision*, Plenum Press, New-York, 1967.
- [YEO 96] YEO A., «World-Wide CHI : Cultural User Interfaces, A Silver Lining in Cultural Diversity», *SIGCHI*, vol. 28, n° 3, juillet 1996.
- [YOU 83] YOUNG R. M., «Surrogates and Mappings : two Kinds of Conceptual Models for Interactive Devices», in Gentner & Stevens eds, *Mental Models*, Hillsdale, N-J, Lawrence Erlbaum Associates, 1983.