

Université Nancy 2

Thèse présentée pour l'obtention du
doctorat de Psychologie

par
Emmanuelle TOGNOLI

**Les conjonctions illusoires en RSVP :
un paradigme pour l'étude
des synchronisations neuronales ?**

Pr. J. Lautrey : Président du jury
Dr. M. Boucart : Rapporteur
Pr. U. Lindenberger: Rapporteur
Dr. A-M. Toniolo
Pr. Paul Dickès : Directeur de thèse



Laboratoire de Psychologie
GRAPCO
BP 3397
54015 Nancy Cedex



The intuitive mind is a sacred gift, and the rational mind is a faithful servant. We have created a society that honors the servant and have forgotten the gift.

Albert Einstein



SECTION 1 : REMERCIEMENTS	6
SECTION 2 : INTRODUCTION	8
SECTION 3 : CADRE THEORIQUE	11
1. PARADIGMES DU CODAGE DE L'INFORMATION	12
<i>Le paradigme dominant de la neuropsychologie classique</i>	13
Le code de l'information : où est le problème ?	13
Les localisations : « où ? » est le problème !	16
Hiérarchie du traitement	17
<i>Ecueils</i>	19
Les contraintes temporelles en vision	19
Le « binding problem » et ses implications neurophysiologiques	21
La réponse des neurones de V1 chez le chat amblyope	25
<i>La théorie des assemblées transitoires synchrones</i>	26
Avant propos	26
Fondements théoriques	26
Vers des évidences empiriques : évolution des méthodes	31
Observations des synchronisations en vision	38
2. LE TRAITEMENT DE L'INFORMATION VISUELLE	53
De la discontinuité : capture et traitement de l'information visuelle	54
La notion d'objet : restauration d'une continuité spatio-temporelle	63
Le décours des opérations perceptives pour l'identification de l'objet	68
<i>Résumé</i>	72
3. PARADIGMES DE LA RECHERCHE, MISE EN PLACE DU DISPOSITIF EXPERIMENTAL ET	
HYPOTHESES OPERATIONNELLES	74
<i>Expériences préliminaires</i>	74
Expérience A : Sujets	75
Matériel et méthodes	75

Résultats	76
<i>Bases théoriques et hypothèses</i>	90
Résultats issues des paradigmes de recherche spatiale.....	92
Conjonctions illusoires en RSVP : recherches classiques	93
Problèmes de terminologie, problèmes théoriques.....	100
Conjonctions illusoires complexes en RSVP : à la recherche des chimères	104
<i>Résumé</i>	107
SECTION 4 : OPERATIONNALISATION	108
4. DES CONJONCTIONS ILLUSOIRES ?	109
<i>Commun à l'ensemble des expériences</i>	109
Aménagements du protocole RSVP	109
Matériel & méthodes	110
<i>Expérience 1 : Modulations temporelles</i>	112
Stratégie	113
Sujets	113
Matériel et méthodes	113
Résultats	115
Discussion.....	118
<i>Expérience 2 : Effets d'exercice & dissociation des intrusions proximales</i>	120
Stratégie	121
Sujets	121
Matériel & méthodes	121
Résultats	122
Discussion.....	127
<i>Expérience 3 : le cadre migrateur ?</i>	130
Stratégie	131
Sujets	131
Matériel et méthodes	131
Résultats	134
Discussion.....	136
<i>Expérience 4 : intégration et interruption ?</i>	138
Stratégie	139
Sujets	140
Matériel & méthodes	140

Discussion.....	136
<i>Expérience 4 : intégration et interruption ?.....</i>	<i>138</i>
Stratégie.....	139
Sujets.....	140
Matériel & méthodes.....	140
Résultats.....	141
Discussion.....	145
<i>Expérience 5 : Sélection des conjonctions illusoires.....</i>	<i>147</i>
Stratégie.....	148
Méthode.....	148
Résultats.....	148
Discussion.....	153
SECTION 5 : CONCLUSION.....	154
SECTION 6 : BIBLIOGRAPHIE.....	160
SECTION 7 : ANNEXES.....	172

Remerciements

En tout premier lieu, je remercie Jean, Josette, Bertrand, Jacqueline, et le reste de ma famille. Vous avez cru en la réalisation de ce projet, vous m'avez soutenu, vous m'avez aidé. Veuillez accepter mes plus sincères remerciements.

Cette recherche fut abritée par trois institutions : L'équipe GRAPCO du laboratoire de Psychologie de Nancy (dirigée par Pr. Paul Dickès puis Pr. André Flieller), le laboratoire de Biologie du Comportement de Vandoeuvre-lès-Nancy (Pr. Bertrand Krafft & Pr. Didier Desor) et le laboratoire de Neurosciences Fonctionnelles & Pathologies de Lille (Dr. M. Boucart).

Pr. Paul Dickès, vous nous avez communiqué une chose infiniment précieuse : « l'esprit de Paul », une curiosité, une ouverture, une tranquillité. Vous nous avez laissée voguer à des recherches dont les soubassements et les règles sont différentes des vôtres, avec respect et tolérance. Grâce à vous, nous pourrions continuer un beau projet sur un « long fleuve tranquille » nommé Boca-Raton. Du fond du cœur, acceptez mes remerciements les plus chaleureux.

Dr. Anne-Marie Toniolo, vous avez soutenu la première partie de ce travail. Vous avez assisté et suscité la progression du travail théorique sur les synchronisations. Vous avez entouré cette recherche d'une pluridisciplinarité, à travers des rencontres choisies. Veuillez acceptez mes remerciements.

Dr. Muriel Boucart, vous nous avez soutenu dans la seconde partie de ce travail. Vous nous avez poussé dans la réalisation de ce projet. Vous nous avez enseigné un pragmatisme et une méthode de travail qui ne nous était pas familière, mais dont nous mesurons pleinement les avantages. Nous avons beaucoup appris à votre contact. Je vous prie d'accepter mes plus sincères remerciements.

Pr. Jacques Lautrey, nous vous avons écouté en plusieurs occasions et spécialement à Bonas en Septembre 2000. Votre ouverture d'esprit et votre curiosité ont été pour nous un exemple et une source d'inspiration. Je vous remercie d'avoir accepté de présider le jury de ce travail.

Pr. Ulman Lindenberger, nous vous avons rencontré à trois reprises, à Bonas, Nancy, et enfin à Sarrebruck, où vous étiez un des maîtres de cérémonie d'une enrichissante conférence sur le binding. Nous avons également été marqué par votre ouverture d'esprit, votre curiosité et votre implication. Je vous remercie d'avoir accepté de figurer parmi les juges de ce travail.

Au laboratoire de Biologie du Comportement, les conversations furent animées et enrichissantes... « de l'or sur lequel notre éveil scientifique était assis ». Chers amis, Didier Desor, Henri Schroeder, Alexandre Bernard, Guillaume Vanakas, Andrée, Daniel, Chantal, Christian, Jules, veuillez recevoir toute ma gratitude.

Parmi les membres du GRAPCO, je souhaiterais adresser des remerciements particuliers à Jean-Luc Kop, ainsi qu'à Jocelyne Tournois, Christine Bocéréan, Claude Houssemand, Romain Martin et Eric Thiebaut.

Parmi les membres du laboratoire de Neurosciences Fonctionnelles et Pathologie, je remercie Pascal Desprez, Fatima Naïli, Pr. Jean-Claude Hache, Danièle Basset, Pr. Pierre Thomas, Claire Rascle, & Delphine Pins, pour leur amical soutien et Pr. Olivier Godefroy, Pr. Patrick Berquin et Pr. Gwénoé Loas pour leurs encouragements.

Les expériences préliminaires figurant dans ce travail ont bénéficié de l'aide de Jean-Luc Bragard, qui nous a permis de nous introduire dans deux écoles de la région de Pont-à-Mousson. Nous remercions également les directeurs, enseignants, élèves et parents d'élèves de ces deux écoles (Ecole Saint-Martin et Ecole Guynemer), pour avoir permis la réalisation de ces expériences.

Nous remercions Bertrand Tognoli, pour nous avoir aidé dans la programmation des expériences préliminaires et pour ses conseils et informations sur le matériel et les logiciels informatique.

Nous avons également eu grand avantage dans des discussions et réunions avec l'équipe Cortex du LORIA, et particulièrement Dominique Martinez et Frédéric Alexandre. C'est en partie de ces interactions qu'a émergé cette conviction de l'intérêt d'une « Psychologie Connexionniste ».

Au cours de cette thèse, nous avons également rencontré Daniel Liévin et Martine François du laboratoire EPAP de l'INRS. Nous vous remercions pour votre amical soutien.

Nous remercions également le réseau Cogniscience EST (Frédéric Alexandre et Georges DiScala) pour le soutien d'un stage en électrophysiologie, et pour le contexte général d'interdisciplinarité dans lequel il nous ont permis de baigner.

Enfin, cette recherche fut soutenue par une bourse de la Fondation Cetelem. Je remercie cette fondation pour son soutien financier et le crédit qu'elle a porté à ce travail.

et...cette dernière ligne pour tous les oubliés de ces remerciements, qu'ils veuillent bien pardonner ma mémoire rouillée et qu'ils acceptent ces remerciements plus anonymes.

Introduction

Comprendre comment se transmet l'information dans le cerveau est d'un intérêt primordial pour les Sciences Cognitives et les Neurosciences Intégratives. Depuis 150 ans, une conception basée sur les découvertes de la Neuropsychologie, l'Histologie et la Neurophysiologie assoit son hégémonie (Finger, 2000). Il s'agit de l'idée selon laquelle l'information se transmet dans le cerveau par un système de voies et modules, et que le code de cette transmission est la modification de la fréquence de décharge du neurone. Cette conception théorique du traitement de l'information recueille à ce point l'unanimité qu'elle ne possède pas une appellation univoque. Pour y faire référence, nous utiliserons le terme de « théorie neuropsychologique classique ».

Depuis une quinzaine d'années, sont apparues de nouvelles propositions théoriques qui concurrencent (ou complètent) la théorie neuropsychologique classique. L'une d'elles défend l'idée selon laquelle l'information peut également se transmettre par la mise en place de synchronisations transitoires au sein d'assemblées de neurones. Cette théorie du « liage par synchronie » (ou *binding* par synchronie) est intéressante, car elle autorise une souplesse et une rapidité de traitement qui, constatées empiriquement, n'avaient pas de support satisfaisant dans la théorie neuropsychologique classique. Les paradigmes de la neuropsychologie classique et du *binding* par synchronie seront discutés au chapitre 1.

La fonction supposée du « *binding* par synchronie » est également susceptible de constituer une avancée significative pour les Sciences Cognitives. Elle répond au « *binding problem* » (problème de l'intégration), posé par Anne Treisman en 1980 (Treisman & Gelade, 1980 ; synthèse dans Treisman, 1996). Partant de la constatation qu'aux étapes précoces, la perception visuelle recourt à de nombreux processus parallèles, Anne Treisman s'interroge sur la manière dont le cerveau « reconnaît » ces différents traitements, topographiquement distribués, comme appartenant à l'analyse d'un même objet. Ce problème d'intégration, au delà de la perception visuelle, est un problème générique concernant tous les aspects du traitement cognitif. Depuis, on a attribué au *binding* par synchronie cette fonction d'intégration des caractéristiques traitées de manière distribuée. La représentation, au sens cognitif, est fondée par le recrutement de tous les neurones impliqués dans une assemblée coalisée par une synchronisation transitoire.

Depuis quelques années, des arguments empiriques viennent étayer la théorie du *binding* par synchronie. Elles concernent le fonctionnement neuronal, et la fonction supposée du *binding* par synchronie : l'intégration. Les recherches sur le fonctionnement neuronal regroupent les enregistrements électrophysiologiques profonds

(essentiellement chez le chat et le singe) et des analyses spectrales du signal EEG de surface (oscillation et cohérence). Ils visent à montrer l'apparition de synchronies dans des populations de neurones lors d'évènements perceptifs, cognitifs ou moteurs. Les analyses concernant la fonction cognitive du *binding* par synchronie concernent essentiellement l'intégration des caractéristiques visuelles précoces traitées en parallèle dans des paradigmes tels que la recherche visuelle ou la *gestalt*.

Les recherches en Neurosciences sur le *binding* par synchronie ont désormais un impact important dans la communauté scientifique. En revanche, les recherches cognitives sur l'intégration des traitements parallèles sont moins influentes, sans doute parce que leur lien avec le *binding* par synchronie reste théorique. Un pas important sera franchi lorsqu'il aura été possible de mettre en relation directe une tâche nécessitant l'intégration de traitements et l'observation de synchronisations de populations de neurones. Le présent projet s'inscrit dans cette démarche.

L'objectif est de construire un instrument de mesure du *binding* par synchronie. Il s'agira d'une situation expérimentale nécessitant l'intégration de caractéristiques visuelles de bas-niveau, et impliquant un réseau neuronal connu pour (1) sa topographie cérébrale, (2) les fonctions cognitives qu'il sous-tend, et (3) de décours temporel des différents processus qui le constituent. Ces éléments de neuropsychophysologie visuelle seront présentés chapitre 2.

Une telle situation expérimentale sera ensuite disponible pour des investigations en imagerie et électrophysiologie, afin de tester l'hypothèse d'un lien entre intégration des traitements parallèles et mise en place de synchronisations transitoires. A cette fin, la présente thèse essayera de montrer que les conjonctions illusoire, erreurs d'identification se manifestant lors d'une recherche visuelle avec distracteurs dans le domaine temporel (paradigme RSVP, *rapid serial visual presentation* ou présentation visuelle sérielle rapide), constituent un cadre d'étude pertinent.

Le terme de présentation visuelle rapide est applicable dès lors que la vitesse de défilement des stimuli successifs surpasse la durée d'une fixation oculaire (soit environ 250 ms par image). Les premières études notables recourant au paradigme RSVP sont celles de Potter & Levy (1969) qui utilisent des images, et Lawrence (1971), qui utilise des mots. Le terme « *Rapid Serial Visual Presentation* » n'apparaît cependant qu'en 1982 dans le titre de la publication de McLean, Broadbent et Broadbent, (1982).

Les conjonctions illusoire, initialement étudiées dans des paradigmes de recherche visuelle spatiale sont envisagées dans le paradigme de recherche visuelle temporel RSVP par Intraub, (1985, 1989) sous l'appellation de dissociation visuelle, mais seulement en 2001 par Botella, Barriopedro & Suero, sous l'appellation de conjonctions illusoire, bien que des publications intermédiaires s'approchent de ce concept sans le nommer (Botella & Ericksen, 1991, 1992 ; Botella, Garcia & Barriopedro, 1992, Boucart, de Visme & Wagemans, 2000 ; Moroni, Boucart, Humphreys, Henaff & Belin, 2000). Les conjonctions illusoire se manifestent lorsque la rapidité de défilement des images rend difficile l'identification de la cible. Elles semblent justement être le résultat d'une intégration inappropriée de certaines caractéristiques de la cible et d'un distracteur. Alors que l'intégration perceptive est usuellement étudiée à

travers des paradigmes « positifs » (l'expérimentateur met le sujet en situation de réussir une perception grâce à l'intégration perceptive), nous proposons de l'étudier à travers ce paradigme négatif (l'expérimentateur provoque une défaillance perceptive chez le sujet, cette défaillance procède d'une intégration perceptive inappropriée, dont elle révèle les règles de fonctionnement). Les études consacrées aux conjonctions illusoire, que ce soit en recherche visuelle spatiale ou temporelle (RSVP), seront présentées au chapitre 3.

La partie expérimentale de cette thèse est présentée au chapitre 4. Elle comportera plusieurs étapes. Premièrement, nous essayerons de dégager un faisceau d'arguments légitimant l'utilisation des conjonctions illusoire en RSVP dans cette perspective d'analyse de l'intégration perceptive. Les expériences 1 à 4 supporteront cet objectif.

En second, nous chercherons à filtrer, parmi toutes les erreurs d'identifications observées dans la situation expérimentale, celles qui sont le produit d'une intégration fallacieuse (conjonctions illusoire), de celles qui résultent d'autres mécanismes (simples intrusions). Les expériences 1, 2 et 4 poseront les premiers éléments de réponse, et l'expérience 5 approfondira cette question.

Ayant retenu des critères pour sélectionner les seules conjonctions illusoire, il s'agira ensuite de faire émerger une modélisation statistique du phénomène. Nous retiendrons en tant que variables indépendantes des mesures de certaines caractéristiques de l'image (eg. couleur, luminosité, taille). Notre variable dépendante sera constituée du degré d'association entre paires de stimuli par les conjonctions illusoire. Cet objectif sera abordé à l'expérience 5.

Enfin, nous serons à même de proposer des séquences RSVP prototypiques, donnant lieu à des conjonctions illusoire de manière stable. Cet objectif clora le chapitre 5. Les séquences d'images RSVP ainsi extraites constitueront un instrument de mesure pour conduire des études en imagerie et en électrophysiologie.

Ce travail exploratoire a pour but de préparer des études ultérieures, qui chercheront à montrer que lorsqu'il se produit une intégration cible-distracteur sur un critère perceptif donné en RSVP, les groupes de neurones sous-jacents à ce traitement entrent en synchronie, alors qu'en l'absence de conjonction illusoire, l'apparition de synchronies n'est pas observée.

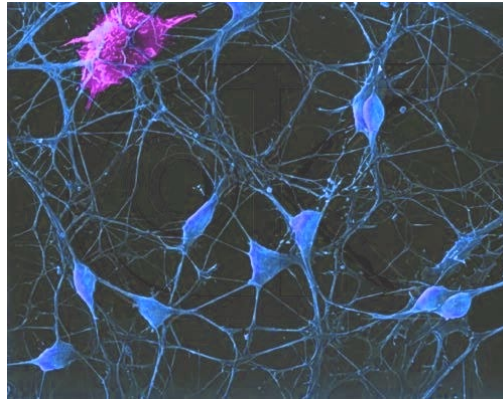
Cadre théorique

Un premier chapitre sera consacré aux paradigmes du fonctionnement cognitif. Dans un premier temps, nous présenterons le paradigme dominant de la neuropsychologie classique. Puis, nous présenterons une sélection de problèmes auquel il se heurte. Enfin, nous détaillerons une des alternatives qui lui est proposé, la théorie des assemblées transitoires synchrones. La polysémie du terme de synchronisation sera soulignée à travers la description des diverses méthodes d'analyse qui l'étudient.

Un second chapitre détaillera la neuro-psycho-physiologie du système visuel. Nous détaillerons la fonction des neurones du système visuel dans les différentes structures de la rétine à la voie ventrale. De cette première partie, il émergera clairement que les informations extraites par les différentes structures sont extrêmement fragmentaires et discontinues, peu en rapport avec l'impression de continuité que nous restitue le système visuel lorsque les traitements sont accomplis. Le versant psychologique permet d'aborder l'organisation perceptive, ou comment le système restaure une continuité spatio-temporelle à partir de ces bribes d'information disparates. Enfin, nous donnerons un modèle du déroulement temporel de ces opérations visuelles, issu de recherches chez le singe.

Un troisième chapitre concernera les bases théoriques et la mise en place de la recherche sur les conjonctions illusoires en RSVP. Des expériences préliminaires seront présentées, au cours desquelles la pertinence du phénomène de conjonction illusoire est apparue comme plausible pour l'étude de l'intégration perceptive.

1. Paradigmes du codage de l'information



Décrypter le fonctionnement du cerveau ne consiste pas en un simple problème de traduction entre deux langages, le langage de la cognition et le langage de la neurophysiologie. Les **relations entre les propriétés de l'activité neuronale et le codage de l'information** doivent être entièrement spéculées, sur base de quelques indices fragiles, et de préférence interdisciplinaires. Ces indices, parfois imposteurs, essayent de nous convaincre qu'ils dépendent plus de l'observé que du dispositif d'observation, et qu'ils sont les auteurs des opérations entreprises, et non qu'ils sont asservis par d'autres mécanismes. Parmi les relations neuro-cognitives stables observées, certaines dévoilent en propre le codage de l'information. Beaucoup d'autres se révéleront être des artefacts ; soit ! elles indiquent de manière consistante les propriétés de l'information en transit, mais elles ne sont que de simples corollaires de la mise en œuvre des mécanismes-sources, eux-seuls supportant en propre le codage de l'information (Cariani, 2001).

Ces spéculations sur le code de l'information dans le cerveau accaparent une multitude de « hackers » depuis plusieurs décennies. Pour comprendre le codage de l'information, ils interrogent les modifications bio-électriques à différents niveaux. Les mesures des états et des changements neuronaux sont réalisées sur des neurones isolés, sur des groupes de neurones ou sur le scalp. Les **méthodes** qui divulguent des indices du fonctionnement cérébral in vivo sont très variées, mais celles qui relèvent directement du codage neuronal sont plus rares ; citons parmi elles, les enregistrements unitaires ou multi-unitaires chez l'animal, les enregistrements locaux chez le sujet humain lors d'interventions en neurochirurgie et les enregistrements de surface, ainsi que toutes les analyses qui dérivent de ces signaux (potentiels évoqués, analyses chaotiques, analyses temps-fréquence...).

Un **premier modèle** du traitement de l'information a été construit durant plus d'un siècle. Basé sur les propriétés individuelles des neurones, ce modèle s'est récemment révélé incapable d'expliquer certaines observations empiriques. Aussi, des propositions **alternatives** ont été formulées. Nous présenterons les grandes lignes et les limites de ce premier modèle. Nous présenterons ensuite une sélection des problèmes qu'il n'a pas su résoudre. Puis nous présenterons une des alternatives proposées, celle du codage par assemblées transitoires synchrones. Cette théorie intégrative motive le projet de recherche dans lequel s'inscrit la présente thèse.

Le paradigme dominant de la neuropsychologie classique

La neuropsychologie a consacré le **modèle boxologique** (en boîtes et flèches), ou modèle en boîtes et en flèches. Les postulats de la théorie apparaissent dans ce modèle de représentation (voir figure 1). La Neuropsychologie étudie les processus et les fonctions. Nous voyons une succession de processus, quelques parallélismes locaux, terminés par des convergences ; tout ceci fonde la **sérialité**. Des associations anatomo-fonctionnelles prédites par d'autres études ; et voici le **localisationnisme**. Une complexification des processus à mesure qu'on avance dans les traitements, et en même temps, un déplacement vers les aires cérébrales plus « centrales », et voilà la **hiérarchie** qui apparaît.

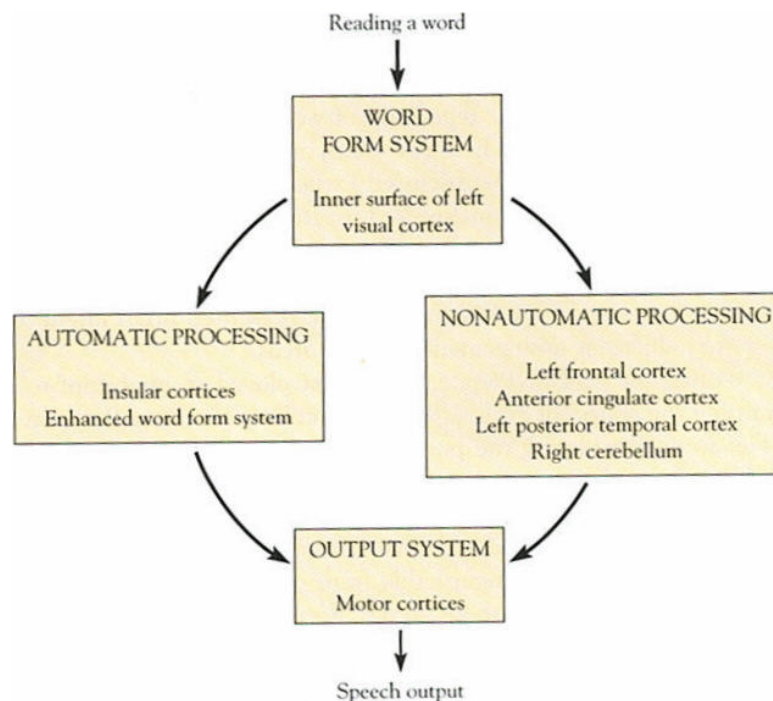


Figure 1 : un modèle boxologique de la neuropsychologie classique (reproduit de Posner & Raichle, 1994)

Nous allons reprendre les différentes propriétés de la théorie, en essayant de dégager les implications des **postulats** qui la sous-tendent et des **simplifications** qu'elle tolère.

Le code de l'information : où est le problème ?

Les modèles du traitement de l'information ont recherché à rendre sérielles les différentes fonctions étudiées (Morin, 1995). A ce titre, les exemples historiques du béhaviorisme (modèles S-R de Watson ; puis S-O-R de Skinner), et de la cybernétique (modèle du canal dans la théorie de la communication de Shannon & Weaver) montrent cette volonté de simplification. Ce choix épistémologique orienté vers des modèles simplifiés a influencé les théories cognitives tout au long du XX^e siècle. On peut considérer, par exemple que les modèles de l'attention (eg. Broadbent, 1982, et le « goulot d'étranglement ») ont été inspirés par les développements théoriques sériels.

Le paradigme de la Neuropsychologie classique inclut cette **sérialité**. Des parallélismes locaux existent, au bout desquels on pose toujours une aire associative, un neurone de décision ou un quelconque mécanisme de convergence (voir Edelman, 2002 ; Driver & Spence, 2000). Ceci est sans doute lié à l'histoire de la discipline : on a posé les fondations de ce paradigme en suivant l'adressage des fibres nerveuses dans le cerveau, c'est-à-dire en suivant des « autoroutes » de traitement de l'information. Que la présence de ces autoroutes soit un fait général, corollaire, ou anecdotique, au regard du **traitement de l'information**¹ n'est pas une question ouverte dans le paradigme.

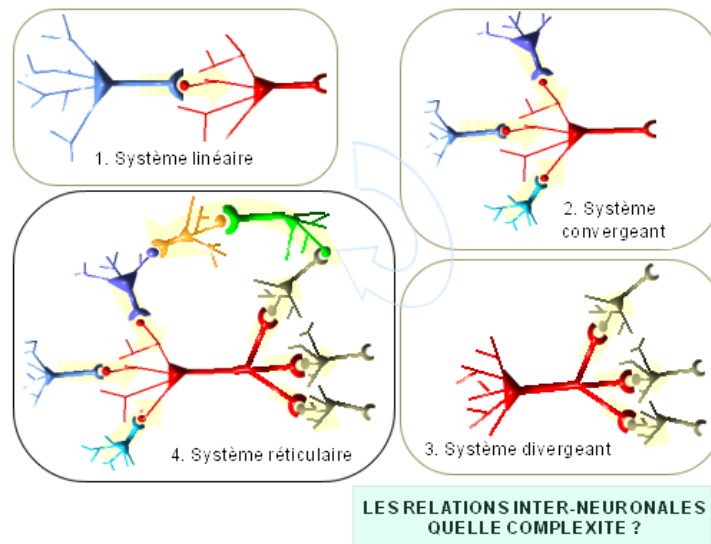


Figure 2 : Prise en compte de la complexité des connexions neuronales.

En conséquence de cette conception sérielle, les théories du traitement de l'information transforment souvent la **transmission** du signal en un modèle **linéaire**. Il nous est familier de considérer comment l'activité d'un neurone agit sur le neurone suivant (voir figure 2 partie 1). C'est une perspective sérielle, très commune. Par exemple, on admet parfaitement l'idée de résumer le traitement visuel par la description des computations réalisées par un ou quelques neurones dans chaque aire visuelle. Par exemple, pour le traitement des contours, les modèles nous expliquent comment on passe de récepteurs rétiniens sensibles à des points, à des neurones sensibles à des segments orientés dans V1 (voir figure 3), puis à des neurones sélectifs à des formes complexes dans IT. Les pas de traitement d'aire en aire sont d'ailleurs souvent mieux connus que les interactions latérales (eg. voir à ce titre l'absence du paragraphe consacré aux connexions latérales dans la revue de Rockland, 2002, qui annonce pourtant les trois types de connexions en introduction, voir aussi les enjeux prospectifs annoncés dans Lamme, Super & Spekreijse, 1998).

¹ Dans un paradigme alternatif, on pourrait par exemple imaginer que les « voies » neuronales ne servent qu'à la mémorisation des informations (ie. à la potentialisation des synapses), un autre mécanisme supportant le traitement cognitif de l'information sur un mode d'action plus rapide.

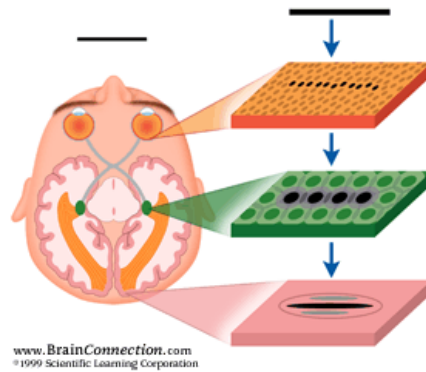


Figure 3 : conceptions « linéarisées » du fonctionnement neuronal : l'exemple des contours visuels (reproduit de <http://www.brainconnection.com>). Les récepteurs rétiniens signalent des points, qui sont ensuite réorganisés en contrastes concentriques noir-blanc dans les cellules ganglionnaires et dans les Corps Genouillés Latéraux, puis en segments orientés dans V1.

Connaissant les principes de la sommation spatiale, nous comprenons comment opèrent plusieurs neurones pour déclencher une réponse du neurone suivant (voir figure 2 partie 2). Cette multi-activité en amont étant rarement différenciée, on pourrait la considérer comme unitaire ; on ne serait pas trop dérangé les multiples sommations des différents afférents d'un neurone par une activité provenant d'un super-neurone ayant davantage de pouvoir d'excitation ou d'inhibition, puisque le problème est d'atteindre un **seuil**. Là encore, il s'agit d'une linéarisation.

A contrario, il est plus rare de considérer la **dispersion** de l'information sortante (voir figure 2 partie 3), et il ne s'est développé aucun outil d'observation spécifique à la **nature réticulaire des réseaux de neurones** (voir figure 2 partie 4), comportant de multiples convergences en entrée de chaque neurone, et de multiples divergences en sortie (eg. Arbib, Erdi & Szentagothai 1997). C'est sans doute à cause de ce vide théorique que la modélisation connexionniste prend une importance grandissante. Elle occupe le terrain laissé vacant par l'absence de dispositifs d'observations « in vivo ».

Cette simplification du problème a conduit à une conception unidimensionnelle du codage de l'information. Usuellement, on représente schématiquement l'activité du neurone par une indexation temporelle de ses potentiels d'action (que l'on nomme raster). On observe sur ce diagramme de l'activité neuronale (figure 4), des modifications de la densité des décharges au cours du temps. On a donc construit les postulats du traitement de l'information autour de la notion de **fréquences de décharge**. L'augmentation de la fréquence des décharges dépendante du stimulus trouve son origine dans les travaux de Adrian, récompensé d'un prix Nobel en 1932, (voir Martin, 2000).



Figure 4 : Diagramme de l'activité d'un neurone au cours du temps : variations de la fréquence de décharge

Tenant à un niveau, une structure en voies et modules ; 'j'achemine - je transforme - j'achemine au suivant...' et à l'autre, un mode de traitement et de convoyage simple et élégant : la **modification de la fréquence de décharge** et l'**intégration synaptique**, les chercheurs du XX^e siècle ont élaboré une conception univoque, mais peut-être partielle, du

codage de l'information. Cette perspective était très heuristique. Dès que l'on a introduit une électrode dans des neurones, et que l'on s'est penché sur les modulations de l'activité de ce neurone liées au stimulus, une pléthore de résultats sont tombés (eg. en vision : les champs récepteurs, Hartline, 1938 ; Kuffler, 1953 ; la sélectivité à l'orientation Hubel & Wiesel, 1959 ; les cellules sensibles à des oppositions de couleurs, DeValois et DeValois, 1993).

Les localisations : « où ? » est le problème !



figure 5 : un buste de Phrénologie, l'ancêtre de la Neuropsychologie : la surface du crâne est découpée en secteurs, auxquels sont attribuées des qualités psychologiques, telles que la sympathie, les aptitudes domestiques ou l'idéalisme.

L'évaluation de ces aptitudes se réalise en examinant les bosses et les creux répercutés à la surface du crâne. Ces reliefs osseux permettent d'apprécier chaque qualité de manière quantitative.

La question du « comment ? » du codage de l'information était entendue. On s'est alors intéressé à la question du « où ? » et du « quoi ? », c'est-à-dire sur la question des **corrélats anatomo-fonctionnels** (voir Farah, 1994 pour discussion). La théorie de la Neuropsychologie classique a engendré une profusion d'hypothèses cherchant à élucider les aires cérébrales (des entités **localisées**¹ ayant une cohérence fonctionnelle), et cherchant à comprendre les computations réalisées par ces groupes de neurones (ladite **fonction**, isolée anatomiquement par un raisonnement circulaire). Des exemples de recherches sur la fonction des neurones et leur localisation apparaissent par exemple pp. 54-62.

Ces recherches ont rempli tout l'espace d'expression scientifique, et ont tue toute proposition discordante sur ses soubassements. Les recherches anatomo-fonctionnelles, par le travail considérable qu'elles représentaient, et par les postulats qu'elles entraînaient, ont sédimenté le statu quo concernant les codes neuronaux.

On est maintenant arrivé à une grande finesse de détail dans ces recherches en localisation, en particulier grâce à la survenue de l'imagerie cérébrale fonctionnelle (Uttal, 2001). Cependant, tout cet édifice présente des risques de réification. Les inférences localisationnistes reposant les unes sur les autres, la propagation des erreurs est problématique. Cependant, la Neuropsychologie a pris soin de s'assurer de la **convergence des résultats** par les nombreuses méthodes. Les lésions et dysfonctionnements cérébraux constitutionnels, acquis, ou transitoires (TMS, *transcranial magnetic stimulation*²) de l'homme et de l'animal sont interrogées, ainsi que le fonctionnement normal, avec de multiples méthodes (tests psychologiques - imageries - électrophysiologie).

¹ Les neurones sont en effet regroupés par unité fonctionnelle. Deux neurones ayant quelque chose à voir l'un avec l'autre sont en général topographiquement proches dans le cerveau (Ringo, 1991)

² La stimulation magnétique transcranienne consiste à appliquer un champ magnétique localisé sur une partie localisée du cerveau. Elle a pour conséquence l'interruption momentanée du fonctionnement de cette zone, et peut en particulier être utilisée pour des investigations neuropsychologiques chez le sujet normal. La stimulation magnétique mime alors, de manière transitoire les lésions acquises du patient neuropsychologique.

Notons en revanche que les recherches en localisation n'ont de sens que si le regroupement fonctionnel des neurones est effectif, c'est-à-dire si les modules¹ existent. Cela n'aurait pas de sens de localiser des fonctions entièrement dispersées. Or, pour nuancer quelque peu ce principe, on remarque que les études en enregistrements unitaires montrent une pluralité fonctionnelle des neurones regroupés au sein d'une même aire cérébrale. Par exemple, les neurones des cortex visuels striés et péristriés se distribuent des rôles de codage de l'orientation, de la couleur, de la disparité rétinienne (Desimone, Schein, Moran & Ungerleider, 1985). On note également que les neurones des différentes couches architectoniques semblent impliqués dans différentes fonctions, et encore que les neurones d'une même portion de cortex, ayant des récepteurs différents (GABAergiques, glutamatergiques, dopaminergiques...) sont impliqués de manière différente dans les traitements (voir Hardcastle, 1999). Plutôt que de modules ayant une fonction unique, ces études suggèrent davantage l'existence de **clusters** dans les aires cérébrales (Tononi, Edelman & Sporns, 1998) liés fonctionnellement de manière distribuée.

Hiérarchie du traitement

Par la définition des modules cérébraux, s'est développée l'idée d'un sens dans le traitement de l'information. Cette notion de **sens**² ne peut exister que parce que ces modules sont définis de manière linéaire. En outre, l'ordre des modules ne peut être défini de manière univoque. Par exemple, la définition classique de la hiérarchie des aires visuelles, obtenue en suivant l'adressage des fibres nerveuses (eg. Felleman & Van Essen, 1991), est sensiblement différente du tableau obtenu par l'analyse des latences de décharges des différentes structures (voir pp.71-72).

A cette notion directionnelle s'additionne un second concept, celui de la **complexification** des traitements successifs (Engel, Fries & Singer, 2001). Si on observe les comportements des neurones dans les modules successifs, on observe en effet une gradation dans la complexité et l'abstraction des traitements réalisés.

De ces deux concepts dérive la notion de **hiérarchie** (voir discussion sur la vision dans Bullier & Nowak, 1995). Cette notion n'est pas triviale ; ses conséquences sur les théories cognitives, elles non plus, ne sont pas triviales. Prenons l'exemple des traitements sensoriels (traitements qui y sont pour beaucoup dans cette notion hiérarchique). Sachant que dans les aires « périphériques », l'information est encore peu transformée (Mesulam, 1998), elle peut être facilement comprise à l'aide des méthodes statistiques traditionnelles corrélant les caractéristiques du stimulus et l'activité neuronale. Ceci ne nous semble pas être vrai pour les aires « plus centrales », lesquelles paraissent échapper à cette statistique corrélationnelle³. L'interprétation de leur fonction consiste essentiellement en des spéculations

¹ Cette conception modulaire s'appuie en particulier sur l'ouvrage influent de Fodor (1983) : The modularity of mind.

² « *The brain [...] perniciously refuses to function in a linear, causal fashion. The hundreds of millions of feedback loops it employs made a travesty of our efforts to establish Newtonian causal sequences.* » Newman, Journal of Consciousness Studies.

³ Prenant l'exemple du système visuel, la découverte des fonctions des neurones de ce système suit une logique « ascendante ». Les fonctions des neurones de V1 ont été plus rapidement élucidés que les fonctions des neurones de V4, de MT ou de IT. Et globalement, la description fine des clusters de neurones et de leurs fonctions propres est bien meilleure dans V1 que dans les aires situées plus haut dans la hiérarchie. Dans V1, au moins 5 formes de sélectivité (position rétinotopique, couleur, orientation, dominance oculaire, fréquence spatiale) ont été mises en évidence, tandis que dans les aires situées en aval, la description se borne parfois à la réponse « moyenne » des neurones les plus nombreux.

théoriques. On a tout naturellement pensé que les aires de traitement plus centrales, manipulant des informations plus complexes et abstraites, et communiquant avec davantage de structures, avaient un rôle de **commandement** sur les aires situées en périphérie. Cette conception entrave les formulations théoriques du cerveau comme système auto-organisé. Elle récusé toute possibilité de traitements simultanés à grande échelle.

Cet hiérarchie des traitements neuronaux a donné lieu à une distinction entre trois formes d'influences neuronales. On parle de modulation *feed-forward*, *feed-back*, et *latérale* (voir figure 6), selon que le neurone exerce une influence sur des cellules situées respectivement plus haut, plus bas, ou au même niveau que lui dans la hiérarchie.

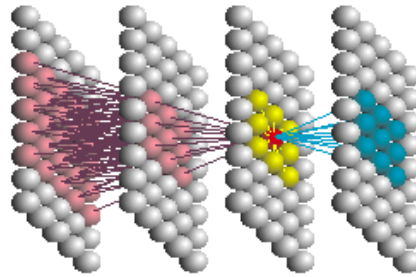


Figure 6 : modulations *feed-forward* (en bleu), *feed-back* (en rose), et *latérales* (en jaune) d'un neurone (en rouge). Reproduit de (Lamme & Roelfsema, 2000).

Cette conception hiérarchique et sérielle suggère que le résultat des **computations neuronales antérieures** est **intégré à la trame de l'activité neuronale ultérieure**. Il en découle une **conception patrimoniale** du traitement de l'information. Le neurone d'une étape donnée reçoit une information comportant, avec plus ou moins de détail (le dépouillement des détails fonde l'abstraction) le produit de tous les traitements antérieurs. A son tour il transmet le contenu informatif de toutes les étapes précédentes, en même temps que le résultat de sa propre computation.

En bout de course, est admise tacitement l'existence de neurones, ou de groupes de **neurones** (aires) de **décision**. On trouve des développements théoriques de ce postulat sous différentes formes (eg. les cellules pontificales de Barlow). Dans sa forme la plus radicale (eg. Barlow, 1972), cette vue est difficilement tenable. En effet, un tel système serait très fragile à la perte neuronale. Si quelques neurones ont un rôle décisif, comment expliquer la pérennité des fonctions cognitives, alors que la perte neuronale n'est pas sélective, et que rien n'empêche la destruction des cellules ayant davantage d'importance dans le fonctionnement mental. Notons cependant que les atteintes des aires considérées comme plus haut dans la hiérarchie, lorsque lésées, se traduisent par des atteintes plus globales (eg. atteintes frontales).

Ecueils



-sans légende-

Les contraintes temporelles en vision

En 1996, un problème théorique est formulé en contradiction avec les postulats de la théorie neuropsychologique classique, par Simon Thorpe et son équipe (Thorpe, Fize et Marlot, 1996). Par une simulation mathématique des temps de traitements disponibles à chaque étape neuronale, les auteurs établissent **l'impossibilité du codage par fréquence de potentiels d'action dans certaines conditions du traitement visuel**.

Dans un premier temps, les auteurs établissent le **temps minimal pour catégoriser un objet**, grâce à un paradigme URVC (Catégorisation Visuelle Ultra-Rapide). Cette latence est obtenue à l'aide de méthodes comportementale (Temps de Réaction, méthode soustractive et calcul du temps de discrimination) et électrophysiologique (potentiels évoqués, voir figure 7). Il ressort de ces études qu'un objet peut être catégorisé en aussi peu de temps que **150 millisecondes**.

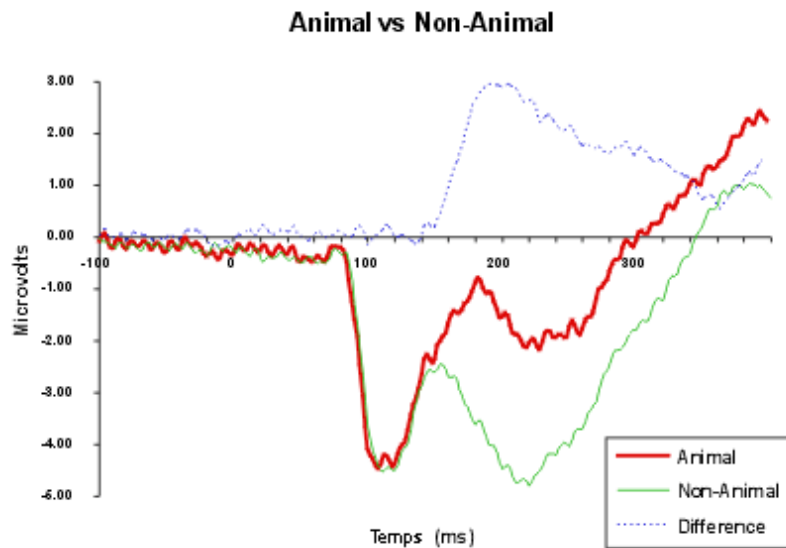


Figure 7 : Potentiels évoqués lors de catégorisations visuelles ultra-rapides, par Thorpe, Fize et Marlot (1996). L'étude démontre qu'une catégorisation peut être réalisée en 150 millisecondes. Les stimuli étaient des photos d'animal (catégorie-cible) ou de distracteurs appartenant à d'autres catégories sémantiques (non-animal : catégorie-distracteur). Ces images étaient présentées brièvement (20 millisecondes). Le sujet devait réaliser une tâche de go no-go¹. Les tracés des potentiels évoqués des catégories cible et distracteur divergent dès 150 millisecondes, suggérant que la catégorisation a été établie à cette latence.

¹ Le go no-go est une situation expérimentale au cours de laquelle le sujet (humain ou animal) répond, ou s'abstient de répondre selon la stimulation qui lui est présentée. Dans l'expérience de Thorpe, Fize & Marlot, (1996), le sujet doit presser sur un bouton lorsqu'il perçoit un animal et doit s'abstenir de presser le bouton lorsque la cible n'est pas un animal.

Dans un second temps, cette latence minimale de catégorisation est utilisée pour calculer le **décalé moyen disponible à chaque étape neuronale**. Compte tenu de la latence d'intégration des récepteurs rétiniens (20 à 30 ms), il reste au signal 120 à 130 millisecondes pour parcourir la dizaine de relais synaptiques séparant ces récepteurs et IT. Chaque étape bénéficie donc d'une disponibilité de 12 ou 13 millisecondes. De cette latence, doivent être soustraits les temps occupés par la conduction axonale (une milliseconde par millimètre). Si on considère que le neurone décharge au maximum à 100 ou 200 Hz (une impulsion toutes les 5 à 10 millisecondes), alors, **seulement un ou deux potentiels d'actions** « informatifs » pourrait être émis à chaque étape de traitement.

Thorpe et coll. (1996) établissent l'impossibilité d'un traitement de l'information visuel qui serait entièrement sériel, et dans lequel le transfert de l'information serait réalisé par relais synaptique et train de potentiels d'action.

Si l'on admet la sérialité des traitements visuels, et si on considère les calculs proposés par Thorpe et collaborateurs, on s'attend à ce que l'information visuelle ne soit codée que par une ou deux décharges à chaque pas de traitement. Or, qui dit « fréquence de décharge » dit que plusieurs décharges doivent être émises par le neurone pour que la modification de fréquence soit décodée par le système.

Suite à la mise en évidence de ce problème théorique, Thorpe a formulé une hypothèse alternative substituant au codage par fréquence de décharge, un codage **asynchrone** (Gautrais & Thorpe, 1998). Cette proposition est basée sur l'idée que l'information transmise à un neurone donné est encodée par le rang de décharge des différents neurones afférents (voir figure 8). Le codage asynchrone est une des alternatives au codage par fréquence. Il fait également partie des codages dits « temporels », à l'instar de l'hypothèse des assemblées neuronales synchrones qui est l'objet de la présente thèse.

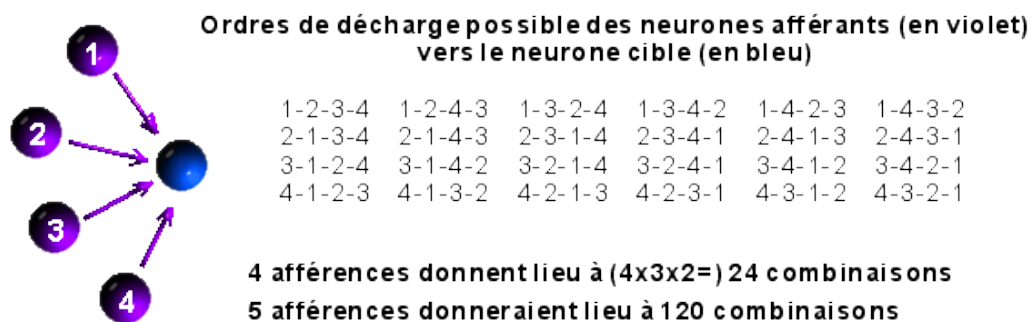


Figure 8 : principe du codage asynchrone : le code reçu par le neurone est contenu dans l'ordre de décharge des afférences synaptiques. Par exemple, le neurone bleu reçoit une information donnée lorsqu'il est successivement activé par les neurones violets 1, puis 2, puis 3, puis 4. Il reçoit une information différente si il est activé successivement par les neurones 1, puis 3, puis 2, puis 4, et ce, même si les autres paramètres d'activation sont similaires. Ce type de codage a été mis en évidence dans MT (aire visuelle sensible au mouvement).

Le « *binding problem* » et ses implications neurophysiologiques

A la fin des années 1970, la neurophysiologie s'est abondamment penchée sur les fonctions unitaires des neurones, mais le problème de l'intégration des computations locales n'est pas posé. C'est Anne Treisman qui soulève ce « *binding problem* » en 1980 (Treisman & Gelade, 1980, Treisman & Schmitt 1982 ; synthèses dans Treisman 1996, Treisman, 1998). Une illustration en est donnée à la figure 9.

Considérant le parallélisme massif de la perception visuelle, comment le cerveau retrouve-t-il, parmi tous les neurones simultanément actifs à un moment donné, et dans différents endroits du cerveau, lesquels codaient pour un objet déterminé ?

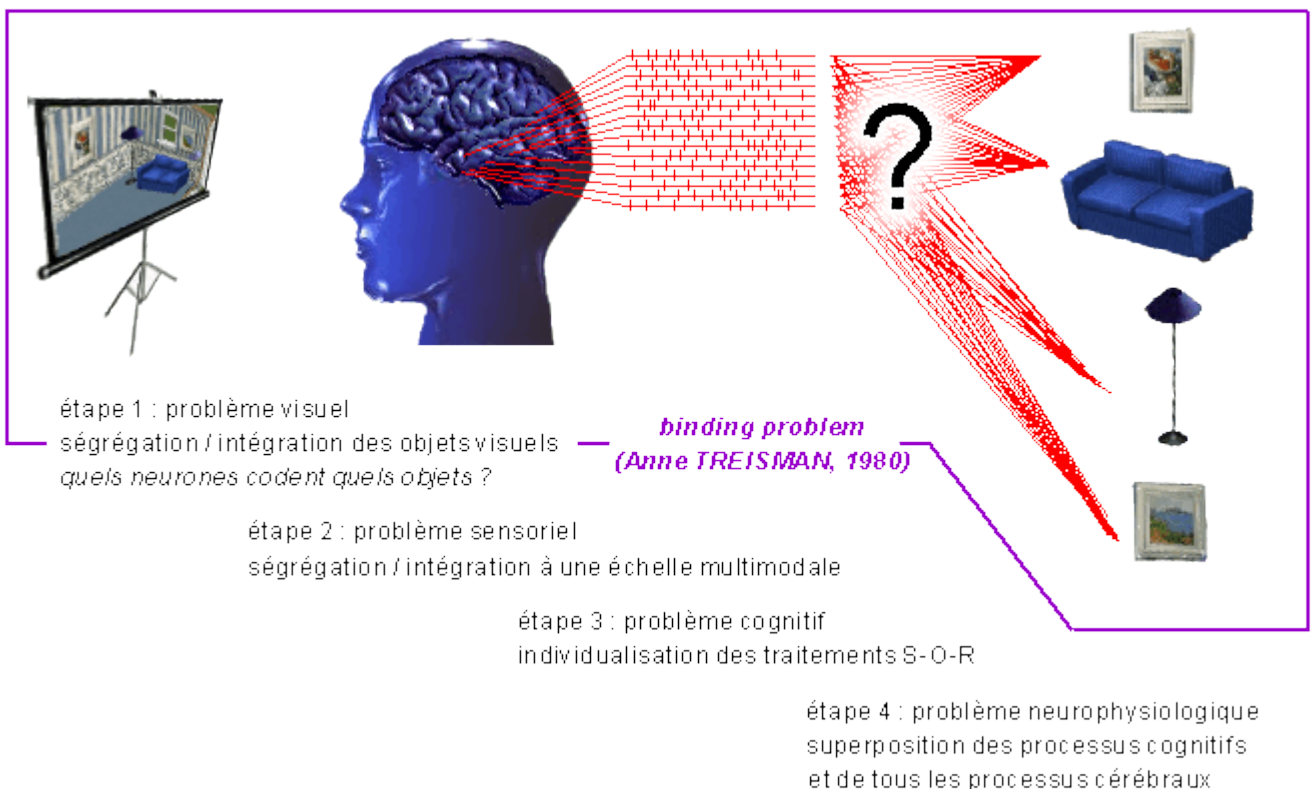


figure 9 : le *binding problem* posé par Anne Treisman, en 1980 : Comment les activités neuronales sont efficacement réassociées à la représentation du bon objet ? Au delà de la perception visuelle, cette question est un problème générique du fonctionnement cérébral.

A la lumière des théories des Neurosciences actuelles, la portée du questionnement n'est pas minime. Au delà de la perception, c'est tout le fonctionnement cérébral qui est interpellé. Cependant, à l'époque, la théorie de la Neuropsychologie classique est incontestée. Bien que les phénomènes mentionnés par Anne Treisman dans son argumentaire (les conjonctions illusoire, assumées comme faillites attentionnelles) sont intrigants, le problème qu'ils posent reste cantonné dans le champ théorique de l'attention sélective. Dans un paradigme dans lequel les postulats de **sérialité** et de **hiérarchie** sont incontestées (voir pp.13-18), l'intégration trans-régionale est sans doute un problème marginal.

Ce contexte scientifique explique sans doute pourquoi cette question théorique n'a déployé tout son impact qu'à partir du début des années 1990. A cette époque, le « *binding problem* » a été exploité par les défenseurs des théories alternatives du traitement de l'information (eg. théorie des assemblées neuronales synchronisées, voir ci-après), et les questions de fond qu'il posait ont été réactualisées.

Fondamentalement, le *binding problem* soulève les deux contraintes interactives de l'intégration et de la ségrégation des objets co-présents dans une scène visuelle. Ces deux concepts sont les deux faces du même problème.

L'intégration en perception visuelle pose le problème suivant. Les neurones visuels codent des propriétés de forme, de taille, de couleur, de texture... Quelque-part dans le cerveau, ces caractères disparates doivent être **unifiés en des entités individualisées**, les objets. A priori, le cerveau ne sait pas que tel fragment de contour (voir illustration figure 10), ou que telle couleur, appartient à un objet plutôt qu'à un autre. Il ne peut pas le savoir sur la base de la position spatiale, car les objets des scènes naturelles se superposent la plupart du temps dans des plans plus ou moins éloignés. Il faudra cependant bien que le système cognitif apprenne à intégrer les composantes de bas-niveau, s'il veut réaliser l'individualisation¹ dudit objet. Il devra intégrer les différents fragments d'une même propriété visuelle (par exemple, les contours du sofa), et il devra combiner toutes les propriétés (contour, couleur...) entre-elles.

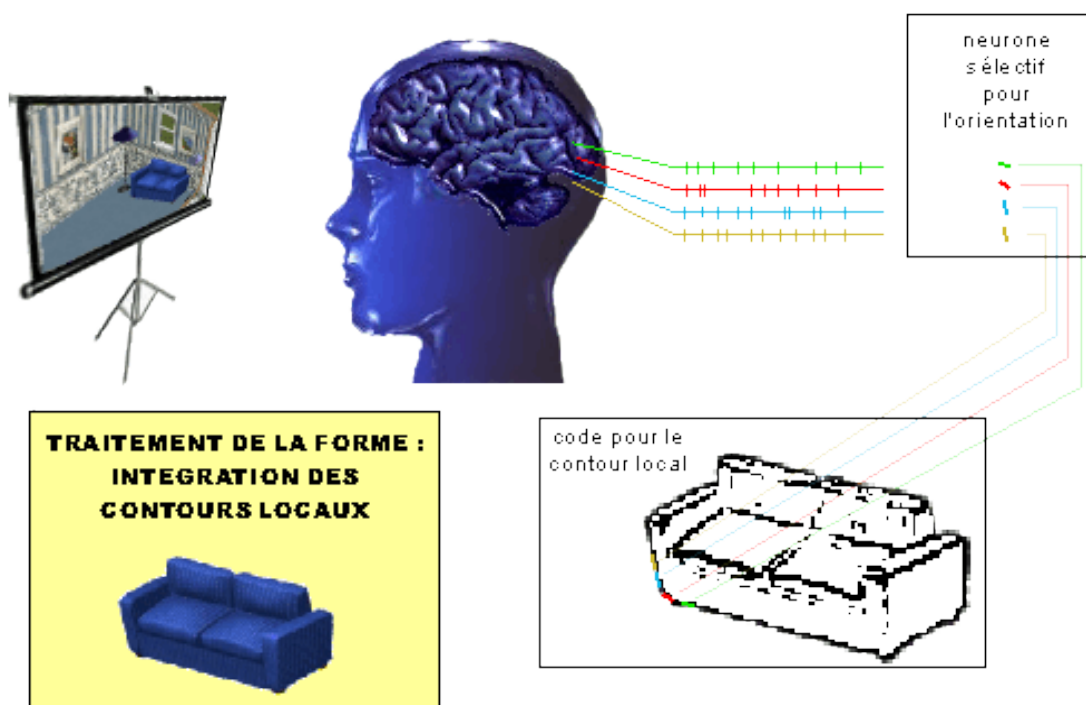


figure 10 : le *binding problem* et l'intégration des contours locaux : les neurones de l'aire V1 codent les fragments de contour. Comment le cerveau sait-il que tels neurones codent pour le même objet ? Par quel moyen assemble-t-il de manière appropriée les contours communs ? Un mécanisme de liage est requis pour cimenter les traitements appropriés.

Le problème de la **ségrégation des objets visuels** va de paire avec le problème de l'intégration. Il se réfère au fait que dans les scènes naturelles, de nombreux objets visuels sont co-présents. Le cerveau doit également affronter le

¹ A notre connaissance, aucune théorie actuelle n'envisage l'identification d'un objet sans une individualisation préalable de cet objet.

problème suivant : différents traitements sont réalisés par différentes aires, concernant les objets respectivement présents dans le champ visuel. Par exemple (figure 11) dans la même aire visuelle, sont codées la position du tableau et la position du sofa. Comment le cerveau prélève-t-il les bonnes computations pour constituer la représentation de chaque objet ? N'arrive-t-il pas que le cerveau associe une caractéristique au mauvais objet ? Dans l'exemple ci-dessous (figure 11), se peut-il que soient associée au tableau la position bas-droite, alors que cet objet est positionné en haut à gauche ? Dans certaines conditions, on observe effectivement de telles erreurs de « recombinaison », les **conjonctions illusoires**.

Ce double-problème d'intégration-ségrégation constitue une difficulté de taille pour la théorie de la neuropsychologie classique. L'**attention sélective**, qui postule le traitement d'un objet à la fois¹, propose la question de l'origine de l'œuf à la place de la question de l'origine de la poule. Certes, un seul objet est perçu à la fois, et le traitement de cet objet est ségrégué dans une fenêtre temporelle propre qui le spécifie, mais qui a indiqué le contenu utile pour la définition a priori de l'objet ?

L'introduction dans les théories cognitives des **mécanismes top-down** pourrait également constituer une solution, mais pour le moment, le contenu et les propriétés de ces processus n'est pas spécifié. Cette béquille théorique s'apparente pour le moment à une nouvelle boîte noire.

¹ Les théories de l'attention sélective assigne celle-ci au rôle de goulot d'étranglement de la cognition (eg. Broadbent, 1982). Un seul objet est traité à la fois.

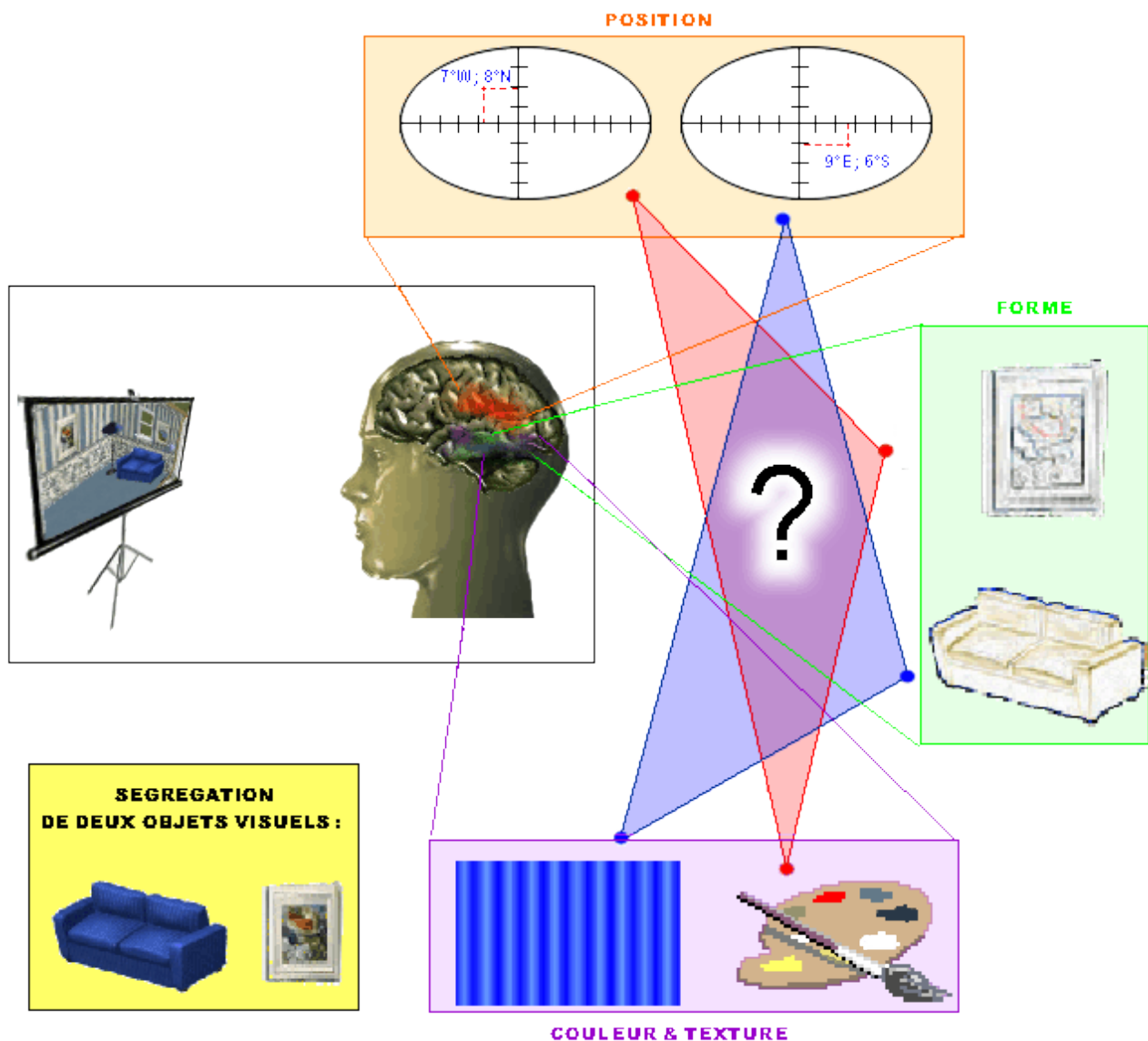


figure 11 : le binding problem et la ségrégation des différents objets visuels : les neurones sont regroupés par fonction, et non par objet-cible. Comment le cerveau « reconnaît-il » les neurones associés à chaque objet ? Quels neurones de la voie dorsale codaient pour la position du sofa, et quels neurones codaient pour la position du tableau ?

Ce problème n'est pas résolu par la neuropsychologie classique : si le traitement est entièrement sériel et hiérarchique, et si un seul objet est traité à la fois (attention sélective), les contraintes temporelles sont appesanties de manière peu crédible. Il est donc peu probable que les critères de ségrégation soient des critères temporels. La rétinitopie n'y remédie pas davantage : certains objets se superposent dans le champ visuel.

L'étiquetage des traitements est résolu dans le binding par synchronie (voir chapitre suivant) : chaque objet est « tagué » par la synchronie des décharges de tous les neurones qui le codent.

Note : dans la théorie originale de Treisman (eg. 1986), la position spatiale n'est pas un simple paramètre de la recherche visuelle, au contraire de ce qui est présenté ci-dessus. Elle est le vecteur du faisceau attentionnel, qui se déplace sérialement dans l'espace rétinotopique. Aussi la représentation de la position spatiale comme un paramètre visuel commun, au côté de la forme, de la texture et de la couleur, se situe au delà de la position théorique de l'auteur. Cependant, plusieurs observations rendent crédible le fait que la position spatiale de l'objet soit également une dimension visuelle passible de conjonction illusoire. Par exemple, dans MT, plusieurs codages de la position (par rapport aux yeux, aux mains, à l'objet lui-même...) sont observés (Boussaoud, 1998), la position n'est donc pas un paramètre unitaire directement dérivé de la position du faisceau attentionnel. D'autre part, certaines études relèvent une modulation des conjonctions illusoires selon la position spatiale (eg. Sohn, Liederman, Reinitz, 1996).

La réponse des neurones de V1 chez le chat amblyope

En 1994, Roelfsema, König, Engel, Sireteanu, & Singer, étudient, chez le **chat amblyope**, **l'activité neuronale de l'aire V1**. Un animal amblyope possède une acuité visuelle moins grande sur l'un de ses deux yeux. Cette étude se proposait d'expliquer le déficit en soumettant les animaux à des réseaux de fréquence spatiale (eg. voir p. 59). Ces réseaux sont perçus comme des stries lorsque la fréquence spatiale est basse, et comme une surface unie lorsque la fréquence spatiale est plus élevée, et dépasse le pouvoir de résolution de l'œil. Ainsi, chez ces animaux, l'œil amblyope échoue à résoudre des réseaux de fréquences élevées que l'œil normal réussit à résoudre. Les auteurs comparent les réponses des neurones des colonnes dominantes pour l'œil amblyope et celles des neurones dominants pour l'œil normal. Ils observent que l'amplitude de réponse des neurones des colonnes dominantes pour l'œil amblyope est tout à fait normale, ainsi que la fréquence de décharge. **Les paramètres classiques (amplitude, fréquence) échouent à expliquer le déficit perceptif.** Il s'avérera que les neurones des colonnes dominantes pour l'œil amblyope ne parviennent pas à coordonner temporellement leurs réponses (à synchroniser leurs réponses), et que cette difficulté est d'autant plus marquée que la fréquence spatiale utilisée pour le réseau est proche de la fréquence de fusion pour l'œil amblyope.

La théorie des assemblées transitoires synchrones

Avant propos

La théorie des assemblées neuronales synchrones doit beaucoup de son impact aux réflexions théoriques et aux modélisations (réseaux de neurones artificiels) qui ont été réalisées par des connexionnistes (eg. Von Der Malsburg, 1981,1995 ; Yen, Menschik, Finkel, 1998, 1999 ; Finkel, Yen & Menschik, 1998 ; Grossberg & Grunewald, 1997 ; Karbowski & Kopell, 2000 ; Wennecker & Palm, 1997^{a, b} ; Hopfield & Herz, 1995 ; Sougné, 1997 ; Choe & Miikkulainen, 1998). Les constats et les postulats des connexionnistes se sont couramment insérés dans les revues de question et dans les théorisations formulées par les psychophysiologistes. Cette pluridisciplinarité a été positive dans l'évolution de ce courant d'idée, mais l'importance du nombre des recherches sur les réseaux de neurones artificiels, et le lieu de leur publication (eg. Journal of Neurosciences) font courir le risque d'amalgame, et pourraient amener à surestimer les manifestations purement biologiques de la synchronisation. Pour cette raison, et cherchant à mesurer l'étendue et les limites du phénomène dans le vivant, nous occulterons les travaux réalisés dans les réseaux de neurones artificiels.

Fondements théoriques

Dès 1949, Hebb avait suggéré l'importance de l'agrégation temporelle de l'activité des neurones. La proposition était restée longtemps en suspens, sans doute tue par l'hégémonie prise par le cognitivisme symbolique aux dépens du connexionisme. Depuis les années 1980, le liage par synchronie (*binding by synchrony*, *temporal binding*, ou théorie des assemblées neuronales transitoires synchrones, ou théorie des assemblées de cellules résonnantes) est apparu comme une alternative possible à la théorie neuropsychologique classique pour le traitement de l'information dans le cerveau. Cette théorie propose que le traitement de l'information se réalise par la création d'**assemblées neuronales transitoires** (Von Der Malsburg, 1981 ; Eckhorn, Bauer, Jordan, Borsch, Kruse, Munk & Reitboeck, 1988 ; Gray, König, Engel & Singer 1989 ; Varela, 1995 ; Bressler, 2002). La proposition récuse le codage de l'information par les propriétés individuelles des neurones, pour proposer un **codage par populations**. Dès lors, des efforts de recherche seront concentrés sur l'observation de **corrélations temporelles** entre les neurones.

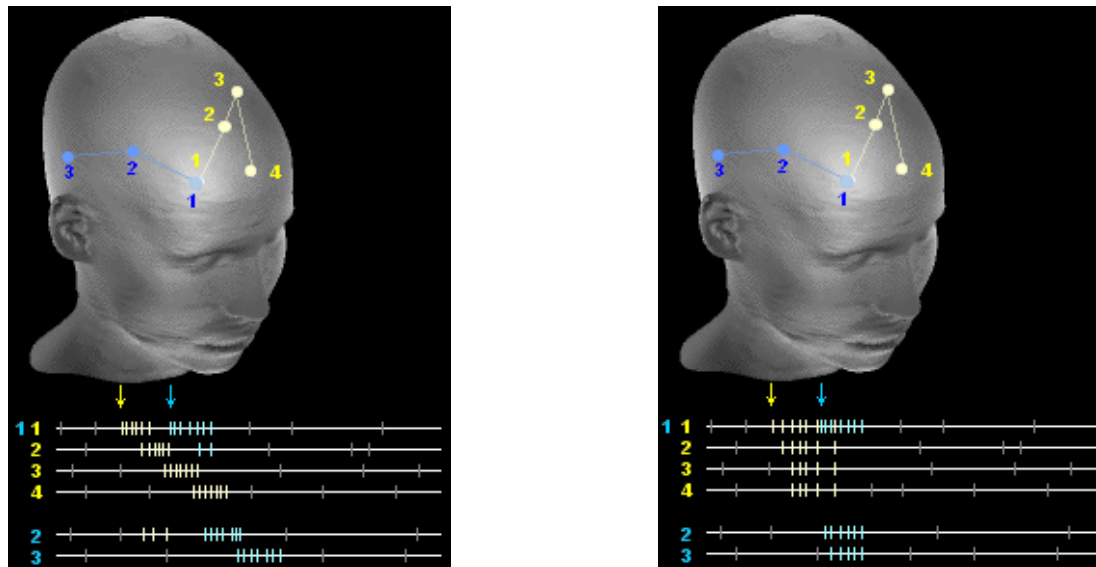


Figure 12 : Principes du codage par fréquence de décharge (à gauche) Vs codage par synchronie (à droite). Deux réseaux hypothétiques (en bleu et en jaune) sont décrits. Ils partagent le neurone numéro 1. L'activité unitaire de chaque neurone est présentée en bas. Les flèches colorées mentionnent respectivement la présentation d'un stimulus engendrant une réponse du réseau jaune et du réseau bleu.

Les diverses théories des assemblées neuronales synchrones ne sont pas toutes formulées à ce niveau neuronal (correspondance terme-à-terme des décharges), mais reposent toutes sur l'idée d'une coordination temporelle des événements bio-électriques.

Les assemblées neuronales se manifestent par la collusion de leur **activité électrique** dans une **synchronisation temporaire** (voir figure 12). Cette théorie satisfait plusieurs contraintes théoriques décisives, jusqu'alors mal expliquées :

- ✧ Flexibilité neuronale : un neurone donné est utilisé par tout un éventail de traitements (*sparse coding*, voir figure 13)
- ✧ Rapidité : le code de l'intégration des activités neuronales est plus économique en temps que celui proposé par la neuropsychologie classique
- ✧ Réversibilité : un neurone peut être rapidement désengagé d'un traitement et réengagé dans un autre traitement. La coactivation, ou engagement simultané d'un neurone dans deux assemblées distinctes, est même théoriquement proposée
- ✧ Parcimonie : peu de neurones sont nécessaires en rapport avec l'immense quantité de traitements possibles
- ✧ Robustesse : La perte de neurones a des conséquences limitées sur les traitements qu'ils réalisaient

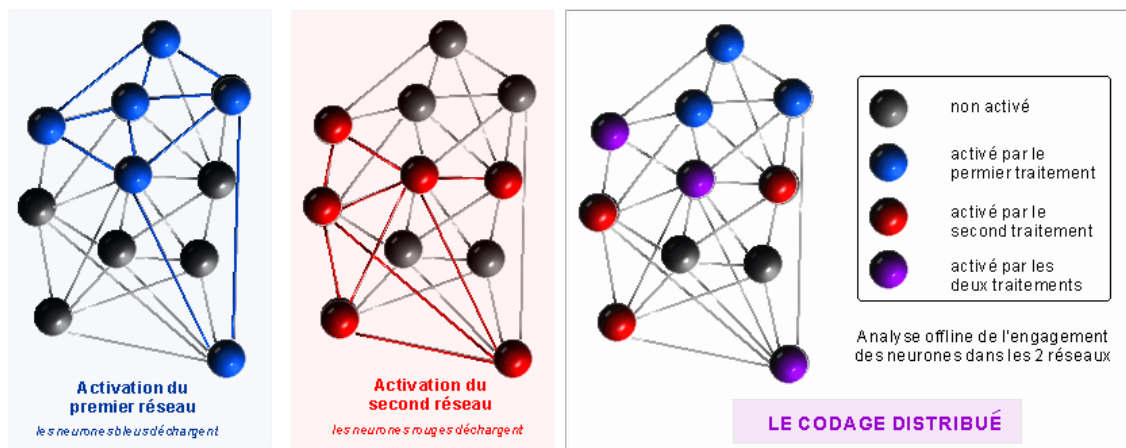


Figure 13 : le codage distribué (sparse coding). Les neurones sont symbolisés par des sphères. L'élaboration d'une représentation implique l'activation de plusieurs neurones. Prenons pour exemple deux représentations, codées respectivement par les réseaux bleu et rouge (à gauche). Sur la droite, se trouve le bilan de l'engagement des neurones dans les deux réseaux. Plusieurs neurones ne sont activés ni par l'une, ni par l'autre représentation, certains neurones ont déchargé uniquement pour l'une des deux représentations, et 3 neurones étaient activés par les deux représentations.

Une comparaison des propriétés du paradigme de la neuropsychologie classique et des assemblées synchrones peut être consultée au tableau 1.

	Neuropsychologie classique	Assemblées synchrones
Code	Modification de la fréquence de décharge des neurones	Synchronisation de l'activité des neurones
Regroupement des neurones	Spatial : théorie localisationniste	Temporel : théorie connexionniste
Effecteurs des traitements	Modules de traitements, innés et appris, reliés par des voies	Assemblées transitoires stabilisées par apprentissage
Synthèse des traitements	Convergence vers des aires d'intégration ou d'association	Emergence cognitive

Table 1 : Comparaison du traitement de l'information postulé dans le paradigme de la neuropsychologie classique et dans le paradigme du binding par synchronie.

La synchronisation : un mécanisme d'intégration pour une hypothèse intégrative

La fonction d'intégration des synchronisations a été formulée théoriquement avant même que lesdites synchronisations ne soient observées (eg. par Hebb, en 1949, Von Der Malsburg, 1981). Autrement dit, le rapprochement entre la théorie des **synchronisations** en neurophysiologie et le **binding problem** posé en Psychologie par Anne Treisman (voir pp 21-24) pré-existait à l'observation in vivo des synchronisations. Cependant, à la survenue des premières études empiriques exhibant des synchronisations dépendantes du stimulus, il était possible de dissocier deux logiques de recherches, selon que le questionnement physiologique prédominait sur le questionnement psychophysiologique, ou vice-versa.

Certaines études mettaient l'accent sur l'**observation des synchronisations**, sans trop s'intéresser aux types de stimuli et aux états mentaux qui les engendraient. Le stimulus était un instrument au service de l'enregistrement des

synchronies¹ ; peu importait sa nature, pourvu qu'il active le neurone. Ces recherches visaient à montrer un fonctionnement des réseaux neuronaux, lié par exemple à la potentialisation des synapses et à l'apprentissage. Elles cherchaient tout particulièrement à savoir **à quelle échelle spatiale** se manifestent les synchronisations et **quels étaient les mécanismes neurophysiologiques** sous-jacents (type de neurones impliqués, neuromédiateurs, structures pacemakers...).

D'autres études posaient comme question centrale la **fonction des synchronisations**. Dans ce cas de figure, on observait quels types de stimuli activaient l'assemblée. Ces recherches convergent vers un **rôle intégrateur de la synchronisation**. Les caractéristiques visuelles ayant quelque chose à voir les unes avec les autres signent leur regroupement par le biais de la synchronie. Le mécanisme semble exister aux niveaux les plus élémentaires (intégration des activités des neurones sélectifs à la direction dans V1, l'ensemble étant supposé coder le contour entier d'une forme), et aux niveaux plus complexes (eg. gestalt, ségrégation figure-fond).

Cette proposition théorique des assemblées neuronales synchrones a été très influente cette dernière décennie, comme en attestent les innombrables publications et la multiplication des colloques dédiés à cette thématique. Cette influence s'explique sans doute en partie par la **nature intégrative de cette hypothèse**. La possibilité d'explications transverses entre les niveaux biologiques et psychologiques a été maintes fois mise en doute ou modérée (eg. LeMoal, 1987). A contre-pied de ces modérations, l'hypothèse reliant le *binding problem* et les assemblées synchrones est une proposition intégrative de ces niveaux psychologique et biologique. Les auteurs connectés à ce courant théorique sont issus de domaines aussi divers que la Psychologie Cognitive, la Neurobiologie Moléculaire et Cellulaire, l'Electrophysiologie, l'Informatique, la Philosophie. Cette thématique forme donc un cadre interdisciplinaire pour les Neurosciences Cognitives.

Notons que **Treisman se défend d'une totale assimilation du psychologique et du biologique** (Treisman², 1998). Après un plaidoyer enthousiaste pour le lien entre « *binding problem* » et « *binding* par synchronie » dans son article de 1996, Treisman évolue vers une segmentation des deux niveaux d'investigation. Elle déclare que les assemblées transitoires synchrones « n'expliquent pas » (ndlr. sur le plan psychologique) le *binding problem*. Le fait neurophysiologique (l'assemblée synchrone) ne peut se substituer à la question psychologique (la notion d'objet), car cela consiste à remplacer une boîte noire par une autre boîte noire. Ayant remplacé le problème psychologique du regroupement des caractéristiques par le constat que des synchronisations se manifestent lorsque des fragments

¹ Le but était initialement d'observer la synchronie, et pour cela, on employait les stimuli préférés des neurones enregistrés (par exemple, pour V1, une barre lumineuse dans le champs récepteur du neurone), à l'instar de ce qui se faisait pour les enregistrements de la fréquence de décharge.

² "It [synchronized neural activity] is an interesting hypothesis, but I don't think it solves the same binding problem that I raised at the beginning of this paper. Synchrony is a possible way of holding on to the solution, of tagging the units that are responding to the same object once they have been identified, but we still need a way of finding which those are." Treisman, 1998, p.1296

d'informations « vont ensemble », nous ne sommes guère plus avancés pour comprendre comment le système cognitif isole et regroupe ces informations communes à l'objet.

On regrette que le concept d'**auto-organisation**, sous-jacent aux recherches sur les assemblées neuronales synchrones, n'apparaisse pas dans les développements théoriques de Treisman. Peut-être des développements significatifs seront-ils fondés sur ce concept dans le futur ? Toujours est-il que, dans l'état actuel de nos connaissances sur la neurophysiologie et la psychologie, ce concept n'est pas opérationnel. Encore loin de pouvoir combler le fossé entre psychologique et biologique, nous sommes réduits à ne pouvoir qu'établir quelques ponts. Les mises en garde de Treisman sont justifiées. Néanmoins, dans le présent chapitre, seront développées des **propositions se situant à l'interface du psychologique et du biologique**. La réduction du fait psychologique à un réseau d'explications biologiques n'est pas notre visée, nous défendons l'hypothèse de deux systèmes de descriptions distincts. En revanche, la légitimation du lien psycho-physiologique permet de **faire collaborer les deux champs de recherches**. Les observations réalisées sur le plan psychologique permettent de **borner les recherches** physiologiques et vice-versa.

Pour illustrer comment la Neurophysiologie peut borner la Psychologie, examinons le concept de « **représentation** », une des grandes boîtes noires de la Psychologie. Ce concept est éclairé par les recherches récentes sur la synchronisation neuronale et la conscience¹. Sur le plan neurophysiologique, la représentation s'entend comme le **produit émergent de l'activité distribuée de neurones** engagés dans une assemblée transitoire synchrone. Cette définition n'est pas une solution définitive pour la psychologie, mais elle est beaucoup plus opérationnelle que celles qui en étaient données auparavant, et permet d'envisager des descriptions de contenu. Par exemple, la représentation d'un objet vu se définirait autour de tous les neurones activés par le traitement visuel, incluant les composantes bottom-up (primitives de contour, de couleur, de position, de mouvement...), les composantes top-down (accès à la mémoire pour l'activation du réseau sémantique lié à l'objet, corrélats émotionnels de l'objet...).

Il était une fois... la synchronisation neuronale

Comme nous l'avons vu, les **problèmes insolubles** posés à la neuropsychologie classique (eg. le *binding problem* de Anne Treisman, les contraintes temporelles en vision posées par Simon Thorpe, voir pp.19-20) ont encouragé l'émergence de **nouvelles formulations théoriques**. Initialement, celle de la synchronisation d'assemblées neuronales transitoires était improbable. En effet, depuis l'invention de l'EEG par Hans Berger en 1929, on avait observé les rythmes cérébraux, et ces derniers avaient été mis en relation avec les états de vigilance. Or le tracé désynchronisé β corrélait avec un état de vigilance et de réflexion, tandis que le tracé synchrone α était la manifestation de la diminution

¹ Le concept de conscience, initialement écarté des recherches des neurosciences cognitives pour cause de mentalisme, subit depuis peu un regain d'intérêt. Il s'agit ici de comprendre quelle information est rendue cognitivement disponible (consciente) par rapport à l'immense masse d'informations partiellement traitées, mais inaccessible à l'esprit. Ces recherches se fédèrent autour de la notion de corrélats neuronaux de la conscience (NCC : *Neural Correlates of Consciousness*)

de la vigilance et de la préparation du sommeil (voir figure 14). Dans un tel contexte, la **synchronie était synonyme de passivité de l'esprit**¹, et la **désynchronisation était synonyme d'activité cognitive**.

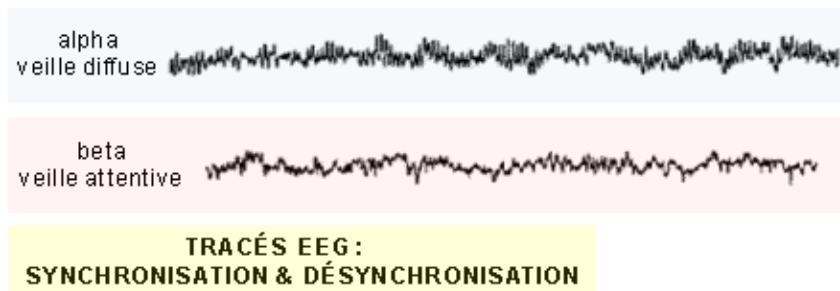


Figure 14 : Les rythmes cérébraux et les états de vigilance. Lorsque le sujet est en veille diffuse, le tracé est synchronisé ; la distance entre deux pics est régulière. Lorsque le sujet est en veille attentive, le tracé devient désynchronisé.

Par ailleurs, on admettait tacitement que les ondes cérébrales répondaient à un ordre fractal² ; on n'admettait pas facilement que ce qui était observé au niveau global du scalp, et ce qui apparaissait aux niveaux plus unitaires, puissent différer. L'idée que le rythme désynchronisé β puisse comporter l'activité de réseaux de neurones synchrones était contre-intuitive.

Ces deux postulats étaient à ce point prégnants que des protocoles ont été élaborés spécifiquement pour **démontrer que l'anesthésiant n'était pas la cause de l'observation des oscillations synchrones** (eg. Gray & Viana Di Prisco, 1997). En effet, les premiers protocoles mis en oeuvre consistaient à administrer des stimuli visuels chez des animaux endormis. Des répliques chez l'animal éveillé ont dues être entreprises. Enfin, Munk, Roelfsema, König, Engel & Singer (1996) montrent que les oscillations synchrones sont favorisées lors de la désynchronisation corticale induite par la formation réticulée activatrice. Cette étude capitale entérine le lien entre les synchronisations et le fonctionnement cognitif.

Vers des évidences empiriques : évolution des méthodes

La formulation de l'hypothèse des synchronisations neuronales est indissociable de l'évolution de certaines méthodes **d'enregistrement et de visualisation des données électrophysiologiques**, évolution qui donna à quelques auteurs la possibilité de réaliser des observations décisives. A l'autre bout du problème, ces nouvelles techniques permettent d'armer la théorie, grâce à l'accomplissement d'observations empiriques.

¹ « L'allure désynchronisée des tracés indiquerait une activité corticale efférente ou associative structurée et intense, tandis que le tracé synchronisé serait le témoin de l'activité simultanée en bouffées rythmiques de grandes populations de neurones, activité vraisemblablement peu compatible avec des efférences ou des associations structurées dans lesquelles chaque neurone serait actif à un moment différent de celui de son voisin en fonction d'un plan d'organisation donné. » (Boisacq-Schepens & Crommelinck., 1996).

² Un système fractal présente des caractéristiques identiques aux différents niveaux d'observation. Cette configuration est un cas particulier, rarissime, des relations existant entre les niveaux d'observation dans les systèmes (voir Von Bertalanffy, 1993). Le chou-fleur est un exemple commun de structure fractale. Le chou-fleur entier est isomorphe d'une branche du légume, elle-même isomorphe d'une fleurette. Les progrès dans la réfutation de cette conception fractale des ondes cérébrales s'expliquent aussi par la généralisation des outils d'analyse fréquentielle (Fourier, Ondelettes, voir annexe I). Sans ces outils, deviner la présence d'oscillations dans un signal complexe est impossible.

Après avoir précisé les différents niveaux d'enregistrement des signaux électrophysiologiques dans le cerveau, nous passerons en revue ces diverses méthodes. Ces observables sont supposés être des indices de la synchronisation. A ce titre, ils seront mis en relation avec une terminologie spécifique de ce domaine.

Les codages temporels : généralités

La conception classique du codage de l'information fait appel à la notion de **fréquence de décharge**. Le neurone transforme le paramètre à représenter en impulsions. A cette notion, concernée par l'activité des neurones isolément, s'oppose la notion de **codage par populations de neurones**. Depuis quelques années en effet, le signal unitaire des neurones laissent apparaître¹ toute une série d'autres dimensions de codage. On spéculait beaucoup sur les patterns temporels (*synfire chain*, séquences de décharge précises), les latences de réponse (codage asynchrone), la synchronie dans une population de neurones. On envisage que ces codes se superposent à la modification de la fréquence de décharge. A la limite, la question même de la légitimité de la fréquence de décharge, comme principal agent du transfert de l'information, est posée. Les modifications de la fréquence ne seraient-elles pas un artefact de la variation d'une autre dimension du signal ?

Ces évolutions théoriques suggèrent que les recherches des dernières décennies ont été exclusivement axées, non pas sur LE code cérébral, mais sur une dimension particulière de ce code, peut-être même pas la plus importante². Toute une série de dimensions informatives n'ont à présent pas été prises en compte. Des informations sur le **codage temporel** des neurones peuvent être trouvées dans Bair (1999), Gestner, Kreiter, Markram, & Herz, (1997).

Les outils traditionnels d'analyse du signal cérébral (voir annexe II, précédée des bases de l'analyse du signal en Annexe I) ne permettaient pas la mise en évidence de ces nouvelles dimensions de codage. Des méthodes spécifiques ont été élaborées ou transposées d'autres domaines scientifiques. Leur fonctionnement est présenté ci-après et détaillé en Annexes IV à VI.

Niveaux d'analyse des signaux cérébraux

L'analyse des signaux cérébraux in vivo peut se réaliser à **différentes échelles**. On distinguera les trois niveaux micro-, méso- et macroscopique (voir figure 15). A ces différents niveaux, sont recueilli plusieurs signaux, les SUA (*single-unit activity*), les MUA (*multi-unit activity*), le LFP (*local field potential*) et l'EEG (électro-encéphalogramme). Au niveau microscopique, on recueille un **enregistrement unitaire** (SUA). Un seul neurone contribue à ce signal. L'enregistrement est obtenu par le biais d'une micro-électrode. Le niveau mésoscopique regroupe deux méthodes d'**enregistrement local** de populations restreintes de neurones. Il se compose des potentiels d'action extra-cellulaires MUA (trouvés dans les hautes fréquences du signal, et traduisant les décharges des neurones jouxtant l'électrode) et

¹ Ces dimensions étaient masquées par la pratique du moyennage des signaux, réalisé au cours d'observations répétées, et destiné à éliminer les composantes singulières pour ne retenir que les composantes liées à la tâche.

² Ce courant théorique n'est pas conçu comme exclusif des précédentes théories du codage de l'information. La question pertinente n'est pas « Quel code ? ». Eggermont (1998) pose la bonne question ; « *How many codes ?* » (« combien de codes ? »).

du potentiel de champ local LFP (trouvé dans les basses fréquences du signal), un analogue mésoscopique de l'EEG. Les enregistrements sont également obtenus par le biais d'électrodes profondes. Enfin, au niveau macroscopique, se trouve les **enregistrements de surface** ou EEG sur le scalp. Ces échelles révèlent plusieurs aspects de l'activité bioélectrique. Ces aspects ne sont pas différents par la précision, mais révèlent des informations qualitativement différentes¹. Des informations complémentaires sur ces niveaux d'enregistrement figurent en annexe III.

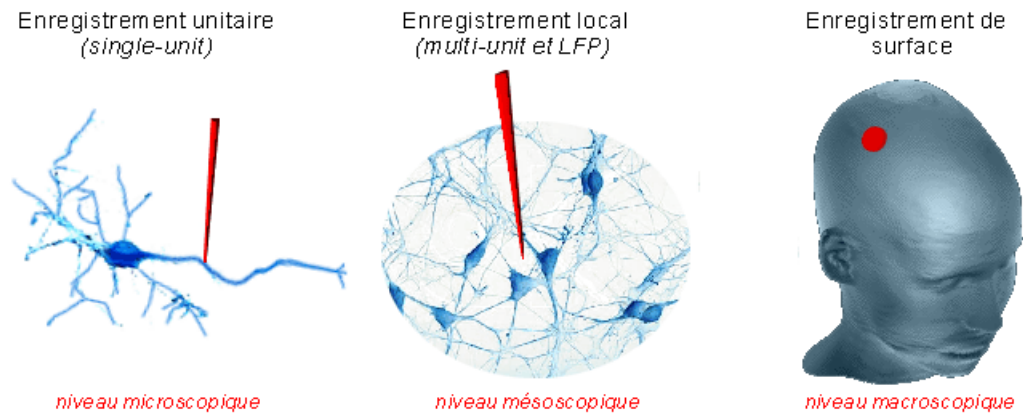


Figure 15 : niveaux d'enregistrement des signaux cérébraux.

Les codages temporels : outils

Au **niveau microscopique**, diverses méthodes permettent l'investigation des patterns temporels. Ces méthodes sont détaillées en annexe IV. L'**inspection visuelle du raster** (un index temporel des décharges) permet de s'assurer de la co-incidence des décharges. L'**autocorrélation** (Gerstein & Perkel, 1969) également appelée ISIH (InterSpike Interval Histogram) permet de déterminer si un neurone émet des décharges de manière périodique (régulière). La **cross-corrélation** (CCH) permet de savoir si les décharges émises par deux neurones présentent une relation temporelle, et éventuellement si cette relation est une synchronisation (co-incidence des décharges).

Aux **niveaux mésoscopique et macroscopique**, sont utilisées deux méthodes analogues des précédentes (annexe V). Les **analyses de composantes fréquentielles** (eg. Fourier ou ondelettes) permettent de dériver la densité spectrale (*PSD : power spectral density*), et indiquent dans quelles bandes fréquentielles le signal comporte des oscillations. Les **analyses de cohérence** quantifient le lien entre deux régions cérébrales, par le degré de synchronisation des phases du signal.

Hormis l'inspection visuelle du *raster* qui est une méthode qualitative, toutes ces méthodes permettent des quantifications. Leurs caractéristiques sont récapitulées sur la figure 16.

¹ « Thus, in practice, intracranial data provide different information, not more information, than is obtained from the scalp, in apparent contrast to the views of some physiologists. » (Nunez P. 2000). « LFPs [...] are not redundant with, or predictable from data on spikes. » (Bullock TH. 1997).

ENTITES CONSIDEREES

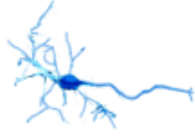
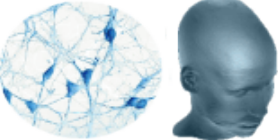
		Une	Deux
NIVEAU D'OBSERVATION Niveau microscopique  Niveaux mésoscopique & macroscopique 	AUTO-CORRELATION	déterminisme temporel dans les décharges d'un neurone <i>oscillations - périodicité</i>	CROSS-CORRELATION
	COMPOSITION FREQUENTIELLE	variations temporelles des oscillations <i>oscillations</i>	COHERENCE FREQUENTIELLE
		coïncidences des décharges de deux neurones <i>synchronisations - relations temporelles</i>	couplage des phases des oscillations de deux régions cérébrales <i>synchronisation des oscillations</i>

Figure 16 : Méthodes de base dans l'analyse du codage temporel. Actuellement, ces méthodes concernent le déterminisme temporel d'un neurone ou d'une zone du cerveau, et les interactions entre deux neurones ou entre deux zones.

Nous observons que les méthodes usuelles de quantification des propriétés temporelles des neurones sont **univariées et bivariées**. Il existe également quelques méthodologies fondées sur l'analyse de triplets, (eg. l'analyse de gravité, Gerstein & Kirkland, 2001), mais à notre connaissance, celles-ci sont peu usitées.

Or, par nature, le réseau cérébral fonde une **problématique multivariée**. Cette **restriction méthodologique bivariée** pose deux problèmes. D'une part, lors de la présentation du stimulus, l'apparition des synchronisations se superpose à l'augmentation de la fréquence de décharges. Cependant par un simple effet mécanique, lorsque davantage d'événements sont condensés dans une brève période, la probabilité de coïncidence temporelle augmente¹. A notre connaissance, aucune estimation n'a été faite sur les effets possibles de ce biais. Or, une méthodologie bivariée semble plus exposée à ce problème qu'une méthodologie multivariée, car la conjonction « accidentelle » de paires d'événements est sensiblement plus probable que la conjonction de trois événements ou davantage. D'autre part, le recours exclusif à des méthodologies bivariées empêche la formulation d'**hypothèses sur la structure et la dynamique des assemblées transitoires**², et inversement, l'absence d'hypothèses sur le recrutement des neurones et leur sortie de l'assemblée empêche l'élaboration de méthodologies multivariées.

De manière plus marginale, on étudie également la synchronisation en **croisant les différents niveaux d'observation** (voir également annexe VI). Cette observation peut se réaliser directement par **inspection visuelle**, sur un diagramme faisant coïncider temporellement LFP et SUA. Cette méthode est utilisée par exemple, par Eckhorn & Obermueller,

¹ N'oublions pas que la bande gamma correspond statistiquement à un événement toutes les 25 millisecondes environ. Or, la permissivité temporelle définissant la synchronisation varie de 1 à 5 millisecondes. Il reste peu de marge pour une erreur systématique dans ce type de modèles.

² Les développements dans cette direction seront tout à fait cruciaux pour l'hypothèse des synchronisations, car ils permettront de répondre à des questions fondamentales, telles que « est-ce que tous les neurones d'un pas de traitement entrent simultanément en synchronie, ou échelonnent-ils leurs contributions dans le temps ? » ou « quand une étape de traitement est résolue, les neurones persistent-ils dans l'assemblée, ou transmettent-ils un contenu informatif à des structures intervenant ultérieurement ? ».

(1993). L'observation peut également être quantitative. Le *STA* (*spike triggered average*) permet de voir si l'incidence des décharges de potentiels d'action se produit en relation avec les phases du signal LFP.

L'objet d'étude : synchronisation ? ...synchronisation ? ...ou synchronisation ?

Beaucoup d'aspects du signal neuronal peuvent recevoir le qualificatif de 'synchrone'. La synchronie (*syn-chronos*: avec/ensemble/en même - temps) dénote de comportements simultanés entre deux ou plusieurs unités.

Si l'on s'adresse à plusieurs **neurones individualisés**, ceux-ci peuvent faire **coïncider leurs décharges** sur une période assez longue (de l'ordre de la centaine de millisecondes) et avec plus ou moins de précision temporelle (généralement 0 à 5 millisecondes : à l'extrême, on parle de synchronie-zéro ou *zero-lag synchrony*, pour la forme de coïncidence la plus stricte des décharges). Cette forme de synchronisation n'est pas indépendante du comportement périodique des neurones. König, Engel, Singer, (1995) suggèrent en effet que des synchronisations sans périodicité sont possibles à très petite échelle. Cependant, dès lors que les neurones sont séparés de plus de quelques millimètres, les synchronisations seraient toujours supportées par des neurones ayant une activité périodique. Ces auteurs défendent l'idée que les oscillations sont le véhicule de la synchronisation. Les **oscillations** entretiennent donc avec les synchronisations une relation étroite, mais la nature de cette relation reste mal assurée à l'heure actuelle. Une présentation des différents tenants de ce problème sera donnée au paragraphe suivant.

Bien que ce comportement soit moins remarquable, on peut poser comme hypothèse l'existence de **coïncidences ponctuelles**. On trouve dans l'étude de Eckhorn et Obermuller (1993) une fraction de neurones qui ne se coordonnent qu'épisodiquement à l'activité oscillante synchrone de la population locale de neurones (c'est-à-dire que ces neurones déchargent en phase avec certaines des oscillations périodiques de la population locale, mais pas avec toutes ces oscillations). Pour le moment, cette possibilité théorique n'est pas plus développée dans la littérature, et ce phénomène est d'ailleurs beaucoup moins visible que les oscillations synchrones.

Enfin, les neurones peuvent **adopter, en simultané, un comportement quelconque** (passer d'une fréquence aléatoire à une fréquence fixe et inversement, baisser ou augmenter en fréquence, changer de dimension de corrélation...). Cette forme de synchronisation n'est pas l'objet de la théorie des assemblées neuronales synchrones.

Si l'on s'adresse à une **population de neurones** enregistrée globalement (formant un signal complexe sur le plan fréquentiel), le concept de synchronisation adopte de nouvelles significations. Comme au niveau unitaire, la synchronisation peut définir la **simultanéité d'un changement** sur une quelconque dimension de ce signal.

Une autre forme de synchronisation de ces signaux est le **couplage en phase de deux régions** cérébrales. Ce phénomène suggère soit une entrée commune, qui génère passivement des patterns similaires, soit une coordination endogène par synchronisation du signal.

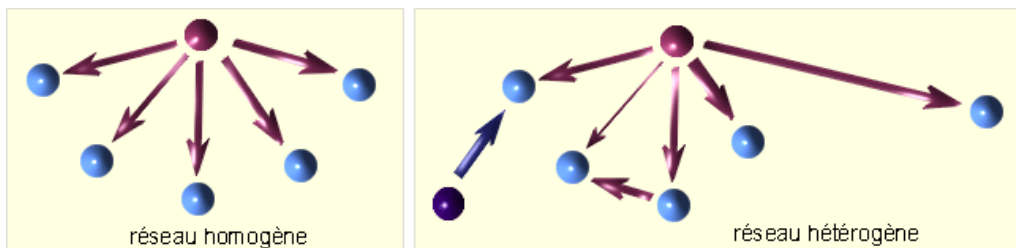
Enfin, le signal peut présenter des **oscillations**. Nous avons déjà mentionné le lien existant entre synchronisation et oscillations au niveau microscopique. Un problème similaire se reporte à ce niveau macroscopique. On observe que les oscillations apparaissent dans les mêmes bandes fréquentielles, et que les mêmes manifestations psychologiques (lien avec les activités cognitives de regroupement perceptif ; eg. réplication de Gray & Singer, 1989 chez l'animal en cross-corrélation par Müller, Bosch, Kreiter, Sosa, Valdes-Sosa & Rockstroh 1996, chez l'homme en analyse des oscillations gamma sur le scalp) accompagnent leur apparition. Les deux équations [même fréquence d'oscillation] ; [mêmes corrélats fonctionnels] ont avalisé l'isomorphisme entre les oscillations gamma des signaux locaux et les oscillations des signaux de surface. Le postulat sous-jacent est, là encore, une relation fractale : les oscillations gamma macroscopiques sont la résultante d'une action de masse, sur la fréquence gamma, d'un nombre critiques d'unités neuronales. Cette relation fractale n'a pourtant rien d'obligatoire, d'autres patterns d'activité peuvent engendrer ces oscillations macroscopiques, sans que le même comportement soit observé au niveau microscopique. Et des oscillations microscopiques peuvent disparaître des niveaux macroscopiques (voir annexe VIII). En résumé, on observe là encore un lien dont la nature et le sens sont mal caractérisés.

Concepts d'oscillation et de synchronisation

En portant le débat sur le plan conceptuel, nous voyons que la **synchronisation** est distincte de l'**oscillation**. L'oscillation se définit comme un comportement auto-centré de l'observé (neurone ou réseau). Au contraire, la synchronisation est une propriété relationnelle. Cependant, dans la pratique, les assemblées synchrones s'appuient sur une coordination fréquentielle (même périodicité) et temporelle (même phase). Dès les premières observations, (Eckhorn & coll., 1988 Gray & Singer, 1989), les deux concepts apparaissent solidaires, puisqu'il s'agit d'étudier des oscillations « cohérentes » ou « synchronisées ». Plus tard, le systématisme du lien entre les deux concepts est questionné ; par exemple König, Engel et Singer, (1995), posent l'hypothèse que **les oscillations facilitent** la mise en place de **la synchronisation** à grande échelle. Les auteurs observent qu'à petite distance, des synchronisations se manifestent sans oscillation, mais dès lors que la distance séparant les deux sites d'enregistrement excède 2 mm, alors la synchronisation s'accompagne toujours d'oscillations. La présence obligatoire des oscillations pour les synchronisations à grande échelle est contestée par d'autres auteurs (Roy, Dear et Alloway, 2001 ; Bressler, Coppola & Nakamura 1993, 1995 , Shumikhina & Molotchnikoff, 1999). Pour ces auteurs, l'apparition des oscillations en complément de la synchronisation est la conséquence des stimuli employés (eg. réseaux sinusoïdaux) mais des stimuli plus communs donneraient lieu à des synchronisations sans oscillations.

Pour les défenseurs du lien entre oscillation et synchronisation, le rôle facilitateur des oscillations sur la synchronisation est justifié de la manière suivante : si la synchronisation est un phénomène émergent simple au niveau conceptuel, en revanche, son implémentation dans un réseau hétérogène n'est pas une affaire simple (voir figure 17), tout particulièrement en raison des délais de conduction variables dans le réseau. Si le réseau neuronal était un réseau homogène (équidistance des neurones cibles et égalité des vitesses de conduction), la synchronisation pourrait émerger passivement, lorsque les neurones-cibles seraient sollicités par un input commun. Or, le **tissu neuronal est**

complexe. Les neurones cibles d'un neurone donné sont situés à des distances variables. Les vitesses de conduction (liées au calibre des neurones et à leur myélinisation) sont différentes. Il existe entre plusieurs neurones « d'un même pas de traitement » des connexions latérales. Ces neurones sont en outre susceptibles de recevoir des activations d'autres traitements, ou d'autres étapes du traitement (modulations top-down par exemple). Pour que tous ces neurones-cibles reçoivent le signal neuronal en même temps, il n'est plus possible de se reposer sur un mécanisme passif¹. Tout ceci implique qu'un **mécanisme endogène maintienne la synchronisation**, menacée de dissolution par cette complexité neuronale. Les données actuelles suggèrent que les oscillations occuperaient ce rôle de préservation de la synchronisation. Dans cette hypothèse, on comprend pourquoi les synchronisations existent sans oscillation à petite échelle. Plus l'assemblée est petite et locale, moins la variabilité peut compromettre sa pérennité.



Implémentation de la synchronisation

figure 17 : Implémentation de la synchronisation. Dans un réseau homogène, la synchronisation peut émerger d'un input commun entre les unités, et s'établir passivement. Dans un réseau hétérogène, le problème est plus complexe. La distance entre les neurones et la vitesse de conduction sont variables, des connexions latérales existent, qui pourraient perturber le système, et les neurones de l'assemblée reçoivent des influences d'autres unités, comme par exemple des modulations top-down. Dans ce cas, la synchronisation doit être activement maintenue. Les oscillations semblent impliquées dans la préservation de ce lien temporel.

Cette chronologie de la théorie des assemblées oscillantes synchrones explique les substitutions fréquentes de ces deux termes d'oscillation et de synchronisation. Cependant, deux **précautions** doivent être prises dans la **généralisation de cette hypothèse**. D'une part, les deux niveaux microscopique et macroscopique doivent a priori être distingués. Si un lien est fermement établi entre les oscillations et les synchronisations des potentiels d'action, rien ne saurait être conclu d'emblée sur le lien entre les oscillations et les cohérences du niveau macroscopique, ni sur les relations transverses des niveaux microscopique et macroscopique. Les études fréquentielles sur le signal EEG semblent postuler que l'émergence d'oscillations gamma au niveau du scalp impliquent qu'une fraction critique de neurones se soit mise à osciller en phase sur cette fréquence, faute de quoi les composantes s'annulent (annexe VIII). Nous n'avons pas trouvé d'études garantissant ce postulat. En outre, quand bien même un lien macroscopique serait établi, les éléments actuels peuvent garantir une relation surjective (lorsque les synchronisations se manifestent à grande échelle, les oscillations sont toujours présentes), mais pas une relation bijective (chaque fois que les oscillations sont présentes, les synchronisations le sont également). Aussi, l'utilisation des oscillations comme outil pour caractériser la synchronisation comporte un risque de faux-positif (prédire l'existence de synchronies alors qu'elles n'existent pas).

¹ Cette hétérogénéité des réseaux neuronaux rend peu probable le fait que la synchronisation soit le résultat d'un input commun. Ceci est d'autant plus critique que l'on s'éloigne des premières étapes de traitement, et qu'alors, des degrés de libertés se sont insinués dans la temporalité initiale du signal.

Conclusion

Cette présentation méthodologique permet tout d'abord de constater **l'hétérogénéité** des observables englobés sous la thématique des assemblées neuronales synchrones. Nous avons vu tout un éventail de méthodes destinées à exhiber ces synchronisations. Tout particulièrement, **trois définitions opérationnelles** sont à retenir de cette présentation dont deux sont applicables aux recherches chez l'homme.

Au niveau unitaire, la **synchronisation** concerne les **décharges** de potentiel d'action, (*spike synchronization*). Sa mise en évidence nécessite une analyse de la cross-corrélation, suite au recueil du signal de décharge par implantation de micro-électrodes. Cette méthodologie invasive est principalement réservée aux recherches chez l'animal (le singe et le chat ont été particulièrement étudiés). C'est ce niveau d'observation qui est le plus directement lié à la théorie des assemblées transitoires synchrones.

Aux niveaux plus globaux, les méthodologies sont applicables à des investigations menées sur l'homme sain et pathologique. Les outils sont la cohérence fréquentielle (**synchronisation des phases** du signal), qui traduit le degré de couplage entre deux zones enregistrées en intra-cortical ou sur le scalp, et l'augmentation des **oscillations dans la bande gamma**, assumée comme étant indirectement liée à l'entrée en synchronisation de nombreux neurones (action de masse).

Le lien entre ces outils et la théorie des assemblées transitoires synchrones est défendu par plusieurs auteurs, mais reste à conforter devant plusieurs **polémiques méthodologiques** (eg. pic de l'auto-corrélogramme artificiellement causé par la période réfractaire des neurones (Bar-Gad, Ritov, Vaadia & Bergman, 2001) ; effet du recours à des MUA non-résolus (Gerstein, 2000) ; biais induits par des neurones temporellement indépendant, mais ayant les mêmes latences moyennes de décharge, ou les mêmes fréquences de décharge intra-essai (Brody, 1999), oscillations gamma artificiellement créées par harmoniques issus de changements dans la bande alpha, (Jürgens, Rosler, Henninghausen & Heil, 1995) oscillation gamma résultant de l'activité musculaire, cohérence fréquentielle créée entre électrodes par la conduction passive du signal dans le volume cérébral ; synchronie artificielle liée à la variabilité des conditions expérimentales ou comportementales (Ben-Shaul, Bergman, Ritov & Abeles, 2001)), **théoriques** (variabilité de décharge des neurones, Shadlen et Newsome, 1998 ; Panzeri, Schultz, Treves, & Rolls, (1999), Shadlen & Movshon, 1999 ; voir aussi Lestienne 2001) **ou fonctionnelles** (réfutation sur la ségrégation figure-fond dans V1, Lamme & Spekreijse, 1998 ; proposition d'interprétations alternatives lors de regroupement perceptif sur des stimuli texturés, Farid & Adelson 2001).

Observations des synchronisations en vision

Parallèlement à la formulation théorique des assemblées neuronales transitoires et à l'évolution des méthodes, il s'agissait d'observer des manifestations de la synchronie dans le cerveau. Des observations ont été réalisées aux trois niveaux opérationnels décrits ci-avant (corrélation des décharges ; oscillations γ ; et couplage en phase). Ces études

cherchaient à comprendre un éventuel rôle de la synchronisation dans les fonctions sensorielles (particulièrement les systèmes olfactif, visuel et auditif), sensori-motrices, mnésiques, et attentionnelles. Notre présentation se concentrera sur les études liées à la fonction visuelle.

Plusieurs études montrent des **synchronisations dépendantes du stimulus**. Ces recherches suggèrent que la synchronisation permet la sélection des neurones participant à un traitement commun (liage perceptif). Au niveau le plus local, il peut s'agir d'une assemblée transitoire formée des activités des neurones de V1 codant pour des segments orientés. A un niveau plus global, il peut s'agir de regroupements perceptifs plus complexes, tels que ceux sollicités dans des tâches de ségrégation figure-fond difficile, dans la perception de « Mooney face¹ », ou de figures de Kanizsa. Nombre de ces études ont été menées par des chercheurs allemands. Ce foyer géographique n'est sans doute pas anodin, puisque l'Allemagne a gardé une très forte tradition de recherche autour de la théorie de la Gestalt. Cette présentation s'arrêtera sur les études les plus emblématiques du lien entre regroupement perceptif et assemblées synchrones. Au terme de cette présentation, l'objectif sera d'être en mesure de choisir, parmi les méthodes applicables à l'homme, celle qui s'avérera la plus pertinente pour la poursuite ultérieure de ce projet.

Investigations en micro-électrode chez l'animal

Les études liées aux oscillations synchrones apparaissent entre 1987 et 1989 (Eckhorn, Bauer, Jordan, Borsh, Kruse, Munk, Reitboeck, 1988 ; Gray & Singer, 1989, préliminaires dans Gray & Singer, 1987). Elles utilisent des enregistrements similaires (SUA, MUA et LFP des aires 17 et 18 du cortex visuel du chat) mais proposent des stratégies d'analyse encore variables. Le résultat de ces études est (1) l'observation d'**oscillations dépendantes du stimulus**, et (2) le **couplage en phase** de ces oscillations, sous certaines conditions lorsque les neurones sont co-activés par le stimulus.

Ces oscillations sont nommées **résonances² gamma** (ou oscillations cohérentes) par Eckhorn et son équipe (1988). Elles se manifestent tout particulièrement **lorsque ces neurones sont activés par leurs stimuli spécifiques** (eg. position préférée, mouvement et vitesse préférés). Elles sont plus évidentes dans le signal LFP que dans les MUA et SUA. Ces résonances sont cohérentes (couplées en phase) dans une même colonne corticale (entre des signaux générés par des couches corticales différentes), entre deux hypercolonnes proches, et entre les deux aires 17 et 18 lorsque les neurones enregistrés codent des propriétés similaires (eg., même préférence d'orientation).

Gray & Singer (1989) observent également des oscillations lors de la présentation de segments lumineux orientés, se déplaçant dans le champ récepteur des cellules enregistrées. Lorsque le **stimulus optimal** est utilisé, on observe un

¹ Les « Mooney face » sont des images de visages dont le contraste est saturé. Le visage est ainsi facilement reconnaissable dans son orientation habituelle, alors qu'il est beaucoup plus rarement perçu lorsqu'il est présenté à l'envers.

² La résonance est également un concept polysémique. Dans une autre acception, le concept désigne une préférence fréquentielle intrinsèque des neurones (eg. Hutcheon & Yarom, 2000), c'est-à-dire une tendance à répondre plus fortement à une excitation d'une fréquence déterminée. Bien que potentiellement lié à l'établissement d'oscillations stables, ce phénomène doit être distingué du couplage en phase de deux signaux oscillants, qu'Eckhorn baptise oscillations résonnantes.

pattern rythmique de décharges (*bursts*) survenant durant les phases négatives des oscillations LFP (voir figure 18). Cette corrélation n'est pas triviale ; usuellement, on observe une absence de corrélation entre le LFP d'une part, et les SUA et MUA d'autre part (Stryker, 1989). L'autocorrélogramme des MUA non résolues montre un pattern oscillant, qui est interprété comme une **action rythmique synchrone de masse des neurones** incrustés dans le signal. Ces auteurs défendent l'existence de coopérations entre des neurones distants sur la base d'oscillations synchrones. Ils présentent alors la « phase du signal » comme un paramètre additionnel, aux côtés des paramètres classique d'amplitude et de durée de la réponse. Cette étude suggère en outre **l'origine intracorticale des oscillations**. Des oscillations gamma sont obtenues dans le cortex visuel, alors qu'aucune oscillation n'est observée dans le LGN. Les auteurs interprètent cette observation comme un argument de la formation intra-corticale des oscillations gamma, puisque l'oscillation ne semble pas s'être transmise passivement des aires antérieures¹.

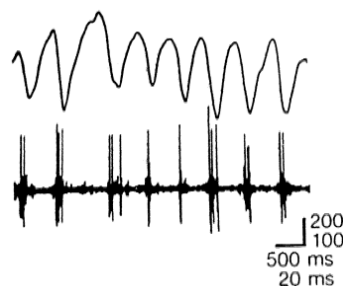


Figure 18 : les bouffées de décharges synchronisées sur le signal LFP (reproduit de Gray & Singer, 1989). Le signal MUA (en bas) est constitué de bouffées de 4 à 5 décharges. Ces bouffées de potentiels d'action sont synchronisé au signal LFP (en haut).

Ces premières séries d'études montraient seulement des oscillations dépendantes du stimulus. On stimulait les groupes de neurones avec leurs stimuli préférés, et on enregistrat l'entrée en oscillation de ces neurones lors de l'événement (« pour un stimulus donné, quels neurones se synchronisent ? »). Par la suite, les études ont utilisé des stimuli présentant une plus ou moins grande cohérence perceptive. Il s'agissait ainsi de démontrer **que plus un stimulus était perceptivement cohérent, plus les neurones impliqués exhibaient des oscillations synchrones** (« quelles propriétés des stimuli sont les plus aptes à générer des oscillations synchrones ? »).

Gray, König, Engel & Singer (1989) enregistrent les signaux MUA dans des colonnes de neurones ayant la même orientation préférée, mais des champs récepteurs disjoints. Ils stimulent les clusters de neurones avec des barres animées de mouvements (deux barres courtes ayant deux mouvements distincts, deux barres courtes ayant le même mouvement, et une seule barre longue, voir figure 19a). Le cross-corrélogramme de ces deux sites d'enregistrement montre une gradation dans l'apparition des oscillations synchrones (figure 19b). Celles-ci sont inexistantes pour les deux barres animées de mouvements opposées, moyennes pour les deux barres animées de mouvements solidaires et fortes pour la barre unique. Cette étude suggère l'existence d'une **relation entre** les opérations cognitives de **cohérence perceptive** et les phénomènes neurophysiologiques **d'oscillations synchrones**.

¹ Notons que cette observation ne compromet pas l'implication des structures thalamiques dans la genèse et l'entretien des oscillations synchrones (le thalamus est réputé être un pacemaker des oscillations). Ces observations démontrent simplement que les oscillations ne sont pas un simple pattern pré-existant, conservé de pas-de-traitement en pas-de-traitement.

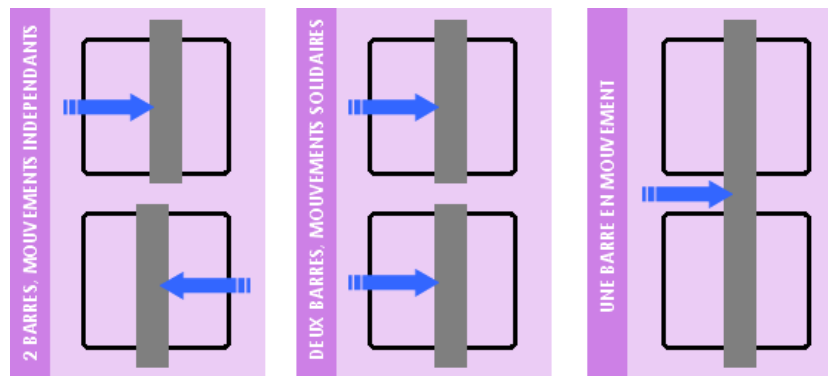


Figure 19a : Stimuli utilisés par Gray et collaborateurs dans leur étude (Gray, König, Engel et Singer, 1989). Les carrés noirs symbolisent les champs récepteurs respectifs des deux clusters de neurones enregistrés. Les stimuli sont des barres en mouvement : soit deux barres animées de mouvements antagonistes, soit deux barres animées d'un mouvement solidaire, soit une longue barre partagée sur les deux champs récepteurs.

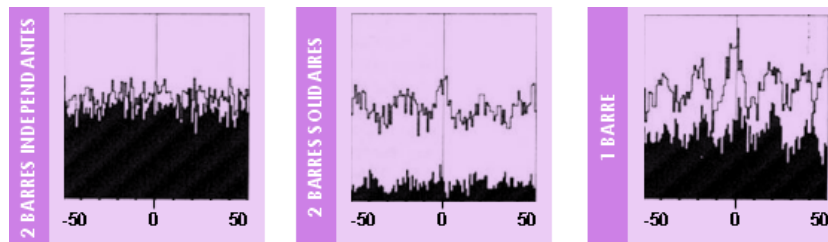


Figure 19b : cross-corrélogrammes des MUA de deux clusters de neurones ayant des champs récepteurs disjoints mais contigus, lors de la présentation des 3 types de stimuli reportés à la figure 19a (adapté de Gray, König, Engel & Singer, 1989).

Un autre pas est franchi, lorsque Engel, König, & Singer (1991) présentent les résultats d'une nouvelle étude, suggérant que la **segmentation perceptive** peut être accomplie par des assemblées synchronisées. Les auteurs procèdent à l'enregistrement de quadruplets de signaux MUA. Ils établissent tout d'abord le comportement de ces clusters de neurones (orientation préférée, corrélation des signaux, voir figure 20). Deux clusters (1, 3) oscillent en synchronie pour un segment orienté à 0 degré. Ces neurones forment une première assemblée. De même, deux clusters (2, 4) oscillent en synchronie pour un segment orienté à 112°. Ces neurones forment une seconde assemblée. La présentation d'un stimulus intermédiaire (représenté en rouge) engage un représentant de chacune des deux précédentes assemblées dans une oscillation synchrone (les clusters 2 et 3).

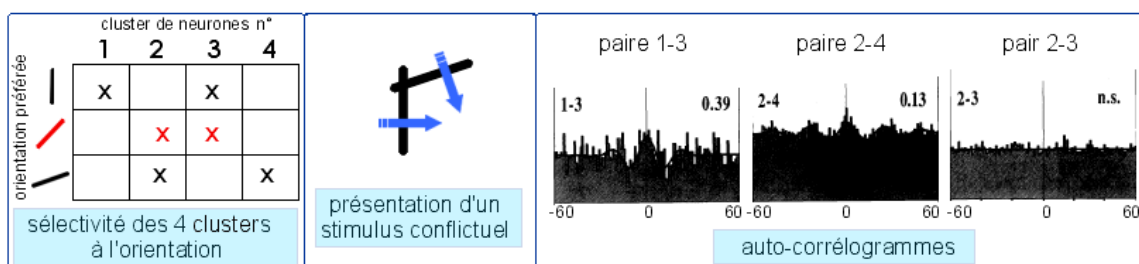


figure 20 : l'expérience de Engel, König, Singer, (1991). Dans une première étape, les auteurs déterminent l'orientation préférée de 4 clusters de neurones. Deux clusters (1, 3) sont coactivés par une barre orientée à 0°. Deux clusters (2, 4) sont coactivés par une barre orientée à 112°. Enfin, deux clusters (2, 3) sont coactivés par une barre d'orientation intermédiaire (en rouge). Un stimulus conflictuel est alors présenté. Il est formé de la présentation simultanée de deux barres en mouvement, une à 0° et une à 112°. Les auto-corrélogrammes sont calculés pour ces trois paires de clusters. Alors que les paires (1, 3) et (2, 4) montrent un comportement oscillatoire en phase, la paire (2, 3) exhibe un auto-corrélogramme plat.

Dans un second temps, les auteurs soumettent ces clusters de neurones à des stimuli conflictuels. Deux barres orientées respectivement à 0° et 112° se déplacent dans le champ récepteur des 4 clusters de neurones. Bien

évidemment, les neurones 1 et 3, d'une part, et 2 et 4, d'autre part, s'engagent dans des oscillations synchrones. Comment se comportent les neurones 2 et 3 ? Ces deux clusters restent indépendants : l'auto-corrélogramme reste plat. Ce résultat suggère que **deux objets distincts mais co-présents sont représentés par deux assemblées différentes**. L'assemblée oscillante synchrone semble ainsi potentiellement impliquée dans la segmentation perceptive.

A ce moment de l'évolution de la théorie, les deux aspects du « *binding problem* », **l'intégration et la ségrégation**, ont été considérés au regard des **assemblées synchrones**. Ces premiers résultats semblent favorables au lien entre le phénomène neurophysiologique et les opérations perceptives liées à la notion d'objet.

La segmentation perceptive est confortée appuyée par cette étude de Castelo-Branco, Goebel & Neueschwander (2000), qui utilisent des plaids pour stimuler les neurones des aires 18 et PMLS du chat (voir figure 21a). Les plaids sont des stimuli constitués de la superposition de deux réseaux (*gratings*), orthogonaux l'un à l'autre. Chaque réseau est animé d'un mouvement perpendiculaire à son orientation. Perceptivement, ces plaids sont interprétés de deux façon, selon le traitement graphique des intersections. Si les intersections sont solides (opaques), alors le sujet (humain ou animal) perçoit une surface unique animée d'un mouvement intermédiaire ; il interprète la scène comme constituée d'une seule surface. Si les intersections sont transparentes, alors le sujet perçoit le mouvement relatif des deux réseaux ; il interprète la scène comme constituée de deux surfaces distinctes glissant perpendiculairement l'une sur l'autre. Des clusters de neurones ayant des champs récepteurs disjoints, mais la même préférence d'orientation, se synchronisent lors de la présentation des stimuli perçus comme une surface unique. En revanche ils restent indépendants lors de la présentation de stimuli perçus comme composés de deux surfaces (voir figure 21b).

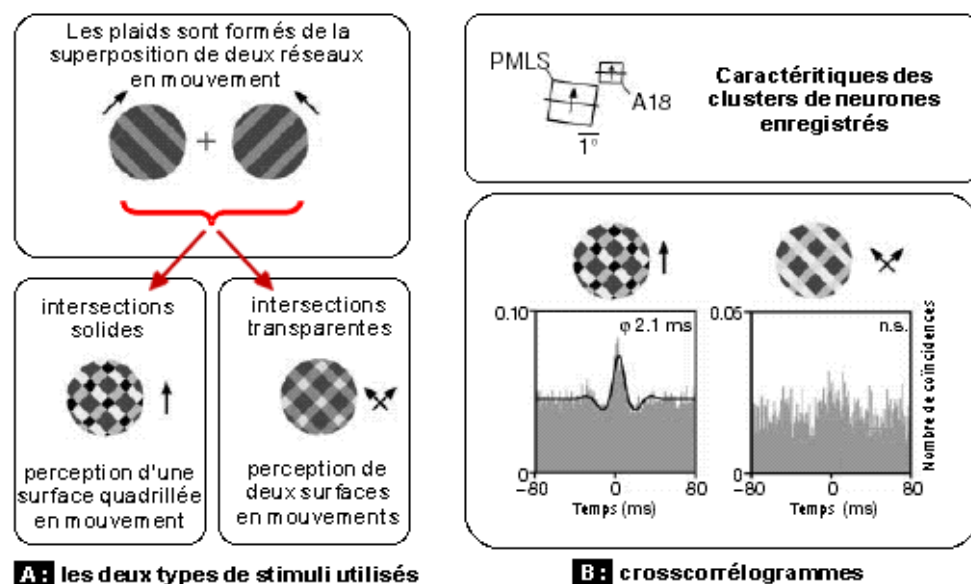


Figure 21 : la ségrégation des surfaces dans l'expérience de Castelo-Branco et coll. (2000). Les stimuli sont des plaids, perçus respectivement comme une seule surface quadrillée en mouvement ou deux surfaces en mouvements. Les clusters de neurones des aires 18 et PMLS du chat se synchronisent lorsque les intersections sont solides (perception d'une seule surface), mais pas lorsque les intersections sont transparentes (perception de deux surfaces en mouvements distincts). (adapté de Castelo-Branco & coll., 2000).

Ces résultats appellent à être confortés par des répliques et des généralisations : d'une part, les résultats ne sont pas suffisamment clairs pour être inattaquables. D'autres équipes utilisant des méthodes similaires trouvent que le

phénomène est de moins grande ampleur (les oscillations synchrones ne concerneraient qu'une toute petite fraction de neurones), d'autres encore ne répliquent pas ces résultats, sans qu'il soit toujours possible de savoir si l'échec de la réplication est liée à des considérations méthodologiques ou à une réelle absence du phénomène (voir discussion dans Lestienne, 2001).

D'autre part, les cross-corrélations sont calculées sur des MUA non-résolus. Ces signaux incrustent une population localisée de neurones, ayant une tendance à donner des réponses similaires certes¹, mais dont les unités ne produisent pas des réponses identiques. Outre des considérations méthodologiques (effectif accru des événements), les premières études en micro-électrode défendaient la **légitimité du recours aux MUA** par des arguments théoriques : les assemblées synchrones étant une problématique de population, il était légitime de les aborder en tant que population (voir Singer, 2000). A ce stade, **il était possible d'objecter que les oscillations synchrones étaient un artefact** lié au recours à des populations de neurones. L'émergence d'une oscillation synchrone au niveau collectif crée un phénomène visible pour l'électrophysiologiste (dans les signaux LFP et MUA), mais les unités neuronales ne sont pas forcément aptes à décoder cette « ambiance » oscillante, car leur seul mode d'expression inter-individuel des neurones est la transmission synaptique entre unités interconnectées². Cela pose le problème de la distinction entre une manifestation neurophysiologique, corollaire d'opérations cognitives (des signaux individuels, qui, considérés sous un angle collectif, montrent des patterns géométriques très reconnaissables avec nos instruments de mesure, mais ne supportant pas en propre le traitement de l'information), et des activités neuronales transmettant réellement de l'information par une structure de décharge rythmique et synchrone.

Eckhorn et Obermuller (1993) apportent quelques clarifications à ce problème. Ils dissocient **trois types de comportement des signaux SUA lorsque le LFP et les MUA présentent des oscillations**. Ces trois types sont :

- un couplage rythmique (*rythmic state* : le neurone est engagé dans une oscillation synchrone du signal de population). Ce couplage suggère que ce neurone est un des contributeurs constants des oscillations globales
- un couplage non-rythmique (*lock-in state* : le neurone est synchronisé sur le signal de population, mais son activité n'est pas périodique). Ce couplage suggère que ce neurone est occasionnellement impliqué dans l'assemblée
- un état de non-participation (*non-participation state* : le neurone n'est pas corrélé à l'oscillation globale).

¹Les neurones sont regroupés fonctionnellement dans le cortex visuel (eg. Ringo, 1991). Par exemple, deux neurones de V1 ayant un champ récepteur similaire et une préférence d'orientation similaire, seront voisins.

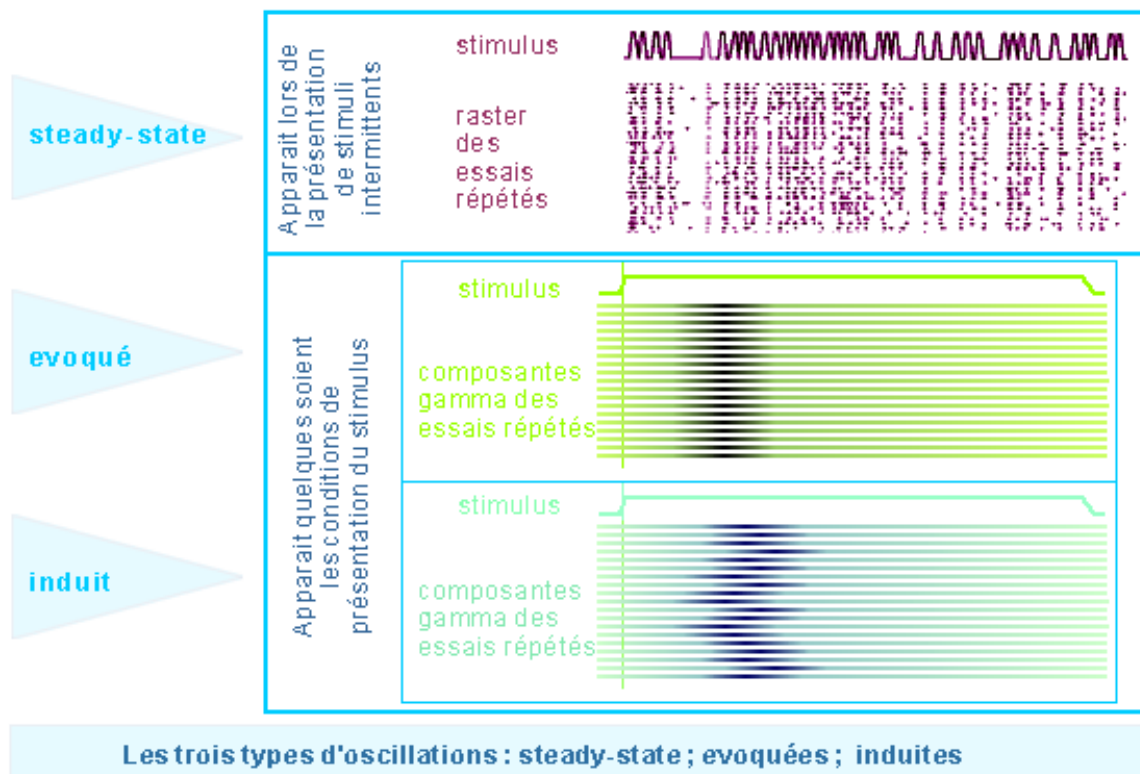
² En d'autres termes, les neurones ne décodent pas des oscillations « interstitielles », mais traitent seulement les paramètres temporels et fréquentiels des afférences synaptiques.

Le couplage est affaibli (on passe d'un couplage rythmique à non-rythmique, puis à une non-participation), lorsqu'on passe d'un stimulus optimal à un stimulus sub-optimal pour le neurone enregistré en SUA. Ces données supportent l'hypothèse que la synchronie est effectivement un code neuronal, et non un artefact émergent.

Les oscillations gamma des niveaux méso- et macroscopique

Pour reprendre la classification de Galambos (1992), nous distinguerons trois types d'oscillations dépendantes du stimulus ; les oscillations steady-state, évoquées et induites (voir figure 22). Les **oscillations « steady-state »**, apparaissent lors de la présentation d'un stimulus intermittent. Les décharges se synchronisent massivement sur le stimulus, et font apparaître des oscillations au niveau macroscopique. Plutôt que générées par le traitement cognitif, ces oscillations sont provoquées par la structure externe du stimulus. Parmi les deux types d'oscillations générées de manière interne dans le cerveau, les **oscillations évoquées** sont celles dont la mise en évidence a été la plus précoce. Elles surviennent en relation temporelle précise avec le stimulus (*phase-locked, time-locked*). Par cette propriété, elles apparaissent dans les signaux moyennés des Potentiels Evoqués¹. Les **oscillations induites** ont été plus tardivement mises en évidence. En léger décalage temporel les unes par rapport aux autres, elles disparaissent des signaux moyennés (voir procédure de mise en évidence annexe IX). Certains auteurs (Bertrand & Tallon-Baudry, 2000) défendent l'idée que ces dernières sont les plus liées aux activités cognitives (eg. à la représentation des objets et aux opérations mnésiques). La topographie de ces composantes gamma semble suggérer, plutôt qu'une activation diffuse et peu spécifique (attention), une relation directe avec la tâche cognitive en cours (Tallon-Baudry et Bertrand, 1999).

¹ Les potentiels évoqués (PE ou ERP, *event related potentials*) consistent à moyenner un grand nombre d'enregistrements EEG, calés sur le moment d'apparition du stimulus. Cette pratique permet d'améliorer le rapport signal sur bruit. Les composantes spécifiques s'annulent, tandis que les composantes liées au stimulus demeurent (voir annexe VII).



Les trois types d'oscillations : steady-state ; évoquées ; induites

figure 22 : les trois types d'oscillations. Le rythme steady-state est provoqué par les intermittences du stimulus et par la synchronisation des décharges sur ces événements répétés (comme le démontre le raster, en haut). Si un stimulus périodique est appliqué, alors un rythme périodique est enregistré en réponse, suscité par les salves de réponses neuronales calées sur le stimulus. Les oscillations générées intracorticalement par le stimulus sont de deux types : les oscillations évoquées (verrouillées en phase sur l'apparition du stimulus), dont le résultat transparait dans les potentiels évoqués et les oscillations induites, qui ne sont pas scrupuleusement verrouillées à l'apparition du stimulus, et qui s'annulent lors de traitements du signal par moyennage.

En 2001, Herrmann réalise une expérimentation en *flicker* (clignotement), afin de tester la réponse EEG lors de la présentation de lumières clignotant à des **fréquences variables**, allant de 1 à 100 Hz. Les potentiels évoqués exhibent des oscillations *steady-state* à toutes les fréquences, au moins jusqu'à 90 Hz. L'auteur rapproche ces données de celles obtenues chez le chat en enregistrement MUA et LFP.

En 1998, Elliot et Müller proposent une ingénieuse expérience d'**amorçage** d'un traitement cognitif par un **pré-masque**¹ clignotant de manière sous-liminaire. Leur objectif est de montrer qu'un pré-masque ayant des caractéristiques communes avec la cible (fonctionnant comme une amorce pour la cible) facilite le temps de réponse sous certaines conditions de fréquence du clignotement de ce pré-masque. Ce pré-masque est sous-liminaire ; à cause de sa haute fréquence, le sujet ne perçoit pas le clignotement (voir figure 23). Les auteurs testent toute une série de fréquence de clignotement du pré-masque. La seule fréquence qui permet d'améliorer significativement la performance cognitive est la fréquence 40 Hz (générant des oscillations steady-state gamma). Cette étude sera complétée par des enregistrements électrophysiologiques dans (Elliot, Herrmann, Mecklinger & Müller, 2000).

¹ Un masque est une forme graphique sans signification appliquée au terme de la présentation de stimuli, afin d'éliminer la persistance rétinienne. Un pré-masque fonctionne selon la même logique, à ceci près qu'il est présenté avant le stimulus.

le sujet perçoit l'affichage en 2 périodes :
un pré-masque constitué de 9 croix, suivi de l'apparition de la cible (un carré de Kanizsa)

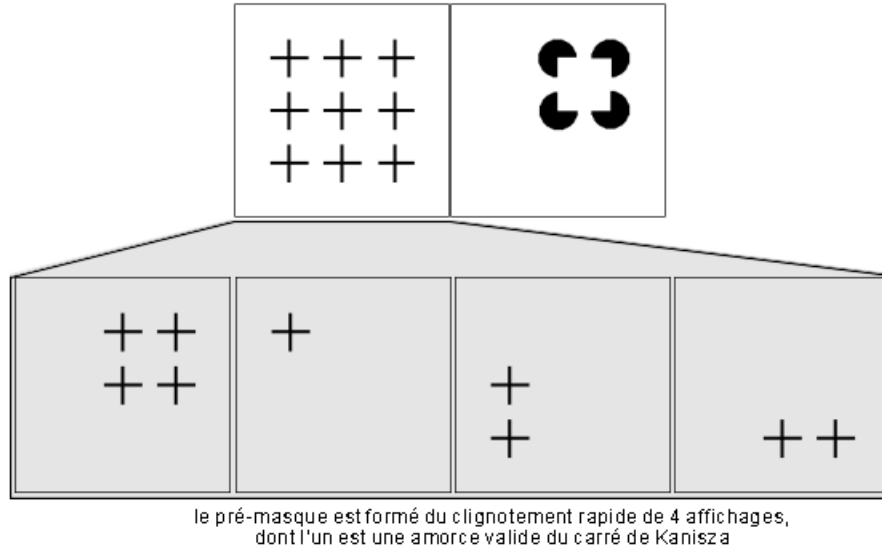


Figure 23 : l'expérience en flicker (en clignotement) de Elliot et Müller, 1998. Lorsque le pré-masque valide est présenté à la fréquence de 40 Hz, les sujets sont plus rapides à détecter le carré de Kanizsa que lorsqu'un pré-masque invalide est présenté. L'amorce n'est efficace qu'à la fréquence générant des oscillations steady-state gamma.

En 1996, Müller, Bosch, Kreiter, Sosa, Valdes-Sosa et Rockstroh transposent l'expérience de Gray & Singer (1989) chez l'humain en enregistrement EEG. Les auteurs rapportent une augmentation significative de la densité des composantes gamma, lors de la présentation d'une seule longue barre, comparée à la présentation de deux barres animées de mouvements opposés (voir figure 19a p.41). Les résultats sont en accord avec ceux obtenus 7 ans auparavant chez l'animal.

Basar-Eroglu Struber, Kruse, Basar & Stadler, (1996) présentent des **stimuli visuels bistables** (le sujet interprète tour à tour le stimulus de deux façons différentes). Dans cette expérience, les stimuli étaient des *SAM* (*stroboscopic alternative motion*, mouvement alternatifs stroboscopiques) ; des stimuli dont le mouvement de rotation est perçu alternativement dans le sens des aiguilles d'une montre ou en sens inverse. Ces stimuli occasionnent une augmentation des **oscillations gamma** sur le scalp entier, et une augmentation significative des activités gamma frontales lors de la perception d'une permutation du sens de rotation. En 2000, Struber, Basar-Eroglu, Hoff et Stadler montrent variabilité inter-individuelle dans l'interprétation des figures multistables. Les sujets ayant une perception très changeante des figures multistables présentent une densité spectrale gamma accrue, relativement aux sujets ayant une perception peu changeante de ces figures.

L'équipe de Tallon-Baudry et Bertrand propose, entre 1996 et 2000, plusieurs études montrant le lien entre la **représentation des objets et les oscillations gamma induites**. Les enregistrements étaient des EEG recueillis sur le scalp, et des LFP enregistrés lors d'interventions chirurgicales de résection de foyers épileptogènes. Deux de ces études ont trait à la perception visuelle et au regroupement perceptif. L'une (Tallon-Baudry, Bertrand, Delpuech & Pernier, 1997) utilisait des triangles de Kanizsa, et mettait en oeuvre un mécanisme de regroupement perceptif dominé

par les processus bottom-up. La seconde (Tallon-Baudry, Bertrand, Delpuech & Pernier, 1997), proposait une tâche de ségrégation figure-fond difficile (figure du dalmatien camouflé), et mettait davantage en oeuvre des processus top-down. Dans les deux cas, on observait, aux alentours de 250 millisecondes après la présentation de la cible, une **augmentation de la densité des composantes gamma lorsque le stimulus était cohérent**. Cette augmentation des oscillations gamma n'apparaissait pas lorsque le stimulus n'était pas cohérent, ou lorsqu'il n'était pas perçu comme tel (voir figure 24).

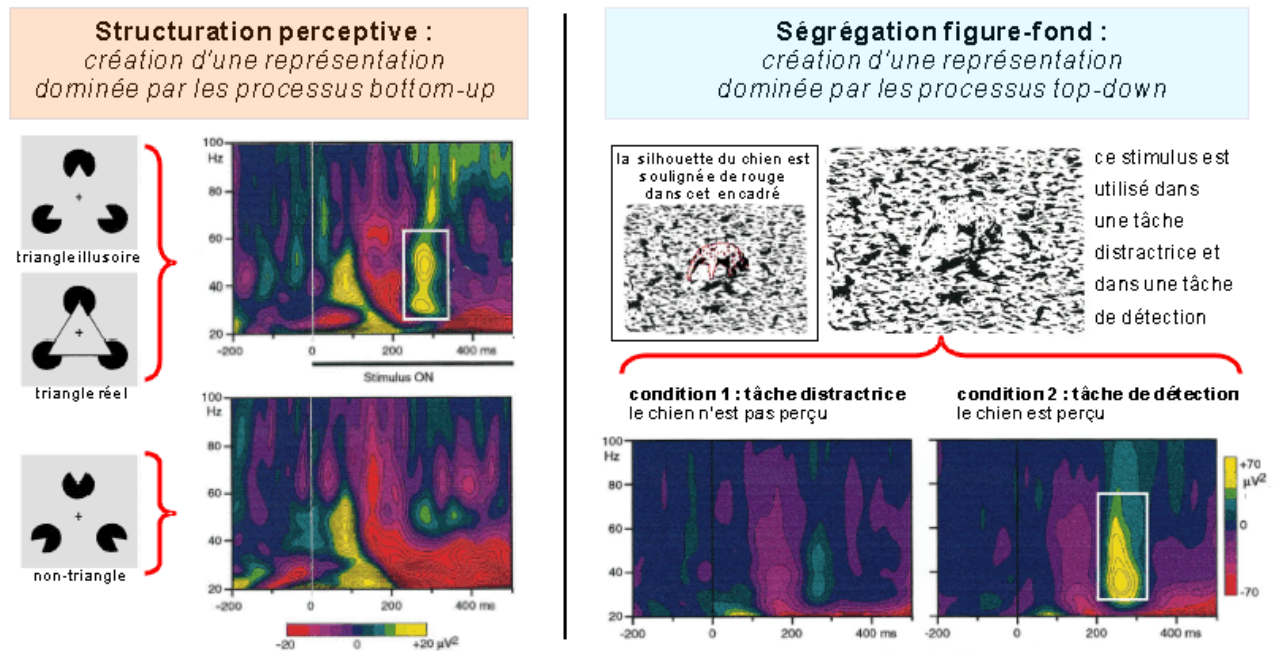


figure 24 : les deux expériences présentées par Tallon-Baudry, Bertrand, Delpuech, Pernier (1997, 1996). Lors de la présentation de triangles de Kaniza (à gauche), les figures donnant lieu à une structuration perceptive (triangle illusoire, ainsi que triangle réel), provoquent une augmentation de la densité spectrale gamma aux alentours de 250 millisecondes après l'apparition de la cible, au contraire de la figure non-triangle. Lors de la présentation de l'image camouflée du dalmatien (à droite), dans deux conditions, naïve et entraînée, on observe une augmentation des oscillations gamma seulement lorsque le chien est perçu (entraîné). Ces résultats suggèrent que la création d'une représentation des objets s'accompagne d'une augmentation des oscillations gamma, que le traitement soit dominé par des processus bottom-up ou par des processus top-down. (adapté de Tallon-Baudry, Bertrand & Fischer, 2001).

D'autres équipes ont échoué à répliquer ou à transposer ces oscillations gamma induites dans des tâches similaires (eg. Herrmann, Mecklinger, Pfeifer, 1999). Il n'est pas impossible que la méthodologie soit à l'origine de ces résultats disparates. La procédure est rendue aléatoire par le choix de nombreux paramètres empiriques à fixer de manière relativement arbitraire (choix de la durée des segments EEG utilisés, filtrage, algorithmes d'analyse temps-fréquence...). Il semble nécessaire d'attendre davantage de réplifications pour éclaircir ces premiers résultats.

Analyses de cohérence

L'analyse de cohérence est surtout utilisée dans le but de mettre en évidence des interactions neuronales à grande distance. D'ailleurs, l'observation d'un couplage en phase entre deux sites enregistrés à moins de deux centimètres de distance n'est pas une preuve absolue d'interactions neuronales, car le signal ayant sa source sous une électrode peut passivement se propager dans le volume de conduction vers la deuxième électrode. Bressler, (1996) et

Lachaux, Rodriguez, Martinerie & Varela, (1999) distinguent la **synchronisation à petite échelle, impliquée dans le liage perceptuel (*binding*)** de la **synchronisation à grande échelle, impliquée dans l'intégration globale**. Cette distinction s'inscrit dans la gradation présentée aux figures 9-11 pp. 21-24. Au niveau le plus local, l'intégration peut consister à assembler les activités des neurones codant pour les segments orientés, pour ainsi produire un contour. A un niveau supérieur, il peut s'agir d'assembler le contour, la texture, la couleur et la position. A un niveau encore supérieur, il peut s'agir d'assembler la représentation visuelle de l'objet avec les traitements réalisés dans d'autres modalités sensorielles, avec des traitements top-down (consignes pour la tâche, stratégie cognitive, motivation, le tout prédominant dans les régions frontales) ou encore avec la préparation et l'exécution motrice. L'étendue du couplage en phase fait écho à cette gradation de l'intégration cognitive.

Une étude par Bressler, Coppola & Nakamura (1993) démontre des **synchronisations à grande échelle, induites par une tâche de discrimination visuelle (go no-go)** chez le singe. La synchronisation se manifeste **dans plusieurs bandes fréquentielles**. Elle se produit entre les électrodes situées au dessus du cortex strié et au dessus des électrodes pariétales d'une part, et entre les électrodes situées au dessus du cortex strié et au dessus du cortex moteur d'autre part. Dans cette étude, la synchronisation ne s'accompagne pas d'oscillations.

Une étude réalisée par Rodriguez, George, Lachaux, Martinerie, Renault & Varela (1999) chez l'humain montre des épisodes de **fort couplage inter-électrode, lors de l'interprétation d'une figure ambiguë**. Les stimuli étaient des « Mooney faces » (voir figure 25). Les figures présentées dans le sens correct sont perçues comme des visages. Elles occasionnent une plus forte densité gamma que les figures présentées à l'envers (pour lesquelles le sujet naïf ne perçoit usuellement pas qu'il s'agit d'un visage).

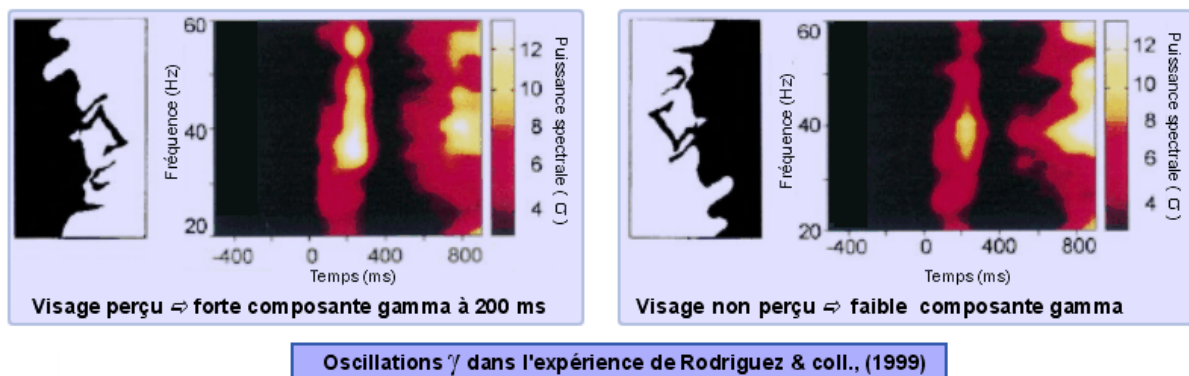
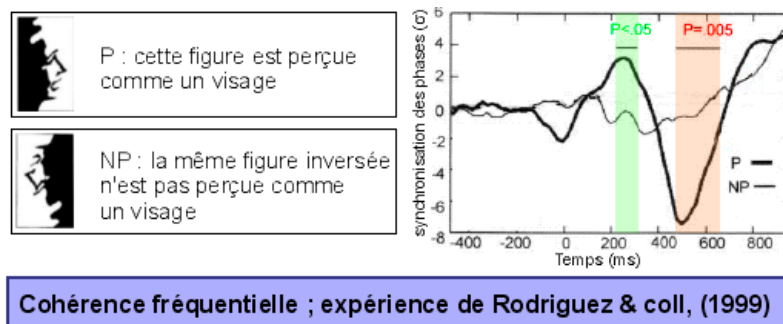


figure 25 : dans l'expérience de Rodriguez & coll, (1999), la présentation dans le sens correct des « Mooney face » s'accompagne de davantage d'oscillations gamma que la présentation du même stimulus présenté à l'envers, à 200 millisecondes après l'apparition du stimulus. Les caractéristiques perceptives ne peuvent pas expliquer ces différences, puisque la figure elle-même est son propre contrôle.

Les auteurs ont également présenté l'évolution temporelle de l'indice de cohérence. L'indice de cohérence permet de différencier les figures perçues des figures non-perçues au cours de deux périodes (voir figure 26, périodes surlignées en vert et en rouge, respectivement). Vers 200 millisecondes, les figures perçues occasionnent une plus forte cohérence inter-électrode que les figures non perçues. A l'inverse, vers 450 millisecondes, les figures perçues sont davantage désynchronisées que les figures non-perçues.



Cohérence fréquentielle ; expérience de Rodriguez & coll, (1999)

figure 26 : évolution temporelle de l'indice de cohérence dans la bande gamma, dans l'expérience de Rodriguez & coll (1999)

Cette évolution de la synchronisation et de la désynchronisation peut être plus précisément observée à la figure 27. Les figures perçues comme des visages s'accompagnent de deux périodes de synchronisation (vers 200-260 millisecondes et vers 645 millisecondes), séparées par un fort découplage (vers 500 millisecondes). Ces données sont interprétées comme la mise en place d'une **première assemblée neuronale, liée à la création de la représentation visuelle** du visage, puis sa **dissolution active**, qui précède la mise en place d'une **seconde assemblée, dédiée à la représentation du comportement moteur** à exécuter.

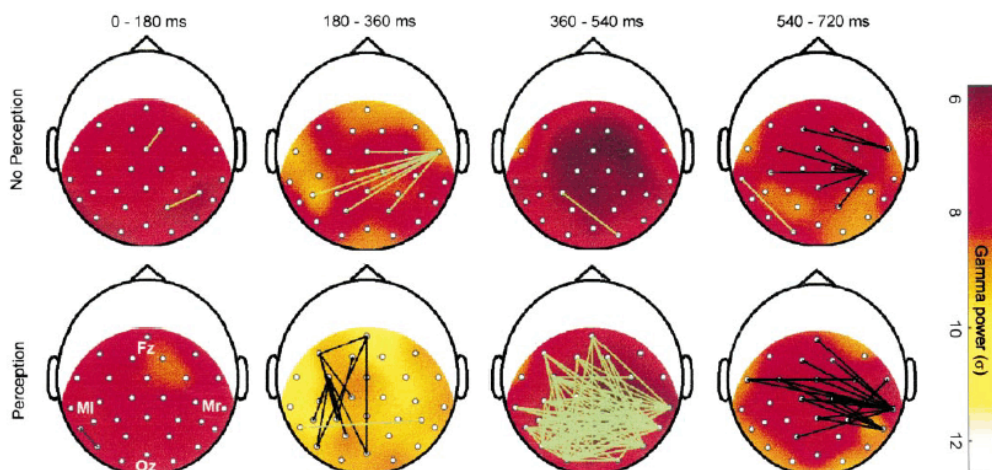


figure 27 : évolution des oscillations gamma, et de leur couplage entre électrodes, dans l'expérience de Rodriguez & coll., (1999) pour les figures non perçues comme des visages (en haut) et pour les figures perçues comme des visages (en bas). Les traits noirs représentent une augmentation du couplage, les traits verts une diminution du couplage, voire une désynchronisation active.

Une autre expérience en électrophysiologie du traitement visuel est présentée par Gruber, Keil & Müller, (cité dans Keil, Gruber & Müller 2001). Les auteurs testent les **effets de l'apprentissage perceptif rapide sur les oscillations gamma**. Les images fragmentées sont présentées une première fois (voir figure 28). Le sujet est incapable de les identifier. Puis, dans 50% des essais, l'image complétée est présentée. Ensuite, l'image fragmentée est présentée une seconde fois. Le sujet est capable d'identifier les images dont il a vu la version complète, au contraire des images n'ayant pas été perceptivement apprises. L'identification de ces images fragmentées s'accompagne d'oscillations gamma à 200 millisecondes, prédominantes sur les électrodes pariétales. Ces oscillations sont davantage

synchronisées sur les électrodes postérieures. L'apprentissage perceptif des images complétées permet de rendre cohérente la seconde présentation de l'image fragmentée (les mécanismes top-down -la mémoire- suppléent à la partialité de l'image pour réaliser l'identification).

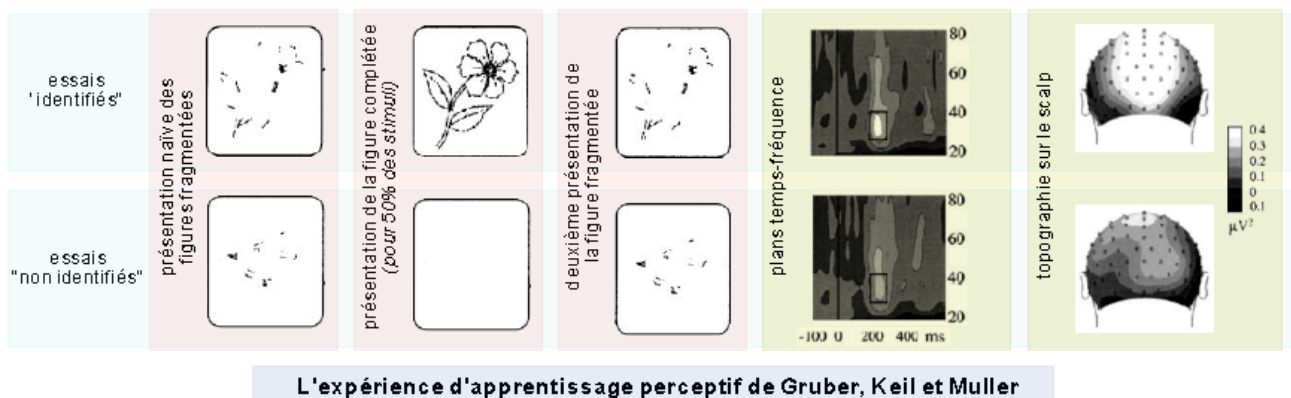


figure 28 : l'expérience d'apprentissage perceptif de Gruber, Keil & Müller (citée dans Keil, Gruber & Müller, 2001). Les auteurs ont utilisé les images fragmentées de l'inventaire de Snodgrass et Vandervort. Les images fragmentées ayant fait l'objet d'un apprentissage perceptif (en haut) donnent lieu à une augmentation de la densité des oscillations gamma, 200 millisecondes après la seconde apparition de l'image fragmentée. Ces oscillations gamma sont bien moins présentes pour les images fragmentées dont la version complète n'a jamais été vue, et qui, par conséquent, ne seront pas identifiées par le sujet (en bas).

Gruber & Müller (2002) montrent que la présentation répétée de dessins au trait provoque une diminution de la densité spectrale gamma, une moindre cohérence entre des électrodes distantes et une modification de la topographie des composantes gamma sur le scalp. Des dessins au trait d'objet sont présentés une seule fois, ou en deux ou trois présentations consécutives. Cette « **habituaton neuronale** » fait écho à des hypothèses neurophysiologiques, suggérant l'existence d'un mécanisme de sauvegarde au sein d'une assemblée représentant un objet. La représentation répétée de l'objet sollicite des assemblées de plus en plus sélectives, et impliquant de moins en moins de neurones.

Lien avec l'attention

De nombreuses études réalisées à différentes échelles ont mis en évidence une **augmentation de la synchronisation lorsque l'attention est sollicitée**. Munk, Roelfsema, König, Engel & Singer (1996) montrent qu'une stimulation de la formation réticulée (structure impliquée dans la vigilance et l'attention) facilite la synchronisation des neurones de l'aire 17 du chat. Ces observations sont complétées par Heculano-Houzel, Munk, Neueschwander, & Singer (1999). Les auteurs réalisent leur étude à la fois sur les variations spontanées de l'état de vigilance, et sur les transitions d'états induits par activation de la formation réticulée. Un résultat similaire est obtenu chez le singe par Steinmetz, Roy, Fitzgerald, Hsiao, Johnson & Niebur (2000) lors de l'accomplissement de tâches visuo-motrices et par Fries, Reynolds, Rorie, & Desimone, (2001) lors de tâches visuo-attentionnelles.

Des études ont également été réalisées chez l'homme. Müller, Gruber & Keil, (2000) montrent une augmentation des oscillations gamma chez l'humain pour une stimulation attendue, relativement à la même stimulation lorsqu'elle est perçue comme un distracteur.

En outre, la synchronisation neuronale semble impliquer le **thalamus**, structure impliquée dans l'attention et la vigilance (eg. Jones, 2001 ; Steriade, 2000). Ces données suggèrent une relation entre l'attention sélective et les oscillations gamma. Ces observations font écho à la formulation théorique de Treisman, faisant de **l'attention une fonction clef dans l'intégration perceptive**.

Les prochaines étapes

Nous avons présenté quelques études supportant le lien entre la synchronisation et le « *binding problem* » en vision. Ces résultats ont été obtenus à des étapes variés du traitement perceptif (contour, surface, discrimination figure-fond, notons cependant l'absence d'études sur la couleur¹), et avec des propagations du traitement de type bottom-up, mais également top-down. Ces études réconcilient, sur le plan fonctionnel, les réserves que nous avons émises au niveau méthodologique quant à l'assimilation totale des oscillations et des synchronisations. Les répliques semblent fonctionner entre les cross-corrélations des enregistrements multi-unitaires et l'observation des oscillations gamma sur le scalp (Müller et coll., 1996). Les mêmes corrélats fonctionnels apparaissent aux différents niveaux.

Les résultats empiriques présentés ci-avant ont cependant été l'objet de contestations : des résultats ont paru manquer de netteté, des répliques ont échoué, des critiques ont été émises sur l'ampleur du phénomène, certains auteurs ont « oscillé » entre adhésion et scepticisme d'une étude à l'autre. Ce scepticisme atteint moins le niveau macroscopique des oscillations gamma (qui ont une histoire indépendante de celle des assemblées synchrones (voir Bullock & Achimovicz, 1993), et dont l'implication dans la remise en question du dogme de la neuropsychologie classique est moins directe).

Le paramétrage des analyses (choix des intervalles temporels, choix des filtres, montage des chaînes d'amplification, spline² des fonctions d'auto-corrélation, choix d'une analyse temps-fréquence stationnaire -Fourier- ou non stationnaire -ondelettes-, mise au point de méthodologies pour des tests de significativité statistique) a possiblement contribué à la confusion. Il semble nécessaire d'accorder un délai supplémentaire pour que le débat se clarifie, et il semble indispensable d'encourager une standardisation de ces méthodes, au moins pour les répliques, en attendant de bien percevoir l'impact des choix méthodologiques sur les résultats.

Concernant le débat théorique, la prochaine étape d'importance consistera sans doute à découvrir les règles de formation et de dissolution des assemblées. Cela permettra de savoir par exemple (1) quand un neurone entre et sort

¹ La couleur est une primitive visuelle à part, ayant un fort pouvoir de discrimination. Par exemple, dans un paradigme de recherche visuelle, si la couleur peut être utilisée comme critère distinctif, alors on observera un « pop-out ».

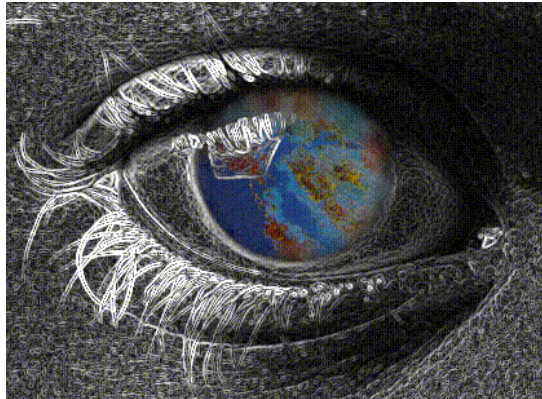
² Pour rendre plus visibles les oscillations dans les cross-corrélogrammes, ces histogrammes étaient surimposés de fonctions splines, sans indicateurs d'adéquation, mais qui faisaient la lecture des données dans le sens attendu.

d'une assemblée (2) si des éléments représentés précocement s'évanouissent avant l'aboutissement du traitement perceptif, ou si ces éléments, une fois acquis, se maintiennent jusqu'à la dissolution complète de la représentation (3) si une hiérarchie « assemblées dans l'assemblée » est observée. Pour le moment, et principalement en raison de la restriction des analyses à des méthodologies bivariées, on ne connaît rien de l'entrée en synchronisation d'un groupe de neurones.

Conclusion

Notre souhait de réaliser des études électrophysiologiques sur le scalp est rendu crédible par la disponibilité de deux méthodes, les analyses de variation de la densité spectrale (oscillations gamma) et les analyses en cohérence (qui permettent de mesurer le couplage à grande distance sur le scalp). Les résultats des études neuro-fonctionnelles obtenus antérieurement suggèrent que le lien entre synchronisation et intégration/ségrégation est très plausible. Ces résultats encouragent la poursuite de notre propre projet concernant les conjonctions illusoires en RSVP, dont la partie exploratoire fonde la présente thèse.

2. Le traitement de l'information visuelle



A l'heure actuelle, les bases théoriques du fonctionnement cognitif humain sont toutes inscrites dans les paradigmes de la neuropsychologie classique. Ce sont ces données et leurs interprétations qui vont être présentées ici, partiellement en contradiction avec le but de ce travail, et à défaut d'autres données.

La perception visuelle d'un objet commence par l'activation des récepteurs rétiniens par la scène visuelle le contenant. Elle s'achève par la création d'une représentation structurale (représentation de la forme), sémantique (représentation de la fonction) et lexicale (représentation du nom) de l'objet individualisé de son arrière-plan (Boucart, 1996).

La fonction visuelle est ici conceptualisée dans un paradigme constructiviste : la représentation visuelle qui s'impose à notre conscience ne pré-existe pas : elle résulte d'un système de contraintes (équipement neurophysiologique humain) et d'hypothèses. La continuité spatio-temporelle qui s'impose subjectivement à nous est le fruit d'une complète reconstruction, faite par le système cognitif humain, à partir des bribes d'informations prélevées de manière discrète par ses récepteurs.

Ce chapitre évoquera la nature des bribes d'informations que saisit le cerveau, et les transformations qu'elles subissent. Puis nous esquisserons une définition du concept d'« objet », abordant quelles sont les principales opérations qui conduisent à son émergence (intégration des traitements) et à sa différenciation du contexte visuel (ségrégation figure-fond). Enfin, nous tenterons de rassembler des données permettant de proposer un modèle du déroulement temporel des opérations perceptives pour l'identification d'un objet.

De la discontinuité : capture et traitement de l'information visuelle

Le signal visuel transite par un certain nombre de structures. Ce chapitre retiendra tout particulièrement les structures évoquées à la figure 29. Cette présentation du fonctionnement visuel sera basée sur la notion de **hiérarchie visuelle** (le traitement est orienté par un sens, il est de plus en plus abstrait dans les aires successives, et les aires situées en haut de la hiérarchie sont nourries par les computations réalisées par les neurones situés dans les aires plus précoces, Felleman & Van Essen, 1991). Les fondements théoriques de la hiérarchie corticale ont été discutés ci-avant (pp.17-18). Cependant, cette notion de hiérarchie est partiellement supportée par le sens des connections cortico-corticales initiées depuis V1 (ou cortex visuel primaire, ou cortex strié, voir figure 29), et par la complexification des sélectivités et des champs récepteurs des neurones. Les latences d'activation des aires visuelles montrent également une certaine hiérarchie (latences minimales), mais ce point est davantage controversé (Lamme & Roelfsema, 2000). Cette hiérarchie suggère que certains processus préalables sont nécessaires à des processus ultérieurs. Elle n'implique pas nécessairement un traitement unidirectionnel, comme le suggèrent les modèles en cascade (Boucart, 1996).

Les données théoriques présentées ci-après ont été obtenues chez le modèle animal (principalement le singe et le chat), et en investigation chez l'humain normal (imagerie) et pathologique (lésion et imagerie). Notons que le cortex visuel du singe (54%) est plus développé que le cortex visuel de l'homme (23%).

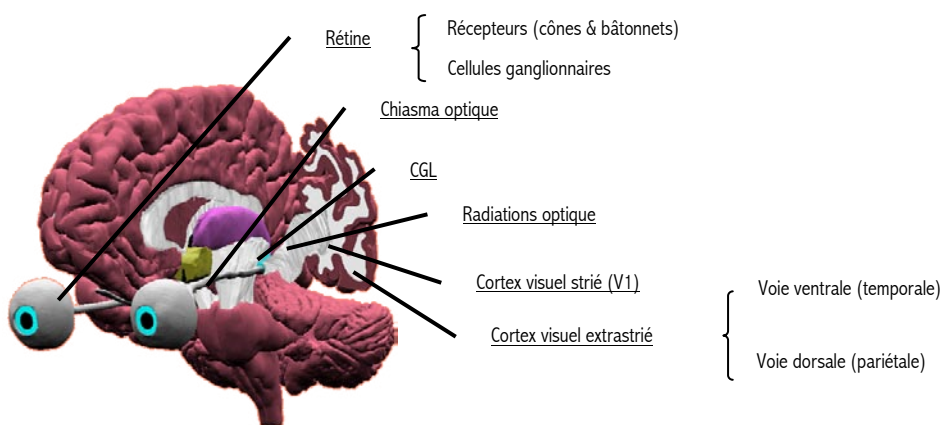


Figure 29 : structures d'intérêt dans le traitement de l'information visuelle (adapté de « Digital anatomist : interactive brain atlas », JW. Sundsten, Washington University)

la rétine et sa fovéa

« Voir » commence par une série d'ajustements (du cou, de la tête et des yeux) pour amener une zone particulière de la rétine, la fovéa, au contact de l'objet ou de la zone d'intérêt. Cette **fovéa** est la zone d'acuité maximale (meilleur pouvoir séparateur de l'œil, ou encore meilleure résolution spatiale).

Elle seule est réputée pouvoir donner un accès conscient à l'information visuelle. Au niveau des théories psychologiques, il est couramment admis que le reste de la scène est interpolé (figure 30). Le « contexte » ou « fond » est effectivement apparent dans notre perception subjective du monde visuel, mais il résulterait du maintien, en mémoire de travail, d'informations provenant de fixations antérieures. Au niveau neurophysiologique, les récepteurs

extra-fovéaux projettent bel et bien vers le cortex, au même titre que les récepteurs fovéaux. En revanche, ils bénéficient d'une représentation corticale proportionnellement moins grande que la fovéa (celle-ci ne représente que 6 degrés d'angle visuel, soit moins de 10% ; elle occupe pourtant 50% des neurones de V1). Les récepteurs extra-fovéaux présentent une résolution spatiale moins bonne (sommation spatiale des signaux). En outre, les récepteurs extra-fovéaux chromatiques sont très peu nombreux. Si la restitution de couleurs dans de grandes portions de surfaces extra-fovéales peut résulter d'un remplissage (*filling-in*) initié par les informations en provenance de récepteurs discrets, en revanche, la perception chromatique dans les détails visuels ne peut résulter que d'une interpolation.



Figure 30 : en a, les informations que capte la rétine en instantané ; en b, l'image que notre système perceptif reconstruit (reproduit de T. Vilis, www.med.uwo.ca/physiology/courses/senseweb/L1eye/l1eye.swf)

La totalité des récepteurs de la fovéa sont des récepteur de type P (ou cellules β , projetant dans la **voie parvocellulaire**). Ces récepteurs initient une voie de traitement lente, et codent des états du signal électromagnétique. Leur réponse est soutenue ou tonique. Ils déchargent aussi longtemps que le signal est présent. A contrario, la **voie magnocellulaire** code des changements d'états de ces signaux, et ce, sur une voie rapide. Sa réponse est transitoire ou phasique. Une dernière voie, mal connue, est la voie koniocellulaire.

Les récepteurs de la rétine sont des cellules sensibles à des ondes électromagnétiques de longueur d'onde de $0,4 \cdot 10^{-6}$ à $0,7 \cdot 10^{-6}$ m, correspondant respectivement à une sensation de couleur bleu à rouge. La fovéa contient exclusivement des **cônes**, alors que le reste de la rétine comporte des cônes et des **bâtonnets** en proportions variables.

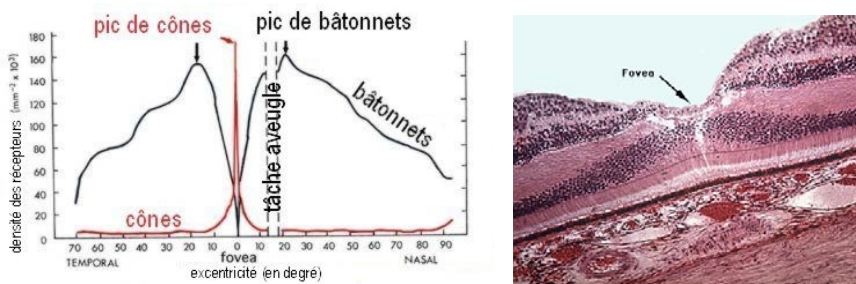


Figure 31 : à gauche, répartition des cônes et des bâtonnets en fonction de l'excentricité. A droite, coupe de la rétine chez le singe, avec au centre, la fovéa (reproduit de « Digital anatomist : interactive brain atlas », JW. Sundsten, Washington University).

Les cônes, densément implantés dans la région fovéale et périfovéale, sont utilisées en vision photopique (vision diurne), tandis que les bâtonnets, très présents en rétine périphérique, sont utilisés en vision scotopique (vision nocturne). A des niveaux d'éclairément modérés, cônes et bâtonnets fonctionnent simultanément. On distingue les cônes S, M, L, codant des **longueurs d'ondes** petites (S, « cône bleu », pic à 437nm), moyennes (M, « cône vert », pic à 533nm), et grandes (L, « cône rouge », pic à 564nm). Les cônes S sont absents de la fovéa. Une seule sorte de bâtonnets existe, qui donne une réponse maximale pour une longueur d'onde de 498 nm. Les bâtonnets sont également absents de la fovéa, et augmentent en densité vers la périphérie (voir figure 31).

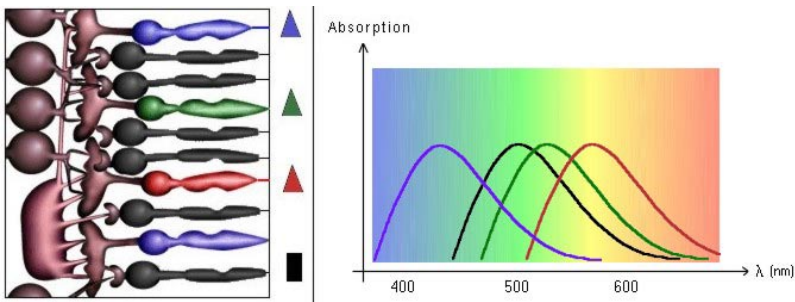


Figure 32 : Réponses des cônes (en bleu, vert et rouge) et des bâtonnets (en noir) en fonction de la longueur d'onde

Dès les cellules ganglionnaires, on observe une sommation spatiale des signaux des récepteurs, ce qui conduit à la notion de **champ récepteur** de la cellule. Ce principe fondamental des neurones visuels est défini par Hartline en 1938. Il s'agit de la portion de l'espace visuel dans laquelle une stimulation engendre une réponse d'un neurone. La taille du champ récepteur des neurones augmente le long des voies visuelles. En 1953, par des études chez le chat, Kuffler détermine un autre principe fondamental des neurones visuels. Une forte proportion d'entre eux montre un **antagonisme de réponse** entre le centre et le pourtour du champ récepteur. Par exemple, une illumination du centre provoque une activation du neurone, tandis que la même illumination appliquée au pourtour provoque une inhibition. Dans cet exemple, il s'agit d'une cellule Center-On

Les cônes prennent en charge les aspects **chromatiques** de l'image, selon un codage de type RVB (Rouge-Vert-Bleu, conforme à la conception théorique de Young-Helmoltz, voir figure 33). En revanche, dès les cellules ganglionnaires (voir figure 33), le codage se réalise sur une opposition rouge-vert ou bleu-jaune (modèle CIE Lab, figure 33), conformément à la conception de Héring (compléments dans Wandell, 2000).

structure	codage		modèle	
RECEPTEUR (cône)				
CELLULES Ganglionnaires	Voie achromatique	Voie chromatique		
		opposition rouge-vert 		opposition jaune-bleu
	← magnocellulaire →	← parvocellulaire →		← koniocellulaires →

Figure 33 : Encodage des ondes électromagnétiques dans la rétine : Traitements différentiels des récepteurs et des cellules ganglionnaires.

Le corps genouillé latéral

Après la décussation partielle des signaux dans le chiasma optique, (l'hémichamp visuel gauche des rétines gauche et droite part dans l'hémisphère droit, tandis que l'hémichamp visuel droit des deux rétines part dans l'hémisphère gauche), l'information visuelle arrive dans une structure thalamique bilatérale ; les **noyaux géniculés latéraux** (ou CGL, corps genouillés latéral, noyau géniculé latéral, *LGN*, *lateral geniculate nucleus*). Ces structures montrent en histologie une structure particulière en nappes, comme on peut le voir figure 34.

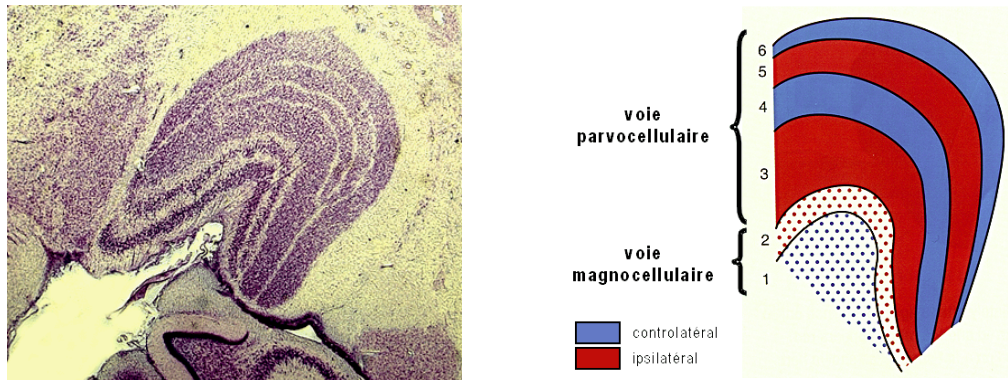


Figure 34 : le Corps Genouillé Latéral. A gauche, l'histologie montre une structure en nappes. A droite, répartition des signaux : voie magnocellulaire en pointillé, et voie parvocellulaire en pleine couleur. Les neurones de la voie koniocellulaire sont situés entre les couches. Les couches 2, 3, 5 véhiculent les signaux de la rétine ipsilatérale et les couches 1, 4, 6 véhiculent les signaux de la rétine controlatérale.

Le rôle des CGL reste mal connu. Historiquement, on leur prêtait une fonction très anodine de triage et de relais. Cependant, on constate de nombreuses projections retour du cortex vers cette structure, qui donne à soupçonner un rôle plus important qu'initialement envisagé.

Cortex Strié

L'encodage du signal visuel dans l'aire V1 (aire 17 de Brodmann, ou cortex visuel primaire ou cortex strié) devient plus complexe. De nombreuses études électrophysiologiques utilisant des stimulations psychophysiques ont été mises en œuvre, avec pour objectif de déterminer quelles dimensions du stimulus sont traitées dans V1. Il ressort de ces recherches que de nombreuses dimensions perceptives modulent l'activité de ses neurones. Ces observations ont été superposées aux connaissances de la cyto-architecture de V1.

Hormis les « quelques » neurones **koniocellulaires** qui projettent directement dans les **blobs** (inclusions de neurones en forme de colonnes), tous les signaux en provenance de CGL projettent dans les couches cytologiques 4C. Il s'agit plus exactement de la couche 4C α pour la voie **magnocellulaire** et de la couche 4C β pour la voie **parvocellulaire**. Les signaux de la couche 4C sont ensuite redistribués dans les autres couches.

V1 présente une sur-représentation de la fovéa ; environ 50% des colonnes de cellules lui sont dédiées. Son tissu conserve les relations topographiques présentes dans l'environnement. On parle de rétinitopie.

Le codage chromatique est imputé à une petite portion du cortex V1. Il concerne des inclusions réparties de manière régulière et formant des colonnes de neurones, les *blobs*, appelées également chevilles de cytochrome-oxydase (voir figure 35). Les *blobs* reçoivent directement les signaux koniocellulaires et indirectement les signaux parvocellulaire, par le biais des couches 4C.

Dans le tissu inter-blob, on distingue des **colonnes** dans lesquelles les neurones exhibent une réponse particulière : lors de tâches perceptives élémentaires, la réponse des neurones est maximale pour certaines dimensions du stimulus. Par exemple, ces colonnes sont sélectives à :

- ✧ L'orientation d'un segment présent dans le champ récepteur
- ✧ La dominance oculaire
- ✧ La fréquence spatiale d'un réseau sinusoïdal

Concernant l'**orientation**, seuls les neurones sélectifs à un angle donné, et ayant leur champ récepteur au lieu de stimulation répondent (Thibos et Levick, 1985). Les autres neurones, ayant un champ récepteur adéquat, mais une sélectivité à une autre orientation restent silencieux (voir figure 35b). Les cellules simples ont un champ récepteur allongé, qui permet une réponse pour un segment à une position particulière. Les cellules complexes agglomèrent les réponses de plusieurs cellules simples. Ils répondent à une ligne d'une orientation donnée avec une certaine invariance à la position spatiale de cette ligne. Ces colonnes de neurones sont organisées de manière radiaire autour des blobs.

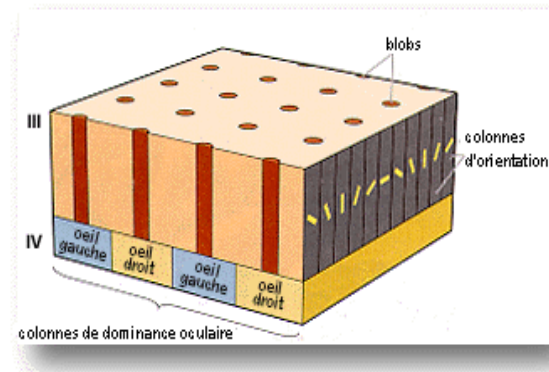
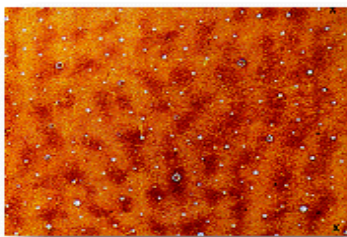


Figure 35 : Le cortex V1. En a, histologie du cortex visuel strié faisant apparaître les blobs en gris : ces inclusions sont formées de neurones sélectifs à la couleur, selon l'opposition rouge-vert ou selon l'opposition jaune-bleu. En b, schéma d'une coupe du cortex V1 : On observe les blobs et le tissu inter-blob. Sur ce dernier, on voit représenté par un segment jaune l'angle d'orientation qui occasionne la réponse maximale du neurone (colonnes d'orientations). On observe également les colonnes de dominance oculaire.

La dominance oculaire consiste en une réponse préférentielle pour un signal émis par l'un des deux yeux. Dans la couche 4, les neurones ne réagissent qu'à la stimulation en provenance d'un seul œil. Dans les autres couches, les neurones sont presque tous binoculaires, cependant ils répondent préférentiellement à une stimulation provenant de l'un des deux yeux, comparativement à la même stimulation reçue par l'autre œil. Ces neurones monoculaires et binoculaires sont rigoureusement alignés dans des colonnes perpendiculaires à la surface. On parle de colonnes de dominance oculaire. Cette spécificité de V1 fonde le traitement de la disparité binoculaire, un des indices de profondeur utilisé par le système visuel pour restituer la vision stéréoscopique.

De même, les neurones de V1 se distinguent selon leurs réponses à la stimulation par des réseaux sinusoïdaux (voir figure 36) de **fréquence spatiale** basse ou haute (DeValois & DeValois, 1988). Cette sélectivité est réputée former un matériel précurseur pour le traitement des bords et des surfaces. Les fréquences spatiales élevées portent en effet l'information sur le détail et sur les contrastes prononcés. Les fréquences spatiales basses portent l'information sur les grandes régions uniformes et sur les transitions progressives. En moyenne, les neurones parvocellulaires réagissent de manière optimale pour des réseaux de fréquence trois fois supérieures aux neurones magnocellulaires. Ces observations confortent les hypothèses selon lesquelles la voie parvocellulaire traite les détails et la voie magnocellulaire traite les aspects plus grossiers.

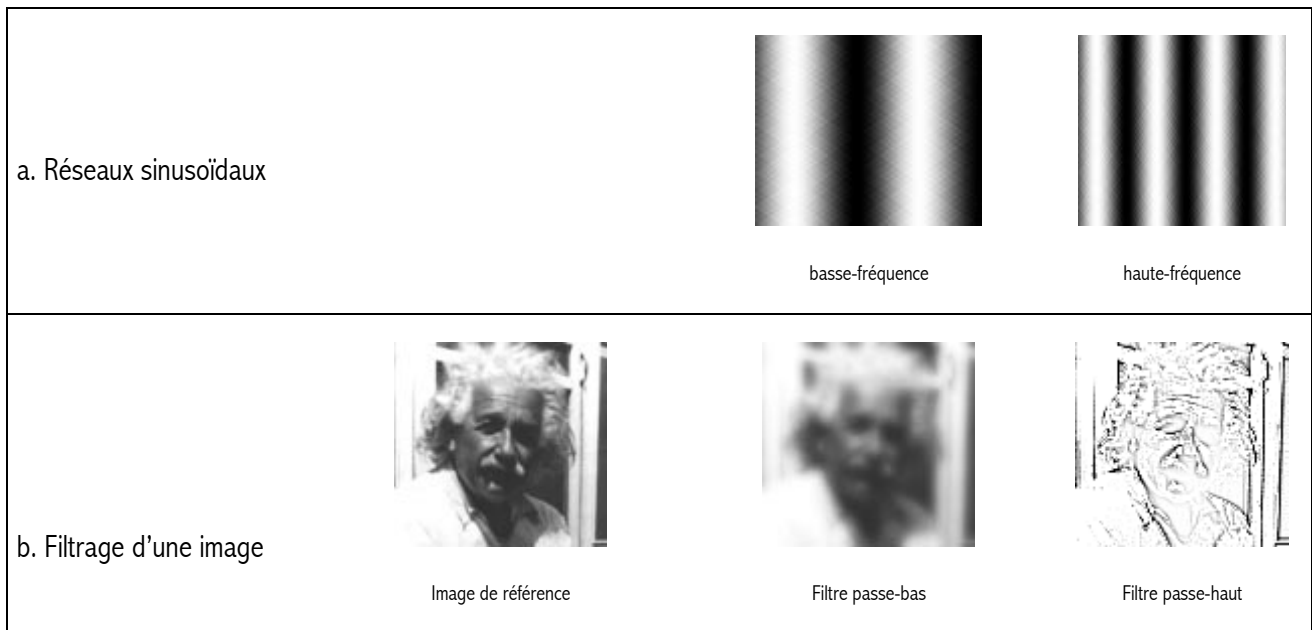


Figure 36 : Les fréquences spatiales en vision. En a, réseaux à basse et haute fréquence utilisés pour étudier la sélectivité aux fréquences spatiales élevées ou faibles. En b, image traitée avec un passe-bas (conserve les basses fréquences) et filtre passe-haut (conserve les hautes-fréquences).

Des informations complémentaires peuvent être trouvées dans l'article synthétique de Hubel et Wiesel, (1979), pour les colonnes de dominance oculaire et les colonnes d'orientation. Concernant le codage en fréquence hautes et basses, le lecteur pourra se référer à l'ouvrage de DeValois & DeValois, (1988).

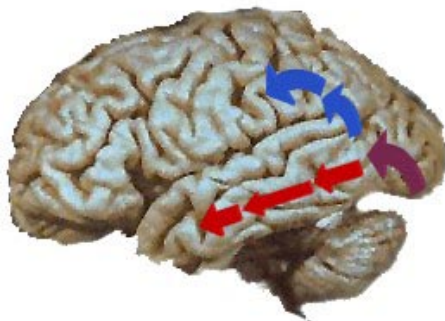
Après ces aménagements et transformations du signal dans V1, on observe des connexions sortantes de deux types. Les couches II et III de V1 projettent vers d'autres aires visuelles, à savoir, les aires 18 et 19. Les couches V et VI rétroprojettent respectivement vers les colliculi et vers les CGL.

Cortex extra-strié : de V2 aux voies dorsale & ventrale

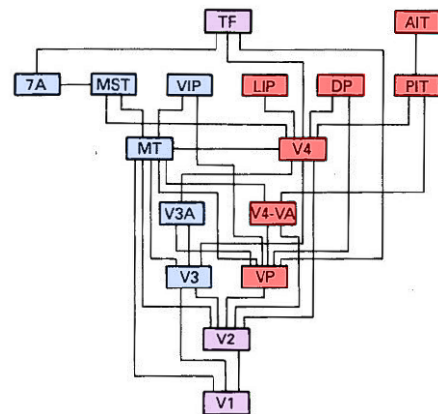
Après V2, on admet que le traitement visuel se subdivise en deux grandes voies anatomiquement distinctes (Ungerleider et Mishkin, 1982 ; Ungerleider & Haxby, 1994). Il s'agit de la voie ventrale ou voie « quoi ? » et la voie dorsale ou voie « où ? », voir figure 37. Cette ségrégation est approximative sur le plan des connexions anatomiques, mais très heuristique sur le plan explicatif ; elle est soutenue par l'existence de deux groupes de neurones visuels distincts, les neurones parvocellulaires (exclusivement voie ventrale) et magnocellulaires (surtout voie dorsale).

V2 est organisée en trois types de structures : bande épaisse, bande fine et inter-bande. Les bandes épaisses reçoivent les afférences de la voie magnocellulaire. Ces neurones projettent à leur tour dans V5 (appelée plus communément MT, l'équivalent humain de V5) et vers le lobe pariétal. Il s'agit de la **voie dorsale**, ou voie occipito-pariétale. Elle est impliquée les interactions visuo-motrices (détection du mouvement et de la position, ajustement visuo-moteur pour la préhension ...). Elle est liée aux mouvements du faisceau attentionnel et à la programmation de la saccade oculaire.

La **voie ventrale** est inscrite dans les deux autres types de structures de V2, les bandes fines et les inter-bandes. Cette voie soutient la reconnaissance des formes et couleurs ; elle est particulièrement dévolue à la vision fovéale. Les bandes fines de V2 reçoivent des afférences des blobs de V1 et projettent vers la voie ventrale, ou voie occipito-temporale. L'information supportée par ce canal semble concerner la couleur et la texture. Par ailleurs, une troisième structure, les inter-bandes, reçoit des afférences de neurones magnocellulaires et parvocellulaires des inter-blobs de V1. Ces neurones sont sensibles aux orientations et semblent être le matériel à la source de l'identification des bordures (ie. variations de l'information visuelle qui indiquent le bord de surfaces). Ces neurones projettent également vers la voie ventrale.



D'après Ungerleider et Mishkin, 1982



D'après Maunsell & Newsome, 1987

Figure 37 : Les deux voies visuelles : en rouge, la voie ventrale et en bleu, la voie dorsale. (Abréviations : VP : ventral parietal ; MT : middle-temporal ; MST : Medial-superior-temporal ; VIP : ventral intraparietal ; LIP : lateral intrapariétal ; DP : Dorsal Prelunate ; PIT : posterior infero-temporal ; AIT : anterior infero-temporal ; TF : parahippocampal area)

Nous allons maintenant examiner plus particulièrement la voie ventrale. Après V2, cette voie continue dans l'aire visuelle **V4**. Dans cette aire, les neurones sont sélectifs à la forme ou la couleur. Certaines cellules de cette aire font preuve de **constance de couleur**. En effet, la composition spectrale de la couleur d'une surface exposée à des éclairagements différents (eg. la lumière du soleil et une ampoule électrique) varie considérablement. La réponse de ces neurones reste constante lorsque les conditions d'éclairage, et donc lorsque la composition spectrale enregistrée par les récepteurs, est modifiée (voir figure 38). On dit que ce n'est pas la couleur réelle, mais la couleur subjectivement perçue de la surface, qui module la réponse de V4, contrairement à V1. Des lésions sélectives de V4 chez les singes démontrent également son implication dans la reconnaissance de **forme** (voir figure 39a) ; on observe d'ailleurs qu'une lésion de V4 dégrade fortement cette aptitude tandis que la perception de la couleur semble relativement maintenue (Schiller, 1993).

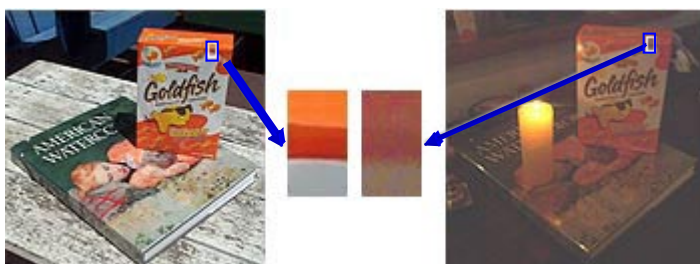


Figure 38 : La constance de couleur dans V4. On observe des changements importants de couleurs selon les conditions d'éclairage. Nous pouvons voir un grossissement de la partie haut-droite du paquet de céréales, qui montre une composition spectrale très différente selon l'éclairage. Les neurones de V4 manifestant une constance de couleur répondent de manière identique dans les deux conditions, alors que les

réponses sont très différentes dans V1.

L'information visuelle transite ensuite dans IT (inféro-temporal), également désigné par TE. Il s'agit de la dernière aire de la voie ventrale exclusivement visuelle. Elle présente une **sélectivité à des formes complexes** (Fujita, Tanaka, Ito & Cheng, 1992 ; voir figure 39b). Cette sélectivité est, à l'instar de V1, distribuée en colonnes. Des neurones proches sont sélectifs à des formes proches.

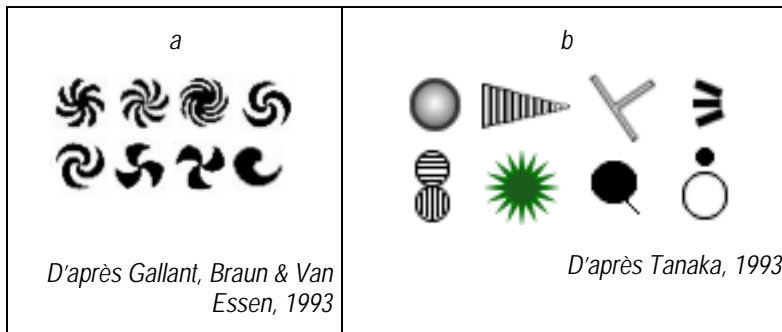


Figure 39 : exemple de formes complexes suscitant des réponses préférentielles des neurones du cortex extrastré : de V4 en a, et de IT en b.

Ces formes ont été testées parmi de nombreuses autres, et elles constituent les stimuli préférés des neurones de V4 et de IT. Les neurones qui sont soumis à ces formes dans leur champ récepteur déchargent davantage que pour les autres stimuli.

Le champ récepteur de ces cellules occupe souvent l'ensemble du champ visuel. La représentation est allocentrée¹. Dans la partie postérieure de IT, la sélectivité semble concerner des éléments simples. Vers les parties antérieures, la sélectivité concerne des instances de plus en plus complexes, par exemple des visages, des maisons, des chaises ou des mains (voir figure 40).

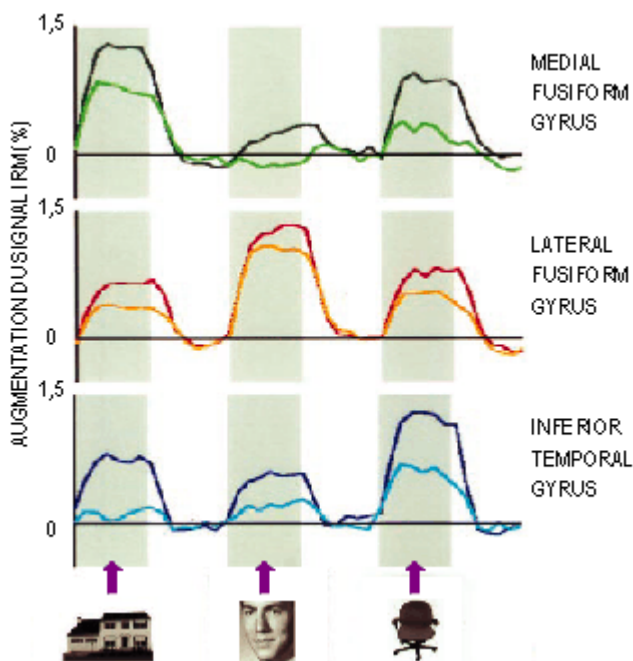


Figure 40 : aires cérébrales ayant des réponses préférentielles pour des maisons, des visages, des chaises. Les stimuli sont présentés sur les plages temporelles représentées en gris. Les stimuli contrôles sont représentés par les plages blanches. Adapté de Ishai, Ungerleider, Martin, Schouten & Haxby, (1999).

On observe que le gyrus fusiforme médian possède une prédilection pour les maisons, que le gyrus fusiforme latéral possède une prédilection pour les visages et que le gyrus temporal inférieur possède une prédilection pour des chaises. Ces trois structures cérébrales font partie de IT.

Prenons l'exemple de cellules « spécialisées » dans la reconnaissance des visages (Desimone, 1991). Ces cellules déchargent pour toute une série de stimuli variés, mais leur réponse est maximale pour la présentation d'un stimulus formé d'un visage. La reconnaissance de cette instance est indépendante de la position, de la taille, de l'angle de vue,

¹ Une représentation est égocentrée lorsqu'elle dépend du point de vue de l'observateur, et allocentrée lorsqu'elle en est indépendante.

et du type d'indice permettant de constituer la forme (contour, couleur, texture et même mouvement). La figure 41 illustre cette propriété.



Figure 41 : Les images a et b sont assez similaires : elles initient un pattern d'activation proche dans V1, mais différencié dans IT. Les images b et c représentent la même personne. Les patterns d'activation seront différents dans V1 mais similaires dans IT.

Les propriétés des cellules de IT avaient été théorisées par Barlow dans les années 1960 (neurones « cardinaux », synthèse dans Barlow, 1972), et furent découvertes peu après par Gross (1969). Leur mise en évidence, et l'observation de leurs capacités plastiques ont initié de multiples théorisations sur la reconnaissance des objets ; on parla alors de **cellules détecteurs de traits** également appelées cellules grand-mère. Dans les premières conceptions théoriques, était développée l'idée selon laquelle chaque objet serait reconnu par un neurone particulier. Il y aurait ainsi un neurone déchargeant à la vue de notre grand-mère, et un autre neurone chargé de « reconnaître » une Volkswagen jaune.

Ces théorisations ont ensuite paru incompatibles avec le nombre de neurones disponibles, et un codage plus parcimonieux a été envisagé. Les théories ont alors évolué vers une **conception connexionniste** (*sparse coding*). Un objet est codé, non pas par un neurone particulier mais par un pattern d'activation dans un groupe de neurones. Ainsi, la stricte sélectivité de ces neurones à des entités uniques (LE neurone qui code pour tel visage, celui qui code pour telle maison, pour telle main ou pour telle chaise) a été remise en question. Les neurones dont la réponse est maximale (augmentation de la fréquence de décharge) pour un stimulus donné, répondent bien souvent, dans une moindre mesure, à toute une série d'autres stimuli. A ce propos, Tanaka (1997) a pu montrer que toutes les configurations complexes et tridimensionnelles qui caractérisaient la sélectivité de certains neurones de IT pouvaient être dépouillées (ie. réduite à une configuration élémentaire de 2 ou 3 caractéristiques), et qu'alors, la réponse du neurone était également déclenchée. Logothetis (1998) conclut donc sa revue sur le sujet en décrétant qu'aucun neurone de IT ne présente une spécificité suffisante pour représenter, à lui seul, une vue d'un objet naturel.

La notion d'objet : restauration d'une continuité spatio-temporelle

Jusqu'à présent dans ce chapitre, nous avons explicité les processus qui conduisent à l'analyse du signal rétinien, processus dit « de vision de bas-niveau ». Ces processus pourraient suffire dans les conditions particulières étudiées en laboratoire, où le système visuel est soumis à des objets isolés. Dans les conditions naturelles, les objets sont nombreux dans le champ visuel, occlus, superposés. Leur énergie nous parvient incomplète et dégradée. Dans ces conditions, les descriptions précédentes semblent insuffisantes pour résoudre les ambiguïté perceptives et former la représentation cognitive des objets.

Les théories de la **vision de haut-niveau** visent à combler ce fossé entre les descriptions neurophysiologiques précédentes et les performances observées dans les situations naturelles de vision. Nous présenterons les conceptions théoriques actuelles de l'identification des objets ; les théories basées sur les vues et les théories admettant l'existence d'une description structurale. Nous aborderons ensuite l'organisation perceptive, essentielle pour l'individualisation de l'objet. Se jouent dans cette opération deux exigences antagonistes : intégrer et ségréguer. Nous évoquerons ensuite un modèle proposant l'existence de plusieurs représentations du même objet, qui cohabitent dans le système cognitif avec des finalités différentes.

Divers modèles de la reconnaissance et de l'identification des objets

Deux courants théoriques s'opposent pour l'identification des objets. La première, plutôt portée par des groupes de recherche en Intelligence Artificielle, suppose que la reconnaissance se réalise par une confrontation entre l'analyse perceptive en cours et un ensemble de vues stockées en mémoire. L'identification se réalise par un appariement entre l'image en cours de traitement et une vue stockée en mémoire. Cette description est dépendante du point de vue de l'observateur (*view-based model*), et l'appariement accepte une interpolation entre les vues pour une meilleure performance (Bülthoff & Edelman, 1992). Des informations complémentaires peuvent être trouvées dans (Riesenhuber M., & Poggio T., 2000). En général, ces modèles sont feed-forward (*object-driven vision*). Au plan neurophysiologique, ils s'appuient sur l'existence d'un « catalogue » d'objets, codés de manière vue-dépendante, qu'ils superposent aux neurones du cortex inféro-temporal (voir pp.61-63), pour une présentation de la sélectivité des cellules de IT). Les détracteurs de cette théorie arguent sur le principe de parcimonie (le nombre de neurones nécessaire pour l'implémentation de ce principe est jugé démesuré). En outre, dans la voie dorsale se trouvent des neurones qui codent les objets de manière allocentrée (système de coordonnées centré sur l'objet), et leur existence pourrait constituer une difficulté pour cette théorie.

Le second courant théorique (« *object-centered model* ») stipule que le traitement visuel aboutit à une description en terme volumétrique, indépendante du point de vue de l'observateur, appelée génériquement **description structurale**. Par exemple, le modèle *RBC* de Biederman (1987) est un représentant princeps de ce courant théorique. Les travaux sur la rotation mentale l'étayent (Shepard & Metzler, 1971 ; Pellizzer & Georgopoulos, 1993 ; Wexler, Kosslyn & Berthoz, 1998), en ce qu'ils proposent un mécanisme d'appariement entre une description volumétrique et une vue de cet objet.

Les deux conceptions « view-based » et « structurale » sont usuellement conçues de manière exclusive. Cependant, à notre connaissance, aucune contrainte ne rend réhibitoire la coexistence des deux processus. Les processus de reconnaissance « basés sur les vues » pourraient être dédiés aux objets très **familiers** que nous reconnaissons de manière très rapide (la plasticité neuronale nécessaire à cet apprentissage est observée dans les aires visuelles de haut-niveau). Les processus basés sur une description structurale seraient réservés aux objets moins familiers. Les premiers constitueraient un niveau de traitement moins achevé, par rapport au seconds.

Quoiqu'il en soit, ces deux théories ont en commun plusieurs points. Elles admettent la nécessité de comparer l'input sensori-moteur actuel avec les **connaissances préalables** du monde visuel. Elles insistent sur la nécessité de **généralisation** pour l'accomplissement d'une identification ou d'une catégorisation. On observe que l'humain est relativement efficace pour l'identification d'exemplaires non prototypiques (portant peu de caractéristiques types de la catégorie à laquelle ils appartiennent) ou non canoniques (prises dans un angle de vue inhabituel) des images scrutées. Elles observent que la reconnaissance procède d'une **prise de décision** (également appelée inférence probabilistique en Intelligence artificielle). Cette prise de décision se réalise après une étape de structuration du matériel visuel, ou organisation perceptive.

Organisation perceptive

De l'information rétinienne sont extraites différentes composantes de bas-niveau (eg. surfaces, contours). Ces bribes d'informations seront attribuées à des instances indépendantes de niveau supérieur, les objets. Cette fonction d'**organisation perceptive** articule plusieurs opérations telles que le *grouping* (regroupement ou liage perceptif), la ségrégation figure-fond et la complétion. L'intégration des différentes composantes dans un objet procède d'opérations heuristiques (qui donnent quelques-fois lieu à des illusions). Elle est régie par un certain nombre de règles élémentaires, dont les Gestaltistes furent les illustres découvreurs.

L'**identification du contour** des surfaces est le premier pré-requis. Les limites des surfaces peuvent se signaler par différents critères. Par exemple une **discontinuité** de la luminance (en d'autres mots, une fréquence spatiale élevée) ou de la couleur, une discontinuité de la profondeur, ou de la texture sont des indices de changement de surface. Mais aucun critère pris isolément n'est décisif, comme l'illustre la figure 42. En outre, la totalité du contour n'est pas toujours apparente (occlusion de certains objets par d'autres, situés en avant-plan), et certaines parties nécessitent une reconstruction mentale (complétion).

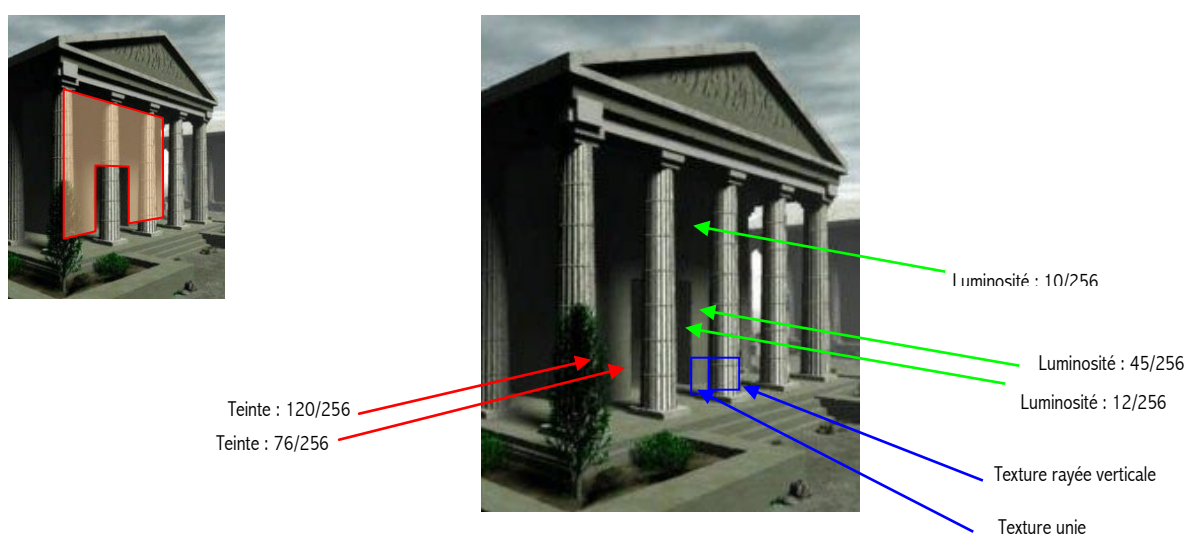


Figure 42 : perception de la surface formée par le mur d'entrée du temple (en rouge dans la miniature de gauche). La discontinuité de luminance indique des changements de surface, mais pas de manière univoque. D'autres indices, comme la couleur, la texture ou la profondeur permettent également d'estimer le contour. Le segment masqué par la colonne à gauche est inféré par des règles géométriques (les données quantitatives sont issues du modèle colorimétrique TSL, Teinte-Saturation-Luminance, voir annexe X)

Au niveau neurophysiologique, il a été suggéré (eg. Eckhorn, Bauer, Jordan, Borsch, Kruse, Munk & Reitboeck, 1988 ; Eckhorn, 1994 ; Gray, König, Engel & Singer, 1989 ; Varela, 1995, voir également chapitre 1) que les neurones codant pour des segments appartenant au même contour (par exemple dans V1) entrent en synchronisation (synchronisation dirigée de manière top-down) lorsque le système visuel a inféré leur appartenance à un contour unique.

Les discontinuités étant repérées, il s'agit ensuite de leur donner une signification dimensionnelle. Les configurations perceptives n'ont en général pas d'interprétation unique, et des règles permettent d'identifier les relations spatiales entre les **surfaces délimitées par un contour commun**. Des recherches semblent indiquer que des portions particulières du contour, les points de courbures, jouent un rôle important dans ces processus (pour une présentation plus détaillée, voir Boucart, 1996). Les points de courbures sur le contour signalent fréquemment des **jonctions** entre deux surfaces contiguës (figure 43). Elles permettent ainsi l'inférence sur les relations spatiales entre les surfaces.

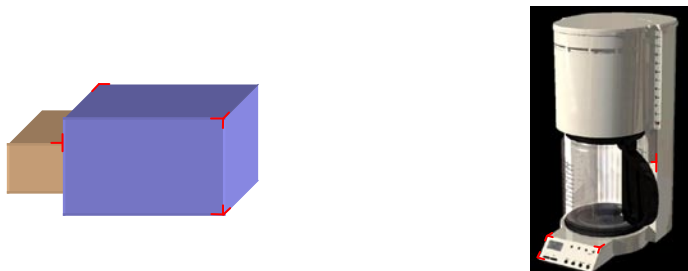


Figure 43 : l'identification des surfaces par les jonctions en L en T, en flèche et en Y (en rouge)

Les surfaces ainsi délimitées peuvent être assimilées aux modèles « fil de fer » des graphismes en 3D par ordinateur. Elles sont associées avec les couleurs et textures déduites des processus de bas-niveau (*filling-in*), par un mécanisme de propagation locale. Si des portions de l'espace sont absentes (eg. portion occlue par un objet en avant-plan, ou portion située dans la tâche aveugle, et donc, non codée par la rétine), elles se trouvent alors **complétées** de la couleur, de la texture adéquate (Pessoa & Neumann, 1998). Notre expérience perceptive est ainsi rendue à une certaine continuité spatiale qui ne pré-existe pas dans les étapes précoces.

Au terme de ces processus, est réalisé l'assemblage des contours locaux, qui fondent l'individualisation de l'objet. Notons que chez la plupart des auteurs, ces processus réalisent la **ségrégation figure-fond**, qui précède la reconnaissance des objets (eg. Vecera & O'Reilly, 1998). Cependant, des opinions divergentes (Peterson, 1994) stipulent que c'est la reconnaissance de l'objet qui préside à la ségrégation figure-fond. Peut-être faut-il considérer là un fonctionnement interactif, aucun argument décisif ne donnant une préséance temporelle à l'une des deux opérations. Notons parallèlement que les théories récentes de la perception visuelle mettent l'accent sur l'interaction entre les flux montants (bottom-up) et descendants (top-down) (eg. Bullier, 1998, 2001 ; Lamme & Roelfsema, 2000).

Bases neurophysiologiques de l'organisation perceptive

Jusqu'il y a peu, un consensus admettait que l'organisation perceptive impliquait des **aires visuelles situées haut dans la hiérarchie du traitement visuel**, et recevant des influences d'aires encore plus intégrées (eg. influences frontales sur les aires visuelles temporales). Des propositions théoriques récentes (Bullier 2001^{a, b}; Lamme & Roelfsema, 2000) insistent également sur l'importance des **connexions latérales et feed-back**¹ dans les aires visuelles précoces. Lamme et Roelfsema décrivent deux étapes dans l'activité des neurones visuels. La première partie du signal constitue le flux ascendant rapide, contenant peu de potentiels d'actions (conforme aux contraintes théorique posées par Thorpe, « une décharge par pas de traitement », voir p.20). Les contraintes temporelles de ce premier signal sont telles que les influences latérales et feed-back sont impossibles. Après ce premier flux, l'activité des neurones ne cesse pas immédiatement, elle se poursuit avec **l'incorporation des modulations latérales et top-down**. Lamme et Roelfsema défendent l'idée que la deuxième partie du signal du neurone diffère en nature de la première partie. Ces auteurs mentionnent des changements dynamiques de sélectivité des neurones de V1 au cours d'une réponse à une stimulation. Ils décrivent également les influences contextuelles affectant le signal ; des stimuli situés hors du champ récepteur du neurone affectent la réponse de ce neurone. Ces modulations feed-back et latérales corrént avec des manifestations psychophysiques (eg. constance de couleur et regroupement perceptif). L'implémentation des phénomènes psychophysiques décrits au paragraphe précédent dépendrait donc de traitements réalisés en haut-niveau, mais également de l'interaction avec les aires visuelles de bas-niveau.

Différents types de représentation

Le concept de **représentation** est l'une des boites noires des Sciences Cognitives. Nous le définissons comme un **construit théorique** visant à décrire les produits temporaires ou finaux du traitement de l'information. Ces « objets » n'existent pas en soit, ils sont des construits théoriques des chercheurs visant à rendre visibles les computations réalisées par le système cérébral. A l'instar d'un appareil photographique sur une scène en perpétuel mouvement, la représentation **fixe** un état du traitement en perpétuelle évolution. C'est avec cette même position épistémologique que nous allons déterminer les différents états représentationnels de l'information visuelle.

Il est relativement bien admis que la distinction entre voie ventrale et dorsale (voir pp.59-60) est heuristique. Nous pouvons donc légitimement dichotomiser deux représentations différentes ; **une représentation « consciente »** de l'information, supportée par la voie ventrale et **une représentation « inconsciente »** de l'information destinée à l'action sur cet objet. Des études expérimentales légitiment cette distinction ; on observe que certaines illusions visuelles

¹ Par exemple, Bullier (2001^b) rapporte une expérience dans laquelle une double stimulation magnétique transcrânienne (TMS) est appliquée à des volontaires sains. Une stimulation de MT provoque usuellement la perception de phosphènes mobiles dans le champ visuel. L'expérience manipule l'intervalle s'écoulant entre cette stimulation de MT et une autre stimulation appliquée à V1 pour inhiber temporairement son fonctionnement. Les résultats inattendus de cette étude montrent que la perception de phosphènes en mouvement n'est pas perturbée par la stimulation antérograde de V1. En revanche, une stimulation de V1, succédant de 5 à 40 ms la stimulation de MT perturbe sévèrement la perception des phosphènes en mouvement. L'auteur suggère que les aires visuelles précoces V1 et V2 se comportent comme des registres actifs, qui intègrent les résultats des computations réalisées à des niveaux hiérarchiques supérieurs par le biais des connexions feedback.

affectent notre jugement sur l'objet, mais pas le comportement de préhension qui est dirigé vers cet objet (eg. Milner & Goodale, 1995 ; Coello, Richaud, Magne & Rosetti, *in press*).

Nous trouvons dans la littérature des modèles qui entrent davantage dans le détail des computations réalisées par le système visuel. Par exemple, Treisman & Kanwisher (1998) distinguent **6 différents types de représentations d'objet** :

- ✧ Objet individualisé (*Object token*) : représentation consciente, dépendant du point de vue actuel de l'objet
- ✧ Description structurale : représentation non-consciente, centrée sur l'objet, permettant de prédire les angles et les distances
- ✧ Type d'objet (*object type*) : représentation du type ou de la catégorie d'appartenance de l'objet
- ✧ Représentation associative : connaissances de ce qu'on peut faire avec l'objet
- ✧ Représentation de la valence émotionnelle et motivationnelle de l'objet
- ✧ Description centrée sur l'action : spécification nécessaires à une réponse motrice en rapport avec l'objet

Certaines de ces représentations sont purement perceptives, d'autres associent des contenus actuels et des informations stockées en mémoire. Ce type de modèle constitue un résumé fidèle des différentes opérations réalisées par le système cognitif. On regrette cependant qu'il y manque la dimension temporelle. Dans quel ordre, et sous quelles latences ces différentes étapes sont-elles achevées ?

Le déroulement des opérations perceptives pour l'identification de l'objet

Tout d'abord, nous tenterons de spécifier le **temps nécessaire pour identifier un objet**. Pour ce faire, les difficultés ne seront pas de trouver des données temporelles, mais de savoir quels processus sont englobés dans le concept d'identification, et à quel stade cette opération est achevée.

Puis, à un niveau de détail supérieur nous donnerons des indications sur le **déroulement temporel des processus** qui mènent à cette identification. Ces données sont intimement liées aux modèles fonctionnels des aires visuelles.

Identifier un objet : combien de temps ?

Jusqu'à il y a une ou deux décennies, la **psychologie** était essentiellement **phénoménologique**. Les questions théoriques se posaient en terme de rendu comportemental des opérations réalisées par la « boîte noire ». Concernant le temps pour percevoir un objet, la question était posée en ces termes : « quel est le temps nécessaire pour réaliser un comportement λ qui dénote que la perception dudit objet a été accomplie ? ». A l'époque, on n'imaginait pas une quelconque utilité à dissocier un traitement cognitif de la séquence S-O-R qui l'incluait.

Depuis, les recherches sont devenues **processuelles**. Le psychologue cherche à comprendre, dans une échelle temporelle drastiquement plus courte, les différentes opérations cognitives réalisées et leurs assignations temporelles. La question du temps d'identification se pose désormais différemment. On cherche désormais à savoir au bout de

combien de temps les neurones visuels permettent d'accéder à un **codage complet de l'objet**, c'est-à-dire à partir de quand les autres traitements cognitifs peuvent utiliser ce produit cognitif pour réaliser leur propres computations.

Basées sur cette différence de perspective historique, les conceptions psychologiques classiques de la perception visuelle indiquaient un temps considérable pour l'accès à l'information d'objet (souvent supérieur à ½ seconde). Les recherches réalisées ces dix dernières années, combinant des aspects comportementaux et neurophysiologiques (temps de réactions, données théoriques sur le traitement neuronal et potentiels évoqués) challengent cette hypothèse.

Le canevas de ces recherches récentes consiste à administrer une tâche perceptive élémentaire (eg. catégorisation ou identification), et à lui appliquer un modèle additif des opérations cognitives afin de réaliser la soustraction des temps occupés par les opérations non perceptives (eg. le temps occupé par la réponse motrice, ou le temps nécessaire pour amener le produit des computations à la conscience). Le résultat de ces expérimentations est une mesure de la perception pure (*covert perception*, par analogie à *covert attention*). Dans cette perspective, la stricte additivité des opérations est une hypothèse faible¹. Sa plausibilité est garantie par la convergence des résultats avec des informations neurophysiologiques (latences de signal en potentiel évoqué ou en enregistrement unitaire chez le singe, modèles théoriques des délais de transmission du signal neuronal).

Le premier problème définitoire que posent ces recherches consiste à définir **à partir de quand nous considérons que l'identification est achevée**. Nous proposons deux critères pour cette définition :

- ✧ Particularisation : l'activité neurophysiologique a permis d'individualiser cet objet des autres interprétations possibles de l'image rétinienne
- ✧ Disponibilité cognitive : le produit de ce traitement est disponible pour utilisation par d'autres processus ultérieurs (assignation d'une valence émotionnelle, mise en œuvre d'une réponse motrice spécifique...)

A lui seul, le critère de particularisation est imparfait. Comme l'a bien démontré Van Rullen² (2000), le traitement neuronal de deux catégories peut être distinctif dès la rétine, mais cela ne signifie pas que la catégorisation soit effective à ce stade. Des raffinements expérimentaux permettent d'améliorer l'estimation (Van Rullen & Thorpe, 2001), mais la dépendance de cette mesure à la catégorie choisie et aux conditions expérimentales reste manifeste.

A notre connaissance, le critère de disponibilité n'a pu être opérationnalisé, aussi, les hypothèses actuelles sur le temps d'identification ou de catégorisation s'appuient, faute de mieux, sur le critère de **particularisation**. Hypothèse vulnérable, ce critère est généralement assisté par des modèles théoriques du traitement visuel, couplées aux données

¹ Les théories actuelles du fonctionnement neurophysiologique conçoivent plus volontiers que le traitement se réalise en cascade, les étapes ultérieures s'initiant avant même que les étapes antérieures soient achevées. Pour une discussion approfondie des limitations des modèles additifs, voir (Sanders, 1997).

² Van Rullen décrit des différences intrinsèques de caractéristiques visuelles (eg. répartition du contraste), qui existent globalement entre deux catégories (ex. la véhicules et les animaux). Ces différences engendrent des profils d'activité neuronale distincts dès la rétine, mais cela ne signifie pas que la rétine « sait » déjà qu'il s'agit de telle ou telle catégorie d'images. Le pattern reflète simplement une différence préexistante dans l'environnement, qui n'a pas encore été cognitivement interprétée.

concernant les latences d'activation des neurones de IT, neurones dont nous avons démontré l'implication critique dans la reconnaissance d'objets (voir pp.60-62). Ces données issues de plusieurs cadres disciplinaires sont suffisamment convergentes pour qu'un crédit sérieux puissent être alloué aux résultats de ces recherches.

Nous avons choisi en première intention de présenter les données issues des travaux extensifs de Fabre-Thorpe, Thorpe & collaborateurs, travaux qui concernent la **catégorisation visuelle ultra-rapide** de scènes visuelles contenant ou non un animal (Thorpe, Fize & Marlot, 1996 ; Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001). La tâche réalisée est un go no-go (cet output a minima réduit - et facilite l'extraction - des composantes motrices). Ces données initialement en Temps de Réaction (TR) présentent l'avantage d'associer plusieurs des méthodes électrophysiologiques décrites ci-avant, et d'être complétées par un paradigme comparatif chez le singe (paradigme offrant la possibilité d'enregistrements unitaires).

Catégoriser n'est pas tout à fait **identifier**, aussi se pose la question de la légitimité de ce modèle temporel pour notre problématique. Nous soutenons la validité de ce modèle, à l'aide de théorisations assimilant l'identification à une catégorisation en classe fine (Riesenhuber & Poggio, 2000). En effet, l'identification et la restitution du nom d'un objet (prenons l'exemple d'un chien) n'est qu'une catégorisation parmi d'autres, ayant elle-aussi des classes supra-ordonnées (eg. un chien est un mammifère) et des classes subordonnées (eg. ce dalmatien, ou, plus bas encore, ce dalmatien particulier qui s'appelle Philomène). Riesenhuber et Poggio présentent les tâches de catégorisation et d'identification comme deux points différents sur le continuum de la **généralisation** perceptive¹. Après ces auteurs, nous soutenons donc le fait que catégorisation et identification ne sont pas différentes qualitativement. Cependant, il reste un doute sur le fait que ces opérations soient quantitativement similaires. Il est probable que chez l'humain, la catégorisation d'objet visuel précède l'identification, cependant, les différences de latences sous-jacentes à ces deux opérations sont sans doute négligeables devant l'incertitude qui persiste intrinsèquement pour la définition de la latence de catégorisation (ce problème est plus largement abordé ci-après).

Les recherches de l'équipe de Fabre-Thorpe et Thorpe indiquent comme délai minimal de catégorisation un temps de **150 millisecondes** (Thorpe, Fize et Marlot, 1996, voir également pp.19-20). Une revue de question complète de ce problème est disponible dans Van Rullen (2000). C'est au bout de ce délai que s'établit le temps de discrimination en TR. C'est à ce délai que les potentiels évoqués se distinguent sans artefacts. C'est également à cette latence que les neurones de IT déchargent.

Ce délai minimal étant posé, il reste une certaine incertitude sur le délai moyen, se traduisant par une imprécision de plusieurs dizaines de millisecondes entre les estimateurs. Cette incertitude peut sembler minime quand on considère le gain en plausibilité de ces recherches comparées à leurs historiques prédécesseurs (voir introduction de ce chapitre). Cependant, rapportée à un temps de 150 millisecondes, cette incertitude constitue, dans l'absolu, une

¹ Toutes les images rétinienne sont uniques, par la combinaison de la distance, l'angle de vue, l'éclairage, et par la nature même de l'objet perçu (par exemple, une tasse peut être rouge, ou bleue, avec une anse carrée ou ronde...). La perception consiste donc à généraliser une interprétation cognitive (tasse), sur un matériel unique et non reproductible.

variabilité considérable. L'évolution à court-terme de ces recherches constituera donc un progrès décisif pour ces problématiques de spécifications temporelles de la perception visuelle.

Décours des opérations perceptives : les étapes intermédiaires

Plus que la latence globale d'identification d'un objet, ce sont les **latences des différentes étapes** du traitement perceptif qui nous intéressent. Nous avons insisté auparavant sur la difficulté d'établir une mesure globale de la perception pure (*covert perception*), c'est dire toute la difficulté de réaliser une estimation de ses sous-composantes, les processus.

Ce questionnement peut être posé au stade de la **vision bas-niveau** et au stade de la vision de haut-niveau. Selon le niveau de description envisagé, cela revient à répondre à des questions telles que : « quand la constance de couleur est-elle achevée ? » ou « au bout de combien de temps après l'apparition du stimulus est réalisée la ségrégation figure-fond ? » ou encore « quand se réalise la différenciation *type-token* (voir p.67) ? ». Pour l'heure, toute estimation sur les latences des opérations de haut-niveau semble déraisonnable. En revanche, des inférences sur les processus de bas-niveau peuvent être tentées, car :

- ❖ elles semblent avoir peu de degré de liberté par rapport aux activités neuronales des aires qui les sous-tendent
- ❖ plusieurs moyens sont disponibles pour caractériser les latences d'activation des différentes structures visuelles (tout particulièrement ERP chez l'homme et le singe et enregistrements unitaires chez le singe)

Seule contrainte impérieuse, les hypothèses anatomo-fonctionnelles doivent être fiables. Or, les données collectées ces dernières décennies sont très fournies et convergentes. Elles nous permettent de présenter **un modèle temporel des activations des aires visuelles**.

Les travaux de Schmolesky, Wang, Hanes, Thompson, Leutgeb, Schall & Leventhal, (1998) analysent les latences de décharges de neurones impliqués dans le traitement visuel. Des stimuli visuels variés étaient flashés devant les yeux de quatre singes anesthésiés. Les enregistrements ont été réalisés sur 558 neurones. Les résultats sont reportés à la figure 44. On observe que les latences de décharges sont plus courtes dans les structures magnocellulaires que dans les structures parvocellulaires. Quarante millisecondes après la présentation des stimuli, presque tous les neurones magnocellulaires des CGL ont déchargé. La plus lente des structures étudiées est V4 ; la totalité de ces neurones a déchargé au bout de 159 millisecondes. On observe également que les activités neuronales de plusieurs structures (V1, V3, MT, MST, FEF et V2) se recouvrent assez largement dans le temps.

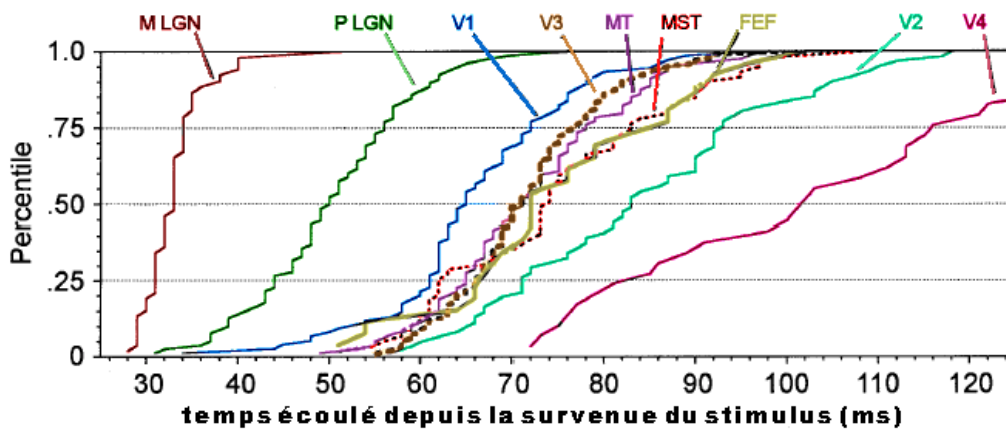


figure 44 : Distributions cumulatives des réponses évoquées visuelles chez le singe anesthésié. (reproduit de Schmolesky & coll., 1998). Près de 600 neurones ont été enregistrés chez 4 singes. (Abréviations : FEF : Frontal Eye Field)

Malheureusement, ces données ne présentent pas les latences d'activation de IT. Des données complémentaires concernant les latences de décharge de ces neurones (et d'autres) peuvent être trouvés dans les revues présentées par (Nowak & Bullier, 1997) et (Lamme & Roelfsema, 2000).

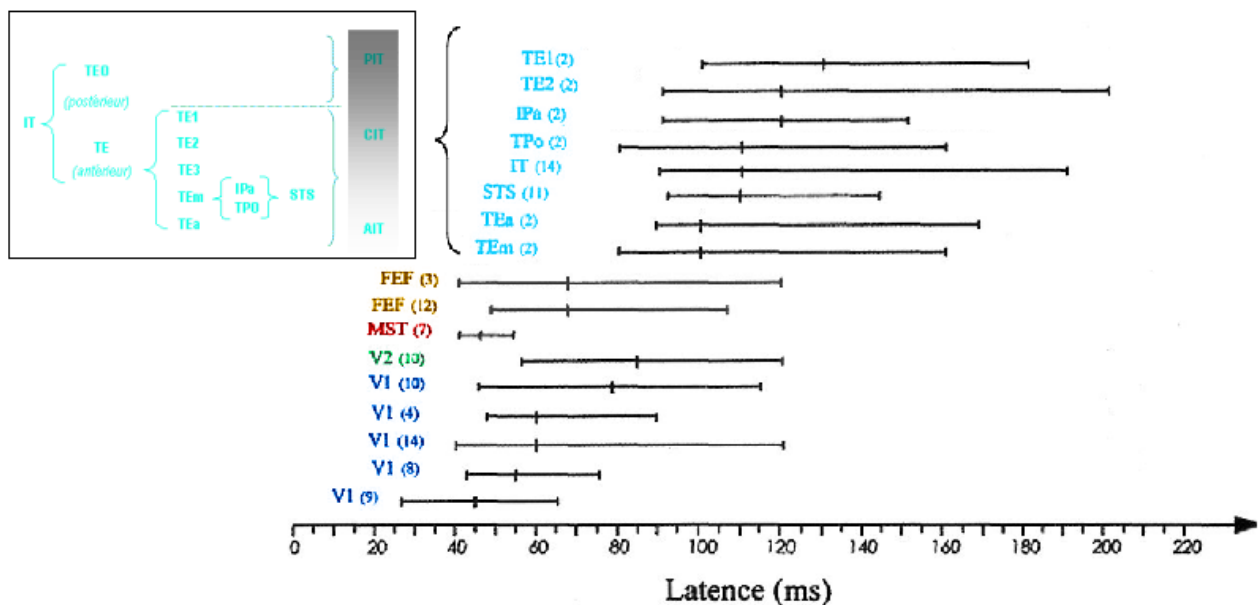


figure 45 : Enregistrements de l'activité neuronale dans différentes structures (adapté de Nowak L, Bullier J., 1997). Les barres représentent les centiles 10, 50 et 90. La plupart des études reportées ici concernaient des animaux éveillés. (Abréviations : IPS : intraparietal sulcus ; STS : superior temporal sulcus ; PIT : posterior infero-temporal ; CIT : central inferotemporal ; AIT : Anterior inferotemporal)

On observe à la figure 45 que les **neurones de IT ont une activité beaucoup plus tardive** que les neurones des structures précédemment décrites. La moitié des cellules enregistrées déchargent au bout d'environ 110 millisecondes. On observe également des disparités dans les mesures temporelles, qui s'expliquent sans doute pour partie par les différences de procédure expérimentale sous-jacentes à ces données, et pour partie par la taille et l'hétérogénéité fonctionnelle de la structure.

Résumé

Ce chapitre a présenté les bases neurophysiologiques et psychophysologiques de la perception des objets. Nous avons démontré que les différents critères visuels (eg. forme, couleur) ne sont pas traités dans des aires spécifiques (eg. aire de la couleur, aire de la forme...), mais que ces traitements sont distribués dans des clusters de neurones à plusieurs stades du traitement visuel. Des expériences récentes (Bullier, 2001^b ; Pascual-Leone & Walsh, 2001) démontrent que les théories entièrement hiérarchiques (Felleman & Van Essen, 1991) sont insuffisantes, puisque l'interruption de l'activité des aires visuelles précoces, après le passage du flux neuronal ascendant perturbe de manière inattendue la formation de la représentation cognitive. Le maintien de l'activité neuronale dans les aires visuelles précoces au moment où les aires plus centrales résolvent la représentation perceptive est nécessaire au bon déroulement de l'interprétation visuelle. Ces données suggèrent une collaboration entre les dimensions perceptives codées dans les différentes aires visuelles (eg. la forme dans V1, dans V4 et dans IT). Un mécanisme de liage étendu, physique (câblage : connexions feed-back) ou fonctionnel (eg. synchronisation), est attendu entre ces différentes activités.

Par ailleurs, un stade avancé du traitement des objets se réalise dans le cortex inféro-temporal. Il est particulièrement important de remarquer que les théories récentes du fonctionnement visuel suggèrent un codage distribué dans cette structure. Un objet est codé par un pattern d'activations dans les cellules de cette aire, et non par l'activation d'un neurone ou d'un groupe de neurones localisé. A ce niveau, un mécanisme de liage local, physique (câblage : connexions latérales) ou fonctionnel (eg. synchronisation), est requis pour faire émerger la représentation cognitive à partir des réponses plus ou moins fortes des neurones codant pour ces objets.

Ces deux mécanismes de liage étendu et local participent à l'organisation perceptive. Ils spécifient des relations intra-groupe (intégration perceptive) et inter-groupe (ségrégation perceptive) aux différents niveaux du traitement visuel.

A l'issue de ce chapitre, nous avons également spécifié les délais temporels au terme desquels ces opérations perceptives se déroulent. Les données issues de l'humain et de modèles animaux (singe) suggèrent que le traitement visuel pur est accompli en 150 millisecondes. Nous avons également souligné les larges périodes de recouvrement des activités neuronales des différentes aires visuelles. Ce modèle temporel sera précieux pour apprécier les mécanismes sous-jacents aux conjonctions illusives, qui vont être abordées ci-après.

3. Paradigmes de la recherche, mise en place du dispositif expérimental et hypothèses opérationnelles

Expériences préliminaires

Le RSVP est un paradigme expérimental au cours duquel le sujet est exposé à un flux rapide d'images (eg. 5-20 images par seconde). Il est utilisé pour l'investigation de la vision, de l'attention, du langage, et de la mémoire à court-terme, car il est à la source de différents phénomènes psychologiques, et en particulier :

- le clignement attentionnel (AB, ou *attentional blink*, Raymond, Shapiro, Arnell, 1992, prémisses dans Weichselgartner & Sperling, 1987), un déficit transitoire de la perception d'une cible lorsque celle-ci succède de moins d'une demi-seconde à une cible préalable en cours de traitement
- la cécité à la répétition (RB, ou *repetition blindness*, Kanwisher, 1987), un déficit du rappel d'une double-occurrence d'une cible
- les conjonctions illusoire (eg. Botella, Garcia, Barriopedro 1992), des erreurs d'identification d'une cible parmi des distracteurs positionnés dans le domaine temporel

Ces expériences préliminaires avaient été initialement élaborées pour étudier le **clignement attentionnel chez l'enfant**. Le clignement attentionnel est un déficit transitoire (250-750 millisecondes) du traitement perceptif lorsque le système cognitif est déjà engagé dans le traitement approfondi d'une cible visuelle préalable. Le phénomène, (Raymond, Shapiro, Arnell, 1992) est étudié lors de flux rapides d'images (autour de 10 images/s, paradigme RSVP : *rapid serial visual presentation*, voir figure 46), lors de la présentation de deux images successives à traiter perceptivement (eg. cible T1 : identification ; cible T2 : détection). Une illustration en est donnée en annexe XI.



Figure 46 : une séquence d'images présentées en RSVP. Les images se succèdent très rapidement au centre de l'écran (10 images par seconde), et le sujet accomplit des tâches perceptives (eg. double tâche d'identification-détection : « quelle image était présentée sur le fond bleu ? » et « le ballon était-il présent dans la séquence ? »).

A l'incidence d'une de ces expériences, nous avons eu l'occasion d'observer les **erreurs d'identification**, se manifestant en simple et en double tâche lorsque le sujet tente d'identifier la cible T1. L'analyse de ces erreurs d'identification a attiré notre attention, car elle ne suggère pas un mécanisme de pari (« *random guessing* », réponse aléatoire pour tenter sa chance), mais semble structurée par des régularités. Ces observations ont encouragé à poursuivre les analyses de ces erreurs. Les résultats de cette première expérimentation sont présentés dans un but

exploratoire. Ils serviront à identifier une éventuelle pertinence du phénomène au regard des problématiques de l'attention et de la vision, et ils permettront de guider l'élaboration des futurs protocoles présentés ci-après.

Expérience A : Sujets

Les sujets étaient 50 enfants âgés de 7 à 9 ans testés dans deux écoles de la région Nancéenne. Les enfants inclus dans l'étude étaient couverts par les accords de l'inspection d'académie, de la direction de l'établissement scolaire et des parents.

Quinze d'entre eux ont été inclus dans un pré-test, initialement destiné à apprécier les intervalles inter-cibles pertinents pour l'étude du clignement attentionnel. Trente-cinq d'entre eux ont réalisé une session expérimentale destinée initialement à mesurer le clignement attentionnel.

Matériel et méthodes

Le matériel utilisé pour ce test était un ordinateur Pentium celeron 300 MHz, équipé d'une carte graphique à 4Mo de mémoire et d'un écran trinitron 15 pouces de pitch 0,28. La résolution utilisée était 1024x768 et la fréquence de rafraîchissement de l'écran était de 75 Hz. Le test était développé sous Visual Basic 6 et l'administration était réalisée en prenant soin d'éliminer les tâches résidentes de Windows afin de libérer un maximum de ressources-mémoire pour la réalisation de la session. Les expérimentations avaient lieu dans des salles borgnes, sous un éclairage à incandescence modéré. Le sujet était assis à environ 53 centimètres devant l'écran de l'ordinateur. L'expérimentateur était situé sur sa droite.

Le test durait environ 2x 45 minutes (installation, consignes, et pauses comprises). Les deux sessions étaient répétées à quelques jours d'intervalles, à la même heure.

Les séquences RSVP utilisées pour ce test ont été construites de la manière suivante :

- elles étaient constituées de quinze images d'objets appartenant à l'une des quatre catégories (voir stimuli en annexe XII).
- les images étaient présentées sérialement au rythme de 14 images par seconde au niveau des instructions machine (58 millisecondes pour l'image et 16 millisecondes pour l'ISI¹). La tachistoscopie était modérément contrôlée. On peut estimer la réalisation effective de l'affichage à 8 ou 9 images par secondes.
- la séquence s'achevait par la présentation d'un masque (une mosaïque de carrés colorés reprenant la composition chromatique des images) destiné à supprimer la persistance rétinienne de la dernière cible de la séquence

¹ L'ISI (*inter-stimulus-interval*) est un écran vide qui sépare la présentation de deux stimuli, et empêche la fusion effective des images durant la transition de trames VGA (voir annexe XV figure XV.C).

- le sujet donnait sa réponse verbalement à l'expérimentateur. Celui-ci cochait la case appropriée dans un formulaire visible sur l'écran de l'ordinateur

Le test se subdivisait en 3 conditions expérimentales, réalisées lors de blocs indépendants :

- une condition en simple-tâche de détection d'une cible présente dans 50% des essais (un ballon, voir figure 46). Cette condition expérimentale n'est pas pertinente pour les présentes analyses.
- une condition en simple-tâche d'identification d'une cible présente dans tous les essais (la seule image apparaissant, non pas sur l'habituel fond gris, mais sur un carré bleu, voir figure 46).
- une condition en double-tâche d'identification et de détection

Les positions des cibles à identifier et à détecter étaient réparties comme présenté à la figure 47 pour le groupe expérimental. Pour le groupe des pré-tests, la cible à détecter (ballon) était positionnée aux rangs 1 à 10 après la cible à identifier. La position des distracteurs dans la séquence, et l'ordre des séquences dans chaque bloc étaient randomisés.

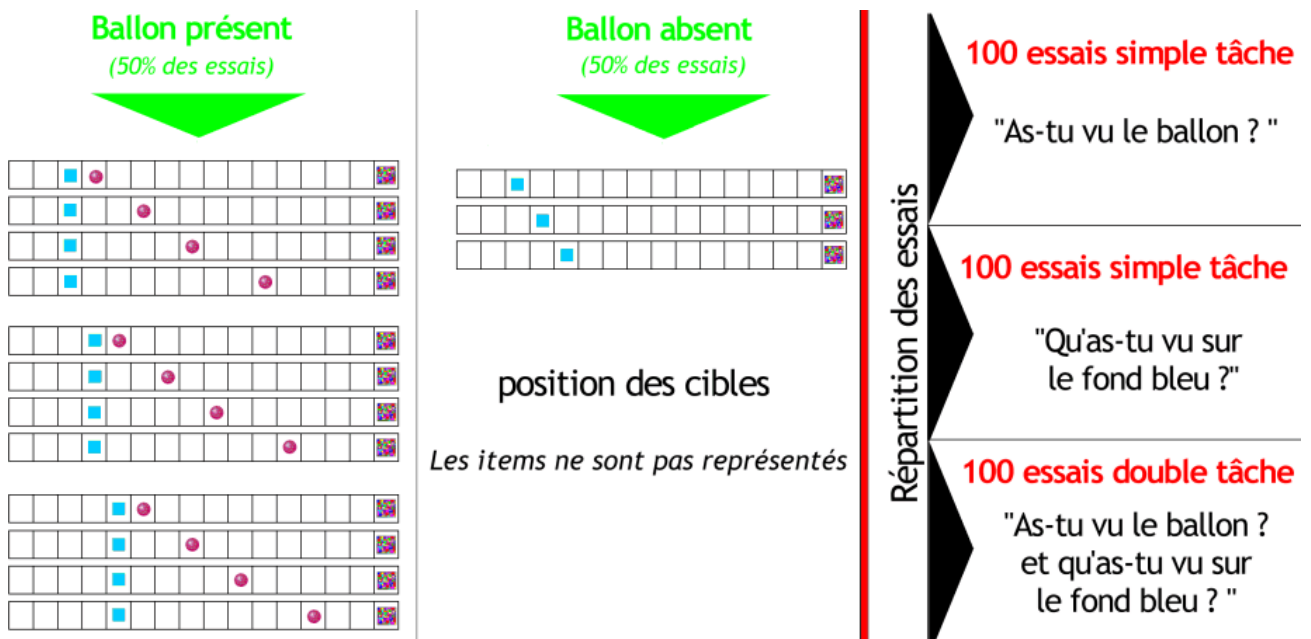


Figure 47 : positions des cibles dans l'expérimentation RSVP utilisée chez l'enfant. La cible à identifier est représentée par le carré bleu. Elle apparaît aux positions 3, 4, ou 5 à partir du début de la séquence. La cible à détecter (ballon rose) est présente dans 50% des essais. Dans ces essais, elle apparaît aléatoirement aux positions 1, 3, 5, ou 8 après la cible à identifier. La séquence s'achève par la présentation d'un masque coloré.

Résultats

Les erreurs ne sont pas aléatoires

Les séquences RSVP étaient constituées de 15 images appartenant à une catégorie donnée (4 catégories ont été utilisées alternativement ; fruits & légumes, objets, animaux et véhicules). Pour le montage de notre protocole de double-tâche, nous avons introduit la seconde cible à détecter (le ballon) dans 50% des essais. Cette cible hors-catégorie remplaçait alors l'une des quinze images de la séquence. Or, le formulaire de réponse continuait à proposer

le nom de l'image de substitution (voir figure 48). Cette particularité du protocole nous permet de tester l'hypothèse suivante : **les erreurs d'identification de la première cible (sur carré bleu) sont-elles le résultat d'un pari sur la réponse (*guessing*), où résultent-elles d'un traitement partiel du matériel visuel auquel le sujet a été exposé ?**

La prédiction dérivée de cette hypothèse est la suivante : si les réponses erronées sont des paris, il doit arriver à l'occasion que le sujet mentionne l'image du formulaire qui n'a pas été présentée dans la séquence (réponse de type 'importation'). Le sujet a statistiquement une chance sur quatorze de donner le nom de cette image lorsqu'il commet une erreur (15 images, moins l'image de la cible, qui, même si elle est mentionnée par erreur, ne peut pas être dissociée des bonnes réponses). Sur l'ensemble des réponses des sujets en identification (1000 essais pour le groupe 'pré-test' et 7000 essais pour le groupe 'expérimental'), nous avons observé 1826 essais comportant la seconde cible (le ballon à détecter) et mentionnant un distracteur. Pour satisfaire **l'hypothèse du pari**, nous devrions dénombrer parmi ces essais, une probabilité de 130 exemplaires mentionnant le distracteur manquant. Nous n'avons relevé **aucune** de ces erreurs.

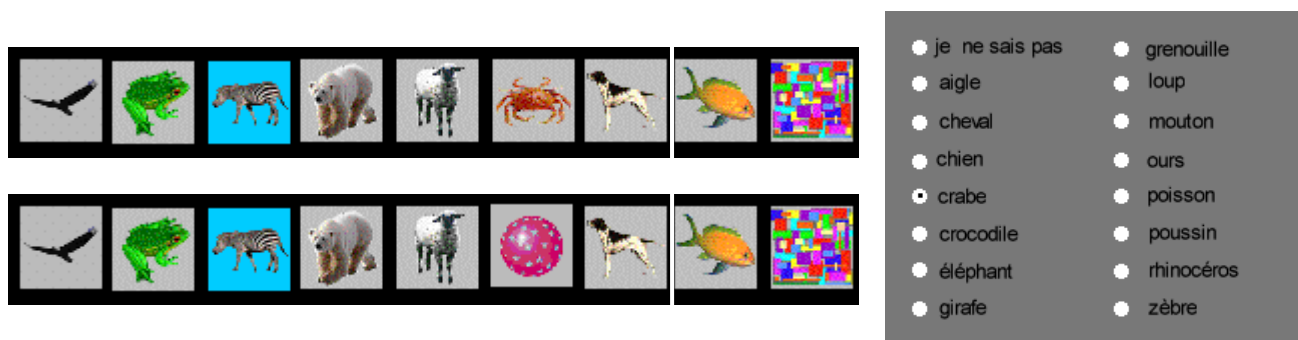


Figure 48 : A gauche, exemples de séquences RSVP de la catégorie 'animaux' avec et sans la seconde cible, le ballon (les séquences sont tronquées, elles comptent usuellement 15 images). A droite, formulaire de réponse. Si le sujet répond aléatoirement par pari, il doit parfois mentionner l'image à laquelle le ballon a été substitué (ici, le crabe). On n'observe pas de tels événements.

Ce résultat suggère que les réponses aléatoires par pari doivent être rares ou absentes. **C'est toujours l'une des images effectivement présentées dans la séquence qui est mentionnée par le sujet.** Ce premier résultat nous a incité à chercher des régularités dans le comportement de ces erreurs. Trois axes d'analyses ont été retenus : la fréquence de ces erreurs, la distribution des erreurs par rapport à la cible (position relative du distracteur par rapport à la cible), et la nature des intrusions (quelle substitution cible-distracteur). La terminologie retenue pour classifier les réponses apparaît à la figure 49.

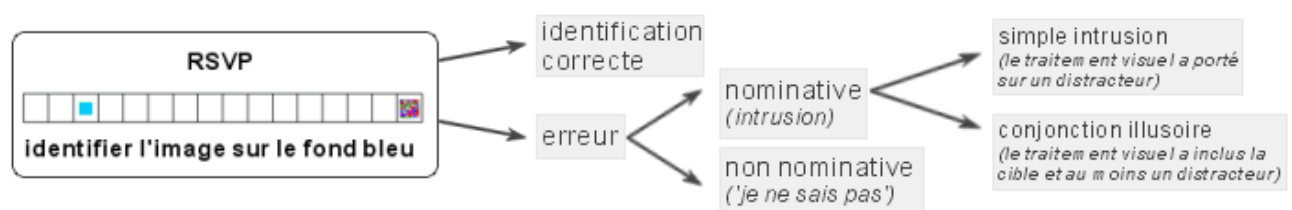


figure 49 : classification des réponses RSVP

Les **erreurs** sont des **intrusions** lorsque le nom d'un distracteur est mentionné en substitution du nom de la cible. D'un point de vue théorique (après Treisman & Schmidt, 1982), ces intrusions sont des **conjonctions illusoires** lorsque la représentation cognitive de la cible s'est formée sur le traitement conjoint d'une cible et d'un distracteur (eg.

mélange indu de caractéristiques visuelles séparables, telles que la forme, la couleur, la luminosité, ou de partie d'objets).

Dans d'autres formulations théoriques (Lawrence, 1971) et particulièrement lors de l'utilisation de mots comme stimuli, des erreurs similaires agrégeant une partie de la cible et une partie du distracteur sont obtenues (réponse 'ramer', alors que la séquence contenait les mots 'radar' et 'fumer'). Elles sont appelées mixture dans ce cadre de recherche. Dans les théories classiques des conjonctions illusoires, l'intégration porte sur des caractéristiques (eg. la couleur d'un stimulus et la forme d'un autre), alors que pour les mixtures, la collision se réalise spatialement entre la partie gauche d'un mot et la partie droite d'un autre mot, chaque partie étant traitée entièrement. Les mécanismes à l'origine des deux phénomènes présentent vraisemblablement des points communs (Botella, Barriopedro & Suero, 2001).

Notons enfin que les expériences usuelles d'identification en RSVP permettent usuellement d'enregistrer des importations (mention d'un stimulus hors du groupe cibles-distracteurs effectivement utilisé). La présence du formulaire de réponse a limité la manifestation de ces erreurs. Seul le stimulus enlevé au profit de la cible à détecter aurait pu être l'objet de ce type d'erreur, et nous avons vu que l'événement ne s'est pas produit.

Similarité et intrusions

La similarité des stimuli est un des candidats à l'explication de la formation des intrusions lors de l'identification. Elle affecte la formation des conjonctions illusoires dans les paradigmes de recherche spatiale (Treisman & Gromican, 1988 ; Ivry & Prinzmetal, 1991 ; Treisman, 1996). Nous suspectons que la formation des conjonctions illusoires en RSVP soit également déclenchée lorsque la cible et le distracteurs partagent une (ou des) caractéristique(s) visuelle(s). La quantification de cette relation est un problème difficile. Nous avons tenté des régressions avec comme variables explicatives la proximité des descripteurs des caractéristiques de bas-niveau de la cible et du distracteur (différence de taille d'image en pixels, différence de moyenne des composantes colorimétriques Lab, différence de compacité, différence de hauteur...) et comme variable dépendante le nombre de confusions entre la cible et le distracteur. Ces modèles présentent une adéquation médiocre (eg. notre meilleur indicateur de Fit était $R^2 = .22$). Cette difficulté conduit à envisager différentes explications :

- Le groupe des descripteurs n'a pas été construit sur une base théorique (primitives visuelles). Nous avons réunis les descripteurs pour lesquels nous maîtrisons les outils de diagnostic (ces descripteurs sont présentés en annexe XIII). Certaines dimensions visuelles importantes ont été laissé de côté par manque d'outil de mesure (eg. comment représenter numériquement la forme globale de l'image ou la texture ?).
- En outre, d'autres facteurs de haut-niveau (comme la similarité sémantique) pourrait également intervenir dans la genèse des intrusions, et n'ont pas été introduits dans les modèles.

- Les descripteurs utilisés moyennaient le caractère à travers l'image (eg. la valeur luminosité était obtenue en moyennant tous les luminosités de tous les pixels de l'image, sans tenir compte des disparités régionales). Quoique légitimé par des recherches sur les fixations oculaires successives (lors de la première fixation, c'est une information holistique qui est extraite ; Loftus, Nelson & Kallman, 1983), ce moyennage est un appauvrissement de l'information, sans doute peu en rapport avec la complexité de l'information visuelle effectivement manipulée par le cerveau.
- Les mécanismes sous-jacents sont hétérogènes : les intrusions observées résultent de diverses défaillances perceptives et attentionnelles.

Dans une perspective exploratoire, nous allons donc exploiter les données de substitution cible-distracteur, en observant les matrices de confusion des stimuli (exemple table 2, pour la catégorie 'fruits & légumes'). En colonnes, nous posons les intruseurs (distracteurs mentionnés en substitution de la cible), et en lignes, nous posons les intrusés (cibles omises au profit d'un distracteur). Par exemple, le chiffre 9 apparaissant au début de la seconde ligne (case jaune) signifie qu'à 9 reprises, la cible 'carotte' a été confondue avec le distracteur 'banane'. Dans ce cas 'carotte' est l'intrusé et 'banane' est l'intruseur. A l'inverse, 'carotte' intruse 'banane' à 5 reprises (case verte).































															
	73	5	8	7	1	3	3	5	1	1	2	1	3	4	2
	9	86	12	4	0	0	3	6	3	5	2	0	4	1	3
	1	2	121	1	3	3	0	3	2	2	1	3	0	0	1
	3	2	12	89	5	7	1	2	3	1	5	0	2	2	3
	4	5	7	5	39	19	5	2	3	4	4	0	2	8	4
	5	3	7	8	16	46	2	2	6	2	6	3	4	2	0
	3	2	13	2	3	10	49	4	6	18	8	2	2	1	3
	3	3	12	3	2	5	1	79	1	1	3	6	3	3	7
	4	2	14	5	5	6	0	4	68	4	5	1	5	1	2
	3	6	12	6	0	4	4	7	5	50	5	7	5	1	8
	2	4	6	5	7	4	2	2	5	4	76	4	5	1	1
	1	2	10	6	4	0	2	12	3	9	3	32	2	2	11
	9	3	4	1	6	8	4	2	4	1	3	2	73	2	6
	1	2	8	5	22	20	3	5	7	4	3	4	6	36	3
	4	6	12	3	4	3	1	17	4	7	8	17	1	1	43

Table 2 : matrice de confusion pour les stimuli de la catégorie 'fruits & légumes'
















															
intruseur	52	47	137	61	78	92	31	73	53	63	58	50	44	29	54
intrusé	46	52	22	48	72	66	77	53	58	73	52	67	55	93	88

Table 3 : effectifs des intrusions causées et subies dans la catégorie 'fruits & légumes'.

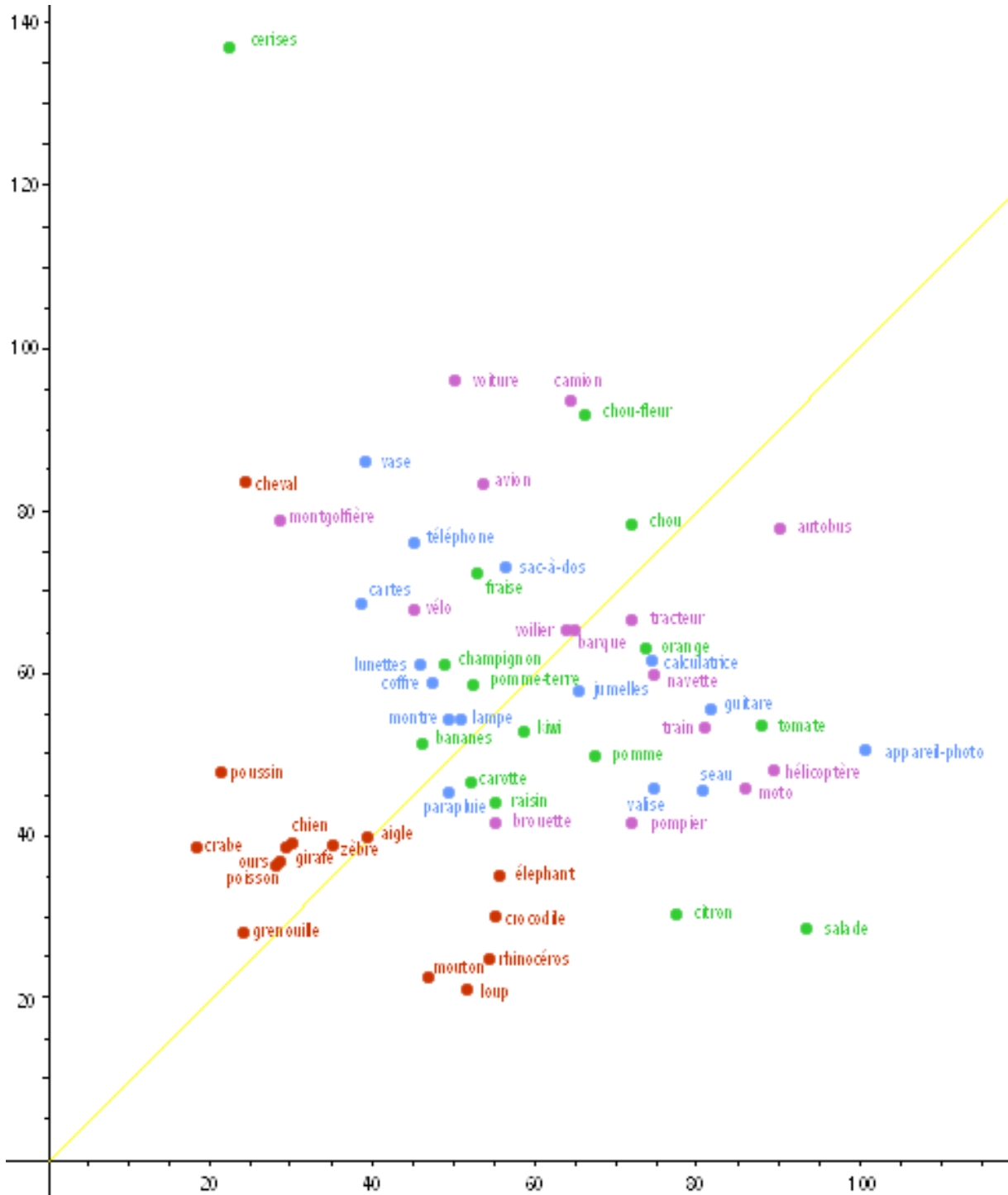


Figure 50 : graphe des intrusions : intrusions subies (abscisses) et causées (ordonnées) par les différents stimuli dans l'expérience RSVP. Notons les stimuli « interrupteurs » dans la partie haut gauche des nuages de points. Ces stimuli sont très souvent intruseurs et rarement intrusés. Leur mode d'action est peu spécifique, ils intrusent assez indifféremment tous les autres stimuli, comme on peut le voir à la table 2 pour l'exemple 'cerises'.

On observe des intruseurs très actifs (eg. cerise), et des confusions se manifestant préférentiellement entre certaines cibles et certains distracteurs (eg. pomme, qui n'est pas un distracteur très actif, mais qui provoque de nombreuses intrusions à la cible tomate). Ces résultats suggèrent l'existence de **deux types d'intruseurs**. Certains intruseurs, comme la pomme dans cette expérience, agissent sur quelques stimuli de prédilection, vraisemblablement

en fonction d'un critère de similarité (eg. forme, couleur, luminance...). D'autres intruseurs fonctionnent quelque soit la cible présentée. Ces stimuli apparaissent en haut à gauche de leurs nuages de points respectifs sur le graphique présenté à la figure 50 ; notons le vase, le cheval, la montgolfière et surtout les cerises). Ces stimuli semblent s'imposer parce qu'ils sont davantage proséxigènes (attractifs). La plausibilité exclusive de l'un ou l'autre des mécanismes en RSVP est questionné par Keysers & Perrett (2002), mais les résultats précédents suggèrent que les deux mécanismes sont possibles. Certains intruseurs semblent se comporter comme des « **interrupteurs** » (ils semblent compromettre le traitement de la cible de manière aspécifique) tandis que d'autres se profilent comme de possibles « **intégrateurs** » (ils semblent perturber sélectivement le traitement des cibles ayant des liens de similarité avec eux-même).

Les stimuli les plus souvent à l'origine d'intrusions sont également ceux qui sont le moins fréquemment intrusés ($r = -.54, p < .0001$), voir aussi table 3, pour la catégorie 'fruits & légumes'. Il s'établit ainsi des relations de dominance entre les stimuli. Certains sont d'efficaces intruseurs. D'autres sont de fréquents intrusés. Les pistes pour l'explication de cette asymétrie sont :

- La qualité picturale de l'image utilisée, qui rend plus ou moins facile l'identification (les images, bien qu'harmonisées pour la résolution et les qualités graphiques, provenaient de sources variées) (Näsänen, Ojanpää, & Kojo, 2001)
- Le caractère plus ou moins proséxigène des stimuli, qui module la priorité cognitive, qui établit une hiérarchie de vitesse de traitement et de qualité de représentation en mémoire de travail
- La fréquence d'usage du nom du stimuli, influant elle-aussi sur le temps de traitement (Content, Mousty & Radeau, 1990)
- Et également en tant que biais inhérent à l'utilisation de photo d'objets comme stimuli, le non-recouvrement d'une partie de l'image-cible par la succession des images de la séquence (voir figure 51), qui induit potentiellement une stratégie d'identification différente (persistance des traitements sur les parties non recouvertes).



Figure 51 : cartes de superposition des stimuli. On remarque un non-recouvrement de certains stimuli (eg. la cerise, la fraise et l'orange pour le premier groupe, le cheval et le poussin pour le second groupe, la montre et la valise pour le troisième groupe, la brouette, la montgolfière et le voilier pour le dernier groupe). Pour des stimuli surpris au cours de nombreux essais, cette partie non masquée peut être l'objet d'une stratégie spécifique d'identification.

Les intrusions se distribuent de manière spécifique

Observer comment se distribuent les intrusions dans la séquence d'image peut nous aider à comprendre les mécanismes sous-jacents. La distribution des intrusions est construite en prenant le rang zéro comme référence de la position de la cible, et en dénombrant les fréquences d'intrusions pour chaque rang relatif à cette référence. Ainsi, une intrusion '-1' se réfère à l'intrusion causée par un distracteur situé juste avant la cible (au rang moins 1 par rapport à la cible).

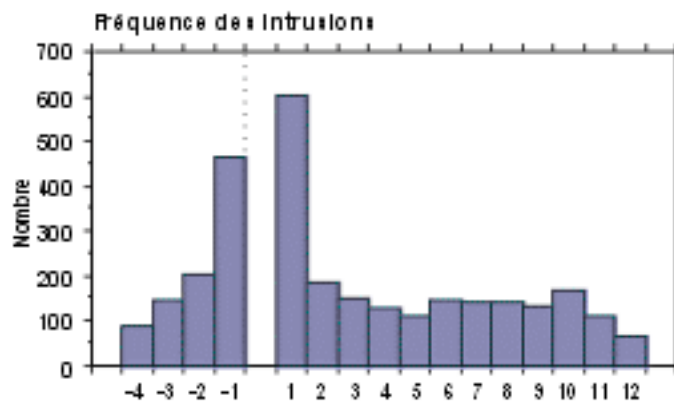


Figure 52 : distribution des intrusions

La distribution des intrusions pour l'ensemble des sujets est présentée à la figure 52. On observe que les intrusions sont nombreuses autour de la position de la cible. Les intrusions '+1' sont les plus fréquentes, les intrusions '-1' sont légèrement moins nombreuses. Les intrusions des positions $[-4 ; -2] \cup [+2 ; +12]$ sont assez régulièrement distribuées, avec une apparente baisse pour les deux bords de la distribution. Cette baisse est un artefact du montage expérimental. Pour que la position de la cible à identifier ne soit pas prédictible, celle-ci variait aléatoirement de la position 3 à 5 dans la séquence. Aussi, les effectifs des intrusions situées au bord de la distribution sont sous-estimés. Par exemple, seul un tiers des séquences (les séquences dont la cible est au rang 5) peut donner lieu à des intrusions '-4'. La distribution corrigée pour cet artefact est présentée à la figure 53. On observe alors un léger effet de primauté (rappel préférentiel de la première cible de la séquence) et un léger effet de récence (rappel préférentiel de la dernière cible de la séquence).

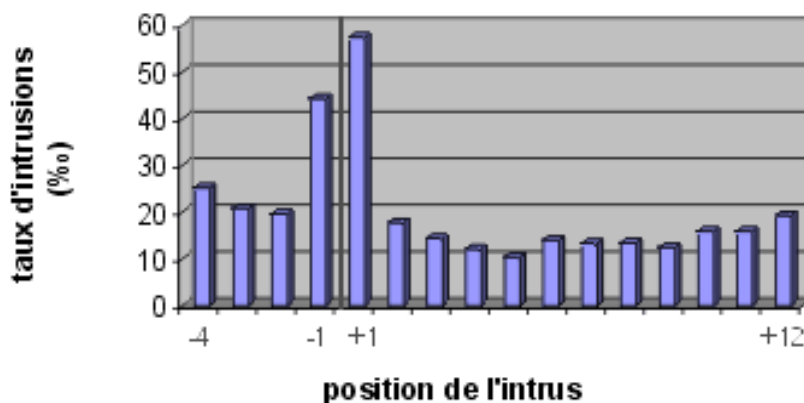


Figure 53 : distribution des intrusions corrigée pour le biais marginal

Cette distribution est assez conforme à celles observées par les autres équipes travaillant sur la distribution des intrusions en RSVP. Ces études relèvent également **de nombreuses intrusions proximales** (situées juste avant et juste après la cible) **et moins d'intrusions distales** (situées à plus d'un rang de la cible). Notons toutefois que les intrusions distales semblent plus nombreuses dans la présente expérience que dans les études présentées par ailleurs (Boucart, Moroni, Fuentes & Belin, 1998 ; Boucart, DeWisme & Wagemans 2000 ; Chun, 1997). Nous verrons à travers une réplication de la présente expérience chez l'adulte (p.86) que cette spécificité semble être liée à la population enfantine. Les enfants semblent commettre davantage d'erreurs de rappel que les adultes.

Concernant les intrusions proximales, nous obtenons une prédominance des intrusions '+1'. Les études utilisant des images d'objets trouvent parfois une prédominance des intrusions '-1' (Boucart, & coll. 1998, Boucart & coll, 2000, dans des expériences similaires à celles utilisées ici), et parfois une variabilité du pattern en fonction des conditions expérimentales (Intraub, 1985, 1989).

Au contraire, les expériences réalisées avec des mots (Lawrence 1971 ; Botella et Villar, 1986, Botella, Barriopedro, et Suero, 1992 ; 2001) et des lettres (McLean, Broadbent, & Broadbent, 1982) montrent plus systématiquement une prédominance des intrusions '+1' (mais voir Kikushi, 1996, avec des mots kanji, une écriture japonaise).

Le centre de gravité de la distribution peut être manipulé par plusieurs facteurs. Le centre de gravité est mesuré par l'indice API (*average position of intrusions*, Botella & Ericksen, 1991). Une postériorisation de la distribution (déplacement du centre de gravité de la distribution vers les valeurs positives de l'API) est obtenue avec :

- des vitesses de défilement plus rapides (Botella et Ericksen, 1991)
- la dimension de réponse (eg. « quelle était la couleur de la cible ? » ; « quel était le nom de la cible ? ») (Botella, Ericksen, 1992, Botella, Garcia, Barriopedro, 1992, McLean, Broadbent & Broadbent, 1982)
- des paramètres affectant les vitesses de traitement relatives du critère de réponse (Botella et Barriopedro & Suero, 2001) et de l'image-cible (Intraub, 1985, Botella, Barriopedro, & Suero, 2001)
- la charge attentionnelle (Chun, 1997, Maki, Couture, Frigen, & Lien, 1997)
- Les disparités apparaissant dans les différentes études rendent également plausible un effet du type de matériel utilisé (mots, chiffres, lettres, images d'objets, scènes)

En référant la présente distribution à celles obtenues par Boucart & coll, (1998), qui utilise un protocole similaire, les explications de la disparité des intrusions '-1' et '+1' seraient à rechercher auprès de deux facteurs de variations :

- Les sujets étaient des enfants. Il s'agit peut-être d'une caractéristique développementale, mais cette interprétation sort du cadre de cette étude
- La vitesse de défilement des images diffère de celle utilisée par Boucart. Ce facteur est réputé modifier le pattern des intrusions (Botella & Ericksen, 1991)

Nous pouvons obtenir des informations complémentaires en analysant différenciellement les sujets en fonction de leur **efficience en identification**. On peut supposer que les sujets efficaces réalisent le traitement plus rapidement. La contrainte temporelle leur est plus légère. Nous postulons que la même source de variation est manipulée lorsque l'expérimentateur diminue la vitesse de défilement ou lorsqu'il invoque des groupes de sujets plus efficaces. Un split a été réalisé sur les sujets du groupe expérimental (15 sujets non efficaces ; 19 sujets efficaces, split réalisé à 44% de bonnes réponses). Les distributions des intrusions des sujets efficaces et peu-efficaces sont reportées à la figure 54.

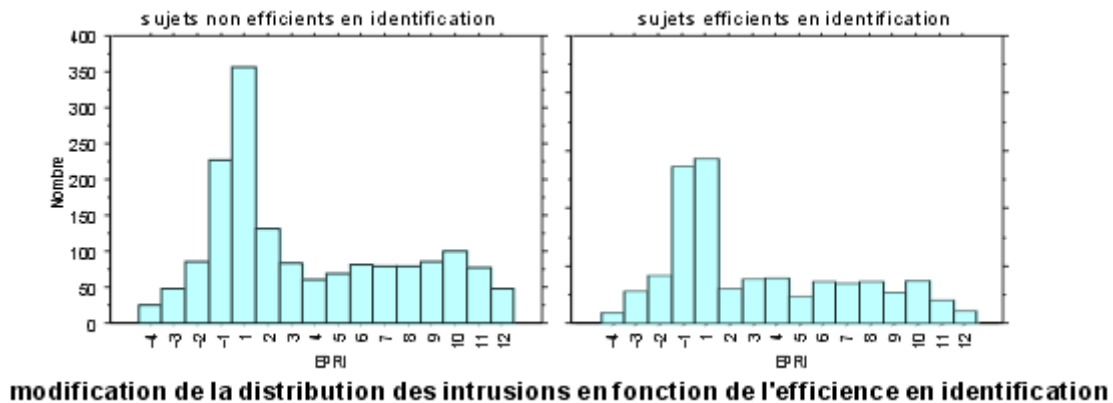


Figure 54 : distribution des intrusions pour les sujets efficaces vs peu-efficaces en identification (distributions agrégées)

Les deux distributions ne se comportent pas de la même manière. On observe à la figure 54 que les intrusions '+1' et '+2' différencient tout particulièrement les deux groupes. Les sujets efficaces ne présentent pas un abaissement régulier de toutes les intrusions. Ils minimisent tout particulièrement les intrusions '+1' et '+2' alors que la plupart des autres intrusions montrent seulement des différences quantitatives (moins d'intrusions distales, révélatrices d'un faible fonctionnement mnésique), et que les intrusions '-1' sont très similaires dans les deux groupes.

Chez l'adulte, une tâche RSVP plus facile (eg. cadence de défilement moins élevée) se traduit par un profil plus antérieur. Chez l'enfant, on retrouve également un déplacement du centre de gravité de la distribution vers la gauche chez les sujets réalisant l'identification plus facilement (voir table 4).

Sujets	API	T-test		
Efficaces	-0,065	T=-3,30	ddl=32	p=.0024
Non efficaces	0,29			

Table 4 : déplacement du centre de gravité vers la gauche pour les sujets efficaces, relativement aux sujets non efficaces. L'API est calculé sur les effectifs d'intrusions des deux rangs précédents et succédant à la cible. Le calcul est celui d'un centre de gravité.

Ces résultats sont en accord avec une modulation de la distribution par la difficulté de la tâche (vitesse de défilement). Bien que les vitesses de défilement utilisées ici et celles utilisées par Boucart & coll., (1998) ne soient pas commensurables, les résultats précédents suggèrent que le « stress » temporel était plus important pour les enfants de cette expérience que pour les adultes testés par Boucart & coll, (1998).

Une seconde analyse a été conduite pour estimer le poids de la charge attentionnelle (double-tâche ou simple-tâche) sur le pattern des intrusions. Cette charge attentionnelle est connue pour modifier le centre de gravité de la

distribution (charge attentionnelle accrue, postériorisation de la distribution, Chun, 1997, Maki, Couture, Frigen & Lien, 1997). Cette analyse n'a en revanche pas abouti aux mêmes conclusions chez l'enfant (voir figure 55). **On n'observe pas d'effet de la charge attentionnelle sur la distribution.**

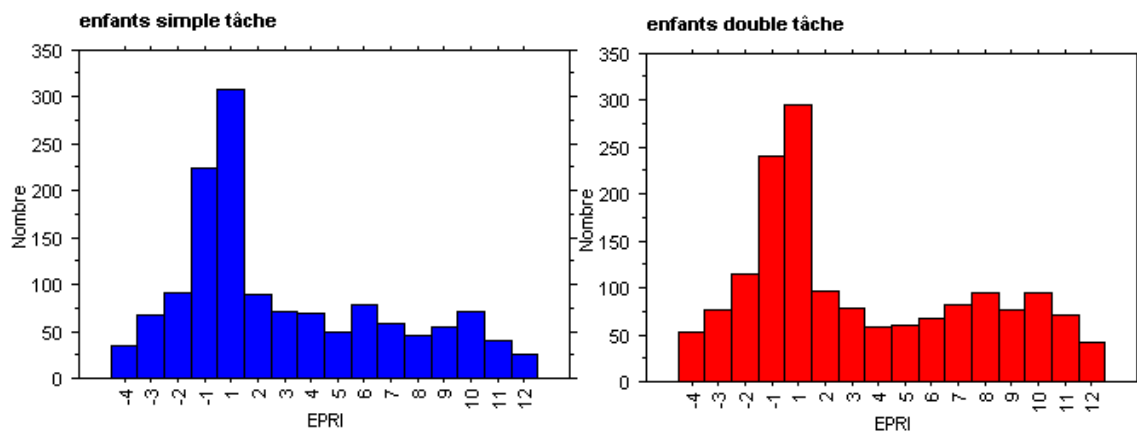


Figure 55 : distribution des intrusions pour les essais en simple et double-tâche chez l'enfant

Une réplication de cette expérience RSVP a été menée chez 30 adultes (expérience préliminaire B, voir caractéristiques de l'échantillon en annexe XIV). La vitesse de défilement était plus rapide pour harmoniser la difficulté dans les deux groupes. Cette réplication montre que les distributions d'intrusions des adultes présentent un effet de la charge attentionnelle (voir figure 56). De manière générale, les sujets adultes présentent moins d'intrusions distales. Ils présentent un pattern pré-cible en simple-tâche (prédominance des intrusions pré-cible) qui n'est pas présent dans le groupe des enfants. Et enfin, on observe chez les adultes une postériorisation de la distribution lors de la double-tâche, conformément aux autres études manipulant le facteur 'charge attentionnelle' (Chun, 1997, Maki, Couture, Frigen, & Lien, 1997). Une différence développementale semble observée ici. Le profil de la distribution des intrusions chez les enfants était déjà très postérieur en simple tâche, peut-être ne peut-il pas se postérioriser davantage ? Dans ce cas, on pourrait trouver chez les sujets efficaces des signes de postériorisation, puisque ceux-ci présentent une distribution plus antérieure. On n'observe pas davantage de modification de la distribution dans ce sous-groupe des enfants efficaces en identification.

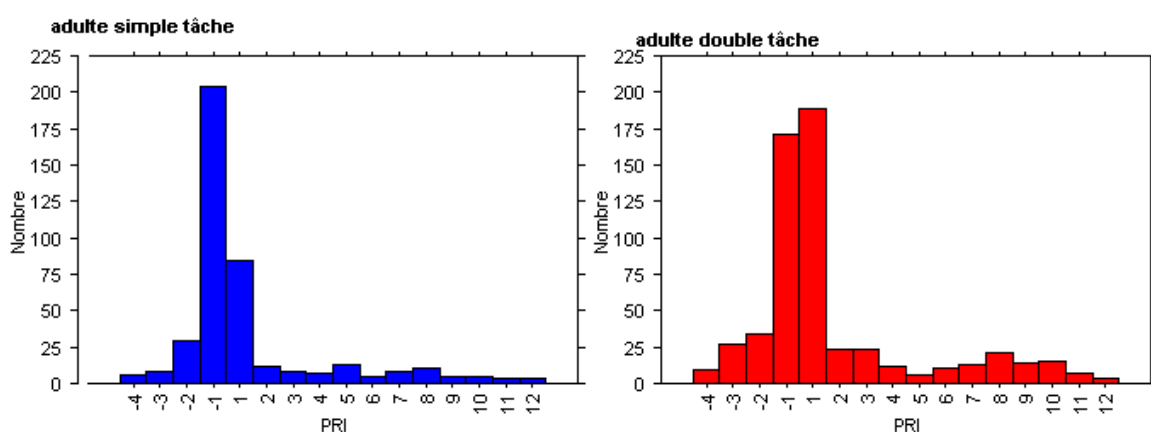


Figure 56 : distribution des intrusions pour les essais en simple et double-tâche chez l'adulte

Un mécanisme systémique semble présider à la distribution des intrusions

Les règles de formations des intrusions sont-elles des règles dans l'absolu (eg. si la proximité est supérieure à une valeur critique, alors la confusion est possible) ou sont-elles régies par les caractéristiques du groupe de stimuli en présence dans l'expérience (eg. considérant les différents stimuli à discriminer, le sujet se focalise sur tel ou tel critère perceptif) ?

L'analyse des nuages de points sur les caractéristiques visuelles montre effectivement que le sujet exploite les caractéristiques les moins représentées dans le groupe de stimuli pour réaliser moins d'erreurs (voir figure 57). La plupart des stimuli sont massés dans la partie centrale des graphiques. Le sujet ne peut pas exploiter ces caractéristiques comme discriminantes, car elles sont partagées par un nombre trop élevé de stimuli. En revanche, un stimulus doté d'une caractéristique anormalement basse ou élevée sera plus facilement discriminé. Nous appellerons ces stimuli, « **distinctifs** » pour le spectre rouge, ou pour la luminosité ...

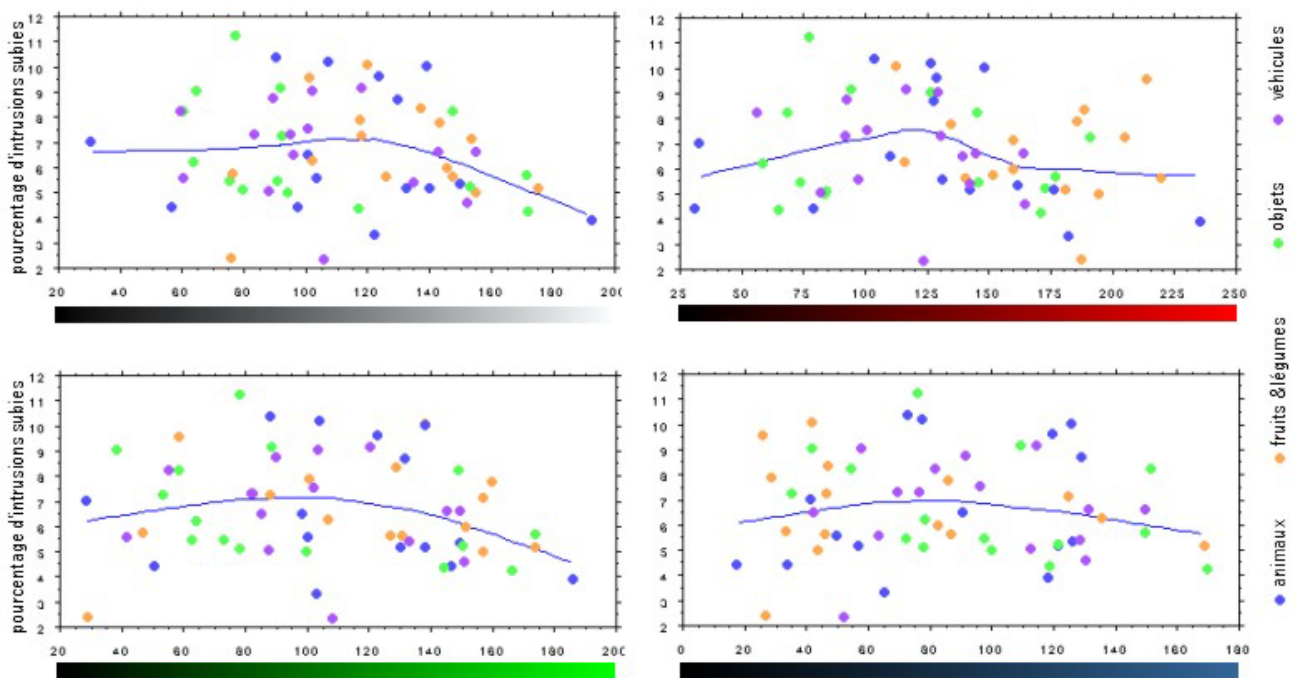


Figure 57 : nuages de points des pourcentages d'intrusions en fonction de caractéristiques chromatiques (modèle RVB). Du haut à gauche en bas à droite, caractéristique RVB, R, V, & B. Une régression non linéaire (spline LOWLESS, tension 66) montre systématiquement qu'une cible ayant une caractéristique dissemblable au reste du groupe (bords des graphes) est moins intrusée qu'une cible ayant un profil central.

L'analyse des distributions d'intrusions donne une seconde information. Nous avons observé que les intrusions se distribuaient de manière spécifique. En dissociant les distributions dans les quatre catégories de stimuli étudiées, nous retrouvons les mêmes profils de distribution (voir figure 58). En revanche, en étudiant les distributions au niveau de chaque stimulus, nous voyons apparaître une différenciation. La distribution spécifique des intrusions discutée p.83 est donc émergente, elle n'existe pas au niveau individuel. L'expliquer n'est cependant pas simple. Nous n'avons trouvé aucune règle récurrente dans les quatre catégories (eg. attribution de rôles spécifiques aux différents stimuli). Les effectifs des observations présentées ici sont faibles, mais la généralité de l'observation (même les stimuli les plus intrusés, dotés de davantage d'observations, se comportent de manière imprédictibles) rendent peu probable un simple artefact d'échantillonnage dans un petit groupe d'observation.

Conclusion : des conjonctions illusoires

A l'incidence d'une recherche menée sur le clignement attentionnel chez l'enfant, nous avons été étonné du comportement des erreurs d'identification de la cible. Ces erreurs ne sont pas apparues comme des tentatives de réponses par sélection aléatoire d'un nom d'objet (pari). Il semble que les erreurs relèvent (1) de la **substitution** du traitement cognitif de la cible par le traitement d'un distracteur, ou (2) du traitement conjoint (**intégration**) de la cible et d'un distracteur (la représentation ainsi formée étant une chimère cognitive, et l'erreur étant une conjonction illusoire).

Il semble que la **similarité** cible/distracteur soit explicative des intrusions (eg. une intrusion est plus probable entre deux stimuli partageant une couleur et/ou une forme... qu'entre deux stimuli ne partageant aucun trait). Aux expériences ultérieures de la présente thèse, nous adressons la question de l'homogénéité de cette relation. Les intruseurs '+1' sont-ils déterminés par les mêmes facteurs de similarités que les intrusions '-1' ou que les intrusions distales ?

Il est apparu que les intrusions proximales se comportent différemment des intrusions distales. Les intrusions distales semblent être des erreurs peu spécifiques (l'efficiency ne les module que quantitativement, mais hormis cela, elles sont équi-réparties). Ces erreurs sont formées a posteriori du traitement de la cible : on n'observe pas de dépression quantitative des intrusions dans le rang du clignement attentionnel (ie. des rangs '+2' à '+5'). Or, si le traitement de ce distracteur était en ligne avec celui de la cible, on devrait observer moins d'intrusions dans les rangs '+2' à '+5' par rapport aux autres rangs distaux, car le traitement de la cible inhiberait l'accès conscient à la représentation de ces distracteurs (Shapiro, Arnell & Raymond, 1997). Cette observation suggère que dans ce cas, les traitements de la cible et du distracteur sont temporellement disjoints (**erreur offline**). En revanche, les intrusions proximales semblent affectées qualitativement (modification de la répartition '-1', '+1'). Elles mettent en coïncidence temporelle directe les traitements de la cible et du distracteur (**erreur online**).

Il semble enfin qu'un mécanisme systémique préside à la formation des intrusions, comme en atteste le pattern émergent de distribution des intrusions. On obtient une distribution des intrusions très stable et reproductible lorsque l'on analyse toutes les intrusions, et lorsque l'on subdivise les distributions des 4 catégories sémantiques étudiées. Cette distribution spécifique s'évanouit lorsque l'analyse est menée sur les stimuli individuellement. Ce phénomène est difficile à comprendre en l'absence d'une attribution de rôles prédéterminés. Comment des distributions individuelles si disparates, ne semblant être dirigées par aucune règle, peuvent-elles former dans chaque groupe sémantique des patterns aussi réguliers ?

Ces questions et hypothèses vont dès à présent faire l'objet d'une série d'expériences. Le protocole RSVP sera aménagé pour servir l'observation des intrusions. Les bases théoriques et empiriques de la suite de ce travail sont présentées au chapitre qui suit.

Bases théoriques et hypothèses

L'expérience précédente nous a incité à penser que certaines des erreurs d'identification en RSVP étaient des **conjonctions illusoires**. Ces erreurs ont été **mise en évidence par Treisman & Schmidt (1982)**. Les auteurs soumettent les sujets à une double-tâche (voir figure 59). Ils devaient d'abord appairer des chiffres bordant l'affichage (tâche distractive), puis ils devaient décider de la présence de certains éléments dans l'affichage. Les sujets commettent alors des erreurs dans la seconde tâche. Ils déclarent par exemple avoir vu les lettres Q, ou R, ou des barres horizontales vertes. Or, ces éléments sont absents de l'affichage, on trouve seulement des parcelles d'informations sur plusieurs objets présents (couleur, orientation, portion de segment d'une lettre).

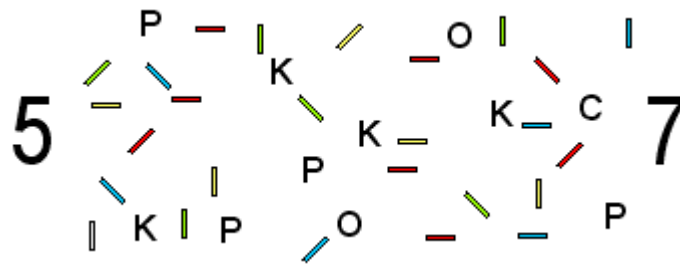


Figure 59 : un des expériences originales de Treisman & Schmidt (1982). Les sujets devaient tout d'abord appairer les deux chiffres situés sur les bords (sont-ils tous les deux impairs ?). Puis ils devaient décider si certains éléments étaient présents dans l'affichage. Les sujets commettent des conjonctions illusoires dans la seconde tâche. Ils reportent des configurations absentes, mais dont les parties sont effectivement contenues dans l'affichage, telles que des lettres R, des barres horizontales vertes (reproduit d'après Wolfe & Cave, 1999)

Ensuite, ces erreurs continuèrent à être étudiées dans des paradigmes de recherche visuelle avec les **distracteurs situés dans le domaine spatial**, mais la double-tâche fut remplacée par une contrainte temporelle (« *brief multi-elements displays* », ou « affichages multi-éléments brefs », Treisman, 1986). Elles ont été définies comme des **faillites dans la recombinaison correcte des caractéristiques visuelles** traitées de manière préattentive. Par exemple, un sujet visionnant brièvement un affichage fait de lettres 'O' vertes et de lettres 'X' rouges, déclarera parfois avoir vu une lettre 'X' verte parmi les éléments présentés (voir figure 60). Il a recombinaison la couleur verte de certains éléments présents dans l'environnement et la forme du 'X'. Cette erreur est possible parce que les caractéristiques sont flottantes tant que le traitement est préattentif.

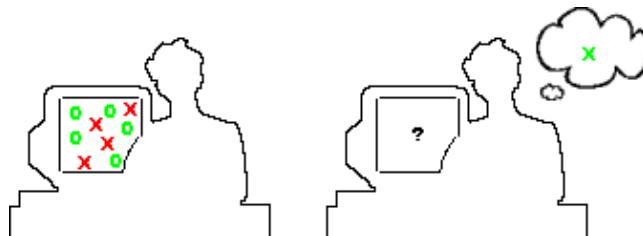


Figure 60 : les conjonctions illusoires lors de recherches visuelles brièvement affichées : le sujet doit décider si l'écran qu'il a vu brièvement contenait ou non un 'X' vert. Lorsque l'affichage n'était constitué que de X rouges et de O verts, il commet parfois des erreurs. Il recombine fallacieusement les caractéristiques flottantes de couleur et de forme.

Les études consacrées aux conjonctions illusoires postulent une dichotomie entre **traitement parallèle** précoce (pré-attentif, automatique) et **traitement sériel** tardif (attentif, contrôlé)¹. Cette dichotomie fut suggérée par l'observation des « pentes de recherches » (*search slope*, calculée par la fonction : temps de réponse TR=f(nombre de distracteurs) eg. Treisman, 1986). Ces pentes sont différentes pour les **recherches simples**, -concernant les caractéristiques indépendantes des objets-, et pour les **recherches conjonctives**, -concernant les associations de caractéristiques- (ces dernières étant la version « laboratoire » des objets). Un exemple de pentes différentes est donné à la figure 61. Lorsque la cible et le distracteur se distinguent par une seule caractéristique, la recherche visuelle occupe le même TR (temps de réponse), quelque soit le nombre de distracteurs. On dit de la cible qu'elle « *pop-out* ». Lorsque la cible et les distracteurs partagent des caractéristiques, et que seul un diagnostic sur leur combinaison permet la décision, alors le TR augmente proportionnellement au nombre de distracteurs. Cette dernière recherche est supposée être réalisée par un balayage séquentiel de tous les objets présents dans la scène, jusqu'à ce que la cible soit découverte.

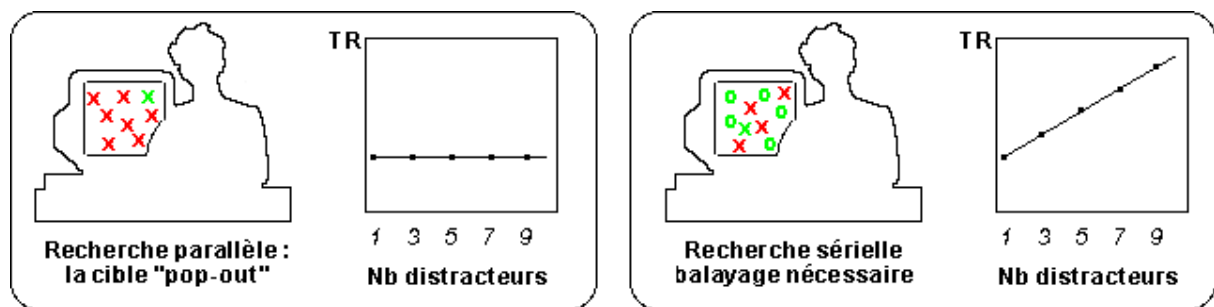


Figure 61 : pentes de recherche pour une recherche simple (à gauche) et pour une recherche conjonctive (à droite). La cible est un X vert. Le sujet doit décider de sa présence en appuyant sur un bouton. L'expérimentateur enregistre l'exactitude et le TR.

A la base de cette conception dichotomique des traitements pré-attentifs vs attentifs, se trouve l'idée d'une restriction quantitative pour l'accès aux traitements cognitifs (Broadbent, 1957 ; Kahneman, 1973). Cette restriction contraint les traitements pré-attentifs à une compétition pour l'accès au traitement cognitif.

Les objets accèdent aux traitements cognitifs par le biais de l'**attention** (Treisman & Gelade, 1980, Treisman, 1986). L'attention agit sur des cartes spatiales² de caractéristiques (*feature maps*). Elle prend la forme d'un faisceau unitaire (spotlight, semblable à celui formulé par Posner (1980), qui va solidariser entre-elles les caractéristiques (on parle d'attention-glu), et fonder la représentation des objets (voir figure 62). Cette hypothèse est contemporaine de la

¹ Cette dichotomie a été modérée depuis les premières publications sur le sujet par Treisman & Gelade en 1980. Certaines dimensions des objets semblent devoir échapper aux traitements sériels, et tout particulièrement des dimensions relevant de la voie magnocellulaires, liées par exemple aux surfaces, à la stéréoscopie (Nakayama & Silverman, 1986).

² Concernant le débat « attention orientée vers l'espace ou vers l'objet » (*space-based or object-based theories of attention*), le lecteur pourra consulter la discussion dans Fernandez-Duque & Johnson, (1999) pp.94-95. Les conceptions de l'attention basées sur l'espace sont surtout supportées par les théories pré-motrices (l'attention 'covert' est un analogue interne du mouvement oculaire, voir Desimone, 1998). Les conceptions de l'attention orientée vers l'objet trouvent leur source dans les travaux de Treisman et d'autres, qui montrent qu'il est plus facile de commuter l'attention sur différentes caractéristiques d'un objet que sur différentes caractéristiques de différents objets. Voir également (Scholl, 2001).

théorie des **deux voies visuelles ventrale et dorsale** (« what and where » de Ungerleider et Mishkin (1982), voir aussi Ungerleider & Haxby (1994), et semble calquée sur ce modèle du fonctionnement visuel. Les corrélats anatomiques du faisceau attentionnel sont d'ailleurs compatibles avec cette théorie, puisqu'ils situent une partie du réseau **d'attention spatiale** dans la voie dorsale (Laberge, 2000 ; Mesulam, 1990 ; Posner & Petersen, 1990 ; Corballis, 1995) Par la suite, des études ont mis en évidence la survenue de conjonctions illusoires alors que l'attention était focalisée, et/ou l'absence de conjonctions illusoires lorsque l'attention est divertie (Tsal, 1989 ; Tsal, Meiran & Lavie ; 1994, Prinzmetal, 1995 ; Navon & Ehrlich, 1995 ; Wolfe et Cave, 1999), affaiblissant les bases théoriques de la dichotomie préattentif-attentif. L'hypothèse alternative faisait appel à la notion de **codage grossier** (*coarse coding*). Le codage visuel grossier est lui aussi attribué à la voie dorsale.

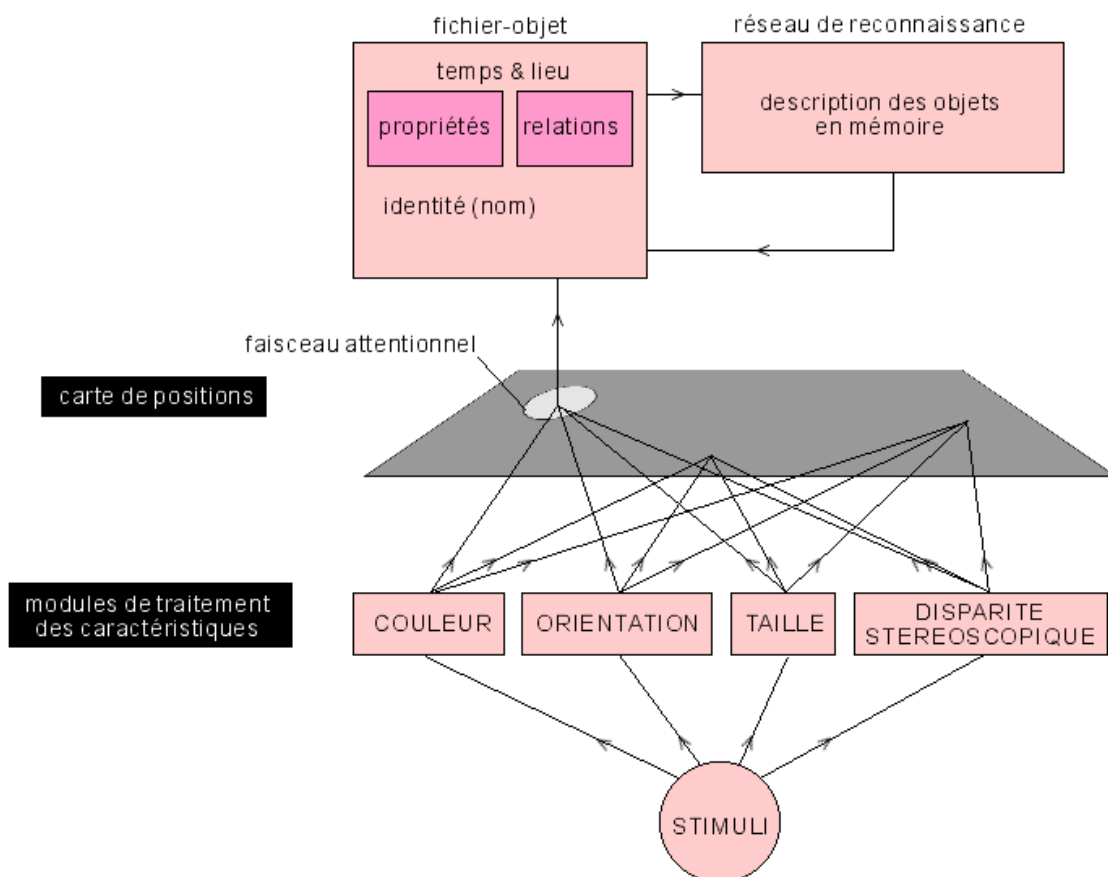


Figure 62 : le modèle FIT (feature integration theory) de Treisman, en 1986 : dans un premier niveau de traitement, parallèle, les caractéristiques de tous les objets présents dans l'environnement sont résolues. Dans un second niveau de traitement, seuls les objets tombant sous le faisceau attentionnels créent une représentation temporaire (fichier-objet ou object-file), permettant la récupération consciente de leurs noms et de leurs propriétés (d'après Treisman, 1986). Ce modèle a évolué au fil des années (voir par exemple Treisman, 1998), mais propose constamment que l'attention crée la glue pour solidariser les caractéristiques visuelles des objets entre-elles. (reproduit de Treisman, 1986)

Résultats issues des paradigmes de recherche spatiale.

Concernant les paradigmes de recherches visuelles spatiales, les résultats antérieurs qui importent pour le présent projet sont résumés ci-après. Une revue peut être trouvée dans Treisman (1996), dont voici les principaux points :

- Les conjonctions illusoires se manifestent lorsque les temps d'exposition aux stimuli sont trop brefs pour permettre des fixations individualisées de chaque objets.
- Les conjonctions illusoires se manifestent même si la cible est située en région fovéale.
- L'orientation préparatoire de l'attention améliore les performances, alors que la situation d'attention divisée aggrave le déficit.
- La similarité des cibles (items appartenant au même groupe perceptuel ou conceptuel) aggrave également les performances.

D'autres auteurs ont établi que la tâche est sensible à l'apprentissage ; des recherches initialement sérielles peuvent devenir parallèles (Sireteanu et Rettenbach, 2000). Des recherches issues de la neuropsychologie (Humphreys, Cinel, Wolfe, Olson, & Klempen, 2000) suggèrent en outre que les différentes caractéristiques ne sont pas assemblées simultanément. Les éléments sont tout d'abord assemblées pour générer la forme. Ensuite, la forme est intégrée avec les détails de la surface. Ce modèle ne propose que deux niveaux de liage, sur la base du profil pathologique d'un patient atteint de lésions pariétales bilatérales. Cependant, on peut présumer l'existence d'un plus **grand nombre d'étapes d'assemblage**, pour peu qu'une variété de dimensions soit insérée dans les protocoles expérimentaux et que la population pertinente soit interrogée. Cette évolution vers différents niveaux d'intégration est visible dans Treisman, (1996), qui décline (1) le liage des propriétés (2) le liage des parties d'objets (3) le liage des étendues de réponses de neurones (4) le liage hiérarchique (5) le liage conditionnel (6) le liage temporel et (7) le liage positionnel. Dans une perspective connexionniste, ce résultat s'entend comme l'émergence progressive des différents assemblages au fur et à mesure de leur élucidation par le système visuel. Cette perspective est compatible avec la fonction présumée de la synchronisation neuronale (voir chapitre 2 pp.26-52), et les différentes échelles (locale, étendue, inter-hémisphérique) auxquelles elle se manifeste.

Notons enfin que des conjonctions illusoires tardives ont été mises en évidence par Virzi & Egeth (1984). Ces auteurs ont observé des conjonctions illusoires basées sur la signification de mots (eg. voir le mot 'Rouge' ou la couleur d'encre 'Brun' dans un affichage qui contenait le mot 'Brun' écrit en rouge). Cette étude étend le phénomène de conjonctions illusoires des étapes perceptuelles précoces aux étapes de traitement plus tardives, postérieures à l'action de l'attention.

Conjonctions illusoires en RSVP : recherches classiques

D'autres auteurs (Potter & Levy, 1969, Intraub, 1985, 1989 ; McLean, Broadbent & Broadbent 1982, Botella & Ericksen, 1992 ; Botella, Gracia, & Barriopedro, 1992 ; Botella, Barriopedro & Suero, 2001 ; Chun, 1997 ; Boucart, Moroni, Fuentes, Belin, 1998 ; Boucart, De Wisme & Wagemans, 2000, Maki, Couture, Frigen, Lien, 1997, Maki, Frigen, Paulson, 1997 ; Subramamiam, Biederman & Madigan, 2000, Keysers, Xiao, Földiák & Perrett, 2001 ; Keysers & Perrett, 2002), étudient les intrusions ou les conjonctions illusoires dans un paradigme de recherche dans lequel les distracteurs sont disposés dans la dimension temporelle. Ces études sont parfois entièrement dédiées au phénomène

(Intraub, 1985, 1989 ; Botella & Ericksen, 1992 ; Botella & coll., 1992 ; Botella & coll., 2001 ; Chun, 1997 ; Subramaniam & coll., 2000 ; Keyser & coll., 2001 ; Keyser & Perrett, 2002), mais plus souvent, ce sont des analyses incidentes réalisées lors de recherches sur le clignement attentionnel. Le paradigme utilisé est le RSVP (*Rapid Serial Visual Presentation* ou présentation sérielle visuelle rapide). La plupart des expériences utilisent des stimuli alphanumériques ou symboliques (mots, lettres, chiffres). Quelques groupes de recherche (Intraub, Boucart, Subramaniam, Keyser) utilisent des images d'objets réels ou des dessins au trait. Les analyses sont conduites sur la fréquence des erreurs, et plus rarement sur la position sérielle des erreurs par rapport à la cible, (Intraub, 1985, 1989 ; Botella & coll., 1992, Botella & coll., 2001 ; Boucart & coll., 1998 ; Boucart & coll., 2000). Aucune étude ne reporte d'analyses sur les types de substitutions cible-distracteur observés. Cette forme d'analyse requiert l'utilisation d'un nombre restreint de stimuli, et les analyses peuvent rarement être menées au niveau inter-individuel, en raison de la difficulté de collecter suffisamment d'informations à ce niveau. Nous tenterons de mettre en oeuvre de telles analyses.

La **terminologie** spécifique des études des conjonctions illusoire en RSVP (après Botella & Ericksen, 1992) est reportée ci-après.

Prenons l'exemple d'une séquence RSVP faite de lettres. La cible est la seule lettre écrite en majuscule, et l'expérimentateur demande au sujet quelle est la couleur de cette lettre. La lettre en majuscule a été présentée dans la couleur bleue. Sur l'ensemble de la séquence, les lettres peuvent apparaître en 5 couleurs différentes (rouge, bleu, vert, jaune, rose).

- Le **critère de réponse** (*key feature*) est la majuscule. Cette caractéristique est extraite de la **dimension du critère de réponse** (*key dimension*) « casse de la lettre », qui comporte potentiellement deux modalités (majuscule et minuscule)
- La **caractéristique de réponse** (*response feature*) est « bleue ». Cette caractéristique est extraite de la **dimension de réponse** (*response dimension*) « couleur de la lettre ». Cette dimension de réponse comporte potentiellement de nombreuses modalités, mais on la circonscrit aux exemplaires effectivement utilisés dans le protocole expérimental (dans l'exemple, cinq modalités : rouge, bleu, vert, jaune, rose)
- On observe une **intrusion post-cible** (*post-target intrusion ; post-intrusion*) si le sujet donne la couleur d'une lettre située après la cible
- On observe une **intrusion pré-cible** (*pretarget intrusion ; pre-intrusion*) s'il reporte la couleur d'une lettre située avant la cible
- A l'issue de l'expérience, on relève un **post-pattern** (id) si la majeure partie des erreurs sont des intrusions post-cible
- on relève un **pré-pattern** (*pre-pattern*) si la majeure partie des erreurs sont des intrusions pré-cible

- et on relève un **pattern symétrique** (*symmetrical pattern*) si les intrusions pré-cible et post-cible sont en nombre égal

Nous ajoutons à cette terminologie les notions d'**intrusions pré-proximales** (intrusions proximales précédant la cible) et **post-proximales** (intrusions proximales succédant à la cible). Dans la présente série d'expériences utilisant des images d'objets, mais également dans Intraub (1985, 1989) avec des photographies d'objets et dans Chun (1997) avec des lettres, ces deux classes d'intrusions proximales sont équivalentes aux intrusions '-1' et '+1', puisque la majoration des intrusions proximales par rapport aux intrusions distales n'est observée que sur ces deux positions. Dans d'autres expériences, elles peuvent englober également les positions '-2' et '+2' (eg. Botella & coll., 1992 ; Botella & Ericksen, 1992).

Les explications des intrusions en RSVP restent très divergentes. Tout d'abord, il se pose la question de l'origine **perceptive-attentionnelle** (précoce) ou **mnésique** (post-identification) des erreurs d'identification. Si on se réfère à la définition des **conjonctions illusoires** (Treisman & Schmidt, 1982), il s'agit d'un déficit de l'intégration des caractéristiques visuelles (forme, couleur...) extraites de manière pré-attentive et parallèle, c'est-à-dire avant l'identification complète de la cible. Si l'on se réfère à la définition des **intrusions** (eg. Koriat, Goldsmith & Pansky, 2000), il s'agit d'une interférence se produisant en mémoire de travail. Le recouvrement entre les bases théoriques du déficit et la terminologie utilisée n'est pas parfaite. Certaines études défendent une mixité des erreurs (Botella & Ericksen, 1992 ; Maki, Couture, Frigen, Lien, 1997), mais attribuent le nom de conjonctions illusoires à la totalité des erreurs (eg. Botella & Ericksen, 1992). Les études cherchent généralement une explication unique au phénomène, et produisent des résultats ambivalents. Aucune étude ne propose explicitement d'assigner ces deux types d'erreurs, en fonction de leurs positions dans la distribution des intrusions ou en fonction de leurs propriétés. Nous proposons que les erreurs formées **offline** (ie. formées une fois que le traitement de la cible est accompli) soient **uniformément distribuées**, tandis que les erreurs formées **online** (ie. formées pendant le traitement de la cible) soient concentrées dans la **proximité temporelle de la cible** (voir aussi les barres bleues dans les parties B, C et D de la figure 64 p.99). Les erreurs online semblent être les plus nombreuses (voir chapitre précédant, et voir Boucart & coll., 2000), mais le ratio des erreurs online vs offline se modifie avec le degré de certitude du sujet (voir figure 63, ces distributions provenant de données réelles, extraites de Intraub, 1989, expérience 2).

di distribution des intrusions en fonction du degré de certitude de la réponse (Intraub, 1989, expérience 2)

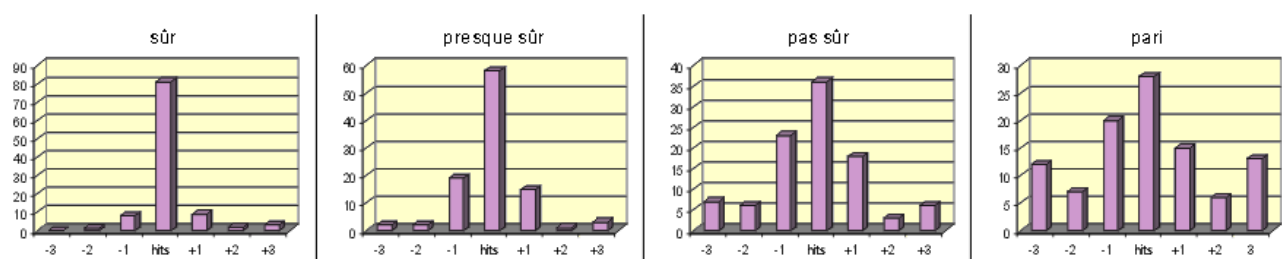


Figure 63 : modification de la distribution des réponses en fonction du degré de certitude du sujet (d'après Intraub, 1989). Moins le sujet est certain de sa réponse, plus la distribution des intrusions est éparse.

Malgré le caractère implicite de la formulation, la plupart des études se concentrent sur les intrusions émanant de distracteurs situés à proximité de la cible (juste après ou/et juste avant la cible). Par exemple, McLean & coll., (1982) ; Intraub, (1985, 1989) ; Botella & Ericksen (1992) ; Botella & coll., (1992) présentent des distributions des intrusions restreintes aux positions [-3 à +3] ou [-2 à +2]. Cette pratique était parfois liée à des compromis méthodologiques¹. Cependant, elle avalise la notion de conjonctions illusoire. Les erreurs sont interprétées comme une malformation de la représentation de la cible, durant la phase perceptive (erreurs online).

La première **théorie explicative** est celle du **filtre**, développée par l'équipe de Broadbent (Broadbent, 1977 ; McLean, Broadbent, & Broadbent, 1982 ; Gathercole & Broadbent, 1984, Broadbent & Broadbent, 1987). Cette vue est basée sur la nature itérative du système visuel, et prône la sérialité des différentes opérations de recherche. Les auteurs proposent par exemple que, lorsque l'on recherche visuellement une Volkswagen blanche, on commence par repérer les véhicules blancs apparaissant dans le champ visuel, puis on tente de décider si ces véhicules comportent les traits distinctifs d'une Volkswagen. Les deux caractéristiques de l'objet (couleur et forme), sont traitées **sériallement**, et la décision intervient à l'issue d'un processus itératif. Dans l'exemple, le processus sériel se réalise en deux étapes, ce qui a valu à cette théorie le nom de « modèle en deux étapes », mais des recherches plus complexes (eg. rechercher une Volkswagen blanche dont le conducteur est mon ami Oscar) peuvent admettre davantage de pas de traitements.

Le modèle de Broadbent & collègues, expliquant les erreurs en RSVP, est basé sur l'observation des intrusions '+1', considérant toutes les autres intrusions comme de l'erreur aléatoire. En effet, les premières études montrent une forte prépondérance des intrusions '+1'. Dans une première étape, le critère de réponse est recherché. Si le critère de réponse est découvert, alors seulement commence l'identification visuelle. Cette identification se réalise sur le matériel visuel en cours dans le système cognitif. Si la cible a déjà été remplacée par le stimulus suivant, c'est ce dernier qui sera rapporté par le sujet. Le mécanisme supposé par ces auteurs est bien résumé par la métaphore de Chun, (1997). Il s'agit d'attraper une balle au lancer (ie. la cible). Si le mouvement du bras est initié tard (si le critère de réponse est résolu tardivement), c'est la balle suivante qui sera capturée au vol.

Ce modèle rencontre de premières difficultés **lorsque certains protocoles exhibent de nombreuses intrusions '-1'** (McLean, Broadbent & Broadbent, 1982, cibles définies par une appartenance catégorielle orpheline dans la séquence). Une béquille est proposée sous la forme d'un **mécanisme additionnel**. Dans la plupart des cas, le sujet réalise la recherche de manière sérielle. Mais lorsque la recherche est plus difficile (eg. recherche catégorielle - nom de l'unique chiffre dans une séquence de lettres), le sujet déploie une stratégie **parallèle**.

Botella & Ericksen (1992) notent l'implausibilité du modèle de McLean, Broadbent & Broadbent, (1982). Le volet sériel de ce modèle prédit qu'aucun stimuli n'est traité hormis celui dont le traitement est ouvert par le critère de

¹ Par exemple, dans McLean, Broadbent & Broadbent, (1982), les caractéristiques des distracteurs sont répétés dans la séquence. Aucune intrusion provenant de plus de 3 rangs autour de la cible n'est reportée par les auteurs, et pour cause, en cas d'hésitation sur l'assignation temporelle des erreurs, l'expérimentateur choisit la position la plus proche de la cible, en vertu de la probabilité plus grande que les intrusions soient proximales.

réponse. Or, la plupart des sujets sont capables de rappeler librement plusieurs stimuli à l'issue d'une séquence RSVP. Plus généralement, remarquons qu'à de telles vitesses de défilement, inhiber les distracteurs est certainement impossible. Aux vitesses de défilement couramment utilisées en RSVP, le temps nécessaire pour lever une inhibition dépasserait sans doute le temps durant lequel le stimulus-cible est disponible. Or, à ces vitesses de défilement, le sujet réussit la tâche de manière correcte : on n'observe pas d'erreurs systématiques.

Ce modèle sériel (traitement du critère de réponse, puis traitement de la cible elle-même) étant réfuté, **Botella & Ericksen, (1992) proposent un modèle exclusivement parallèle¹** (simultanéité des traitements du critère de réponse et de la cible, puis ré-appariement temporel des deux événements). Cette théorie introduit la notion de **décision** (quelle cible est apparue en même temps que le critère de réponse), comme le montre bien la condition 1 de Botella & Ericksen (1992). Les auteurs observent une prédominance des intrusions '+1' dans la condition de base. Mais lorsque le formulaire de réponse ne propose plus le nom de la cible, un nombre important d'erreurs se reportent sur la position '-1', suggérant que **le sujet dispose de plusieurs candidats-cibles au moment de donner sa réponse**. L'opportunité de décider quelle était la cible parmi un choix de plusieurs candidats était aussi défendue par Intraub, (1981, 1985). L'auteur proposait un système de stockage à très court-terme (une sorte de mémoire iconique qu'Intraub appelle *short-term conceptual store*) ayant une capacité de 3 objets. Ce système de stockage était insuffisant pour maintenir les objets pendant plusieurs secondes jusqu'à la réponse verbale, mais il était le lieu où se formaient les conjonctions illusoires (**dissociations visuelles²**). La théorie d'**Intraub, (1981, 1985)** est rendue improbable par la forme de la distribution des intrusions. Lorsque les stimuli visuels sont des images ou des dessins d'objets, les intrusions proximales se différencient des intrusions distales sur une latence de 100 millisecondes avant l'affichage de la cible à 100 millisecondes après son remplacement. Ces délais suggèrent davantage que c'est le **traitement perceptif lui-même qui est impliqué, et non sa consolidation en mémoire**. Ce point sera plus largement exposé ci-après.

Les études précédentes n'exploitaient que les intrusions '+1' (voir figure 64, A). Les interprétations de Botella & Ericksen (1992) englobent dès lors les intrusions proximales ('+1' et '-1', voire, '+2' et '-2') et ne considèrent désormais comme de l'erreur que les intrusions distales. L'indice API proposé par **Botella & Ericksen (1991) suggère que ces intrusions proximales sont un phénomène unitaire**, et que les conditions expérimentales modulent la répartition pré-post des intrusions. La figure 64 partie B illustre ce postulat, **mais on peut lui opposer d'autres alternatives** (64, parties C & D), assumant une **complexité encore plus grande** des erreurs d'identification en RSVP, comme le suggèrent

¹ Par l'expérience 1 de leur étude, McLean, Broadbent & Broadbent (1982) suggèrent que le modèle parallèle n'est pas suffisant à lui seul. Présupposant que dans le modèle parallèle, les traitements des dimensions indépendantes occupent plus ou moins de temps, les auteurs prédisent que l'inversion des critères de réponse (chercher une forme d'une certaine couleur ou chercher la couleur d'une certaine forme) doit donner lieu à une inversion du pattern. Or il observe le même pattern de réponses dans les deux conditions. Botella & Ericksen (1992) opposent les résultats d'une simulation Monte-Carlo, montrant que les résultats de cette expérience peuvent être reproduits dans une modélisation parallèle du phénomène.

² Selon Intraub, (1985, 1989), les erreurs en RSVP sont formées par la migration du critère de réponse (un cadre noir au pourtour de l'image cible). Ce cadre est conceptuellement indépendant de l'image. Or la signification de l'image compte pour beaucoup dans l'assignation temporelle des objets contenus dans le système de stockage. Puisque la signification de l'image ne contraint pas l'attribution du cadre à une image plutôt qu'à une autre, ce cadre migre parfois sur l'un des deux autres images contenues dans le buffer.

également Botella & Ericksen, 1992). Ces mécanismes distincts pour les intrusions proximales sont supportés par les verbalisations des sujets. Deux types de verbalisations sont récurrentes lorsque le sujet hésite à donner sa réponse :

- « je vais dire le poisson mais c'était l'image juste avant » ; « je ne sais pas si c'est la fraise, ou le citron qui était situé juste après » : **difficulté d'attribution temporelle du critère de réponse**. Notons d'une part que le sujet éprouve des difficultés à se représenter l'ordre effectif des stimuli, comme le montre l'expérience 3 de Intraub (1985). Dans cette expérience, le sujet doit reporter l'image située juste après le cadre. Or, ils relatent fréquemment l'image située dans le cadre. D'autre part, nous supposons que lorsque la cible et le distracteurs accèdent à un traitement conjoint, mais qu'une (ou des) caractéristique(s) signe(nt) l'incongruité de l'association, alors il se construit deux représentations distinctes¹, et le sujet ignore à laquelle attribuer le critère de réponse.
- « c'était bleu..., » ; « c'était orienté à 45° » ; « c'était un animal dont la tête était dirigée vers la gauche », réponses en rapport avec des caractéristiques (eg. couleur, orientation globale) ou des constructions partielles d'objet (eg. direction de la tête). Ces verbalisations sont compatibles avec le modèle des conjonctions illusoire par Treisman & Gelade (1980). Elles suggèrent que **la représentation définitive et exclusive d'un unique objet n'est pas formée** au moment de ces réponses. Seule une collection éparses et incomplète de caractéristiques est disponible (les caractéristiques flottantes de Treisman, 1986, non encore solidarisées par la glue de l'attention). A partir de ces informations éparses, le sujet reconstruit une réponse (voir aussi paragraphe ci-après, la voie de pari sophistiqué de Botella, Barriopedro & Suero, 2001). Les expériences proposant un petit nombre de stimuli, présentés avant la session et répétés de nombreuses fois au cours du test se prêtent tout particulièrement à la formation de ces erreurs, puisque le sujet a appris la combinaison de caractéristiques propre à chaque image (eg. il sait que le pantalon est violet ; s'il se représente perceptivement la caractéristique violette, il en déduira qu'il s'agissait du pantalon). En revanche, ce comportement est moins probable dans les protocoles au cours desquels les stimulations ne sont que peu ou pas répétées.

¹ Cette construction de deux représentations distinctes est également rencontrée en attention auditive dans le paradigme d'écoute dichotique. Si une représentation unique ne peut être élaborée lorsque deux messages différents sont présentés auditivement (un à chaque oreille), alors deux représentations latéralisées sont construites (Cherry, 1953 ; Moray, 1969)

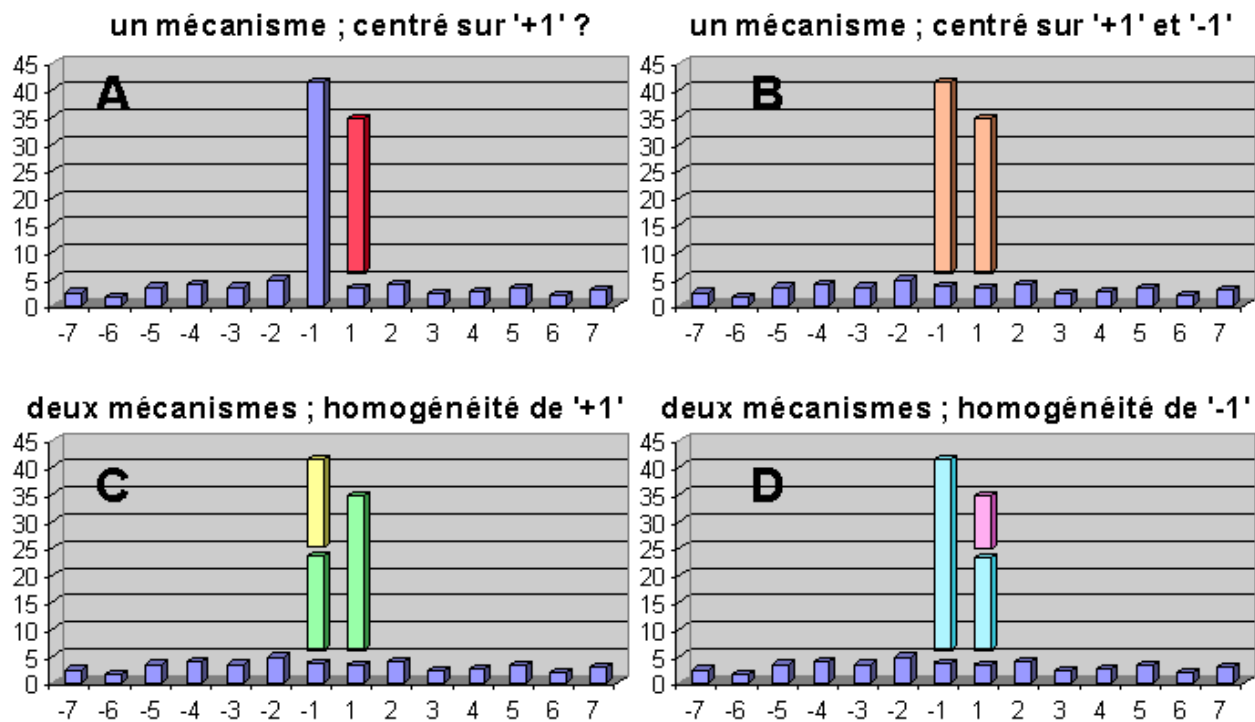


Figure 64 : hétérogénéité des conjonctions illusoires : quels postulats ? On considère usuellement les intrusions distales comme de l'erreur (le mécanisme n'est pas perceptif-attentionnel mais mnésique). En A, le postulat de McLean, Broadbent & Broadbent (1982), qui focalise son interprétation sur les intrusions '+1'. Ce modèle fut réfuté par les expériences exhibant de nombreuses intrusions '-1'. En B, le postulat de Botella & Ericksen, (1992). Ces auteurs considèrent un mécanisme unitaire pour les intrusions '+1' et '-1', (voire '-2' et '+2'), puisqu'ils proposent de calculer la modulation de la répartition pré-post par l'indice API. Nous verrons ultérieurement que deux alternatives (un mécanisme additionnel porté sur '-1', en C, et porté sur '+1' en D) doivent également être envisagés.

Botella, Barriopedro & Suero, (2001) proposent ensuite un modèle complet et spécifique pour expliquer les conjonctions illusoires en RSVP. Ce modèle propose que les caractéristiques de la cible soient extraites de manière pré-attentive en l'attente de la détection du critère de réponse. Suite à la détection du critère de réponse, le traitement perceptif peut adopter **deux voies différentes** (mutuellement exclusives) :

- La voie de traitement par **attention focalisée** (que l'auteur estime à 36% des réponses) : réponses basées sur le traitement complet du matériel présent dans l'environnement. Dans un second temps, les caractéristiques traitées de manière pré-attentives sont assemblées, sous couvert de l'attention focalisée.
- La voie de traitement par **pari sophistiqué**¹ (que l'auteur estime à 31% des réponses) : réponses basées sur un traitement partiel du matériel présent dans l'environnement. L'attention n'est pas complètement focalisée et un module permet la réponse sur la base d'informations partielles

Ce modèle repose sur le lien entre les conjonctions illusoires et l'attention, posé originellement par Treisman & Gelade (1980), mais nous avons vu que ce lien avait été affaibli (Wolfe & Cave, 1999). On remarque également que le

¹ Cette voie de traitement par pari sophistiquée est hautement plausible, car le système visuel est performant dans le traitement des stimuli dégradés, bruités, incomplets, occlus, comme en attestent de très nombreuses études (eg. Riesenhuber & Poggio, 2000)

modèle repose sur des **présupposés modulaires** (eg. les traitements de la dimension de réponse et de la cible sont encapsulés dans deux modules distincts), certainement justifiés par le souhait de l'auteur de réaliser une modélisation formelle du phénomène (simulation Monte-Carlo, formant la seconde partie de l'étude, Botella, Barriopedro & Suero, 2001). Cette modularité ne cherche sans doute pas à être biologiquement plausible (on peut raisonnablement assumer que la même voie de traitement visuelle véhicule le contenu des modules K et R, et supporte les traitements par attention focale et par pari sophistiqué). Si les modules permettent de construire les équations de la simulation, en revanche, du point de vue neurophysiologique, il est plus probable que ces deux modes de réponse soient situés sur un **continuum**. La réponse est une décision cognitive, basée sur une information plus ou moins complète, selon les conditions de son extraction (eg. difficulté d'identification générales liées à la rapidité du flux d'images et difficultés particulières liées à la configuration perceptive à interpréter) et selon les capacités de traitement du sujet (efficacité du traitement perceptif, qualité du réseau stockant les représentations sémantiques acquises, niveau d'attention...). Ce continuum s'intègre parfaitement à l'alternative théorique des conjonctions illusoires appelant la notion de codage grossier (Wolfe et Cave, 1999). Notons enfin que dans cette théorie, la conjonction illusoire désigne la confusion entre un distracteur et le critère de réponse, et non entre un distracteur et la cible.

Problèmes de terminologie, problèmes théoriques

Bien que les erreurs d'identification en RSVP soient couramment désignées comme des conjonctions illusoires (eg. Intraub, 1985, 1989 ; Botella, Barriopedro & Suero, 2001), il n'a pas encore été fermement établi que ces erreurs d'identifications répondent à la définition de conjonction illusoire telle qu'elle est envisagée dans les études de recherche visuelle spatiale. Cette circonspection terminologique est suscitée par deux problèmes : la mixité des erreurs d'identification, et le problème de la définition de la cible.

Il est nécessaire de **distinguer la réponse et le mécanisme ayant présidé à sa formation**. Au niveau de la réponse, nous obtenons ce qu'il convient d'appeler des intrusions¹ : le sujet est invité à donner une unique réponse, et quelquefois, en substitution du nom de la cible, il reporte le nom d'un distracteur. Au niveau du mécanisme, les choses sont moins simples. L'intrusion du nom du distracteur en remplacement du nom de la cible peut se produire parce que :

- Le traitement cognitif du distracteur a entièrement remplacé celui de la cible (intrusion - online)
- Le traitement cognitif de la cible et celui du distracteur se sont mêlés, et le sujet peut variablement décider de reporter le nom de la cible ou du distracteur (conjonctions illusoires - online), donnant quelquefois lieu à des réponses « correctes », et parfois à des réponses erronées
- Le sujet a oublié la cible et tente sa chance par une réponse aléatoire (pari aléatoire - offline)

¹ L'intrusion désigne le remplacement d'un objet par un autre. Il conviendrait d'appeler ces réponses « conjonctions illusoires » s'il était permis au sujet de mentionner deux stimuli, et si le sujet mentionnait la cible et un distracteur proche. Nommer les intrusions « conjonctions illusoires » procède de l'inférence de l'expérimentateur, qui attribue ces erreurs à un dysfonctionnement de l'intégration des caractéristiques.

- Lors de l'élaboration de la réponse, le sujet commet une erreur de rappel, car une interférence s'est produite en mémoire à court-terme (intrusion - offline)

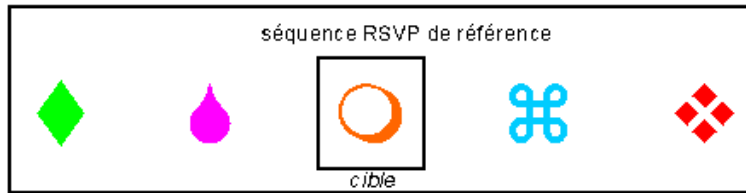
Comme postulé précédemment, on s'attend à ce que les **réponses offline** soient assez **uniformément distribuées** : toutes les images ayant eu accès au buffer visuel sont relativement équiprobables. Maki, Couture, Frigen & Lien, (1997), recherchant une covariation des fréquences d'erreurs avec le clignement attentionnel, observent effectivement que certaines classes d'erreurs sont uniformément réparties. En revanche, les **intrusions online** sont bornées par les **périodes de recouvrement temporel des traitements respectifs de la cible et du distracteur**. Pour des stimuli constitués d'images ou de dessins d'objets, on observe usuellement que les intrusions ne dépassent leur niveau de base (supposé être formé par les intrusions offline) qu'aux positions [-1 à +1] (Boucart, DeWisme et Wagemans, 2000, Boucart, Moroni, Fuentes, Belin, 1998 ; Chun, 1997, voir également chapitre précédent). A des défilements moyens de 8 à 10 images par seconde, cela représente une proximité temporelle de 100 à 125 millisecondes, soit justement le temps durant lequel se réalise l'analyse perceptive d'une image (eg. Thorpe, Fize & Marlot, 1996), à l'exclusion de la consolidation en mémoire (Tovee & Rolls, 1995, Keysers & Perrett, 2002, Subramaniam, Biederman & Madigan, 2000), qui occupe quelques 200 à 350 millisecondes supplémentaires. Ces intervalles temporels suggèrent, contrairement aux assertions de Intraub, (1985, 1989), que les **intrusions online procèdent d'un mécanisme perceptif pur, et non d'un problème de consolidation de la représentation perceptive** ou de promiscuité des objets dans un système de stockage (*visual short-term conceptual store*, Intraub, 1985). Dans le cas contraire, on observerait un étalement des intrusions proximales sur les trois positions situées avant et après la cible. Cet étalement n'apparaît dans aucune des études contrôlant correctement la tachistoscopie (ie. dont les latences prescrites expérimentalement sont conformes à l'exécution-machine ; et dont la présentation des images successives n'est pas saccadée, ce qui provoquerait des compressions et distensions des délais d'affichages, et donc une mauvaise assignation temporelle des intrusions).

On peut prédire que les erreurs formées online par intégration des caractéristiques de la cible et des distracteurs (conjonction illusoire) présentent un effet de la similarité cible/intruseur (l'intégration est sensible à la similarité, Treisman, 1996, Ivry & Prinzmetal, 1991). En revanche, les erreurs par substitution (online ou offline) pourraient être moins dépendantes des caractéristiques relatives de la cible et du distracteur, jusqu'à ne dépendre que des caractéristiques propres du distracteur. Nous chercherons à isoler ces deux configurations dans les études à suivre.

Nous voyons d'ores et déjà qu'une partie seulement des erreurs d'identification en RSVP semble être des conjonctions illusoires. Une étude amalgamant toutes les erreurs d'identification, au titre des conjonctions illusoires, sur-généralise le phénomène. Il n'est pas certain que la variance d'erreur (les intrusions provenant d'autres mécanismes que l'intégration) soit négligeable, ni même minoritaire.

Mais le malentendu provient aussi d'un **problème de définition de la « cible »**. Intraub (1985, 1989) évoque les conjonctions illusoires comme l'intégration fallacieuse du cadre situé au pourtour de la cible (critère de réponse) et d'un distracteur. Le cadre migre car il est conceptuellement indépendant de la cible. A l'incidence de cette étude, Intraub introduit une distinction sur laquelle il vaut la peine de s'arrêter. Elle suggère que **la cible** (l'objet dont le sujet

doit reporter le nom) et le **critère de réponse** (le cadre situé au pourtour de la cible) **sont deux objets distincts**. Cette distinction est légitime du point de vue de la psychophysiologie visuelle (ségrégation figure-fond ; théories de l'attention basées sur l'objet), mais dans la plupart des études (y compris Intraub, 1985, 1989), le postulat sous-jacent est que les « objets » sont **équivalents aux affichages entiers** (un plan d'affichage=un objet visuel).



définition de la cible	contenu de la représentation pour une réponse mentionnant le distracteur '+1'
<p>on considère la cible et le critère de réponse comme un tout unique</p>	<p>→ conjonction illusoire faible (weak illusory conjunction) : objet-distracteur intègre, seule la cible est altérée il n'en reste que le pourtour (le cadre noir)</p>
<p>on considère la cible et le critère de réponse comme deux objets visuels distincts</p>	<p>→ intrusion : objet distracteur intègre représenté seul</p> <p>→ conjonction illusoire forte : mélange de caractéristiques de la cible (couleur) et du distracteur (forme). L'intégrité des deux objets visuels a été altérée</p> <p>→ conjonction illusoire complète : des caractéristiques de la cible et du distracteur sont simultanément présentes dans la représentation (chimère). La réponse bifurque alternativement sur la cible ou sur l'intruseur</p>

Figure 65 : les problèmes de définition posés par les conjonctions illusoires en RSVP. Doit-on considérer que la cible est formée du groupe « objet à reporter et critère de réponse », où doit-on considérer que la cible est unilatéralement formée par l'objet à reporter. Cette question importe pour classifier les erreurs. Dans un cas, la représentation du distracteur suivant la cible sera interprétée comme une conjonction illusoire, dans l'autre cas, elle sera interprétée comme une intrusion. Les inférences sur le contenu de la représentation permettent de dissocier des conjonctions illusoires procédant d'un échange plus ou moins complet de caractéristiques.

Définir comme conjonction illusoire une erreur relevant du traitement conjoint d'une partie de la cible (le cadre) et d'un distracteur (l'image entière) est irréprochable tant que l'on considère le plan d'affichage entier comme l'objet-cible (voir figure 65). Nous avons qualifié ces **conjonctions illusoires** de « **faibles** », car, contrairement aux études menées avec des paradigmes de recherche spatiale, **l'intégrité de l'un des deux objets** liés (le distracteur '+1') **n'a pas été altérée**. Le second qualificatif utilisé par Intraub (1985) pour désigner ces erreurs d'identification, la « **dissociation visuelle** » semble mieux décrire ce phénomène. Une partie de la cible (le cadre noir) se dissocie du reste, et va se « coller » à l'image suivante. Notons qu'Intraub (1985, expérience 3) a établi qu'il ne s'agissait cependant pas d'un masquage par interruption dû au distracteur '+1'.

A contrario, si l'on considère que la cible et le critère de réponse constituent deux objets distincts (après Intraub, 1985, 1989), alors il convient de classifier les erreurs différemment. Le critère de réponse n'est plus considéré comme

un objet à représenter¹, il est un simple déclencheur du traitement cognitif de la cible (Intraub, 1985, le nomme marqueur visuel ou *visual marker*). Si le distracteur '+1' est représenté seul, il s'agit alors d'une **intrusion**, et non plus d'une conjonction illusoire (figure 65). Il ne s'agit plus de la recombinaison fallacieuse de caractéristiques flottantes de deux objets (comme c'est le cas dans les paradigmes de recherche visuelle spatiale), mais il s'agit de **l'association indue de deux objets complets, aucun n'ayant été altéré dans son intégrité**. Nous hésiterions à qualifier ces erreurs de conjonctions illusoires, non pas à cause de la signification du terme (il s'agit bien de la conjonction illusoire de deux objets), mais en regard de l'histoire de ce concept (les conjonctions illusoires issues des paradigmes de recherche visuelle spatiale ont toujours impliqué une atteinte de l'intégrité des objets, puisque les erreurs sont formées avant que les caractéristiques ne soient combinées correctement entre-elles).

Considérant malgré tout qu'il s'agit d'une conjonction illusoire, Intraub (1985, p.440) évoque les différences par rapport aux conjonctions illusoires obtenues dans les paradigmes de recherche spatiale : « [...] *it should be noted that these results neither support, nor refute Treisman & Gelade's (1980) position that certain classes of features (e.g., color and form) are initially registered independently in the visual system and are conjoined by focal attention. The illusory conjunctions obtained in the present research certainly seem to occur at a later stage of processing than initial sensory registration. They seem to be based on meaning, expectation and identification time, and they support the view that a very short-term memory store is involved. These types of illusory conjunctions [...] show that a relatively high level of analysis mediates some types of visual dissociation* ».

Ce point de vue n'exclut pas l'existence de conjonctions illusoires semblables aux paradigmes de recherche visuelle spatiale en RSVP. Les conjonctions illusoires seront avérées si la représentation associe des caractéristiques de la cible et du distracteur, comme c'est le cas dans les deux derniers exemples présentés à la figure 65. Cette possibilité de **conjonctions illusoires complexes** est déniée par Intraub, (1985), mais la logique à l'œuvre dans ses recherches était d'interpréter unitairement la totalité du phénomène. Nous postulons que des conjonctions illusoires complexes sont disséminées dans le corpus d'erreurs issues des paradigmes RSVP, et que des règles peuvent être découvertes, qui permettraient de les isoler des autres erreurs d'identification.

Cette pratique consistant à déployer un effort de recherche pour un événement supportant très peu de variance est assez inhabituelle en Psychologie. Du point de vue épistémologique, il ne nous semble pas insensé de consacrer des recherches à un événement mineur et difficile à filtrer. C'est bien cette même logique qui permet d'isoler un principe actif dans une plante pour réaliser un médicament, ou qui permet de collecter un élément radioactif dans un minéral. Le botaniste et le géologue n'analysent pas seulement la matière principale qui compose la plante ou le minéral.

¹ Deux différences sont notables quant-au sort des deux objets « cadre » et « cible ». dans les expériences d'Intraub. Premièrement, le cadre est à détecter, tandis que la cible est à identifier. Corollairement, le traitement du cadre, dont les caractéristiques sont connues à l'avance, est dirigé par des composantes top-down, tandis que le traitement de la cible, imprévisible, est dirigé par les composantes bottom-up. Intraub assume que le traitement de ce critère de réponse est plus long (Intraub, 1985, p.433), en raison de sa discordance conceptuelle avec l'image. Il nous semble plus probable que ce traitement préparé par les attentes du sujet et moins complexe soit réalisé en moins de temps. Ces caractéristiques sont entièrement connues à l'avance, et on peut considérer qu'il fonctionne plus comme un déclencheur du traitement visuel que comme un objet à représenter, c'est-à-dire duquel donner une effigie cognitive d'après le flux bottom-up.

Conjonctions illusoirs complexes en RSVP : à la recherche des chimères

La plupart des études en RSVP postulent que la conjonction illusoire est formée par association induite du critère de réponse et d'un distracteur (conjonctions illusoirs faibles dans notre proposition figure 65). Aussi, les conjonctions illusoirs sont systématiques (hormis les intrusions offline, toutes les erreurs confondent le critère de réponse et le distracteur). En conséquence, trouver des règles de ségrégation entre des conjonctions illusoirs et d'autres formes d'intrusions ne s'avère pas pertinent. Notre visée est d'établir l'existence de conjonctions illusoirs plus complexes (des chimères). Aussi, nous chercherons à dégager des règles permettant classifier les erreurs.

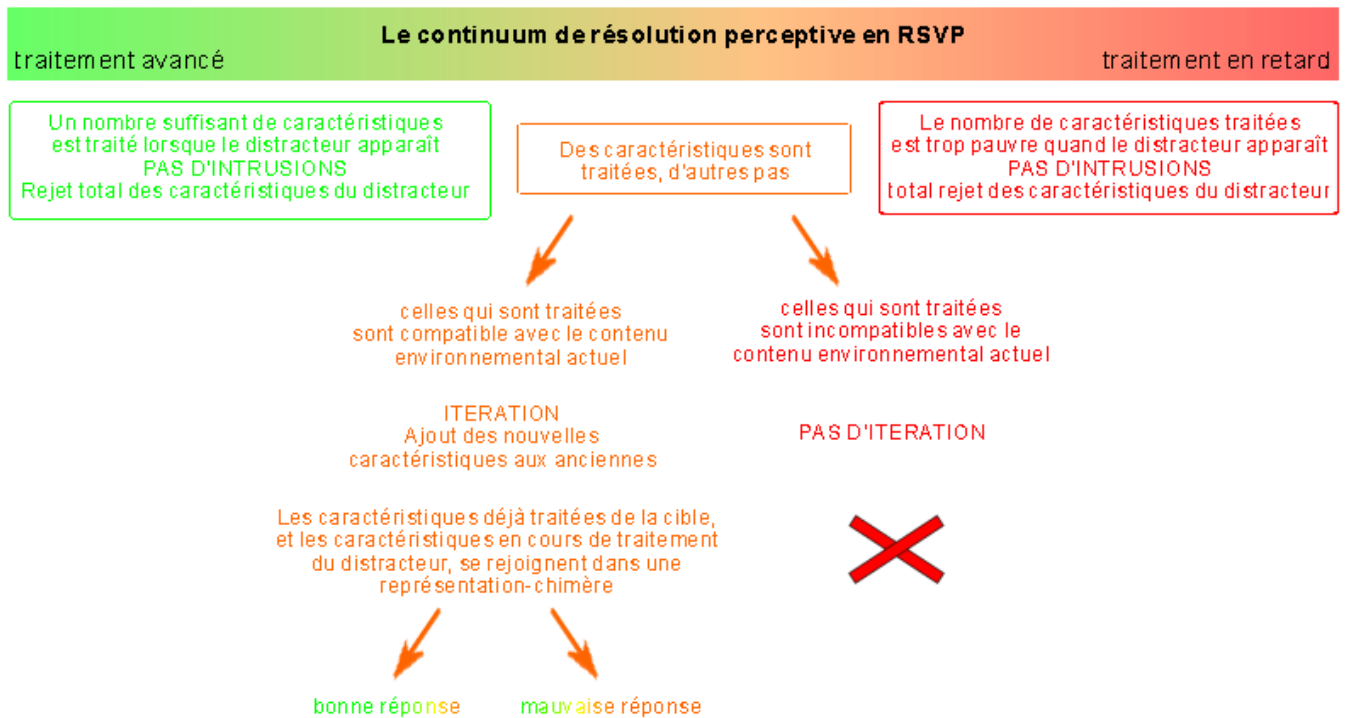


figure 66 : état d'avancée du traitement perceptif et conjonctions illusoirs : modèle hypothétique. Ce modèle basé sur un continuum prédit que le traitement perceptif peut avorter à toutes les étapes du traitement.

Nous postulons à priori que les erreurs dépendent de l'état d'avancée du traitement visuel (voir figure 66). Le modèle est fondé sur la nature itérative du fonctionnement visuel. Le système visuel prélève de l'information, par vagues de traitement successives, jusqu'à l'identification de l'objet d'attention. Prenons l'exemple d'une possible intrusion '+1'. Si le traitement est suffisamment avancé, alors la représentation sera invulnérable à ce distracteur '+1', car trop de caractéristiques divergentes signifieront l'incongruence d'une itération (on observe effectivement un déclin rapide de la fréquence des intrusions proximales). De même, si le traitement est trop peu élaboré, alors le sujet ne peut pas créer une représentation. Il produira une réponse d'indétermination (« je ne sais pas »). Entre ces deux extrêmes, se trouvent les traitements intermédiaires. Le traitement de la cible a permis la résolution de certaines caractéristiques, mais pas de certaines autres. Dans cette situation, deux cas de figures sont possibles :

- Soit l'image qui se présente ensuite est compatible avec le contenu en cours de traitement, et les caractéristiques du distracteur seront ajoutées à celles de la cible, formant une **chimère cognitive** (conjonction illusoire sur deux objets réels).
- Soit il y a incompatibilité, et le sujet sera dans l'impossibilité de former une représentation (ce point étant à réévaluer, il se peut également que lorsque la cible a entièrement échoué, le distracteur la remplace entièrement).

Ce modèle engendre la prédiction que les erreurs d'indétermination croissent régulièrement avec la vitesse de défilement, alors que les conjonctions illusoires croissent à partir d'un seuil critique puis décroissent lorsque le temps disponible ne permet plus d'amorcer la représentation de la cible. Cette prédiction semble contraire aux résultats de Keyzers & Perrett, (2002, voir figure 67), mais les auteurs de cette étude utilisent une procédure de choix forcé, qui ne permet pas de discerner les indéterminations des intrusions. Nous proposons que la performance se comporte comme reporté à la figure 68.

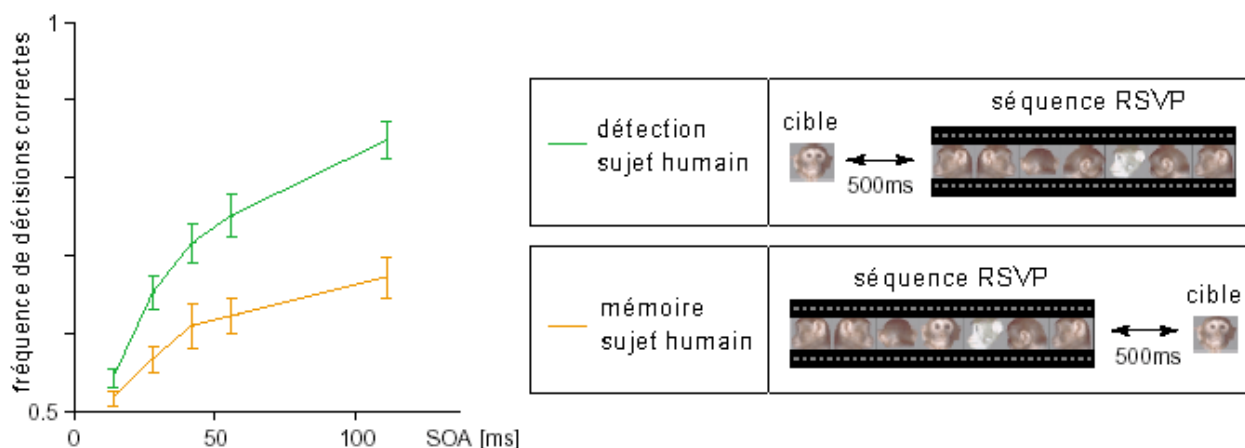


Figure 67 : performances en RSVP chez l'humain dans deux tâches de détection (spécification de la cible avant la présentation de la séquence) et de mémoire (spécification de la cible après la présentation de la séquence) en fonction de la durée de présentation des stimuli, par Keyzers, Xiao, Földiák & Perrett (2001). Les durées de présentation des images étaient de 14, 28, 42, 56, et 111 millisecondes. Le sujet devait dire si oui ou non, il avait vu la cible dans la séquence. La ligne de hasard est à 0,5. (modifié de Keyzers & Perrett, 2002).

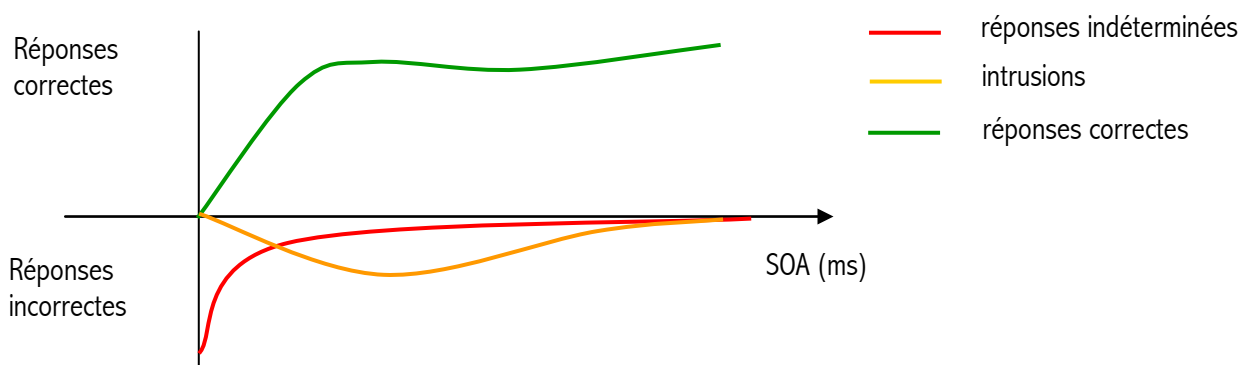


Figure 68 : prédictions sur le comportement des erreurs d'identification et des réponses exactes en fonction de la vitesse de défilement, étendant les observations de Keyzers & Perrett (2002)

Les caractéristiques visuelles de la cible et du distracteur sont donc associées, lorsque les éléments déjà traités de la cible et ceux du distracteurs sont compatibles. Cette hypothèse suscite le modèle suivant (voir figure 69). Compte tenu du parallélisme des étapes visuelles précoces, on doit envisager que les durées de résolution des différentes

caractéristiques élémentaires sont indépendantes entre-elles, et dépendent essentiellement de la configuration perceptive étudiée. Si la couleur de la cible est résolue précocement, et qu'un objet de même couleur se présente dans l'environnement, alors le traitement du distracteur va s'associer au traitement de la cible, et créer une chimère (représentation cognitive constituée des contributions partielles de deux objets de l'environnement). Les intrusions '+1' s'expliquent donc par l'incrustation de caractéristiques du distracteur '+1' dans le traitement de la cible.

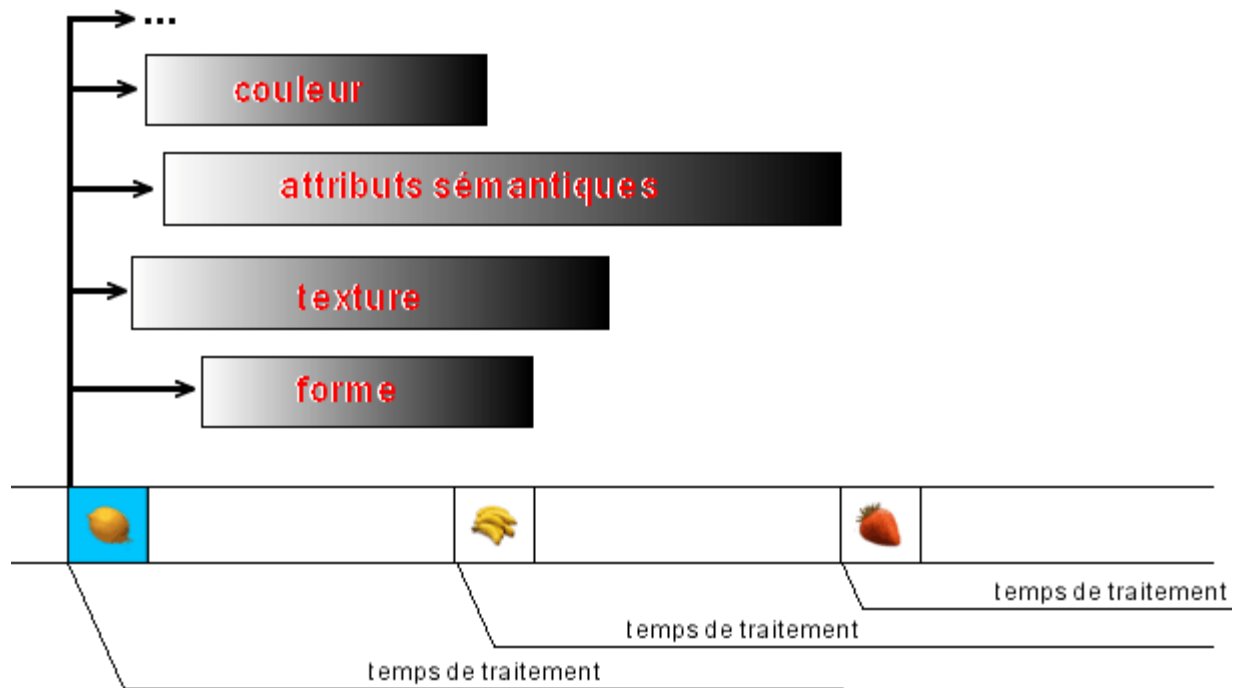


Figure 69 : mécanisme supposé des conjonctions illusoires en RSVP. Lorsque la cible (le citron) apparaît, le traitement cognitif s'initie, par le biais de différentes routines parallèles. Le traitement de la cible n'est pas terminé lorsque le distracteur suivant apparaît, mais la couleur est déjà résolue. Si le distracteur est d'une couleur compatible avec celle de la cible (c'est le cas des bananes), alors son traitement est adjoint à celui de la cible, par le mécanisme d'itération perceptive. Le résultat est la fondation d'une représentation hybride entre la cible et le distracteur, une chimère. Cette chimère est une version complexe des conjonctions illusoires. Remarquons que si les bananes étaient apparues après la résolution de la forme de la cible, il y aurait eu incompatibilité, et l'itération aurait été rejetée.

Il se produit également des intrusions par le distracteur précédant la cible. Ces intrusions peuvent-elles également comporter des conjonctions illusoires complexes ? Ce cas de figure serait explicable par le mode de résolution de la tâche du sujet. Pour réussir la tâche, le sujet est contraint à laisser accéder aux traitements, les distracteurs précédant la cible. Puisque tous les distracteurs précédant la cible initient un traitement perceptif, des conjonctions illusoires sont également possibles au rang précédent la cible. **Les caractéristiques du distracteur '-1', dont le traitement est avancé, se mêlent aux caractéristiques de la cible** (le concept de compétition développé par Keysers & Perrett, 2002, illustre bien cette orientation bipolaire des conjonctions illusoires). Il restera à déterminer si ce sont les caractéristiques résolues du distracteurs qui permettent ou refusent l'itération sur la cible, ou si au contraire, la détection du critère de réponse atténue le traitements des caractéristiques incongruentes du distracteur en préservant celui des caractéristiques congruentes. Quoiqu'il en soit, nous pouvons prédire que d'éventuelles conjonctions illusoires '-1' et '+1' ont un comportement asymétrique, c'est-à-dire que la relation cible/distracteur '+1' ne sera pas égale à la relation cible/distracteur '-1'. En effet, la détection du critère de réponse se produit après que le traitement du

distracteur soit initié (Botella & Ericksen, 1992 ; Botella, Barriopedro, Suero, 2001). Les stades de traitements respectifs de la cible et du distracteurs sont donc différents en intrusion pré-cible et post-cible. Sous couvert de cette hypothèse parallèle du traitement du critère de réponse et de l'objet, on s'attend à ce que l'organisation des conjonctions illusoires soit différente des deux cotés de la cible.

Résumé

Nous avons présenté ces expériences préliminaires dans le but d'introduire le lien entre erreurs d'identifications en RSVP et intégration perceptive. Ce lien à approfondir dans la série d'expérience suivante justifiera l'emploi du terme de « conjonctions illusoires » pour qualifier ces erreurs. De ces expériences préliminaires, nous retenons qu'elles ne résultent pas (ou pas uniquement) d'un mécanisme de pari. Nous suspectons que la similarité des dimensions physiques de la cible et du distracteur participe à la formation des erreurs.

Concernant la distribution des erreurs, nous admettons sur des bases théoriques que les erreurs distales concernent un mécanisme de récupération de l'identité de la cible en mémoire de travail. Ce mécanisme n'est pas pertinent pour la présente étude. En revanche, il demeure que les intrusions proximales sont les possibles conséquences de l'intégration perceptive. Nous questionnons la répartition pré-post des intrusions : ces deux types d'intrusions, impliquant les distracteurs situés respectivement avant ou après la cible sont-elles co-dépendantes ou indépendantes ?

Les hypothèses théoriques formulées pp.90-107 vont être éprouvées dans la section suivante, à travers la réalisation d'une série de 5 expériences.

Opérationnalisation

Une partie de ce travail a consisté à faire progresser la méthodologie sous-jacente à cette étude en RSVP. Nous présenterons tout d'abord ces aménagements du protocole par rapport aux séries d'expériences préliminaires A et B.

Ensuite, nous conduirons une série de cinq expériences destinées à justifier l'existence de conjonctions illusoires en RSVP. L'expérience 1 cherchera à comprendre l'effet de modulations temporelles sur les intrusions. L'expérience 2 invoquera les effets d'exercice pour démontrer une dissociation des intrusions proximales. L'expérience 3 testera l'hypothèse d'Intraub (1985, 1989) faisant des erreurs d'identification en RSVP des conjonctions illusoires faibles ou dissociations visuelles (voir p.102-103). L'expérience 4 approfondira les deux concepts d'intégration et d'interruption. Enfin, l'expérience 5 permettra la sélection des conjonctions illusoires parmi les simples intrusions. Elle permettra de tester le lien existant entre la fréquence des conjonctions illusoires et la similarité physique de la cible et du distracteur.

4. Des conjonctions illusoires ?

Commun à l'ensemble des expériences

Aménagements du protocole RSVP

Des aménagements ont été réalisés pour améliorer le recueil des données dans les prochains protocoles. Nous avons amélioré la tachistoscopie sous-jacente aux expériences. La conduite de ce projet a rendu ces précautions nécessaires, car la validité des résultats sera cruciale pour s'opposer au problème de la faiblesse des effectifs d'observation (une expérience RSVP ne peut être administrée pendant plus d'une demi-heure consécutive à un sujet donné). Les détails de nos motivations et choix méthodologiques sont reportés en Annexe XV.

L'objectif premier est d'obtenir des effectifs de substitutions [cible ; distracteurs] suffisamment nombreux pour pouvoir analyser les matrices de confusions¹ dans le détail. Il s'agit donc de concentrer les observations sur un petit nombre de stimuli. Tout d'abord, nous avons réduit la **longueur de la séquence** (et donc le nombre d'images dans la catégorie) de 15 à **10 images**. Cette mesure réduit la dispersion des observations. Dans la même logique, nous avons réduit le nombre de catégories étudiées de quatre à une ou deux (une catégorie naturelle et une catégorie artefactuelle). Nous redoutions à l'origine la lassitude des sujets sur une tâche monotone, mais les sessions courtes, doublées de la difficulté de la tâche, empêche le sujet de se démobiliser.

Nous avons remplacé le **critère de réponse** initial (un carré bleu) par un **carré gris** (plus neutre dans sa relation colorimétrique avec la cible). Nous avons eu recours à un **fond noir**. Le fond noir présente l'avantage de contraindre l'orientation de la luminance des stimuli de manière unipolaire (il n'existe pas des images plus claires que le fond et des images plus sombres ; toutes les images sont nécessairement plus claires). Cependant, il nécessite un dispositif particulier. Avec un équipement classique, si un sujet arrive dans la salle d'expérience en portant un pull clair, un reflet apparaît au centre de l'écran et masque l'affichage. Aussi, l'instauration du fond noir s'accompagnait de la conduite de l'expérience en salle obscure avec un affichage au travers de lunettes de réalité virtuelle. Ces mêmes lunettes suppriment en grande partie les mouvements relatifs des yeux du sujet par rapport au dispositif d'affichage, et favorisent ainsi considérablement la fidélité des résultats.

Nous avons varié plus systématiquement la **position de la cible**. Elle apparaît désormais aux positions **3 à 8** de la séquence de 10 images. La distribution de la position des cibles devient ainsi **symétrique**. Nous avons cependant évité

¹ Les matrices de confusion sont des tableaux de fréquence croisant les stimuli intruseurs et les stimuli intrusés. Un exemple est présenté p.80.

d'inclure les deux premières et les deux dernières positions, car le mécanisme serait alors différent, impliquant également les effets mnésiques de primauté et de récence, et brouillant ainsi les résultats par deux sources de variations.

Nous avons suspecté, dans la première série d'expériences, que le regroupement chromatique des stimuli d'une catégorie donnée accapare la majeure partie des conjonctions illusoires. Si 3 ou 4 images sont très similaires en couleurs, il se produit de nombreuses conjonctions illusoires amorcées par la couleur, et la manifestation d'autres sortes de conjonctions illusoires (sur la forme, sur la luminosité, sur la texture ?) est entravée. Dans certaines catégories, telles que les fruits ou les légumes, un grand nombre de représentants sont de couleurs verte ou rouge. Lors de la sélection des stimuli, il se crée ainsi des clusters de stimuli similaires sur la couleur, qui interfèrent ensuite avec la logique exploratoire développée dans la présente thèse. Pour compenser ce problème, nous avons réalisé une **assignation chromatique** : les dix stimuli utilisés seront un exemplaire rouge, orange, jaune, vert, bleu indigo, violet, rose, grège et brun, voir annexe XVI. Cette opération minimisera peut-être le nombre d'intrusions, car la couleur est un critère très puissant de discrimination visuelle, comme l'attestent les études sur les asymétries de recherche visuelle (eg. Treisman & Gromican, 1988). Elle facilite le traitement en améliorant la segmentation de l'image et favorise l'encodage (Gegenfurtner & Rieger, 2000, Tanaka, Weiskopf & Williams, 2001). En outre, l'information de couleur survit à l'inhibition endogène imposée par le clignement attentionnel (Shapiro, Raymond & Taylor, 1993). Atténuer la similarité chromatique existant entre quelques images nous permettra de voir si d'autres primitives visuelles peuvent amorcer une conjonction illusoire. Enfin, nous avons systématisé le recours à des **mesures répétées**.

Matériel & méthodes

Conditions de test

Les tests ont été administrés dans une salle borgne et sans éclairage principal. Le sujet était confortablement installé sur une chaise, et équipé de lunettes de réalité virtuelle (dispositif peu immersif Sony Glasstron PLM-S700E), voir figure 70). L'expérimentateur était présent dans la salle (dans le dos du sujet¹), et saisissait les réponses du sujet sur un clavier, éclairé par une ampoule rouge, et situé hors du champ visuel du sujet. Sauf mention contraire, les passations duraient entre 15 et 30 minutes, installation et consignes comprises. La durée des épreuves est contrainte par le caractère pénible de la tâche. Le défilement rapide des images et son caractère continu occasionne des gênes oculaires et des maux de tête chez certains sujets. Les sessions étaient systématiquement précédées d'un entraînement, comportant la dénomination de chaque image et l'administration de 12 essais d'entraînement.

¹ Le sujet en cours de test est usuellement très désireux de ratification sur sa performance. Aussi, il a tendance à orienter son attention visuelle vers l'expérimentateur lorsque celui-ci est à portée du regard, et il latéralise ainsi son profil de performance (comme le montrent des protocoles de recherche visuelle, Tognoli, Toniolo, et Kop, données non publiées). Pour juguler ce biais, nous avons retenu à la fois un dispositif de réalité virtuelle, et une position de l'expérimentateur hors de la vue du sujet.

Les sujets étaient des étudiants ou des membres du personnel de plusieurs institutions (Universités Nancy I, Nancy 2 et USTL, Hôpital Roger Salengro et clinique Fontan à Lille). Ils étaient volontaires pour la passation des épreuves, avait une vision normale ou corrigée à la normale, et ne reportaient pas d'histoire d'accident neurologique.



Figure 70 : lunettes de réalité virtuelle Sony Glasstron utilisées pour les expériences. Le dispositif supprime les mouvements de la tête relativement à l'image et limite les distracteurs visuels provenant de l'environnement dans lequel est réalisé le test.

Hardware

Les tests ont été administrés sur un ordinateur Pentium celeron, cadencé à 300 MHz, doté d'une carte graphique à 4Mo de mémoire, (ATI rage Pro).

Software

Les tests ont été développés sous ERTS (Beringer, version 3.26), pour s'assurer de la réalisation des affichages tachistoscopiques et administrés en rafraîchissement 70 Hz, en résolution 640 x 480 pixels.

Matériel et méthodes

Hormis pour l'expérience 1 (pp.112-119) manipulant ce facteur, la vitesse de défilement était de 8,8 images/s (soit 114 ms/image, répartis en 71 ms d'exposition à l'image et 43 ms d'intervalle inter-stimuli). Les stimuli sont reportés en annexe XVI.

Expérience 1 : Modulations temporelles

Les **erreurs d'identification visuelle** confondant un objet présent dans l'environnement avec un autre sont rares chez le sujet normal. On les obtient en laboratoire en appliquant un **stress temporel** au sujet, dans des paradigmes de recherche visuelle spatiale (*brief multi-element displays*) ou temporelle (RSVP). De cette façon, il ne lui est plus possible de réaliser des fixations individualisées (Treisman, 1996) ou d'itérer pour lever une ambiguïté perceptive, et certaines tâches perceptives échouent, parfois lorsque le sujet est très confiant en sa réussite perceptive (Intraub, 1989). Alors même que le temps est un paramètre clef dans la formation des intrusions, peu d'études le manipulent expérimentalement pour comprendre, dans la structure temporelle, quels sont les déterminants de la formation des erreurs.

Des **manipulations temporelles de fort contraste** ont été appliquées dans les premières études pour démontrer les points communs et les différences entre les performances visuelles en présentation normale et en présentation rapide (Potter & Levy, 1969 ; Intraub, 1981). De ces études, il a été déduit que **l'intervalle entre les stimuli** (ISI : *inter-stimulus-interval*) **est le principal facteur occasionnant les erreurs d'identification** en RSVP. Ces expériences étaient alors réalisées avec un projecteur cinématographique 16 mm. Elles montraient que les performances des sujets restaient excellentes lorsque l'image était brièvement présentée (eg. 1/3 du temps nécessaire pour réaliser une fixation oculaire, soit 114 ms), pour peu qu'un intervalle suffisant sépare les images successives. En revanche, dès lors que l'intervalle entre les stimuli était « supprimé »¹, on observait un fort **masquage** par les images situées juste avant ou juste après la cible. Notons que lorsque l'intervalle inter-stimuli est supprimé, la pression temporelle sur les traitements est plus forte (à temps d'exposition constant des stimuli, un ISI plus grand se traduit par une vitesse de défilement plus faible). Une **manipulation temporelles de fin contraste** a été réalisée par Botella & Ericksen (1991), mais la **variable indépendante était encore une fois l'ISI**. Cette étude conclut à une antériorisation de la distribution des intrusions lorsque l'ISI est augmenté, et à une diminution concomitante des erreurs. Encore une fois, l'étude ne contrôlait pas la covariation de l'ISI avec la pression temporelle globale.

A notre connaissance, aucune expérience n'a permis de **dissocier les différents paramètres temporels** se manifestant en RSVP. Il n'est alors pas possible de déterminer si les intrusions observées sont le fait de la **proximité des deux images**, de la **durée d'exposition au stimuli**, ou de la **pression temporelle**.

En outre, certains des délais utilisés dans ces séries d'expériences (Potter & Levy, 1969 ; Intraub, 1981) sont trop importants pour que l'on puisse considérer être en situation RSVP. Ces recherches visaient tout particulièrement à établir qu'une identification performante est obtenue en dessous du délai de fixation oculaire sous certaines conditions (absence de masquage). Ce résultat permettait d'établir que les images présentées en RSVP sont complètement

¹ Bien que ces auteurs annoncent l'utilisation d'ISI nuls dans certaines conditions expérimentales, le fonctionnement du projecteur cinématographique devait cependant introduire un délai sans image (*blank field*) lors de la transition de deux plans successifs (très probablement, une pale passait devant l'obturateur lors de l'avancement du film). Ce délai de quelques millisecondes était considéré comme négligeable à l'époque, mais doit être considéré comme important à la lumière des théories actuelles du fonctionnement visuo-attentionnel.

comprises, mais ne sont pas maintenues en mémoire. Depuis lors, des recherches ont montré qu'une vingtaine de millisecondes sont suffisantes pour réaliser des tâches perceptives lorsque la cible est isolée (eg. Thorpe, Fize et Marlot, 1996). Et lorsque la cible est insérée dans un flux d'images indépendantes (RSVP), les phénomènes de cécité à la répétition et d'amorçage ont permis d'établir qu'elle accède à un niveau de traitement sémantique, même lorsque l'attention est focalisée sur une cible préalable et que son identification échoue (Shapiro, Driver, Ward, Sorensen, 1997 ; Maki, Frigen & Paulson, 1997).

Stratégie

L'objectif de cette expérience était d'observer l'effet des paramètres temporels de défilement des images sur les intrusions. Nous avons réalisé une expérience en modulant indépendamment et conjointement les temps de présentation et les intervalles inter-stimuli. En outre, nous avons distingué les erreurs d'indétermination (« je ne sais pas ») et les intrusions, pour permettre de faire la part de ce qui est lié à la difficulté de la tâche et de ce qui est lié aux intrusions proprement-dit. La terminologie des paramètres temporels est reportée à la figure 71.

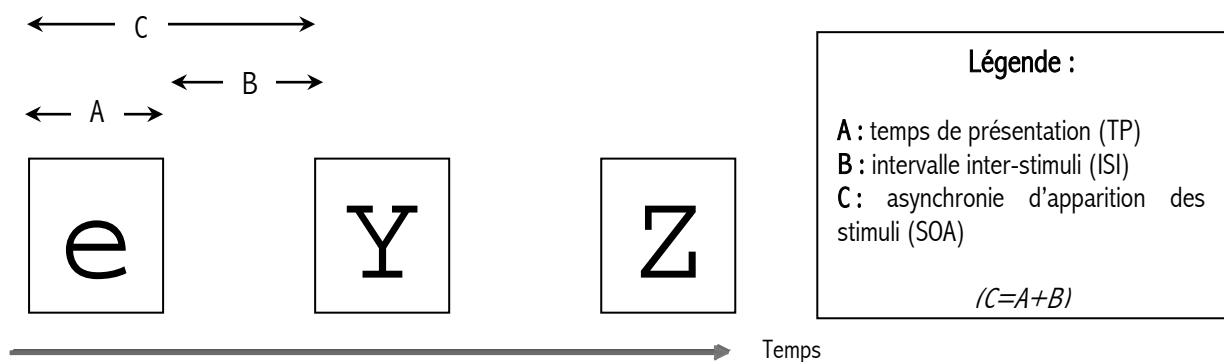


Figure 71 : terminologie des paramètres temporels en RSVP

Sujets

Les sujets étaient 23 adultes volontaires pour l'expérience (11 femmes, 12 hommes), âgés de 20 à 49 ans (moyenne d'âge : 28 ans). Les sujets ont reporté une vision normale, ou corrigée à la normale.

Matériel et méthodes

Les sujets ont complété une session RSVP de 168 essais, durant environ 25 minutes, consignes et entraînement compris. Les essais variaient quant-à la durée de présentation des images (TP) et quant-à la durée des intervalles inter-stimuli (ISI). Nous avons retenu un plan expérimental en mesures répétées avec trois vitesses de défilement des images :

- L_5_4, vitesse perçue la plus lente, 5 trames d'affichage TP et 4 trames d'ISI
- M_5_3, vitesse perçue comme intermédiaire, 5 trames d'affichage TP et 3 trames d'ISI
- R_4_4, vitesse perçue comme la plus rapide, 4 trames d'affichage TP et 4 trames d'ISI).

Les paramètres de ces trois vitesses sont présentés à la table 6. Ces trois conditions permettent de tester les effets suivants :

- les deux conditions L_5_4 et M_5_3 maintiennent constant le temps d'exposition à l'image. La comparaison entre ces deux conditions permet de mesurer l'effet conjoint de la durée de l'ISI et de la pression temporelle
- les deux conditions L_5_4 et R_4_4 maintiennent constant l'ISI. La comparaison entre ces deux conditions permet de mesurer l'effet conjoint de la durée d'exposition à l'image et de la pression temporelle
- enfin, les deux conditions M_5_3 et R_4_4 maintiennent constante la pression temporelle. La comparaison entre ces deux conditions permet de mesurer l'effet conjoint de la durée d'exposition à l'image et de la durée de l'ISI

trames	1	2	3	4	5	6	7	8	9
L_5_4	TP : 71 ms					ISI : 57 ms			
M_5_3	TP : 71 ms					ISI : 43 ms			
R_4_4	TP : 57 ms					ISI : 57 ms			

Table 6 : latences utilisées pour les trois conditions expérimentales. Dans la condition L_5_4 (perçue comme la plus lente), les images sont présentées pendant 5 trames VGA (soit 71 ms), et sont séparées par un ISI de 4 trames VGA (57 ms). La condition M_5_3 (perçue comme intermédiaire) utilise également un TP de 5 trames (71 ms) mais un ISI de 3 trames (43 ms). Enfin, la condition R_4_4 (perçue comme la plus rapide) utilise un TP et un ISI de 4 trames VGA (57 ms). Notons que la vitesse réelle de défilement est identique à la condition précédente.

La présentation des essais par bloc de même vitesse de défilement était nécessaire. En effet, la transition aléatoire entre des essais lents et rapides risquait de déstabiliser les sujets et compromettre la validité des résultats. Cependant, un design expérimental en trois blocs (un par condition) aurait posé problème. Les blocs d'essais relégués à la fin de la session expérimentale auraient présenté, outre la variation expérimentale, des effets d'apprentissage. Pour juguler les effets de l'ordre d'administration, nous avons fractionné chaque condition expérimentale en 7 blocs de 7 essais (blocs A à G, voir figure 72), et proposé une alternance fréquente des conditions expérimentales. Les mesures répétées étaient disséminées dans la session, avec au minimum cinq blocs d'écart entre deux répétitions du même bloc. Par ailleurs, trois variantes de l'expérience ont été utilisées, qui alternaient l'ordre d'administration des blocs. Le montage expérimental est reporté à la figure 72.

Version A

A	C	E	B	D	F	C	E	G	D	F	A	E	G	B	F	A	C	G	B	D
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Version B

C	E	G	D	F	A	E	G	B	F	A	C	G	B	D	A	C	E	B	D	F
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Version C

E	G	B	F	A	C	G	B	D	A	C	E	B	D	F	C	E	G	D	F	A
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Figure 72 : ordre des blocs dans les trois versions de l'expérience. Chaque bloc comporte 7 essais. Les blocs des cases blanches sont présentés à la plus faible vitesse, les blocs des cases gris clairs sont présentés à une vitesse intermédiaire et les blocs des cases gris foncé sont présentés à la vitesse la plus rapide. L'écart minimal entre deux blocs similaires est de 5 blocs. L'ordre des essais dans les blocs est randomisé pour davantage limiter l'effet de la familiarité.

Résultats

Fréquences des erreurs

Trois types de réponses ont été obtenues : des réponses correctes, des intrusions (nom d'une autre image de la séquence) et des réponses indéterminées (je ne sais pas). La fréquence de ces trois types de réponses est reportée à la figure 73.

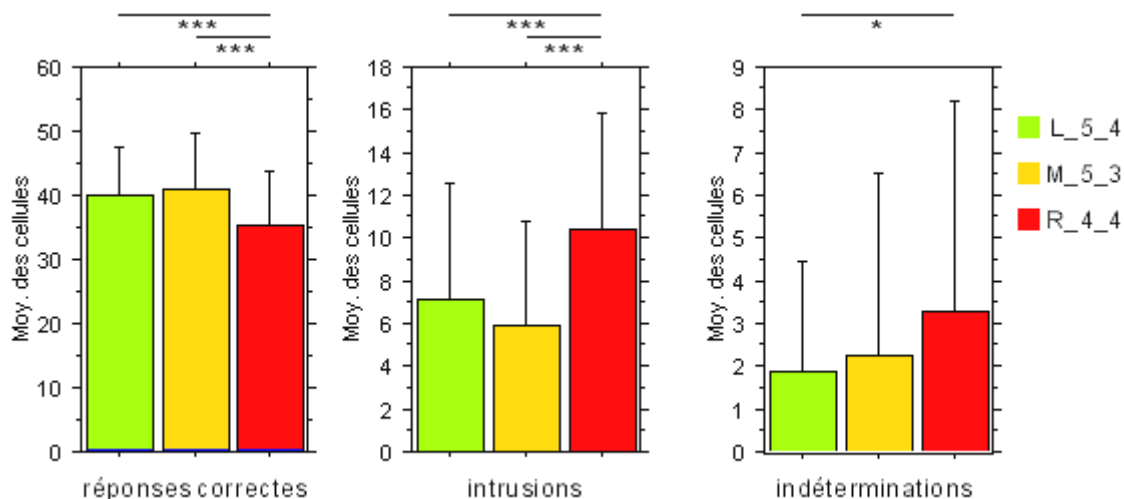


Figure 73 : performances d'identification lors des 3 conditions temporelles (L_5_4 : vitesse perçue comme la plus lente, TP=5 trames, ISI =4 trames ; M_5_3 : vitesse perçue comme intermédiaire, TP=5 trames, ISI =3 trames ; R_4_4 : vitesse perçue comme la plus rapide, TP=4 trames, ISI = 4 trames)

Le nombre moyen de réponses correctes est significativement différent dans les trois conditions ($F(2,22)=37,71$ $p<.0001$). Les sujets obtiennent significativement moins de réponses correctes à la vitesse perçue comme la plus rapide (R_4_4) par rapport aux deux autres conditions ($p<.0001$ pour le contraste R_4_4 et M_5_3 ; $p<.0001$ pour le contraste R_4_4 et L_5_4 ; test post-hoc PLSD de Fisher). De même, le nombre moyen d'intrusions est significativement différent dans les trois conditions ($F(2,22)=25,35$ $p<.0001$). Les sujets obtiennent significativement plus d'intrusions à la vitesse perçue comme la plus rapide (R_4_4) par rapport aux deux autres conditions ($p<.0001$ pour le contraste R_4_4 et M_5_3 ; $p=.0001$ pour le contraste R_4_4 et L_5_4 ; test post-hoc PLSD de Fisher). Le nombre moyen de réponses indéterminées est significativement différent dans les trois conditions ($F(2,22)=3,28$ $p=.047$). Les sujets obtiennent significativement moins de réponses correctes à la vitesse perçue comme la plus rapide (R_4_4) par rapport à la vitesse perçue comme la plus lente (L_5_4) ($p=.017$ pour le contraste R_4_4 et L_5_4 ; test post-hoc PLSD de Fisher).

Concernant le nombre d'intrusions, la différence significative entre les conditions R_4_4 et L_5_4 permet de suspecter soit un effet de la durée d'exposition au stimulus, soit un effet de la pression temporelle. La différence significative entre les conditions R_4_4 et M_5_3 permet de suspecter soit un effet de la durée d'exposition au stimulus, soit un effet de l'intervalle inter-stimuli. **Si un seul facteur explique la structure de ces données**, alors ces résultats suggèrent que **c'est la durée d'exposition au stimulus** qui explique les intrusions (voir table 7).

Comparaison \ Effet	Durée d'exposition au stimulus	Pression temporelle	Durée de l'ISI
R_4_4 et L_5_4	?	?	
R_4_4 et M_5_3	?		?
Bilan	✓		

Table 7 : Bilan des résultats sur la fréquence des intrusions. La durée d'exposition au stimulus est le point commun des deux contrastes significatifs.

Ces résultats sont en contradiction avec les interprétations données par les études de Intraub (1981), et Botella & Ericksen (1991) qui prêtaient à l'ISI la difficulté d'identification en RSVP. Ces études ne dissociaient pas la pression temporelle des autres facteurs. La présente expérience démontre que si le sujet n'a pas été suffisamment exposé à l'image, alors la probabilité d'intrusions augmente. Les mêmes résultats furent mis en évidence dans des paradigmes de recherche visuelle avec les distracteurs situés dans le domaine spatial (Prinzmetal, Henderson & Ivry, 1995). Lorsque le temps d'exposition diminue, quelles intrusions sont mises en cause ? Nous allons maintenant étudier cette question à travers les distributions des intrusions.

Distribution des intrusions

La distribution composite des intrusions est présentée à la figure 74. On observe que les intrusions '+1' et '-1' reproduisent le pattern de résultat obtenu sur les fréquences globales d'erreurs (paragraphe précédent), à savoir que la fréquence d'erreurs est spécifiquement influencée par le temps de présentation des stimuli. Les comparaisons de moyennes sont reportées à la table 8.

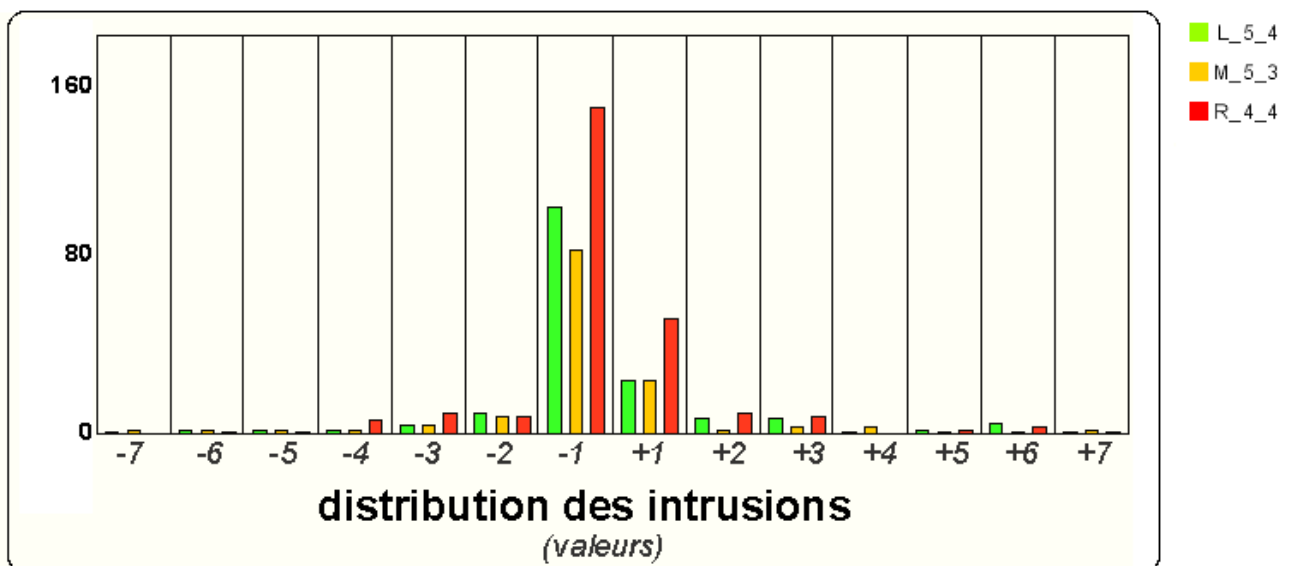


Figure 74 : distribution des intrusions dans les trois conditions expérimentales. On observe que les intrusions surnuméraires de la condition R_4_4, par rapport aux deux autres conditions, se portent tout particulièrement sur les positions '-1' et '+1'

Les deux premières conditions (images exposées pendant 71 ms) présentent un nombre similaire d'intrusions '+1' et '-1', tandis que la troisième (images exposées pendant 57 ms) présente un nombre accru d'intrusions. La taille de l'effet est considérable en '+1'. **En diminuant le temps de présentation de 20%, le nombre d'intrusions '+1' a plus que doublé (+117%).** Les intrusions '-1' sont déjà très nombreuses dans les conditions L_5_4 et M_5_3.

L'augmentation absolue est importante dans la condition R_4_4 (+52 intrusions), mais en proportion, l'effet est deux fois moins fort (+58%) que pour les intrusions '+1'.

Intrusions	ANOVA	Tests a posteriori
Moins 1	F(2,22) = 12,21 p < .0001	R_4_4 vs M_5_3 p = .0015 R_4_4 vs L_5_4 p < .0001
Plus 1	F(2,22) = 9,384 p = .0004	R_4_4 vs M_5_3 p = .0005 R_4_4 vs L_5_4 p = .0005

Table 8 : comparaison de moyennes pour les intrusions proximales. Les interactions significatives différencient la condition M_4_4 des deux autres conditions.

Les augmentations des intrusions '-1' et '+1' dans la condition R_4_4 n'étant pas proportionnelles, nous suspectons qu'un **mécanisme mixte sous-tende les intrusions '-1'**. Nous avons cherché un système de prédiction de ces intrusions '-1'. L'hypothèse utilisée était que les intrusions '-1' supportent à la fois le mécanisme responsable des intrusions '+1' et le mécanisme responsable des intrusions distales (voir aussi expérience 2 pp.122-129). Pour simplifier, nous avons supposé que les conjonctions illusoires étaient en nombre égal en '-1' et en '+1'. Nous avons cherché à voir si le reliquat des intrusions '-1' était compatible avec les modulations expérimentales affectant les intrusions distales.

Ce système de regroupement des fréquences coïncide parfaitement avec nos observations (voir figure 75). Les intrusions '-1' sont contraintes à reproduire les fréquences observées en '+1'. Le reste des intrusions '-1' et les intrusions distales présentent la même structure à travers les conditions expérimentales, à ceci près que les intrusions '-1' sont plus nombreuses d'un facteur 1,8. Ce système de regroupement des fréquences laisse un résidu de 2,2% seulement (fréquences théoriques moins fréquences observées, rapportées à l'effectif total des intrusions '-1'). Cette coïncidence est un indice en faveur de l'hypothèse mixte des intrusions '-1'.

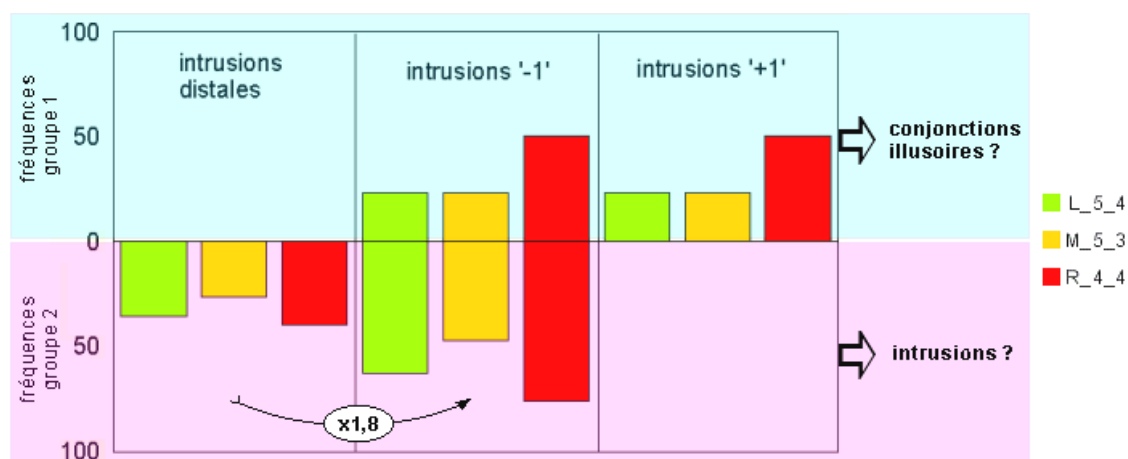


Figure 75 : regroupement hypothétique des fréquences d'intrusions observées en '-1'. En haut, pour former le premier groupe d'erreurs, nous avons forcé les fréquences des intrusions '-1' à adopter le même profil que les intrusions '+1'. En bas, on observe que le reliquat coïncide parfaitement avec la forme des intrusions distales. Les intrusions '-1' sont simplement 1,8 fois plus nombreuses que les intrusions distales, mais se répartissent de la même manière à travers les conditions expérimentales.

Discussion

La présente expérience a réservé deux observations contre-intuitives. D'une part, deux vitesses de défilement absolument identiques (M_5_3 et R_4_4) sont subjectivement perçues comme différentes. Et d'autre part, deux vitesses de défilement différentes (M_5_3 et L_5_4) génèrent le même nombre d'intrusions et le même nombre de réponses correctes.

Les difficultés subjectivement perçues des trois conditions s'ordonnent de la même manière que les fréquences des réponses indéterminées ($R_{4_4} > M_{5_3} > L_{5_4}$), a contrario des **intrusions** ($M_{5_3} \approx L_{5_3} < R_{4_4}$). Cela suggère que **le sujet n'est pas toujours conscient de ce second type d'erreurs**, corroborant ainsi les observations de Intraub (1989) sur le niveau de confiance excessif des sujets en leurs réponses en RSVP. Notons que des conjonctions illusoire (fondées de manière pré-attentionnelles) comptent parmi les types d'erreurs pour lesquelles le sujet n'est pas conscient d'avoir perceptivement défailli.

Les conclusions directes de cette expérience sont que **les intrusions se produisent en RSVP spécifiquement parce que l'image n'a pas été exposée suffisamment longtemps**, contredisant ainsi les travaux de Botella & Ericksen (1991), Potter & Levy (1969) et Intraub (1981). Les modulations expérimentales affectent spécifiquement les intrusions proximales. L'effet est considérable pour la position '+1', mais moindre pour la position '-1' suggérant que cette seconde classe d'erreurs est mixte.

L'interprétation du phénomène dans le cadre de la théorie du masquage (eg. Keysers & Perrett, 2002) n'est pas supportée. Si le phénomène répond bien à la définition d'un masquage (« Au niveau le plus général, le masquage fait référence à une réduction de la visibilité d'un objet (la cible) occasionnée par la présentation d'un second objet (le masque) contigu dans l'espace ou dans le temps. » traduit de Enns & DiLollo, 2001, p.345), en revanche, certaines propriétés rendent cette terminologie impropre. En effet, à des vitesses de défilement peu contrastées (rendant la difficulté comparable), cet effet se produit quelque soit le délai d'apparition du distracteur à l'issue de l'exposition de la cible (rétroactif) ou quelque soit le délai d'apparition de la cible après le distracteur (proactif). Or, le masquage est affecté par l'asynchronie d'apparition des stimuli (voir Humphreys & Bruce, 1989 pour revue). Corollairement, il est peu probable que l'interférence des traitements de la cible et du distracteur dans les structures visuelles sous-corticales et corticales explique **directement** ces résultats. En revanche, il est plausible que le nombre de salves envoyées au système visuel par les cellules ganglionnaires soit en ligne directe avec ce résultat. Les modèles actuels du fonctionnement visuel (voir chapitre 2) n'insistent pas sur l'alimentation répétée du système visuel par ses récepteurs¹, mais ce résultat semble en rappeler l'importance. Le système visuel poursuit son traitement au delà de l'existence effective de l'objet dans l'environnement (Colheart, 1980 ; Keysers & Perrett, 2002). Mais en présence de distracteurs,

¹ Les modèles actuels du fonctionnement visuel se focalisent sur la hiérarchie des aires de traitement, et simplifient la description de l'information périphérique à un flux unique issu des récepteurs (*one-pass process*), bien que de récente date, on permette plus de finesse dans les descriptions des niveaux de traitements ultérieurs, par exemple des flux asynchrones et/ou bidirectionnels (eg. Sanocki, 1993 ; Ullman, 1995 ; Bullier, 2001^a). Un consensus existe autour du caractère itératif du fonctionnement visuel, mais cette propriété n'est pas bien représentée dans les modèles.

la robustesse de ce traitement semble primitivement liée à la quantité d'information prélevée pendant que la cible était présente. Cette conclusion supporte le lien entre conjonctions illusoires et itération visuelle proposé figure 66 p.104.

A l'incidence de cette expérience, nous avons soulevé l'hypothèse d'une mixité des intrusions '-1'. Ce point va être plus largement examiné à travers une situation expérimentale d'apprentissage, laquelle a révélé une dissociation des intrusions proximales.

Expérience 2 : Effets d'exercice & dissociation des intrusions proximales

Dès les premières études en RSVP, les théories explicatives des erreurs d'identification se concentrent sur les intrusions proximales. A la suite de Lawrence (1971), les premières études sont consacrées aux intrusions '+1' (Broadbent, 1977, McLean, Broadbent, & Broadbent, 1982 ; Gathercole & Broadbent, 1984, Broadbent & Broadbent, 1987) jusqu'à ce que Botella & Ericksen (1992) proposent un modèle fondé sur les **transactions entre intrusions pré-proximales et post-proximales** ¹.

Botella & Ericksen proposent d'expliquer l'ambivalence des intrusions pré- et post-proximales par les temps relatifs de traitement du critère de réponse et de la cible. D'un point de vue opérationnel, nous observons effectivement que la distribution des intrusions est modulée par le choix spécifique des stimuli utilisés (voir expérience préliminaire A p.88), et donc qu'elle est probablement liée à la difficulté propre à chaque traitement.

Puisqu'il semble exister un système de transaction entre intrusions pré-cibles et intrusions post-cibles fondé sur la sélection des stimuli, le calcul **API** proposé par Botella & Ericksen, (1991) est **empiriquement légitimé**. Il en dérive implicitement que les intrusions proximales résultent d'un mécanisme unitaire, en contradiction avec les remarques de plusieurs études y compris de Botella & collaborateurs (Botella & Ericksen, 1992 ; Maki, Couture, Frigen & Lien, 1997 ; Maki, Frigen & Paulson, 1997) sur l'hétérogénéité et la complexité du phénomène.

Ce constat n'est qu'empirique. Il **ne légitime pas d'emblée l'unicité du phénomène sur le plan théorique** (voir p.99). Lorsque nous observons, par exemple, une postériorisation de l'API, nous ne savons pas si elle résulte d'une augmentation des intrusions post-proximales, d'une diminution des intrusions pré-proximales ou d'une transaction entre les intrusions pré- et post-proximales (cette dernière fondant l'hypothèse de Botella & Ericksen, 1992). Supposons que deux mécanismes distincts soient responsables des intrusions pré- et post-proximales respectivement, certains stimuli étant l'objet d'intrusions uniquement par le premier mécanisme, et d'autres étant l'objet d'intrusions par le second. Le dosage de ces deux types de stimuli dans le protocole expérimental fondera la répartition initiale du centre de gravité de la distribution. Ensuite, les manipulations expérimentales modifieront effectivement le pattern d'intrusions, mais en affectant uniquement celle des deux positions qui lui est directement liée². Dans ce cas, l'autre position sera inutile dans le calcul de l'API, sauf à considérer qu'elle norme la fréquence des intrusions sur une autre dimension, comme par exemple le niveau de difficulté de la tâche.

¹ Les travaux d'Intraub (1985, 1989) s'intéressaient également aux intrusions pré-cibles et post-cibles, mais les règles de répartition n'étaient pas réfutables : « selon que le sujet attend d'abord le critère de réponse ou la cible, le cadre sera parfois interprété comme faisant partie du traitement de l'image précédente ou du traitement de l'image suivante dans le système de stockage conceptuel à court-terme ». (Intraub, 1985, p.433).

² Dans notre série d'expériences utilisant des mesures répétées (expériences 1, 2 et 4), nous observons fréquemment que la manipulation expérimentale n'affecte qu'une seule des deux positions proximales lorsque les distributions d'intrusions sont corrigées pour le niveau de difficulté. On observe une modulation des intrusions '-1' pour l'expérience 1 et une modulation des intrusions '+1' pour l'expérience 4 ; l'expérience 2 ne présentant pas de différences en fonction des conditions expérimentales.

Stratégie

La présente expérience vise à **tester l'hypothèse d'unicité ou d'hétérogénéité des intrusions proximales**. Dans nos séries d'expériences, les intrusions proximales sont restreintes aux positions '-1' et '+1', car la vitesse de défilement est modérée¹: empiriquement on n'observe une majoration des intrusions par rapport au niveau de base qu'à ces positions.

Si les intrusions '-1' et '+1' résultent de mécanismes distincts, l'apprentissage est une situation expérimentale apte à les dissocier. En effet, rien n'oblige les intrusions provenant de mécanismes distincts à présenter la même courbe d'apprentissage, et on s'attend à observer des évolutions différentes. Ces évolutions distinctes seront rapportées aux profils d'évolution des sessions répétées, et en particulier à l'apprentissage inter-session et à l'apprentissage intra-session (voir ci-après).

Sujets

Les sujets étaient 22 adultes volontaires pour l'expérience (11 femmes, 11 hommes), âgés de 19 à 44 ans (moyenne d'âge : 26 ans). Les sujets ont reporté une vision normale, ou corrigée à la normale.

Matériel & méthodes

Les sujets étaient soumis à 4 sessions de 120 essais RSVP, réalisés à la même heure à quelques jours d'intervalle. La procédure était identique au protocole générique décrit pp.109-111. Deux catégories d'images ont été étudiées, les animaux et les vêtements (voir répartition des essais figure 76).

¹ Le choix de l'intervalle pertinent pour définir les intrusions proximales est lié au recouvrement temporel des traitements de la cible et des distracteurs contigus. Si la vitesse de défilement est lente, alors les intrusions se concentrent davantage sur la (les) position(s) immédiatement contiguë(s) à la cible, car ces intervalles définissent une plus grande plage temporelle.

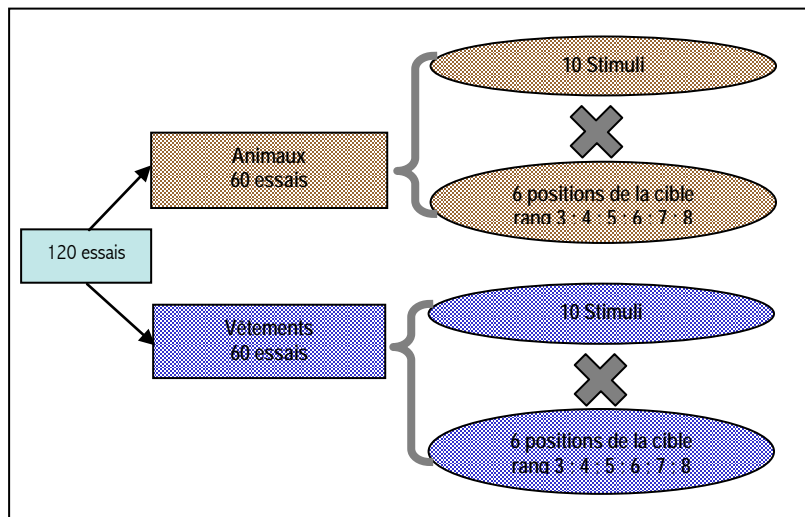


Figure 76 : répartition des essais de l'épreuve 'apprentissage'. Les distracteurs sont assignés aux différents rangs par randomisation. L'ordre des essais est géré par randomisation réelle.

Résultats

Evolution générale

L'évolution des performances au fil des essais est reportée à la figure 77. On observe que la performance augmente régulièrement lors des 3 premières sessions pour se stabiliser à la quatrième session. Les performances semblent alors avoir atteint un plafond.

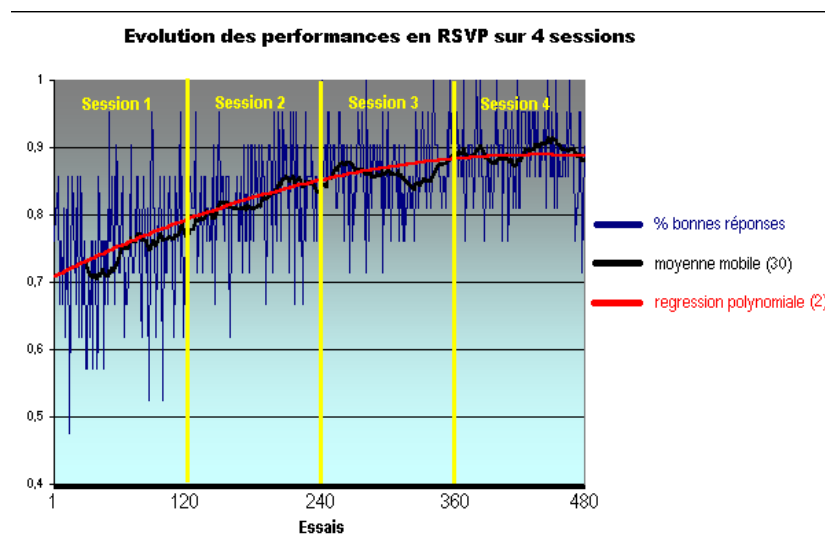


Figure 77 : évolution des performances au cours des essais réalisés lors de 4 sessions RSVP. Les données sont ajustées par une moyenne mobile d'incrément 30 en noir, et par une régression polynomiale d'ordre 2, en rouge.

Deux sources des effets d'apprentissage sont à envisager. La première est liée à la tâche proprement dite, c'est-à-dire à l'**ajustement** du sujet **aux conditions de flux rapide d'image** (optimisation de la détection du critère de réponse, identification et consolidation de la cible). Dans cette perspective, plus le sujet sera exposé à des essais en RSVP, plus il progressera. Aussi, nous examinerons tout particulièrement l'**évolution intra-session**.

La seconde est liée à l'exposition des stimuli **au début de chacune des quatre sessions** (avant le test, le sujet voit chaque stimulus pendant quelques secondes et en reporte le nom). On s'attend à ce que cet apprentissage se traduise par une élévation du niveau initial d'une session par rapport à la session précédente. C'est tout particulièrement **l'évolution inter-session** qui sera examinée dans cette perspective. Selon Subramaniam, Biederman & Madigan, (2000), la **mémorisation des caractéristiques des images** ne peut se réaliser que lors de présentations prolongées, car en condition RSVP, la consolidation en mémoire est interrompue par le flux d'images. Dans les verbalisations des sujets, nous avons pu identifier une difficulté liée à associer un traitement partiel avec l'identité de la cible (voir p.98). Par exemple, le sujet se rappelle que la cible était rouge, mais il échoue à en déduire son identité. Ce type d'erreurs est amené à diminuer au fil des sessions, car la présentation initiale des stimuli améliore la connaissance qu'à le sujet de leurs caractéristiques, et maximise les réponses basées sur des informations partielles (voie de pari sophistiqué, Botella & coll., 2001).

Evolution intra-session

L'évolution intra-session traduit l'**adaptation du sujet à la situation RSVP**. On observe que la pente d'apprentissage est de plus en plus faible à chaque session (voir figure 78). Le gain en performance est maximal lors de la première session, puis perd progressivement en intensité pour devenir nul dès la troisième session. Autrement dit, **l'apprentissage perceptif en RSVP atteint son point de saturation à la troisième session**.

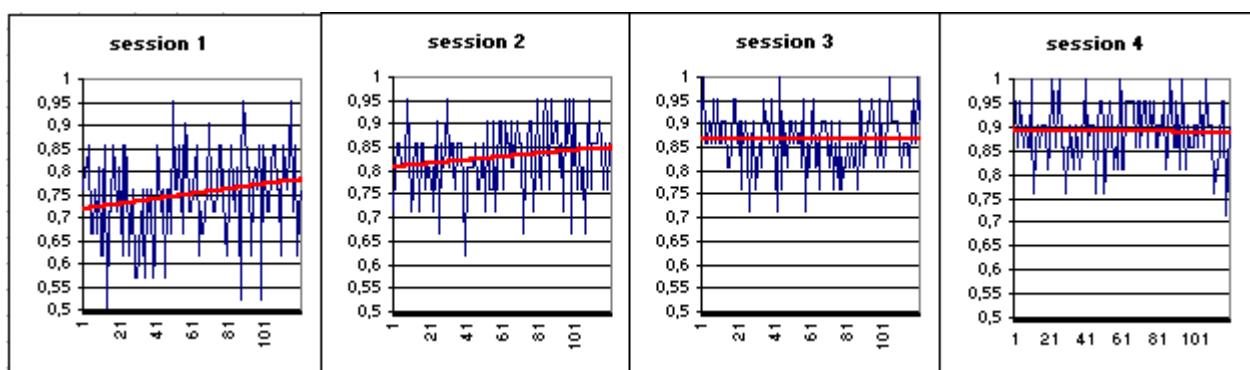


Figure 78 : évolution des performances pour chacune des 4 sessions de RSVP. A la première session, le coefficient de régression est de 0,1% (gain en performance de 12% sur la session). A la seconde session, le coefficient de régression n'est plus que de 0,036% (gain de 4,35%). A la troisième session, le coefficient de régression tombe à -0,00076% (perte de 0,09%) et à la quatrième session, le coefficient de régression est de -0,005% (perte de performance de 0,60% au cours de la session).

Evolution inter-session

Si les performances intra-session cessent d'évoluer dès la troisième session, en revanche les effets inter-sessions restent visibles sur l'ensemble des quatre sessions. Les intercepts des régressions linéaires (tendance de la performance en début de session) des quatre sessions sont respectivement de 0,716 ; 0,806 ; 0,867 ; et 0,893. Lors de la seconde session, le sujet débute l'épreuve à un niveau de performance supérieur au début de la première session, mais inférieur à la fin de la première session, suggérant que le sujet doit « se remettre dans le bain », c'est-à-dire que les premiers essais sont perturbés par le manque de familiarité à la tâche. Lors des sessions 3 et 4, les niveaux de performance en entrée de session sont toujours supérieures aux niveaux de performance de sortie de la

session précédente. Deux interprétations (non exclusives) sont possibles pour ces données. Soit le niveau réel de performance en fin de session est masqué par une interaction antagoniste entre apprentissage et fatigue. Il serait alors sous-estimé. Soit des apprentissages de la tâche se réalisent en dehors de la série d'essais. Cette deuxième possibilité corrobore l'hypothèse de Subramaniam, Biederman & Madigan (2000) concernant l'**apprentissage des caractéristiques des stimuli lors de la présentation des stimuli juste avant chaque session**.

Cet apprentissage inter-session se poursuit sur chacune des 4 sessions étudiées (gain de 9% entre les deux premières sessions ; gain de 6,1% entre la session 2 et la session 3, et gain de 2,6% entre les deux dernières sessions), **mais son effet diminue linéairement** à chaque session selon la fonction affine $y = -3,2x + 12,2$ où x représente le numéro de la session. Si la progression continue linéairement comme le suggèrent ces quatre séries de données, on peut estimer qu'elle **plafonne après la quatrième session**. Puisque l'apprentissage intra-session est lui aussi devenu nul, il est probable que les performances moyennes stagnent aux alentours de 90% de réponses correctes.

Classification des réponses

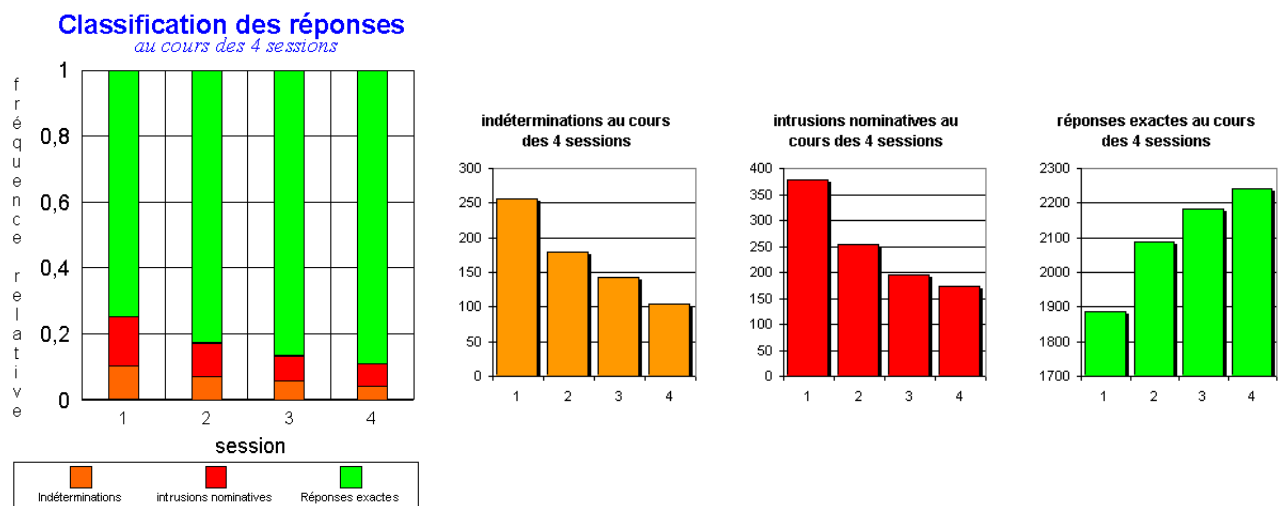


Figure 79 : Evolution des performances au fil des sessions : nombre d'indéterminations, nombre d'erreurs nominatives, nombre de réussites

L'évolution des performances est reportée à la figure 79. La performance passe de 1887 (75% des essais) à la session 1 à 2243 (89% des essais) à la session 4, soit une augmentation de 19 % pour les **essais réussis**.

Le nombre d'**indéterminations** pour l'ensemble des sujets passe de 256 (10% des essais) à 105 (4% des essais) au cours des quatre sessions, soit une baisse de 59%. La **diminution** est relativement **linéaire**, elle **est compatible avec la progression inter-session**, ce qui suggère que plus les sujets connaissent les caractéristiques des stimuli, moins ils commettent d'indéterminations.

Le nombre d'**erreurs nominatives (intrusions)** décroît lui-aussi. Il passe de 377 (15% des essais) à la session 1 à 172 (7% des essais) à la session 4, soit une baisse de 54%. La diminution du nombre d'intrusions est marquée entre les deux premières sessions, et semble s'amenuiser ensuite. La **progression** est différente de celle des

indéterminations. Elle semble **résulter d'une adjonction de l'apprentissage intra-session** (diminution sur les deux premières sessions puis stagnation) **à l'apprentissage intersession** (diminution linéaire). Nous observons effectivement à la figure 80 que les intrusions décroissent régulièrement au cours de chaque session, suggérant que l'exposition continue aux essais RSVP favorise la performance. Le paragraphe suivant précisera les positions d'intrusions qui sont le plus liées à cette progression.

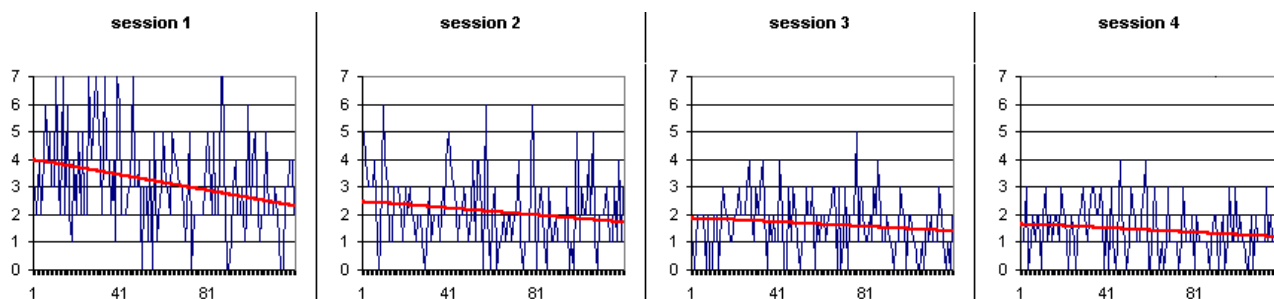


Figure 80 : évolution des intrusions au cours des quatre sessions expérimentales.

Evolution des distributions d'intrusions

Entre la première et la quatrième session, le nombre d'intrusions distales diminue de 64%, le nombre d'intrusions '-1' diminue de 67%, et le nombre d'intrusions '+1' ne diminue que de 25%. Au cours des quatre sessions, les **intrusions '-1' se réduisent** en proportion équivalente aux intrusions distales, mais **les intrusions '+1' diminuent plus lentement**. La répartition pré-post des intrusions proximales se modifie avec l'exercice, comme on l'observe à la figure 81. Lors de la première session, le nombre d'intrusions est maximal pour les distracteurs situés juste avant la cible. Au fil des sessions, l'écart entre intrusions '-1' et intrusions '+1' se minimise. A la session 4, on observe même un avantage pour les intrusions '+1'.

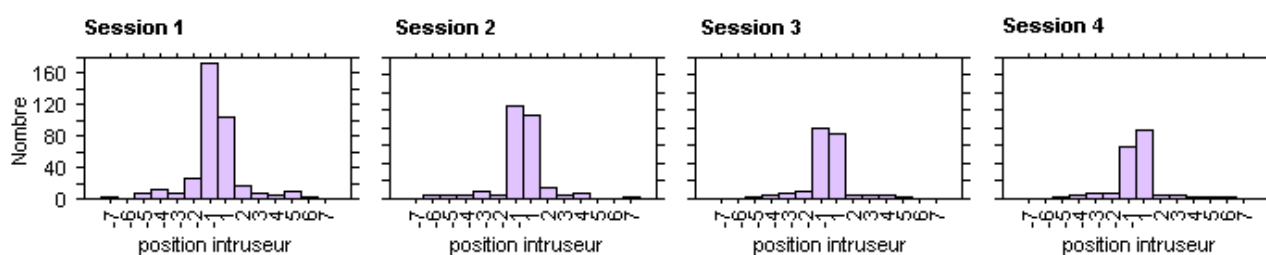


figure 81 : évolution de la distribution des intrusions au fil des 4 sessions.

Corollairement, **l'API est devenu plus postérieur avec l'exercice**. Il passe de -0,25 à la première session à -0,06 à la seconde, 0 à la troisième et 0,15 à la dernière. Or, la plupart des auteurs ont mis en évidence une postériorisation de la distribution lorsque la tâche RSVP est d'une plus grande difficulté (eg. vitesse de défilement plus rapide, Botella et Ericksen, 1991 ; charge attentionnelle plus grande, Chun, 1997 ; moindre familiarité des stimuli, Botella & coll. 2001). Le présent résultat est en apparence **contradiction** avec cette série d'études, et **questionne l'intérêt de l'indice API lorsque les intrusions proximales sont concentrées aux rangs '-1' et '+1'**. Ici, la majeure partie de la variation de l'API est supportée par les seules intrusions '-1', tandis que les précédentes études manipulant la difficulté de la tâche via diverses contraintes expérimentales affectaient probablement (1) les intrusions post-proximales ou (2) la balance des

intrusions pré- et post-proximales. Effectivement, on peut voir dans Chun, (1997) le profil inverse de celui-ci. La charge attentionnelle se manifeste par une variation des intrusions '+1'. Les intrusions '-1' quant-à elles restent relativement stables en fonction de la charge attentionnelle, et ne se différencient pas selon que le sujet réalise une simple ou une double tâche.

En résumé, la diminution de la fréquence des intrusions est donc due essentiellement aux intrusions '-1' qui, principaux contributeurs des intrusions à la première session, subissent une diminution constante au fil des sessions. Nous obtenons là une dissociation des intrusions pré-proximales et post-proximales. Les intrusions '-1' évoluent de manière similaire aux intrusions distales (voir figure 82), tandis que les intrusions '+1' se différencient.

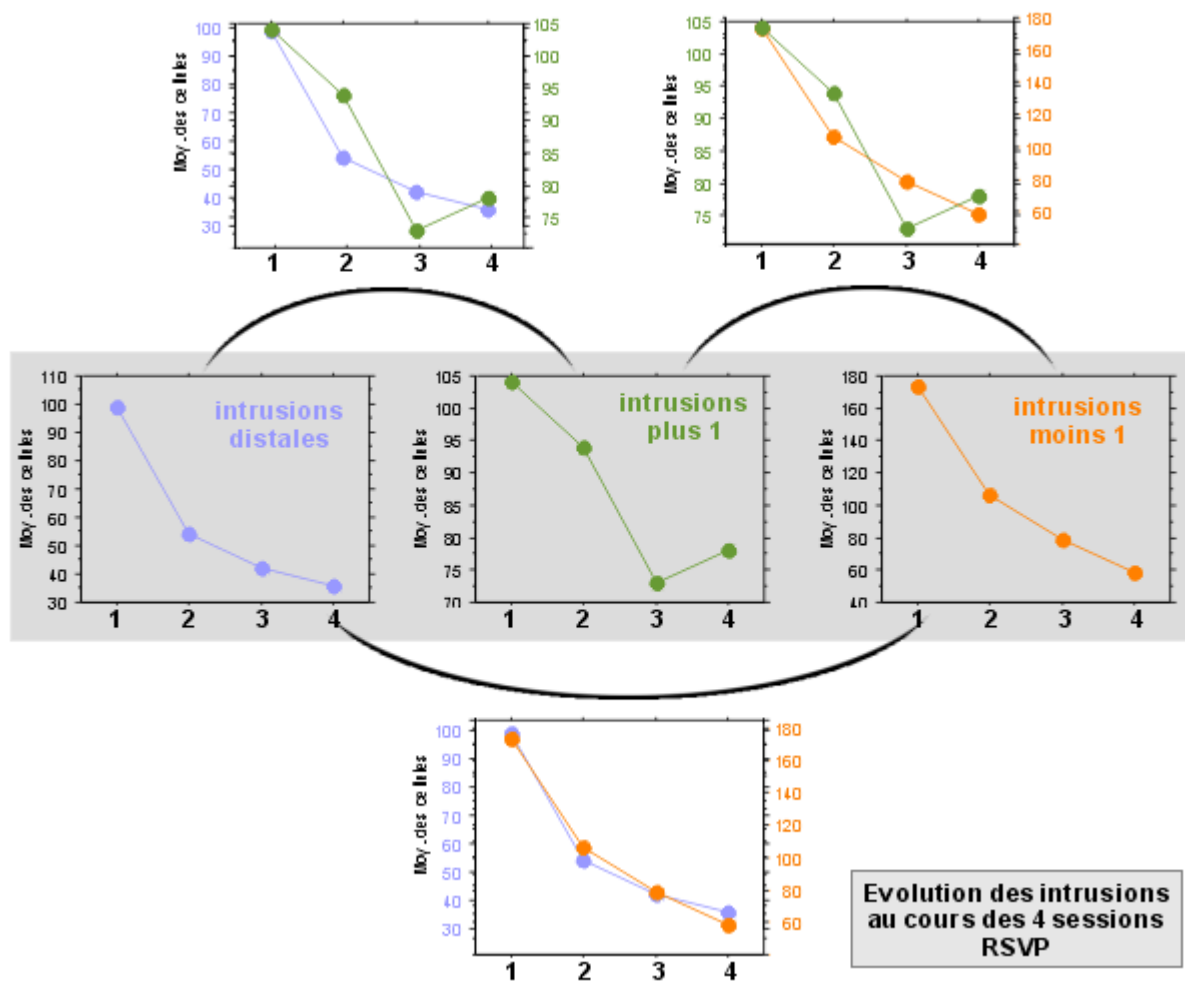


figure 82 : comparaison des courbes d'apprentissage des intrusions pré-proximales (en vert), post-proximales (en orange) et distales (en bleu). Les intrusions '-1' et distales se comportent de manière similaire, tandis que les intrusions '+1' ont une évolution spécifique.

Les intrusions '+1' sont peu sensibles à l'exercice. On n'observe que de faibles variations inter-session (figure 83). Lorsque l'on observe les variations intra-session, on note des modifications inconsistantes sur les deux premières sessions (sans doute due à la variabilité des réponses dans la phase de familiarisation avec la tâche), suivies d'une stagnation sur les deux dernières sessions (figure 83). Le comportement de ces courbes est très différent de celui des intrusions '-1' et distales, qui subissent des évolutions intra-session semblables à celles reportées à la figure 80 pour l'ensemble des intrusions.

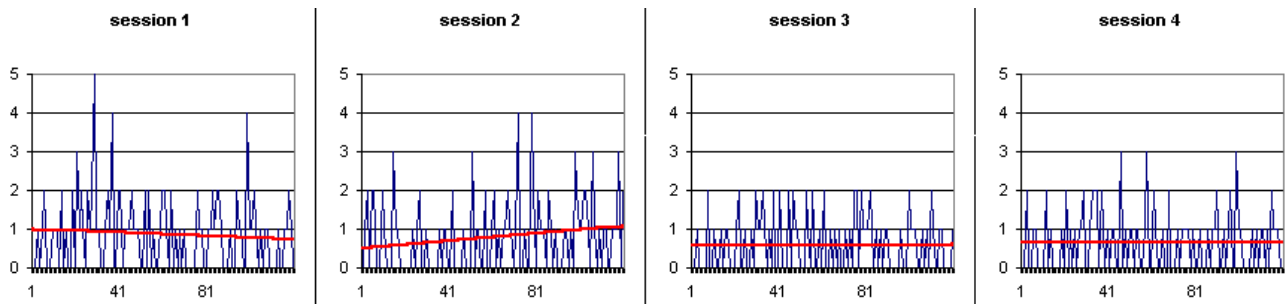


Figure 83 : évolution de la fréquence des intrusions '+1' au cours des 4 sessions

Discussion

Botella & collaborateurs proposent de synthétiser le comportement des intrusions proximales par un calcul de centre de gravité (indice API, Botella & Ericksen, 1991), considérant ainsi que ces intrusions sont homogènes, et que les manipulations expérimentales se traduisent par une transaction entre les différentes positions dans la distribution. Au contraire, cette expérience démontre que **les intrusions pré-proximales et post-proximales se comportent différemment**, et que l'exercice affecte sélectivement les premières. Nous n'observons pas d'augmentation des intrusions '+1' concomitamment à la diminution des intrusions '-1', suggérant qu'il ne se produit pas de transaction entre ces deux classes de stimuli, dans les conditions expérimentales manipulées ici.

La diminution des intrusions '-1' semble en partie suscitée par l'apprentissage de la tâche RSVP. Elle présente une évolution similaire à celle des intrusions distales. La baisse linéaire de l'apprentissage inter-session évoque la voie de pari sophistiquée. Mais l'apprentissage intra-session évoque également une adaptation au flux d'image rapide.

Malgré cette diminution drastique, il reste encore, à l'issue de la quatrième session, un fort contingent d'intrusions '-1', alors que les indicateurs de progression de la performance suggèrent une stagnation. Ce contingent d'intrusions '-1' semble irréductible. Il s'ajoute au contingent d'intrusions '+1', qui lui-aussi, évolue peu. Nous avons formulé plusieurs **hypothèses sur l'hétérogénéité des intrusions** (figure 64 p.99), et il semble que la solution C soit supportée par nos données (voir également figure 84). Une partie des intrusions '-1' et les intrusions '+1' fondent un premier complexe d'erreurs situé dans la proximité temporelle de la cible. Ces erreurs sont peu sensibles à l'apprentissage, et appartiennent plus probablement au domaine de la perception et de l'attention qu'au domaine de la mémoire.

La seconde partie des intrusions '-1' fonde un second groupe d'erreurs spécifiquement manipulé dans cette expérience. Ces résultats sont en accord avec ceux de l'expérience 1, qui suggéraient déjà une mixité des intrusions '-1' (voir figure 75 p.117).

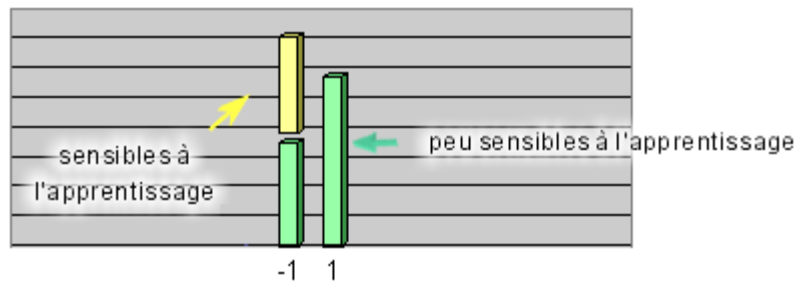


Figure 84 : constitution des intrusions : hypothèse supportée par l'expérience 2. C'est la partie jaune des intrusions '-1' qui semble exposée à l'apprentissage, alors que les parties vertes des intrusions '-1' et '+1' évoluent peu.

Pour expliquer le complexe d'intrusions '-1' et '+1', nous devons donc chercher un mécanisme peu sensible à l'apprentissage (à la connaissance plus ou moins grande des stimuli et à l'habitude de réaliser des identifications en flux visuel rapide), se manifestant de manière proactive et rétroactive, et se traduisant a priori par une intégration (conjonction illusoire) ou par une substitution (intrusion) du traitement de la cible par celui d'un distracteur.

Dans les verbalisations des sujets, nous avons pu identifier deux types de difficultés d'identification de la cible perçues par les sujets¹ et donnant lieu à des confusions cible-distracteur. La première était une difficulté à se remémorer, parmi le groupe de stimuli, celui qui était porteur d'une caractéristique identifiée (eg. « c'était rouge mais je ne me souviens plus de quoi il s'agit »). On peut supposer que ce n'est pas cette source d'erreur qui se manifeste ici ; cette source d'erreur est certainement réduite par l'apprentissage, et spécialement par la présentation prolongée des stimuli avant la session de test. Ces erreurs se distribuent pour part dans les indéterminations (le sujet refuse de formuler une réponse) et pour part dans les intrusions (le sujet essaie de réaliser un pari sophistiqué, comme le suggèrent Botella & coll, 2001). Lorsque des intrusions sont produites par ce mécanisme, rien n'oblige à ce qu'elles soient distribuées dans la proximité temporelle de la cible.

La deuxième source de confusion était une difficulté d'attribution du critère de réponse à une image parmi deux. Nous avons suggéré (p.98) que ces intrusions se produisaient lorsque les **traitements de la cible et du distracteur** avaient été **intégrés**, et qu'un trait signalait l'incongruence de l'association. A l'instar du système auditif en écoute dichotique, le système visuel, lorsqu'il se retrouve devant un dilemme, dissout probablement l'association, et restitue en tant que cible deux représentations **dissociées**. Ainsi, des sujets rapportent parfois voir le fond gris sur deux images et non une seule et s'étonnent lorsque l'expérimentateur confirme qu'une unique image porte le critère de réponse. Ainsi, la migration du cadre décrite par Intraub, (1985, 1989) ne consiste probablement pas en une conjonction illusoire entre un distracteur et le critère de réponse, mais plus probablement en une conjonction illusoire dissolue de la cible et d'un distracteur.

¹ Il se produit également des erreurs dont les sujets n'ont absolument pas conscience. Intraub (1989) souligne que des erreurs sont commises alors même que le sujet est très confiant dans sa réponse.

Les erreurs d'identification formant le complexe '-1' et '+1' décrit à l'issue de cette expérience, consistent vraisemblablement en des conjonctions illusoires. Certaines, dont le liage n'a pas persisté jusqu'au terme du traitement, sont perçues par les sujets comme des perceptions défaillantes. A celles-ci s'ajoutent les conjonctions illusoires complètes, dont le sujet n'a visiblement pas conscience, puisque dans ce cas, il manifeste une grande certitude quant-à sa réponse (Intraub, 1989).

Nous allons maintenant interroger les modèles explicatifs de ces intrusions proximales, et en particulier le modèle de migration du cadre (Intraub, 1985, 1989). L'expérience suivante nous permettra de prendre position entre les hypothèses précoces et tardives des intrusions. Selon les hypothèses tardives, l'erreur se produit après que l'image ait été entièrement identifiée. Selon les hypothèses précoces, les erreurs se produisent de manière pré-attentionnelle, durant le traitement des caractéristiques visuelles séparables (primitives visuelles). Parmi les résultats précédents, plusieurs suggèrent que l'erreur se produit de manière pré-attentionnelle.

Expérience 3 : le cadre migrateur ?

La théorie d'Intraub (1985) s'inscrit dans la lignée des travaux de Potter & Levy, (1969) et Potter, (1976), visant à établir que les déficits d'identification observés lors de flux rapides d'images ne s'expliquent pas par un déficit perceptif, mais par une incapacité à retenir l'information (« immédiatement perçu, immédiatement oublié »). Lorsque des images indépendantes sont présentées en séries rapides, nous n'avons pas l'impression de percevoir des lignes et des tâches sans signification, mais nous sommes capable de percevoir des tous organisés. La perception des images est possible, mais c'est la rétention qui est déficitaire. Dans la continuation de ces travaux, Intraub (1985, 1989) élabore un modèle des intrusions en RSVP construit autour du concept de **mémoire à très court-terme**.

Les deux points cruciaux de la théorie d'Intraub sont :

- la formation des conjonctions illusoirs dans un buffer visuel. Ce buffer est un système de **mémoire à très court terme**. Intraub l'appelle variablement *short-term conceptual store* ou *very short-term buffer* (Intraub, 1985) ou *very short-term store* ou encore *very short-term integrative buffer* (Intraub, 1989)¹.
- La **migration du cadre** (*frame migration*) dans ce système de stockage en raison de l'indépendance conceptuelle de ce dernier avec l'image-cible. Ce point vise à établir que les conjonctions illusoirs sont formées dans **une mémoire post-catégorielle**, au contraire des conjonctions illusoirs de Treisman (Treisman & Gelade, 1980, Treisman, 1986, 1996, 1998), affectant le fichier-objet avant l'intégration perceptive, et réalisée durant l'action du faisceau attentionnel sur la trame perceptive. Intraub (1985) développe ainsi des arguments concernant les modulations top-down (contexte, signification de la scène visuelle) des étapes visuelles précoces. Ce postulat conduit à des résultats mitigés (Intraub, 1989, expériences 3 et 4).

Intraub (1985) utilise un paradigme RSVP avec des images d'objets réels, et fonde sa théorie sur l'explication des intrusions '-1' et '+1'. Dans ces expériences, le critère de réponse est constitué par un cadre carré, placé au pourtour de l'image. Le cadre est conceptuellement indépendant de l'image (sa signification ne s'intègre pas avec celle de l'image), aussi le système cognitif le désolidarise temporellement de son contexte (ie. de l'image qui apparaît en même temps que lui). Selon Intraub, le cadre, l'image-cible et une ou deux autres images sont situés dans un système de mémoire conceptuelle à très court-terme. Puisque le cadre n'est pas solidarisé à la cible, il peut migrer et aller

¹ Les différentes formulations de ce buffer par Intraub sont en conflit avec les théories de la mémoire. Selon ces théories (eg. Wheeler & Treisman, 2002), deux formes de mémoire visuelle se succèdent. La mémoire iconique (Neisser, 1967 ; Sperling, 1960) est pré-catégorielle, est dotée d'une grande capacité pour les matériels visuels présentés de manière isolée (sans masquage) et semble reposer sur l'activité additionnelle des neurones de IT (durant 350 millisecondes) après le traitement perceptif pur (les 50 premières millisecondes) selon Tovee & Rolls (1995) Subramaniam, Biederman & Madigan (2000). La participation d'autres structures visuelles n'est bien sûr pas exclue. La mémoire visuelle à court-terme (VSTM) qui lui succède est conceptuelle, de capacité limitée et rend possible un rappel différé de quelques secondes. Or la mémoire décrite par Intraub, mélange ces attributs (voir aussi Potter, 1999). Quoique catégorielle et disponible sur quelques secondes jusqu'au rappel, elle est active dans les 300 millisecondes autour de la cible (Intraub, 1989 avance que l'action de ce buffer soit online). Seules les équipes de Intraub et Potter (Intraub, 1985, 1989, 1993 ; Potter, 1976, 1993, 1999 ; O'Connors & Potter, 2002) défendent la nécessité de cette forme intermédiaire de mémoire à court-terme.

librement se fixer sur une des autres images présentes dans ce système de stockage. Intraub parle alors d'**effet de dissociation** (une forme d'illusion visuelle dans lequel le cadre n'est pas perçu comme accompagnant la cible).

Intraub (1985) observe deux directions de migration, se traduisant par des intrusions '-1' et '+1'. Elle explique cette ambivalence de la position des intrusions par l'ordre de traitement de l'image et du cadre (Intraub, 1985), puis par l'ordre de focalisation de l'attention (Intraub, 1989) :

- Si le sujet a donné la primauté au traitement du cadre, alors c'est le traitement de l'image qui va tarder, et on observera une intrusion par un distracteur précédant la cible (eg. une intrusion -1).
- Au contraire, si le sujet traite l'image en premier, c'est le traitement du cadre qui va être repoussé. Le cadre sera ainsi associé à l'image suivant la cible, et on observera une intrusion par un distracteur succédant à la cible (eg. intrusion +1).

Dans cette formulation, les conjonctions illusoires consistent en l'appariement fallacieux entre le cadre et une image présentée à un autre moment. C'est une forme de conjonctions illusoires beaucoup plus élémentaire (un seul échange) que celles que nous suspectons dans notre corpus d'expériences, au cours desquelles des caractéristiques visuelles complexes de la cible et d'un distracteur seraient associées.

Stratégie

La présente expérience vise à tester la théorie de « migration du cadre ». Intraub (1985, 1989) estime que la migration du cadre se produit parce que la cible et le critère de réponse sont conceptuellement indépendants. Selon elle, les conjonctions illusoires se produisent parce qu'il n'est pas possible au système visuel d'associer la signification de la cible et du critère de réponse. Notre stratégie expérimentale consiste à comparer deux conditions, l'une utilisant un critère de réponse extrinsèque à l'image-cible (le classique carré gris), et l'autre utilisant un critère de réponse intrinsèque (la cible est la seule image de la séquence présentée en noir et blanc). Dans cette seconde condition, le critère de réponse est conceptuellement dépendant de la cible. Le sujet détecte que l'image est en noir et blanc après avoir inspecté la distribution chromatique de l'image. De la théorie de Intraub (1985) dérive la prédiction que nous n'observerons pas de conjonctions illusoires dans la seconde condition.

Sujets

Les sujets étaient 25 adultes volontaires pour l'expérience (12 femmes, 13 hommes), âgés de 20 à 56 ans (moyenne d'âge : 34 ans). Les sujets ont reporté une vision normale, ou corrigée à la normale.

Matériel et méthodes

Les sujets ont complété une session RSVP de 120 essais, durant environ 25 minutes, consignes et entraînement compris. Nous avons retenu un plan expérimental en mesures répétées avec deux conditions expérimentales. Dans

l'une d'elle (fond gris), la cible est l'unique image de la séquence présentée sur un carré gris. Dans la seconde (achromatique), la cible est l'unique image de la séquence présentée en noir et blanc. La cible est positionnée aux rangs 3 à 8 de la séquence.

Les conditions expérimentales sont alternées par blocs de 10 essais. La moitié des sujets commencent la session par un bloc d'essais de la condition 'fond gris' et l'autre moitié commence par un bloc d'essais de la condition 'achromatique' (voir figure 85). Les blocs appariés sont distants d'au moins 3 blocs, et l'ordre des essais dans les blocs est aléatoire pour limiter l'effet de la familiarité.

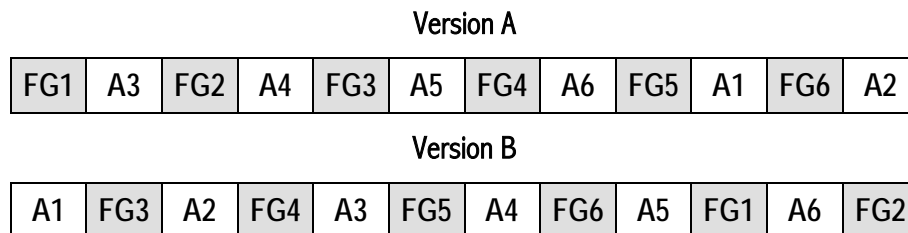


Figure 85 : ordre des blocs dans les deux versions de l'expérience. Chaque bloc comporte 10 essais. Chaque condition comporte 6 blocs numéroté de 1 à 6. L'écart minimal entre deux blocs appariés est de 3 blocs. L'ordre des essais dans les blocs est randomisé pour davantage limiter l'effet de la familiarité.

Pour que le sujet ne puisse pas utiliser stratégiquement la couleur pour identifier la cible dans la condition 'fond gris' (et biaiser ainsi les résultats en faveur de cette condition), les stimuli sont choisis dans des gammes chromatiques similaires et monotones. De cette façon, le contraste est similaire pour les cibles des deux conditions expérimentales. Il n'y a pas davantage d'information visuelle dans l'image sur carré gris et dans l'image en noir & blanc. En effet, la segmentation perceptive, préalable de l'identification visuelle, se réalise principalement sur la base des variations de luminance (voir chapitre 2 pp.53-73), mais elle peut être complétée par d'autres plans de segmentation, basés par exemple sur des variations de teinte (eg. Gegenfurtner & Rieger, 2000, Tanaka, Weiskopf & Williams, 2001). Si une information chromatique riche avait été présente dans les stimuli, l'identification des surfaces et des contours serait devenue inégale dans les deux conditions.

Les distracteurs sont des animaux choisis dans des teintes « brun moyen ». Les cibles de la condition 'fond gris' sont les mêmes stimuli que précédemment (distracteurs), mais ceux-ci sont présentés sur un carré gris de taille 250 x 250 pixels, soit environ 3° x 3° en affichage 640 x 480 pixels avec le dispositif d'affichage utilisé. Les cibles de la condition 'achromatique' sont également les mêmes stimuli que les distracteurs, mais ceux-ci ont été convertis en noir et blanc. Ils apparaissent directement sur le fond noir (voir exemple de séquence figure 86).



figure 86 : En haut, exemple de séquence RSVP de la condition 'fond gris'. La cible est l'escargot. En bas, la même séquence en condition 'achromatique'.

Nous supposons que ce dispositif manipule effectivement la dépendance conceptuelle de la cible et du critère de réponse. Dans la condition achromatique, qui supporte la dépendance conceptuelle de la cible et du critère de réponse, ces deux événements sont spatialement superposés. Mais surtout, la détection du critère de réponse n'est possible qu'à l'issue de l'analyse perceptive de l'image entière. En effet, pour pouvoir détecter que l'image est en noir et blanc, le sujet ne peut pas se contenter d'une analyse locale et partielle de l'image. S'il cherche du brun, le sujet doit parcourir l'image entière pour garantir qu'il ne s'en trouve nulle part. S'il procède en recherchant les parties achromatiques, alors il ne suffit pas de détecter du noir (tous les distracteurs en contiennent), du gris ou du blanc (id, voir figure 87) pour décider qu'une image est la cible. C'est l'opposition noir et blanc, à l'exclusion de toute information chromatique, qui permet de décider que l'image est la cible. Aussi, la détection du critère de réponse oblige à un traitement étendu de l'image. Ce traitement garantit que la cible et le critère de réponse seront conceptuellement liés, les distributions de luminance (précurseur de l'information sur la forme de l'objet) étant inscrit dans le critère de réponse.

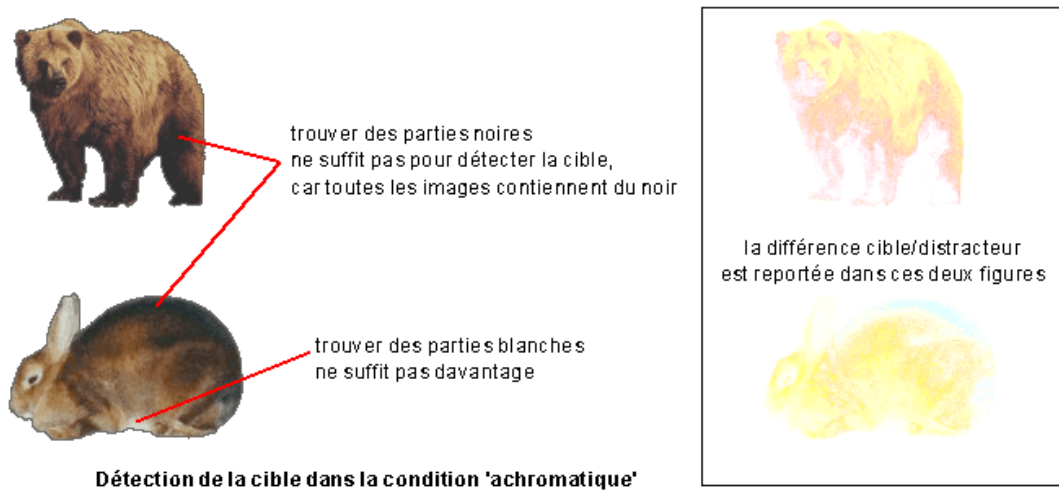


Figure 87 : détection du critère de réponse dans le flux d'image de la condition achromatique. Si le sujet cherche localement du noir, du gris ou du blanc, la tâche est irréalisable. Toutes les images comportent des parties de faible saturation chromatique, et dont la position est imprévisible. Il semble obligatoire que le sujet traite l'image globalement. Nous voyons dans la partie droite de l'image que la différence entre la cible et le distracteur tient aux parties de luminance moyenne (ni trop claires, ni trop sombres).

Résultats

Fréquence des erreurs

Trois types de réponses ont été obtenues : des réponses correctes, des intrusions (nom d'une autre image de la séquence) et des réponses indéterminées (je ne sais pas). La fréquence de ces trois types de réponses est reportée à la figure 88.

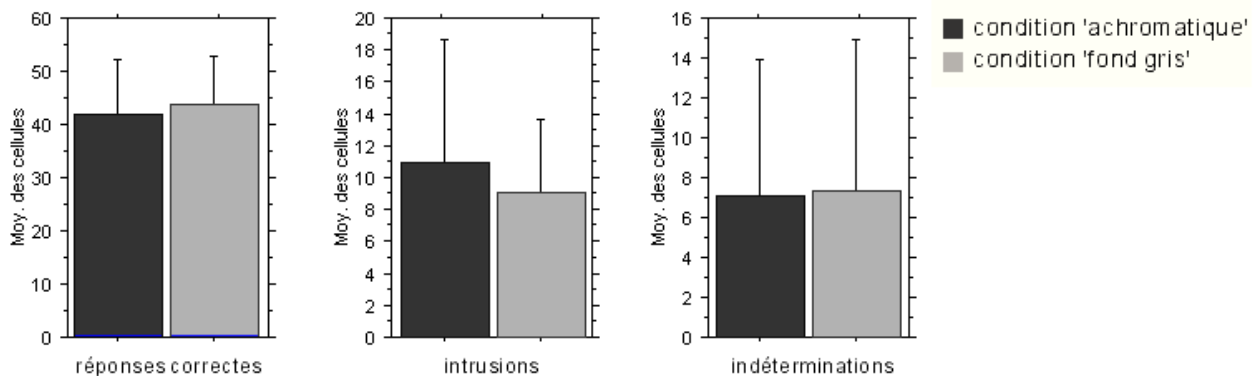


Figure 88 : performances d'identification lors des 2 conditions expérimentales. Aucune des différences n'est significative.

Les sujets obtiennent 41,9 réponses correctes (70%) en condition 'achromatique' et 43,6 réponses correctes (72,7%) en condition 'fond gris'. Ces performances ne sont pas significativement différentes dans les deux conditions ($F(1,24)=0,97$ $p=.33$). Les sujets fournissent en moyenne 7,36 réponses indéterminées dans la condition fond gris et 7,12 réponses indéterminées dans la condition 'achromatique'. La différence n'est pas significative ($F(1,24)=0,056$ $p=.82$). Ils commettent en moyenne 9,04 intrusions dans la condition fond gris et 10,96 intrusions dans la condition 'achromatique'. La différence n'est pas significative ($F(1,24)=2,131$ $p=.16$).

De la théorie d'Intraub (1985), dérive la prédiction que davantage d'erreurs seront obtenues dans la condition 'fond gris'. Nous avons montré que les fréquences de réponses correctes, de réponses indéterminées et d'intrusions ne sont pas significativement différentes dans les deux conditions. Ceci constitue le premier élément empirique pour contester l'hypothèse d'indépendance conceptuelle proposée par Intraub (1985). Nous allons maintenant approfondir les analyses à travers la distribution des intrusions.

Distribution des intrusions

L'analyse de la distribution des intrusions ne peut être conduite sur la base des données individuelles, car trop peu d'informations sont collectées à ce niveau. Cependant, cette série de données constitue un élément d'importance pour l'explication des intrusions, et nous souhaitons tout de même examiner son comportement. Nous avons retenu une stratégie de description des données composites. Les intrusions de tous les sujets sont agglomérées, et des analyses descriptives sont menées sur ces données. La distribution des intrusions pour l'ensemble des sujets est

reportée à la figure 89. On observe classiquement que le nombre d'intrusions est plus important aux positions '-1' et '+1'.

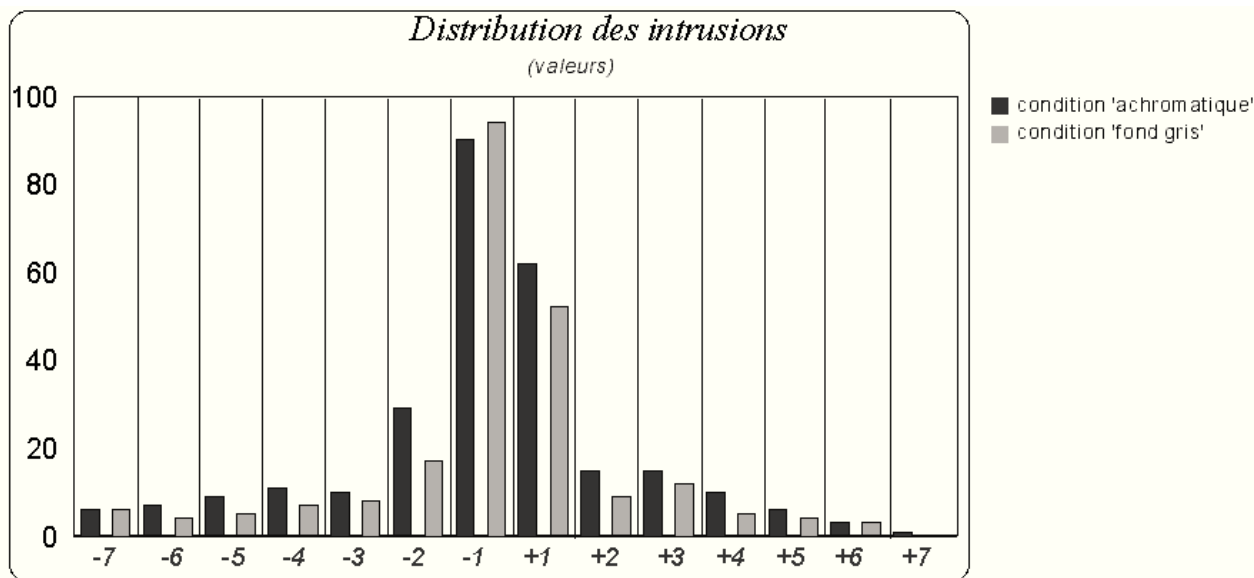


Figure 89 : distribution des intrusions dans les deux conditions expérimentales. Fréquences absolues des intrusions en fonction de la position.

La forme de cette distribution (sur quelles positions se portent les intrusions ?) est biaisée par les légères différences de fréquences d'intrusions (un peu plus d'intrusions ont été commises dans la condition 'achromatique'). Pour voir si les distributions sont structurellement identiques dans les deux conditions expérimentales, nous avons transformé les effectifs en pourcentages (voir figure 90). Ainsi, nous voyons que dans chacune des trois conditions expérimentales, environ 25% du nombre total des intrusions se portent à la position '+1'.

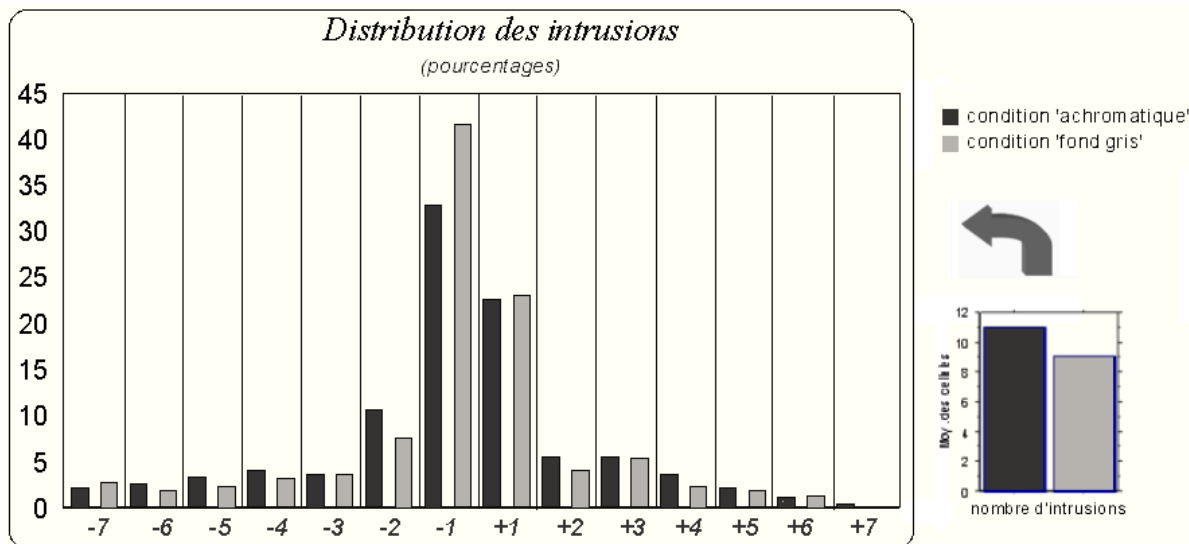


Figure 90 : distribution des intrusions dans les deux conditions expérimentales. Les distributions des deux conditions ont été normalisées sur la fréquence globale des erreurs, afin de voir si la répartition des intrusions dans les différentes positions est structurellement modifiée. On observe que les intrusions '-1' et '-2' se différencient tandis que les intrusions '+1' restent stationnaires dans les deux conditions expérimentales.

Les distributions des deux conditions présentent quelques différences. On observe le même pourcentage d'intrusions '+1'. En revanche, on observe légèrement moins d'intrusions '-1' et légèrement plus d'intrusions '-2' dans la condition 'achromatique', relativement à la condition 'fond gris'. Tout se passe comme si certaines des intrusions '-1'

de la condition achromatique s'étaient reportées sur la position '-2'. Cette légère variation est peut-être liée au niveau d'activité suscitée par les cibles des deux conditions expérimentales. La cible sur 'fond gris' est plus lumineuse que la cible 'achromatique' (voir figure 86 p.133). Elle déclenche peut-être une réponse plus forte des aires visuelles.

Par ailleurs, les sujets reportaient une plus grande difficulté dans la condition 'achromatique', qui n'est pas répercutée dans les performances. Si on observe le ratio [intrusions distales sur intrusions proximales], on note que davantage d'intrusions distales sont produites dans cette condition (ratio de 0,80 en condition 'achromatique', et de 0,55 en condition 'fond gris'). Cette augmentation des intrusions distales par rapport aux intrusions proximales et l'émoussement global de la distribution sont typiques d'une difficulté accrue de la tâche, poussant les sujets à donner davantage de réponses hasardeuses (voir figure 63 p.95). Enfin, la difficulté de la tâche s'accompagne usuellement d'une postériorisation de la distribution. On observe effectivement une légère postériorisation dans la condition 'achromatique'. L'API (voir page 84) de la condition 'fond gris' est de -0,337, et celui de la condition 'achromatique' est de -0,286. Tous ces éléments suggèrent que la seule différence affectant les deux distributions est liée à une légère disparité entre les difficultés des deux conditions.

A l'incidence de cette expérience, nous observons également que les deux distributions se différencient de celles obtenues dans les autres expériences de la présente thèse. Usuellement, les intrusions proximales ('-1' et '+1') représentent environ 80% du total des intrusions, dont environ 60% pour les intrusions '-1' et 20% pour les intrusions '+1'. Dans la présente expérience, les intrusions proximales ('-1' et '+1') représentent environ 60% du total des intrusions, dont 37% pour les intrusions '-1' et 23% pour les intrusions '+1'. La différence entre la présente expérience et les autres (expériences 1, 2, 4 et 5) affecte sélectivement les intrusions '-1'. Encore cette fois, la difficulté pourrait être la première explication, et l'absence d'information chromatique serait la seconde piste à explorer. L'absence de focalisation sur l'information de couleur pourrait modifier structurellement le comportement des intrusions.

Discussion

Cette expérience établit que les intrusions sont aussi nombreuses, et se distribuent de manière très semblable, que le critère de réponse soit conjoint ou disjoint de la cible. Les légères différences dans les distributions semblent exclusivement liées à la tendance des sujets à produire des réponses hasardeuses, en raison de la difficulté accrue de l'une des deux tâches. La présente expérience suggère que les conjonctions illusoire ne peuvent pas être expliquées par l'indépendance conceptuelle de la cible et de son critère de réponse. La réfutation de cette hypothèse n'exclut pas, a priori, que les conjonctions illusoire (toutes ou parties) résultent d'un mécanisme de confusion dans un buffer de très brève persistance. Mais si buffer il y a, il ne peut être que pré-catégoriel (eg. mémoire iconique¹ de Neisser, 1967 ;

¹ Les travaux de Sperling, (1960) établissent que davantage d'information est disponible pendant le traitement perceptif, que le sujet n'en reporte quelques secondes plus tard au moment du rappel immédiat. Il en dérive le concept de mémoire iconique (Neisser, 1967), c'est-à-dire d'une mémoire dotée d'une grande capacité, mais pré-catégorielle (encore appelé registre sensoriel). Cette forme de mémoire dure quelques centaines de millisecondes. Elle précède la consolidation en mémoire court-terme, permettant quant-à elle le rappel du nom de l'image, plusieurs secondes après sa présentation (Wheeler et Treisman, 2002).

Sperling, 1960). Ces résultats plaident davantage en faveur de Treisman (Treisman & Gelade, 1980, Treisman, 1986, 1996, 1998) et son fichier-objet (une représentation formée sous l'action du faisceau attentionnel sur la trame perceptive, voir figure 62 p.92).

L'examen des plages temporelles sur lesquelles se trouvent des intrusions proximales permet de poursuivre la discussion sur un éventuel buffer. Dans les expériences RSVP d'Intraub (1985, 1989), ne sont reportés que les intrusions des 3 positions avant et après la cible. Il n'est par conséquent pas possible d'apprécier sur quelles positions les intrusions proximales se détachent des intrusions distales. Dans nos expériences, on observe que les intrusions sont concentrées sur les positions occupées par le distracteur '-1', la cible et le distracteur '+1' (eg. voir figures 74, 81, 89 et 98, pp.116, 125, 135 & 144). Cette concentration se réalise sur un intervalle temporel d'environ **350 millisecondes**, soit justement l'ancrage perceptif d'un buffer de 3 objets centré sur la cible, lorsque les images sont présentées à 9 ou 10 images/seconde. Si les conjonctions illusoires résultaient d'un **mécanisme unitaire** (sélection aléatoire de l'image porteuse du cadre parmi un buffer de 3 images, Intraub, 1985), alors ce délai supporterait la théorie de la mémoire conceptuelle d'Intraub.

Cependant, des données suggèrent que les intrusions '-1' et '+1' se comportent en relative **indépendance** (voir expériences 1, 2 et 4). Dans cette perspective, nous devons considérer deux plages temporelles distinctes pour la formation des conjonctions illusoires, occupant chacune **228 millisecondes** (les 114 ms du distracteur '-1' et les 114 ms de la cible d'une part ; ainsi que les 114 ms de la cible et les 114 ms du distracteur '+1' d'autre part). Cette plage temporelle des intrusions proximales recouvre principalement le traitement perceptif pur (150 millisecondes selon Thorpe, Fize et Marlot, 1996), déduction faite d'éventuelles latences de consolidation en mémoire. Elle est donc compatible avec une explication perceptive ou attentionnelle des conjonctions illusoires (Treisman & Gelade, 1980). Remarquons toutefois qu'Intraub n'exploite pas l'hypothèse unitaire pour soutenir sa théorie, et qu'au contraire, elle argumente autour de cette seconde proposition. Evoquant alternativement les intrusions '-1' comme résultant d'un traitement prioritaire du critère de réponse, et les intrusions '+1' comme résultant d'un traitement prioritaire de la cible, elle concède à l'existence de deux plages distinctes de formation des intrusions, ayant une latence de 228 millisecondes (Intraub utilise les mêmes SOA que ceux utilisés dans la présente expérience). Or, si les conjonctions illusoires étaient liées à la consolidation en mémoire, la plage temporelle pertinente devrait englober 200 à 250 millisecondes supplémentaires (eg. Tovee & Rolls, 1995), et on devrait observer une dispersion des intrusions proximales sur les rangs '-3', '-2', '+2' et '+3', ce qui n'est pas le cas. **Ces résultats suggèrent que les conjonctions illusoires en RSVP, tout comme les conjonctions illusoires en recherche spatiale, résultent d'une intégration perceptive déficiente.**

Les récents travaux de Keysers & Perrett (2002) replacent le liage perceptif en RSVP au sein d'un continuum intégration/interruption. L'expérience suivante nous permettra d'identifier ces deux types d'erreurs et leur position dans la distribution des intrusions.

Expérience 4 : intégration et interruption ?

Keysers & Perrett (2002) présentent une étude en RSVP réalisée à la fois chez le singe (enregistrements unitaires des réponses neuronales) et chez l'homme (étude comportementale), et inscrite dans la théorie du masquage. Les auteurs proposent de rendre compte des deux formes de masquage, par intégration et par interruption (Turvey, 1973, Breitmeyer, 1984, Enns & DiLollo, 2001, revue dans Humphreys & Bruce, 1989) grâce à la notion plus générale de compétition neuronale (voir figure 91). Les stimuli étaient différents visages de singes orientés de face, de profil, de dos, ou selon différents angles intermédiaires. Les neurones d'une structure visuelle, STS (*superior temporal sulcus*) semblent refléter les profils de performance. Une partie de ce sulcus qui borde la partie antérieure de IT (Logothetis, 1998) semble impliquée, chez le singe, dans le codage de haut-niveau des visages et la direction du regard (Perrett & Oram, 1993). Dans la tâche RSVP, la représentation vainqueur de la **compétition** est celle qui a engendré l'activité neuronale la plus massive dans STS pendant la période de recouvrement temporel des décharges. Puisque la phase initiale de l'activité neuronale est intense, il existe une prédisposition pour le nouveau stimulus à emporter la compétition, mais « la compétition diffère de l'interruption en ce qu'elle n'a aucun a priori sur l'ordre temporel des stimuli. [...] Le nouveau stimulus entre en compétition avec l'ancien (masquage rétro-actif), et l'ancien stimulus est lui aussi en compétition avec le nouveau (masquage pro-actif). » (Keysers & Perrett, 2002, pp.120-121). Ainsi, Keysers & Perrett établissent un continuum entre les **mécanismes d'interruption et d'intégration**.

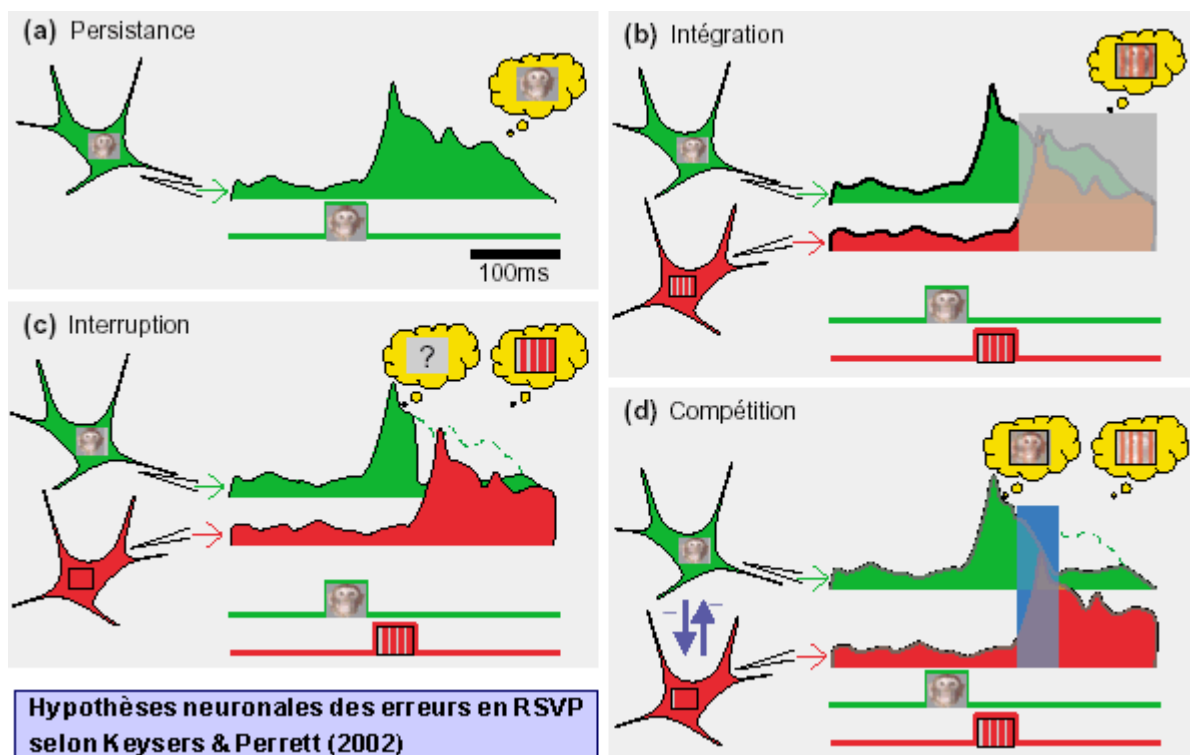


figure 91 : hypothèses neuronales pour le masquage par Keysers & Perrett, (2002). Partant du constat que les neurones du système visuel persistent à décharger au delà du temps réel d'affichage de l'image (a), les auteurs proposent d'unifier les deux concepts d'intégration (b) et d'interruption (c) dans un unique concept, celui de compétition neuronale (d). Reproduit de Keysers & Perrett, (2002)

L'originalité de l'étude de Keyzers & Perrett (2002) est d'évoquer le mécanisme d'interruption aux cotés du mécanisme d'intégration. La plupart des autres études en RSVP ne s'expliquent les intrusions que par le mécanisme d'intégration, sous différentes terminologies (dissociations visuelles, Intraub, 1989 ; intrusions, erreurs de liage, ou conjonctions illusoire Botella & Ericksen, 1992 ; Botella, Garcia, Barriopedro, 1992 ; Chun, 1997 ; Botella, Barriopedro & Suero, 2001).

En revanche, nos précédentes analyses (expérience préliminaire A) suggéraient l'existence d'un mécanisme d'interruption, en marge du mécanisme d'intégration (voir pp.81-82). Au cours des expériences préliminaires, on observait que la plupart des stimuli développent des affinités spécifiques avec un groupe restreint de stimuli. Cette affinité sélective se traduit par quelques fréquences élevées dans les lignes (intrusions subies) ou dans les colonnes (intrusions causées) de la matrice d'association (eg. table 2 p.80) et suggèrent un mécanisme **d'intégration**. En revanche, un petit nombre de stimuli se comportent en **interrupteurs**. Ils intrusent les autres stimuli de manière peu sélective, et on observe des fréquences élevées dans toute une colonne de la matrice d'association. En conséquence de cette activité de masse, ils provoquent davantage d'intrusions que les distracteurs n'agissant sélectivement que sur un ou deux stimuli. On les retrouve donc isolés en haut à gauche sur le graphe des intrusions (voir figure 50 p.81). Dans ces analyses préliminaires, **les deux mécanismes semblent assignés à des stimuli particuliers**.

Stratégie

L'objectif de la présente expérience est d'**apprécier l'existence des deux mécanismes d'intégration et d'interruption dans le corpus d'erreurs relevé lors d'expérimentations RSVP**. La stratégie consiste à présenter deux réseaux de stimuli distincts, ne comportant qu'une différence (voir figure 92). La condition 'normale' utilisera un groupe de 10 stimuli (le même que celui utilisé dans les expériences 1, 2 et 5). La condition 'substituée' utilisera 9 des stimuli précédents, mais, en substitution du dixième stimulus, le scorpion, elle proposera un zèbre (cette substitution résultant d'un tirage au sort).

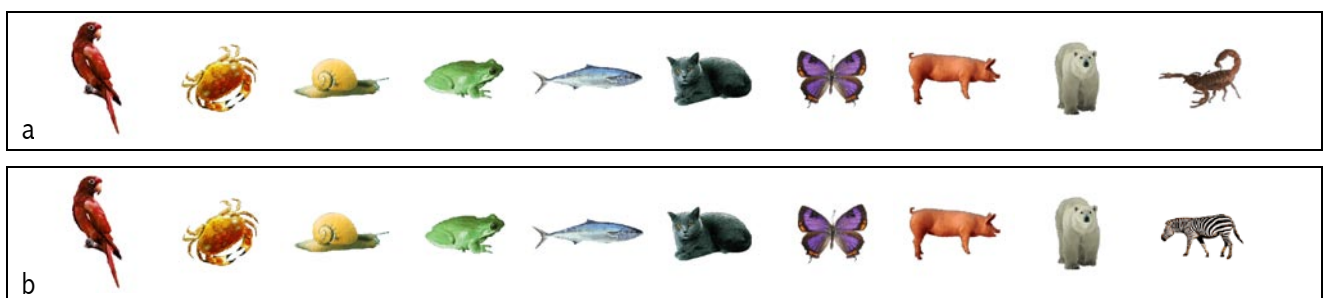


Figure 92 : les deux groupes d'image utilisés dans cette expérimentation. En a, la condition 'normale' ; en b, la condition 'substituée'.

Selon nos observations préliminaires (expérience A), les **stimuli interrupteurs** intrusent indifféremment tous les stimuli. Aussi, ce type d'erreurs sera **insensible à la modification du réseau de stimuli**. Pour identifier ce mécanisme, nous chercherons les intrusions qui restent stables d'une condition à l'autre.

En revanche, les **stimuli intégrateurs** ont des affinités sélectives avec quelques stimuli bien précis. Aussi la modification du réseau de stimuli est supposée les affecter. Nous chercherons à établir si **les stimuli les plus affins** au couple de substitution **seront les plus affectés** (voir figure 93).

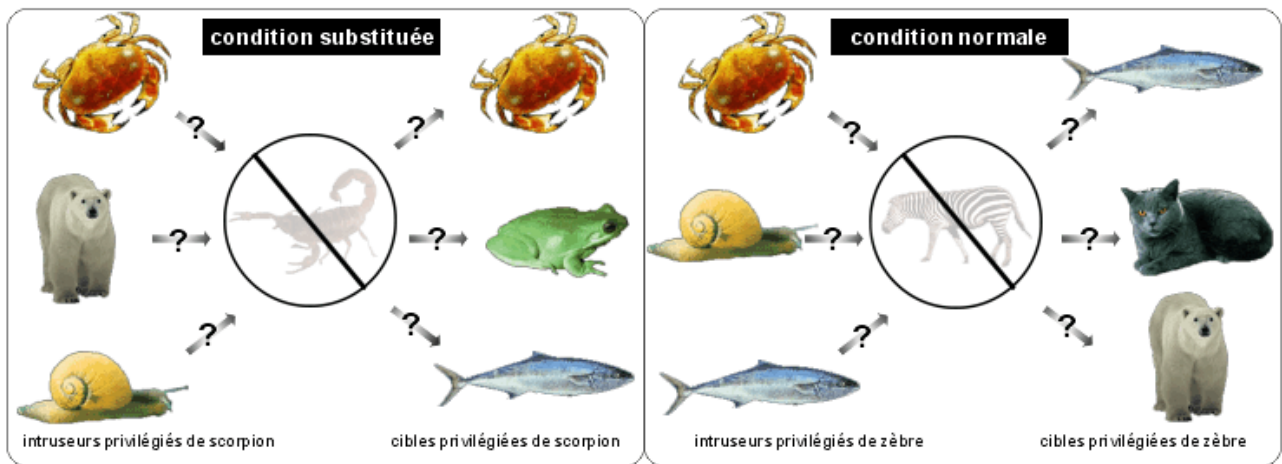


Figure 93 : perturbation des stimuli affins au couple de substitution dans les deux conditions expérimentales. Dans la condition 'substituée', les intruseurs de prédilection du scorpion ne trouvent plus leur cible par rapport à la condition « normale ». De même, les cibles préférées de scorpion cessent de subir ces intrusions. De même, dans la condition normale, le groupe de stimuli affins de zèbre cesse de fonctionner, perturbant ainsi le profil fin des résultats.

Sujets

Les sujets étaient 24 adultes volontaires pour l'expérience (12 femmes, 12 hommes), âgés de 21 à 57 ans (moyenne d'âge : 31 ans). Les sujets ont reporté une vision normale, ou corrigée à la normale.

Matériel & méthodes

Les sujets étaient soumis à 110 essais RSVP, voir exemple figure 94. La procédure était identique au protocole générique décrit pp.109-111 . Les essais étaient répartis de la manière suivante : 10 essais avec le zèbre en tant que cible, et 10 essais avec le scorpion en tant que cible, 45 essais avec le zèbre en distracteur, 45 essais avec le scorpion en distracteur (voir figure 95). Le zèbre et le scorpion sont moins souvent visionnés par le sujet (elles n'apparaissent chacune que dans 50% des essais) mais elles sont davantage présentées en tant que cible. Nous tentons ainsi de créer une exposition comparable aux stimuli invariants et aux stimuli de substitution.

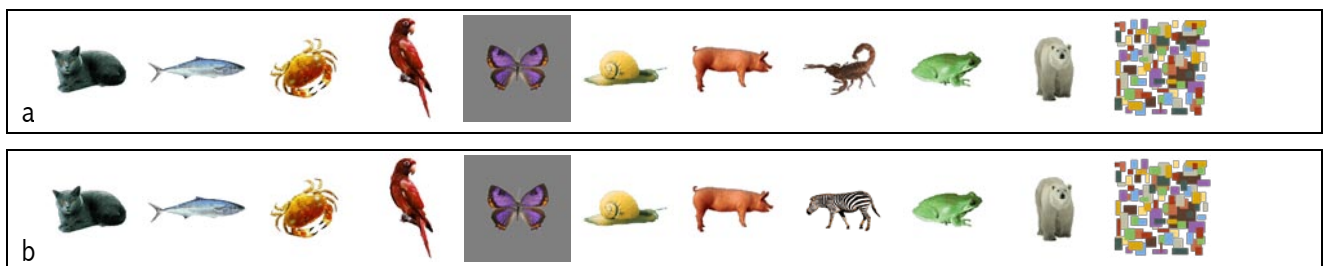


Figure 94 : exemples d'essais utilisés : en a, séquence d'images d'un essai de la condition « scorpion ». En b, son apparié de la condition « zèbre ».

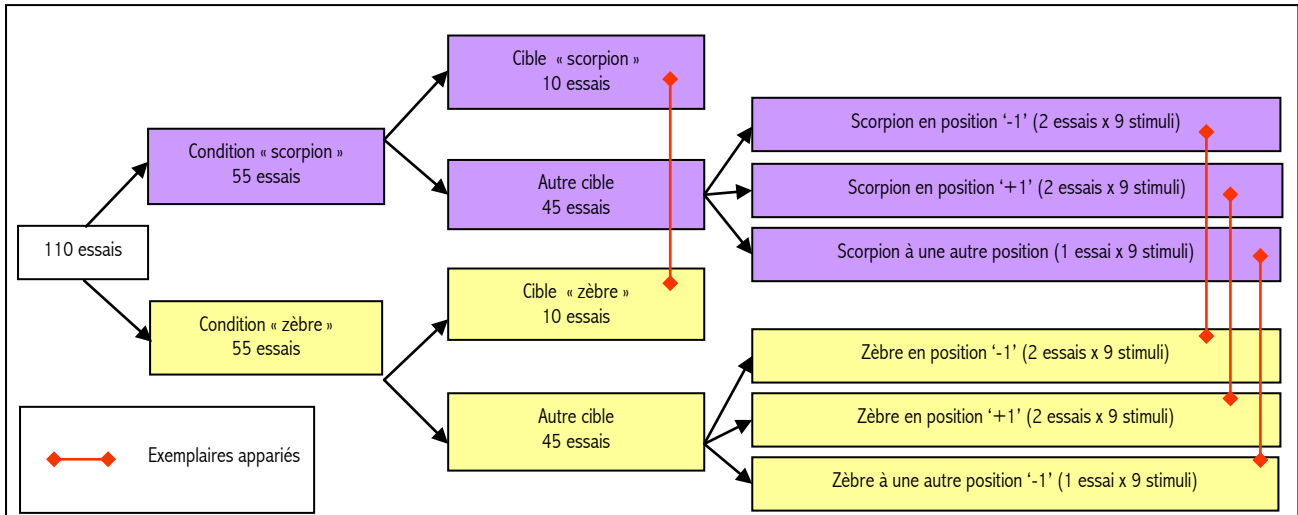


Figure 95 : répartition des essais de l'épreuve 'systémique'. L'ordre des essais est géré par randomisation réelle. Les essais des deux conditions expérimentales sont structurellement identiques. Seule varie la présentation du zèbre ou du scorpion.

Résultats

Le cas des intégrateurs : changement des intrusions en fonction de l'affinité au couple de substitution

Cette analyse se déroule en trois étapes. Premièrement, nous avons calculé l'**affinité** de chacun des 9 stimuli invariants au couple de substitution (scorpion/zèbre). Deuxièmement, nous avons placé les stimuli dans le graphe des intrusions, pour chaque condition expérimentale. Nous avons ainsi mesuré la **distance** entre les deux positions de chaque stimuli. Enfin, nous avons calculé la **corrélation** entre l'affinité au couple de substitution et la distance dans le graphe des intrusions.

Le calcul de l'affinité au couple de substitution est réalisé de la manière suivante : pour chacun des 9 stimuli invariants, on calcule sa distance au centre du référentiel, dans l'espace de dimension 4 [intrusions causées à scorpion ; intrusions subies par scorpion ; intrusions causées à zèbre ; intrusions subies par zèbre]. Plus la distance est grande, plus le stimulus partage des intrusions avec le couple de substitution. Le résultat de cette analyse est reporté à la figure 96.

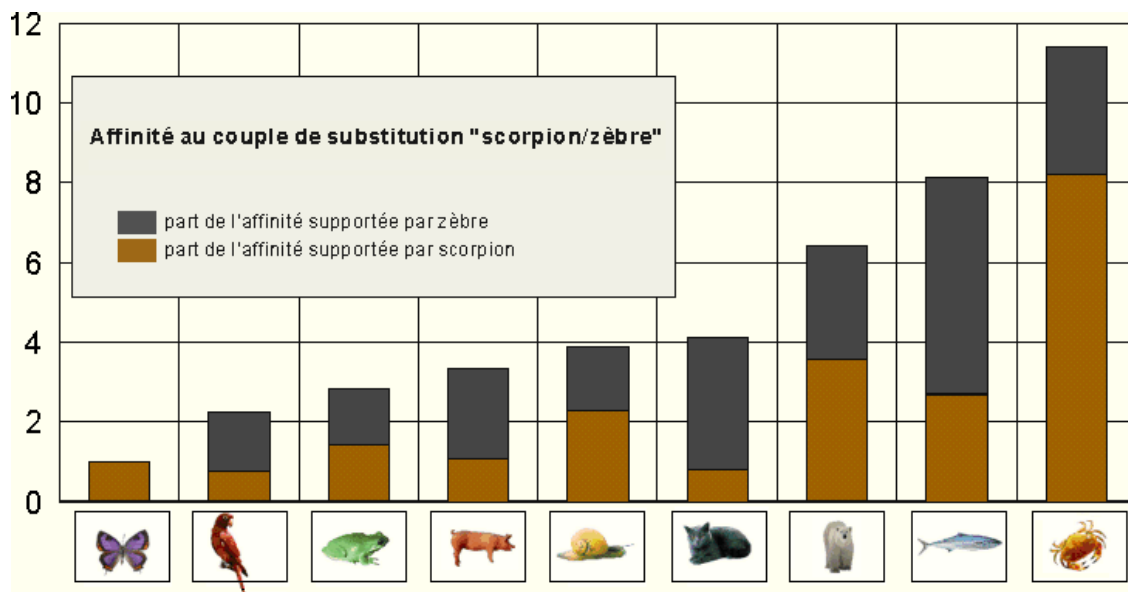


Figure 96 : affinité des stimuli invariants au couple de substitution

Ensuite, nous avons réalisé le graphe des intrusions pour les deux conditions expérimentales. Le graphe est reporté à la figure 97. On en déduit la distance entre la position de chaque stimulus dans la condition normale et dans la condition substituée.

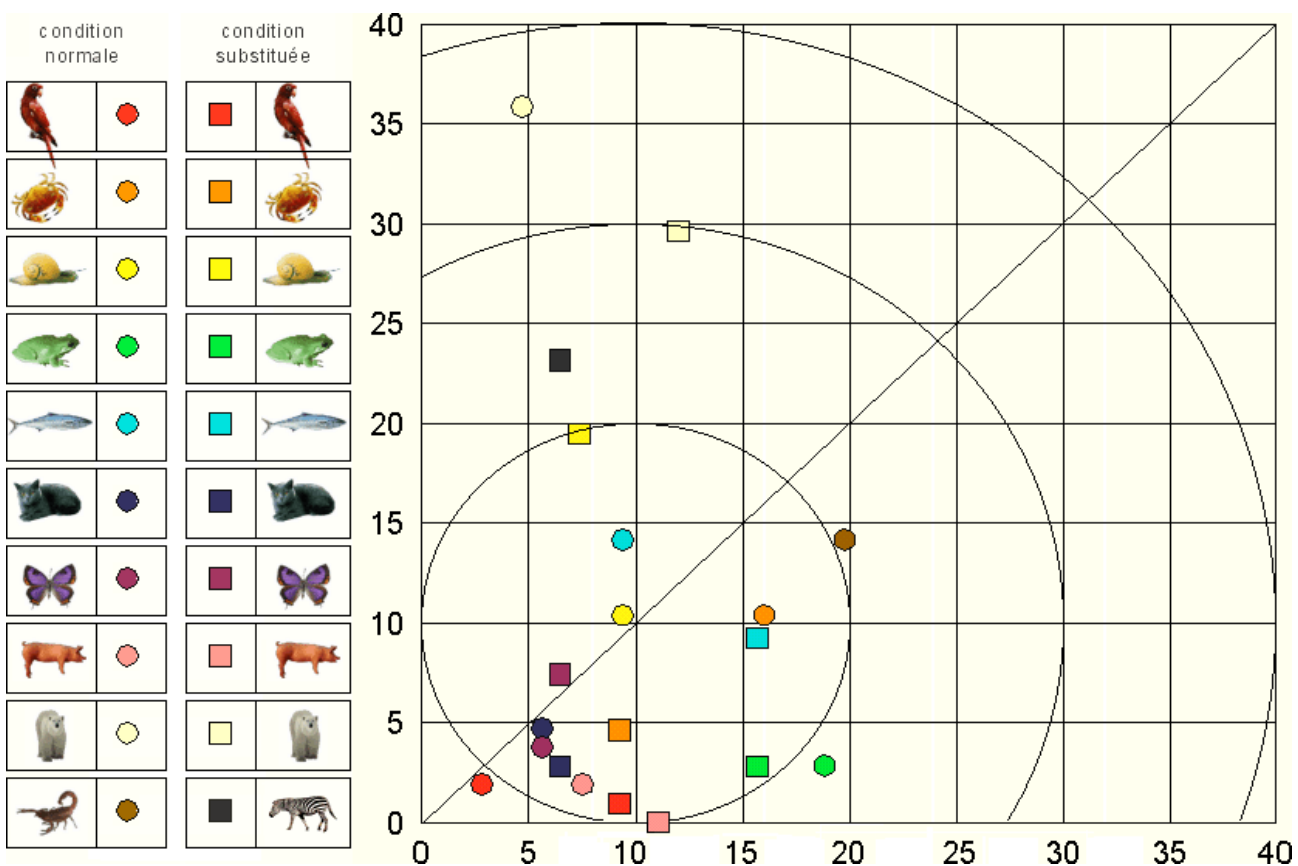


figure 97 : graphe des intrusions pour les deux conditions expérimentales (intrusions subies en abscisses et intrusions causées en ordonnées). On observe qu'une partie des stimuli n'est pas affectée par la condition expérimentale (les deux exemplaires du stimuli sont proches l'un de l'autre) tandis que d'autres sont affectés (ours, crabe, escargot, poisson)

Cette distance corrèle avec l'affinité au couple de substitution ($r=.61$). Elle s'élève même à .77 si on ne calcule l'affinité que sur la base des intrusions subies par le couple de substitution. Cette fréquence des intrusions subies était le paramètre qui différenciait le plus le scorpion du zèbre (voir annexe XVII). Cette corrélation élevée suggère qu'**une partie des intrusions est effectivement supportée par un mécanisme d'intégration**. Plus un stimulus interagit avec les deux stimuli de substitution, plus il subit des modifications lors du changement de conditions expérimentales.

Le cas des interrupteurs : permanence des intrusions à travers les conditions expérimentales

fréquences observées	hits	indet	intr	fréquences théoriques	hits	indet	intr	résidus	hits	indet	intr
hits	1031	37	71	hits	986	54	99	hits	45	-16	-28
indet	50	14	10	indet	64	3	6	indet	-14	11	4
intr	62	11	34	intr	93	5	9	intr	-31	6	25

Table 9 : répartition des réponses de la condition substituée (en colonne) en fonction des réponses de la condition normale (en ligne). Les résidus positifs de la diagonale montrent que les sujets ont tendance à donner deux réponses de même classe en condition normale et en condition substituée. (hits : réponses correctes ; indet : réponses indéterminées ; intr : intrusions)

Les réponses des sujets sur les essais appariés ont tendance à se reproduire comme le montre la table 9. Les résidus convergents (même classe de réponse pour les deux essais appariés) représentent 4% des effectifs observés pour les réponses correctes, 79% pour les intrusions et 74% pour les réponses indéterminées. Si la convergence est peu significative pour les réponses correctes, en revanche, elle représente une fraction importante des observations pour les essais échoués.

Un sujet ayant formé une intrusion pour un essai de la condition normale aura **tendance** à former à nouveau une intrusion dans la condition substituée. Cette **répétition d'intrusion** ne peut être expliquée par une stabilité des conjonctions illusoires, car la manipulation expérimentale a affecté le réseau de stimuli, comme l'a démontré le paragraphe précédent. Malgré tout, on observe, pour 31% des couples d'essais appariés dont l'un au moins présente une intrusion, une répétition de l'intrusion dans la seconde condition.

Vingt de ces double-intrusions (soit 59% des double-intrusions) mentionnent deux fois le même intruseur. Cette fréquence représente 18% des essais appariés comportant au moins une intrusion. La distribution de ces intrusions (voir figure 98) est différente de la distribution globale. On observe **exclusivement des intrusions '-1'**.

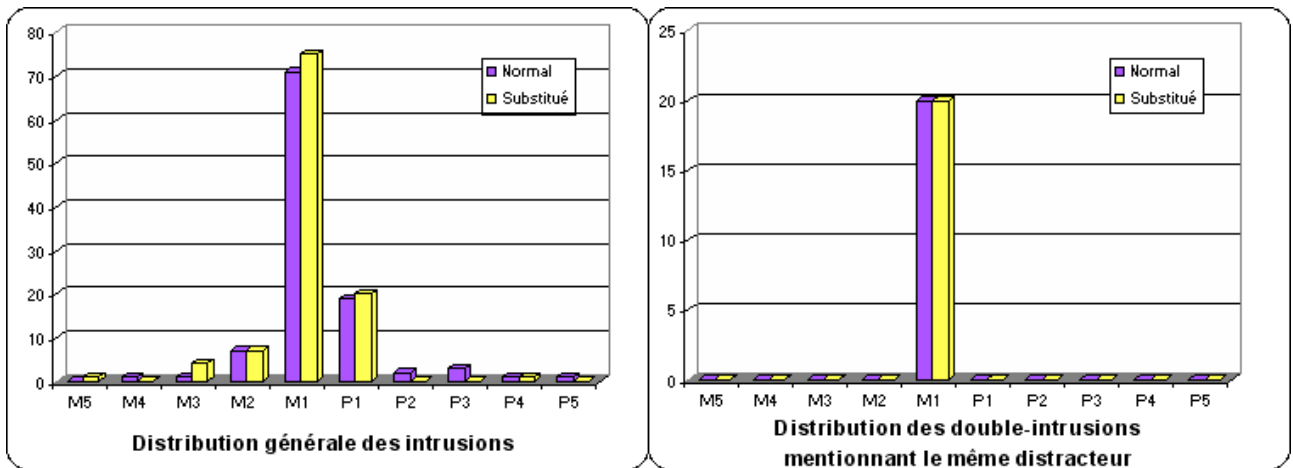


Figure 98 : distribution des intrusions pour les intrusions et pour les conjonctions d'intrusions

Les distracteurs ne semblent pas avoir d'égales chances de produire ces double-intrusions. On retrouve comme intruseurs de ces essais, **l'ours** (11 essais), **l'escargot** (4 essais) et **le poisson** (3 essais). Ces mêmes stimuli occupent les positions haut-gauche dans le graphe des intrusions (voir figure 99). Ces positions traduisent un comportement très fréquemment intruseur et rarement intrusé, incompatible avec une affinité sélective à quelques cibles, mais compatible avec une action de masse sur tous les stimuli. Nous nommons ces stimuli interrupteurs, car ils empêchent la formation de toute représentation dans leur proximité. A contrario des présupposés théoriques du masquage, ces interrupteurs fonctionnent de manière proactive : ils empêchent le traitement de la cible située juste après eux. Notons que ces trois stimuli sont parmi les 4 plus clairs du groupe de stimuli (voir annexe XVII), alors que le fond de l'écran est noir. Une première explication, périphérique, serait que la forte activité des cellules rétiniennes demande une certaine latence pour retourner à son niveau de base. Une seconde explication plus centrale serait que la forte activité générée par ces interrupteurs dans les structures visuelles corticales demande davantage de temps pour diminuer et biaise ainsi la compétition neurale en leur faveur. Mais ces explications ne semblent pas suffire à elles seules, un autre stimulus clair, la grenouille, ne se comportant pas du tout comme un interrupteur.

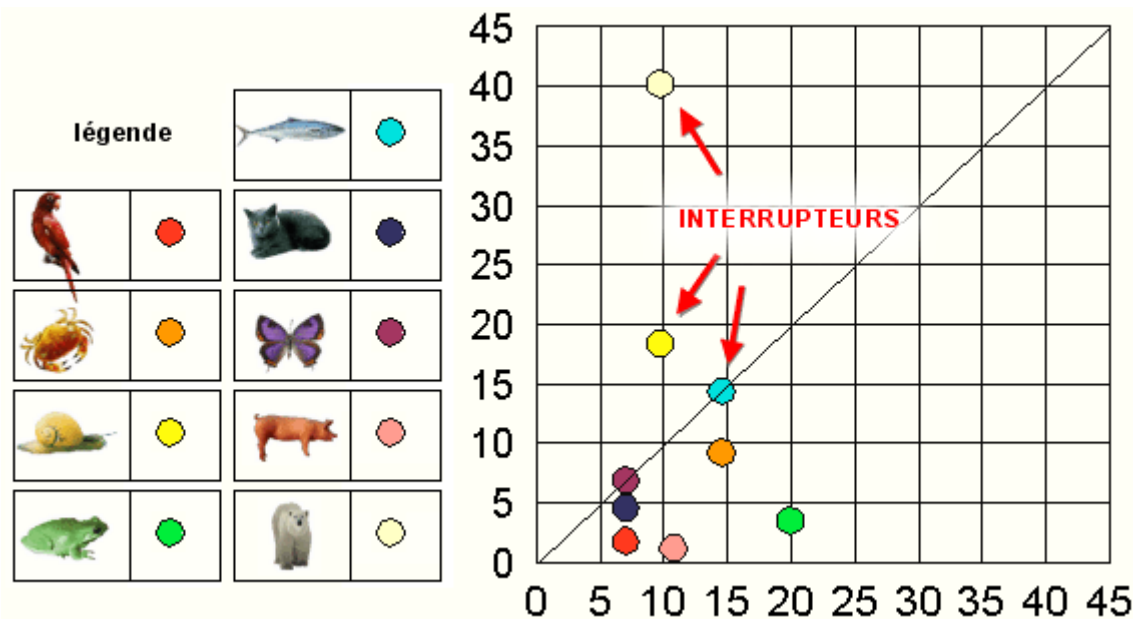


Figure 99 : position des intruseurs-interrupteurs dans le graphe des intrusions. Le graphe est calculé pour les neuf stimuli invariants seulement.

Discussion

Cette expérience suggère que les deux mécanismes d'intégration et d'interruption sont à l'œuvre dans la tâche RSVP. Dans cette expérience, **le mécanisme d'interruption était estimé à un cinquième des intrusions**. Les résultats suggéraient en outre que les interrupteurs sont indépendants du contexte visuel. Ils agissent quelque soit le réseau de stimuli utilisé. En conséquence, ils présentent une activité d'intrusion beaucoup moins différenciée dans la matrice d'association. Alors que les intégrateurs intrusent sélectivement une ou deux cibles, les interrupteurs agissent sur la plupart des stimuli. Cette activité massive sépare ces interrupteurs des autres stimuli dans le graphe des intrusions. Ils occupent le quadrant haut-gauche de ce graphe. Certaines propriétés visuelles, telles que la **saillance** par rapport au fond visuel, semblent nécessaires, mais ne semblent pas suffisantes. Dans la présente expérience utilisant un fond noir, les stimuli saillants étaient les stimuli clairs, tandis que dans les expériences préliminaires utilisant un fond gris clair, les stimuli saillants étaient porteurs de couleurs vives. Enfin, nos résultats suggèrent que le mécanisme se produit **exclusivement de manière proactive**. Ces deux derniers résultats sont potentiellement en rapport avec le temps nécessaire aux cellules rétinienne ou corticales pour retourner à leur niveau de base après une forte activation.

L'intégration et l'interruption ont été conçues par Keysers & Perrett (2002) comme les deux pôles d'un continuum (voir figure 100). A contrario, dans cette expérience, nous les avons circonscrit à des stimuli particuliers, mais le petit échantillonnage obtenu dans cette expérience ne rend visible que les exemples les plus marquants, et il est possible que les autres stimuli soient impliqués à un degré moindre dans la formation de réponses par interruption.

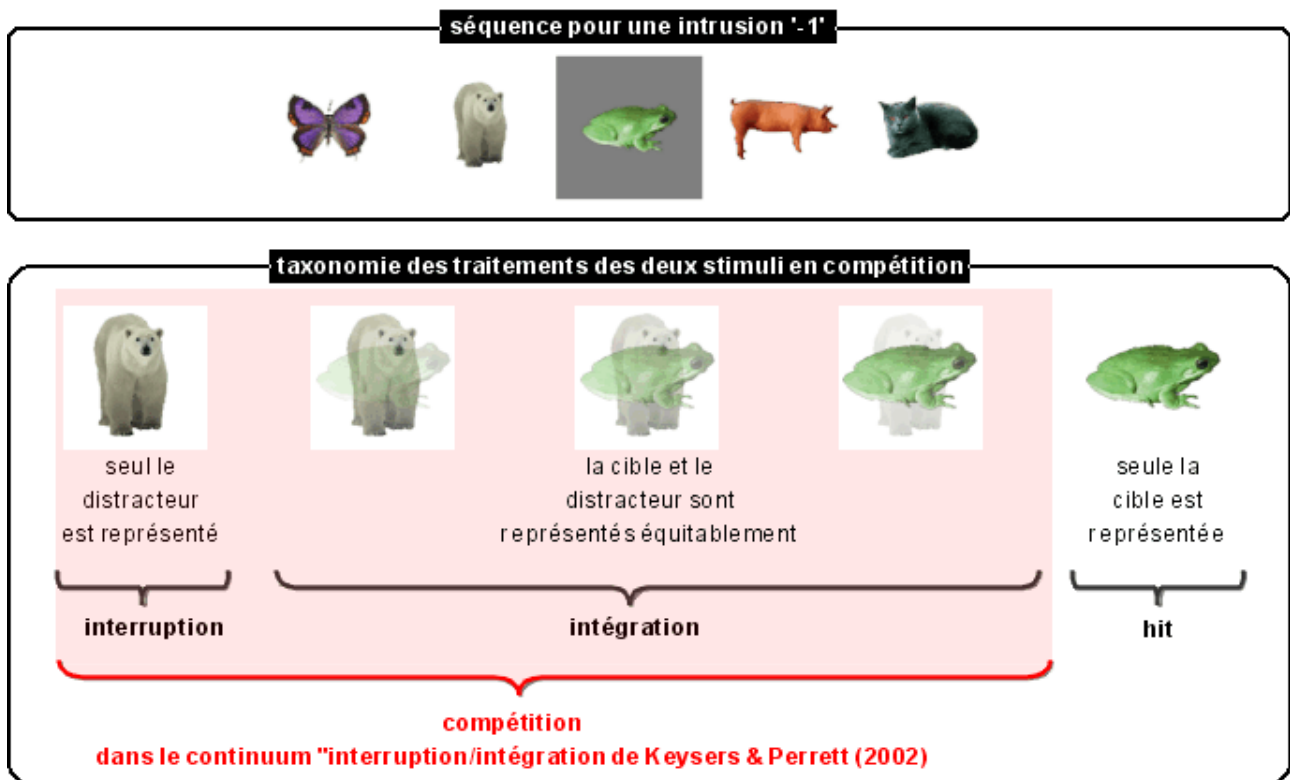


Figure 100 : le continuum interruption/intégration d'après Keyzers & Perrett (2002). Nous avons étendu le continuum au cas où la cible est représentée seule. Les deux pôles idéaux sont ainsi la représentation unique du distracteur (interrupteur) et la représentation unique de la cible (hit). Entre ces deux pôles se trouvent les intégrations.

Cette hypothèse d'un continuum entre interruption et intégration éclaire plusieurs observations restées en suspens jusque lors. La compétition entre la cible et le distracteur explique en partie l'asymétrie entre les intrusions pré- et post-cible. Lorsque l'interrupteur (eg. l'ours) est présenté en '-1' et qu'un autre stimulus (eg. la grenouille) est présenté en tant que cible, alors la probabilité d'une intrusion est forte. En revanche, lorsque l'interrupteur est la cible, et qu'un stimulus quelconque est présenté en '+1', alors la probabilité que ce distracteur interfère avec le traitement de la cible est très faible. Cette règle explique pourquoi les stimuli les plus intruseurs sont également les moins intrusés (voir p.82). **La part des intrusions liée au mécanisme d'interruption est fondamentalement asymétrique et repose sur les propriétés du distracteur.**

Par ailleurs, cette hypothèse permet de délimiter le rôle de la similarité de la cible et du distracteur sur les intrusions. On peut désormais s'attendre à ce que les caractéristiques communes expliquent les conjonctions illusoire lorsqu'un maximum de caractéristiques des deux stimuli sont co-présentes (ie. au centre du continuum de la figure 100). En revanche, **lorsque le produit de la compétition est très asymétrique** (représentation quasi-exclusive du distracteur), alors **on peut s'attendre à ce que la similarité cible-distracteur soit peu explicative des intrusions**. Ceci explique pourquoi les précédentes régressions multiples n'expliquaient pas une grande part de variance (voir p.78 et annexe XIII).

Il nous reste maintenant à sélectionner des configurations perceptives (couples cibles/intruseurs) aptes à présider à la formation de fréquentes conjonctions illusoire. Cette sélection sera réalisée au paragraphe suivant.

Expérience 5 : Sélection des conjonctions illusoires

La survenue des conjonctions illusoires, dans les différents dispositifs empiriques permettant de les observer, semble être un évènement stochastique. La présentation répétée du même matériel chez un sujet donné donnera quelques-fois lieu à une conjonction illusoire et quelques-fois non, sans qu'il soit possible de le prévoir. Cette imprédictibilité est attribuée à la nature labile et instable de l'attention, (eg. Treisman, 1996, 1998). Si l'attention a solidarisé les caractéristiques entre-elles, alors la réponse du sujet sera correcte, alors que si le faisceau attentionnel n'a pu solidariser les caractéristiques, une conjonction illusoire est susceptible de se produire. Cette **hypothèse localisationniste** consensuelle repose sur des données neuropsychologiques (eg. patient RM, Friedman-Hill, Robertson & Treisman, 1995) suggérant que l'intégrité de la partie **pariétale du réseau attentionnel** supporte **le liage des caractéristiques des objets**. Dans cette perspective, ne sont conjonctions illusoires que les erreurs avérées. Ainsi, ces paradigmes utilisant des objets simplifiés (souvent conjonctions forme + couleur) dichotomisent les traitements visuels. Un liage perceptif ne peut être que correct (association de la bonne couleur à la forme) ou incorrect. En d'autres termes, il est implicitement admis que la formation de conjonctions illusoires dépend autant, sinon plus, de l'état interne du sujet que du matériel visuel soumis à sa perception visuelle, rendant vaine la recherche des conditions externes affectant la formation des conjonctions illusoires (mais voir Ivry & Prinzmetal, 1991).

Dans une **perspective distribuée** du « binding problem », l'imprédictibilité de la survenue des conjonctions illusoires peut être reliée à la notion de **décision perceptive**, l'attention devenant alors une condition nécessaire, mais non suffisante. Par exemple, dans le paradigme RSVP utilisant des images d'objets, la compétition des objets visuels implique une cible et un seul distracteur. Lorsqu'une telle compétition se déroule, un traitement « défaillant » (liage incertain) donne parfois lieu à une erreur, mais parfois également à une réponse correcte¹. Le mécanisme de conjonction illusoire est donc sous-estimé, car **une partie des réponses correctes reflète une difficulté de liage** ayant finalement aboutit « fortuitement » à la bonne réponse. Botella & Ericksen (1992) soulèvent ce même problème en manipulant la présence de la cible et des distracteurs dans le formulaire-réponse, et en suggérant qu'une réponse unique n'est peut-être pas le reflet fidèle des mécanismes sous-jacents. Plus récemment, cette hypothèse est bien décrite dans Keyesers & Perrett (2002) qui généralisent la notion de compétition cible/distracteur en RSVP.

Alors que les autres études en RSVP, contraintes par l'hypothèse dichotomique, ne considèrent les conjonctions illusoires que sous un angle holistique (fréquence des intrusions ; distribution des intrusions), il ne nous semble pas vain d'essayer de **sélectionner des configurations perceptives meilleures que d'autres pour enregistrer des conjonctions illusoires**, en vue d'investigations neurophysiologiques. On observe effectivement que les distracteurs ne

¹ Dans les paradigmes utilisant des objets réduits à une forme et une couleur, le liage des caractéristiques a peu de degrés de libertés. Il est peu plausible qu'un liage incorrect se traduise par une réponse correcte. De cette restriction dérive la vue dichotomique sur le binding. En revanche, pour des objets complexes, un liage partiellement incorrect peu parfois aboutir à une décision perceptive correcte. On conçoit alors le binding sur un continuum.

sont pas tous aussi efficaces à générer des intrusions et que les cibles ne sont pas toutes aussi susceptibles d'en recevoir (voir expériences préliminaires A et B, et expérience 4).

Cependant, nous avons mis en évidence, dans les expériences précédentes, la mixité des erreurs, qui avait déjà été suggérée auparavant par Botella & Ericksen, (1992) ; Maki, Couture, Frigen, Lien (1997). Cette mixité des erreurs s'oppose à notre projet de sélection des configurations perceptives. Un travail de tri doit être réalisé pour minimiser les manifestations autres que les conjonctions illusoires.

Stratégie

Cette dernière expérience vise à repérer les séquences RSVP donnant lieu à de fréquentes conjonctions illusoires, et à établir des prédictions sur les mécanismes responsables des erreurs et les structures visuelles impliquées. En premier lieu, nous aborderons le problème de la structure des données. Nous avons identifié trois groupes d'intrusions significatifs sur la base de la position des erreurs, respectivement les intrusions 'plus 1', 'moins 1' et distales. Nous avons pu constater que la position de l'intruseur n'est pas un critère suffisant de ségrégation des différents types d'intrusions, car différents mécanismes occasionnent des intrusions sur une même position. La structure des intrusions sera identifiée par une analyse factorielle, en vue de conforter nos précédentes hypothèses.

Parmi ces différents types d'intrusions, nous cherchons à identifier ceux qui résultent d'une intégration entre la cible et le distracteur. Les erreurs pouvant sans ambiguïté être imputées à un dysfonctionnement de l'intégration perceptive seront présélectionnées, puis nous retiendrons celles qui présentent une forte probabilité de survenue.

Méthode

Le protocole et les sujets sont les mêmes que ceux utilisés dans l'expérience 2. Pour mémoire, les sujets étaient 22 adultes volontaires pour l'expérience (11 femmes, 11 hommes), âgés de 19 à 44 ans (moyenne d'âge : 26 ans). Ils étaient soumis à 4 sessions de 120 essais RSVP, réalisés à la même heure à quelques jours d'intervalle. Deux catégories d'images ont été étudiées, les animaux et les vêtements (voir Annexe XVI).

Résultats

Structure des intrusions : analyse en composantes principales

Nous avons analysé une matrice de données composites centrée sur les stimuli. Les observations étaient les 180 couples 'cible/intruseur' (90 pour la catégorie animaux, et 90 pour la catégorie vêtements). Les variables utilisées étaient :

- Proba_M1 : pourcentage de séquences intruseur/cible ayant donné lieu à une intrusion 'moins 1' (eg. sur 58 essais présentant le poisson en 'moins 1' et la grenouille en cible, nous avons obtenu 5,5% d'intrusions 'moins 1')

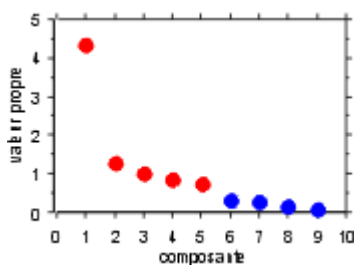
- Proba_P1 : pourcentage de séquences cible/intruseur ayant donné lieu à une intrusion 'plus 1' (eg. sur 52 essais présentant le chat
- D : nombre d'intrusions distales formées par ce couple cible/intruseur
- ISC_D : nombre global d'intrusions distales subies par la cible de ce couple cible/intruseur
- ICD_D : nombre global d'intrusions distales provoquées par l'intruseur de ce couple cible/intruseur
- ISC_M1 : nombre global d'intrusions 'moins 1' subies par la cible de ce couple cible/intruseur
- ICD_M1 : nombre global d'intrusions 'moins 1' provoquées par l'intruseur de ce couple cible/intruseur
- ISC_P1 : nombre global d'intrusions 'plus 1' subies par la cible de ce couple cible/intruseur
- ICD_P1 : nombre global d'intrusions 'plus 1' provoquées par l'intruseur de ce couple cible/intruseur

La matrice de corrélation est reportée à la figure 101.

	PROBA_M1	PROBA_P1	D	ISC_D	ICD_D	ISC_P1	ICD_P1	ISC_M1	ICD_M1
PROBA_M1	1	0,445	0,565	0,366	0,243	0,43	0,18	0,507	0,461
PROBA_P1	0,445	1	0,473	0,448	0,219	0,539	0,59	0,459	0,214
D	0,565	0,473	1	0,457	0,52	0,389	0,187	0,323	0,26
ISC_D	0,366	0,448	0,457	1	0,36	0,851	0,387	0,706	0,274
ICD_D	0,243	0,219	0,52	0,36	1	0,373	0,359	0,205	0,5
ISC_P1	0,43	0,539	0,389	0,851	0,373	1	0,429	0,848	0,286
ICD_P1	0,18	0,59	0,187	0,387	0,359	0,429	1	0,245	0,347
ISC_M1	0,507	0,459	0,323	0,706	0,205	0,848	0,245	1	0,174
ICD_M1	0,461	0,214	0,26	0,274	0,5	0,286	0,347	0,174	1

Figure 101 : matrice de corrélation des variables incluses dans la factorisation.

La factorisation a été réalisée par une Analyse en Composantes Principales avec rotation Varimax. Nous avons retenu une solution en 5 facteurs, expliquant 91,2% de variance. La matrice factorielle après rotation est reportée à la figure 102.



	1	2	3	4	5
PROBA_M1	0,30	0,06	0,82	-0,02	0,39
PROBA_P1	0,30	0,76	0,47	0,04	-0,06
D	0,19	0,12	0,70	0,62	-0,09
ISC_D	0,85	0,20	0,11	0,28	0,05
ICD_D	0,17	0,14	0,04	0,87	0,33
ISC_P1	0,91	0,26	0,13	0,16	0,11
ICD_P1	0,19	0,89	-0,09	0,17	0,24
ISC_M1	0,90	0,09	0,26	-0,05	0,07
ICD_M1	0,10	0,15	0,17	0,24	0,91

Figure 102 : chute des valeurs propres et matrice factorielle après rotation.

Le premier facteur est saturé pour les intrusions globalement subies par la cible et ce, quelque soit la position 'plus 1', 'moins 1' ou 'distale'. C'est un **facteur de vulnérabilité générale de la cible**.

Le second facteur est saturé par la probabilité d'intrusion 'plus 1' du couple cible/intruseur, mais également par l'activité d'intrusion en 'plus 1' du distracteur. C'est un **facteur spécifique des intrusions 'plus 1'**. Nos précédents résultats suggèrent que ce facteur est lié à l'intégration cible/distracteur.

Le troisième facteur est fortement saturé par la probabilité d'intrusions 'moins 1' du couple et par la fréquence des intrusions distales du couple. Ce facteur fait écho aux résultats des expériences 1 et 2. Ces expériences exhibaient un groupe d'intrusions réparties sur les positions '-1' et distales. On remarque la saturation intermédiaire de PROBA_P1, suggérant la contribution d'une partie des intrusions 'plus 1' à ce groupe, ce qui n'était pas apparu clairement au cours des expériences précédentes. Dans l'expérience 2, ce groupe d'erreurs paraissait sensible à l'apprentissage. Une interprétation plausible est celle d'un **facteur de pari sophistiqué** (Botella & coll., 2001).

Le quatrième facteur représente la fréquence des intrusions distales et l'activité d'intrusion du distracteur en distal. La présence de cette composante suggère que les intrusions distales sont elles-aussi déterminées par deux mécanismes, puisque un autre facteur (le troisième) est lui-aussi fortement saturé par les variables liées aux intrusions distales. Celui-ci semble plus spécifique dans la mesure où seules les intrusions distales sont représentées. Nous l'interprétons comme un **facteur spécifique des intrusions distales**, représentant peut-être les erreurs typiques du rappel.

Le dernier facteur est fortement saturé par l'activité d'intrusions 'moins 1' du distracteur et dans une moindre mesure, par la probabilité d'intrusions 'moins 1' dans le couple cible/intruseur. La prééminence de l'activité globale sur l'activité spécifique suggère que ce facteur dépend fortement des propriétés du distracteur. Il fait écho au comportement des interrupteurs (voir expérience 4) affectant spécifiquement la position 'moins 1' et maximisant la représentation du distracteur dans la compétition cible/intruseur. Il s'agit donc d'un **facteur compétition 'moins 1'**.

Il reste des **questions en suspens autour de ce facteur compétition 'moins 1'**. Nous avons caractérisé une partie de ce facteur à l'expérience 4. Il s'agissait d'intrusions affectant spécifiquement la position 'moins 1' et se manifestant, indépendamment du contexte formé par les autres stimuli, par des intruseurs ayant une forte saillance visuelle. Sur le continuum de compétition cible/distracteur, ces intrusions maximisent la représentation du distracteur au détriment de celle de la cible.

Il reste à savoir si des intégrations cible/distracteur s'ajoutent à ces interruptions dans ce facteur de compétition 'moins 1'. Dans l'expérience 4, les interruptions manifestes étaient estimées environ au cinquième des intrusions 'moins 1', laissant une place éventuelle pour un mécanisme d'intégration. Mais la présente analyse factorielle ne montre aucun partage de variance des deux intrusions proximales dans un facteur commun. Si intégration il y a, sa nature sera nécessairement différente de celle des intrusions 'plus 1'. Une hypothèse plus plausible est celle d'une **spécialisation de la compétition cible/distracteur**, les distracteurs de la position 'moins 1' supportant uniquement un mécanisme d'interruption et les distracteurs de la position 'plus 1' supportant uniquement un mécanisme d'intégration. Cette

spécialisation peut refléter une **latéralisation** telle que celle décrite à la figure 103. Ces deux formes de compétition sont sans doute sensibles aux mêmes manipulations expérimentales, ce qui explique leur co-variation dans l'expérience 1 (variation des paramètres temporels). Mais la nature des confusions qu'elles forment sont différentes, ce qui explique l'absence d'un regroupement des deux variables proba_M1 et proba_P1 dans un facteur commun au cours de l'analyse en composantes principales.

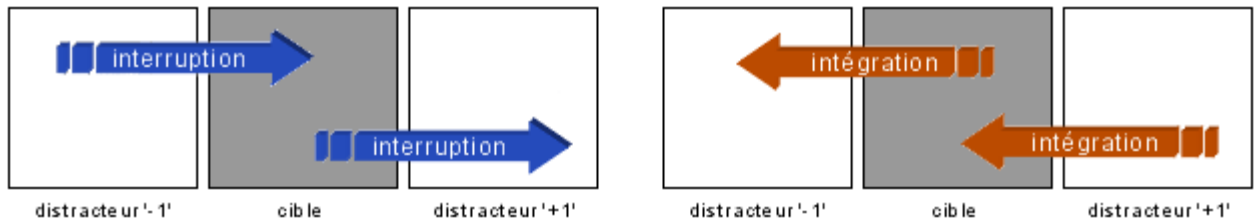


Figure 103 : latéralisation de la compétition cible/distracteur : si l'interruption fonctionne de manière proactive, la compétition distracteur '-1'/cible se traduira par une intrusion, alors que la compétition cible intruseur '+1' se traduira par une bonne réponse, invisible dans nos analyses. De même, si le mécanisme d'intégration fonctionne de manière rétroactive, alors la compétition distracteur '-1'/cible se traduira par une bonne réponse et la compétition cible/distracteur '+1' par une intrusion. On n'observera le mécanisme d'interruption qu'à travers les intrusions '-1' et le mécanisme d'intégration qu'à travers les intrusions '+1'.

Cette structure factorielle confirme l'existence de plusieurs mécanismes sous-tendant les intrusions, comme l'avaient déjà suggéré Botella & Ericksen (1992), Maki, Couture, Frigen & Lien (1997). La principale source de mixité semble supportée par le facteur de pari sophistiqué. La compétition cible/distracteur se traduit par des intrusions aux positions 'plus 1' et 'moins 1'. En l'absence d'une identification des intégrations 'moins 1' sur la base des propriétés des intégrations 'plus 1', il semble que nous devions renoncer à employer cette classe d'intrusion pour étudier le *binding*. Il reste qu'une partie des intrusions 'plus 1' concorde avec l'hypothèse d'une intégration cible/distracteur. C'est en détaillant ces intrusions que nous allons sélectionner des configurations perceptives.

Le travail de sélection que nous devons entreprendre consiste à supprimer les configurations perceptives relevant peut-être d'une intégration cible/intruseur ou peut-être d'un pari sophistiqué, pour ne retenir que les intrusions 'plus 1' fréquentes relevant sans ambiguïté d'une intégration.

Séparation probabiliste des intrusions et des conjonctions illusoires

La compétition entre la cible et le distracteur se traduit par deux manifestations : des interruptions du traitement de la cible par le distracteur en 'moins 1' et des intégrations du traitement du distracteur au traitement de la cible en 'plus 1'. Dans le continuum de compétition cible/distracteur, la représentation cognitive conjointe de la cible et du distracteur est faible pour les interruptions, mais plus forte pour les intégrations. Ceci nous oriente vers les intrusions '+1', comme étant le lieu où chercher des conjonctions illusoires. Cependant, **la position des intrusions ne suffit pas à séparer les conjonctions illusoires des intrusions** car le facteur de pari sophistiqué (estimé par Botella & coll, 2001, au tiers du total des intrusions) est également représenté à la position 'plus 1' d'une manière non négligeable. En

conséquence de la rotation orthogonale appliquée à l'analyse en composantes principales, les facteurs 'spécifique plus 1' et 'pari sophistiqué' sont indépendants. Certains couples cible/intruseur ont des scores élevés sur les deux facteurs ; dans ce cas, il n'est pas possible de déterminer si l'intrusion relève d'un pari sophistiqué ou d'une conjonction illusoire. D'autres couples présentent un score élevé sur le facteur 'spécifique plus 1', en même temps qu'un score faible au facteur de 'pari sophistiqué'. Pour ces couples-là, la probabilité d'enregistrer une intrusion au lieu d'une conjonction illusoire est modérée. Ce sont ces couples cible/intruseur que nous retenons pour les études ultérieures. Les couples sélectionnés à priori possédaient un score supérieur à la médiane au facteur 'spécifique plus 1' et un score inférieur à la médiane au facteur 'pari sophistiqué'. Ils sont représentés en bleu sur le nuage de point de la figure 104. La probabilité moyenne d'observer une intrusion 'plus 1' parmi les couples présélectionnés est de 3,5%¹, la probabilité la plus élevée étant de 9,4%. Les plus efficaces de ces couples cibles/distracteurs sont reportés à la figure 105.

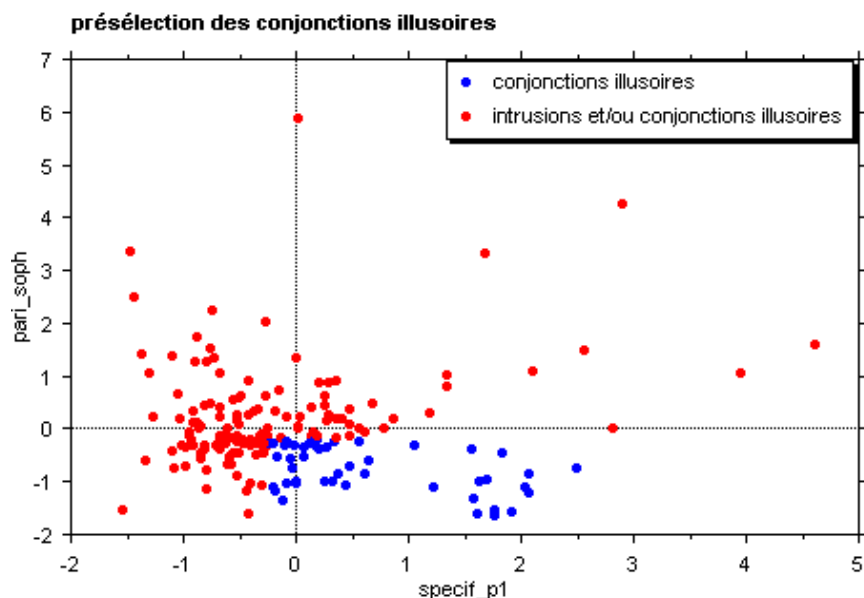


figure 104 : présélection des conjonctions illusoires non-ambiguës. (abréviations : pari_soph : facteur de pari sophistiqué ; specif_P1 : facteur spécifique des intrusions plus 1)

¹ Sur la totalité des couples, la probabilité d'intrusions 'plus 1' s'échelonne de 0% à 27,8% des essais. Mais les essais les plus propices à l'observation des intrusions 'plus 1' sont également d'interprétation ambiguë : si deux mécanismes se superposent pour la formation des erreurs, la fréquence est mécaniquement plus élevée. Aussi, la probabilité d'observer une conjonction illusoire dans les couples présélectionnés est-elle plus modeste.



Figure 105 : couples cibles/distracteurs ayant une forte probabilité de conjonctions illusoires. Le chiffre reporté à coté du couple traduit la probabilité d'une intrusion plus 1. Les index gris sont des échelles de similarité (de très similaire en noir à peu similaire en blanc) pour la couleur (curseur rouge) et la luminosité (curseur vert) dans le modèle Lab.

Discussion

Les analyses préliminaires (pp.78-82) suggéraient un effet systématique de la similarité critérielle globale (eg. forme, couleur, texture, luminosité) sur les conjonctions illusoires. Ayant permis de dissocier les intrusions et les conjonctions illusoires, la présente analyse modère cette hypothèse. Des deux similarités globales quantifiées (couleur, liée à la voie parvocellulaire et luminosité, liée à la voie magnocellulaire), aucune ne montre de lien systématique avec la survenue des conjonctions illusoires (voir figure 106).

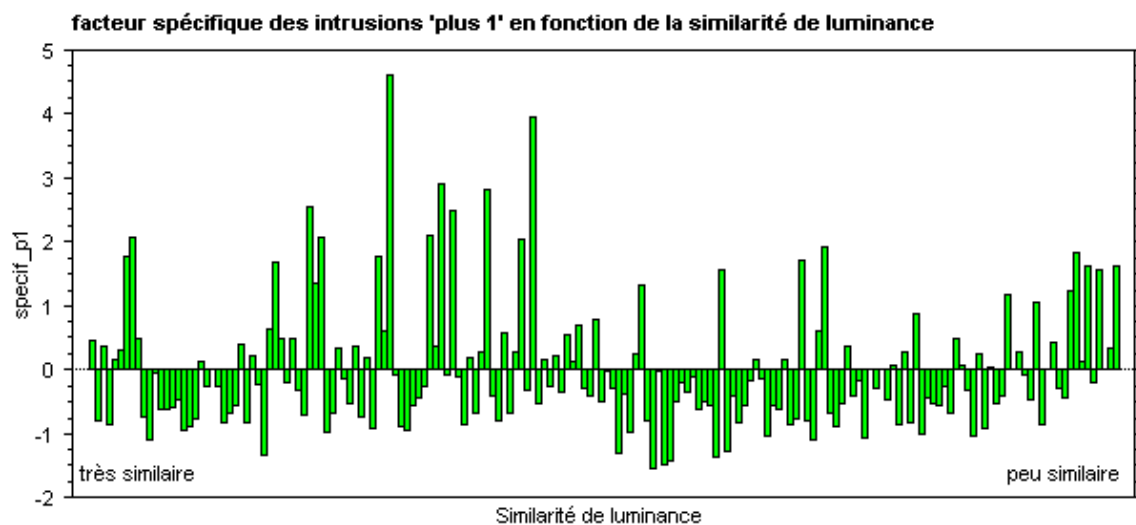
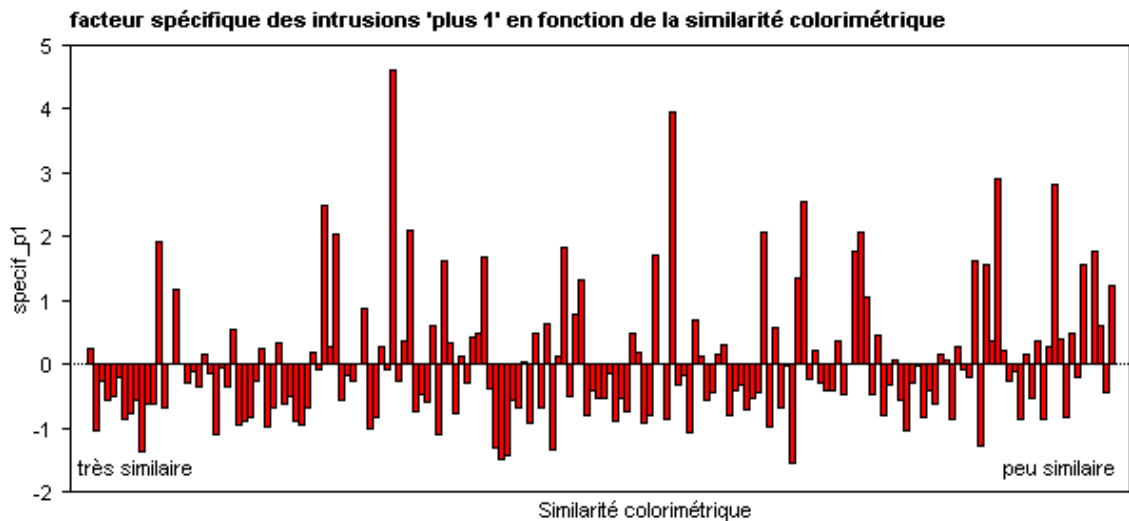


Figure 106 : le facteur spécifique des intrusions 'plus 1' ne montre pas de relation systématique avec la similarité colorimétrique (modèle Lab, composantes a et b) ou avec la similarité de luminance (modèle Lab, composante L).

Certains des couples sélectionnés pour leur lien non-ambigu avec les conjonctions illusoires sont **similaires pour la couleur**, d'autres pour la **luminosité**, et d'autres pour les deux paramètres, mais aucune de ces primitives visuelles n'explique univoquement les résultats. On observe que des objets très dissimilaires pour la couleur (eg. tee-shirt et chaussettes), pour la luminosité (jupe et pull) ou pour les deux paramètres (salopette et tee-shirt) peuvent néanmoins donner lieu à des conjonctions illusoires. De même, la figure 107 présente des objets très similaires sur les deux dimensions de couleurs et de luminosité, mais ne donnant pas lieu à des conjonctions illusoires.



Figure 107 : exemples de stimuli proches sur les dimensions de couleur et de luminance, mais ne donnant pas lieu à des conjonctions illusoires en 'plus 1'.

En fait, il semble que l'**effet de la similarité globale (couleur, luminosité) sur les intrusions suggérée par les analyses préliminaires était surtout supporté par la voie de pari sophistiquée**, permettant la récupération de l'identité du stimulus par un traitement partiel (eg. la cible sur le carré gris était rouge) puis un appariement en mémoire (eg. le seul animal rouge était un perroquet). Puisque ces intrusions ne manifestent pas une intégration cible/distracteur, ils sortent donc du domaine d'investigation de la présente thèse.

Même si ces analyses laissent ouverte la possibilité d'une action de la similarité locale ou partielle de la luminosité ou de la couleur, elles réfutent l'hypothèse d'une action de la similarité globale. Notons que les hypothèses de la prééminence du global sur le local suggèrent que c'est à la similarité globale

Par ailleurs, il reste que la similarité de forme, difficile à apprécier par des méthodes quantitatives, peut intervenir dans l'intégration cible/intruseur, comme le suggèrent de manière exemplaire les trois premiers couples de la figure 105 (bonnet/jupe : veste/tee-shirt et robe/salopette). La similarité de forme peut faciliter la manifestation des conjonctions illusives de deux manières :

- plus deux formes sont similaires, plus le **recouvrement spatial** des deux images est grand, et plus la **compétition neuronale** décrite par Keysers et Perrett (2002), liée à la persistance neuronale peut s'installer. Ainsi, les objets présentés à la figure 107 sont peut-être trop peu superposés pour que cette compétition neuronale intervienne dans la décision perceptive. On observe cependant des contre-exemples (eg. robe/chaussettes) dans les couples présentés à la figure 105. Ces couples présentant peu de recouvrement spatial ont néanmoins une forte probabilité de conjonction illusoire.
- mais également, la similarité de forme peut être le critère de regroupement de la cible et du distracteur, dans une perspective itérative du fonctionnement visuel.

Cette hypothèse d'un effet de la similarité de forme reste ouverte pour de futures investigations.

Conclusion

Les erreurs d'identification en RSVP sont appelées conjonctions illusoires sans que cette appellation ne prête à discussion. La solution basse consiste à dire que la conjonction concerne le liage incorrect du critère de réponse associé à la cible et de l'objet-intruseur. La solution haute consiste à dire que les caractéristiques de la cible et du distracteur sont mêlées dans une représentation illusoire. L'expérience 3 a montré que **les conjonctions illusoires sont des erreurs plus complexes qu'une simple dissociation visuelle du critère de réponse**, s'associant indûment avec le distracteur. Deux arguments supportent cette hypothèse. En premier lieu, que le critère de réponse participe à la construction cognitive formée par le sujet est rendu peu probable par les résultats incidents de cette expérience. Nous remarquons que le critère de réponse n'est qu'un simple marqueur visuel (un déclencheur du traitement). Il ne fait pas pleinement partie de la construction cognitive qu'élabore le sujet. Dans le cas contraire, il devrait se manifester des différences entre les deux conditions expérimentales de cette expérience 3. En effet, le critère de réponse extrinsèque n'est pas recouvert par le flux d'images. Il n'est donc pas susceptible de masquage. En revanche, le critère de réponse intrinsèque est susceptible de subir un masquage, car il est spatialement recouvert par les images du flux RSVP. Cependant, on n'observe aucune différence entre les post-intrusions des deux conditions. Plus directement, la présentation du critère de réponse intrinsèque de la cible (cible en noir et blanc parmi des distracteurs colorés) n'empêche pas la formation des conjonctions illusoires, démontrant que le binding concerne des caractéristiques de la cible et des caractéristiques du distracteur, plutôt que le critère de réponse attaché à la cible et un distracteur. Ce résultat supporte l'hypothèse de **conjonctions illusoires complexes (chimères cognitives)**. Les résultats de l'expérience 3 réfutent l'hypothèse d'un déficit de la consolidation et orientent l'interprétation du dysfonctionnement vers un pur problème perceptif.

L'expérience 1 renforce cette hypothèse d'une chimère cognitive en mettant en ligne directe les conjonctions illusoires et **l'itération perceptive**. Les conditions de perception RSVP ne permettent pas l'accomplissement total du traitement visuel, comme le démontre la voie de pari sophistiqué, représentant la récupération de l'identité de la cible sur la base d'information partielle. La finalité de l'itération perceptive est de compléter l'information sur la cible. L'asynchronie entre la cible et le distracteur se trouve en dessous de la latence nécessaire pour compléter l'identification et supporte l'hypothèse que le système visuel n'a pas encore abouti à une décision perceptive lorsque le distracteur envahit les structures visuelles corticales (voir également, à titre indicatif, les latences d'activations des structures visuelles chez le singe, reportées figures 44 et 45 pp.71-72). Le système visuel se trouve alors **cerné par**

deux contraintes. Il faut **acquérir plus d'informations** sur la cible pour permettre son identification. Mais par ailleurs, il faut **segmenter** le flux visuel, c'est-à-dire cesser d'acquérir de l'information lorsque la stimulation a changé. Puisque le temps d'exposition du stimulus est la seule variable temporelle explicative des conjonctions illusoires, l'hypothèse d'une défaillance de ce processus de segmentation est rendue plausible. L'amplitude du phénomène est importante. Les expériences 2, 4 et 5 démontrent que les intrusions 'plus 1' sont fortement saturées par le mécanisme de conjonction illusoire. Or, l'expérience 1 montre qu'une diminution de 20% du temps d'exposition des stimuli se traduit par une augmentation de plus de 100% des intrusions 'plus 1', celles là mêmes qui semblent le plus fortement saturées par le mécanisme de conjonctions illusoires.

Les règles de cette inclusion du distracteur dans le traitement de la cible ont échouées à être circonscrites statistiquement (expérience 5). La similarité cible-distracteur était fortement suspectée, et avait été intégrée aux hypothèses explicatives (eg. figure 69 p.106). **La couleur et la luminosité globales, apparues comme de puissant vecteurs d'erreurs dans les analyses préliminaires, se sont avérées être des leurres supportés par la voie de pari sophistiqué.** Ces erreurs formées offline sur la base d'un traitement perceptif partiel, sortent de la définition d'une conjonction illusoire. Une similarité globale de forme reste une interprétation possible pour les conjonctions illusoires, mais la quantification de ce paramètre est difficile. La théorie des géons de Biederman (1987) offre un support pour de futures investigations. Il faut peut-être considérer que le problème, orienté par la suggestion d'un effet de la similarité de couleur et de luminosité, a été posé à l'envers. Nous avons supposé que le système visuel permet l'itération lorsqu'il constate un appariement possible entre le contenu antérieur et le contenu actuel. A contrario, peut-être qu'il refuse l'itération lorsqu'il constate une discordance entre le contenu antérieur et le contenu actuel. Ainsi, des caractéristiques résolues tardivement ne participent pas à la segmentation visuelle en RSVP, car elles sont trop rarement résolues lorsque le distracteur accède aux aires corticales.

Nous continuons à soutenir l'existence de critères d'association cible/intruseurs, car les fréquences d'intrusions dans les matrices d'associations ne peuvent être correctement prédites sur la base de la puissance du distracteur et de la vulnérabilité de la cible (eg. table 2 p.80). Ces estimations laissent un résidu important. Le lien sélectif du distracteur avec la cible reste une des observations clef pour les conjonctions illusoires, mais des expériences complémentaires sont nécessaires pour tester cette hypothèse.

La thèse supportée dans le présent document est que les conjonctions illusoires en RSVP se manifestent par un liage transitoire de l'activité des neurones des structures visuelles. Nous avons décrit de nombreuses études ayant mis en évidence le lien entre la synchronisation neuronale et la structuration perceptive (gestalt, regroupement perceptif). Ces études suggèrent que la représentation cognitive d'un objet est cimentée par la coordination temporelle des activités neuronales à petite et à grande échelle. Les **conjonctions illusoires**, quoiqu'étant des erreurs, **traduisent un effort d'intégration perceptive** produit par le système visuel sur le matériel présent dans l'environnement. En cela, elles se rapprochent davantage d'heuristiques visuelles que de la stricte définition d'une erreur (eg. traitement

inachevé). Cet effort d'intégration perceptive est une des formes de structuration perceptive corrélant avec la synchronisation neuronale. Ayant conforté le lien entre les conjonctions illusoires en RSVP et l'intégration perceptive, nous pouvons donc conclure que les conjonctions illusoires en RSVP constituent un paradigme approprié pour l'étude du *binding* par synchronisation. Cependant, l'échec des modélisations statistiques du phénomène empêche de formuler des prédictions spécifiques sur le type de synchronisation impliqué (eg. synchronisation à petite échelle, entre les neurones de IT ou synchronisation étendue entre des aires de différents niveaux hiérarchique, voir pp.72-73). Les latences d'activités neuronales laissent la porte ouverte aux deux hypothèses.

A notre connaissance, les conjonctions illusoires, intensivement étudiées par des paradigmes de recherche visuelle spatiale, n'ont pas fait l'objet de recherches électrophysiologiques. Elles s'y prêtent probablement mal, car le liage concerne des petits objets temporellement coïncidents, mais spatialement distincts. On peut s'attendre que la conjonction illusoire se manifeste brièvement au sein de populations localisées de neurones et d'une manière très semblable à un liage correct. Comme nous l'avons vu au chapitre 1, pp.26-52, ce type de manifestation neuronale n'est pas visible par les méthodes non-invasives d'observation de la coordination neuronale.

En revanche, le paradigme RSVP se prête probablement mieux à cette investigation, car les objets sont spatialement coïncidents, mais temporellement distincts. En outre, les objets sont de taille angulaire suffisante (5°), et activent une masse suffisante de neurones pour être **visible à travers les méthodes non-invasives**. Les résultats de l'expérience 1 suggèrent que lorsque conjonction illusoire il y a, le **liage neuronal est prolongé**. Il s'étend de la plage temporelle de la cible à celle du distracteur car le système visuel semble exploiter le flux d'information ascendant pour poursuivre le traitement lorsque la première série d'itérations n'a pas permis la décision perceptive (expérience 1). C'est cet événement qui sera observé en comparant les réponses correctes et les conjonctions illusoires.

Les modèles actuels du fonctionnement visuel (Lamme, Supèr & Spekreijse, 1998 ; Lamme & Roelfsema, 2000 ; Bullier, 2001^{a, b} ; Pacual-Leone & Walsh ; 2001) donnent une importance particulière aux flux rétroactifs vers le cortex visuel primaire. A ce titre, la coordination de l'aire visuelle striée et des aires temporales et pariétales seront examinées par une analyse de cohérence. En effet, lorsque le flux ascendant du distracteur arrive dans les aires visuelles striées, la cible a déjà envahi intensivement les structures visuelles extrastriées, et **la chimère cognitive s'accompagne peut-être alors d'une synchronisation des aires dominées par l'information-cible et des aires dominées par l'information-distractrice**. Sur la base de méthodes électrophysiologiques, il nous est possible de savoir si les manifestations de la synchronisation (gamma et cohérence) concernent la cible isolément ou le couple cible/intruseur, en rapportant les observations aux estimations du moment où l'information-distractrice envahit le cortex strié et extrastrié (entre 30 et 45 millisecondes après l'apparition du distracteur pour V1, en interpolant sur la base des latences observées chez le singe, voir pp.71-72).

Ayant dégagé, d'une part la pertinence du phénomène de conjonctions illusoires pour les recherches sur la synchronisation neuronale, et d'autre part, ayant circonscrit les erreurs en RSVP qui relèvent des conjonctions illusoires parmi la totalité des erreurs enregistrées, nous avons maintenant pour projet de recherche de réaliser ces investigations électrophysiologiques. La faible fréquence d'observation des conjonctions illusoires (entre 5 et 10% dans les configurations perceptives pré-sélectionnées) sera compensée en minimisant la durée d'exposition des stimuli. L'expérience 1 ayant montré que cette manipulation affecte quantitativement, mais non qualitativement le phénomène. Ces erreurs échappant en partie à la conscience du sujet (Intraub, 1989), et étant peu sensibles à l'apprentissage (expérience 2), le paradigme peut facilement être adapté pour une investigation électrophysiologique.

Bibliographie

- Bair W., (1999). Spike timing in the mammalian visual system. *Current opinion in Neurobiology*, 9 : 447-453.
- Bar-Gad I., Ritov Y., Vaadia E., Bergman H., (2001). Failure in identification of overlapping spikes from multiple neuron activity causes artificial correlations. *Journal of Neuroscience Methods*. 107 : 1-13.
- Barlow HB., (1972). Single units and sensation : a neuron doctrine for perpetual psychology ? *Perception*, 1 : 371-394.
- Basar-Eroglu C., Struber D., Kruse P., Basar E., Stadler M., (1996). Frontal gamma-band enhancement during multistable visual perception. *International Journal of Psychophysiology*, 24 : 113-125.
- Ben-Shaul Y., Bergman H., Ritov Y., Abeles M., (2001). Trial to trial variability in either stimulus or action causes apparent correlation and synchrony in neuronal activity. *Journal of Neuroscience Methods*, 111 : 99-110.
- Berger H., (1929). Ueber das Elektroenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87 : 527-570.
- Bertrand O., Tallon-Baudry C., (2000). Oscillatory gamma activity in humans : a possible role for objects representation. *International Journal of Psychophysiology*, 38 : 211-223.
- Biederman I., (1987). Recognition-by-components : a theory of human image understanding. *Psychological Review*, 94 : 115-147.
- Boisacq-Schepens N., Crommelinck M., (1996). *Neuro-psycho-physiologie. 2. Comportement*. Paris, Masson.
- Botella J., Barriopedro M., Suero M., (2001). A model of the formation of illusory conjunctions in the time domain. *Journal of Experimental Psychology : Human Perception and Performance*, 27 (6) : 1452-1467.
- Botella J., Ericksen CW., (1991). Pattern changes in rapid serial visual presentation tasks without strategic shifts. *Bulletin of the Psychonomic Society*, 29 (2) : 105-108.
- Botella J., Ericksen CW., (1992). Filtering versus parallel processing in RSVP tasks. *Perception & Psychophysics*, 51 (4), 334-343.
- Botella J., Garcia ML., Barriopedro M., (1992). Intrusion patterns in rapid serial visual presentation tasks with two response dimensions. *Perception and Psychophysics*, 52 (5) : 547-552.
- Botella J., Villar MV., (1986). Identificación de palabras en presentación rápida de series visuales : una replica del efecto Lawrence. *Conocimiento y acción*, 1 : 7-21.
- Boucart M. (1996). *La reconnaissance des objets*. Grenoble, Presses Universitaires de Grenoble.
- Boucart, M., de Visme, P., Wagemans, J. (2000) Effect of benzodiazepine on temporal integration in object perception. *Psychopharmacology*, 3, 249-255.
- Boucart, M., Moroni, C. Fuentes, L., Belin, C., (1998). Variation of Attentional Blink as a function of task and semantic information processing. *Abstracts from the Psychonomic Society*, 3 : 685.
- Boussaoud D., (1998). Un immense chantier neuronal. *La Recherche*, 309 : 58-61
- Breitmeyer BG., (1984). *Visual masking : an integrative approach*. Oxford University Press.

Bressler SL., (1995). Large-scale cortical networks and cognition. *Brain Research Reviews*, 20 : 288-304.

Bressler SL., (1996). Interareal synchronization in the visual cortex. *Behavioral and Brain Research*, 76 : 37-49.

Bressler SL., (2002). Understanding cognition through large-scale cortical networks. *Current directions in psychological Sciences*, 11 (2) : 58-61.

Bressler SL., Coppola, R., Nakamura R., (1993). Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature*, 366 : 153-156.

Bressler SL., Kelso JAS, (2001). Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5 (1) : 26-36.

Broadbent DE., (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, 54, 205-215.

Broadbent DE., (1982), Task combination and selective intake of information. *Acta Psychologica*, 50 (3) : 253-290.

Broadbent DE., Broadbent MH., (1987). From detection to identification: response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics*, 42 (2) : 105-113.

Broadbent, D.E. (1977). The hidden preattentive process. *American Psychologist*, 32, 109-118.

Brody CD., (1999). Correlations without synchrony. *Neural Computation*, 11 : 1537-1551.

Bullier J. Nowak LG., (1995). Parallel versus serial processing : new vistas on the distributed organization of the visual system. *Current Opinion in Neurobiology*, 5 : 497-503.

Bullier J., (1998). Architecture fonctionnelle du système visuel, in M. Boucart, MA. Hennaff, C. Belin, (Eds) *La vision : aspects perceptifs et cognitifs*. Solal : collection neuropsychologie.

Bullier J., (2001a). Integrated model of visual processing. *Brain Research Reviews*, 36 : 96-107.

Bullier J., (2001b). Feedback connections and conscious vision. *Trends in Cognitive Sciences*, 5 (9) : 369-370.

Bullock TH., (1997). Signals and signs in the nervous system : the dynamical anatomy of electrical activity is probably information-rich. *Proceedings of the National Academy of Sciences (USA)*, 94 : 1-6.

Bullock TH., Achimovicz JZ., (1993). A comparative survey of oscillatory brain activity, especially gamma-band rhythms. *Symposium on oscillatory event related dynamics*, Tecklenburg, "Münsterland", 2-6 Sept.

Bülthoff HH., Edelman S., (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Science (USA)*, 89 : 60-64.

Burke-Hubbard B., (1995). *Ondes et ondelettes*. Paris, Bibliothèque Pour la Science.

Cariani PA., (2001). Neural timing nets. *Neural Networks*, 14(6) : 737-53.

Castelo-Branco M., Goebel R., Neuenschwander S., (2000). Neural synchrony correlates with surface segregation rules. *Nature*, 405 : 685-689.

Cherry EC., (1953). Some experiments on the recognition of speech, with one and with the two ears. *Journal of the Acoustical Society of America*, 25 : 975-979.

Choe Y., Miikkulainen R., (1998). Self-organization and segmentation in a laterally connected orientation map of spiking neurons. *Neurocomputing*, 21 : 139-157.

Chun MM., (1997) Temporal binding errors are redistributed by the attentional blink. *Perception and Psychophysics*, 59 (8) : 1191-1199.

Coello Y., Richaud S., Magne P., Rosetti Y., (in press). Vision for spatial perception and vision for action: A dissociation between azimuth and depth dimension. *Neuropsychologia*.

- Colheart M., (1980). Iconic memory and visible persistence. *Perception and Psychophysics*, 27 : 183-228.
- Content A., Mousty P., Radeau M., (1990). Brulex : une base de données lexicales informatisées pour le français écrit et parlé. *L'année Psychologique*, 90 : 551-566.
- Corballis MC., (1995). Visual integration in the split brain. *Neuropsychologia*, 33 (8) : 937-959.
- Decharms C., Zador A., (2000). Neural representation and the cortical code. *Annual Review of Neurosciences*, 23 (1) : 613-647.
- Desimone R., (1991). Face-selective cell in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, 3 : 1-8.
- Desimone R., (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transaction of the Royal Society of London, Serie B*, 353 : 1245-1255.
- Desimone R., Schein SJ., Moran J., Ungerleider LG., (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Research*, 25 (3) : 441-452.
- DeValois RL., DeValois KK., (1993). A multi-stage color model. *Vision Research*, 33 (8) : 1053-1065.
- DeValois RL., DeValois KK., (1988). *Spatial vision*. New-York, Oxford University Press.
- Driver J., Spence C., (2000). Multisensory perception: beyond modularity and convergence. *Current Biology*, 10 (20) : R731-R735.
- Eckhorn R., (1994). Oscillatory and non-oscillatory synchronizations in the visual cortex and their possible roles in associations of visual features. *Progress in Brain Research*, 102 : 405-426.
- Eckhorn R., Bauer R., Jordan W., Borsch M., Kruse W., Munk M., , Reitboeck HJ., (1988). Coherent oscillations : a mechanism of feature linking in the visual cortex. Multiple electrode correlation analyses in the cat. *Biological Cybernetics*, 60 (2) : 121-130.
- Eckhorn R., Obermueller A., (1993). Single neurons are differently involved in stimulus-specific oscillations in cat visual cortex. *Experimental Brain Research*, 95 (1) : 177-182.
- Edelman S., (2002). Constraining the neural representation of the visual world. *Trends in Cognitive Sciences* 6 : 125-131.
- Eggermont JJ., (1998). Is there a neural code ? *Neuroscience and Behavioral Reviews*, 22 (2) : 355-370.
- Elliot MA., Herrmann CS., Mecklinger A., Müller HJ., (2000). The loci of oscillatory visual-object priming : a combined electroencephalographic study and reaction-time study. *International Journal of Psychophysiology*, 38 : 225-241.
- Elliot MA., Müller HJ., (1998). Synchronous information presented in 40-Hz flicker enhances visual feature binding. *Psychological Science*, 9 : 277-283.
- Engel AK., Fries P., Singer W., (2001). Dynamic predictions : oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2 (10) : 704-716.
- Engel AK., König P., Singer W., (1991). Direct physiological evidence for scene segmentation by temporal coding. *Proceedings of the National Academy of Science (USA)*, 88 : 9136-9140.
- Enns JT., DiLollo V., (2001). What's new in visual masking. *Trends in Cognitive Sciences*, 4 (9) : 345-352.
- Fabre-Thorpe M., Delorme A., Marlot C., Thorpe, S., (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13 (2) : 1-10.
- Farah MJ., (1994). Neuropsychological inference with an interactive brain : A critique of the locality assumption. *Behavioral and Brain Sciences*, 17 : 43-61.

- Farid H., Adelson EH., (2001). Synchrony does not promote grouping in temporally structured displays. *Nature Neuroscience*, 4 (9) : 875-876.
- Felleman DJ., Van Essen DC., (1991). Distributed hierarchical processing in primate visual cortex. *Cerebral Cortex*, 1 : 1-47.
- Fernandez-Duque D., Johnson ML., (1999) Attention Metaphors : How Metaphors Guide the Cognitive Psychology of Attention. *Cognitive Science*, 23 (2) : 83-116.
- Finger S., (2000). *Minds behind the brain : a history of the pioneers and their discoveries*. New-York, Oxford University Press.
- Finkel LH., Yen SC., Menschik E., (1998). Synchronization : the computational currency of cognition. In *Proceedings of the 8th International Conference on Artificial Neural Networks Skovde, Sweden, 2-4 Sept.*
- Fodor J., (1983). *The modularity of mind*. Cambridge, MIT Press
- Friedman-Hill SR., Robertson LC., Treisman A., (1995). Parietal contributions to visual feature binding. Evidence from a patient with bilateral lesions. *Science*, 269 : 853-855.
- Fries P., Reynolds JH., Rorie AE., Desimone R., (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291 : 1560-1563.
- Fujita I., Tanaka K., Ito M., Cheng K., (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, 360 : 343-346.
- Galambos R., (1992). A comparison of certain gamma-band (40 Hz) brain rhythms in cats and man. In Basar E., Bullock T., (eds.). *Induced rhythms in the brain*. Brelin, Springer, 103-122.
- Gallant JL., Braun J., Van Essen DC., (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science*, 259 (5091) : 100-103.
- Gathercole SE., Broadbent DE., (1984). Combining attributes in specified and categorized target search : Further evidence for strategic differences. *Memory & Cognition*, 12, 329-337.
- Gauthrais J., Thorpe S., (1998). Rate coding versus temporal order coding : a theoretical approach. *Biosystems*, 48 : 57-65.
- Gegenfurtner KL., Rieger J., (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, 10 : 805-808.
- Gerstein GL., (2000). Cross-correlation measures of unresolved multi-neuron recordings. *Journal of Neurosciences Methods*, 100 : 41-51
- Gerstein GL., Kirkland KL., (2001). Neural assemblies : technical issues, analysis and modeling. *Neural Networks*, 14 : 589-598.
- Gerstein GL., Perkel DH., (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science*, 164 (881) : 828-830.
- Gestner W., Kreiter AK., Markram H., Herz AVM., (1997). Neural codes : firing rates and beyond. *Proceedings of the National Academy of Science (USA)*, 94 : 12740-12741.
- Gray CM., König P., Engel AK., Singer W., (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchroniation which reflects global stimulus properties. *Nature*, 338 (6213) : 334-337.
- Gray CM., Singer W., (1987). Stimulus-specific neuronal oscillations in the cat visual cortex : a cortical functional unit. *Abstract from the Neuroscience Society*, 13 : 404.3.

Gray CM., Singer W., (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Science (USA)*, 86 (5) : 1698-1702.

Gray CM., Viana Di Prisco G., (1997). Stimulus-dependent neuronal oscillations and local synchronization in striate cortex of the alert cat. *The Journal of Neurosciences*, 17 (9) : 3239-3253.

Gross CG., Bender DB., Rocha-Miranda CE., (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science*, 166 : 1303-1306.

Grossberg S., Grunewald A., (1997). Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience*, 9, 117-132.

Gruber T., Müller MM., (2002). Effects of picture repetition on induced gamma band responses, evoked potentials, and phase synchrony in the human EEG. *Cognitive Brain Research*, 13 : 377-392.

Hardcastle V., (1999). What we don't know about brain. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 30 (1) : 69-89.

Hartline HK., (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *American Journal of Physiology*, 121 : 400-415.

Hebb D., (1949). *The organization of behavior : a neuropsychological theory*. New-York, Wiley.

Heculano-Houzel S., Munk MHJ., Neueschwander S., Singer W., (1999). Precisely synchronized oscillatory firing patterns require electroencephalographic activation. *The Journal of Neurosciences*, 19 (10) : 3992-4010.

Herrmann CS., (2001). Human EEG responses to 1-100 Hz flicker : resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, 137 : 346-353.

Herrmann CS., Knight RT., (2001). Mechanisms of human attention : event-related potentials and oscillations. *Neuroscience and Behavioral Reviews*, 25 : 465-476.

Herrmann CS., Mecklinger A., Pfeifer E., (1999). Gamma responses and ERPs in a visual classification task. *Clinical Neurophysiology*, 110 : 636-642.

Hopfield JJ., Herz AV., (1995). Rapid local synchronization of action potentials : toward computation with coupled integrate-and-fire neurons. *Proceedings of the National Academy of Science (USA)*, 92 : 6655-6662.

Hubel D., Wiesel T., (1979). Les mécanismes cérébraux de la vision. *Pour la Science*, 25 : 79-93.

Hubel DH., Wiesel TN., (1959). Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, 148 : 574-591.

Humphreys GW., Ciel C., Wolfe J., Olson A., Klempen N., (2000). Fractionating the binding process : neuropsychological evidence distinguishing binding of form from binding of surface features. *Vision research*, 40 : 1569-1596.

Humphreys, GW., Bruce, V. (1989). *Visual cognition. Computational, experimental and neuropsychological perspectives*. Mahwah, NJ : Lawrence Erlbaum.

Hutcheon B., Yarom Y., (2000). Resonance, oscillation and the intrinsic frequency preference of neurons. *Trends in Neurosciences*, 23 : 216-222.

Intraub H., (1981). Rapid Conceptual Identification of Sequentially Presented Pictures. *Journal of Experimental Psychology : Human Perception and Performance*, 7 : 604-610.

Intraub H., (1985). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology : Human Perception and Performance*, 11 : 431-442.

Intraub H., (1989). Illusory conjunctions of forms, objects and scenes during RSVP search. *Journal of experimental Psychology : Human Perception and Performance*, 11 : 431-442.

- Ishai A., Ungerleider LG., Martin A., Schouten JL., Haxby JV., (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Science (USA)*, 96 (16) : 9379-9384.
- Ivry RB., Prinzmetal W., (1991). Effect of feature similarity on illusory conjunctions. *Perception and Psychophysics*, 49 : 105-118.
- Jones EG., (2001). The thalamic matrix and thalamocortical synchrony. *Trends in Neurosciences*, 24 (10) : 595-610.
- Jürgens E., Rosler F., Henninghausen E., Heil M., (1995). Stimulus-induced gamma oscillations : harmonics of alpha activity ? *Neuroreport*, 6 (5) : 813-816.
- Kahneman D., (1973). *Attention and effort*. Londres, Prentice Hall.
- Kaneoke Y, Vitek JL., (1996). Burst and oscillation as disparate neurophysiologic properties : methods of detection. *Journal of Neuroscience Methods*, 68 : 211-223.
- Kanwisher NG., (1987). Repetition blindness : Type recognition without token individuation. *Cognition*, 27 : 117-143.
- Karbowski J., Kopell N., (2000). Multispikes and synchronization in a large neural network with temporal delays. *Neural Computation*, 12 : 1573-1606.
- Keil A., Gruber T., Müller MM., (2001). Functional correlates of macroscopic high-frequency brain activity in the human visual system. *Neuroscience and Behavioral Reviews*, 25 : 527-534.
- Keysers C., Perrett DI., (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences*, 6 (3) : 120-125.
- Keysers C., Xiao DK., Földiák P., Perrett DI., (2001). The speed of sight. *Journal of Cognitive Neuroscience*, 13 : 90-101.
- Kikushi T., (1996). Detection of Kanji words in a rapid serial visual presentation task. *Journal of Experimental Psychology : Human Perception and Performance*, 22 (2) : 332-41.
- König, Engel, Singer, (1995). Relation between oscillatory activity and long-range synchronization in cat visual cortex. *Proceedings of the National Academy of Science (USA)*, 92 : 290-294.
- Koriat A., Goldsmith M., Pansky A., (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, 51 : 483-539.
- Kuffler SK., (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16 : 37-68.
- LaBerge D., (2000). Networks of attention. In M. Gazzaniga (Ed.) *The new cognitive neuroscience*. Cambridge, MIT Press.
- Lachaux J.P., Rodriguez E., Martinerie J. Varela F. (1999) Measuring phase-synchrony in brain signals. *Human Brain Mapping*, 8, 194-208.
- Lamme VAF., Roelfsema PR., (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience*, 23 : 571-579.
- Lamme VAF., Supér H., Spekreijse H., (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current opinion in Neurobiology*, 8 : 529-535.
- IArbib M., Erdi P., Szentagothai J., (1997). *Structure, function and dynamics, an integrated approach to neural organization*. Cambridge, MIT press.
- Lawrence DH., (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics*, 10 : 85-89.

- LeMoal M., (1987). Entre le comportement et la molécule : les neurosciences intégratives peuvent-elles exister ? 8In Séron X., (dir.) *Psychologie et cerveau*. Paris PUF, 195-221.
- Lestienne R., (2001). Spike timing, Synchronization and information processing on the sensory side of the central nervous system. *Progress in Neurobiology*, 65 : 545-591.
- Loftus GR., Nelson WW., Kallman HJ., (1983). Differential acquisition rates for different types of information from pictures. *Quarterly Journal of Experimental Psychology A*, 35 (1) : 187-98.
- Logothetis N., (1998). Object Vision and visual awareness. *Current Opinion in Neurobiology*, 8 : 536-544.
- Maki WS., Couture T., Frigen K., Lien D., (1997). Sources of the attentional blink during rapid serial visual presentation : perceptual interference and retrieval competition. *Journal of Experimental Psychology : Human Perception and Performance*, 23 (5) : 1393-1411.
- Maki WS., Frigen K., Paulson K., (1997). Associative priming by targets and distractors during rapid serial visual presentation : does word meaning survive the attentional blink ? *Journal of Experimental Psychology : Human Perception and Performance*, 23 (4) : 1014-1034.
- Martin, KAC, (2000). The pope and grandmother – a frog's-eye view of theory. *Nature Neuroscience*, 3 (suppl.) : 1169.
- Maunsell JHR., Newsome WT., (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10 : 363-401.
- McLean JP., Broadbent DE., Broadbent MHP., (1982). Combining attributes in rapid serial visual presentation tasks. *Quarterly Journal of Experimental Psychology*, 35A : 171-186.
- Mesulam MM., (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann. Neurol.*, 28 (5) : 597-613.
- Mesulam MM., (1998). From sensation to cognition. *Brain*, 121 (6) : 1013-1052.
- Milner AD., Goodale MA., (1995). *The visual brain in action*. Oxford University Press.
- Molotchnikoff S., Shumikhina S., (2000). Relationships between image structure and gamma oscillations and synchronization in visual cortex of cats. *European Journal of Neurosciences*, 1 : 1440-1452.
- Moray N., (1969). *Attention selective processes in vision and hearing*. London, Hutchinson.
- Morin, E., (1995). Vers un nouveau paradigme. *Sciences Humaines n°spécial « penser la complexité »*, 047 : 20-23.
- Moroni C., Boucart M., Humphreys G.W., Henaff M.A., Belin C. (2000) Failure to identify the target does not prevent the attentional blink effect. *Neuro Report*, 11, 2775-2780.
- Müller MM., Bosch J., Elbert T., Kreiter A., Sosa M., Valdes-Sosa P., Rockstroh B., (1996). Visually induced gamma-band responses in human electroencephalographic activity : a link to animal studies. *Experimental Brain Research*, 112 (1) : 96-102.
- Müller MM., Gruber T., Keil A., (2000). Modulation of induced gamma band activity in the human EEG by attention and visual information processing. *International Journal of Psychophysiology*, 38 : 283-299.
- Munk MHJ., Roelfsema PR., König P., Engel AK., et Singer W., (1996). Role of reticular activation in the modulation of intracortical synchronization. *Science*, 272 : 271-274.
- Nakayama, K., Silverman, GH., (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, 320 (6059) : 264-265.
- Näsänen R., Ojanpää H., Kojo I., (2001). Effect of stimulus contrast on performance and eye movements in visual search. *Vision Research*, 41 : 1817-1824.
- Navon D., Ehrlich B., (1995). Illusory conjunctions : does inattention really matter ? *Cognitive Psychology*, 29 : 59-83.

- Neisser U., (1967). *Cognitive Psychology*. New-York, Appleton Century Crofts.
- Newman J., Reticular-thalamic activation of the cortex Synchronous Oscillations and the Emperor's New Clothers. *Journal of Consciousness studies online* : <http://www.imprint.co.uk/online/new1.html>
- Nowak LG., Bullier J., (1997). The timing of information transfer in the visual system. In KS. Rockland, JH. Kaas, A. Peters, (Eds). *Extrastriate visual cortex in primates*, Vol. 12. New-York, Plenum Press.
- Nunez PL., (2000). Toward a quantitative description of large scale neocortical dynamic function and EEG. *Behavioral and Brain Sciences*, 23 (3) : 371-398.
- O'Connors KJ., Potter MC., (2002). Constrained formation of object representations. *Psychological Science*, 13 (2) : 106-111.
- Pacual-Leone A., Walsh V., (2001). Fast Backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, 292 : 510-512.
- Panzeri, S., Schultz, S. R., Treves, A., and Rolls, E. T. (1999). Correlations and the encoding of information in the nervous system. *Proceedings of the Royal Society of London serie B : Biological Sciences*, 266 : 1001-1012.
- Pellizzer G., Georgopoulos AP., (1993). Common processing constraints for visuo-motor and visual mental rotation. *Experimental Brain Research*, 93 : 165-172.
- Perkel DH., Gerstein GL., Moore GP., (1967). Neuronal spike trains and stochastic point processes. II. Simultaneous point processes. *Biophysics Journal*, 7 : 391-418.
- Perrett DI., Oram MW., (1993). Neurophysiology of shape processing. *Image Vision and Computation*, 11 : 317-333.
- Pessoa L., Neumann H., (1998). Why does the brain fill in ? *Trends in Cognitive Sciences*, 2 (11) : 422-424.
- Peterson MS., (1994). Object recognition can and do operate before figure-ground organization. *Current Directions in Psychological Science*, 3 (4) : 105-109.
- Posner MI., (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32 : 3-25.
- Posner MI., Petersen E., (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13 : 25-42.
- Posner MI., Raichle ME., (1994). *Images of mind*. New-York, Scientific American Library.
- Potter MC., (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology : Human Learning and Memory*, 2 (5) : 509-522.
- Potter MC., (1993). Very short term conceptual store. *Memory & Cognition*, 21 : 156-161.
- Potter MC., (1999). Understanding sentences and scenes : the role of conceptual short-term memory. In V., Colheart., *Fleeting memories : cognition of brief visual stimuli*. Cambridge, MIT Press : 13-46.
- Potter MC., Levy El., (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, 81 (1) : 10-15.
- Prinzmetal W., (1995). Visual feature integration in a world of objects. *Current Direction in Psychological Science* 4 : 90-94.
- Prinzmetal W., Henderson D., Ivry R., (1995). Loosening the constraints on illusory conjunctions : the role of exposure duration and attention. *Journal of experimental Psychology : Human Perception and Performance*, 21 : 1362-1375.
- Raymond JE., Shapiro KL., Arnell KM., (1992). Temporary suppression of visual processing in an RSVP task : an attentional blink ? *Journal of Experimental Psychology : Human Perception and Performance*, 18 : 849-860.
- Riesenhuber M., Poggio T., (2000). Models of object recognition. *Nature Neuroscience*, 3 (suppl.) : 1199-1204.
- Ringo JL., (1991). Neuronal interconnection as a function of brain size. *Brain, Behavior and Evolution*, 38 (1) : 1-6.

- Rockland KS., (2002). Visual cortical organization at the single axon level: a beginning. *Neuroscience Research*, 42 (3) : 155-166.
- Rodriguez E., George N., Lachaux J.P., Martinerie J., Renault B. Varela F. (1999) Perception's shadow : long-distance synchronization in the human brain activity. *Nature*, 397, 430-433.
- Roelfsema PR., König P., Engel AK., Sireteanu R., Singer W., (1994). Reduced synchronization in the visual cortex of cats with strabismic amblyopia. *European Journal of Neurosciences*, 6 : 1645-1655.
- Roy SA., Dear SP., Alloway KD., (2001). Long-range cortical synchronization without concomitant oscillations in the somatosensory system of anesthetized cats. *The journal of Neuroscience*, 21 (5) : 1795-1808.
- Sanders AF., (1997). *Element of Human performance : reaction processes and attention in human skill*. Mahwah, NJ : Lawrence Erlbaum.
- Sanocki T., (1993). Time course of object identification : evidence for a global-to-local contingency. *Journal of Experimental Psychology : Human Perception and Performance*, 19 878-898.
- Schiller PH., (1993). The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Visual Neuroscience*, 10 : 717-746.
- Schmolesky MT., Wang Y., Hanes DP., Thompson KG., Leutgeb S., Schall JD., Leventhal AG., (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology*, 79 : 3272–3278.
- Scholl BJ., (2001). Objects and attention : the state of the art. *Cognition*, 80 (1) : 1-46.
- Shadlen MN., Movshon JA., (1999). Synchrony unbound : a critical evaluation of the temporal binding hypothesis. *Neuron*, 24 (1) : 67-77, 111-25.
- Shadlen MN., Newsome WT., (1998). The variable discharge of cortical neurons : implications for connectivity, computation and information coding. *The journal of Neuroscience*, 18 (10) : 3870-3896.
- Shapiro KL., Arnell KM., Raymond JE., (1997). The attentional blink. *Trends in Cognitive Sciences*, 1 (8) : 291-296.
- Shapiro KL., Raymond JE., Taylor TL., (1993). The attentional blink suppresses size and shape but not colour information. *Annual meeting of the Association for Research in Vision and Ophthalmology (ARVO)*. Sarasota, Florida.
- Shapiro, K., Driver, J., Ward, R., Sorensen RE., (1997). Priming from the attentional blink : a failure to extract visual tokens but not visual types. *Psychological Science*, 8 : 95-100.
- Shepard RN., Metzler J., (1971). Mental rotation of three-dimensional objects. *Science*, 171 : 701-703.
- Shumikhina, S., Molotchnikoff, S., (1999). Pulvinar participates in synchronizing neural assemblies in the visual cortex in cats. *Neuroscience Letters*, 272 : 135-139.
- Singer W., (2000). Why use more than one electrode at a time ? *New Technologies for Life Sciences: A Trends Guide*. Special Issue, 12-17.
- Sireteanu, R., Rettenbach, R., (2000). Perceptual learning in visual search generalizes over tasks, locations and eyes. *Vision Research*, 40 (21) : 2925-2949.
- Sohn YS., Liederman J., Reinitz MT., (1996). Division of inputs between the hemispheres eliminates illusory conjunctions. *Neuropsychologia*, 34 : 1057-1068.
- Sougné J., (1997). Variable oscillation frequencies for solving the problem of multiple instantiation. *Proceedings of the Nineteenth conference of the Cognitive Science Society*. Mahwah, NJ : Lawrence Erlbaum.
- Sperling G., (1960). The information available in brief visual presentations. *Psychological Monographs*, 74 (498).
- Steinmetz PN., Roy A., Fitzgerald PJ., Hsiao SS., Johnson KO., Niebur E., (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*, 404 : 187-190.

- Steriade M., (2000). Corticothalamic resonance, states of vigilance and mentation. *Neuroscience*, 101 (2) : 243-276.
- Struber D., Basar-Eroglu C., Hoff E., Stadler M., (2000). Reversal rate dependent differences in the EEG gamma band during multistable perception. *International Journal of Psychophysiology*, 38 : 243-252.
- Stryker MP., (1989). Is Grandmother an oscillation ? *Nature*, 338 : 297-298.
- Subramaniam S., Biederman I., Madigan S., (2000). Accurate identification but no priming and chance recognition memory for pictures in RSVP sequences. *Visual Cognition*, 7 (4) : 511-535.
- Tallon-Baudry C., Bertrand O., (1999). Oscillatory γ activity in humans and its possible role in object representation. *Trends in Cognitive Science*, 3 : 151-162.
- Tallon-Baudry C., Bertrand O., Delpuech C., Pernier J., (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, 16 (13) : 4240-4249.
- Tallon-Baudry C., Bertrand O., Delpuech C., Pernier J., (1997). Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *Journal of Neuroscience*, 17 (2) : 722-734.
- Tallon-Baudry C., Bertrand O., Fischer C., (2001). Oscillatory synchrony between human extrastriate areas during visual short-term memory maintenance. *Journal of Neuroscience*, 21, RC177 : 1-5.
- Tanaka JW., Weiskopf D., Williams P., (2001). Of color and objects : The role of color in high-level vision. *Trends in Cognitive Sciences*, 5, 211-215.
- Tanaka K., (1993). Neuronal mechanisms of object recognition. *Science*, 262 : 685-688.
- Tanaka K., (1997). Mechanisms of visual object recognition : monkey and human studies. *Current Opinion in Neurobiology*, 7 : 523-529.
- Thibos LN., Levick, WR., (1985). The orientation bias of brisk-transient (Y) ganglion cells is different for alternating and drifting gratings. *Experimental Brain Research*, 58 : 1-10.
- Thorpe S., Fize D., Marlot C., (1996). Speed of processing in the human visual system. *Nature*, 381 : 520-522.
- Timsit M., (2001). Les potentiels en relation avec les événements (ERPS) Problèmes méthodologiques et problèmes d'interprétation. In *Les potentiels évoqués : une fenêtre sur la cognition*. Atelier de formation INSERM N°123.
- Tononi G., Edelman GM., Sporns O., (1998). Complexity and coherency : integrating information in the brain. *Trends in Cognitive Sciences*, 2 : 474-484.
- Tovee MJ., Rolls ET., (1995). Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex. *Visual Cognition*, 2 : 35-39.
- Treisman A., (1986). Features and objects in visual processing. *Scientific American*, 255 (5) : S114-125.
- Treisman A., (1996). The binding problem. *Current Opinion in Neurobiology*, 6 (2) : 171-178.
- Treisman A., (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London, Serie B*, 353 : 1295-1306
- Treisman A., Gormican S., (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95 (1) : 14-48.
- Treisman A., Schmidt H., (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14 (1) : 107-141.
- Treisman AM., Gelade G., (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12 (1) : 97-136.

- Treisman AM., Kanwisher NG., (1998). Perceiving visually presented objects : recognition, awareness, and modularity. *Current Opinion in Neurobiology*, 8 : 218-226.
- Tsal Y., (1989). Illusory conjunctions support the feature integration theory ? A critical review of theory and findings. *Journal of Experimental Psychology : Human Perception and Performance*, 15 : 394-400.
- Tsal Y., Meiran N., Lavie N., (1994). The role of attention in illusory conjunctions. *Perception and Psychophysics*, 55 (3) : 350-358.
- Turvey MT., (1973). On peripheral and central processes in vision : inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Reviews*, 81 : 1-52.
- Ullman S., (1995). Sequence seeking and counter streams : a computational model for bidirectional information flow in the visual cortex. *Cerebral Cortex*, 5 (1) : 1-11.
- Ungerleider LG., Haxby JV., (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, 4 : 157-165.
- Ungerleider LG., Mishkin M., (1982). Two cortical visual systems. In DJ., Ingle, MA., Goodale et RJW. Mansfield (Eds.), *Analysis of visual behavior*, Cambridge, MIT Press, 549-586.
- Uttal W., (2001). *The new phrenology : the limits of localizing cognitive processes in the brain*. Cambridge, MIT Press.
- Van Rullen R., (2000). Une première vague de potentiels d'action : une première idée de la scène visuelle. Thèse de doctorat de l'Université Paul Sabatier.
- Van Rullen R., Thorpe SJ., (2001). The time course of visual processing : from early perception to decision making. *Journal of Cognitive Neuroscience*, 13 (4) : 454-461.
- Varela FJ., (1995). Resonant cell assemblies : a new approach to cognitive functions and neuronal synchrony. *Biological Research*, 28 (1), 81-95.
- Varela FJ., Lachaux JP., Rodriguez E., Martinerie J., (2001). The brainweb : phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2 (4) : 229-239.
- Vecera SP., O'Reilly RC., (1998). Figure-ground organization and object recognition processes : An interactive account. *Journal of Experimental Psychology : Human Perception and Performance*, 24 : 441-462.
- Virzi RA., Egeth HE., (1984). Is meaning implicated in illusory conjunctions ? *Journal of Experimental Psychology : Human Perception and Performance*, 10 (4) : 573-580.
- Von Bertalanffy L., (1993). *Théorie générale des systèmes*. Paris Dunod.
- Von Der Malsburg C., (1981). The correlation theory of brain function. Internal Report 81-2, Göttingen, Germany : Max Planck Institute for Biophysical Chemistry.
- Von Der Malsburg C., (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology*, 5 (4) : 520-526.
- Wandell BA., (2000). Computational neuroimaging : color representation and processing. In M. Gazzaniga (Ed.) *The new cognitive neuroscience*. Cambridge, MIT Press.
- Weichselgartner E., Sperling, G., (1987). Dynamics of automatic and controlled visual attention. *Science*, 238 : 778-780.
- Wennecker T., Palm G., (1997a). Cell assemblies, Associative Memory and temporal structure in brain signals. In R. Miller (Ed.) *Time and the brain. Conceptual Advances in Brain Research*, Vol. 2. Harwood Academic Publishers.
- Wennecker T., Palm G., (1997b). On the relation between neural modelling and experimental neuroscience. *Theory in Biosciences*, 116 : 273-289.

- Wexler M., Kosslyn SM., Berthoz A., (1998). Motor processes in mental rotation. *Cognition*, 68 : 77-94.
- Wheeler ME., Treisman A., (2002). Binding in short-term visual memory. *Journal of Experimental Psychology : General*, 131 (1) : 48-64.
- Wolfe JM., Cave KR., (1999). The psychophysical evidence for a binding problem in human vision. *Neuron*, 24 : 11-17.
- Wright JJ., (1997). EEG Simulation : Variation of Spectral Envelope, Pulse Synchrony and ~40 Hz Oscillation. *Biological Cybernetics*.
- Yen SC., Menschik ED., Finkel LH., (1998). Extraction of Perceptually Salient Contours by Striate Cortical Networks. *Vision Research* 38 (5) : 719-741.
- Yen SC., Menschik ED., Finkel LH., (1999). Perceptual grouping in striate cortical networks mediated by synchronization and desynchronization. *Neurocomputing*, 26-27 (1-3) : 609-616.

Annexes

ANNEXE I : ELEMENTS DE BASE DE L'ANALYSE DU SIGNAL.....	172
ANNEXE II : ANALYSE DE BASE DES SIGNAUX EEG.....	175
ANNEXE III : NIVEAUX D'ANALYSE DES SIGNAUX CEREBRAUX	176
ANNEXE IV : LES CODAGES TEMPORELS : OUTILS DES NIVEAUX MICROSCOPIQUES.....	178
ANNEXE V : LES CODAGES TEMPORELS : NIVEAUX MESOSCOPIQUE ET MACROSCOPIQUE	181
ANNEXE VI : OUTILS TRANSVERSES DES NIVEAUX MICROSCOPIQUES ET MESOSCOPIQUES	184
ANNEXE VII : SYNCHRONISATION MICROSCOPIQUE : QUELS EVENEMENTS NEURONAUX ?.....	185
ANNEXE VIII : LES CODAGES TEMPORELS : VISIBILITE MACROSCOPIQUE DES OSCILLATIONS.....	186
ANNEXE IX : MISE EN EVIDENCE DES OSCILLATIONS INDUITES.....	187
ANNEXE X : VARIETE DES MODELES COLORIMETRIQUES.....	188
ANNEXE XI : LE CLIGNEMENT ATTENTIONNEL CHEZ L'ENFANT	190
ANNEXE XII : STIMULI UTILISES EN RSVP CHEZ L'ENFANT	191
ANNEXE XIII : DESCRIPTEURS UTILISES POUR LA REGRESSION, EXPERIENCE PRELIMINAIRE A	193
ANNEXE XIV : CARACTERISTIQUES DE L'ECHANTILLON ADULTE (EXPERIENCE PRELIMINAIRE B).....	196
ANNEXE XV : LOGISTIQUE DE L'EXPERIMENTATION EN TACHISTOSCOPIE.....	197
ANNEXE XVI : STIMULI DES EXPERIENCES 1, 2, 4 ET 5.....	202
ANNEXE XVII : PROPRIETES DES STIMULI DE SUBSTITUTION	203

Annexe I : Eléments de base de l'analyse du signal

Les variations du courant intra-cellulaire dans un neurone, tout autant que les courants bioélectriques enregistrés dans une population de neurones, peuvent être approximés par des **ondes** pseudo-périodiques. Nous allons donc aborder les éléments de base de l'analyse du signal. Des informations plus complètes peuvent être obtenus dans l'ouvrage vulgarisé de Barbara Burke-Hubbard (1995).

Pour commencer, récapitulons les deux paramètres de base de l'analyse des ondes. On les observe sur le tracé des variations du signal au cours du temps. Pour une onde parfaitement périodique (une fonction sinusoïdale), ces deux paramètres sont **l'amplitude et la fréquence**. L'amplitude pic-à-pic se caractérise par l'écart entre l'intensité minimale et l'intensité maximale (voir figure I.A). La fréquence (inverse de la période) indique le nombre de répétitions d'un schéma élémentaire (une oscillation) au cours d'une unité de temps (généralement la seconde). Elle s'exprime en Hertz ou cycles par seconde.

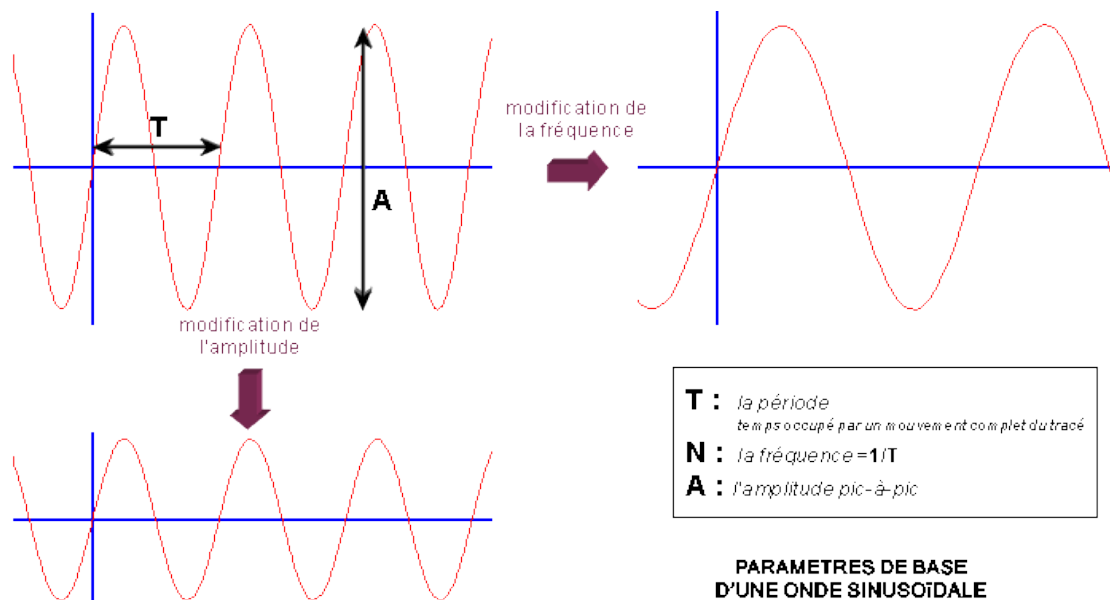


figure I.A : paramètres de base d'une onde sinusoïdale

Les ondes observées dans le cerveau n'obéissent pas une régularité aussi saisissante que cette onde mathématique $f(x) = \sin(x)$. Mais un arsenal d'outils d'analyse existe. Leur origine est due à Fourier, en 1807, qui caractérise les fonctions périodiques. **Fourier** a découvert que des fonctions périodiques ordinaires peuvent être décomposées en une somme de composantes sinusoïdales. Un exemple de décomposition est donné à la figure I.B. Ces ondes ordinaires peuvent donc être résumées par des paramètres similaires de fréquence et d'amplitude, à ceci près que ces ondes comportent de multiples composantes, ayant leurs paramètres propres. La fonction initiale, ou série de Fourier, exhibe les informations temporelles, tandis que la fonction transformée, ou « transformée de Fourier », exhibe les informations de fréquence et masque les informations temporelles. Il s'agit d'une **décomposition en fréquence**.

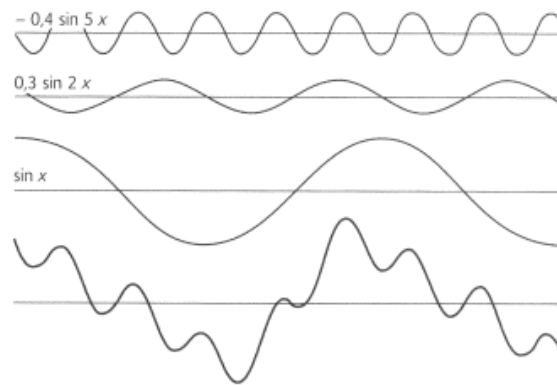


Figure 1.B : décomposition d'une onde ordinaire en composantes sinusoïdales (reproduit de Burke-Hubbard B., 1995) Une onde périodique contient donc des fréquences diverses, ayant un poids (une amplitude) plus ou moins grand. Ainsi, le signal présenté en bas est la somme des trois sinusoïdes situées au dessus.

Pour un signal périodique, le temps n'est pas une information primordiale, puisqu'il ne se produit pas de changement ; le schéma de base se reproduit de façon immuable. Pour les signaux non linéaires (comportant des discontinuités et des changements non prévisibles), il fallait restaurer la dimension temporelle. Ce fut chose faite en 1946, lorsque Dennis Gabor proposa l'analyse de **Fourier à fenêtres**. Il s'agit d'une **décomposition temps-fréquence**. On choisit une taille de fenêtre, selon que l'on souhaite observer les hautes ou les basses-fréquences, et on réalise la décomposition en déplaçant la fenêtre le long du signal.

Des sophistications et des révolutions ont eu lieu depuis. L'objectif principal était de juguler au mieux les difficultés liées au principe d'Heisenberg. Ce principe postule l'impossibilité de localiser simultanément un signal en temps et en fréquence. Le produit des **incertitudes en temps et en fréquence** ($\Delta t \Delta f$) possède une borne inférieure. Plus le signal est bien caractérisé en temps, moins il est bien caractérisé en fréquence, et vice-versa. En corollaire, un signal de basse-fréquence est moins localisable en temps qu'un signal de haute-fréquence. Pour faire simple, disons que pour avoir la possibilité d'observer une oscillation de basse-fréquence, il faut laisser se dérouler davantage de temps. Le paramètre fréquentiel obtenu caractérise une plage temporelle plus grande (plus indéterminée) que s'il avait été question d'une haute fréquence.

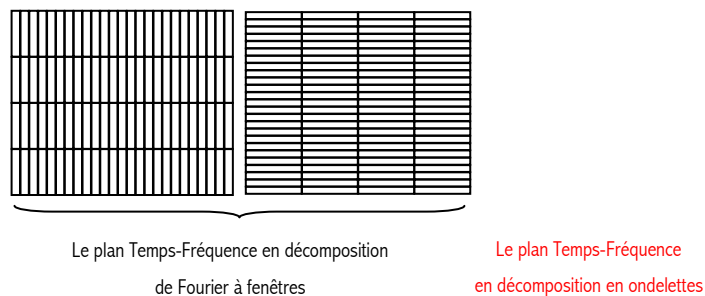


Figure 1.C : Les plans Temps-Fréquence en analyse de Fourier à fenêtre et en analyse en ondelettes

L'analyse en **ondelettes**, une **décomposition temps-échelle** (eg. Jean Morlet, vers 1975, cité dans Burke-Hubbard B., 1995), permet ensuite d'exhiber simultanément les hautes et les basses fréquences, grâce à l'introduction d'échelles dans le plan temps-fréquence (voir figure 1.C). L'analyse de Fourier à fenêtres tenait constante la taille de la fenêtre, et

dans cette fenêtre, le nombre d'oscillations variait. Morlet propose une méthode dans laquelle le nombre d'oscillations reste constant, et la taille de la fenêtre varie. Cette méthode d'analyse, conçue initialement comme un artifice ayant de bons résultats empiriques, se révéla être une méthode exacte de représentation du signal : l'erreur y est nulle, c'est-à-dire que les transformations et reconstructions successives du signal n'engendrent pas de résidu. En outre, cette méthode est robuste, car de petites variations ne forment pas des modifications démultipliées dans la transformation ou dans la reconstruction.

Aujourd'hui, de nouvelles classes d'outils sont disponibles, pour une **décomposition temps-fréquence-échelle** (paquets d'ondelettes, ondelettes de Malvar, *matching pursuits*) dont certaines sont prometteuses pour l'analyse du signal cérébral.

Annexe II : Analyse de base des signaux EEG

Les plus fréquentes des analyses quantitatives appliquées au signal EEG sont les analyses fréquentielles. Elles reposent sur les méthodologies présentées en Annexe I. Les gammes de fréquences des signaux cérébraux ont été découpées comme présenté à la table II.A.

Gamme de fréquences	appellation
0-4 Hz	delta
4-8 Hz	theta
8-12 Hz	alpha
12-20 Hz	beta
20-80 Hz	gamma
80-200 Hz	« ripples »

Table II.A : les bandes fréquentielles du signal EEG

Ont ensuite été développées des méthodologies destinées à comprendre le comportement du cerveau soumis à des stimulations auditives, visuelles ou autres. Ces méthodes sont les potentiels évoqués (PE, ou *ERP* : *event related potentials*).

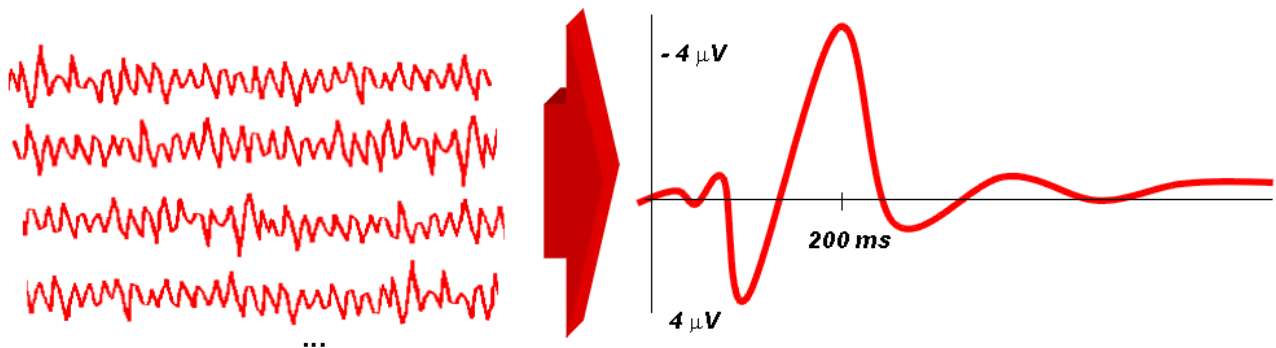


figure II.A : les potentiels évoqués (à droite) sont calculés par moyennage d'un grand nombre de tracés EEG (à gauche).

Les PE sont construits de la manière suivante. Les séquences EEG à moyennner sont calées sur le moment d'apparition du stimulus (voir figure II.A). Les fluctuations aléatoires sont éliminées par le moyennage, et il ne reste dans le tracé du potentiel évoqué que les composantes systématiques liées au traitement du stimulus. Ce tracé est analysé en terme d'ondes (composantes de forte amplitude apparaissant à un moment donné), de leur polarité (positive ou négative), de leur latence (survenue à 50, 100 ms, 300, 400 millisecondes...), de leur amplitude. Le tracé fictif présenté à la figure II.B montre une onde P1 (ou P100), onde positive apparaissant 100 millisecondes après la présentation du stimulus, et une onde N2 (ou N200), onde négative apparaissant 200 millisecondes après la présentation du stimulus.

Annexe III : Niveaux d'analyse des signaux cérébraux

On parle d'**enregistrements unitaires** (*SUA : single-unit activity*) lorsqu'une micro-électrode bipolaire est placée de part et d'autre de la membrane pour enregistrer les différences de potentiel entre l'extérieur et l'intérieur de la cellule. Le signal est alors saturé par l'activité propre du neurone, c'est-à-dire par les flux ioniques entrants et sortants. En particulier, il révèle les potentiels d'action (*spikes*) émis par le neurone. Cet enregistrement peut être discrétisé en un *raster*, un diagramme qui exprime l'occurrence de potentiels d'actions au cours du temps (voir figure III.A). Au cours d'un unique essai, ou au cours de multiples répétitions de cet essai, peut être construit le diagramme de l'évolution temporelle de la fréquence de potentiels d'action ou *PSTH* (voir figure III.A).

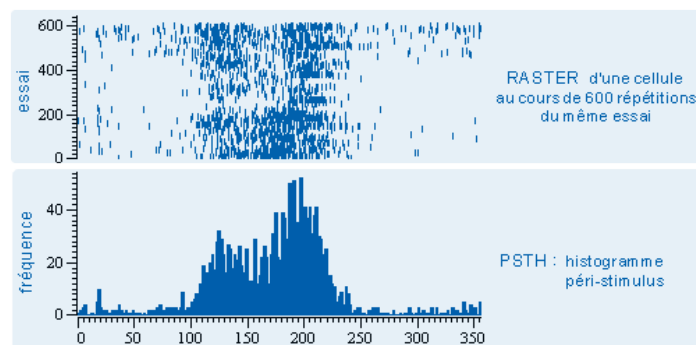


Figure III.A : les potentiels d'action d'une cellule neuronale au cours de mesures répétées (*Raster*), et l'histogramme de fréquence de ces potentiels d'action (*PSTH*). Dans l'exemple, on observe que la fréquence de décharge du neurone est maximale 200 millisecondes après l'apparition du stimulus.

On parle de **potentiels de champs locaux** (*LFP : local field potential*) lorsque l'enregistrement est réalisé en profondeur dans le cerveau (l'électrode est placée au milieu d'une population de neurones), et qu'un filtre passe-bas (0-100 Hz) est appliqué au signal recueilli. Le signal est alors saturé par l'activité des neurones situés à proximité immédiate de l'électrode, mais comporte également les contributions décroissantes de neurones situés à plus grande distance. Notons que la taille de l'électrode influe sur le nombre de neurones simultanément incrustés dans le signal. Ce sont les dendrites, et tout particulièrement les dendrites pyramidales, qui forment ce signal (Wright, 1997).

On parle d'**enregistrement extracellulaire des potentiels d'action** (*MUA ; multi-unit activity ou multi-unit spikes*) lorsque l'enregistrement est réalisé au milieu d'une population de neurones par le biais d'électrodes profondes, et que le signal est conditionné pour révéler les potentiels d'action émis à proximité de l'électrode. Le conditionnement consiste à appliquer un filtre passe-haut (eg. 500-3000 Hz), et on relève de manière indissociée les potentiels d'action de plusieurs neurones. On parle de « *dirty data* », ou d'enregistrement multi-neurones non résolus (*unresolved MUA*) lorsque l'on calcule le raster de ce signal en laissant indissociées les contributions de plusieurs neurones. On parle d'enregistrements multi-neurone résolus (*resolved MUA*) lorsque l'on isole les rasters de chaque neurone. Ce résultat s'obtient par des traitements statistiques (eg. recherche de valeurs pic-à-pic, K-Mean, ACP). On se retrouve alors dans la situation d'enregistrement unitaire (voir figure III.B). Cette configuration d'enregistrement multiple des potentiels d'action est préférée à la situation d'enregistrement unitaire, car elle permet le recueil de davantage d'événements, et améliore ainsi les possibilités de calcul statistique. Il est ainsi possible de reconstruire les *rasters* de plusieurs cellules,

recueillies simultanément lors d'un unique enregistrement. Notons que certains auteurs préfèrent utiliser des MUA non résolus plutôt que des MUA résolus, argumentant que cette stratégie est mieux adaptée à l'analyse de codages par populations (eg. voir Singer, 2000). Elle préserve les composantes d'oscillations synchrones (collectives), et annule les composantes spécifiques de chaque neurone.



Figure III.B : Principe de reconstruction des trains de potentiels d'action de 3 neurones par conditionnement du signal mésoscopique. Des neurones situés à proximité de l'électrode diffèrent en taille, en forme, et en distance à l'électrode. Lorsqu'ils génèrent des PA, l'onde résultante diffère en forme et en amplitude. Grâce à ces variations, il est possible de ségréguer les signaux provenant des différents neurones contribuant au signal.

Enfin, on parle d'**EEG** lorsque l'enregistrement est réalisé par amplification des courants recueillis à travers le scalp. Globalement, le signal est saturé par l'activité des neurones corticaux situées en dessous de l'électrode, mais comportent les contributions d'impact décroissants de neurones situés à distance respectable. On estime qu'une seule électrode comporte les activités synaptiques de quelques 10 millions à 1 milliard de neurones, et qu'elle enregistre une surface d'environ 50 à 100 cm² (Nunez P., 2000).

De par les neurones qui les constituent, les couches externes du cortex diffusent peu ; leur activité ne se propage pas sur plusieurs électrodes. En revanche, les couches plus profondes diffusent fortement, et la fusion des sources est plus grande pour ces signaux. Les neurones pyramidaux, ayant une **diffusion** axonale assez large, et une distribution parallèle, permettent une action de masse et contribuent fortement au signal (Timsit M., 2001). Les hautes fréquences sont filtrées à travers l'os crânien, et subsistent essentiellement les basses-fréquences. L'échantillonnage de ces signaux se réalise usuellement dans la bande de fréquence [0 Hz ; 100 Hz].

Ces trois niveaux d'observation sont soumis à diverses **restrictions expérimentales**. Les niveaux microscopique et mésoscopique sont réservés à des expérimentations menées in vitro ou sur l'animal, ou à de rares cas d'intervention neurochirurgicale chez l'homme. Ils impliquent une logistique et une expertise lourde. Les protocoles réalisés sur le signal EEG sont de mise en oeuvre plus facile, ce qui explique la recherche de ponts entre les différents niveaux d'observation.

Annexe IV : Les codages temporels : outils des niveaux microscopiques

L'inspection visuelle du raster permet d'apprécier l'existence de **synchronisations des potentiels d'action**, en inter-essai au niveau d'un neurone (voir exemple de la synchronisation « *steady-state* » figure IV.A), ou entre plusieurs neurones. Les **coïncidences** des décharges apparaissent visuellement par l'alignement des décharges.

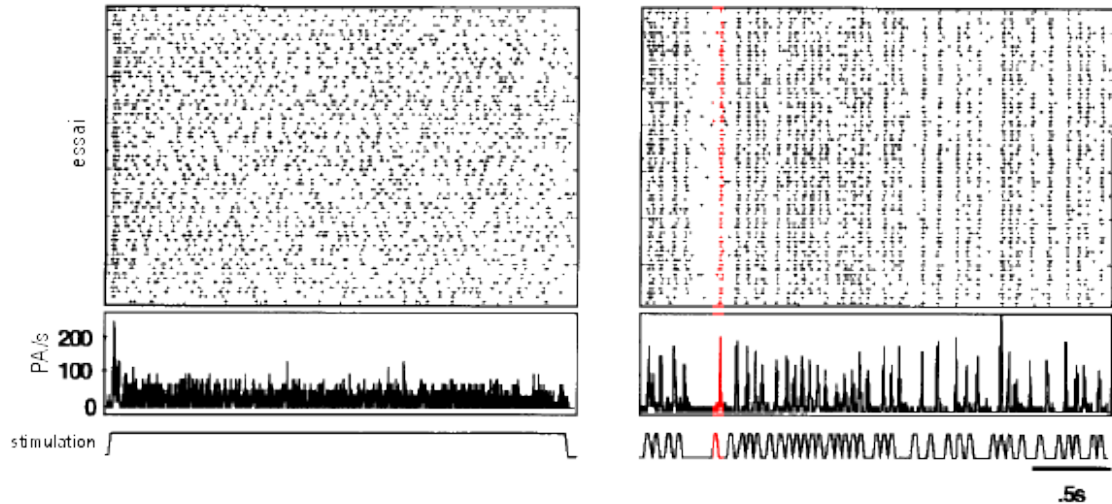


figure IV.A : synchronisation inter-essai des potentiels d'action d'un neurone. Lorsque la stimulation est maintenue en continu, le neurone décharge de manière aléatoire d'un essai à l'autre. Si la stimulation est périodique, les potentiels d'action montrent des relations temporelles avec la survenue du stimulus. Elles se synchronisent au stimulus, et les PA (potentiels d'action) s'alignent dans le raster. On parle alors de synchronisation « *steady-state* » (reproduit de DeCharms & Zador, 2000)

L'**autocorrélation**¹ (également appelée *IHH* : *inter-event histogram* ou *ISI* : *inter-spike interval*) permet de mesurer les relations causales existant entre les potentiels d'action (PA) d'un neurone. Si la position (temporelle) d'un potentiel d'action du neurone a une incidence sur le positionnement temporel du potentiel d'action suivant, et si cette relation est linéaire, alors, des pics et des dépressions apparaissent sur ce tracé. Le mode de construction d'un autocorrélogramme est reporté à la figure IV.B.

¹ Ce terme est polysémique. On parle également d'autocorrélation (ACH) pour des histogrammes de cross-corrélation comparant le signal à lui-même, pour découvrir des composantes oscillant à une certaine fréquence. Cette seconde méthode consiste à vérifier si le signal est auto-semblable de moment en moment. L'autosimilarité est estimée par une cross-corrélation (voir ci-après) comparant le signal à une version plus ou moins décalée (un shift) de lui-même. Cette méthode est identifiable d'emblée apr l'axe des abscisses : si l'axe est symétrique autour de zéro, il s'agit d'une auto[*cross*]-corrél¹ation (ACH). Si l'axe est positif, il s'agit de l'autocorrélation (ISIH).

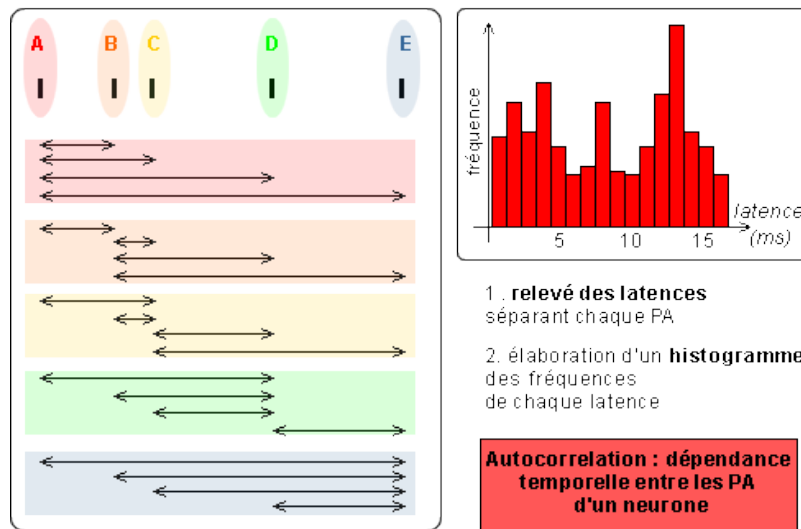


Figure IV.B : construction de l'autocorrélogramme d'un neurone (Perkel, Gerstein & Moore, 1967). Le raster (en haut à gauche, exemple de cinq décharges, de A à E) permet de relever toutes les latences inter-décharges (symbolisées par la longueur des flèches en bas à gauche). Ces latences sont reportées dans un histogramme (en haut à droite) qui traduit la fréquence des délais entre les PA émis par ce neurone.

L'autocorrélation permet éventuellement d'observer, au niveau microscopique, les **oscillations** des décharges d'un neurone. Ici, 'oscillation' caractérise primitivement l'aspect sinusoïdal de l'autocorrélogramme. Il s'agit plus exactement d'un comportement périodique dans la probabilité de décharges (voir exemple figure IV.C). Ces oscillations traduisent un déterminisme linéaire dans la dynamique neuronale, c'est-à-dire une tendance à adopter une fréquence de décharge régulière (voir avertissement annexe VII). Cette fréquence unique (ou une composition fréquentielle plus complexe) peut être estimée à partir du **périodogramme** (figure IV.C).

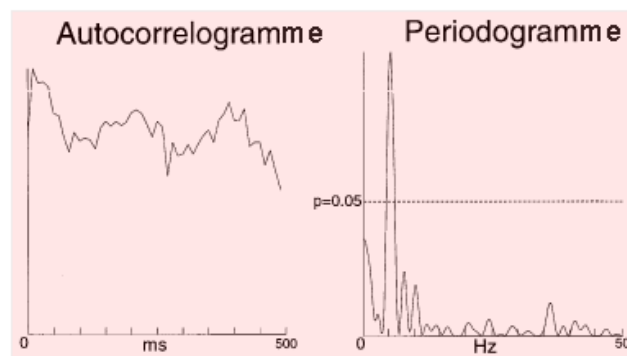


Figure IV.C : exemples d'auto-corrélogramme et de périodogramme (adapté de Kaneoke & Vitek, 1996)

La **crosscorrél**ation (XCG, XCH, ou CCH : **cross-correlation histogram**) traduit la dépendance linéaire entre les potentiels d'actions de deux cellules neuronales (figure IV.D). Il permet de répondre à la question « Est-ce que ce neurone A présente ses PA en relation temporelle avec ceux du neurone B ? ». Si la réponse est négative, l'histogramme est plat. Dans le cas contraire, l'histogramme présente un pic spécifique (figure IV.E). Ce pic indique une tendance des neurones à émettre leurs décharges en relation temporelle, que ces décharges soient simultanées ou décalées dans le temps. Si ce pic est centré sur zéro, cela suggère en outre que les décharges des neurones sont simultanées (les neurones synchronisent leur décharges ; font coïncider leur décharges).

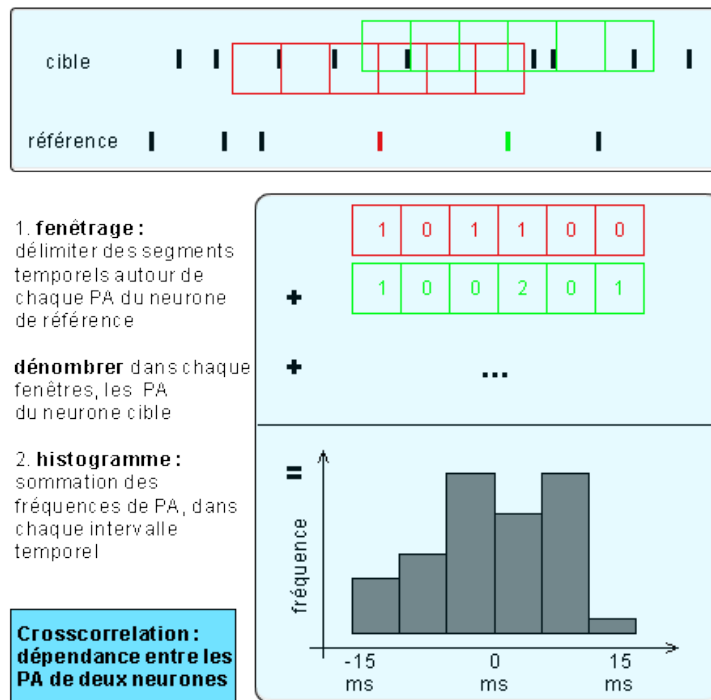


Figure IV.D : construction du cross-corrélogramme de deux neurones. Sur le raster (en haut), on repère individuellement chaque décharge du neurone de référence (eg. décharge représentée en rouge et en vert). L'axe temporel du neurone cible est discrétisé (en haut, fenêtrages rouge et vert, correspondant respectivement aux décharges rouge et verte). On dénombre les potentiels d'action du neurone cible situé dans chaque segment temporel autour des PA du neurones de référence. Ces fréquences sont reportées dans un histogramme, qui traduit la dépendance des deux neurones.

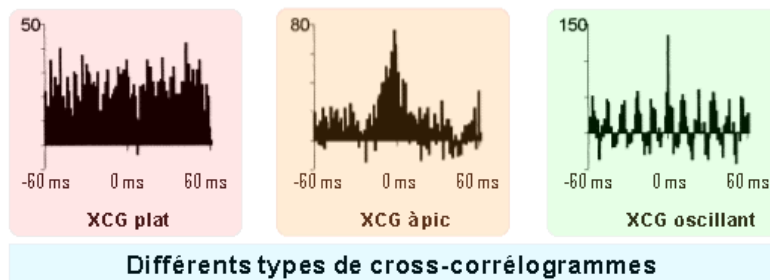
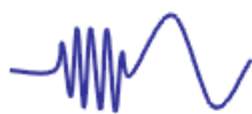


Figure IV.E : exemples de cross-corrélogrammes (reproduit de Mollotchnikoff & Shumikhina, 2000). A gauche, un cross-corrélogramme plat, signifiant que les deux neurones ne présentent pas de relations temporelles dans leurs activités. Au centre un cross-corrélogramme avec pic central, signifiant que les neurones ont tendance à décharger de manière simultanée, mais sans comportement périodique coordonné. A droite, un cross-corrélogramme oscillant (un pic centré sur zéro et des satellites), signifiant que les neurones déchargent en même temps, et qu'ils adoptent la même fréquence périodique de décharge.

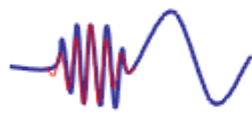
Annexe V : Les codages temporels : niveaux mésoscopique et macroscopique

A l'instar de l'auto-corrélogramme, qui traduit la dépendance temporelle des décharges de Potentiels d'action, il est possible d'étudier les changements de **comportement oscillatoires** des signaux molaires. Les oscillations peuvent concerner le signal dans sa globalité, ou des composantes fréquentielles spécifiques. Si le **signal global devient oscillant**, la périodicité devient visible sur le tracé. En revanche, si les oscillations concernent seulement une ou quelques composantes fréquentielles, leur observation directe sera compromise ; les comportements périodiques sont alors masqués par la complexité des composantes fréquentielles. Dans ce cas, on recourt à des analyses fréquentielles du signal (voir l'exemple d'une analyse en ondelette, figure V.A).

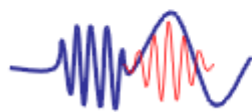


Ce signal comporte une partie oscillant à 45 Hz et une partie oscillant à 8 Hz.

Pour étudier les composantes de ce signal oscillant à 45 Hz, on déplace une ondelette de la fréquence appropriée le long du signal.



Lorsque l'ondelette se retrouve alignée en phase avec une composante de fréquence similaire, le produit (ondelette.signal) donne toujours une valeur positive : le coefficient d'ondelette sera élevé.



Lorsque l'ondelette est alignée avec une composante de fréquence dissimilaire, le produit est tantôt positif, tantôt négatif. Ces valeurs s'annulent en grande partie, le coefficient d'ondelette sera faible.

Pour étudier les autres composantes fréquentielles, on renouvelle l'opération avec des ondelettes de fréquences différentes.

figure V.A : principes de l'analyse des composantes fréquentielles d'un signal par les ondelettes.

De cette analyse, peut être extraite une visualisation quantitative des composantes fréquentielles (*PSD* : *power spectral density* ou puissance spectrale) au cours du temps. Par exemple, figure V.B, on observe une augmentation des composantes à 40-45 Hz, se produisant 300 millisecondes après la présentation du stimulus.

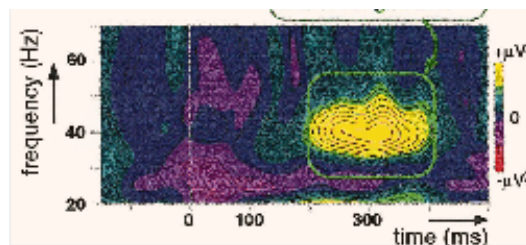


figure V.B : Variations du contenu fréquentiel du signal EEG : représentation des variations de la puissance spectrale dans le plan temps-fréquence (adapté de Tallon-Baudry & Bertrand, 1999). Cette représentation permet de visualiser les variations des composantes fréquentielles du signal, par rapport à un niveau de base.

La **cohérence fréquentielle**, appelée simplement ‘cohérence’, est une mesure des relations spécifiques existant entre deux sites du cerveau. A l’instar du crosscorrélogramme qui quantifie le lien entre deux neurones par les relations temporelles entre leurs potentiels d’action, la cohérence quantifie le lien entre deux sites, par les relations temporelles (relations de phase) entre paires de signaux molaires (deux signaux LFP, ou deux signaux EEG distincts). Pour chaque composante fréquentielle du signal EEG ou LFP, est mesurée la consistance des phases des deux signaux. Les variations de cet indice MSC (*magnitude squared coherence*) en fonction de la fréquence sont reportées dans un diagramme (voir figure V.C). Si les deux sites enregistrés sont en phase dans une bande fréquentielle, cela suggère soit une connexion directe entre eux, soit la présence d’un input commun. Notons cependant qu’en raison de la diffusion des signaux EEG et LFP dans le volume de conduction, une cohérence forte est observée de manière passive entre deux sites situés à faible distance. La cohérence mesure donc une **synchronisation fréquence-spécifique** (*frequency specific synchronization*), c’est-à-dire des **oscillations transitoirement couplées en phase** entre deux régions cérébrales. D’autres indices, versions améliorées ou alternatives de celui-ci, sont disponible pour la quantification des synchronisation de phase des signaux molaires, par exemple le PLV de (Lachaux, Rodriguez, Martinerie & Varela, 1999), ou l’AMVAR (*adaptive multivariate auto-regressive*, voir description dans (Bressler & Kelso, 2001).

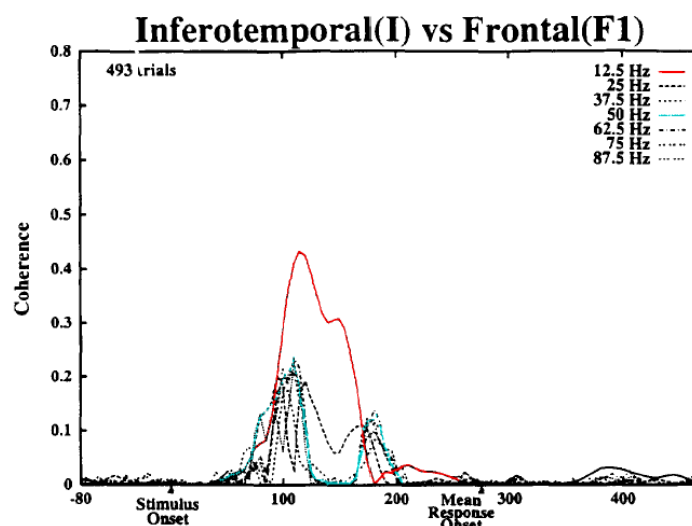


Figure V.C : variation de l’indice de cohérence au cours du temps (adapté de Bressler, 1995) entre les sites inféro-temporal et frontal, dans différentes bandes de fréquence. Observons par exemple l’augmentation de la cohérence bêta à 100 millisecondes après la présentation du stimulus (en rouge) ou l’augmentation de la cohérence gamma, à 100 et 185 millisecondes (en bleu).

Il est également possible de reporter à la fois la présence des oscillations dans une bande de fréquence donnée, et le couplage en phase de cette fréquence, comme indiqué à la figure V.D. La couleur du scalp indique la densité spectrale des rythmes gamma (plus la teinte est jaune claire, plus il y a de composantes gamma dans le signal par rapport au niveau de base). Les traits reliant les électrodes indiquent les signaux augmentant ou baissant en cohérence dans cette même bande gamma.

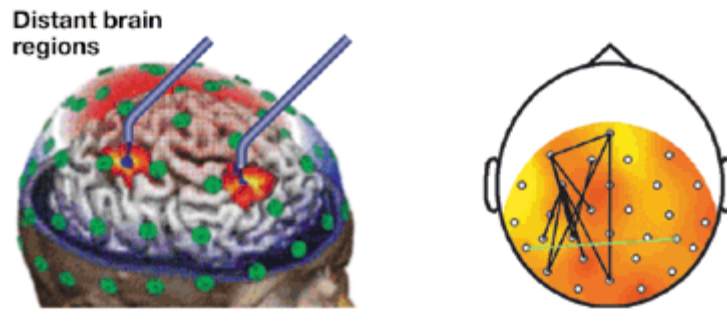


Figure V.D : cohérence entre deux sites neuronaux (reproduit de Varela, Lachaux, Rodriguez & Martinerie, 2001). Les électrodes qui augmentent significativement en cohérence sont reliées par des traits noirs. Les électrodes qui diminuent en cohérence sont reliées par des traits verts.

Outre la cohérence fréquentielle décrite ci-dessus, qui impliquent que les neurones oscillent sur la même bande fréquentielle, il existe d'autres modes de couplage des signaux. Brièvement, nous présentons figure V.E trois configurations possibles. Soit les deux signaux sont de même fréquence, soit l'un des signaux oscille sur une fréquence harmonique de l'autre, soit encore, la relation est d'ordre supérieur (ex. couplage quadratique des phases). Une illustration de ces trois cas de figure est donnée ci-après.

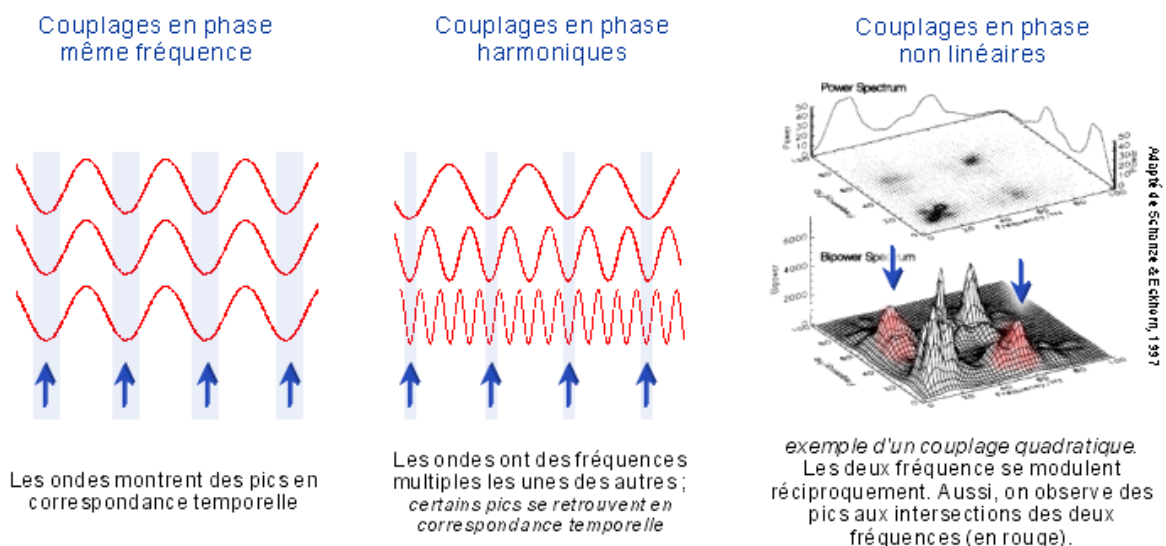


Figure V.E : trois types de couplages en phase. Le couplage linéaire, le couplage harmonique et le couplage quadratique

Annexe VI : Outils transverses des niveaux microscopiques et mésoscopiques

Certaines études ont mis en évidence des épisodes de synchronisation entre les signaux MUA et LFP. Ces deux signaux ne sont pas suscités par les mêmes mécanismes électrophysiologiques, les MUA révélant les potentiels d'action et le LFP étant davantage révélateur de l'intégration dendritique des potentiels post-synaptiques. Aussi, leur corrélation n'est pas triviale (Stryker, 1989). Les premières observations mettaient en correspondance les deux séries de signaux, et on observait directement qu'à une certaine latence après l'apparition du stimulus, les potentiels d'action du raster MUA se produisent durant les phases du signal EEG (voir figure VI.A).

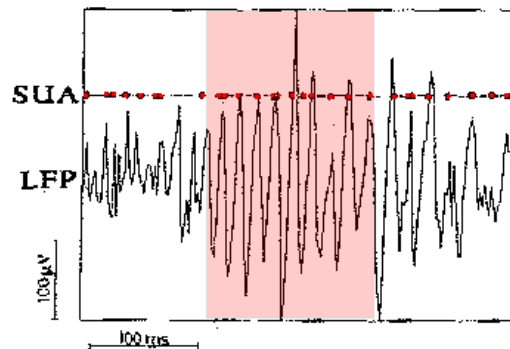


Figure VI.A : synchronisation des signaux LFP et SUA, à 100 millisecondes après la présentation du stimulus. On observe dans la partie rouge que les phases du LFP et les potentiels d'action du MUA coïncident transitoirement (adapté de Eckhorn & coll., 1988).

LE STA (*spike triggered average*) est une méthode quantitative reposant sur la même logique. Il permet de voir si **un neurone décharge en relation avec les phases du signal LFP**. Le STA procède de la même logique que les potentiels évoqués (voir annexe II), par un moyennage des ondes sur un grand nombre d'enregistrements. La méthode permet d'annuler le bruit et de faire ressortir les composantes systématiques. A contrario des Potentiels Evoqués qui calent les différents tracés EEG sur le moment d'apparition du stimulus, le STA cale les tracés LFP sur les décharges d'un neurone (voir figure VI.B). Si le neurone émet à un moment particulier de l'oscillation du signal LFP, on observe un pic caractéristique centré sur zéro. Des auteurs (eg. Fries & coll. 2001) arguent que cette technique est plus sensible que le cross-corrélogramme pour détecter les synchronisations oscillatoires.

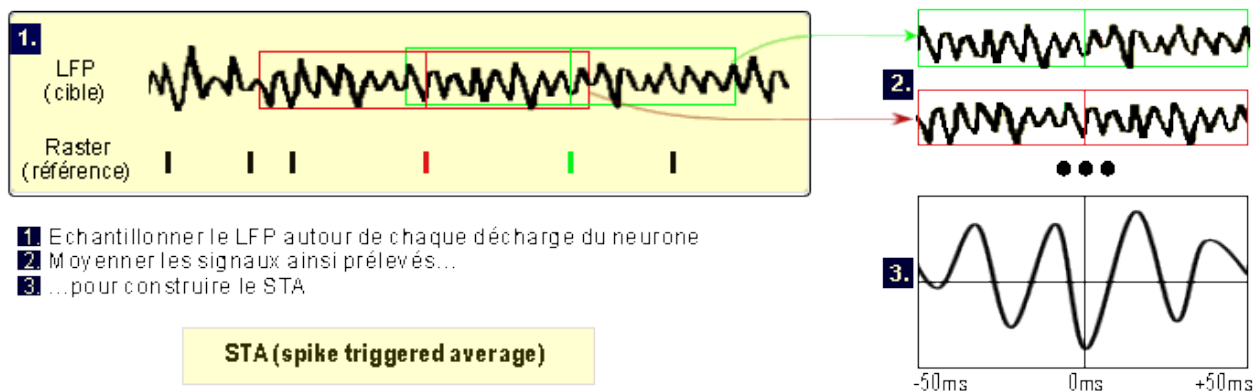


figure VI.B : construction du STA. Des segments du signal LFP sont échantillonnés autour des décharges du neurone cible. Les segments LFP ainsi extraits sont moyennés. Cela permet de construire un graphique révélant le comportement du LFP dans la proximité temporelle des décharges. Si ce graphe est plat, les deux signaux sont indépendants, en revanche, si ce graphe présente des oscillations, alors les deux signaux interagissent.

Annexe VII : Synchronisation microscopique : Quels événements neuronaux ?

Dans ses fondements théoriques, le concept d'assemblée neuronale synchrone est basé sur le niveau microscopique, et tout particulièrement sur le raster (voir annexe IV). Le phénomène neurophysiologique sujet à synchronisation peut être le **potentiel d'action** (*spike*) ou la **bouffée de potentiels d'action** (*burst*). Les études empiriques de la synchronisation ne donnent généralement que des indications très imprécises sur cette unité d'observation.

Un problème d'importance, lié à cette transformation du signal en raster, est la double-signification de la fréquence des signaux neuronaux. On distingue la fréquence des décharges (leur rareté au cours du temps) et la fréquence de l'onde engendrée ponctuellement par cette décharge (voir figure VII.A). La sélection d'outils d'analyse appropriés à chacun des deux problèmes est indispensable. Les outils devenant plus sophistiqués, les justifications sur ce sujet sont absentes des principales publications empiriques sur les synchronisations.

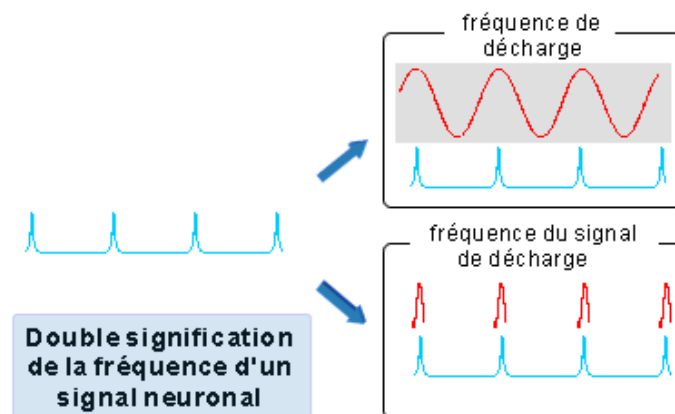


Figure VII.A : les deux fréquences du signal neuronal. En haut, la fréquence de l'événement. Ici la fréquence est périodique. Bien que cette pratique soit impropre, on la représenterait sinusoïdalement par l'onde rouge située au dessus. Elle ne doit pas être confondue avec la fréquence du signal de décharge (en bas, en rouge) qui quantifie sur des intervalles discrets, les propriétés spécifiques d'un potentiel d'action.

Annexe VIII : Les codages temporels : visibilité macroscopique des oscillations microscopiques

L'impact macroscopique des oscillations et des synchronisations microscopiques peut être estimé en considérant la sommation linéaire des signaux (voir figure VIII.A). Si chaque neurone contribuant au signal présente un comportement périodique (oscillation, comme c'est le cas pour les quatre exemples de la figure VIII.A)), il n'est pas certain que nous observions une **oscillation** résultante au niveau macroscopique. Quelque-soit la fréquence des neurones, si les activités ne sont pas en **phase**, le comportement oscillatoire¹ disparaît au niveau macroscopique (figure VIII.A, en haut).

En revanche, si les décharges d'un grand nombre de neurones sont en phase, on observera un effet au niveau macroscopique. Si des neurones ayant des fréquences de décharges diverses sont transitoirement mis en phase, un effet ponctuel sera observé, mais par définition cet événement n'est pas oscillatoire². A contrario, on observera une oscillation (un comportement périodique) seulement si ces décharges sont **synchrones en phase et de même fréquence**. A cette condition drastique seulement, la relation fractale entre les niveaux microscopique et macroscopique apparaît. Et à cette condition seulement, les outils fréquentiels du niveau macroscopique dévoileront de manière fractale les synchronies (simultanéités des décharges) du niveau microscopique.

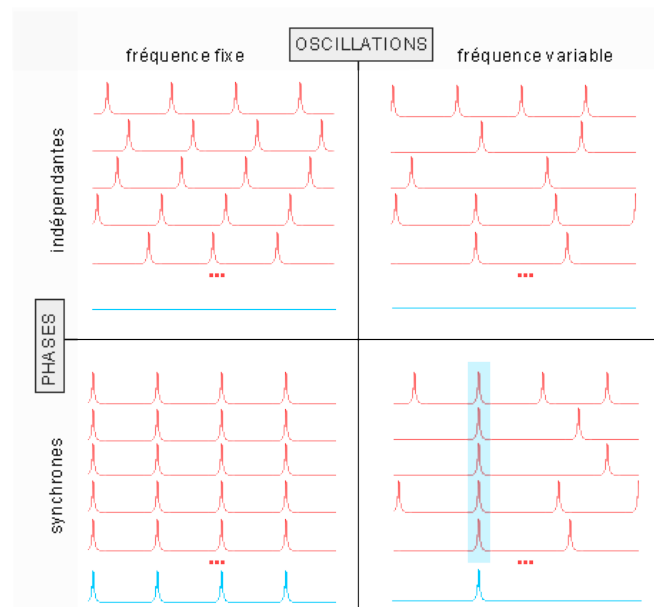


Figure VIII.A : Comportements oscillatoires au niveau microscopique (en rouge) et impact au niveau macroscopique (en bleu) de neurones ayant des décharges périodiques. Quelque-soit la fréquence, si les phases des activités neuronales sont indépendantes (en haut), on n'observera pas d'oscillations au niveau macroscopique. En revanche, si les phases des activités neuronales sont synchrones, on observera une oscillation macroscopique lorsque la fréquence est fixe (bas-gauche), ou un pic isolé lorsque la fréquence varie de neurones en neurones (transient phase-locking, en bas à droite). Notons que la synchronisation des phases de neurones ayant diverses fréquences oscillatoires ne peut être que ponctuelle, puisque par nature, les différences fréquentielles déphasent le signal au cours du temps.

¹ Cela ne signifie pas que la synchronisation des fréquences des neurones (le fait que plusieurs neurones adoptent la même fréquence de décharge) ne constitue pas un événement d'importance. Seulement, cet événement ne peut être enregistré au niveau macroscopique.

² Les analyses fréquentielles du signal impliquent une répétition du motif. Ces outils sont peu sensibles à un événement unique, c'est-à-dire un événement non oscillatoire par nature.

Annexe IX : Mise en évidence des oscillations induites

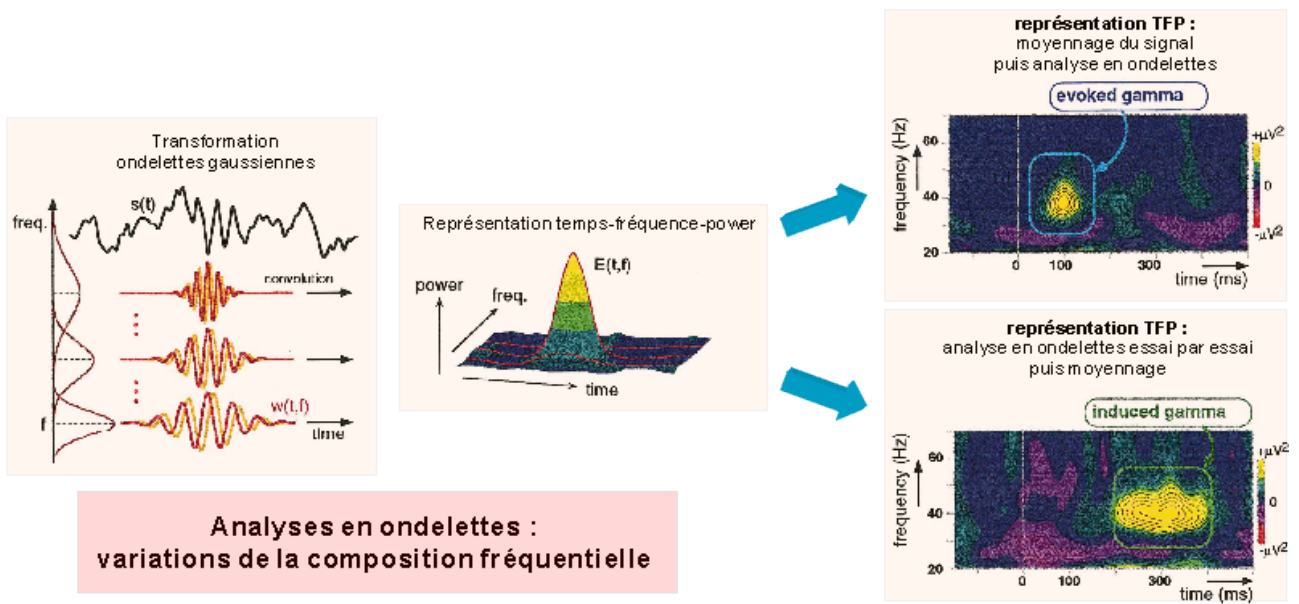


figure IX.A : les oscillations induites disparaissent des signaux moyennés. Aussi, l'analyse en ondelettes doit être appliquée directement sur les segments EEG, et ce sont les plans temps-fréquence qui sont moyennés (en haut, à droite). En revanche, les oscillations évoquées résistent au moyennage. L'analyse en ondelettes est réalisée directement sur le tracé moyenné des potentiels évoqués (en bas à droite) (adapté de Tallon-Baudry et Bertrand, 1999).

Annexe X : Variété des modèles colorimétriques

La longueur d'onde : éléments de base

La couleur est une qualité des objets liée à l'absorption et la réfraction de la lumière. Les surfaces des objets absorbent un certain nombre de longueurs d'ondes et en réfléchissent d'autres. Celles qui sont réfléchies peuvent être captées par notre rétine, si leurs longueurs d'onde sont situées entre 0,4 et 0,7 nm. En deçà et au delà, notre équipement ne nous permet pas de transformer la caractéristique physique en qualité psychologique. La source d'information qui nous parvient est donc une longueur d'onde (surface parfaitement unie¹) ou une composition spectrale de longueurs d'ondes allant du rouge au bleu (voir figure X.A).

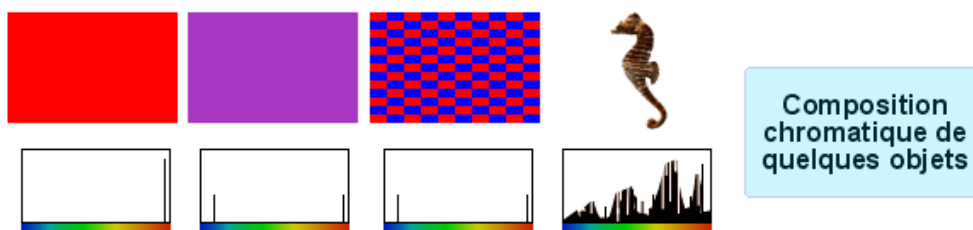


figure X.A : composition chromatique de quelques objets : une surface de vraie couleur (exemple, le rouge) et parfaitement unie, qui serait mesurée dans des conditions idéales d'illumination, présenterait une raie unique au spectromètre. Une fausse couleur telle que le violet présenterait deux raies, de même qu'une texture équi-répartie bleue et rouge. Une image d'objet présente usuellement une composition chromatique plus complexe.

Modèle RVB

A la rétine, nous avons vu que les récepteurs chromatiques étaient au nombre de trois, et présentaient des pics de réponse pour des longueurs d'onde de 437 nm, 533 nm et 564 nm. Ces réponses sont plausiblement reproduites par le modèle RVB (voir figure X.B), qui donne les trois valeurs numériques de quantité de rouge, de vert, et de bleu contenues dans une image.



Figure X.B : le modèle RVB. En a , les trois dimensions, en b, représentation dimensionnelle du modèle.

On observe sur la représentation dimensionnelle que le modèle superpose la couleur et la luminosité (un des sommets du cube code le noir et le sommet diamétralement opposé code le blanc). Un **objet** ayant un codage RVB de **(30;50;180)** sur une échelle de 256 contient « davantage » de bleu qu'un **objet** codé **(30;50;70)**. Mais il est également plus clair que le second objet. Cette caractéristique risque de nous empêcher de dissocier les effets de la couleur et de la luminosité.

¹ Une surface parfaitement unie est possible en théorie, mais n'existe pas dans les conditions naturelles

Composantes a & b du modèle Lab

Dès les cellules ganglionnaires de la rétine et jusqu'au cortex visuel strié, la couleur est essentiellement codée sur l'opposition rouge-vert de certaines cellules, et sur l'opposition jaune-bleu de certaines autres. Le modèle colorimétrique proche de cette caractéristique est le modèle Lab (prôné par la CIE : Commission Internationale de l'Eclairage). Outre la composante de luminance, il contient les deux composantes chromatiques a (antagonisme rouge-vert), et b (antagonisme jaune-bleu), voir figure X.C.

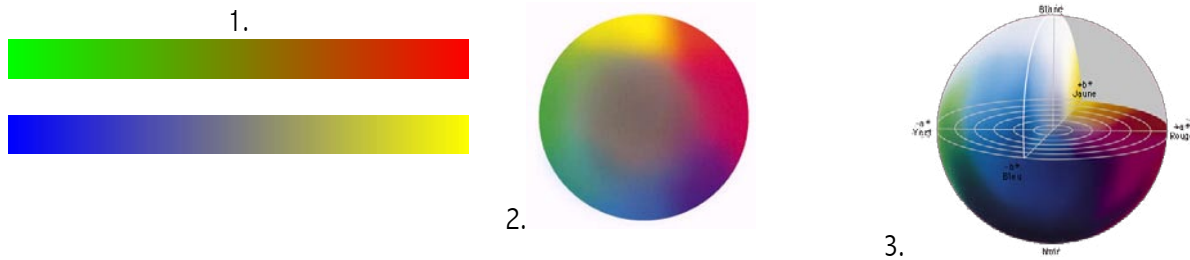


Figure X.C : le modèle Lab. En 1, les deux dimensions chromatiques du modèle Lab, en 2, représentation dimensionnelle des composantes chromatiques du modèle. En 3, représentation dimensionnelle du modèle complet.

Composante T du modèle TSL

Le dernier modèle que nous allons exposer s'appuie sur notre perception subjective d'une continuité entre les deux extrémités du spectre visible (le rouge et le bleu) via les fausses couleurs (le violet). Le modèle TSL (teinte - saturation - luminance) également nommé HSL (*hue, saturation, luminance*) arrange les différentes couleurs sur un anneau allant du rouge (0°) passant par le jaune (90°) puis par le vert-bleu (180°) et enfin par l'indigo (270°). Au contraire du modèle Lab, il s'agit d'une représentation en coordonnées polaire.

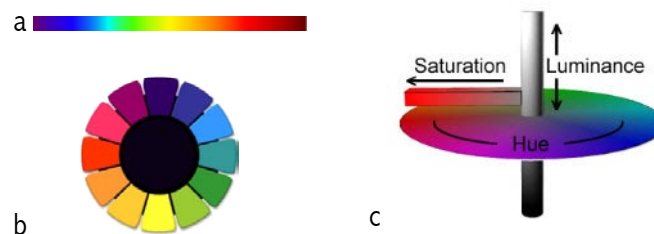


Figure X.D : le modèle TSL. En a, le spectre des couleurs visible déplié., en b, le repliage des deux extrémités autour des fausses couleurs violettes forme la roue chromatique du modèle TSL. En c, représentation dimensionnelle du modèle TSL.

Annexe XI : Le clignement attentionnel chez l'enfant

Le clignement attentionnel est un déficit de la perception d'une cible succédant dans des délais très brefs à une première cible à identifier. Les résultats présentés à la figure XI.A ont été obtenus chez des enfants âgés de 7 à 9 ans, et soumis au protocole de l'expérience préliminaire A (voir chapitre 3).

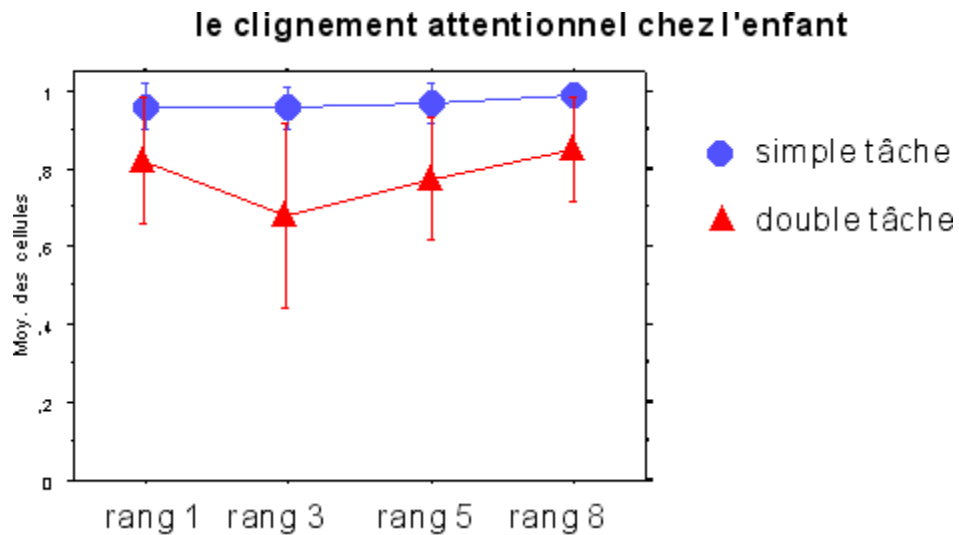






























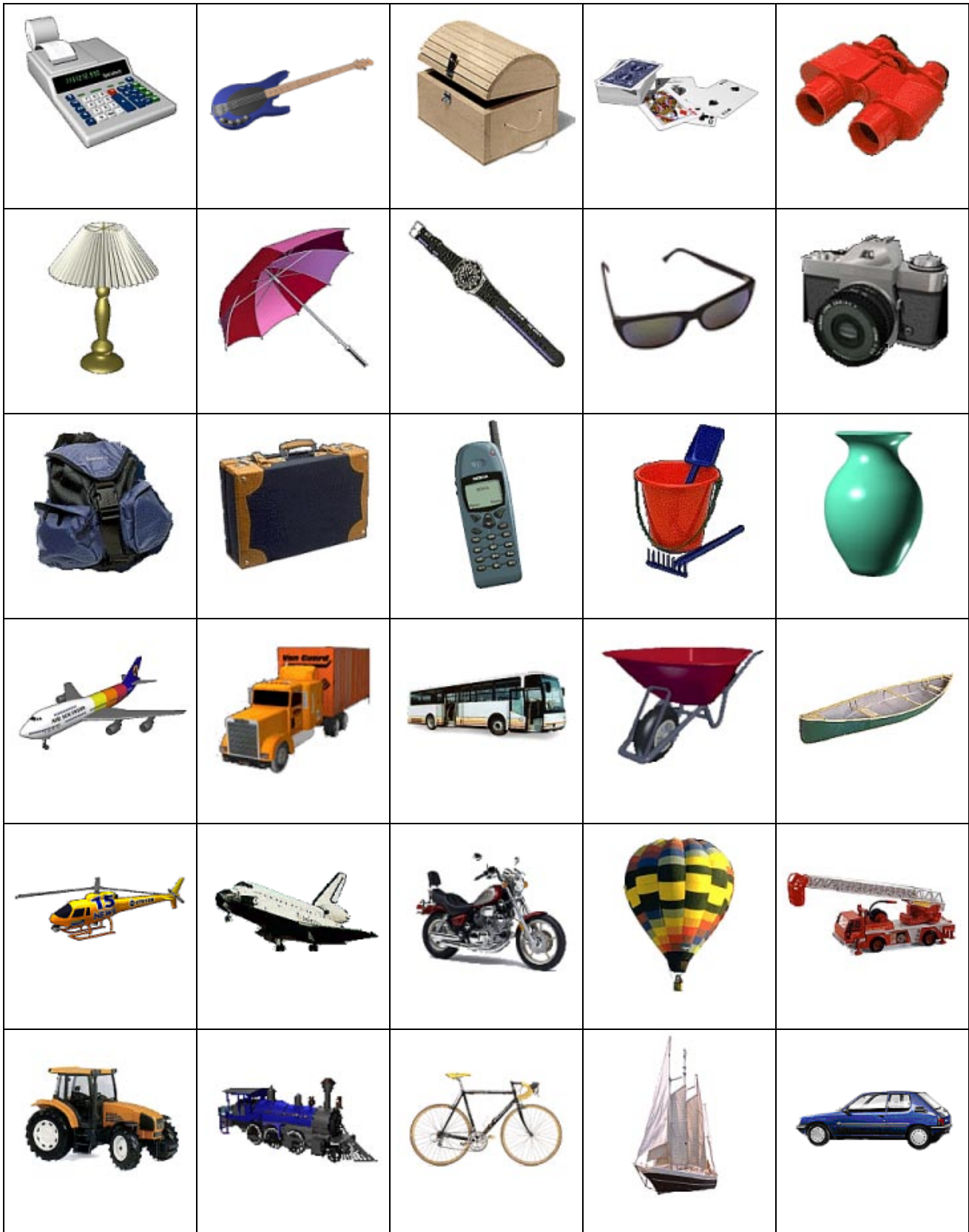


Figure XI.A : le clignement attentionnel chez l'enfant. En abscisses, se trouvent les positions de la cible à détecter T2 par rapport à la cible à identifier T1. En ordonnées, se trouvent les performances moyennes de détection de la cible T2 lorsque celle-ci est présente. En double tâche, la détection de la seconde cible est altérée durant environ 8 rangs après la présentation de la première cible à identifier. Ce déficit est absent lorsque le sujet réalise une simple tâche. Le déficit forme une courbe caractéristique en U.

Annexe XII : Stimuli utilisés en RSVP chez l'enfant



Annexe XIII : descripteurs utilisés pour la régression, expérience préliminaire A

Stratégie pour la représentation des qualités visuelles par des variables numériques

Il est difficile de trouver un modèle des primitives visuelles, (c'est-à-dire des composantes indépendantes du traitement visuel, également appelées caractères visuels, ou dimensions visuelles) :

- qui soit d'un unique niveau de description, par exemple :
 - niveau psychologique : rendu des opérations cognitives au niveau comportemental
 - niveau psychophysologique : opérations supportées par des aires neuronales
 - niveau neuronal : opérations sous-tendues par des neurones
- qui soit exhaustif (couvre toutes les dimensions analysées par le système visuel)
- qui soit intègre (sans recouvrement : une dimension n'est pas représentée de manière redondante par deux indicateurs distincts).

De même, il est délicat de réaliser la conversion d'un caractère visuel (eg. la composition chromatique d'une image) en données numériques (eg. un ou quelques chiffres résumant cette composition chromatique) (voir figure XIII.A). Plusieurs des critères envisagés sont fondamentalement multidimensionnels (au minimum, deux dimensions, ladite qualité visuelle, et la position de cette qualité dans le champ visuel). Notre projet implique que nous transformions ces caractères visuels en variables dépendantes. Pour les opérationnaliser, nous serons obligés de tolérer l'approximation. Ne perdons cependant pas de vue que la transformation « caractère visuel - données numérique » n'est pas une transformation anodine.

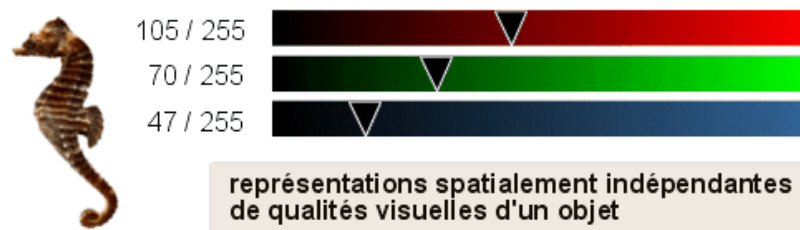


figure XIII.A : un mode de représentation des caractéristiques chromatiques de l'image de l'hippocampe (ici, le modèle RVB). Cette image comporte beaucoup de rouge, peu de vert et très peu de bleu. Les valeurs numériques associées à ces qualités sont moyennées à travers l'image entière, réalisant une représentation grossière.

Une stratégie d'extraction de ces données consiste à **moyenner le caractère à travers l'image entière**. Ces mesures ont été réalisées en utilisant les analyseurs des logiciels graphiques, et font donc abstraction du contexte physique dans lequel l'image est présentée (distorsions engendrées par le système d'affichage, par ses paramètres, par les conditions d'illumination de la pièce, par la position du mobilier engendrant des réflexions...). La qualité des mesures n'est donc pas maximale (par rapport à l'utilisation de moyens tels qu'un photomètre). L'équipement photométrique est un investissement trop coûteux par rapport au cadre institutionnel de cette recherche, aussi nous lui avons substitué

une stratégie de mesure moins perfectionnée. Compte tenu des conditions expérimentales (salle semi-obscur, lunettes de réalité virtuelles), le caractère ordinal des données obtenues peut cependant être défendu.

Cette méthode est relativement facile à mettre en oeuvre, mais présente une faible validité. Des objets sur-appris que nous utilisons dans nos protocoles sont très probablement identifiés sur la base de la résolution de quelques points discriminants, plutôt que sur la résolution de la totalité des caractères visuels (voir figure XIII.B). Ce traitement restrictif pourrait consister à ne traiter qu'une portion de l'image (par exemple l'œil très reconnaissable d'une grenouille), ou à ne traiter qu'un caractère (par exemple la couleur, si elle est suffisamment discriminative, faisant ainsi abstraction des autres caractères visuels de forme, de texture...).



Figure XIII.B : Des chunks dans le traitement de l'information visuelle ? La description objective de l'image nous dira sans doute que celle-ci contient beaucoup de vert. Imaginons alors que le traitement visuel du sujet se soit restreint à une partie significative, l'œil : ce n'est pas de vert dont il serait alors question, mais de rouge et de noir.

Le but final de cette opérationnalisation des caractères visuels est de mesurer leur pouvoir explicatif dans les erreurs d'identification commises par les sujets (« Est-ce la similarité de couleur, de forme, de texture, ... qui explique la confusion entre la cible et un distracteur ? »). Ces caractères visuels seront donc des variables indépendantes dans les modélisations statistiques que nous tenterons de mettre en oeuvre. Mais si la caractérisation de l'image entière est sans commune mesure avec les traitements que réalise effectivement le système visuel d'un sujet, nos hypothèses ne pourront pas être éprouvées ?

Si la restriction concerne le choix stratégique du traitement d'un caractère visuel, et le délaissement des autres, nos méthodes statistiques seront efficaces. En revanche, si la restriction concerne des portions de l'image, alors nous risquons de passer à côté des résultats attendus : dans l'exemple figure XIII.B, si un traitement restrictif est réalisé sur l'œil de l'animal, le critère de couleur pertinent est le rouge. Cependant, lors de la caractérisation de la couleur de l'image entière, cette portion rouge est noyée par la quantité de vert présente dans l'image. Dans cette logique, il est à craindre que des similarités effectives s'avèrent peu prédictives dans certains cas, et que dans d'autres cas, d'apparentes dissimilarités (globales) occasionnent des erreurs sans qu'il nous soit donné de comprendre pourquoi au travers des résultats.

Mise en oeuvre de la régression pour l'expérience préliminaire A.

La régression a été réalisée sur la catégorie sémantique bénéficiant du plus grand nombre d'intrusions, la catégorie fruits & légumes.

Variables utilisées

Moyenne de la distribution de luminance	Ecart-type de la distribution de luminance
Moyenne de la composante a du modèle Lab	Ecart-type de la composante a du modèle Lab
Moyenne de la composante b du modèle Lab	Ecart-type de la composante a du modèle Lab
Largeur de l'image	Hauteur de l'image
Ratio Largeur sur Hauteur	Surface de l'image
Périmètre de l'image	Compacité de l'image

Résultat de la régression

$R^2 = .216$ Erreur standard de l'estimation = 5,048

F= 4,523 ddl = 12 p<.001

Significativité des régresseurs

Moyenne de la composante a T=5,461 p<.001

Ecart-type de la composante a T=2,98 p=.003

Moyenne de la distribution de luminance T=2,2 p=.029

Autres variables non significatives

Annexe XIV : caractéristiques de l'échantillon adulte (expérience préliminaire B)

Les sujets étaient 30 adultes (19 femmes et 11 hommes) âgés de 18 à 54 ans (moyenne : 28 ans). L'appariement entre les vitesses de défilement des deux protocoles a été réalisé sur les performances en détection de la cible T2 (le ballon, voir table XIV.A) et sur les différences de score entre simple et double tâche.

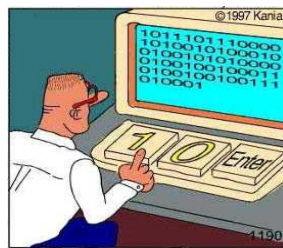
Pourcentage de réponses correctes	Identification fond bleu simple tâche (ST)	Identification fond bleu double tâche (DT)	Détection Ballon simple tâche (ST)	Détection Ballon double tâche (DT)
adultes	76%	68%	99%	86%
Différence ST-DT	8%		13%	
Enfants	48%	41%	98%	85%
Différence ST-DT	7%		13%	

Table XIV.A : variables d'ajustement des deux protocoles A et B : l'appariement des difficultés des deux tâches (vitesse de défilement) a été réalisée sur les performances en détection du ballon.

Annexe XV : Logistique de l'expérimentation en tachistoscopie

Le mode de programmation retenu pour les expérimentations 1 à 5 est un mode de programmation dit « de haut-niveau » (programmation intuitive, d'apprentissage rapide), sous le système MS-DOS qui garantit une bonne exécution des instructions de programmation, en particulier pour des affichages brefs de l'ordre de une ou quelques trames VGA. L'interface de programmation utilisée est ERTS de Beringer (<http://www.beringer.de>). Nous allons maintenant présenter les contraintes et les possibilités que nous donne ce cadre de programmation.

Systèmes d'exploitation et contrôle des ressources en programmation de haut-niveau



Le vrai programmeur ...

-sans légende-

Le langage natif de l'ordinateur est un langage binaire. Celui-ci s'exprime par des états 0 ou 1 et par des changements d'états, gérés par le processeur¹. Interagir avec un tel système implique soit de donner des instructions proches de ce niveau binaire (programmations/instructions de bas-niveau), soit de donner des instructions dans un langage plus intuitif, proche du langage parlé de l'utilisateur (haut-niveau). Dans ce deuxième cas, une/des couche(s) intermédiaire(s) de fonctionnement apparai(ssen)t : le programme interprète la requête de l'utilisateur. Il transforme les instructions intuitives de l'utilisateur en des commandes compréhensibles pour l'ordinateur.

C'est dans cet esprit d'ergonomie qu'ont été élaborés les systèmes d'exploitation (Windows, MS Dos, Linux, Mac OS, ...), interfaces intuitives entre l'ordinateur et l'utilisateur. Ils permettent par exemple d'explorer les supports de stockages (disques durs, disquettes, disques réseaux, CDrom...), de lancer les applications logicielles de manière simple, ils assurent de manière autonome la maintenance et l'organisation des différentes tâches concomitantes (affichage vidéo, veille antivirus, son, applications α , β , χ , gestionnaire d'accès aux disques, instructions de commande des périphériques d'impression, de scanner... et ceci n'est qu'un tout petit exemple de tout ce que supporte le système d'exploitation). Ces systèmes d'exploitation sont plus ou moins gourmands en ressources mémoire, et sont plus ou moins permissifs sur le « multi-tâche ». Ainsi, la consommation de ressources systèmes et les tâches de fond (opérations réalisées par le système en arrière plan) sont minimales dans le système MS-dos, et sont nombreuses et difficilement contrôlables² dans les systèmes d'exploitations standard (Windows & Mac-OS en tête). Dans la perspective

¹ (un regroupement de transistors ; si le courant passe à travers, il s'agit d'un un, si le courant ne passe pas, il s'agit d'un zéro)

² La plupart des systèmes actuels de programmation d'expérimentations sont défaillants pour les affichages brefs, en raison de l'utilisation sans contrôle de systèmes d'exploitation « gourmands ». Des développements résolvant ce problème existent cependant (eg. P. Desprez, Laboratoire de Neuropsychologie Expérimentale - Lille), apportant des solutions opérationnelles à ce problème.

d'un rigoureux contrôle des affichages brefs, nous avons donc choisi un système de programmation d'expérimentations fonctionnant sous le système d'exploitation MS-DOS.

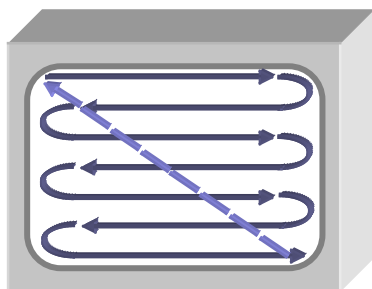
L'affichage

ERTS est doté de procédures de contrôle visant à vérifier avant la mise en oeuvre de l'expérimentation, que les impératifs de programmation inhérents à l'affichage (taille des objets en mémoire vive et temps de présentation) pourront être respectés lors des sessions expérimentales.

Trame VGA et fréquence de rafraîchissement

Sur un moniteur d'ordinateur comme sur une télévision, lorsque l'image change, les différents endroits de l'image ne sont pas modifiés en même temps. Il se réalise un balayage, qui « rafraîchit » les unités d'affichage (les pixels) de manière séquentielle¹ (voir figure XV.A, partie a).

On appelle trame l'opération de rafraîchissement de la totalité de l'écran, et également le temps nécessaire à cette opération. La durée de la trame dépend du paramétrage du moniteur de l'ordinateur. Le paramètre en question est la fréquence de rafraîchissement ou fréquence verticale (voir figure XV.A, partie b).



fréquence de rafraîchissement	Durée d'une trame
50 Hz	20 ms
60 Hz	16,67 ms
70 Hz	11,67 ms
85 Hz	11,67 ms

Figure XV.A : trame VGA. En a, rafraîchissement d'une matrice VGA : les pixels de chaque ligne sont modifiés de manière séquentielle. Lorsque le faisceau arrive au dernier pixel en bas à droite, il se déplace vers le premier pixel en haut à gauche et initie une nouvelle trame. En b, temps d'affichage des trames en fonction de la fréquence de rafraîchissement.

Notons que notre système visuel ne perçoit pas l'évolution de la trame VGA (il y a fusion de l'image à partir de 20 Hz, et ces phénomènes se déroulent entre 50 et 100 Hz).

¹ Cette séquentialité est obligatoire sur tous les systèmes d'affichage à tube cathodique ; elle est liée au déplacement du faisceau d'électrons. Les mêmes instructions sont utilisées pour les écrans TFT.

Initialisation d'un nouvel affichage

Deux types d'affichages sont possibles. En mode « trame-synchrone », l'initiation d'un nouvel affichage (le remplacement d'une image par une autre) se réalise seulement durant les transitions de trames (lorsque le faisceau quitte le dernier pixel en bas à droite pour rafraîchir le premier pixel en haut à gauche). Cela permet un meilleur contrôle des temps d'affichage des stimuli et une standardisation de la présentation.

En « trame-asynchrone », on observe des différences dans la mise en place de l'image à l'écran (voir figure XV.B). Les différences d'analyse perceptive induites par ce biais ne sont, à notre connaissance, pas connues. On peut toutefois s'interroger sur l'effet de la capture de l'attention des premiers signaux affichés. Si l'attention est capturée en bas de l'image, l'analyse perceptive pourrait ne pas être la même que lorsque l'attention est capturée en haut ? Par précaution, une standardisation des conditions d'affichage peut être recommandable.

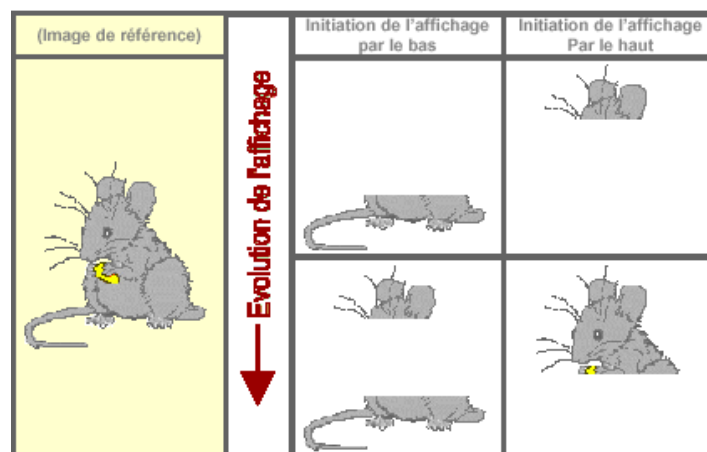


Figure XV.B : évolution de l'affichage d'une image en mode trame-asynchrone dans deux conditions : l'image s'initialise par la partie basse (colonne 3) ou par la partie haute (colonne 4).

En outre, le mode d'affichage asynchrone introduit une variabilité des temps d'affichage réels pour une petite image située au centre de l'écran. En effet, si les instructions vidéo initient un changement d'image lors de l'affichage du fond, il peut se passer plusieurs millisecondes avant que l'affichage réel des pixels de l'image ne démarre. En outre, ce délai varie systématiquement d'un essai à l'autre. Le tableau ci-après évalue cette variabilité des latences pour des images occupant respectivement 10 et 25% de la hauteur de l'écran, et affichées en 60 Hz. A contrario, pour ces deux mêmes images présentées en synchronisation sur la trame, les délais seraient respectivement de 7,5 millisecondes, et de 6,25 millisecondes, sans aucune variance.

Délai entre l'initialisation d'affichage et le début réel de changement des pixels (signal)	Petite image	Image moyenne
Délai moyen	6,8 ms	4,7 ms
Pourcentage en affichage immédiat	10%	25%
Délai maximal	15 ms	12,5 ms
Écart-type	4,71 ms	4,16 ms

Table XV.A : en trame asynchrone, estimation des délais pour l'affichage réel de deux images (petite et moyenne) après l'initialisation de l'affichage. Lors de certains essais, il peut se passer une quinzaine de millisecondes entre l'initialisation et l'apparition réelle des premiers pixels de l'image. Dans d'autres essais, l'initialisation des premiers pixels est immédiate. Lorsque l'expérimentateur prescrit un affichage de quelques dizaines de millisecondes seulement, cette variance est considérable.

Pour tous les affichages brefs (eg. inférieurs à 4 trames VGA), ce problème introduit une variabilité inter-essais considérable au niveau des temps de présentation des stimuli. Pour juguler ce problème, nous adopterons le mode « trame-synchrone ». Il nous contraint à utiliser des pas de présentation fixes (multiples de la durée de la trame, voir figure XV.A). Ainsi, nous n'avons pas toute latitude pour sélectionner les durées d'affichage des stimuli. En revanche, la standardisation des conditions d'affichage est garantie, et ce point nous a semblé crucial, eu égard aux contraintes engendrées par notre dispositif expérimental.

La mise en place d'un écran blanc (*ISI*, *inter-stimulus interval*) entre deux affichages successifs est également un point important du dispositif. Il évite qu'à certains moments dans l'évolution de l'affichage, il ne se produise une fusion physique des parties haute et basse des deux images respectives (voir figure XV.C). L'incidence de cet événement serait regrettable, car il justifierait la perception de conjonctions illusoirs, alors que le phénomène est sensé refléter une transformation cognitive.

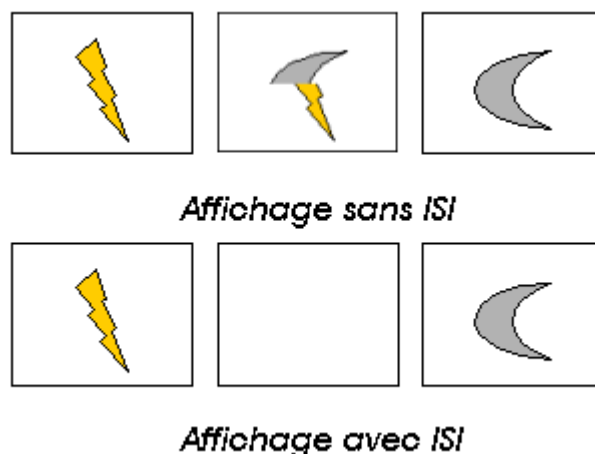


Figure XV.C : comparaison des affichages avec et sans ISI. Les trois cadres montrent les états successifs de l'affichage. Lorsque l'ISI est absent, la fusion des images se produit lorsque l'initialisation du nouvel affichage balaye l'ancien affichage. L'introduction d'un écran blanc entre les deux stimuli permet de juguler ce problème.

La mémoire et l'inscription des données sur le disque

Un ordinateur réalise usuellement de multiples tâches de manière simultanée. Des gestionnaires de priorité fonctionnant à différents niveaux (eg. IRQ : *Interrupt Request*) organisent les différentes tâches en fonction des ressources disponibles. Dans une session expérimentale tachistoscopique, le gravage des informations sur le disque








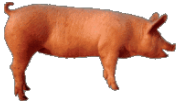













dur (réponses du sujet) peut compromettre la conduite des affichages brefs. ERTS ne réalise l'inscription des informations sur le disque qu'entre les essais, préservant ainsi toutes les ressources disponibles pour les opérations vidéo pendant la présentation des stimuli.

Possibilités de synchronisation de signaux extérieurs

Enfin, ERTS est doté d'interfaces permettant de coupler les évènements expérimentaux avec différents appareillages externes (IRM, EEG...). Cela permet de synchroniser les mesures comportementales supportées par le programme avec des mesures d'imagerie/électrophysiologie, et ce, avec une rigueur temporelle satisfaisante. Cette possibilité rend le dispositif expérimental viable pour des transpositions vers les protocoles mixtes envisagés dans la suite du projet (comportement+imagerie ; comportement+électrophysiologie).

Dotés de cette logistique, nous escomptons pouvoir mener des analyses sur la nature des intrusions (ces analyses sont absentes de la littérature, potentiellement parce que la variance d'erreur compromet la validité des résultats), et obtenir des résultats stables. Nous avons également fait évoluer le paradigme RSVP dans le but de le rendre plus adéquat au recueil des erreurs d'identifications.

Annexe XVI : Stimuli des expériences 1, 2, 4 et 5

Stimuli communs aux quatre expériences				
				
				
Stimuli additionnels pour les expériences 2 et 5				
				
				
Stimulus additionnel pour l'expérience 4				
				

Les assignations chromatiques ont été réalisées pour les stimuli des expériences 1, 2, 4 et 5. Le résultat est reporté à la figure XVI.A.

PERROQUET	CRABE	ESCARGOT	GRENOUILLE	POISSON	CHAT	PAPILLON	COCHON	OURS	SCORPION
SALOPETTE	CHAUSSETTES	ROBE	JUPE	TEE SHIRT	PULL	PANTALON	BONNET	ECHARPE	BLOUSON

Figure XVI.A : assignation chromatique des stimuli. Chaque catégorie sémantique comporte un stimulus rouge, orange, jaune, vert, bleu, indigo, violet, rose, grège et marron.

Annexe XVII : Propriétés des stimuli de substitution

Les deux stimuli 'scorpion' et 'zèbre' se comportent différemment quant-aux intrusions qu'ils subissent et quant-aux intrusions qu'ils causent à d'autres stimuli. Le scorpion est souvent intrusé (rang 1/10 des intrusions subies), et souvent intruseur (rang 2,5/10 des intrusions causées). Il est celui des stimuli qui occasionne le plus de réponses indéterminées. Le zèbre est très rarement intrusé (rang 10/10 des intrusions subies), et souvent intruseur (rang 2/10 des intrusions causées). Il occasionne un nombre intermédiaire de réponses indéterminées (rang 5/10).

Le profil des caractéristiques visuelles des stimuli est reporté à la figure @. Sur la plupart des critères analysés (hormis la taille pour les deux, et la compacité pour le scorpion), les stimuli de substitution occupent une position intermédiaire. Ils sont assez proches l'un de l'autre, mais également assez proche de la majorité des stimuli occupant ces positions centrales.

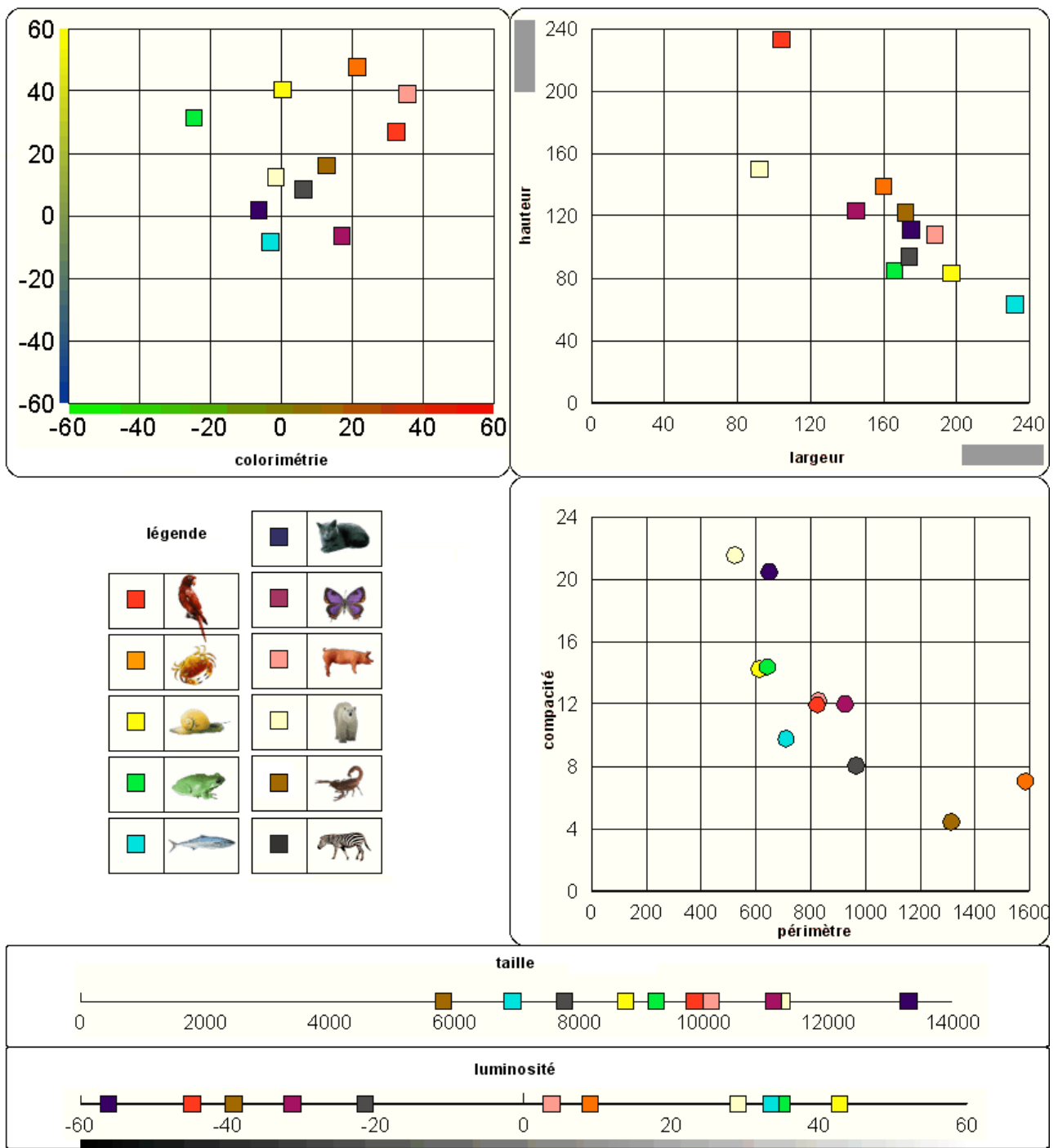


Figure @ : profil des caractéristiques visuelles des stimuli. En haut à gauche, le profil colorimétrique Lab (composante a en abscisses et composante b en ordonnées). En haut à droite, le ratio largeur sur hauteur. Au milieu à droite, les caractéristiques de périmètre et de compacité. En bas, les caractéristiques de taille (surface en pixel) et de luminosité (ces données sont dérivées des algorithmes de Corel PhotoPaint 9).

Résumé

Le liage perceptif (*binding problem*) est un problème important auquel se heurte le système cognitif lorsqu'il réalise des traitements. Il consiste à assembler les fragments d'informations sensorielles disparates en une représentation unifiée de l'objet. Plus de vingt ans après la formulation de ce problème par Anne Treisman, [Treisman & Gelade, *Cognitive Psychology*, 12 (1), 1980] peu de données sont disponibles pour comprendre comment le système cognitif résout ce problème.

Le *binding* par synchronie est une hypothèse neurophysiologique récente répondant au *binding problem*, qui propose que cet assemblage se réalise par coopération temporelle des activités neuronales. Les neurones participant à un traitement commun sont transitoirement impliqués dans une assemblée neuronale synchronisée, et émettent leurs décharges de manière coïncidente. Cette assemblée neuronale transitoire signe l'appartenance des traitements neuronaux disparates à une représentation commune.

Le but de cette thèse était d'élaborer un protocole expérimental dans le but d'étudier les rapports entre le liage perceptif (*binding problem*) et le *binding* par synchronie. A cette fin, le paradigme RSVP (*rapid serial visual presentation*) a été évalué pour sa capacité à répondre à ce projet. Ce paradigme est une tâche d'identification visuelle d'une cible parmi des distracteurs, mais a contrario des protocoles classiques, les distracteurs sont positionnés dans la dimension temporelle, et non dans la dimension spatiale. Des études récentes ont suggéré que des conjonctions illusoire figuraient parmi les erreurs observées à l'incidence de ce paradigme expérimental. Les conjonctions illusoire sont des erreurs résultant d'une intégration erronée entre une cible et un distracteur.

Nos expériences ont démontré que parmi les erreurs d'identification en RSVP, certaines relevaient effectivement d'une intégration perceptive erronée, mais ces erreurs se sont révélées être moins nombreuses qu'initialement envisagé. Nous avons localisé ces conjonctions illusoire dans la dimension temporelle. Nous avons sélectionné des configurations perceptives ayant une forte probabilité de donner lieu à ces conjonctions illusoire.

En conclusion de cette thèse, nous soutenons que le paradigme RSVP est un outil d'étude adéquat pour confronter le *binding problem* formulé par les Psychologues et le *binding* par synchronie formulé par les Neurophysiologistes.

