

**ANALYSE DES MOUVEMENTS OCULAIRES
DURANT L'OBSERVATION D'UNE IMAGE PUBLICITAIRE:
DONNÉES OCULOMÉTRIQUES
ET RÉPONSES À UNE QUESTION**

[OCULOMÉTRIE]

*P. LOSLEVER**
*F. ROUSSEAU**
*J.-C. ANGUE**

1 Du problème aux analyses

L'image, fixe ou mobile, est devenue le principal vecteur de communication. L'image publicitaire est l'exemple par excellence: accrocheuse, informationnelle, émotionnelle... Ce type de communication utilise essentiellement le canal visuel. Notre travail présente, après un rappel de différentes méthodes utilisées dans l'étude des images publicitaires, l'analyse de données oculométriques *stricto sensu* recueillies au rythme de 60 par seconde, mais, éventuellement cumulées par tranches d'une seconde; puis une correspondance entre les mouvements oculaires, assimilables à ce qui est observé, et la réponse à une question composée, assimilable à ce qui est retenu.

1.1 L'étude des images publicitaires

Sans faire un exposé exhaustif des méthodes d'analyse d'une image publicitaire, nous en traçons les grands axes afin de situer notre travail.

L'image publicitaire obéit à certaines lois de perception; on lui attribue, par rapport aux autres moyens de communication, la propriété d'universalité, de multiplicité tant en forme qu'en signification (polysémie), de globalité, de quasi-instantanéité,... (BATICLE 1985). D'une part, une image peut être considérée comme un support physique: on parle alors de luminosité, de contraste, de couleur, de forme,...; dans ce cas, seul le support de

(*) Laboratoire d'Automatique Industrielle et Humaine. URIAH. U.A. C.N.R.S. n°1118;
Université de Valenciennes et du Hainaut Cambrésis;
Le Mont Houy, B.P. 311 - 59304 Valenciennes Cédex - France.

communication est étudié et sa fonction première de vecteur d'information disparaît. Mais d'autre part, quant au contenu, les méthodes d'analyse peuvent être classées en deux familles: approche par expertise et approche expérimentale.

L'approche par expertise fait appel à la sémiotique et à la rhétorique. Parmi les travaux dans ce domaine, effectués dans les années soixantes dix, on citera DURANT 1970, ECO 1970, MOUNIN 1974, GAUTIER et CHEBAT 1978). Plus récemment, VETTRIANO SOULARD (1987a, b, 1988) s'est basé sur l'analyse du contenu d'un ensemble d'annonces publicitaires afin d'en déduire des règles, des invariants, ou des indices susceptibles de les qualifier. Cet ensemble est réalisé en groupant les annonces suivant un ou plusieurs facteurs, les autres étant constants. Ces méthodes ont l'inconvénient de ne pas prendre en compte les fonctions de perception des récepteurs.

Pour tenir compte de ces remarques, le receveur, ici la personne qui regarde l'image, est observé par nous à deux niveaux. D'une part, on enregistre ses mouvements oculaires durant la phase d'observation; d'autre part, on s'enquiert, par questionnaire, de ce qu'il a retenu.

1.2 Rappels sur la vision centrale et périphérique

Dans les mouvements oculaires, il faut considérer deux distinctions: dans la cinématique de l'œil, entre fixations et saccades; dans le champ visuel, entre vision centrale et périphérique.

Bien que les images perçues consciemment par l'homme lui paraissent stables et non discontinues, l'observation des mouvements oculaires décèle des saccades plus ou moins régulières (de 2 à 5 fois par seconde), d'amplitude parfois importante (jusqu'à une quinzaine de degré) et très rapides: environ 1s/400 (LÉVY-SCHØEN 1969). Entre deux saccades, le regard semble se poser en des endroits particuliers de l'image; ce sont les fixations dont la durée est très variable. Selon LOFTUS (1981), une reconnaissance peut apparaître à partir d'un seuil minimum de 50ms. La durée de fixation peut aller de quelques centièmes à quelques dixièmes de seconde. Elle donne une idée du temps mis par l'homme pour décoder, enregistrer et retenir l'information visuelle. C'est donc une variable fréquemment utilisée dans l'analyse du comportement visuel (ACIER 1990).

On ne voit les détails avec précision que lorsqu'on les fixe; ce qui s'obtient quand on oriente l'œil de façon à amener l'image sur une région particulière de la rétine appelée *fovea centralis*. Cette vision, dite centrale, correspond à un angle solide de vue faible relativement au champ visuel total. À l'extérieur de cet angle, la vision est peu sensible aux détails; c'est la vision périphérique. C'est sur cette distinction que reposent de nombreuses

recherches sur l'analyse des mouvements oculaires, dans des domaines aussi divers que l'ergonomie cognitive, la médecine, le sport, la lecture,...

Le problème qui se pose alors est d'enregistrer et d'analyser l'évolution spatio-temporelle des fixations.

Certains auteurs, utilisant les techniques de la statistique descriptive classique, visent à modéliser :

le comportement spatial : "Où ?" (point de fixation suivant)

le comportement temporel : "Quand ?" (instant de fin de la fixation).

D'autres recourent à des théories complexes: théorie des files d'attente (CARBONELL 1966); théorie de l'information (STRIEB et GROVE 1979); probabilités de transition d'une zone vers une autre (PAPIN 1981, MOLNAR 1981 et 87, ELLIS & STARK 1986).

Dans l'analyse des données oculométriques, on a utilisé deux entités de base pour décrire la chronologie de la lecture: la fixation, considérée par MOLNAR, par exemple dans l'étude de l'observation d'œuvre d'art (MOLNAR 1981 et 87); ou l'échantillon temporel, adopté par TOLE (1982) pour l'évaluation ergonomique d'un poste de pilotage d'avion.

Quant à nous, bien que vision centrale et vision périphérique ne puissent être dissociées, nous retenons ici l'hypothèse suivante:

Si le regard se pose sur un élément précis de l'image (tête d'un personnage, texte, objet ...), il y a vision centrale; si le regard n'est pas sur un élément précis, l'image est regardée dans son ensemble, en vision périphérique.

1.3 Dispositif expérimental: présentation des images et enregistrement des données oculométriques

Les images publicitaires sont présentées par un projecteur de diapositives. Elles s'inscrivent dans un rectangle d'environ $1,2 \times 0,8$ m² et le sujet est situé à environ 2m de cette surface. La dimension et l'orientation des images étant variables, les coordonnées de la position du regard sont normées par rapport à la taille de celles-ci: le point (0, 0) correspond au bord inférieur gauche, le point (100, 100) au bord supérieur droit. En raison des erreurs de mesure, cette position peut être en deçà de 0 ou au delà de 100 dans l'une ou l'autre des directions (l'intervalle étant de -10 à + 110 environ).

Trente images sont proposées au sujet, chacune pendant 5 secondes. La durée de présentation doit être assez longue pour qu'on puisse reconnaître une stratégie de lecture semblable à celle de l'observation d'une affiche publicitaire. Des images présentées, une seule nous intéresse particulièrement (cf. *infra*, §1.4) et fera l'objet du questionnaire.

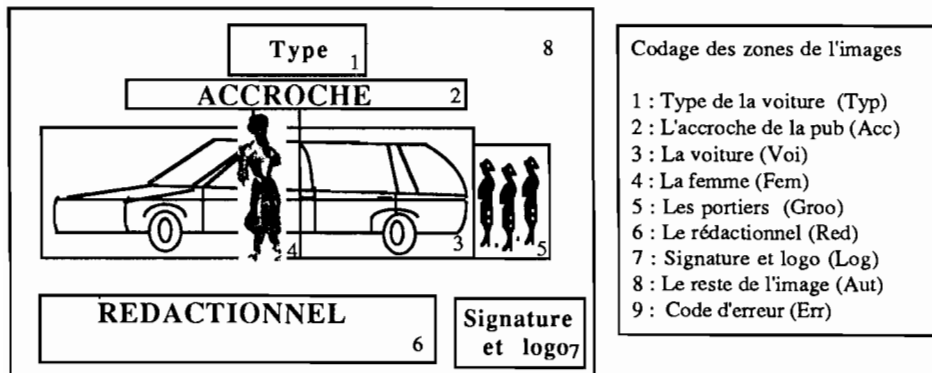
La mesure, très délicate, des mouvements oculaires repose sur des propriétés optiques de l'œil (méthode du vecteur cornéen (YOUNG & SHEENA 1975). L'œil est éclairé par une source lumineuse non visible (infra rouge). Le système de mesure comprend divers éléments dont :

Une caméra sensible aux faibles luminosités, placée devant le sujet, qui filme les mouvements de l'œil.

Une seconde caméra, située à côté de la tête du sujet, qui enregistre la scène observée.

Un système, calculant, en temps réel, la direction du regard dans un repère lié à l'image scène; et cela à une fréquence de 60 échantillons par seconde.

Un réticule incrusté dans l'image scène matérialise la zone observée dans l'image présentée au sujet. L'expérimentateur peut donc vérifier le déroulement normal du protocole. De plus, la mesure s'effectue sans contact avec l'œil ou la tête du sujet et nécessite peu de coopération de sa part. Les perturbations liées à l'instrumentation sont donc très limitées.



1.4 L'image retenue: découpage en huit zones

De la série de trente images présentées à chaque sujet, une seule a été étudiée assez minutieusement pour pouvoir faire l'objet de la présente étude. Sur la figure, cette image est décomposée schématiquement en huit zones que nous décrirons,... en ajoutant des couleurs au schéma.

Au centre, une voiture rouge; debout devant la porte de la voiture, une femme, vêtue de rouge; à l'arrière de la voiture, chargés de bagages, trois grooms dont on ne doutera pas qu'ils soient en rouge. Et au-dessus, en guise d'accroche, la phrase: Le rouge va bien aux voitures de sport, surmontée d'une référence de catalogue précise, le type. Au bas de l'image, dans une bande dépourvue d'agressivité, un rédactionnel plaide la cause du modèle proposé, censé allier charme et utilité; avec, à droite, le logo de la marque.

A toute orientation du regard est affectée une modalité z de l'ensemble Z des huit zones sémantiques de l'image. Une 9-ème modalité recense les erreurs du système de mesure que l'on n'a pu corriger.

De façon précise, les intervalles de mesure des positions horizontales, x , et verticales, y , sont découpés pour suivre le schéma des 8 zones. Afin de nuancer le contraste des boîtes rectangulaires, on utilise un codage flou; codage d'ailleurs adapté au système visuel humain. Les variables x et y sont découpées en des modalités dont les fonctions d'appartenance sont en forme de trapèze. En croisant les modalités de x et y , et en regroupant ces croisements suivant les 8 zones, on obtient, pour chaque direction, un niveau d'appartenance à chacune des 8 modalités de Z . Généralement, le regard r tombe dans le corps d'une zone, e.g. celle de la carrosserie: $k(r, \text{voit}) = 1$ et les autres $k(r, \text{zone})$ sont nuls; mais vers l'arrière, on peut avoir: $k(r, \text{voit}) = 0,5$; $k(r, \text{groo}) = 0,5$; le reste étant nul; etc. Dans la suite, afin d'abrégier le discours, nous parlerons, e.g., du nombre de visées du sujet s tombant dans la zone z pendant la t -ème seconde de présentation de l'image; sans rappeler qu'il s'agit d'une quantité continue et non d'un simple dénombrement.

1.5 Format des données et enchaînement des analyses

La description formelle des comportements enregistrés met en correspondance plusieurs ensembles.

S : ensemble des 47 sujets auxquels a été présentée la série des 30 images.

I : ensemble des 300 instants de mesure en lesquels est subdivisé l'intervalle de 5 secondes consacré à la présentation d'une image; la direction du regard est déterminée à chacun de ces instants.

T : ensemble $\{A, B, C, D, E\}$ de cinq tranches d'une seconde en lesquelles est partagé I . La durée moyenne d'une fixation sur zone étant de l'ordre de 250 à 350ms, la largeur d'une fenêtre doit être supérieure à cette durée, sans pour autant être trop large de façon à ne pas trop filtrer la mobilité du regard. En découpant des tranches de 0,5s ou de 1s, on a obtenu des résultats équivalents quant à la succession des profils de distribution des regards de l'ensemble des 47 sujets, depuis le début de la présentation jusqu'à la fin de celle-ci.

X, Y : ensemble $\{0, 1, \dots, 100\}$ des valeurs des coordonnées horizontale et verticale introduites sur le champ d'une image (cf. *supra*, §1.3).

Z : ensemble des 8 zones sémantiques en lesquelles a été découpée l'image, avec une neuvième modalité d'erreur (cf. §1.4).

Q : ensemble des 6 modalités de réponse à la question composée servant à déterminer, après passage des images, ce qui a été retenu de l'image particulière qui nous intéresse. Cet ensemble de modalités sera décrit au §4.1, en étudiant les relations entre réponse et comportement de lecture.

Au §2, les données oculométriques sont analysées en conservant l'échelle de temps I des 300 instants d'observation; mais les transitions de $X \times Y$ à $X \times Y$ ou de Z à Z sont considérées sur l'ensemble de la population, sans distinguer les comportements des divers sujets.

À partir du §3, l'échelle de temps est l'ensemble T des 5 tranches; et l'image est réduite à l'ensemble Z des zones; mais les différences entre sujets font l'objet de certaines analyses.

Au §4, on met en relation les comportements de lecture avec l'ensemble Q des modalités de réponse au questionnaire.

La description précise des tableaux considérés sera donnée au fur et à mesure des analyses.

2 Analyse des données oculométriques instantanées

saut de la ligne de visée entre deux mesures successives											
l1	G10	G8	G6	G4	G2	h0	D2	D4	D6	D8	D \geq 10
H \geq 10	4	1	1	1	1	1	2	1	1	0	4
H8	0	0	0	0	1	1	0	1	0	0	1
H6	0	0	0	0	1	1	1	1	0	0	0
H4	0	0	0	0	1	2	1	2	1	0	1
H2	1	1	0	1	3	10	9	4	2	0	1
v0	1	1	2	4	21	558	144	10	3	1	1
B2	0	1	0	2	9	90	27	3	2	0	1
B4	0	0	1	1	2	9	4	0	0	0	0
B6	0	0	0	0	1	4	1	0	0	0	0
B8	0	0	0	1	1	3	1	0	0	0	0
B10	4	0	0	1	2	4	3	1	0	0	3

2.1 Mobilité du regard

D'une part, d'une fixation de zone à une autre, la ligne de visée peut, de façon quasi instantanée, faire un saut dont l'amplitude, mesurée sur l'image en Δx et en Δy , est de plusieurs dizaines. D'autre part, au cours d'une fixation, la visée n'est pas rigoureusement immobilisée. On a, entre deux mesures consécutives (i et $i+1$, i.e. t et $t+(1/60)$), des $(\Delta x, \Delta y)$ variant de 0 à quelques unités. Le tableau publié ici offre l'histogramme bidimensionnel d'un échantillon de la distribution des sauts: y prédominent les petits déplacements correspondant aux sauts (décalages) contenus dans les fixations.

Considérons successivement les cases les plus chargées.

À l'intersection de la ligne v0 et de la colonne h0 on lit: saut(v0, h0) = 558: dans 558 cas, le déplacement a été 0 ± 1 , dans la direction verticale comme dans la direction horizontale. Vient ensuite saut(v0, D2) = 144: déplacement vertical nul, mais saut vers la droite de 2 ± 1 ; le saut vers la gauche est beaucoup moins fréquent: saut(v0, G2) = 21. Le mouvement vertical le plus fréquent est saut(B2, h0) = 90 : saut, de 2 ± 1 , sans décalage horizontal (autre que 0 ± 1); le mouvement opposé est rare: saut(H2, h0) = 10. Le seul mouvement diagonal de fréquence notable est saut(B2, D2) = 27.

2.2 Analyse des transitions instantanées entre zones

Les fréquences des changements de zone entre deux échantillons successifs constituent une matrice de transition, qui peut être soumise à l'analyse des correspondances (Van der HEIJDEN 1985). Mais, dans le cas présent, la diagonale requiert un traitement particulier.

Tableau des transitions entre zones: $t_{ZZ} : 9 \times 9$

9	type	accr	voit	femm	groo	réda	logo	autr	erre
Type	1095	176	0	0	0	1	1	60	2
Accroch	173	3618	121	115	33	1	2	281	11
Voiture	2	131	1738	218	125	1	2	106	5
Femme	0	118	230	1634	1	2	0	176	6
Groom	0	35	126	1	678	0	0	5	8
Rédact	0	1	1	2	0	121	9	64	0
Logo	0	2	0	0	1	11	112	48	0
Autre	59	268	110	188	8	60	46	1561	8
Erreur	3	10	8	8	4	0	0	7	295

2.2.0 Examen du tableau brut et reconstitution d'une diagonale

Dans la matrice de transition t_{ZZ} , 9×9 , la diagonale est prépondérante (77% du total du tableau). Ceci s'explique par le fait que la durée séparant deux mesures successives est très courte relativement aux durées des fixations oculaires. De plus, cette matrice est quasi symétrique: la différence entre $k(i, i')$ et $k(i', i)$ ne dépassant jamais les fluctuations d'échantillonnage. À une visée sur la mention du type (au haut de l'image) succède dans 1095 cas une visée sur cette même zone; le nombre des passages du type à l'accroche ou de l'accroche au type n'est que de ≈ 175 ; restent quelque 60 passages dans chaque sens entre type et fond (autre).

Notre but étant d'analyser les changements d'une zone à une autre, les éléments de la diagonale n'ont pas de sens direct. Mais pour l'analyse factorielle, ni le tableau brut ni le tableau avec diagonale mise à zéro ne donnent des résultats parfaitement clairs. On reconstitue donc une diagonale, comme il est d'usage (cf. [TRAC. MANQ. FLUX.], in *CAD*, Vol.XVII, n°3), de telle sorte que soit minima la trace du tableau obtenu.

En bref, ne retenant du tableau de transition, t_{ZZ} , que les termes extradiagonaux de son symétrisé, s_{ZZ} , on calcule d'après ceux-ci une loi marginale b_Z ; on cherche une autre loi sur Z , a_Z , dont le produit par elle-même, $a_Z \times a_Z$, est un tableau dont les termes extradiagonaux ont précisément pour loi b_Z ; et on attribue au tableau de départ, t_{ZZ} , la diagonale de $a_Z \times a_Z$ (à un coefficient près). Cette méthode fournit la diagonale exacte si t_{ZZ} est un tableau carré symétrique, produit d'une loi par elle-même. Ultérieurement, on peut parfaire la reconstitution de la diagonale en y ajoutant un terme proportionnel à a_Z , afin que la trace soit minima.

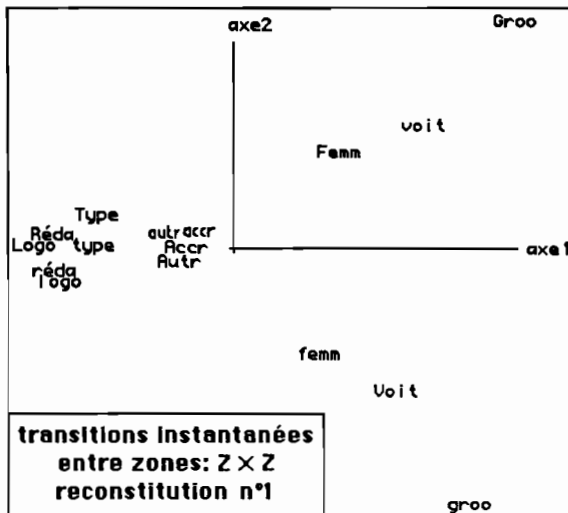
Il y a donc deux reconstitutions possibles de la diagonale; et, de plus, l'on peut, ou non, prendre en compte la modalité d'erreur: ce qui n'a ici aucune

influence notable. Nous donnons les résultats obtenus avec la reconstitution 1, avec 9 modalités; et ceux avec la reconstitution 2 et 8 modalités. Dans le premier cas, comme dans le second, l'analyse porte sur un tableau principal 8×8 , sans mettre *erreur* en ligne ni en colonne principale.

2.2.1 Analyse des transitions: première reconstitution de la diagonale

Transitions entre zones: 9×9 : reconstitution 1									
9	type	accr	voit	femm	groo	réda	logo	autr	erre
Type	16	176	0	0	0	1	1	60	2
Accr	173	244	121	115	33	1	2	281	11
Voit	2	131	133	218	125	1	2	106	5
Femm	0	118	230	102	1	2	0	176	6
Groo	0	35	126	1	8	0	0	5	8
Réda	0	1	1	2	0	2	9	64	0
Logo	0	2	0	0	1	11	1	48	0
Autr	59	268	110	188	8	60	46	252	8
Erre	3	10	8	8	4	0	0	7	0

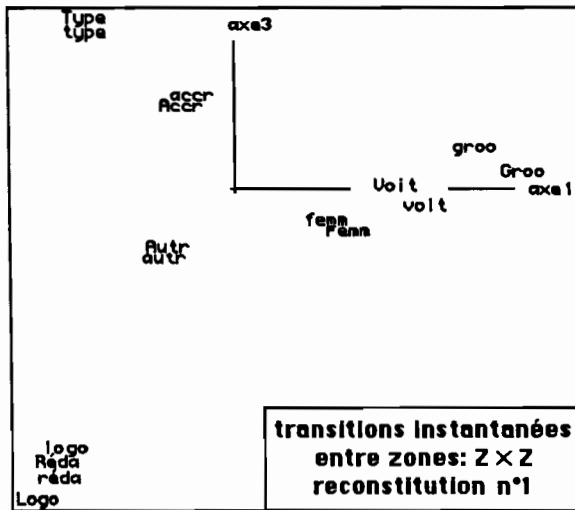
L'inertie sur le premier axe dépasse nettement celle sur les axes 2 et 3; il y a ensuite une nette dénivellation entre les rangs 3 et 4. Comme la matrice de transition est quasi symétrique, on peut dans l'interprétation parler de facteur direct ou inverse, selon que coïncident ou s'opposent les deux projections d'une même zone considérée comme ligne (sigle avec capitale) ou colonne (sigle en minuscule).



Le facteur 1 est direct; il oppose à tout le reste {Femm, Groo, Voit}: la voiture et les personnages.

La plupart des transitions se font entre deux zones de l'image qui, si on les considère comme lignes, se projettent, sur l'axe 1, d'un même côté de l'origine: soit toutes deux sur ($F1 > 0$): e.g. de Femm à voit; soit sur ($F1 < 0$).

Le facteur 2 est inverse: il intéresse le groupe {Femm, Groo, Voit} qu'il scinde en deux {Voit} \neq {Femm, Groo}. Ce qui correspond à la disposition spatiale, le passage de Femm à Groo ne pouvant guère se faire sans l'intermédiaire de Voit. Pour les autres zones, on a $F2 \approx 0$.



Le facteur 3 est direct: il oppose la partie supérieure de l'image {Type, Accr}, à la partie inférieure et au fond {Réda Logo Autr}.

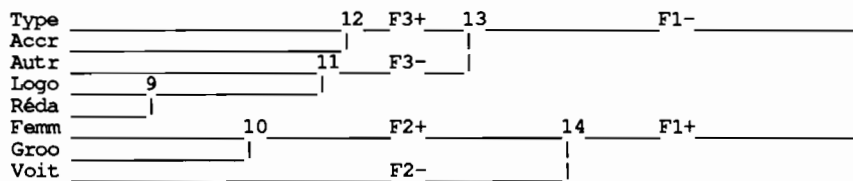
Ceci indique, que, pour les zones dont les sigles, comme ligne et colonne, se projettent sur (F1<0), les transitions se font surtout entre zones de la partie supérieure, ou entre zones de la partie inférieure de l'image

Transitions entre zones: 9 x 9 : reconstitution 1

trace :	5.564e-1 (erreur est en suppl.)						
rang :	1	2	3	4	5	6	7
lambda :	2074	1211	1072	612	311	165	119 e-4
taux :	3727	2176	1926	1100	559	296	215 e-4
cumul :	3727	5904	7830	8930	9489	9785	10000 e-4

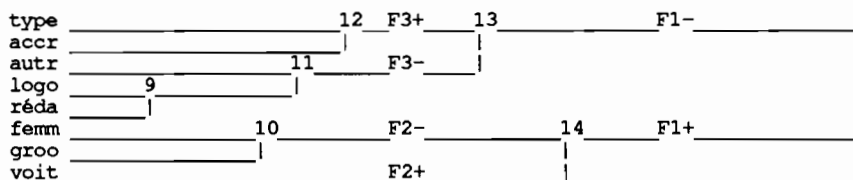
La classification des deux ensembles en correspondance confirme pleinement l'interprétation des facteurs.

reconstitution 1: Classification des antécédents



On notera que les structures hiérarchiques mises sur les deux ensembles sont quasi identiques.

reconstitution 1: Classification des conséquents



Transitions entre zones: 8x8 : reconstitution 2

8	type	accr	voit	femm	groo	réda	logo	autr
Type	36	176	0	0	0	1	1	60
Accr	173	321	121	115	33	1	2	281
Voit	2	131	191	218	125	1	2	106
Femm	0	118	230	151	1	2	0	176
Groo	0	35	126	1	22	0	0	5
Réda	0	1	1	2	0	8	9	64
Logo	0	2	0	0	1	11	6	48
Autr	59	268	110	188	8	60	46	334

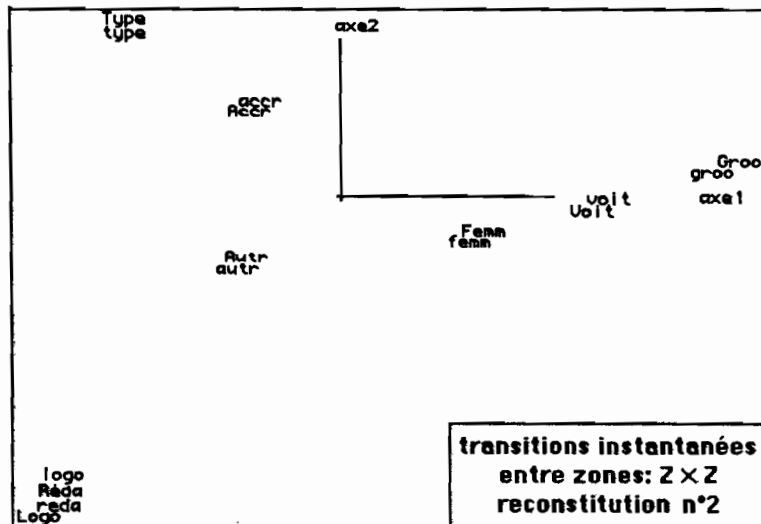
2.2.2 Analyse des transitions: deuxième reconstitution de la diagonale

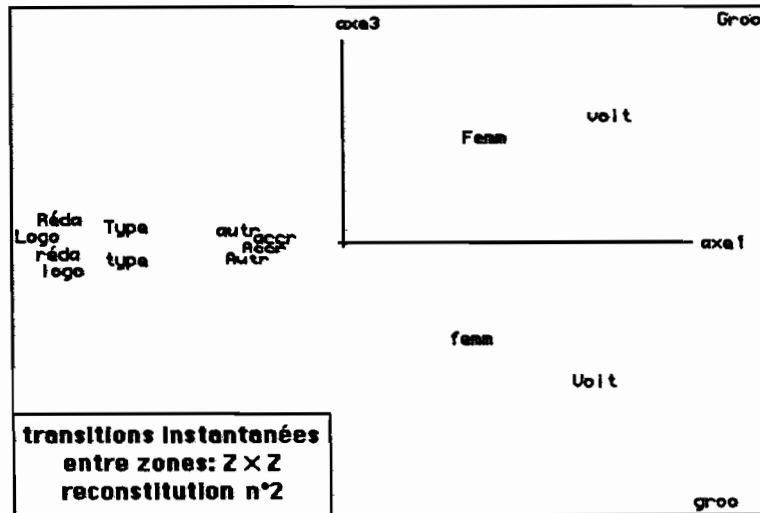
Transitions entre zones: 8x8 : reconstitution 2

trace :	5.157e-1						
rang :	1	2	3	4	5	6	7
lambda :	2452	1424	623	322	243	77	15 e-4
taux :	4755	2761	1208	625	472	150	29 e-4
cumul :	4755	7516	8724	9349	9820	9971	10000 e-4

Avec cette reconstitution de la diagonale, qui comporte une étape au delà de celle déjà analysée (cf. §2.2.1), la trace est un peu plus faible, la décroissance des valeurs propres est plus rapide. [Ainsi est plus parfaitement corrigé l'effet de la lourde diagonale du tableau brut, considéré au §2.2.0: l'analyse de celui-ci, non publiée, produit des valeurs propres : {0,7741; 0,74; 0,65; 0,55; 0,55 ...} qui rappellent la suite de 1 issue d'un tableau réduit à sa diagonale; avec des facteurs difficiles à interpréter.]

L'interprétation subsiste, à ceci près que les facteurs 2 et 3 sont échangés; les deux facteurs directs précédant le facteur inverse. Le plan (1x2) offre une partition des zones en trois classes: voiture et personnages, (F1>0); type et accroche, (F1<0; F2>0); rédactionnel et logo, avec le fond, (F1<0; F2<0).





L'axe 3 rend compte des transitions directes entre personnages et voiture: d'où les deux associations {Femm, voit, Groo} et {femm, Voit, groo}.

Comme au §2.2.1, la CAH confirme l'analyse factorielle: La partition en trois classes vue dans le plan (1 x 2) est aussi celle définie par les deux nœuds les plus hauts.

reconstitution 2 : Classification des antécédents

Type	10	F2+	14	//	F1-	
Accr						
Autr						
Logo	9					
Réda						
Femm		11	F3+	13	//	F1+
Groo						
Voit			F3-			

On notera que, les niveaux des nœuds 11 et 12 étant très proches, il y a, entre les CAH des deux ensembles, une différence de numérotage des classes. Il n'en reste pas moins que l'écart entre la matrice de transition et le modèle symétrique n'est que fluctuation.

reconstitution 2 : Classification des conséquents

type	10	F2+	14	//	F1-	
accr						
autr						
logo	9					
réda						
femm		12	F3-	13	//	F1+
groo						
voit			F3+			

3 Correspondance entre sujets, zones et tranches temporelles

3.0 La correspondance ternaire $S \times T \times Z$

Au cours de la tranche de temps t , le sujet s , a dirigé vers la zone z de l'image un nombre $k(s, t, z)$ de fixations instantanées. D'où un tableau de correspondance ternaire $S \times T \times Z$. Ce tableau ne peut être analysé tel quel; mais on en tire des tableaux rectangulaires.

D'abord, les tableaux de marge rectangulaires: $S \times T$, $T \times Z$, $S \times Z$. Le premier de ces tableaux est trivial, car, aux erreurs de saisie près, on a $k(s, t) = 60$ pour tout sujet s et toute tranche t . Restent $T \times Z$ et $S \times Z$, objets des analyses des §§3.1 et 3.3; d'où résultent deux typologies des zones, fondées, respectivement, sur la correspondance avec les tranches temporelles et avec les sujets.

Ensuite, il y a les tableaux croisant l'un des ensembles avec le produit des deux autres. Au §3.2, le tableau $(S \times T) \times Z$ est adjoint en supplément à l'analyse du tableau $T \times Z$: on apprécie ainsi la diversité des comportements visuels (s, t) qui peuvent être adoptés par un sujet au cours d'une tranche t .

Le tableau $S \times (Z \times T)$, quant à lui, donne du comportement de chaque sujet s le bilan le plus détaillé qu'on puisse faire, en rapportant les fixations aux tranches temporelles. Nous ne publions pas d'analyse de $S \times (Z \times T)$, mais au §4, on cherchera dans ce tableau une explication possible du comportement de réponse verbale des sujets, interrogés sur ce qu'ils ont retenu de la présentation des images.

T x Z : 5.tranches x 9.zones

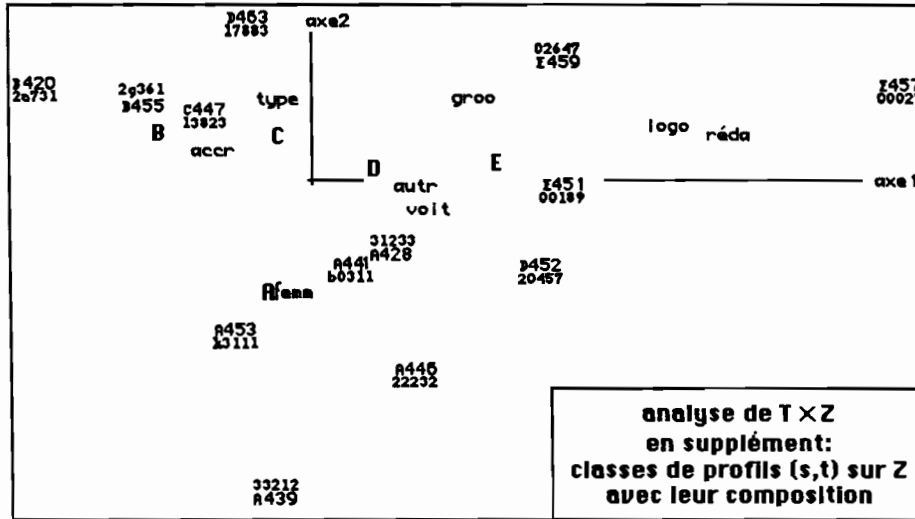
9	type	accr	voit	femm	groo	réda	logo	autr	erre
tA	116	1303	898	1535	26	5	0	736	87
tB	644	2424	231	586	152	0	10	503	146
tC	559	1741	768	448	429	8	6	599	145
tD	621	1050	958	580	264	80	105	914	138
tE	282	773	1055	470	552	241	173	1113	45

3.1 Correspondance entre tranches temporelles et zones de l'image

T x Z : 5.tranches x 8.zones (erreur en suppl)

trace :	2.304e-1				
rang :	1	2	3	4	
lambda :	1289	826	104	86	e-4
taux :	5593	3585	450	372	e-4
cumul :	5593	9178	9628	10000	e-4

La lecture du tableau, confirmée par l'analyse factorielle et la CAH, montre qu'en moyenne les sujets sont d'abord attirés par le personnage féminin: un tiers du temps au cours de la première seconde, A. Puis par l'accroche: 50% pour la deuxième seconde, B. Les parts relatives de la voiture et celle du fond (autr), minima en B, ont leur maximum en fin d'observation: E. Le rédactionnel et le logo, qui n'ont jamais un poids majeur, ne comptent



qu'à partir de D et surtout en E. Pour les grooms, également, les fixations, sinon l'attention véritable, croissent au cours de la présentation.

Sur le plan (1 x 2), on a marqué, outre les sigles des éléments des deux ensembles en correspondance, T et Z, des classes de couples (s, t), dont chacun est décrit par le bilan du comportement visuel d'un sujet s pendant une tranche de temps t; le numéro d'une classe, e.g. A453, est accompagné d'une suite de cinq caractères, telle que k3111, qui en indique la composition temporelle (cf. *infra*, §3.2).

tA				fem++++	8
tD	acc-	6	=CdG	7	
tC	acc+				
tB	acc+++	voi-			
tE	réd++	log+	gro+	acc-	

Il apparaît qu'en somme, la classification des zones d'après la correspondance avec l'ensemble T, des tranches d'une seconde, a peu de rapport avec celle, obtenue au §2.2, fondée sur les transitions instantanées: seules les classes 11 {type, accro} et 9 {réda, logo} peuvent être redevables d'une part de leur cohésion à la proximité sur l'image.

type	A-	11	B++	
accr	B+++			
femm			A++++	14
autr	10	13	B- E+	
voit				
groo		12	E++	
logo	9	E++++		
réda				

CAH des zones d'après leur correspondance avec les tranches {A, B, C, D, E}

c	ST × Z en suppl à T × Z: Partition de ST: Sigles des (s,t)		
420	01A 12B 18C 20B 29B 39B 37B 45B 10D 06B 10C 33D 30C 08C 13C 45C 02E 19B 02A 10B 05D 27C 43B	2a731	B
455	02B 26D 07C 05B 25B 41B 21B 38B 18B 11B 19D 09C 13B 09D 17A 24B 17B 28B 46D 36B 16B 06C 07B 40D 22B 45D 26E 30A	2g361	B
447	15C 46E 28E 25C 13D 01C 44B 31B 46C 40E 11C 45A 17C 35C 27D 34B 47C	13823	C
459	35E 34C 30E 32C 26C 38C 14D 27E 41E 33E 21E 31E 35D 22C 08B 42D 03C 09B 02D	02647	E
463	27A 37D 43D 06D 41D 15B 39C 40B 27B 21C 28C 17E 41C 08D 40C 04B 07E 28D 23B 30B 36C 29C 43C 18D 32B 23D 23E	17883	D
457	22D 29E 44E 37E 03E 16E 39E 03D 22E	00027	E
453	12A 24A 07A 42A 47B 29A 39A 16A 40A 42B 33B 30D 11A 37A 06A 43A 09A 23A 20A 45E 16C 15A 05A 26A 25A 19A	k3111	A
439	14A 14B 26B 34D 42E 14E 35A 01B 33C 44C 47A	33212	A
451	16D 01D 23C 17D 06E 36D 34E 47E 05E 47D 43E 07D 36E 32D 01E 20D 20E 08E	00189	E
452	19E 44D 09E 24E 11E 31C 11D 20C 46A 15E 12C 13A 39D 25E 14C 13E 25D 04D	20457	D
446	18E 46B 35B 38D 42C 38E 02C 34A 33A 24D 12D	22232	A
428	03B 21D 10E 24C 29D 32E 31A 04A 04C 12E 15D 08A	31233	A
441	28A 36A 05C 04E 37C 10A 19C 21A 38A 44A 31D 32A 41A 22A 03A 18A	b0311	A

3.2 Diversité des profils temporels d'attention des sujets sur les zones

Avec 47 individus et 5 tranches de temps, il y a $47 \times 5 = 235$ paires (s, t), dont le tableau ST × Z donne le profil d'attention sur l'ensemble des zones. On pourrait analyser ce tableau pour lui-même; mais nous préférons ici l'adjoindre en supplément à l'analyse du tableau T × Z, afin d'apprécier dans quelle mesure le profil sur Z d'une paire (s, t) est déterminé par la tranche temporelle t indépendamment du sujet s.

Plutôt que d'observer directement le nuage des (s, t) décrit chacun par ses coordonnées sur les axes factoriels, on effectue une classification de l'ensemble S × T, et on se borne à considérer le nuage des centres de gravité des classes.

Pour choisir la partition à retenir pour ST, on a procédé en deux étapes. Premièrement, on a considéré la partition de ST en 8 classes définie par les 7 nœuds les plus hauts. L'une des classes, numérotée 457, comprend 7 paires (s, E) et deux paires (s, D); dans les autres classes, il y a plus de diversité dans les tranches de temps. Descendant dans la hiérarchie au dessous des 7 nœuds les plus hauts, on s'est appliqué à prendre en compte les subdivisions se signalant par un profil sur T concentré sur une tranche, voire deux.

Ainsi a été retenue une partition en 13 classes dont nous donnons, en un tableau, le contenu explicite; avec, pour chaque classe, son bilan sur T. Par

exemple, la classe 455, comprend 2.A, 16.B, 3.C, 6.D, 1.E; ce qu'on écrit: 2g361; où la lettre g, qui vient après f=15 du code hexadécimal, signifie 16.

(47.Sujets x 5.tranches) x 8.zones

13	B420	B455	C447	E459	D463	E457	A453	A439	E451	D452	A446	A428	A441
type	43	355	41	25	1563	30	32	1	105	0	0	11	16
accr	2050	1770	1083	269	405	36	760	12	133	74	16	250	433
voit	13	64	301	431	164	21	197	134	284	1050	361	221	669
femm	24	144	42	81	65	41	1135	820	77	176	502	230	282
groo	17	6	155	905	5	89	2	11	20	77	113	5	18
réda	5	0	0	0	5	224	0	2	82	6	2	8	0
logo	0	5	0	14	3	192	5	0	73	0	2	0	0
autr	58	444	44	135	420	265	421	43	993	403	44	444	151

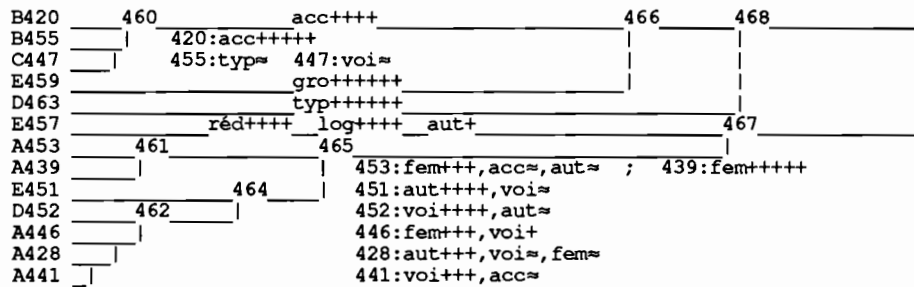
Un tableau de contingence donne le bilan sur Z de chaque classe.

De plus, par un programme usuel d'analyse discriminante, on a déterminé, pour chaque classe, celle des tranches temporelles dont elle est le plus proche. On ne s'étonnera pas de voir affectée à B la classe 455, déjà citée: on écrit donc B455; etc.

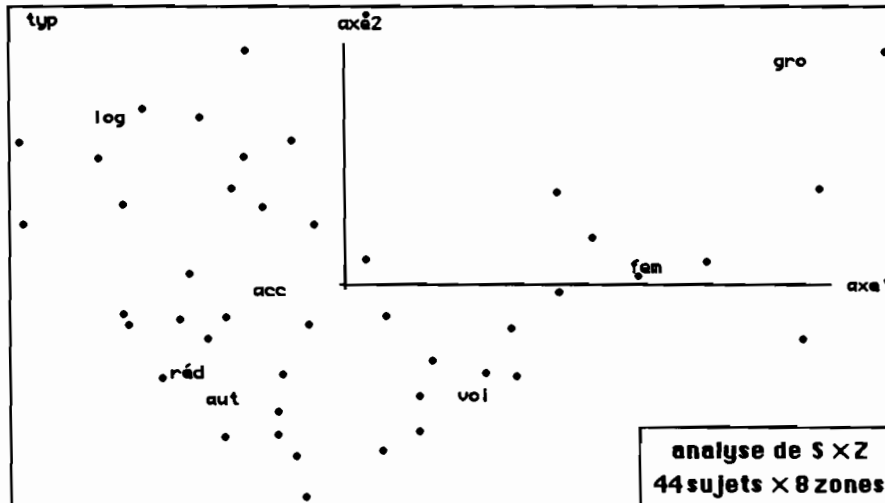
affect	des cl	de	(s,t)	à T	
5	->A	->B	->C	->D	->E
A	39	4	1	3	0
B	9	26	3	7	2
C	10	10	8	12	7
D	9	9	2	13	14
E	9	2	3	10	23

D'après ces affectations, on a créé un petit tableau 5 x 5: on lit, e.g., à l'intersection de la ligne B et de la colonne ->A le nombre 9, total des paires (s, B) rentrant dans une des classes affectées à la tranche A.

Enfin, l'arbre de la CAH est étiqueté en terme de zones d'après le listage VACOR.



De ces divers documents, on conclut que la tranche de temps A, 1-ère seconde, est la mieux caractérisée; par son association avec femm (classe A453) ou avec voit (A441). La 2-ème seconde, B, a également un caractère net, étant liée à l'accroche dans B420 et B455 (cette classe-ci allant également quelque peu avec type et voit). Les tranches C et D, aux affectations discordantes, font la transition vers E. Cette 5-ème tranche est d'abord, dans 457 liée au rédactionnel et au logo; on trouve aussi, dans 451 des paires (s, D) et (s, E) dont les regards se perdent dans le fond; et il y a encore, dans 459, de surprises fixations sur les grooms.



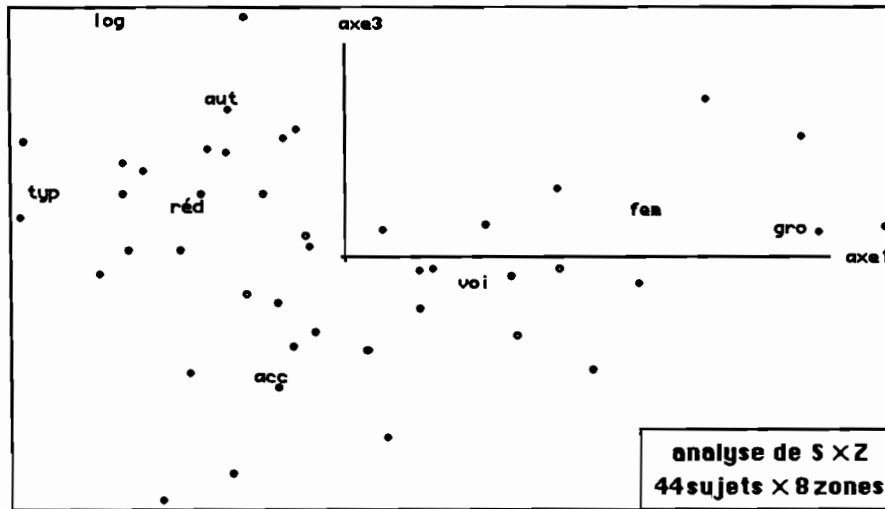
3.3 Affinité globale entre sujets et zones

S x Z	: 44 (+3) .Sujets x 9.zones							
trace	: 4.634e-1							
rang	1	2	3	4	5	6	7	
lambda	: 1794	878	676	482	330	294	180	e-4
taux	: 3871	1895	1459	1041	711	634	389	e-4
cumul	: 3871	5766	7225	8266	8977	9611	10000	e-4

Diverses analyses ont signalé que, durant les secondes {B D E}, le sujet 23 ne s'intéresse pas du tout à la zone centrale, mais maintient son regard en haut de l'image; le sujet 22 se fixe sur le logo au cours de la quatrième seconde, D; le sujet 03 est lié au bas de l'image durant {D E}. En raison de ces différences par rapport au profil moyen, ces sujets apporteraient des contributions très fortes à l'analyse de S x Z; ils sont donc mis en supplément dans l'analyse rapportée ici.

La première valeur propre est double de la deuxième; mais celle-ci est proche de la troisième: on publie donc les plans (1 x 2) et (1 x 3).

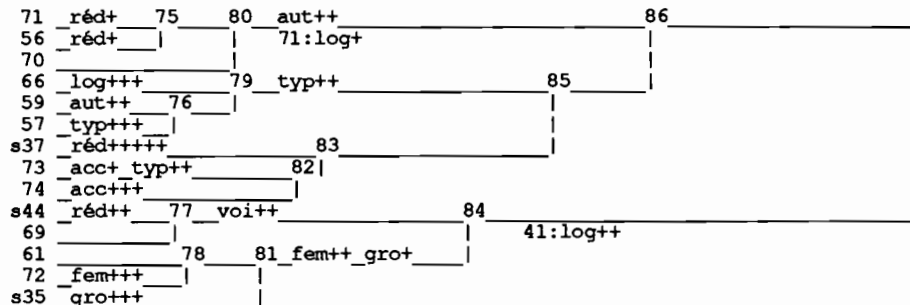
01 05 47 20 16	75	80	F2-	F3+	86	//	F1-
08 07 39	F1--						
15 06 04 12 25	F2--						
29 36 43	F2++	F1--	79	F2+	F3+	85	
41 32	.76						
17 40 28	F1---						
37	F4--		83	F3-			
18 45 27 30	F2++		82				
21 46 02 13 10	F3--						
44	F4--	77			84	//	F1+
11 24 19 31 09 38	F2-						
26 33	78	81	F1++	F2+			
34 14 42	F1+++						
35	F2+++						

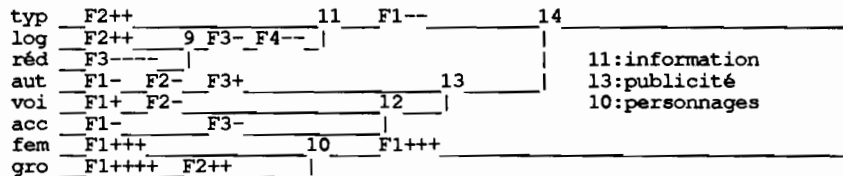


Les sujets {03, 22, 23} ayant été écartés, le nuage N(S) n'a pas de point isolé; et recouvre assez uniformément N(Z). Sur (F1>0), le point le plus écarté est gro: on sait déjà que le groupe des grooms est l'objet d'un nombre assez faible de fixations que la CAH permet d'attribuer à certains sujets.

Pour plus de précision, on a muni la CAH des sujets de deux étiquetages, l'un par les facteurs, l'autre par les zones. On doit signaler que réd et log n'ont de poids non négligeable que pour les classes explicitement étiquetées par ces zones.

Sur l'arbre de S étiqueté par les facteurs, on a donné la composition explicite des classes: ce qui offre matière à des comparaisons avec la classification des profils (s, t), par tranches de temps, donnée au §3.2: ainsi la fixation du sujet 35 sur les grooms, est propre aux tranches E et D (secondes 4 et 5): c'est pourquoi, au §2.2, 35D et 35E sont dans la classe E459; ici, s35 est étiqueté gro+++; plus généralement, l'étiquette gro+ vaut pour la classe 81.





La classification de l'ensemble Z des zones offre une partition en 3 classes qui admet une interprétation sémantique simple. Au sommet de la hiérarchie, se sépare la classe 10{fem gro}, qui comprend les deux personnages.

Le reste, 14, se divise en 11{typ log réd} et 13{aut voi acc}: 11, qui sur l'axe 1 s'oppose à 10, renferme les informations précises que comporte l'image; 13, {voiture et slogan avec le fond}, accroche les regards sur l'objet même de la publicité.

Cette classification ne reproduit aucunement l'une ou l'autre des deux précédentes (des §§2.2 et 3.1); seule l'agrégation de réd avec log subsiste partout.

4 Relation entre comportements oculaires et réponses verbales

4.1 Le questionnaire

On a dit que, d'une série de trente images présentées au sujet, celle retenue pour notre étude particulière présente une voiture de luxe, un break rouge, avec des personnages: une femme et des portiers d'hôtel (grooms) vêtus de rouge. Pour rappeler cette image à la mémoire, on pose la question suivante:

Parmi les images publicitaires qui viennent de vous être présentées, y en avait-il une comportant une voiture rouge et des personnages ?

avec les trois modalités de réponse:

Je ne m'en souviens plus [] ; NON [] ; OUI [] .

Est ensuite posée, s'il y a lieu, une question complémentaire conditionnelle:

Si vous avez gardé un tel souvenir, dites:

quelle était la marque ?

quel était le modèle ?

Pour tenir compte de l'ensemble du comportement de réponse de chaque sujet, on a retenu un code en six modalités; dont la dernière s'est trouvée être sans objet:

??? : ne se rappelle plus s'il a vu ou non une telle image
 mmF : dit se souvenir de l'image mais donne des informations complémentaires fausses
 mmN : dit n'avoir pas vu d'image de cette sorte
 mmO : dit avoir vu l'image, mais ne répond rien à la question complémentaire
 mrq : dit avoir vu; et a bien retenu la marque de la voiture
 mtt : dit avoir vu; et a bien retenu marque et modèle.

Aucun sujet ne rentrant dans la modalité mtt, il ne reste qu'un ensemble, noté Q, cinq modalités, dont les fréquences respectives sont données ci-après:

Q : {??? mmF mmN mmO mrq} : {18 5 8 13 3} .

Ainsi, il n'y a que 21 sujets sur 47 qui déclarent avoir vu une image telle que celle que suggère la question; et 3 seulement ont retenu la marque.

4.2 Correspondance entre réponses et fixations sur zones, sans distinction temporelle

Afin de mettre en rapport comportement oculaire et comportement de réponse verbale, on construit un tableau de correspondance $Q \times Z$, dont chaque ligne, afférente à une modalité de réponse q, est obtenue à partir du tableau $S \times Z$ (cf. supra §3.3) par cumul des lignes des sujets s rentrant dans la modalité q.

S.Z+QLC9& : 5.réponses × 8.zones								
8	type	accr	voit	femm	groo	réda	logo	autr
???	642	2524	1501	1577	745	175	184	1400
mmF	440	797	377	300	117	0	0	425
mmN	438	1265	772	658	205	106	23	518
mmO	666	2164	1036	846	308	33	75	1160
mrq	36	541	224	238	48	20	12	362

On voit sur ce tableau de cumul par réponse que pour les 5 sujets rentrant dans la modalité mmF (répondre qu'on a vu une telle image, mais en donnant des réponses complémentaires fausses) aucune fixation n'a été relevée sur le rédactionnel ni sur le logo; mais le caractère réda = logo = 0 appartient à 24 des 47 sujets; ce qui interdit de l'utiliser pour discriminer l'attitude mmF.

affectation aux s de modalités de Q					
5	???	mmF	mmN	mmO	mrq
???	8	0	2	4	0
mmF	3	2	2	4	0
mmN	0	0	1	0	1
mmO	4	1	1	1	0
mrq	3	2	2	4	2

Une autre voie s'offre pour déterminer la réponse à partir du comportement visuel cumulé sur T: analyser le tableau $Q \times Z$ comme principal, $S \times Z$ étant en supplément; et recourir à l'analyse discriminante; i.e., dans l'espace engendré par les 4 axes factoriels obtenus, affecter à chaque sujet s la réponse q dont il est le plus proche: la matrice de confusion obtenue ne se distingue pas de ce qu'on pourrait obtenir par tirage au sort.

4.3 Correspondance entre réponses et fixations sur zones, ventilées selon les tranches temporelles

On précise l'analyse en distinguant les 5 tranches temporelles: $T = \{A, B, C, D, E\}$; d'où un tableau $S \times ZT$, $47 \times (8.5)$, avec $k(s, zt) =$ nombre de visées de s sur z pendant la tranche de temps t . On crée (comme pour $S \times Z$) un cumul $Q \times ZT$, avec $k(q, zt) =$ total des visées sur z enregistrées pendant la tranche de temps t pour un sujet s ayant fourni la réponse q .

affectation aux s de modalités de Q					
5	???	mmF	mmN	mmO	mrq
???	11	0	0	1	0
mmF	2	3	1	2	0
mmN	1	0	7	2	0
mmO	4	2	0	8	0
mrq	0	0	0	0	3

Ici, l'analyse discriminante, fournit un excellent résultat; la somme des termes diagonaux étant 32, on a 32 affectations justes sur 47. Par exemple, des 8 sujets rentrant dans la modalité mmN, 7 reçoivent de l'analyse discriminante cette même modalité; un seul reçoit la modalité mmF; cependant que mmF n'est attribué faussement qu'à 3 sujets: dont un a répondu ???; et deux autres mmO.

Mais ce succès n'est qu'illusoire; n'étant dû qu'à ce que le nombre de dimensions prises en compte est de l'ordre du nombre des sujets à discriminer ($\text{card}S \approx \text{card}ZT$): cf. e.g. [SÉP. CORR.] in *CAD*, Vol.II, n°4, 1977; et [SÉP. ALÉAT.] in *CAD*, Vol.VII, n°4, 1952.

Reprenons en effet l'analyse en ne comprenant dans la ligne de cumul afférente à un mode de réponse q que les sujets qui non seulement ont fourni la réponse q mais encore ont un numéro d'identification non divisible par 3. Ceci revient à réserver comme échantillon d'épreuve les 15 sujets dont le numéro est divisible par 3. La discrimination réussit bien pour les autres sujets; mais échoue pour l'échantillon d'épreuve; même si, au lieu de chercher à prédire q d'après le comportement oculaire on se contente de considérer la dichotomie entre les sujets qui ne répondent pas ou disent n'avoir pas vu {???, mmN} et les autres.

En l'état de l'expérimentation, le comportement oculaire, certes cohérent, ne fournit donc que de modestes indices sur l'information que le sujet a retenue du défilé des images. Peut-être la discrimination serait-elle meilleure si, au lieu de ne considérer que le comportement oculaire devant l'image publicitaire à laquelle on s'intéresse particulièrement, on prenait en compte le comportement devant toutes les images; et s'enquerrait, par questionnaire, de ce qui a pu être retenu non d'une, mais de plusieurs, des images de la suite présentée.

5 Perspectives et conclusions

Les analyses des §§2.2, 3.1 et 3.3 ont fourni, des parties, ou zones, de l'image, plusieurs schémas successifs cohérents allant de l'espace à la signification.

Découper la surface de l'image en zones floues et analyser les taux relatifs d'observations issus de ce découpage est un des points essentiels de la méthode utilisée ici. D'une part, ce découpage respecte l'imbrication des zones de l'image publicitaire (zones: slogan, objet, sujet, descriptif, logo...); d'autre part, il permet de prendre en compte l'erreur de mesure.

Il reste que si le capteur nous informe de la direction du regard, nous n'avons aucune connaissance sur la taille de la zone observée par le sujet; ce qui laisse en question le rayon effectif de diffusion spatiale de l'attention durant une fixation; rayon qui varie nécessairement selon qu'il s'agit d'une vision fovéale ou périphérique.

Les petits mouvements de l'axe de visée (cf. §2.1) pourraient, toutefois, en suggérer un ordre de grandeur. Et il vaudra la peine d'étudier sur toute une série d'images, et non seulement sur une seule, la distribution de ces mouvements; en distinguant entre images et entre sujets, notamment pour apprécier dans quelle mesure ce paramètre de mobilité conditionne l'information retenue; dont on a vu au §4 qu'elle ne pouvait être estimée valablement d'après le comportement visuel des sujets devant une seule image.

En tout état de cause, une description quantitative du comportement visuel, par exemple basée sur des modèles probabilistes de taux d'occupations ou de transitions spatiales, dans le cadre de l'observation d'images complexes et fortement connotatives telles que les images publicitaires, doit prendre en compte la valeur sémantique des zones. De ce point de vue, la classification des zones obtenue au §3.3, a le mérite de montrer qu'échappant, au moins en partie, aux suggestions de la contiguïté, les sujets distribuent les préférences de leurs regards selon des critères sémantiques.

Cependant, selon les analyses des §§3.1 et 3.2, les zones de l'image se distinguent quant à la priorité temporelle autrement que par la stricte signification. Cette priorité, sans être universelle, est partiellement commune à l'ensemble des sujets. Les profils de fixation des tranches A, B et E se groupent bien en classes, même si E offre plusieurs types bien distincts: certains sujets, entraînés à apprécier le temps de présentation, se concentreraient sur E; d'autres se désintéresseraient de l'image. Les tranches C et D sont des passages s'agréant à divers types.

Nous concluons que l'analyse des données oculométriques doit être approfondie aussi bien dans le domaine publicitaire (images fixes ou dynamiques) que dans la communication homme-machine, par exemple.

Références bibliographiques

B. ACIER : *Contribution à la conception et à l'évaluation d'interfaces hommes-machines par l'analyse du comportement visuel d'un superviseur humain*; Thèse, Université de Valenciennes et du Hainaut-Cambrésis; Janvier 1990.

Y. BATICLE : *Clés et codes de l'image*; Editions Magnard, Paris, 1985.

J. R. CARBONELL : "A queuing model of many-instrument visual sampling"; in : *IEEE Transactions on Human Factors in electronics*, Vol HFE-7, n°4, pp.157-164; December 1966.

B. CATHELAT : *Socio styles système*; Editions d'organisation, 1990.

J.-C. CHEBAT & B. GAUTIER : "La rhétorique au service de la publicité"; in: *Communication & langages*, n°38, pp.103-116; 1977.

J. DURAND : "Figures de rhétorique et image publicitaire"; in: *Communication*, n°15, pp.70-95; Paris, Seuil; 1970.

U. ECO : "Sémiologie des messages visuels"; in: *Communication*, n°15, pp.11-51; Paris, Seuil; 1970.

S.R. ELLIS, L. STARK : "Statistical dependency in visual scanning"; in: *Human factors*, Vol. 28, n°4, pp421-483; 1986.

A. LÉVY SCHÖEN : "L'étude des mouvements oculaires"; in: *Revue des techniques et des connaissances*; Paris, Dunod; 1969.

LOFTUS : "Tachistoscopic simulation of eye fixations on picture"; in: *Journal of experimental psychology : human learning and memory*; volume 7, n°5, pp.369-376; 1981.

F. MOLNAR : "About the role of visual exploration in aesthetics"; in: *Advances in intrinsic motivation and aesthetics*; H.I. DAY ed., Plenum Publishing corporation; 1981.

F. MOLNAR : "Some æsthetical aspect of visual exploration"; in: *Eye movements : from physiology to cognition*; J.K.O'Regan et A. Lévy-Schoen, edd.; Elsevier Science Plublishers B.V, North-Holland 87; pp367-374; 1987.

PAPIN : *L'exploration visuelle de radiographies du thorax*; Rapport de recherche; C.E.R.M.A; Juillet 1981.

M.I. STRIEB & W. GROVE: "Visual performance modeling in the human operator simulator"; in: *Model of Human Operators in vision dependent task*; C. MARVIN ed., NASA conference publication, n° 2013, pp.19-32, 1979.

Van der HEIJDEN : "Transition matrices, model fitting and correspondence analysis"; in: *IV-èmes journées internationales, Analyse des données et Informatique*, Versailles, pp.221-226; 1985.

M.-C. VETTRIANO-SOULARD : "La publicité des voitures de luxe"; in: *Communication & langages*, n°74, pp.90-109; 1987.

M.-C. VETTRIANO-SOULARD : "L'image publicitaire des boissons de luxe"; in: *Communication & langages*, n°71, pp.90-111; 1987.

M.-C. VETTRIANO-SOULARD : "L'image publicitaire des stylos de luxe"; in: *Communication & langages*, n°78, pp.72-90; 1988.

YOUNG & SHEENA : "Survey of eye movement recording methods"; in: *Behavior Research Methods and Instrumentations*, Vol 7(5), pp.397-429; 1975.