

Université de Lyon

Laboratoire d'Etude des Mécanismes Cognitifs (EMC), EA 3082

Ecole doctorale Neurosciences et Cognition

THESE

Présentée en vue de l'obtention du grade de

DOCTEUR DE L'UNIVERSITE DE LYON

Discipline : SCIENCES COGNITIVES

**LES INTERACTIONS ET INTEGRATIONS MULTI-SENSORIELLES SOUS
L'ANGLE D'UN MODELE DE MEMOIRE A TRACES MULTIPLES**

Par

Mathieu LESOURD

Soutenance prévue le 22 Septembre 2011

Sous la direction du professeur Rémy VERSACE

Composition du Jury

Denis Brouillet, Professeur à l'Université Paul-Valéry, Rapporteur

Micah M. Murray, Associate professor, Centre Hospitalier Universitaire Vaudois and
University of Lausanne, Rapporteur

Stéphane Rousset, Maître de conférence à l'Université Pierre Mendès-France, Examineur

Olivier Koenig, Professeur à l'Université de Lyon, Examineur

Rémy Versace, Professeur à l'Université de Lyon, Directeur

*« C'est souvent seulement par
manque d'esprit créateur
qu'on ne va pas assez loin
dans la souffrance. »*

Marcel Proust

A mes parents, Anne-Marie et Rémy Lesourd,

Et à ma grand mère

Remerciements

Je tiens à remercier le professeur Olivier Koenig qui m'a accueilli au sein du laboratoire EMC et qui m'a permis de réaliser ce travail de thèse dans les meilleures conditions possibles. Le laboratoire EMC a représenté durant ces quatre années, un bon cadre de travail. Son haut niveau de technicité et l'effort d'intégration des doctorants au sein de sa structure ne peuvent qu'aider au bon épanouissement professionnel.

Je remercie chaleureusement le professeur Rémy Versace qui a suivi mon travail depuis le Master 2 et qui a accepté d'encadrer ce travail de thèse. Je remercie sa patience face aux multiples activités que j'ai menées durant ma thèse. Ses conseils avisés et son haut niveau de compétence m'ont permis de grandir intellectuellement.

Je remercie sincèrement Denis Brouillet d'avoir accepté d'être rapporteur de mon travail de thèse. Je remercie également Micah M. Murray dont les travaux ont inspiré une grande partie de ma thèse et qui a eu la gentillesse d'être rapporteur de ce travail. Je remercie enfin Stéphane Rousset et Olivier Koenig qui ont accepté de faire partie de mon jury de thèse.

Je remercie les membres de mon équipe, Lionel Brunel, Guillaume Vallet et Benoit Riou qui ont toujours été d'un soutien indéfectible. Nos nombreuses interactions scientifiques ont toujours été riches et m'ont souvent redonné confiance en moi. J'espère pouvoir continuer à collaborer avec eux.

Je remercie mes collègues et amis du laboratoire, Mélaïne, Nina, Anna, Adrien, Arnaud, Hélène, Sarah-Lise, Emilie et Audric pour la bonne humeur et l'entente qui a régné durant ces dernières années.

Je remercie Julien Besle pour l'aide qu'il m'a apporté lors de l'analyse des données du Race Model. Je le remercie également de m'avoir donné les lignes de code de son programme, ce qui a rendu plus aisée ma tâche de programmation.

Je remercie également Rip Rockwell pour la finesse de ses conseils en analyses statistiques.

Je tiens à remercier aussi les titulaires du laboratoire Jordan Navarro, Stéphanie Mazza, Emanuelle Reynaud, Pascale Colliot et Georges Michael qui en plus d'être des chercheurs de qualité avérée, sont des personnes sincères, intègres et qui ont une

considération pour les doctorants et pour les étudiants. La recherche est un milieu difficile et pourtant, elle gagnerait à être plus représentée par ce type de personnes.

Je remercie également mes amis avec qui j'ai partagé des moments géniaux et parfois plus difficiles durant ces 4 années. Je me rends compte aujourd'hui que l'amitié est une véritable richesse. J'ai une pensée pour mes amis nancéiens : Djoul, Mary, Toufik, Fred, Pok, Jules & Julie, Romich, Nico, Lolo & Alban. J'ai une pensée pour mes amis parisiens : Flo & Aline, Oliv' & Pauline et mon ami Boulette. Enfin j'ai une pensée pour mes amis de Lyon : Céline, Laurence et Damien.

Je tiens à remercier aussi Elodie qui a, quoi qu'il en soit, dû supporter beaucoup de choses pendant ces quatre années.

Je tiens à remercier mes parents adorés qui ont toujours été pour moi d'un soutien sans limites. Je les remercie de tout mon cœur de m'avoir permis de m'épanouir où bon me semblait.

J'ai également une pensée particulière pour ma grand-mère qui ne comprenait pas vraiment ce que je faisais car elle pensait que je serai médecin après ma thèse. Quoi qu'il en soit, elle a toujours fait partie des gens qui ont cru en moi et cela n'a pas de prix. Elle me manque aujourd'hui.

Pour finir sur une note joyeuse, je pense aussi à une rencontre pleine de fraîcheur qui m'a rappelée que quand on est jeune, on est vite retapé.

Table des matières

LISTE DES FIGURES	1
LISTE DES TABLEAUX	5
RESUME	8
ABSTRACT	10
INTRODUCTION.....	12
1. CONSIDERATIONS THEORIQUES	15
1.1 Conception multi-système vs système unique de la mémoire.....	15
1.1.1 Différentes conceptions de la mémoire et de son organisation.....	15
1.1.1.1 Débat autour de la mémoire sémantique.....	15
1.1.1.2 Consensus autour de la mémoire épisodique	17
1.1.2 Que nous dit la neuropsychologie à propos du système sémantique?	19
1.1.2.1 La démence sémantique	19
1.1.2.2 Déficit modalité-spécifique et catégorie spécifique	21
1.2.3 Les limites de l'approche abstractionniste	23
1.2 Articulation d'un système conceptuel basé autour de représentations perceptives	23
1.2.1 Des représentations sensori-motrices dynamiques	24
1.2.2 Le modèle des systèmes symboliques perceptifs	26
1.2.2.1 Création et réactivation de symboles perceptifs	26
1.2.2.2 Nature des symboles perceptifs.....	27
1.2.2.3 Le simulateur à la base de l'activité mentale conceptuelle	28
1.2.2.4 Les zones de convergence.....	29
1.3 Les modèles de mémoire à traces multiples.....	30
1.3.1 Le concept de redintégration.....	30
1.3.2 Approche computationnelle des modèles de mémoire à traces multiples	31
1.3.3 Le modèle Act-in (Versace et al, 2002 ; Versace et al, 2009)	34
1.3.3.1 Cadre théorique.....	34
1.3.3.2 L'émergence des connaissances.....	35

1.4 Problématique générale	38
1.4.1. De l'interaction à l'intégration.....	39
1.4.2. Intégration perceptive et intégration mnésique	39
1.4.3. Influence d'un encodage multimodal sur la récupération à long terme	41
2. L'INTEGRATION MULTI-SENSORIELLE EN PERCEPTION	42
2.1 Introduction	42
2.1.1 Facilitation et conflit perceptif.....	42
2.1.2 Facteurs modulant l'intégration multi-sensorielle chez l'Homme	43
2.1.2.1 La contiguïté spatio-temporelle.....	43
2.1.2.2 La congruence	44
2.1.2.3 L'attention	45
2.2 Vers une conception interactive des systèmes sensoriels	46
2.2.1 Qu'en est-il de la conception modulaire des systèmes sensoriels ?.....	46
2.2.1.1 Les illusions perceptives	46
2.2.1.2 Argument neuropsychologique.....	50
2.2.1.3 Au-delà de la conception modulaire des systèmes sensoriels.....	50
2.2.2 Acquisition de nouvelles associations multi-sensorielles	51
2.2.2.1 L'apprentissage perceptif multi-sensoriel.....	51
2.2.2.2 Corrélats cérébraux des acquisitions de nouvelles associations multi-sensorielles	54
2.3 Les modèles de l'intégration multimodale	54
2.3.1 Modèles de compétition, modèle de coactivation et modèle de compétition interactif.....	54
2.3.2 Méthodes d'étude de l'intégration multi-sensorielle	56
2.3.2.1 Le Race Model	56
2.3.2.2 Mesures d'amplitude de la violation.....	60
2.3.3 Peut-on différencier intégration et interaction multimodales ?.....	61
2.3.4 Liens fonctionnels possibles entre interaction et intégration selon Act-in	64
2.4 Conclusion	65
2.5 Problématique	66
2.5.1 Etude de la création d'une trace multi-sensorielle.....	66
2.5.2 Intégration/interaction multi-sensorielle et Race Model	67
2.6 Partie expérimentale	68
Partie 1.....	68
Expérience 1	68
Expérience 2	72

Expérience 3	76
Conclusion	83
Partie 2.....	84
Expérience 1 Tâche de catégorisation multimodale.....	85
Expérience 2 Tâche de détection latéralisée sans congruence sémantique.....	89
Expérience 3 Tâche de détection latéralisée avec congruence sémantique.....	95
Expérience 4 Tâche de catégorisation latéralisée.....	100
Conclusion	105
CHAPITRE 3. L'INTEGRATION EN MEMOIRE A LONG TERME	110
3.1. Modèle de récupération en mémoire à long terme	110
3.1.1. Généralité sur les processus de recollection/familiarité.....	111
3.1.2. Un ou plusieurs processus pour rendre compte des performances en reconnaissance ?	111
3.1.3. Les méthodes de mesure	113
3.1.3.1. Les méthodes de dissociation de tâche.....	113
3.1.3.2. Les méthodes de dissociation de processus	114
3.2. Aspect fonctionnel et structurel de la récupération en MLT.....	115
3.2.1. Rôle des structures temporelles médianes	115
3.2.1.2. Le cortex rhinal	115
3.2.1.3. Lobe temporal médian et « binding »	116
3.2.2. Mécanismes oscillatoires et processus mnésiques	119
3.2.2.1. Introduction.....	119
3.2.2.2. Les oscillations gamma	120
3.2.2.3. Une vision interactive: l'hypothèse du « process binding ».....	122
3.3. Interaction multimodale en mémoire.....	123
3.3.1. Récupération liée au contexte multimodal.....	123
3.3.2. Effet d'un encodage multimodal sur la récupération	125
3.4 En résumé	131
3.5 Problématique	132
3.6 Partie expérimentale	135
Expérience 1	135
Expérience 2	141
Expérience 3	146
Expérience 4	150
Expérience 5	156

Expérience 6	166
Expérience 7	171
3.7 Synthèse	175
4. DISCUSSION GENERALE	179
4.1 Rappel de notre problématique	179
4.2 Rappel des principaux résultats expérimentaux	180
4.2.1 Expériences portant sur l'intégration multi-sensorielle en perception.....	180
4.2.2 Expériences portant sur intégration et récupération en mémoire à long terme	182
4.3 Implications théoriques.....	183
4.3.1 Multimodalité et conception distribuée de la mémoire	183
4.3.1.1 Spécificité de l'association multimodale	183
4.3.1.2 Réinterprétation des effets d'interférence sensorielle en mémoire	185
4.3.2 Intégration interaction et activation	188
4.4 Ouverture et perspectives.....	190
4.1 Des hypothèses supplémentaires.....	190
4.2 Mémoire et perception, vers une nouvelle conception.....	192
4.2.1 Une nouvelle représentation du lobe temporal médian	192
4.2.2 D'une vision top-down vers une vision horizontale des relations perception/mémoire	194
4.5 Quelques mots pour terminer	198
BIBLIOGRAPHIE	199
A. ANNEXES – MATERIEL EXPERIMENTAL	214
B. ANNEXES – DONNEES BRUTES.....	241

Liste des figures

FIG. 1 - Organisation de la mémoire sémantique pour les conceptions modales (A) et amodales (B).....	16
FIG. 2 – Les voies afférentes et efférentes du système de mémoire du lobe temporal interne, représentées ici sur le cerveau du singe, semblent déterminantes dans la transition de la perception à la mémoire. Pour que l'activité dans les aires TE et PG – qui est influencé par l'activité dans le cortex frontal (CF) – donne lieu à un souvenir à long terme, l'activité neuronale doit se produire, au moment de l'apprentissage, le long des projections de ces régions au lobe temporal interne – d'abord dans le cortex parahippocampique, le cortex péri-rhinal et le cortex entorhinal, et ensuite en plusieurs étapes, à travers l'hippocampe. L'ensemble des informations traitées sort finalement de ce circuit par le subiculum et le cortex entorhinal et retourne aux aires TE et PG. Adapté de Squire et Zola-Morgan (1991).	18
FIG. 3 - Schéma simplifié de la modélisation de l'organisation des connaissances sémantiques en mémoire, selon Farah & McClelland (1991).....	21
FIG. 4 - Création de symboles perceptifs, selon Barsalou (1999)	27
FIG. 5 - Illustration de la formation et de la réactivation d'un simulateur dans le modèle Perceptual Symbols System, PSS de Barsalou (1999)	28
FIG. 6 - Représentation schématique du modèle Minerva II.....	32
FIG. 7 - Structure connexionniste minimale implémentant une conception non abstractive de la mémoire, selon Rousset (2000)	34
FIG. 8 – Représentation schématique du modèle Act-in inspiré par le modèle MINERVA II (Hintzman, 1986).....	37
FIG. 9 - Illustration d'un cas de synesthésie. La perception de chaque caractère induit la sensation d'une couleur spécifique	49
FIG. 10 - Schémas des processus impliqués dans un apprentissage unisensoriel vs multi-sensoriel. Durant l'encodage unisensoriel (a), seules les structures visuelles seraient activées, alors que durant l'entraînement multi-sensoriel (b), une gamme plus importante de structures serait activée (b). Par conséquent l'apprentissage unisensoriel n'entraîne de modification que sur les représentations visuelles (c), alors qu'un apprentissage multi-sensoriel peut modifier les relations entre modalités à différents niveaux : au niveau des structures unisensorielles (d) ou des structures associatives (e), (f).....	52
FIG. 11 - Modèle de compétition et modèle de coactivation (D'après Mordkoff & Yantis, 1991)	55
FIG. 12 - Modèle de compétition interactif (D'après Mordkoff & Yantis, 1991).....	56

FIG. 13 - Illustration graphique de la falsification des modèles de compétition. De gauche à droite : la première courbe (rouge) correspondant à la fonction de répartition des TR bimodaux est graphiquement supérieure à la seconde courbe (noire) prédite par un modèle d'activations séparées. Les troisième et quatrième courbes correspondent aux fonctions de répartition des TR unimodaux. D'après Molholm et coll. (2002)	59
FIG. 14 – Formation d'un percept cohérent faisant appel à deux mécanismes : <i>combination</i> et <i>integration</i> pour lever l'ambiguïté sensorielle d'une situation donnée, selon Ernst & Bühlhoff (2004).....	62
FIG. 15 - Représentation graphique de l'interaction des facteurs CONTEXTE et BLOCS sur les temps de réponse (ms)	71
FIG. 16 – Représentation graphique de l'interaction entre les facteur BLOCS et ASSOCIATION sur les temps de réponse (ms).....	75
FIG. 17 – Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V) pour les sujets placés dans la condition « avec association » (à droite) et « sans association » (à gauche).....	80
FIG. 18 - Représentation graphique des courbes de violation (AV – (A+V)) des sujets placés dans la condition « sans association » et des sujets placés dans la condition « avec association ». Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un alpha < .05.	81
FIG. 19 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V) pour les sujets jeunes et les sujets âgés	87
FIG. 20 - Représentation graphique des courbes de violation (AV–A+V) des sujets jeunes et des sujets âgés. Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un alpha < .05.....	88
FIG. 21 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V) pour les sujets jeunes et les sujets âgés	92
FIG. 22 - Représentation graphique des courbes de violation (AV–A+V) des sujets jeunes et des sujets âgés. Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un alpha < .05.....	93
FIG. 23 - Courbes de violation des deux groupes recalées avec pour origine $p(\text{TR} < t) > 0$	93
FIG. 24 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V).	98

FIG. 25 - Représentation graphique des courbes de violation (AV–A+V) des conditions « détection sémantique » et « détection non sémantique ». Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un $\alpha < .05$	99
FIG. 26 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V).	103
FIG. 27 - Représentation graphique des courbes de violation (AV–A+V) des conditions « catégorisation latéralisée » et « catégorisation non latéralisée ».....	104
FIG. 28 - Représentation graphique des courbes de violation (AV–A+V) des quatre conditions « catégorisation latéralisée », « catégorisation non latéralisée », « détection sémantique » et « détection non sémantique ».	107
FIG. 29 - Représentation graphique des courbes de violation (AV–A+V) des conditions « détection » et « catégorisation ». Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un $\alpha < .05$	108
FIG. 30 – Représentation des opérations réalisées par le cortex rhinal, selon Fernández & Tendolkar, 2008	116
FIG. 31 – Représentation schématique du codage neuronal par deux évènements distincts mais similaires	118
FIG. 32 - Procédure utilisée dans l'expérience de Murray et collaborateurs (2004)	130
FIG. 33 - Représentation graphique de l'effet du facteur ITEM sur les temps de réponse (ms).....	140
FIG. 34 – Représentation graphique de l'interaction entre les facteurs ITEM et MODALITE sur les temps de réponse (ms)	144
FIG. 35 – Représentation graphique de l'effet du facteur ITEM sur les pourcentages de réponses correctes (%)	149
FIG. 36 – Représentation graphique de l'effet du facteur ITEM sur les temps de réponse (ms).....	154
FIG. 37 – Représentation schématique du paradigme expérimentale. La phase d'encodage et la phase de récupération étaient séparées par un délai de 5mn pendant lequel les sujets réalisaient un Sudoku.....	159
FIG. 38 – Représentation graphique de l'effet du facteur MODALITE sur les pourcentages de réponses correctes (%).....	161
FIG. 39 - Représentation graphique de l'effet du facteur MODALITE sur le d'	162
FIG. 40 – Représentation graphique de l'effet du facteur MODALITÉ en considérant séparément les pourcentages de réponses correctes pour les temps de réponses rapides et pour les temps de réponses lents	163

FIG. 41 - Représentation graphique de l'effet du facteur MODALITÉ sur le degré de certitude.....	164
FIG. 42 – Effet du facteur MODALITE sur la proportion totale d'items rappelés par condition.....	169
FIG. 43 - Représentation graphique de l'effet du facteur ITEM sur les temps de réponse.....	174
FIG. 44 – Représentation schématique de l'activation d'une trace mnésique dans un réseau neuronal suite à la présentation d'un stimulus visuel représentant un concept X et d'une stimulation tactile non signifiante. A -A gauche, la présentation des stimuli a entraîné l'activation de régions dédiées au traitement de ces modalités. B -A droite, l'activation automatique (en gris) des régions auditives suite à la stimulation visuelle.....	186
FIG. 45 - Représentation schématique de l'activation d'une trace mnésique dans un réseau neuronal suite à la présentation d'un stimulus visuel représentant le concept X et d'une stimulation auditive non signifiante. A - A gauche, la présentation des stimuli a entraîné l'activation de régions dédiées au traitement de ces modalités. B - A droite, l'activation automatique (entouré en rouge) des régions auditives suite à la stimulation visuelle est superposée à l'activation par le stimulus auditif non signifiant.....	187
FIG. 46 – Représentation du modèle d'organisations des représentations visuelles dans la voie visuelle ventrale (VVV) imaginé par Bussey et Saksida (2007). A – Les traits simples (A, B, C et D) sont représentés dans les régions caudales de la VVV alors que des représentations plus complexes sont stockées dans des régions plus rostrales. B – Diagramme du modèle connexionniste propose par Bussey et Saksida (2002).....	193

Liste des tableaux

TAB. 1 – Moyennes des temps de réponses et erreurs types en fonction des facteurs BLOC et CONTEXTE	70
TAB. 2 – Moyennes des temps de réponses et erreurs types en fonction des facteurs BLOC et ASSOCIATION	74
TAB. 3 – Répartition des items de l’expérience en fonction des facteurs ASSOCIATION et MODALITE et de la position d’apparition	78
TAB. 4 – Moyennes des temps de réponses et des pourcentages de réponses correctes et erreurs types en fonction des facteurs MODALITE et ASSOCIATION	79
TAB. 5 - Tableau de significativité des 6 premiers percentiles pour les groupes avec association systématique et sans association	81
TAB. 6 – Moyennes des temps de réponses et des pourcentages de réponses correctes et erreurs types en fonction des facteurs GROUPE et MODALITE	87
TAB. 7 - Test de Significativité des six premiers percentiles pour la distribution des sujets jeunes et celle des sujets âgés. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués italique	88
TAB. 8 – Moyennes des temps de réponses et pourcentages de réponses correctes et erreurs types en fonction des facteurs MODALITE et GROUPE	91
TAB. 9 - Test de Significativité des six premiers percentiles pour la distribution des sujets jeunes et du septième au douzième percentile pour la distribution des sujets âgés. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués en italique.	92
TAB. 10 – Moyennes des temps de réponses et pourcentages de réponses correctes et erreurs types en fonction des facteurs MODALITE et DETECTION	97
TAB. 11 - Test de Significativité des six premiers percentiles. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués en italique.	98
TAB. 12 – Moyennes des temps de réponses et pourcentages de réponses correctes et erreurs types en fonction du facteur MODALITE	103
TAB. 13 - Test de Significativité des six premiers percentiles. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués en italique.	104

TAB. 14 – Tableau récapitulatif des résultats obtenus sur le gain intégratif en fonction des expériences réalisées.....	106
TAB. 15 – Répartition des stimuli en fonction des facteurs GROUPE et ITEM	138
TAB. 16 – Moyenne des temps de réponses et erreurs types en fonction des facteurs GROUPE, ITEM et CATEGORIE.....	138
TAB. 17 – Moyennes des temps de réponses et des pourcentages de réponses correctes et erreurs types en fonction des facteurs ITEM, CATEGORIE et MODALITE	144
TAB. 18 – Répartition des stimuli de la phase d’encodage en fonction des facteurs CATEGORIE et MODALITE.....	146
TAB. 19 – Répartition des stimuli lors de la phase de récupération en fonction des facteurs CATEGORIE et ITEM	147
TAB. 20 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et ITEM	148
TAB. 21 – Moyennes des poucentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et ITEM.....	148
TAB. 22 – Répartition des stimuli dans les listes expérimentales en fonction des groups de sujet	152
TAB. 23 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et ITEM	153
TAB. 24 - Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et ITEM.....	153
TAB. 25 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et MODALITE.....	160
TAB. 26 - Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE	160
TAB. 27 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE	161
TAB. 28 – Moyennes des degrés de certitude et erreurs types en fonction des facteurs MODALITE et CATEGORIE.....	164
TAB. 29 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et MODALITE.....	168

TAB. 30 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE	168
TAB. 31 – Moyennes des pourcentages d’items rappelés et erreurs types en fonction des facteurs CATEGORIE et MODALITE	168
TAB. 32 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et MODALITE.....	172
TAB. 33 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE	173
TAB. 34 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et ITEM	173
TAB. 35 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et ITEM.....	173

Résumé

Notre travail s'inscrit dans une conception fonctionnaliste de la mémoire qui se donne pour objectif de décrire le fonctionnement de la mémoire plus que sa structure. Notre assise théorique est basée sur le modèle Act-in qui postule que la mémoire est représentée par un ensemble de traces multidimensionnelles. Ces traces comporteraient l'ensemble des dimensions d'une expérience (*e.g.*, motrices, sensorielles, émotionnelles). Pour rendre compte de l'émergence des connaissances, ce modèle repose sur deux mécanismes : l'activation et l'intégration. Notre travail avait pour objectif de spécifier les relations fonctionnelles entre activation et intégration à travers les mécanismes perceptifs et les mécanismes mnésiques. Nous avons également proposé un mécanisme d'interaction, permettant de rendre compte de la formation des traces sensorielles en mémoire. Comme le modèle Act-in conçoit les traces en mémoire comme étant multidimensionnelles, nous avons étudié les différentes relations entre mécanismes d'intégration, interaction et activation via le prisme de la multimodalité. Ce travail s'est organisé autour de deux grands axes expérimentaux, un premier tourné vers l'intégration en perception et l'autre vers l'intégration en mémoire.

Dans un premier axe, nous avons introduit la notion d'interaction multimodale pour rendre compte des relations entre modalités sensorielles dans des tâches perceptives. Nous avons montré que lorsque des interactions sensorielles étaient systématiquement répétées, les performances de sujets étaient meilleures que celles observées dans un contexte multimodal dans lequel les modalités sensorielles n'entretenaient pas de lien. Nous avons mesuré à l'aide d'un outil mathématique, le Race Model, le gain multi-sensoriel dans diverses conditions. Nous avons montré que le Race Model mesurait un gain multi-sensoriel essentiellement lorsque la tâche n'impliquait pas de processus mnésiques (*i.e.*, détection). Selon nous, cela montre que cet outil mathématique est plus sensible à l'interaction multimodale, lorsque les composants sensoriels ne sont pas intégrés en mémoire.

Dans un second axe, nous avons développé un paradigme expérimental permettant de mesurer les effets d'un encodage sensoriel sur une récupération ultérieure (*i.e.*, catégorisation, reconnaissance et rappel libre). Nous avons montré qu'une association multimodale congruente lors d'un encodage (*e.g.*, image de chien et aboiement du chien) améliorait significativement les performances dans diverses tâches de mémoire. Selon nous, l'activation

des dimensions auditives associées au sein de la trace explique cette facilitation lors de la récupération. Nous avons également montré qu'une association multimodale non congruente lors de l'encodage (*e.g.*, image d'un chien et bruit blanc) perturbait la récupération en mémoire à long terme. Ce résultat apporte un argument supplémentaire en faveur des modèles de mémoire qui postulent que les connaissances sont avant tout sensorielles.

Pour résumer, ce travail a permis dans un premier temps de proposer un mécanisme d'interaction multimodale nécessaire à l'intégration des composants sensoriels en mémoire. Dans un second temps, nous avons montré que des modalités sensorielles associées lors de l'encodage pouvait être réactivées ultérieurement et faciliter le traitement uniquement si les modalités sensorielles étaient sémantiquement congruentes. Ce résultat renforce la conception d'une mémoire organisée autour de connaissances avant tout sensorielles.

Enfin, les résultats que nous obtenons suggèrent que la mémoire et la perception partagent des traitements communs car un trait perceptivement absent peut influencer un trait perceptivement présent et réciproquement. Cela nous permet d'envisager une relation horizontale entre mémoire et perception et remet en question la conception traditionnelle *top-down* des liens entre mémoire et perception.

Mots clés : modèle de mémoire à trace multiple, intégration multi-sensorielle, Race Model, perception/mémoire

Abstract

Our conception of memory has the aim to describe the memory process more than the memory architecture. Our theoretical approach is based on the Act-in model, postulating that memory is represented by a multidimensional set of traces. These traces should include all dimensions of an experience (*e.g.*, sensory-motor and emotional components). This model relies on two mechanisms (*i.e.*, activation and integration) in order to reveal the emergence of knowledge. The aim of this work was to specify the functional relationships between activation and integration through memory and perception mechanisms. We also proposed a mechanism of interaction, to understand the formation of sensory memory traces. The model Act-in is composed by multidimensional memory traces. Thus, we studied the relationship between integration, interaction and activation through multimodality. This work was organized around two main axis. On one hand perceptual integration and on the other hand the memory integration.

In the first axis, we have introduced the concept of multimodal interaction to show the relationship between sensory modalities in perceptual tasks. Our results revealed that when sensory interactions were systematically repeated, the performance of subjects were better than those observed in a context in which multimodal sensory modalities did not maintain a link. We measured using a mathematical tool, the Race Model, multi-sensory gain under various conditions. We have shown that the Race Model measured multi-sensory gain, mainly when the task did not involve memory processes (*i.e.*, detection). This mathematical tool seems to be more sensitive to multimodal interaction, when the sensory components are not integrated into memory.

In the second axis, we used an experimental paradigm to measure the effects of sensory encoding on subsequent retrieval tasks (*i.e.*, categorization, recognition and free recall). Our results showed that a congruent multimodal association at encoding (*e.g.*, image of dog and the dog barking) improved performance in various memory tasks. It seems that the activation of the sensory dimensions could explain the facilitation during retrieval. Moreover, an incongruent multimodal association at encoding (*e.g.*, image of a dog and white noise) disturbed performance from retrieval in long-term memory. This result provides a further argument supporting sensory-based memory model.

In summary, we propose a multimodal interaction mechanism for the integration of sensory components in memory. Secondly, we have shown that sensory modalities associated during encoding could be reactivated later and facilitate the processing only when the sensory modalities were semantically congruent. This result improves the assumption that memory is sensory-based.

Finally, these results suggest that memory and perception share common features. The reactivation of a sensory dimension can influence the perceptual dimension processing and vice versa. This allows us to consider a horizontal relationship between memory and perception. In such case, the traditional relationship between memory and perception, as assumed to be top-down, could be questioned.

Key words: multitrace memory modelization, multisensory integration, Race Model, perception/memory

Introduction

La manière dont nous récupérons les souvenirs et les connaissances sur le monde qui nous entoure a beaucoup été étudiée dans le champ de la psychologie et notamment dans celui de la psychologie cognitive. De nombreux modèles ont ainsi tenté d'expliquer et de comprendre ces mécanismes de récupération depuis le milieu du XIX^{ème} siècle à l'aide d'une méthodologie expérimentale. Deux types de conception s'opposent dans la description des mécanismes mnésiques, la première, dite structuraliste, tente d'expliquer la mémoire de manière hiérarchique, à l'aide de systèmes et de sous-systèmes, en insistant sur la structure de la mémoire (*e.g.*, Tulving, 1991). L'autre conception, dite fonctionnaliste, envisage la mémoire comme un système unique et essaie d'en comprendre le fonctionnement (*e.g.*, Versace, Nevers & Padovan, 2002). Notre approche est issue de ce dernier type de conception, dans lequel les souvenirs sont modélisés sous forme de traces mnésiques comprenant toutes les dimensions de ce souvenir, appelé épisode. Une trace serait donc par définition multisensorielle car elle serait définie par l'ensemble des dimensions qui définissent un épisode. Concernant la récupération d'un épisode, l'accès à une des dimensions de la trace pourrait permettre de récupérer l'ensemble de la trace et par conséquent l'épisode. Déjà dans des réflexions anciennes, certains auteurs parlaient de « redintégration », c'est à dire la possibilité de récupérer l'ensemble à partir d'une des parties du tout (James, 1890). Par exemple, voir un citron peut activer automatiquement la composante gustative très acide, mais aussi la composante tactile avec la texture très particulière des agrumes, etc.

Dans ce travail de thèse, nous allons étudier les interactions entre les modalités sensorielles en mémoire. Nous essaierons de caractériser le lien qui existe entre les différentes modalités sensorielles en mémoire, de la création d'une trace jusqu'à la récupération de celle-ci. Nous nous intéresserons à la manière dont les composants sensoriels de la trace interagissent et sous quelles conditions ils peuvent faciliter ou perturber la récupération d'une trace en mémoire. Pour comprendre ces phénomènes, nous nous placerons dans une conception fonctionnaliste de la mémoire. En effet, ce type de modélisation répond à certaines contradictions que les modélisations structuralistes peinent à résoudre. Par exemple, selon la logique des doubles dissociations, le système cognitif pourrait être décomposé en autant de modules que l'on observerait de dissociations. Une architecture unique régit par des processus est plus à même

de rendre compte du fonctionnement cognitif. Nous tenterons de caractériser les interactions multi-sensorielles en mémoire à l'aide du modèle de mémoire à traces multiples Act-in (Versace *et al.*, 2002 ; Versace, Labeye, Badard & Rose, 2009), qui repose sur deux mécanismes, l'activation et l'intégration, pour rendre compte de l'émergence des connaissances et des souvenirs. Pour expliquer l'émergence d'une trace, Act-in propose une étape d'activation des composants sensoriels en lien avec un percept, puis l'intégration progressive des composants d'une trace pour reconstruire un souvenir par exemple. Notre travail aura pour but de spécifier les mécanismes d'activation et d'intégration qui sont définis selon le modèle Act-in, en s'appuyant sur l'étude du lien entre les modalités sensorielles.

Ce travail va s'articuler autour de quatre parties. Dans une partie d'introduction théorique, après avoir présenté succinctement le débat qui existe entre les conceptions structuralistes et fonctionnalistes de la mémoire, nous présenterons la notion d'intégration, et notamment la manière dont elle a été utilisée dans certains modèles, afin d'essayer d'en dégager des caractéristiques principales. Nous aborderons également le lien qui existe entre la notion d'activation et celle d'intégration tout au long de ce chapitre introductif. Enfin nous présenterons le modèle Act-in qui supporte ces deux notions d'activation et d'intégration. Nous développerons ensuite notre problématique en deux chapitres, l'un reposant sur l'intégration multi-sensorielle perceptive et l'autre sur l'intégration multi-sensorielle en mémoire à plus ou moins long terme. Dans ces deuxième et troisième chapitres, nous discuterons des mécanismes intégratifs perceptifs et mnésiques.

Le deuxième chapitre comporte une partie théorique et une partie expérimentale qui portent essentiellement sur l'intégration multi-sensorielle en perception. Nous essaierons de voir comment des modalités sensorielles peuvent se combiner, s'influencer lors de tâches perceptives et enfin comment elles peuvent s'intégrer au sein d'une trace mnésique. Dans cette partie, nous développerons la notion d'interaction multimodale et nous essaierons de spécifier les liens entre activation et intégration lors de la formation de traces sensorielles. L'interaction peut être définie comme l'influence d'une modalité sur une autre alors qu'elles restent toutes deux individualisées (Driver & Noesselt, 2008). A l'opposé, l'intégration nécessite que les modalités sensorielles soient associées au sein d'une trace. Ces deux conceptions étant proches, on peut se demander s'il existe un lien entre interaction et intégration. On peut aussi se demander s'il est possible de distinguer d'une part des interactions multi-sensorielles qui aboutiraient à la création d'une trace en mémoire et d'autre part des interactions qui n'aboutiraient pas à un tel résultat. Nous utiliserons également une

méthode mathématique, le Race Model (Miller, 1982) pour tenter de quantifier l'intégration multi-sensorielle associée à la création d'une trace mnésique.

Le troisième chapitre est construit de la même manière que le chapitre précédent avec une introduction théorique puis une partie expérimentale. Ici, nous allons nous intéresser à la manière dont les modalités sensorielles se combinent en mémoire à long terme et quel est l'effet de cette combinaison lors d'une phase de récupération ultérieure. Nous manipulerons la nature du lien entre modalités sensorielles pour des items ayant une représentation sémantique afin de voir comment des associations multimodales congruentes ou non peuvent influencer la récupération de la trace. Nous étudierons également ces effets multimodaux en fonction de la nature de l'accès en mémoire (implicite/explicite). Tandis que dans le chapitre précédent, il s'agissait d'étudier la formation d'une trace en mémoire, dans ce chapitre, il s'agit d'étudier la récupération d'une trace multimodale.

Dans une quatrième partie, nous discuterons des différents résultats obtenus dans les deux chapitres expérimentaux puis nous discuterons des implications théoriques en essayant de spécifier plus encore les liens entre les mécanismes d'activation et d'intégration du modèle Act-in. Enfin, nous nous interrogerons sur la nature des mécanismes intégratifs en perception et en mémoire, en d'autres termes, nous nous demanderons si les mécanismes intégratifs engagés lors des opérations perceptives et mnésiques sont similaires.

1. Considérations théoriques

1.1 Conception multi-système vs système unique de la mémoire

1.1.1 Différentes conceptions de la mémoire et de son organisation

Les théories non abstraites conçoivent la mémoire de manière unitaire, elles ne distinguent plus perception (ou traitement) et mémoire (Charnallet, Rousset, Carbonnel & Pellat, 1996), d'autre part, elles rejettent l'existence d'une mémoire sémantique vue comme une base permanente de connaissances qui auraient été abstraites de nos expériences antérieures (voir Versace *et al.*, 2002 ; Versace *et al.*, 2009). Les théories abstraites de la mémoire se donnent quant à elles pour objectif principal de décrire la structure, l'organisation de la mémoire (Tulving, 1972, 1991). Dans cette conception, les informations sont stockées dans des sous-systèmes de traitement distincts. Cette approche prend essence dans les nombreux travaux réalisés en psychologie cognitive au début des années 1980, mais également suite aux observations faites dans le domaine de la neuropsychologie. Ces deux disciplines ont conduit les théoriciens à ne plus concevoir le système mnésique de manière unitaire, mais à envisager un découpage plus fin de la mémoire.

1.1.1.1 Débat autour de la mémoire sémantique

Dans le champ d'étude de la mémoire, il existe un débat sur la nature des connaissances sur le monde et la manière dont celles-ci sont stockées dans le système sémantique. Le débat repose sur la nature des représentations mais aussi sur la manière dont celles-ci sont construites et récupérées en mémoire sémantique. La conception modale postule que les connaissances reposent sur des traits sensori-moteurs (Barsalou, 1999, 2003; Versace *et al.*, 2002) alors que la conception amodale postule l'existence de représentations qui par essence ne sont plus liées aux aspects sensoriels de l'expérience, ces connaissances seraient décontextualisées et invariantes. On pourrait résumer en disant que la controverse porte directement sur l'existence ou non d'un stock sémantique.

Pour certains auteurs, la mémoire sémantique est basée sur la formation de représentations amodales, l'existence d'un système cognitif autonome et indépendant permettrait de prendre en charge ces représentations. Le modèle SPI est un des modèles cognitifs de référence dans l'approche structuraliste de l'étude de la mémoire. Dans une version de ce modèle, Tulving (1995) répertorie cinq systèmes hiérarchisés: la mémoire procédurale, le système de

représentations perceptives, la mémoire à court terme, la mémoire sémantique et la mémoire épisodique. Selon l'auteur, cette hiérarchie correspondrait approximativement à l'ordre de développement phylogénétique présumé des systèmes. Dans cette conception du système mnésique, les connaissances sur le monde sont amodales (Rogers *et al.*, 2004 ; voir FIG. 1 - B), elles sont abstraites de leur contexte d'encodage. De plus, le système sémantique est largement individualisé et n'entretient pas de relation directe avec le système de représentations perceptives. Ce système sémantique amodal serait supporté par les lobes temporaux bilatéraux antérieurs (Lambon Ralph & Patterson, 2008) et l'altération de ces structures produirait une dégradation voire une perte des concepts. Cette dégradation conceptuelle s'expliquerait par la forte interconnexion existant entre les cortex associatifs modalités-spécifiques et le pôle temporal antérieur (Gloor, 1997).

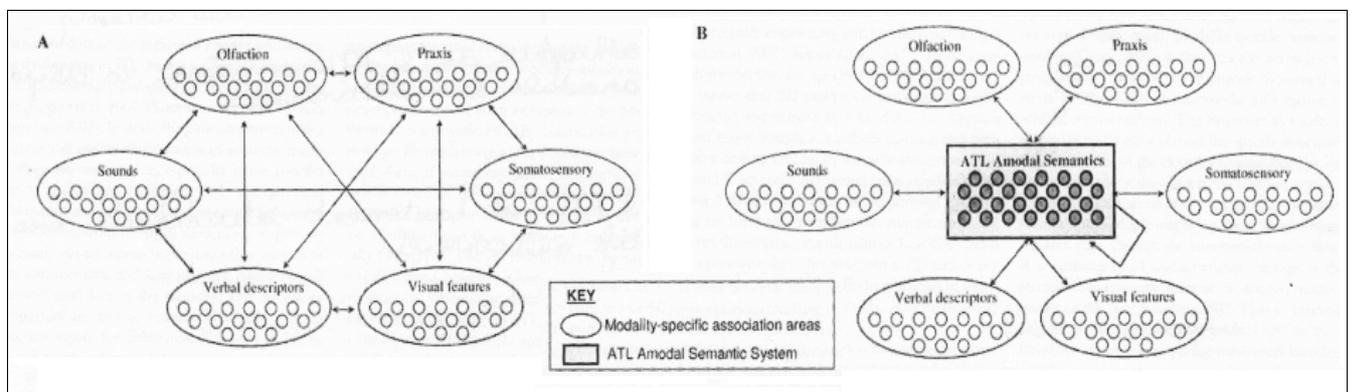


FIG. 1 - Organisation de la mémoire sémantique pour les conceptions modales (A) et amodales (B)

Dans la conception modale (voir FIG. 1–A), Wernicke et Meynert s'intéressaient déjà au dix-neuvième siècle, au processus qu'ils appelaient *conceptualisation*, c'est-à-dire la manière dont se formaient et se réactivaient les concepts (voir Eggert, 1977). Leur modèle de conceptualisation repose sur ces différentes affirmations : (a) les éléments fondateurs des concepts sont des engrammes modalités-spécifiques localisés dans les aires corticales responsables du domaine sensoriel, moteur ou verbal ; (b) ces engrammes modalités-spécifiques sont hautement interconnectés ; et (c) ce réseau de connexion est à la base de ce que les auteurs appellent *conceptualisation*. Un concept spécifique serait ainsi représenté par la coactivation de tous les engrammes associés.

S'il est présomptueux d'assimiler *conceptualisation* et *intégration*, il ne paraît pas aberrant de penser que récupérer un concept en mémoire, c'est réactiver des unités sensorielles ou symboliques selon la vision théorique que l'on adopte. Dans la vision amodale, l'intégration est portée par un système sémantique amodal (Lambon Ralph & Patterson, 2008; Rogers *et*

al., 2004), on pourrait parler de structure intégratrice qui aurait un corrélat anatomique. Dans la conception modale, l'intégration n'est pas une structure, le mécanisme intégratif serait plutôt le résultat d'une coactivation de structures sensorielles et motrices.

1.1.1.2 Consensus autour de la mémoire épisodique

Dans la conception non abstractionniste, la mémoire est conçue uniquement comme épisodique, comme un système qui conserve uniquement les traces des épisodes ou événements vécus par un individu (*e.g.*, Carbonnel, 2000; Rousset, 2000; Versace *et al.*, 2002). Un épisode est constitué de l'ensemble des stimulations vécues par le sujet lors d'une rencontre avec un référent quelconque et possède généralement plusieurs composantes (visuelle, auditive, motrice, olfactive, etc...) reflétant l'activité des différents canaux sensoriels mis en jeu au cours de l'interaction avec ce référent. L'évocation du sens d'un item consiste alors à recréer momentanément, sur la base de l'épisode actuel, une partie des épisodes antérieurs ayant mis en jeu cet item. Dans le modèle de Tulving, il existe une distinction entre stock épisodique et stock sémantique, ce qui représente déjà une première divergence avec les modèles non abstractifs de la mémoire. De plus, pour Tulving, le système sémantique est subordonné au système épisodique, ce qui signifie dans un modèle hiérarchique, que l'encodage de nouveaux souvenirs dépend de la qualité d'encodage dans le système sémantique. Ceci n'est vrai que pour l'encodage, car selon Tulving, la récupération est indépendante et autonome dans chaque système. Pour les modèles non abstractifs, ce serait plutôt l'inverse qui se produirait, c'est-à-dire que les nouvelles traces créées à partir d'un épisode seraient la base de la construction ou réémergence de nouvelles connaissances. C'est l'activation d'une multitude de traces concernant un référent qui devrait entraîner l'émergence de connaissances de type « sémantique » à propos de ce référent (Versace *et al.*; 2002). Le point de divergence ne repose donc pas sur la nature d'un épisode en tant que tel, cette conception étant aujourd'hui partagée mais c'est surtout l'existence d'un stock sémantique en opposition à un stock épisodique qui est sujette à débat.

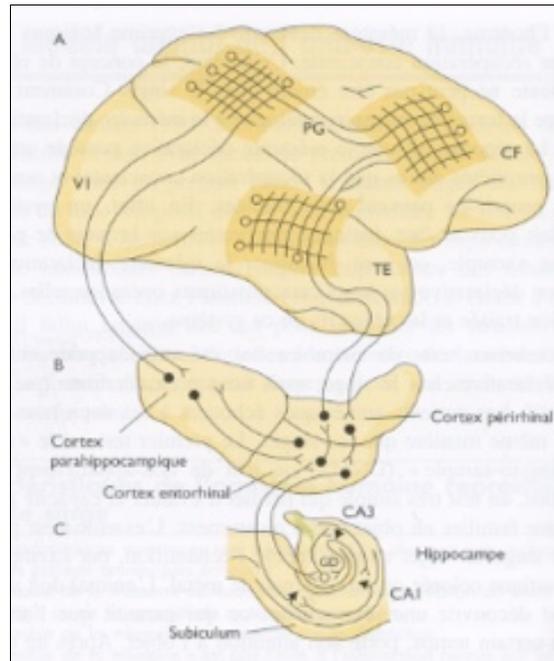


FIG. 2 – Les voies afférentes et efférentes du système de mémoire du lobe temporal interne, représentées ici sur le cerveau du singe, semblent déterminantes dans la transition de la perception à la mémoire. Pour que l'activité dans les aires TE et PG – qui est influencé par l'activité dans le cortex frontal (CF) – donne lieu à un souvenir à long terme, l'activité neuronale doit se produire, au moment de l'apprentissage, le long des projections de ces régions au lobe temporal interne – d'abord dans le cortex parahippocampique, le cortex péri-rhinal et le cortex entorhinal, et ensuite en plusieurs étapes, à travers l'hippocampe. L'ensemble des informations traitées sort finalement de ce circuit par le subiculum et le cortex entorhinal et retourne aux aires TE et PG. Adapté de Squire et Zola-Morgan (1991).

Il existe également des points de divergence concernant le modèle neuro-anatomique de consolidation en mémoire, toutefois le débat concerne uniquement les souvenirs de type épisodiques. Dans la conception classique de la consolidation en mémoire (Squire, Cohen, & Nadel, 1984; Squire & Zola-Morgan, 1991), un épisode est encodé dans des neurones néocorticaux et hippocampiques (voir FIG. 2). Dans le néocortex, les différents éléments (sons, odeurs, images) d'un épisode sont représentés physiquement dans des populations neuronales distinctes. Cependant, cette séparation physique crée un problème de « binding » au niveau de la récupération de cet épisode. C'est le complexe hippocampique qui jouerait un rôle critique dans la résolution de ce problème, c'est lui qui fournirait le mécanisme de liage entre les différents fragments néocorticaux. Dans cette conception, l'hippocampe permettrait des changements synaptiques rapides support d'apprentissages rapides, ce qui permettrait au néocortex de se charger des apprentissages plus lents. C'est essentiel que l'apprentissage soit lent dans le néocortex, car des modifications trop rapides dans les poids synaptiques pourraient avoir comme conséquence d'interférer avec des informations stockées précédemment. L'ajout d'informations ou l'apprentissage de nouvelles connaissances se ferait donc de manière progressive. Dans l'hippocampe, à l'inverse, les représentations peuvent être formées de telle manière qu'une interférence ne vienne pas détruire une donnée nouvellement

encodée. Les représentations seraient en fait encodées orthogonalement, dans des assemblées neuronales différentes. En d'autres termes, le recours à des populations neuronales distinctes permettrait d'éviter que des apprentissages trop proches dans le temps n'entraînent des oublis rapides des informations nouvellement acquises. Ce mécanisme a pour conséquence de limiter les interférences, et permettrait ainsi d'acquérir rapidement de nouvelles traces mnésiques sans perdre les anciennes, au moins à court terme.

Le modèle « Multiple Trace Theory of Human Memory » (Nadel, Samsonovich, Ryan & Moscovitch, 2000) tente d'apporter des indications plus fines concernant des aspects critiquables de la conception classique. Bien que ce modèle, dont la traduction pourrait être « Théorie des Traces Multiples » (TTM), fasse référence à la notion de traces multiples, il est important de remarquer que le modèle TTM distingue mémoire épisodique et mémoire sémantique contrairement à d'autres modèles de mémoire dit « à traces multiples » (Versace *et al.*, 2002; Versace *et al.*, 2009). Dans ce modèle alternatif, les auteurs critiquent principalement le fait que dans la conception classique, l'hippocampe ne soit plus impliqué dans la récupération des souvenirs, une fois ceux-ci stabilisés dans le néocortex. Selon les auteurs, le complexe hippocampique serait toujours impliqué dans le stockage et la récupération des souvenirs épisodiques. A l'inverse, les connaissances de type sémantique seraient stabilisées au sein du néocortex, comme le préconise la vision classique du modèle de consolidation de la mémoire, et n'auraient plus besoin de l'hippocampe pour être réactivées (Squire *et al.*, 1984 ; McClelland, McNaughton & O'Reilly, 1995). Ainsi les deux types de mémoire épisodique et sémantique seraient dépendants du complexe hippocampique mais un seul type deviendrait indépendant au cours du temps. De plus, les auteurs suggèrent que chaque réactivation d'un souvenir conduit à un nouvel encodage au sein de l'hippocampe (Nadel & Moscovitch, 1997, 1998 ; Moscovitch & Nadel, 1998). Ce nouvel encodage aurait pour conséquence la création d'une nouvelle trace mnésique, ce qui rendrait les traces beaucoup plus résistantes à un déficit touchant le système hippocampique.

1.1.2 Que nous dit la neuropsychologie à propos du système sémantique?

Autant l'existence d'une mémoire épisodique ne fait plus vraiment débat aujourd'hui, l'existence d'un stock sémantique est plus controversée. La neuropsychologie cognitive apporte des informations quant aux dissociations que l'on observe dans la pathologie et par la même, de rendre compte de l'existence ou non d'un système de mémoire sémantique

1.1.2.1 La démence sémantique

Pour nombre d'auteurs, la démence sémantique (DS) remet en cause le fait que la mémoire sémantique soit basée sur un système largement distribuée de représentations modalité-spécifique et que le lien entre ces modalités ne requiert pas la présence de représentations localisées dans une région particulière du cerveau. Deux caractéristiques importantes de la DS sont établies (Lambon Ralph & Patterson, 2008) : (a) le déficit sémantique est un phénomène graduel dans lequel les concepts et les frontières entre les concepts disparaissent progressivement (Patterson & Hodges, 2000) ; (b) la dégradation des concepts est amodale, le patient va être dans l'incapacité de produire ou comprendre des informations quel que soit le domaine verbal ou non verbal, incluant les mots, les images, les objets, les sons, les odeurs, le toucher, etc. (Rogers *et al.*, 2004). Ces caractéristiques font dire aux auteurs que la DS serait vraisemblablement liée à l'atrophie focale du lobe temporal antérieur, qui supporterait la formation des représentations amodales sémantiques. Carbonnel (2000) ajoute les caractéristiques suivantes : (c) les seules informations conceptuelles préservées sont déterminées par l'expérience personnelle récente des patients. Snowden Griffiths et Neary (1995) ont en effet montré que, chez des patients atteints d'une DS, l'expérience personnelle déterminait non seulement quels concepts étaient préservés et lesquels étaient perdus, mais aussi la nature des informations conceptuelles disponibles. Les auteurs soulignent que, chez ces patients, la compréhension des concepts n'est pas générale mais, au contraire, « limitée aux aspects du sens directement dérivés de l'expérience personnelle » et défendent l'idée selon laquelle ces données « suggèrent une relation beaucoup plus interactive entre le système de mémoire sémantique et l'expérience autobiographique » (Snowden *et al.*, 1995, p.241). (d) Les patients ayant une DS présentent un gradient atypique d'amnésie rétrograde. Celle-ci est totale pour la période la plus ancienne alors que les informations concernant les événements récents sont relativement préservées. Concernant le point (b), Moscovitch et Nadel (1999), sur la base d'une étude de cas, constate que même si la DS se traduit par une perte des connaissances conceptuelles tant dans la modalité verbale que dans la modalité visuelle, cette dégradation conceptuelle serait plus sévère dans le domaine verbal. Les aspects (c) et (d) ne sont pas non plus incompatibles avec une thèse non abstractive de la mémoire, au sens qu'ils peuvent tout fait s'expliquer par une atteinte des connaissances épisodiques. L'amnésie rétrograde massive pour les éléments passés pourrait traduire une dégradation des traces épisodiques de ces événements, conduisant ainsi au déficit sémantique sévère qui est observé. A l'inverse, la préservation des traces épisodiques relatives aux événements contemporains rendrait compte du maintien de certaines connaissances conceptuelles « dérivées de

l'expérience personnelle récente » comme cela est décrit dans l'étude de cas de Rousset et collaborateurs (1996).

1.1.2.2 Déficit modalité-spécifique et catégorie spécifique

Pour illustrer notre propos, nous prendrons comme exemple le déficit qui peut toucher spécifiquement la catégorie vivant/non vivant. L'observation de ce type de déficit peut nous inciter à interpréter l'organisation de la mémoire sémantique en catégorie taxonomique. Le modèle de Farah et McClelland (1991) est un modèle connexionniste de type abstraitif qui suppose l'existence d'un stock de représentations sémantiques et qui rend compte de l'existence d'une telle dissociation.

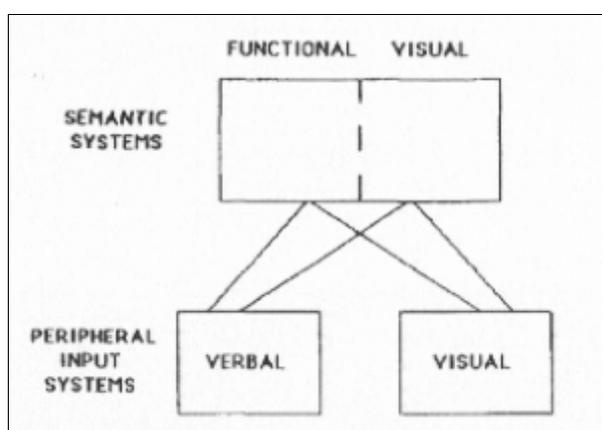


FIG. 3 - Schéma simplifié de la modélisation de l'organisation des connaissances sémantiques en mémoire, selon Farah & McClelland (1991)

Dans ce modèle (voir FIG. 3), la mémoire sémantique comporte deux types d'unités, visuelles et fonctionnelles qui contribuent différemment à définir les animaux et les objets (Warrington & McCarthy, 1987). Les animaux sont définis préférentiellement par les unités visuelles et par peu d'unités fonctionnelles, alors que les objets sont définis par un nombre équivalent d'unités visuelles et fonctionnelles. L'étude des troubles du patient EC (Charnallet & Carbonnel, 1995; Carbonnel, Charnallet, David & Pellat, 1997) mettent en évidence une agnosie visuelle d'objet massive sans déficit perceptif, une absence complète de connaissances concernant l'apparence visuelle des items concrets et un déficit de la compréhension des mots concrets, spécifique aux noms d'items vivants. Lorsque le patient est invité à fournir une définition de 20 stimuli (10 animaux et 10 objets) et à répondre à une liste de questions relatives à leurs aspects visuels et fonctionnels, les résultats montrent l'absence de connaissances visuelles sur les objets et les animaux, mais aussi une dissociation relative aux connaissances non visuelles, satisfaisantes pour les objets mais quasi inexistantes pour les

animaux. Carbonnel (2000) tente de montrer à partir de cette étude de cas, l'intérêt d'une modélisation non abstraite de la mémoire. Il note que dans le modèle de Farah et McClelland, une « lésion » du système sémantique visuel pourrait expliquer l'absence de connaissances visuelles concernant les animaux, il précise que les unités fonctionnelles ne pourraient pallier ce déficit, du fait de l'absence d'activations des unités visuelles. Toutefois, l'altération des unités visuelles dans le modèle de Farah et McClelland provoque un déficit catégorique spécifique sur entrée verbale et imagée, ce qui n'est pas le cas chez EC. Carbonnel précise que l'agnosie visuelle d'EC touchait tous les types d'items. A moins de postuler un autre déficit, ce qui n'est pas le cas d'EC, ce modèle abstraitif ne permet pas de rendre compte du tableau clinique de ce patient.

Il existe des travaux qui ont montré que la mémoire sémantique pouvait être organisée sur la base des modalités sensorielles (Martin, Haxby, Lalonde, Wiggs & Ungerleider, 1995; Martin, Wiggs, Ungerleider & Haxby, 1996). L'interaction avec un objet de notre environnement devrait donc entraîner une activité dans des régions cérébrales visuelles, mais aussi dans des régions (motrices) impliquées dans l'usage de ces objets. Alors que les animaux seraient définis uniquement par des traits visuels dans la plupart des épisodes les mettant en jeu. Sur cette base, Carbonnel postule qu'il est possible de rendre compte de l'ensemble du tableau d'EC par la présence d'un seul déficit : l'indisponibilité de la composante visuelle. Il ajoute que : (a) si l'on présente une image, celle-ci ne devrait s'apparier avec aucune trace épisodique, car privée de composante visuelle ; (b) si l'indice de rappel est un mot, les traces réactivées ne comporteront pas de composantes visuelles, aucune connaissance visuelle ne pourra donc être évoquée à partir d'un nom d'item ; (c) si l'indice de récupération est un nom d'objet, il devrait s'apparier avec des traces faisant référence à un vécu avec l'objet, c'est cette fois la composante fonctionnelle qui devrait être fortement activée. A partir de la démonstration de Carbonnel, nous sommes en mesure de montrer qu'un modèle non abstraitif de la mémoire peut être plus puissant qu'un modèle abstraitif pour rendre compte d'un ensemble de troubles neuropsychologiques.

Toutefois, si l'on suppose que l'organisation de la mémoire sémantique est basée sur les modalités sensorielles, on peut penser qu'une lésion dans une aire sensorielle devrait entraîner un déficit spécifique à cette modalité, alors qu'en réalité, un déficit dans une catégorie peut intervenir dans plusieurs modalités (Farah, Hammond, Mehta & Ratcliff, 1989). En fait une théorie reposant sur une organisation sensorielle de la mémoire sémantique est possible si l'on considère les représentations mentales distribuées et interactives (Thompson-Schill, Aguirre,

D'Esposito & Farah, 1999). Ainsi chaque composant d'un objet devrait permettre l'activation d'un autre composant de l'objet. Si l'un des composants de la représentation est endommagé suite à une lésion par exemple, il aura plus de difficulté à réactiver d'autres parties de la représentation. Pour reprendre l'exemple de la catégorie des animaux, si les traits visuels de la représentation sont endommagés, alors il sera plus difficile de réactiver les traits fonctionnels de la représentation. Cette hypothèse de « *interactive modality-specificity* » suppose donc que certains traits ont plus d'importance dans les représentations, et qu'une perturbation de ces traits ou composants suffirait à entraîner un déficit de la représentation associée, alors que les autres composants sensoriels sont intacts. Cette hypothèse permet entre autres d'expliquer l'interprétation de Carbonnel concernant le déficit d'activation des unités visuelles pour le vivant chez le patient EC. Ces unités visuelles ne seraient plus capables d'activer les unités fonctionnelles de la représentation, alors que ce sont ces unités visuelles qui sont majoritairement impliquées dans la récupération de la catégorie vivant. Les unités fonctionnelles n'auraient ainsi pas la capacité de s'auto-activer pour pallier la perturbation des unités visuelles.

1.2.3 Les limites de l'approche abstractionniste

La critique la plus fréquente adressée aux modèles abstractionnistes repose sur la tendance à augmenter le nombre de systèmes, a posteriori, pour rendre compte des nouvelles dissociations observées. Deux autres critiques peuvent aussi être évoquées : la première relève de la difficulté à décrire de manière cohérente les différentes formes de connaissances stockées en mémoire à long terme et associées à différentes modalités sensorielles, ainsi que la difficulté à décrire les nombreuses interactions et intégrations associées à ces différentes formes de connaissances. Les modèles multi-systèmes peinent à rendre compte de l'intégration des propriétés d'une expérience et de la nature des connaissances qui émerge de cette intégration. Le concept d'émergence n'existe pas dans les modèles abstractionnistes, le résultat de la récupération correspond à une connaissance en mémoire, l'activation est directe. La deuxième critique adressée tient au grand nombre de modélisations proposées pour décrire les connaissances stockées en mémoire à long terme (Versace *et al.*, 2009). Il semblerait plus judicieux de définir les représentations sur la base de l'architecture cérébrale et donc des contraintes imposées par celles-ci.

1.2 Articulation d'un système conceptuel basé autour de représentations perceptives

1.2.1 Des représentations sensori-motrices dynamiques

Des travaux réalisés en neuroimagerie ont permis de mettre en lumière les régions cérébrales impliquées lors du traitement conceptuel des catégories. Martin et collaborateurs (1996) ont montré que nommer des animaux et des objets activaient les lobes temporaux ventraux bilatéraux ainsi que l'aire de Broca, de plus ils ont mis en évidence que nommer des animaux activaient sélectivement le lobe occipital médian alors que nommer des outils activaient préférentiellement l'aire prémotrice gauche ainsi que le gyrus temporal médian. Les animaux activeraient préférentiellement les régions visuelles alors que les outils activeraient des régions impliquées dans la motricité, il est alors facile de faire l'analogie avec les travaux de Farah et McClelland (1991) qui définissent les animaux principalement par des unités visuelles et les objets par des unités fonctionnelles. Les régions cérébrales activées durant l'identification d'objets sont donc dépendantes, en partie, des propriétés intrinsèques des objets présentés (Martin *et al.*, 1995).

Des travaux en psychologie cognitive ont également apporté des arguments en faveur d'une représentation sensorielle des connaissances. Ainsi Zwaan, Stanfield et Yaxley (2002) ont demandé à des sujets de lire des phrases qui représentaient des animaux ou des objets dans certaines positions (par exemple un œuf entier ou un œuf au plat). Après chaque phrase, un dessin d'objet ou d'animal était présenté et les sujets devaient dire si oui ou non ils l'avaient vu précédemment. Les expériences menées par Zwaan et collaborateurs ont montré que les réponses étaient plus rapides lorsque la forme de l'objet évoquée dans la phrase était congruente avec celle présentée sur le dessin. Cette étude démontre que le traitement d'un matériel linguistique est accompagné d'une activation automatique de composants perceptifs. Brunel, Lesourd, Labeye et Versace (2010) ont réalisé une expérience en deux phases dans le but de démontrer le caractère sensoriel des connaissances. Lors d'une première phase de catégorisation de figures géométriques, chaque figure était présentée durant 500ms et pouvait ou non être associée à un son (*i.e.* bruit-blanc) de même durée. Dans une deuxième phase, les sujets devaient catégoriser des images d'objets cibles en fonction de leur taille typique (petit ou grand), les images cibles étant systématiquement précédés par une amorce (*i.e.* une figure géométrique issue de la première phase). Les auteurs avaient manipulé le caractère typiquement bruyant ou non des objets, c'est à dire, les objets qui étaient présentés avaient la particularité de produire du son (*e.g.*, sèche-cheveux) ou non (*e.g.*, lampe halogène). Le résultat princeps des auteurs est un effet d'amorçage engendré uniquement par les amorces « sonores » (*i.e.*, les formes géométriques qui ont été vues avec un bruit-blanc lors de la

première phase, voir aussi Brunel, Labeye, Lesourd & Versace, 2009) et limité à la catégorisation d'objets typiquement « sonores ». De plus, cet effet d'amorçage était modulé en fonction du SOA. Les auteurs ont observés un effet facilitateur pour un SOA de 500 ms, c'est à dire équivalent au temps de présentation de l'association figure géométrique/son lors de la première phase. A l'inverse, lorsque le SOA était de 100 ms, les auteurs ont mis en évidence un effet perturbateur. Selon les auteurs, alors que l'effet facilitateur devait être due à une préactivation par l'amorce de la modalité auditive partagée par l'amorce et la cible (voir aussi Pecher, Zeelenberg & Barsalou, 2004), l'effet d'interférence pourrait être le résultat d'un recouvrement entre la réactivation du bruit blanc associé à l'amorce et le traitement de l'image à catégoriser qui elle-même implique un composant auditif. Les résultats de cette expérience sont difficiles à expliquer dans le cadre de l'approche structurale, car ils montrent clairement que l'existence de propriétés sensorielles auditives peut influencer la catégorisation des connaissances. Ces arguments remettent en cause la nature amodale des connaissances.

D'autres travaux ont révélé l'importance des représentations motrices même dans les activités ne nécessitant pas explicitement l'utilisation de ces représentations motrices (Tucker & Ellis, 1998; van den Bergh & Vrana, 1990). Par exemple, Tucker et Ellis (1998) ont tenté de démontrer que la vision d'un objet devrait activer automatiquement des représentations motrices cohérentes avec la position, la forme, la taille et l'orientation de l'objet. Pour cela, ils ont présenté des objets pouvant être orientés correctement ou inversés, avec l'orientation permettant la prise de l'objet soit avec la main gauche ou avec la main droite. Les sujets devaient indiquer le plus rapidement possible si les objets étaient orientés correctement ou inversés. Les auteurs ont observé que les réponses étaient plus rapides lorsque la main utilisée pour donner la réponse était la même que celle dont la représentation devant être activée en premier pour saisir l'objet.

Ces différents arguments mettent ainsi en évidence l'aspect sensoriel des connaissances mais aussi l'aspect moteur de celles-ci. Toutefois si l'on accepte que les expériences qui sous tendent la construction de traces mnésiques sont dynamiques car elles représentent l'aspect interactif que le sujet entretient avec son environnement, alors l'aspect dynamique de ces expériences doit être encodé au niveau des traces mnésiques. Cette hypothèse a été testée par Zwaan, Madden, Yaxley et Aveyard (2004), pour ce faire, les auteurs ont utilisé un paradigme impliquant la présentation auditive d'une phrase dans laquelle était décrit le mouvement d'un objet se rapprochant du sujet ou au contraire s'en éloignant. Après une période de 750ms,

deux images séparées par un masque de 175ms étaient présentées au sujet. Les deux images représentaient deux objets identiques ou différents, lorsque les objets étaient différents, le second était soit plus petit soit plus grand que le premier. Les sujets devaient indiquer si les deux objets étaient identiques ou différents. Les résultats montrent un effet de congruence entre le mouvement apparent suscité par les deux images et celui indiqué dans la phrase.

1.2.2 Le modèle des systèmes symboliques perceptifs

Le modèle de Lawrence Barsalou propose un cadre théorique pour expliquer le système conceptuel humain (Barsalou, 1999, 2003 ; Barsalou, Simmons, Barbey, & Wilson, 2003).

1.2.2.1 Création et réactivation de symboles perceptifs

Lorsque Barsalou décrit son modèle, il propose dans un premier temps une description de l'activité neuronale qui a lieu pendant l'observation d'un objet : la confrontation avec cet objet déclenche une activation neuronale dans les aires sensorielles. Par exemple en vision, des neurones répondent aux bordures, d'autres aux couleurs, aux orientations, aux directions des mouvements, etc. L'ensemble de ces activations distribuées dans les aires sensorielles visuelles correspondent à une représentation visuelle de l'objet, mais il existerait également d'autres formes de représentations rattachées aux autres modalités (représentations auditives, olfactives, motrices, etc.). En même temps que les représentations sensorielles s'activent, des neurones dans des aires associatives capturent les configurations des patterns neuronaux correspondant, à la fois au niveau intra-modal et inter-modal. Par la suite, les liaisons réciproques entre zones sensorielles et zones associatives permettent des phénomènes de réactivations sensorielles : en l'absence de l'objet lui-même, une réactivation partielle des aires sensorielles est possible par l'intermédiaire de ces zones associatives (voir FIG. 4).

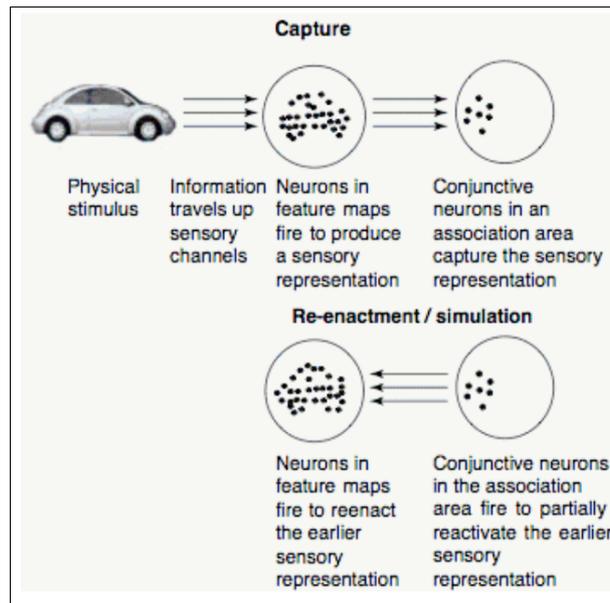


FIG. 4 - Création de symboles perceptifs, selon Barsalou (1999)

La représentation mentale créée à partir de la perception d'un objet correspond à l'activation d'un pattern de neurones primaires particulier, enregistré par les zones associatives (Conjunctive neurons). Les liaisons réciproques entre ces zones et les aires sensorielles permettent par la suite la réactivation du pattern sensoriel en l'absence de l'objet lui-même.

1.2.2.2 Nature des symboles perceptifs

Le processus de formation des symboles perceptifs peut s'opérer sur n'importe quel aspect d'une expérience. Il s'opère en vision, mais il peut aussi s'opérer sur d'autres modalités sensorielles (audition, toucher, olfaction, goût). Dans toutes ces modalités, l'attention va se focaliser sur certains aspects de l'expérience et va les stocker en mémoire à long terme, ces aspects vont ensuite devenir des symboles. Ainsi de nombreux symboles vont être stockés. Par exemple, au niveau de l'audition, nous allons acquérir des symboles concernant le langage, ainsi que les nombreux sons que nous entendons chaque jour. Au niveau du toucher, nous allons acquérir des symboles perceptifs sur les textures et les températures. Chaque type de symboles serait stocké dans les régions cérébrales correspondantes, par exemple, les symboles visuels seraient établis dans les régions cérébrales visuelles. De nombreux arguments ont ainsi établi que la perturbation d'une région sensori-motrice entraîne la perturbation du traitement conceptuel des catégories qui s'appuient sur ces régions lors de l'expérience perceptive.

Pour Barsalou, les symboles perceptifs sont donc multimodaux, ils sont directement liés à l'expérience perceptive et ils sont largement distribués sur les aires cérébrales modalité-spécifique.

1.2.2.3 Le simulateur à la base de l'activité mentale conceptuelle

À partir de ces considérations, Barsalou propose la notion de simulateur qui permet d'expliquer la formation de connaissances plus conceptuelles c'est-à-dire plus générales, à partir d'un ensemble de situations particulières. Pour lui, le fait de rencontrer plusieurs exemplaires d'une même catégorie (d'un concept) va activer des patterns neuronaux primaires très similaires, se chevauchant en plusieurs endroits. Ces renforcements neuronaux vont permettre par la suite la construction de représentations à modalité spécifique (symboles perceptifs) et de représentations multimodales du concept, impliquant à la fois les zones primaires et les zones associatives. Les simulateurs seraient alors des représentations mentales multimodales correspondant à tout ce qui est commun entre les exemplaires d'une même catégorie (voir FIG. 5).

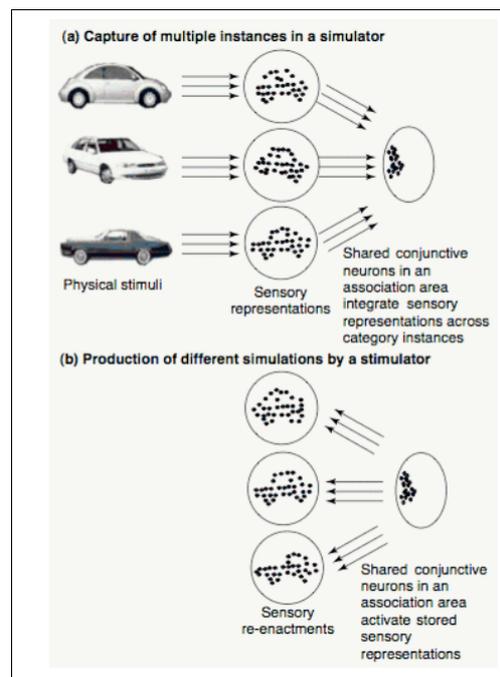


FIG. 5 - Illustration de la formation et de la réactivation d'un simulateur dans le modèle Perceptual Symbols System, PSS de Barsalou (1999)

Dans ce modèle, l'importance du simulateur n'est pas moindre. Pour Barsalou il est à l'origine des simulations mentales spécifiques que l'individu peut faire à propos d'une catégorie générale. Par exemple en fonction du contexte, la simulation sera le résultat d'une

réactivation exhaustive ou partielle des neurones correspondant au simulateur. Un exemplaire pourra émerger mentalement ou bien une « moyenne » des différents exemplaires.

1.2.2.4 Les zones de convergence

Le modèle de Barsalou offre donc aux aires associatives un rôle particulièrement important dans l'émergence d'une représentation mentale. Pour Simmons et Barsalou (2003), les zones associatives appelées zones de convergence, sont hiérarchiquement organisées : les zones de convergence intra-modales enregistrent les patterns des neurones primaires activés en présence d'un objet, tandis que les zones de convergence intermodales enregistrent l'activité simultanée de l'ensemble de ces zones intra-modales. L'importance que confère Barsalou aux zones de convergence n'est pas sans rappeler le modèle « Timelocked multiregional retroactivation » (Damasio, 1989). Dans ce modèle neuro-anatomique de 1989, Damasio fournit un cadre théorique pour la compréhension des processus intégratifs. Selon Damasio, l'architecture cognitif/neuronal implique que les représentations soient situées et fortement distribuées. La représentation d'entités abstraites et d'évènements dérive tous deux de représentations concrètes. Damasio ajoute qu'une zone de convergence existe en tant que pattern synaptique dans une assemblée de neurone dans les cortex associatifs et doit satisfaire les conditions suivantes : (1) ces zones reçoivent des projections de multiples régions corticales ; (2) en réponse à des informations de type feed-forward, elles peuvent renvoyer des informations de type feed-back ; (3) elles possèdent des connexions inhibitrices de type feed-forward/feed-back avec d'autres structures corticales et sous corticales. Une zone de convergence est amodale par définition, en ce sens qu'elle reçoit des signaux d'une ou plusieurs modalités sensorielles mais elle ne préserve pas les relations topographiques ou topologiques entre les traits telles qu'elles étaient dans l'expérience psychologique. Le rôle d'une zone de convergence est de permettre la reconstitution d'une représentation momentanée des entités ou évènements dans les cortex sensoriels et moteurs. Damasio distingue deux types de zones de convergence. Les zones de convergence de type I inscrivent les coïncidences temporelles et visent à les reproduire. Les zones de convergence de type II réagissent aux séquences et produisent des activations ordonnancées dans les cortex cibles. Les zones de convergence de type I se situeraient dans les cortex sensoriels associatifs et seraient assistées dans l'apprentissage par le système hippocampique. Les zones de convergence de type II sont liées aux cortex moteurs et sont assistées dans l'apprentissage par les ganglions de la base et le cervelet. Dans des conditions normales, les deux types de zone de convergence interagissent à différents niveaux, c'est pourquoi l'apprentissage relatif à une

entité ou un évènement recrute les deux types de zone. Le rappel ou la reconnaissance provoquent des opérations dans les deux types de zones de convergence, mais le stimulus déclencheur active seulement un type de zone de convergence au début du processus.

Barsalou se propose donc à travers son modèle PSS d'expliquer l'émergence des concepts à l'aide d'un mécanisme distribué qui pourrait trouver son corrélat anatomique dans un vaste ensemble d'interactions neuronales comme les a décrites Damasio. Nous allons maintenant aborder les modèles de mémoire à traces multiples qui tentent d'expliquer l'émergence des connaissances sémantiques mais aussi des souvenirs épisodiques à partir de la notion de traces.

1.3 Les modèles de mémoire à traces multiples

Les modèles de mémoire à trace multiples donnent une grande importance à la notion de traces. Ces modèles postulent l'existence d'une seule mémoire à long-terme conservant une trace de toutes les expériences d'un individu. La trace se caractérise par deux propriétés essentielles : la trace est épisodique car elle code les caractéristiques de l'épisode de traitement dans laquelle elle s'est constituée ; elle est multidimensionnelle car elle reflète les multiples propriétés des épisodes. La mémoire à long-terme est constituée par l'accumulation de ces traces.

1.3.1 Le concept de redintégration

Le terme « redintégration » que l'on trouve dans les écrits de William James (1890) et dans ceux d'Hamilton (1859), est utilisé pour caractériser une forme de récupération en mémoire. Si l'on considère une trace mnésique, c'est la capacité d'un des éléments de la trace de pouvoir réactiver l'ensemble de celle-ci. A partir de la présentation d'un seul des éléments de la trace, la trace peut être réactivée et récupérée. Horowitz et Prytulak (1969) ont essayé de caractériser plus précisément ce qu'ils appellent « redintegrative memory », par ce terme, on comprend que « redintégration » est un type de mémoire. Pour les auteurs, dans une tâche de mémoire quelconque, le stimulus proposé peut être relativement distinct de la réponse demandée au sujet, par exemple une tâche d'appariement de mots et de chiffres. Par contre dans d'autres tâches, le stimulus fait partie de la réponse attendue, il fait partie intégrante de la réponse. Ce type de mémoire a été nommé « redintegrative » (Hamilton, 1859). Pour caractériser une situation « redintegrative », ils supposent qu'un sujet S étudie une situation AB et qu'ensuite, on propose à S le stimulus A. Si le sujet répond B, il existe deux

interprétations possibles, soit la partie A a permis de récupérer la partie B, soit la partie A a permis de récupérer l'ensemble AB et ensuite la partie B. Est-ce que pour autant n'importe quelle partie A pourrait permettre de récupérer la partie B, cela ne semble pas si évident, en effet il semble que les noms soient mieux rappelés que les adjectifs dans des tâches de rappel libre et qu'ils représentent de meilleurs indices dans des tâches de rappel indicé (Lockhart, 1969). Il existerait ainsi une sorte de pondération en fonction de la nature du stimulus considéré, Horowitz et Prytulak appelle cette pondération « puissance ».

Les auteurs ont également utilisé un calcul basé sur les probabilités conditionnelles afin de dire si oui ou non une tâche donnée pouvait être considérée comme « redintegrative » en fixant une valeur arbitraire. Supposons qu'un sujet S voit une liste de phrases et qu'ensuite on lui administre une tâche de rappel libre. Notons W_i , chaque élément de la liste composé des parties A_i et B_i . $p(A_i)$ sera la probabilité que le sujet S rappelle A lors du rappel libre, et $p(B_i)$ la probabilité de rappeler la partie B. Enfin, $p(W_i)$ représentera la probabilité que le sujet rappelle l'ensemble. La probabilité conditionnelle que S rappelle la phrase entière, étant donné qu'il ait donné une partie de l'ensemble, s'écrira $p(W_i|A_i)$. Pour se situer dans une condition « redintegrative », la probabilité conditionnelle $p(W_i|A_i)$ doit être supérieure à .60 selon Horowitz et Prytulak. Dans leur article de 1969, les auteurs vont ainsi distinguer des tâches à caractère « redintegrative » et d'autres qui ne le sont pas. Même si la manière de distinguer ces deux types de tâches semble arbitraire, la façon de formaliser le concept de « redintegration » est intéressante et nous devrions peut être considérer qu'une situation est plus ou moins « redintegrative », plutôt que « redintegrative » ou non.

Ce concept plutôt ancien est intéressant car la philosophie des modèles de mémoire à traces multiples semble en être directement inspirée. En effet selon le principe de « redintegration » un composant de la trace est capable de réactiver le tout ou une partie du tout, ce qui correspondrait à un des postulats de ce type de modèle. En revanche, le mécanisme de « redintegration » est considéré comme une forme de récupération en mémoire, ce qui signifie qu'il existerait d'autres processus mnésiques à la base de l'émergence des souvenirs et des connaissances alors que dans les modélisations à traces multiples, ce mécanisme est central.

1.3.2 Approche computationnelle des modèles de mémoire à traces multiples

Dans la modélisation MINERVA 2 (Hintzman, 1986), chaque expérience, ou chaque traitement de l'information est représenté dans une trace par une configuration de ce que Hintzman appelle des propriétés primitives. Ces propriétés primitives peuvent être des

tonalités émotionnelles ou des traits sensoriels ou bien des propriétés plus abstraites. Ce modèle a pour objectif de donner un cadre au mécanisme d'émergence des connaissances à partir de l'activation de multiples traces. Hintzman fait l'hypothèse que chaque expérience produit une trace indépendante en mémoire, il fait également l'hypothèse que chaque trait d'une trace est encodé indépendamment avec une probabilité qui s'élève avec le temps de présentation ou avec le nombre de répétitions de l'information. La mémoire est ainsi représentée par une matrice dans laquelle chaque ligne correspond à une trace et chaque colonne à une primitive (voir FIG. 6). A chaque cellule est attribuée une valeur, -1, 0 ou +1, signifiant que le trait est absent, non pertinent ou présent.

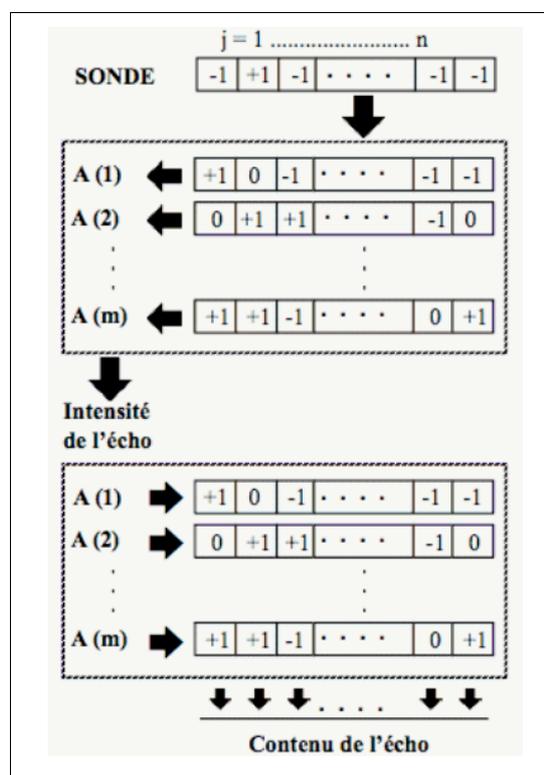


FIG. 6 - Représentation schématique du modèle Minerva II

La récupération ou émergence des connaissances se fait à partir de l'activation de traces par la représentation de l'épisode actuel de traitement, qu'Hintzman appelle une « sonde ». Cette émergence se fait en deux étapes : premièrement, chaque trace est activée en fonction de sa similarité avec la sonde ; deuxièmement, les traces activées renvoient en retour un vecteur « écho » reflétant la similarité ou synergie (Rousset, 2000) entre la sonde et la mémoire. L'écho est lui-même représenté sous la forme d'un vecteur dont chaque cellule est égale à la somme des valeurs des traces pour la colonne correspondante pondérée par l'activation de chaque trace. Cette propriété de l'écho permet ainsi au système de retrouver ou de compléter

une trace à partir d'un élément de l'épisode, ce qui correspond à une situation de rappel. Elle permet aussi de rendre compte des effets de contexte sur la récupération d'épisodes, si l'on admet que le contexte est codé dans la trace. De plus, l'intensité de l'écho résultant de la somme des valeurs d'activation de toutes les traces peut être considéré comme un indicateur de la familiarité et servir de base à la prise de décision dans une tâche de reconnaissance. De plus ce modèle peut rendre compte, sans postuler l'existence de plusieurs stocks de mémoire, d'un mécanisme d'abstraction. A partir d'un ensemble d'épisodes, il est possible d'évoquer des caractéristiques générales d'un objet sans faire référence explicitement aux spécificités de chaque épisode comme on peut le retrouver à un niveau cognitif dans la modélisation PSS de Barsalou.

Bien que ce modèle rende compte de nombreux phénomènes mnésiques dans une architecture non abstractive, il est critiquable sur sa simplicité, et notamment sur le grand nombre de calcul devant s'effectuer sur un nombre de traces a priori illimitées. Une proposition de modèle connexionniste minimaliste (voir FIG. 7) a été faite par Rousset (2000), ce modèle s'inspire de MINERVA 2, mais aussi du modèle neuro-anatomique de Damasio (1999) que nous avons présenté précédemment. Cette modélisation a l'avantage par rapport à la précédente de donner aux propriétés primitives, une véritable tonalité sensorielle, mais a également l'avantage de s'abstraire des critiques faites à MINERVA 2. Chaque ellipse représente une assemblée neuronale et les flèches représentent des faisceaux de connexions qui relient ces assemblées de neurones. A chaque présentation d'un stimulus multimodal, le système modifie progressivement les valeurs des connexions de manière à répliquer la configuration des stimulations sur les échos. Si le stimulus présenté est déjà connu du système, il sera directement recréé en écho, du fait des adaptations précédentes. Dans le cas contraire, de nouvelles adaptations seront nécessaires pour obtenir un écho égal à l'entrée. Rousset précise que la rapidité d'obtention d'un écho égal à l'entrée pourrait donc contribuer à un sentiment de familiarité. La re-création d'éléments contextuels antérieurement associés au stimulus présenté correspondrait à l'évocation précise de l'épisode d'encodage. Enfin la réinjection des échos en entrée (flèche pointillée) permet de simuler une réelle dynamique des processus mnésiques, en permettant un fonctionnement du système à partir de ses propres évocations.

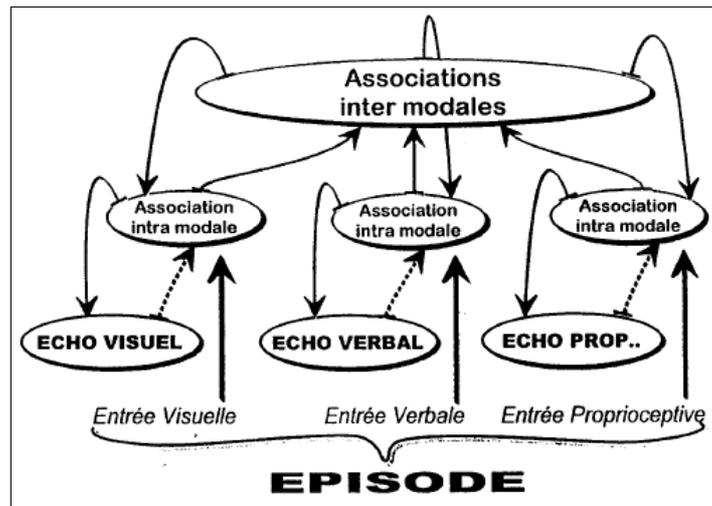


FIG. 7 - Structure connexionniste minimale implémentant une conception non abstraactive de la mémoire, selon Rousset (2000)

1.3.3 Le modèle Act-in (Versace et al., 2002 ; Versace et al., 2009)

1.3.3.1 Cadre théorique

Nous allons maintenant présenter Act-in, un modèle de mémoire à traces multiples. Versace et collaborateurs (2002) donnent un ensemble de caractéristiques pour spécifier l'architecture de ce modèle, bon nombre d'entre eux permettent d'ailleurs de répondre aux limites des modèles abstraectifs de la mémoire. Dans ce modèle, les traces mnésiques stockées en mémoire ne sont pas localisées, ni indépendantes les unes de autres, mais distribuées sur un ensemble de composants. Chaque module code de manière distribuée les différents composants élémentaires, ou dimensions, des expériences associées à chaque trace. Ces dimensions sont essentiellement sensorielles, motrices et émotionnelles, on retrouve ici l'importance accordée aux aspects sensoriels de l'expérience. Les traits codés au niveau de chacune des dimensions ne sont pas des invariants, mais dépendent de nos expériences passées. L'intégration des dimensions est nécessaire pour la constitution de traces unifiées, bien qu'une trace puisse se constituer au niveau d'une seule dimension. La trace d'une expérience serait le résultat d'une synchronisation d'activations au sein des différents modules. Une trace n'est pas conçue comme une copie conforme d'une expérience passée et les connaissances émergent des états d'activation du système. De plus une connaissance de type souvenir correspond à un état d'activation très proche d'un état antérieur spécifique, alors qu'une connaissance catégorielle ou sémantique est issue de multiples traces antérieures. Cette caractéristique permet entre autres de ne pas avoir à postuler l'existence de deux systèmes, un système épisodique et un

système sémantique. Nous allons détailler maintenant les mécanismes impliqués dans l'émergence des différents types de connaissance.

1.3.3.2 L'émergence des connaissances

1.3.3.2.1 Emergence comme résultante de l'interaction de multiples activations

L'encodage des différents composants d'une expérience au sein d'une trace requiert la participation de nombreux systèmes distribués sur de multiples régions cérébrales. Ce sont ces mêmes régions qui sont à la base du traitement des différentes modalités sensorielles. Par conséquent, toute forme de connaissance émerge de l'activation de ces différents systèmes, qui résulte de l'interaction entre les individus et leur environnement. Cependant, il n'existe pas de description précise dans la littérature des mécanismes impliqués dans l'émergence des connaissances. Pour rendre compte d'une telle émergence, le modèle Act-in propose l'existence de deux mécanismes, un mécanisme d'activation et un mécanisme d'intégration. L'exposition à un environnement va se traduire par l'activation en parallèle de multiples structures qui encodent les dimensions sensorielles. Par exemple, voir un réveil va activer très rapidement les aires visuelles correspondantes. Ces activations initiales se propagent ensuite très rapidement vers d'autres composants sensoriels et moteurs (activations inter- et intra-modales). Cette propagation d'activation doit plutôt être conçue comme se déroulant en cascade que de manière séquentielle. De plus les activations multimodales peuvent interagir très tôt, via de nombreuses connexions parallèles et réciproques entre les systèmes neuronaux sensorimoteurs. Il est clair que les activités produites dans ces aires doivent être intégrées ou synchronisées à une certaine étape du traitement pour permettre l'émergence de connaissances cohérentes et de comportements adaptés. Les activations multimodales initiales, qui sont spécifiques aux propriétés élémentaires des objets, sont nécessairement intégrées progressivement, permettant ainsi l'accès à des connaissances de plus en plus élaborées.

Une expérience menée par Labeye, Oker, Badard et Versace (2008) permet d'illustrer ces mécanismes d'activation/intégration. Les auteurs ont utilisé un paradigme d'amorçage à court terme en lien avec une tâche de catégorisation sémantique incluant des images d'objets. La cible représentait un objet appartenant à la catégorie des outils ou à celle des ustensiles de cuisine et était directement précédée par une amorce représentant l'image d'un objet. L'objet utilisé en amorce, lorsqu'il était utilisé, pouvait nécessiter un geste très similaire ou un geste complètement différent de celui suscité par la cible. L'amorce et la cible pouvait appartenir toutes deux ou non à la même catégorie sémantique. Pour un SOA de 100ms entre l'amorce et

la cible, les auteurs ont observé un effet facilitateur de la similarité gestuelle et de la catégorie entre l'amorce et la cible. Au contraire, pour un SOA de 300ms, les auteurs ont observé une interaction due aux effets facilitateurs de la similarité gestuelle uniquement lorsque la cible et l'amorce appartenaient à la même catégorie. En considérant ce résultat en terme de dynamique d'émergence des connaissances, une plus longue période d'exposition de l'amorce a certainement rendu possible l'intégration des différents composants de l'amorce, et par la même occasion empêchée l'expression indépendante de chacun de ces composants pris individuellement.

1.3.3.2.2. Activations intra- et inter-traces

Nous avons vu précédemment que l'émergence d'une connaissance correspondait à un état du système mnésique et que plus il y avait de traces activées moins l'information disponible serait spécifique. En effet, si le nombre de traces réactivées est essentiel pour différencier entre l'émergence d'une connaissance spécifique ou catégorielle, ces différentes formes de connaissances diffèrent également au niveau des mécanismes impliqués. Pour illustrer les mécanismes qui seraient impliqués dans la construction des traces et l'émergence des connaissances, nous pouvons utiliser la représentation matricielle de Minerva II (voir FIG. 8). Chaque ligne représente une trace et chaque colonne un composant (visuel, auditif,...) de l'expérience, encodé au sein de la trace. Dans une situation donnée, la nature de la connaissance susceptible d'émerger dépend de deux mécanismes, un mécanisme d'activation intra-trace et un mécanisme d'activation inter-trace. L'activation inter-trace fait référence à la diffusion d'activation entre les différentes traces (verticalement dans la matrice) sur la base de l'expérience actuelle. Par exemple, voir une image de chien entraîne l'activation des traces mnésiques correspondant aux situations dans lesquelles l'individu a été confronté à stimulus visuel similaire (par exemple, d'autres chiens). Cependant la vue de ce chien doit aussi activer, au sein de chaque trace, toutes les autres propriétés associées aux dimensions visuelles déjà activées, comme des propriétés tactiles, olfactives, auditives, etc. (*e.g.*, aboiement, l'odeur du chien, son pelage). Cette activation, qui se produit entre les dimensions au sein de chacune des traces correspond à l'activation intra-trace.

Les activations inter-traces et intra-traces permettent aux individus de produire des comportements appropriés et efficaces dans les situations dans lesquelles ils se trouvent. La plupart du temps, les connaissances et les comportements adaptés à une situation sont ceux qui ont été préalablement associés à de nombreuses situations antérieures similaires à la

situation présente. L'important n'est pas la spécificité de la situation présente, mais au contraire sa similarité avec d'autres situations antérieures. Ainsi, lorsque l'on se trouve dans une situation similaire à de nombreuses autres situations antérieurement vécues, l'importante diffusion d'activation inter-traces permet l'émergence automatique et rapide de connaissances de type catégorielle. La diffusion d'activation inter-traces semble donc être la règle générale de fonctionnement associé au couplage entre expérience présente et expérience passée.

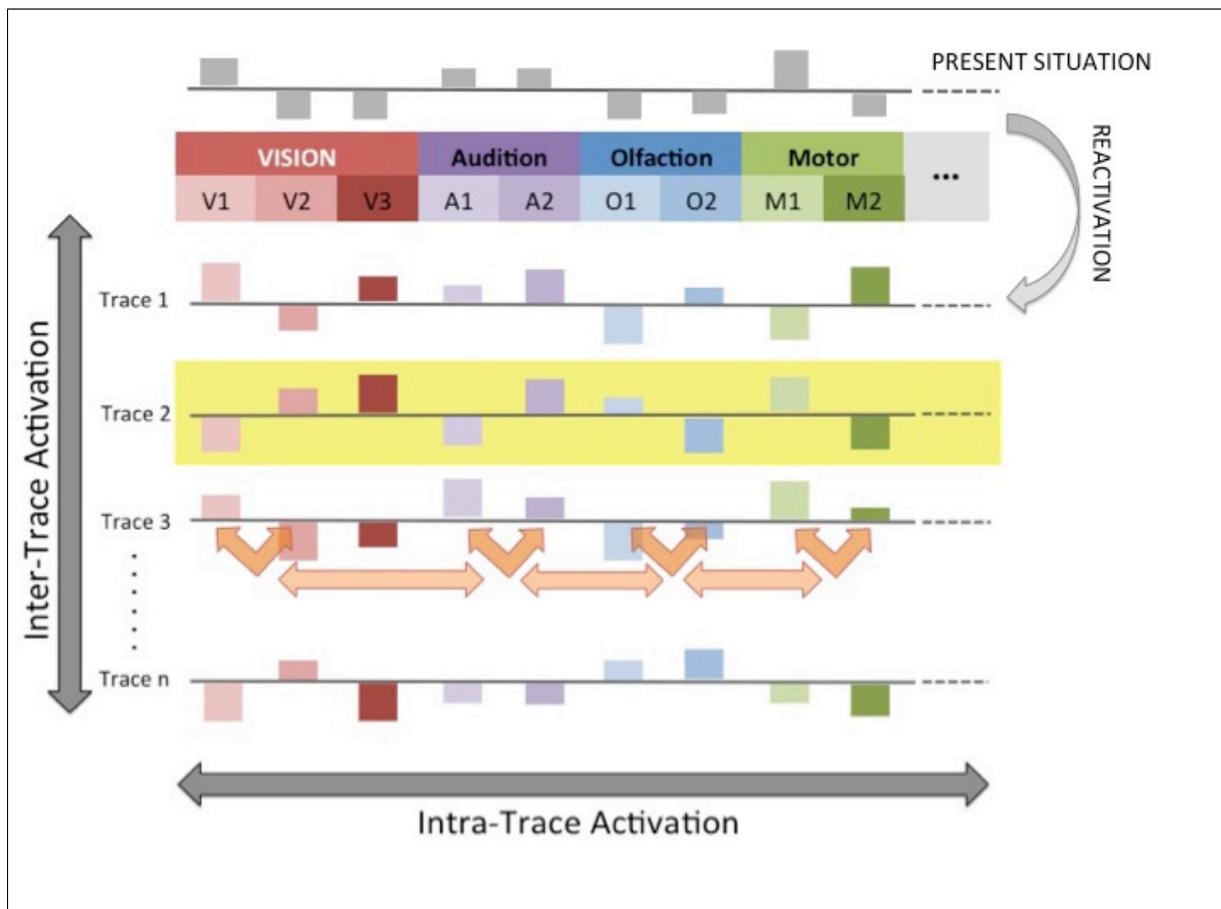


FIG. 8 – Représentation schématique du modèle Act-in inspiré par le modèle MINERVA II (Hintzman, 1986)

Dans le cas de situations nécessitant par contre de récupérer une connaissance spécifique, le couplage doit intervenir entre la situation présente et les traces mnésiques d'une situation antérieure précise. Cela est possible, seulement si la situation présente active un nombre limité de traces se référant à la situation passée ou encore si la diffusion d'activation inter-trace est limitée. L'efficacité d'une récupération de connaissances spécifiques dépendrait d'une part du potentiel qu'a une trace d'être spécifiquement réactivée et d'autre part, du potentiel de la situation actuelle de réactiver des traces spécifiques. Deux facteurs vont influencer le fait qu'une trace peut être spécifiquement réactivée : sa distinctivité et le niveau d'intégration de ses composants. Plus la trace sera distinctive, plus cela la rendra spécifique, donc différente

des autres traces. Cette observation fait référence à la spécificité de la trace dans son intégralité et non à la spécificité de chacun des composants de la trace. La spécificité d'un composant, lorsqu'il est activé, devrait rendre toute la trace plus accessible. Cela expliquerait pourquoi, plus les composants de la trace sont intégrés et forme une unité quasi-indissociable, plus il sera alors aisé de réactiver la trace dans son intégralité.

Concernant la capacité d'une situation à réactiver une trace spécifique, la présence de caractéristiques spécifiques devrait agir comme un indice fort pour réactiver spécifiquement les traces qui possèdent ces mêmes caractéristiques. Ainsi la présence d'indices pertinents dans l'environnement devrait réduire le nombre de traces activées (réduction de la diffusion d'activation inter-trace).

1.4 Problématique générale

Nous avons vu dans cette première partie que les connaissances en mémoire pouvaient s'articuler autour de représentations perceptives. Il a été mis en avant l'importance des modalités sensorielles dans l'organisation des connaissances. Les modèles de mémoire à traces multiples qui postulent l'existence de traces multimodales semblent donc constituer un cadre explicatif intéressant pour modéliser l'émergence des connaissances en mémoire. Cette émergence s'appuierait, notamment dans le modèle Act-in autour des mécanismes d'activation/intégration des modalités sensorielles en mémoire. Nous avons également introduit les notions de mécanisme d'activation intra et inter-traces pour rendre compte de l'émergence des connaissances dans ce cadre théorique. Dans les chapitres théoriques qui suivent, nous allons aborder l'intégration dans sa composante perceptive puis dans sa composante mnésique, nous essaierons également d'établir un lien entre ces deux types d'intégration.

Les objectifs de ce travail sont multiples mais s'inscrivent tous dans le même cadre explicatif. Le cadre théorique dans lequel nous allons situer notre problématique est représenté par le modèle de mémoire à trace multiple Act-in (Versace *et al.*, 2002 ; Versace *et al.*, 2009) et nous étudierons plus précisément les deux mécanismes constitutifs du modèle qui sont l'intégration et l'activation. Bien qu'il existe des données dans la littérature qui tentent d'expliquer les relations entre ces deux mécanismes (Brunel *et al.*, 2009; Brunel *et al.*, 2010; Vallet, Brunel & Versace, 2010), il reste essentiel de spécifier l'aspect fonctionnel entretenu par ces deux mécanismes. L'intégration et l'activation sont intrinsèquement liées et sont mises en œuvre par le système mnésique dès lors qu'il s'agit de faire émerger une connaissance ou

de créer une trace à partir d'une expérience. En d'autres termes, l'activation et l'intégration interviennent tant dans des processus de perception que dans des processus mnésiques, toutefois cette distinction entre système mnésique et perceptif peut rendre obscur le lien entre ces deux mécanismes. Dans ce travail, nous allons nous efforcer de comprendre comment l'activation et l'intégration se combinent dans la perception mais aussi dans la mémoire. Pour cela nous allons introduire une notion importante qu'est l'interaction (Driver & Noesselt, 2008 ; Versace *et al.*, 2009, Ernst & Bühlhoff, 2004). Selon nous l'interaction est un traitement nécessaire à la formation d'une trace, en d'autre termes, il est le prérequis aux mécanismes mentionnés ci-dessus.

1.4.1. De l'interaction à l'intégration

Dans le deuxième chapitre, nous allons essayer d'introduire la notion d'interaction à l'aide d'un paradigme simple d'associations bimodales. Nous pensons que la présentation conjointe de deux percepts sensoriels doit aboutir à la création d'une trace mnésique. La constitution d'une telle trace est progressive et devrait se traduire par la création d'une nouvelle entité, dans laquelle les modalités sensorielles préexistantes sont devenues indissociables (Goldstone, 2000). Selon nous l'interaction multimodale serait un prérequis à l'intégration progressive des composants en mémoire. Nous allons utiliser une phase d'apprentissage proche de celle utilisée par Brunel et collaborateurs (2009). Cela va nous permettre de tester le caractère automatique de l'interaction multimodale. En dehors de toute attention portée volontairement par le sujet sur les deux stimuli simultanément, nous devrions tout de même observer une interaction puis l'intégration de ces deux traits.

Nous allons donc utiliser deux groupes de sujets pour lesquels l'association entre les traits visuels et auditifs va varier. L'apprentissage devrait être meilleur pour le groupe pour lequel l'association entre le trait visuel et le trait auditif est systématique par rapport au groupe où il n'y a aucune association. Nous devrions donc observer une différence entre la pente d'apprentissage des deux groupes, une pente plus importante pour le groupe d'association systématique traduirait probablement le renforcement de l'association entre les deux traits.

1.4.2. Intégration perceptive et intégration mnésique

Dans la suite du deuxième chapitre nous allons reprendre la problématique de la partie précédente mais en changeant la méthode d'étude ainsi que le matériel. L'intérêt de cette partie est d'étudier l'intégration dans ses versants perceptif et mnésique. Nous posons

l'assertion selon laquelle l'intégration perceptive peut être étudiée directement alors que l'intégration mnésique ne peut l'être qu'à travers le mécanisme d'activation. En effet, si deux éléments sont intégrés en mémoire au sein d'une trace alors la présentation d'un des éléments va entraîner l'activation de l'autre élément. C'est la mise en œuvre du mécanisme d'activation qui statue sur l'intégration des deux composants. En d'autres termes, la force de l'association existant entre deux traits conditionne la capacité d'un des traits à activer l'autre. Nous allons donc comparer deux types de matériels, un matériel bimodal sémantiquement congruent et un matériel bimodal accessoire. Le matériel bimodal sémantiquement congruent sera constitué de traits audiovisuels déjà associés en mémoire, par exemple une forme géométrique pour le trait visuel et son étiquette verbale pour le trait auditif. Comme ces deux traits sont déjà associés en mémoire (intégrés), la présentation de l'un devrait activer l'autre. A l'inverse le matériel bimodal accessoire pourrait être constitué d'une forme géométrique et d'un son aigu par exemple. Ainsi aucun des deux traits ne peut activer l'autre, car il n'existe aucune trace en mémoire pour cette association. Avec ces deux matériels, nous pouvons donc directement tester la distinction entre intégration perceptive et mnésique en s'abstrayant d'une phase d'apprentissage préalable.

Nous allons utiliser la méthode du Race Model (Miller, 1982) pour étudier les mécanismes d'intégration perceptive et d'intégration mnésique. Le Race Model permet de mettre en évidence la présence d'une intégration multi-sensorielle significative et d'écarter l'hypothèse de simples mécanismes facilitateurs. De plus, le Race Model ne devrait mesurer que de l'intégration perceptive et non mnésique. En effet, seule l'interaction entre composants sensoriels devrait s'accompagner d'une intégration multi-sensorielle, comme le mesure le Race Model. Afin de faciliter la compréhension de notre propos, nous ajouterons que le Race Model devrait mesurer une intégration entre composant sensoriels plus forte lorsque les composants sensoriels n'entretiennent pas de liens de congruence sémantique entre eux. Cette hypothèse repose sur le postulat du modèle Act-In selon lequel les composants fortement intégrés en mémoire s'activent réciproquement. Dans le cas d'un stimulus composé de traits sémantiquement congruents, la présentation de l'un des traits réactive l'autre et dans ce cas, la présentation des deux traits simultanément n'apporte pas plus de gains que la présentation bimodale elle-même. En effet, si un trait de la trace est capable de réactiver automatiquement et très précocement l'autre trait de la trace, alors on devrait se trouver dans une situation proche de la présentation bimodale. Par conséquent, si la présentation unimodale est proche

de la présentation bimodale, on devrait trouver moins d'intégration dans la condition de congruence que dans celle de non congruence.

1.4.3. Influence d'un encodage multimodal sur la récupération à long terme

Dans le troisième chapitre, nous allons observer l'effet d'un encodage multimodal sur la récupération à long terme. Il existe des études qui ont montré qu'un encodage multimodal pouvait apporter un bénéfice lors de la récupération (Murray *et al.*, 2004; Lehmann & Murray, 2005; Giard & Perronnet, 1999; Vallet *et al.*, 2010). Selon le modèle Act-in, les unités sensorielles sont l'élément constitutif des connaissances en mémoire, il est donc logique de penser qu'une manipulation sensorielle lors d'une phase d'encodage pourra moduler la récupération en mémoire à long terme. Nous allons donc essayer d'interpréter les effets sensoriels en mémoire à l'aide du modèle Act-in, notamment à l'aide des mécanismes d'activation et d'intégration. Pour cela, nous allons mettre au point un paradigme en deux phases permettant de mesurer les effets de l'encodage sur la récupération. Afin de tester les effets sensoriels, nous utiliserons une condition multimodale congruente (*e.g.*, image du chien et aboiement du chien), une condition multimodale non congruente (*e.g.*, image du chien et bruit blanc) et une condition unimodale (*e.g.*, image du chien). Nous testerons notre paradigme dans des tâches nécessitant un accès implicite ou explicite à la trace, afin de voir quelle est l'influence de la manipulation sensorielle sur le type d'accès en mémoire

2. L'intégration multi-sensorielle en perception

2.1 Introduction

Dans le champ de l'intégration multi-sensorielle et notamment audiovisuelle, des chercheurs ont catégorisé les variables qui influencent les relations entre vision et audition à l'aide de deux facteurs (Spence, 2007). D'une part, les facteurs structuraux (*stimulus-driven*) qui incluent des éléments comme la contiguïté spatio-temporelle ou l'intensité du stimulus, d'autre part les facteurs cognitifs qui dépendent de connaissances antérieures ou de processus top-down et incluent notamment la congruence sémantique entre stimuli.

L'étude de l'intégration multimodale peut être envisagée sous deux angles : l'étude des perceptions et des réactions comportementales de sujets face à des situations sensorielles conflictuelles qui engendrent des illusions perceptives et d'autre part, l'étude des performances de sujets face à des entrées sensorielles multimodales qui portent des informations complémentaires d'un point de vue structurel ou sémantique.

2.1.1 Facilitation et conflit perceptif

On observe classiquement une modulation de la réponse à la présentation de stimuli de différentes modalités. Ainsi, le seuil de détection dans une modalité sensorielle donnée serait influencé par la présentation concomitante ou proche dans le temps d'un stimulus d'une autre modalité. Les recherches menées pour mieux comprendre ces mécanismes intégratifs dans des situations non-illusoires ont utilisé deux types d'approches : la comparaison du traitement de cibles bimodales à celui de cibles unimodales, on parle alors d'*effet de cibles redondantes* ("*redundant target effect*") et l'étude de la présentation d'un stimulus accessoire sur le traitement d'un stimulus cible de modalité différente (stimulus principal), on parle de *facilitation intersensorielle*. La présentation simultanée ou proche dans le temps d'un stimulus d'une autre modalité induit une diminution du temps de réaction à une cible de modalité sensorielle différente. Toutefois cette facilitation peut être influencée par différents facteurs que nous détaillerons par la suite. Des résultats similaires, mettant en évidence une facilitation, ont été retrouvés pour la localisation de cibles visuelles accompagnées de stimuli sonores (Molholm *et al.*, 2002). Par exemple, les temps de réaction peuvent être diminués de plusieurs dizaines de millisecondes lorsque le signal sonore est présenté à partir de la même source spatiale (McDonald, Teder-Sälejärvi & Hillyard, 2000). Toutefois, l'étude de la dépendance spatio-temporelle de cet effet révèle que plus la stimulation sonore est éloignée

du stimulus principal (dans le temps ou dans l'espace), moins l'effet de facilitation est marqué.

Cependant, une présentation multimodale n'entraîne pas systématiquement une facilitation, elle peut au contraire être à l'origine d'un ralentissement des temps de réponse ou d'une diminution des pourcentages de bonnes réponses. Par exemple, Colavita (1974) présentait des séquences composées de flashes et de sons à des sujets qui devaient répondre le plus rapidement possible sur une touche aux stimulations visuelles et sur une autre touche pour les stimulations auditives. Des essais bimodaux, composés d'une présentation simultanée des stimuli auditifs et visuels, étaient également présentés une fois tous les six items. Étonnamment, malgré des latences de réponse rapides et des pourcentages de réponse correctes pour les stimuli unimodaux, les sujets répondaient exclusivement à la modalité visuelle lors des présentations bimodales.

2.1.2 Facteurs modulant l'intégration multi-sensorielle chez l'Homme

2.1.2.1 La contigüité spatio-temporelle

La facilitation liée à l'intégration multi-sensorielle n'est cependant pas systématique. La survenue de cette facilitation dépend de nombreux facteurs qui modulent l'amplitude de l'effet tant au niveau comportemental (Lewald & Guski, 2003) qu'au niveau neuronal (Stein & Meredith, 1993). Deux propriétés fondamentales ont été mises en évidence : la contigüité temporelle et spatiale. Les premiers travaux réalisés chez l'animal ont montré que la réponse cellulaire était plus grande pour des stimuli bimodaux présentés dans une même fenêtre temporelle, par rapport à une condition où un délai était introduit entre la présentation des stimuli. De même, la réponse cellulaire était plus ample lorsque les stimuli bimodaux étaient présentés dans la même fenêtre spatiale.

Teder-Sälejärvi et collaborateurs (2005) se sont intéressés récemment à l'importance de la contigüité spatiale sur l'intégration multi-sensorielle. Dans leur expérience, les sujets étaient situés au milieu d'un dispositif expérimental comprenant face à eux, une enceinte et une LED sur la gauche et sur la droite. Les sujets devaient répondre à des stimuli rares auditifs, visuels ou audiovisuels qui étaient diffusés parmi des stimuli fréquents. Les stimuli rares se caractérisaient par une intensité sonore ou/et lumineuse plus importante que les stimuli fréquents. Toutes les combinaisons d'items étaient possibles : stimulus visuel à gauche (Vg) et à droite (Vd), auditif à gauche (Ag) et à droite (Ad), audiovisuel congruent (AVc) à gauche

et à droite et non congruent (AVn) avec un visuel à gauche et un auditif à droite et un visuel à droite et un auditif à gauche. Les résultats ont montré que pourcentages de détection étaient plus élevés et les latences de réponse plus rapides pour les stimuli bimodaux que pour les stimuli unimodaux quelle que soit la condition congruente ou non congruente. Peu de différences entre ces deux conditions ont été retrouvées en utilisant le Race Model (Miller, 1982). Malgré l'absence d'effet de congruence spatiale sur les données comportementales, les activités neuronales diffèrent clairement entre ces deux conditions.

L'importance de la simultanéité de présentation des stimuli tant au niveau spatial que temporel est attestée, toutefois il semble néanmoins que dans certaines conditions une facilitation inter-sensorielle soit possible, lorsqu'il existe un délai entre les présentations auditives et visuelles par exemple. Toutefois, les effets sur les activités neuronales semblent plus importants lorsqu'il existe une congruence spatiale (Teder-Sälejärvi, Di Russo, McDonald & Hillyard, 2005 ; Talsma, Doty & Woldorff, 2007).

2.1.2.2 La congruence

La congruence entre les stimuli provenant des différents sens peut également favoriser la survenue d'une facilitation multi-sensorielle. C'est ce qu'on appelle l'effet de redondance: les performances sont meilleures lorsque les stimuli bimodaux sont présentés systématiquement ensemble. On observe cet effet avec des stimuli bimodaux composés de formes sans signification et de sons purs (Giard & Perronnet, 1999) et ainsi qu'avec des stimuli constitués d'objets de la vie courante présentés avec le son qu'ils produisent habituellement (Molholm, Ritter, Javitt & Foxe, 2004). Les données comportementales indiquent qu'un objet bimodal est mieux et plus rapidement traité que le même objet défini par une seule de ses composantes quelle que soit la tâche à effectuer sur l'objet (Giard & Peronnet, 1999 ; Miller, 1982). Cet effet semble plus marqué lorsque la combinaison des composantes unimodales forme une unité perceptuelle. Miller (1991) a montré que l'effet est plus important pour une combinaison audio-visuelle fréquente (présentation répétée d'une combinaison audio-visuelle particulière) que pour une combinaison rare (présentation ponctuelle d'une combinaison audio-visuelle particulière). A l'inverse, la présentation de stimuli bimodaux non congruents peut perturber le traitement de stimuli bimodaux avec ou sans significations (Fort, Delpuech, Pernier & Giard, 2002 ; Schneider, Debener, Oostenveld & Engel, 2008). Fort et collaborateurs ont fait apprendre à des sujets des paires de stimuli bimodaux composés d'items visuels, auditifs et audiovisuels sans signification. Les auteurs ont montré que le traitement était perturbé lorsque les sujets étaient amenés à voir par la suite des stimuli composés d'items qui n'avaient pas été

associés lors de la phase d'apprentissage. Certains auteurs suggèrent que l'intégration multi-sensorielle pourrait faciliter la perception d'une modalité sensorielle et inhiber ou perturber la perception d'une autre modalité (Sinett, Soto-Faraco & Spence, 2008). Cette dissociation peut être mise en lien avec les travaux de Miller et D'Esposito (2005) qui ont différencié deux réseaux neuronaux, un premier réseau permettrait une coïncidence intermodale en contribuant à la détection des stimulations bimodales correspondantes et un second réseau permettrait la fusion ou le liage intermodal.

2.1.2.3 L'attention

L'attention peut également moduler les effets de l'intégration multi-sensorielle sur les performances des individus. Des études réalisées en EEG ont montré que l'attention dirigée vers une modalité est associée à une augmentation de l'activité neuronale des régions cérébrales correspondantes, par exemple, attendre un stimulus visuel augmente l'activité dans les régions visuelles (Eimer & Schröger, 1998). En d'autres termes, porter son attention sur une modalité sensorielle augmente l'activité cérébrale dans les régions dédiées au traitement de cette modalité. Des travaux réalisés en IRMf ont également montré une diminution de l'activité dans le cortex visuel durant des phases où l'attention devait être dirigée vers le stimulus auditif (Mozolic *et al.*, 2008b). Les auteurs ont montré que cette diminution d'activité dans le cortex correspondant à la modalité non attendue n'était pas accompagnée d'une augmentation d'activité dans le cortex correspondant à la modalité attendue. Cette diminution d'activité peut être expliquée par la suppression de cette modalité car la suppression sensorielle est normalement corrélée à des diminutions d'activité hémodynamique dans les régions sensorielles de la modalité ignorée (Laurienti *et al.*, 2002). Concernant l'intégration audiovisuelle, des auteurs ont montré qu'elle intervenait seulement lorsque l'attention était partagée entre deux modalités mais pas lorsque celle-ci était dirigée sur l'une ou l'autre des modalités (Mozolic, Hugenschmidt, Peiffer & Laurienti, 2008a).

D'autres études ont indiqué que la co-occurrence spatiale ou temporelle de stimuli auditifs et visuels durant une tâche où il fallait diriger son attention vers l'une ou l'autre des modalités pouvait entraîner un traitement attentionnel non intentionnel de la modalité à ignorer (Driver & Spence, 2000). D'autres travaux ont également révélé l'existence d'un traitement conjoint des modalités sensorielles alors qu'une seule des modalités devait être traitée. Dans une étude récente réalisée en EEG (Talsma *et al.*, 2007), les auteurs montrent que lorsqu'une seule des modalités est attendue, la co-occurrence temporelle des deux modalités cause l'association des stimuli à des étapes de traitement plus tardives. Busse et collaborateurs (2005) ont mis en

évidence en EEG et en IRMf que porter son attention vers un stimulus visuel augmentait l'activité dans les aires auditives suite à la présentation simultanée, mais spatialement éloigné d'un son devant être ignoré. L'intégration des composants sensoriels pourrait donc intervenir même dans des conditions où le composant de la modalité à ignorer n'est pas spatialement congruent. D'autres auteurs ont manipulé le caractère systématique ou non de l'association entre des traits audiovisuels dans un paradigme où seule une modalité ne doit être traitée (Baier, Kleinschmidt & Müller, 2006). Ils ont proposé à un groupe de sujets une tâche où un trait visuel était associé systématiquement à un trait auditif et à un autre groupe de sujets, la même tâche où le trait visuel et le trait auditif n'entretenaient aucune relation. Les sujets connaissaient la règle d'association mais ils devaient répondre uniquement à une des deux modalités qui était indiquée à l'aide d'un cadre de couleur qui apparaissait quelques secondes avant le stimulus. Les auteurs ont montré que lorsque les sujets attendaient une association non systématique, l'activité dans le cortex sensoriel indicé augmentait significativement alors qu'il n'existait pas d'activité dans le cortex sensoriel non indicé. Cependant, lorsque les stimuli auditifs et visuels étaient systématiquement associés, l'activité augmentait dans les deux systèmes sensoriels quelle que soit la modalité indiquée. D'autres auteurs ont montré que se focaliser sélectivement sur la modalité d'un objet entraînait une facilitation du traitement des traits d'autres modalités de cet objet qui devaient être ignorées (Molholm, Martinez, Shpaner & Foxe, 2007). Les auteurs concluent que se focaliser sur une modalité pour le traitement d'un objet entraîne la coactivation des représentations de cet objet dans les modalités ignorées. De plus, Molholm et collaborateurs mettent en évidence un effet de dominance visuelle, c'est-à-dire que les représentations visuelles des objets ont une plus grande influence sur les représentations auditives que l'inverse (Colavita, 1974).

Il semble au regard de ces données que l'intégration entre modalités sensorielles peut se faire, sous certaines conditions, de manière automatique, car bien qu'un sujet n'ait à traiter qu'une modalité sur les deux qui lui sont présentées, dans des paradigmes impliquant aussi bien des stimuli sans signification que des objets réels, un traitement est tout de même réalisé dans les régions cérébrales de la modalité que le sujet est censé ignorer.

2.2 Vers une conception interactive des systèmes sensoriels

2.2.1 Qu'en est-il de la conception modulaire des systèmes sensoriels ?

2.2.1.1 Les illusions perceptives

2.2.1.1.1 L'effet McGurk

Ce sont les données issues des illusions perceptives intermodales qui ont réellement permis de montrer la nature de l'intégration multi-sensorielle. L'un des exemples les plus frappants a été mis en évidence par McGurk et McDonald (1976) dans le domaine de la parole. L'effet McGurk est dû à l'influence de la perception visuelle du mouvement des lèvres sur la perception auditive de la parole. Dans leur expérience princeps, les auteurs montrent que la perception auditive de la syllabe /ba/ est modifiée par la perception simultanée du mouvement des lèvres articulant la syllabe /ga/, la syllabe finalement perçue étant /da/. Cet effet peut s'expliquer par le fait que les informations auditives et visuelles de la parole dans les conditions habituelles sont complémentaires et concordantes. Or acoustiquement et visuellement, les syllabes /ba/ et /ga/ sont très différentes l'une de l'autre, mais d'un point de vue acoustique la syllabe /ba/ est plus proche de /da/ que de /ga/, et d'un point de vue visuel la syllabe /ga/ est plus proche de /da/ que de /ba/. Compte tenu des proximités perceptuelles différentes de ces syllabes dans les deux modalités, et afin de satisfaire à la concordance habituelle des informations bimodales, l'interprétation la plus cohérente faite par le système cognitif est /da/. Le système cognitif va donc intégrer deux sources sensorielles contradictoires en un percept différent. Pour appuyer l'hypothèse d'un traitement multi-sensoriel, des auteurs ont montré que chez des patients atteints de démence de type Alzheimer, l'effet McGurk n'était plus présent que chez un tiers des patients testés (Delbeuck, Collette & Van der Linden, 2007). Or la maladie d'Alzheimer peut se caractériser par un syndrome de déconnexion cérébral (pour revue, voir Delbeuck, Van der Linden & Collette, 2003). Autrement dit, les communications entre les aires du cerveau sont détériorées. Par conséquent, cette diminution de l'effet McGurk dans la MA proviendrait bien d'une déconnexion sensorielle. L'isolation des modalités sensorielles conduirait à une perturbation de l'intégration issue du traitement sensoriel (Delbeuck *et al.*, 2007). Il semble donc au regard de l'effet McGurk et de l'hypothèse de déconnexion sensorielle que ce genre de phénomènes pourrait s'expliquer par une hypothèse de dégradation sensorielle plutôt que par une hypothèse de dégradation des représentations phonologiques en mémoire.

2.2.1.1.2 L'effet du flash illusoire

Les influences intermodales peuvent également être illustrées par l'existence d'illusions visuelles induites par des stimulations sonores. Shams, Kamitani, et Shimojo (2000) ont mis en évidence une illusion robuste dans laquelle la présentation d'un simple flash lumineux pouvait être perçue comme une succession de flashes lumineux lorsque celui-ci était

accompagné de plusieurs bips sonores. Cette illusion illustre le fait qu'un son peut moduler la perception visuelle. Afin de déterminer si ces effets intermodaux intervenaient à des étapes précoces/tardives des aires visuelles ou dans les aires associatives, les auteurs ont réalisé le même type d'expérience en potentiels évoqués (Shams, Kamitani, Thompson & Shimojo, 2001). Les résultats indiquent que les potentiels évoqués visuels sont modulés par la présentation du son dans les essais illusoires mais qu'ils ne sont pas influencés dans les essais non illusoires. Les auteurs ont montré que cette modulation des potentiels évoqués visuels par un son intervenait vers 170ms. En considérant que les potentiels évoqués avant 200ms post-stimulus sont dus à l'activité dans les aires modalité-spécifiques, ces modulations doivent vraisemblablement provenir des aires visuelles. De plus, les auteurs ont montré que les activités cérébrales dans les aires visuelles associées au second flash illusoire et celles associées au second flash physique étaient similaires. Cette similarité suggère que les mêmes mécanismes cérébraux sous-tendent le traitement de ces deux percepts. Cela signifie que, comme la présentation du flash physique induit une activité dans les aires visuelles, la représentation du flash illusoire implique également une activité dans le cortex visuel.

2.2.1.1.3 La Synesthésie

Il est peut être étonnant de placer la synesthésie dans les illusions perceptives, car ce n'est pas la perception du stimulus qui est modifiée dans ce phénomène, en effet la perception de ce stimulus engendre l'association systématique d'une autre perception. Nous la classerons tout de même ici, car ce phénomène est induit par une stimulation réelle. La synesthésie intervient lorsqu'un individu éprouve des sensations produites par un stimulus externe, mais qui ne sont pas normalement associées à ce stimulus. Chez les sujets synesthètes, des sensations correspondant à un sens (appelé inducteur) évoquent spontanément des sensations liées à un autre sens (appelé concurrent). La synesthésie est une expérience perceptuelle consciente, concrète et involontaire et n'est pas un produit de l'imagination. Si des sensations synesthésiques peuvent survenir après la consommation de certaines drogues (haschisch, mescaline) ou des troubles neurologiques (crise épileptique, lésion), la synesthésie est généralement de nature développementale et apparaît dès le plus jeune âge.

Il existe de nombreuses formes de synesthésie qui mettent en jeu différentes combinaisons inducteur-concurrent avec des perceptions très précises. Par exemple, chez un sujet particulier, un goût peut induire la sensation d'une forme géométrique spécifique dans sa main gauche. Certains synesthètes pourront parler de la couleur, de la forme ou de la saveur de la voix d'un locuteur ou d'une musique. Toutefois, les formes les plus répandues de synesthésie

sont celles où un mot ou une lettre entendue induit la perception d'une couleur (voir FIG. 9). Par ailleurs, l'expérience peut être intra-modale : par exemple, la forme graphémique de mots, lettres ou chiffres induit une perception de couleur.



FIG. 9 - Illustration d'un cas de synesthésie. La perception de chaque caractère induit la sensation d'une couleur spécifique

La synesthésie est un phénomène durable et robuste dans le temps : chez un sujet donné, la perception de la couleur rouge à l'entente (ou à la vue) de la lettre "B" sera reproductible. C'est également un phénomène unidirectionnel : si la perception de la lettre "B" induit la couleur rouge, la couleur rouge n'induit pas la perception de la lettre "B".

Bien que les bases neuronales de la synesthésie soient encore inconnues aujourd'hui, deux théories neurobiologiques ont été proposées pour rendre compte de ce phénomène. La première suggère que la synesthésie pourrait être due à une connectivité anormale entre les aires cérébrales des modalités sensorielles de l'inducteur et du concurrent. Chez de nombreuses espèces, on constate l'existence à la naissance, de connexions abondantes et exubérantes entre les différentes régions cérébrales sensorielles dont la plupart disparaît avec la maturation du système nerveux. Ainsi, nous serions tous synesthètes à la naissance, et nous perdriions cette expérience perceptuelle avec la perte de connexions intermodales. L'expérience synesthésique pourrait être due au maintien de ces connexions exubérantes au cours du développement ou à la création de connexions neuronales "anormales" (Baron-Cohen, Harrison, Goldstein & Wyke, 1993). En revanche, Grossenbacher et Lovelace (2001) suggèrent que la synesthésie n'est pas due au maintien ou à la création de connexions exubérantes, mais à un dysfonctionnement de connexions présentes chez tout individu. Il s'agirait plus précisément de désactivation de connexions en retour (*feedback*) dans les voies neuronales "normales" provenant de zones de convergence intra- ou inter-modale. Cette hypothèse s'appuie notamment sur l'observation mentionnée plus haut, que le phénomène de synesthésie peut être induit chez les non-synesthètes par la consommation de drogues qui agissent sur le cerveau en modifiant l'activité de réseaux neuronaux déjà existants.

2.2.1.2 Argument neuropsychologique

Des données issues de la neuropsychologie témoignent également de phénomènes intégratifs multimodaux. Ainsi, des auteurs ont testé l'intégration audiovisuelle chez des patients ayant un déficit visuel (Frassinetti, Bolognini, Bottari, Bonora & Làvadas, 2005). Le phénomène de blindsight consiste pour certains patients à répondre à des stimuli visuels dans une partie de leur champ visuel pourtant aveugle. Certaines structures intervenant dans ce phénomène tels que le colliculus supérieur (CS), sont des structures « hétéromodales », c'est-à-dire qu'elles reçoivent des projections provenant de différentes modalités sensorielles. Une majeure partie des efférences des CS sont des neurones multi-sensoriels, ce qui suggère que l'interaction entre des inputs sensoriels différents intervenant dans le CS pourrait affecter le traitement unimodal de l'information visuelle dans l'hémi champ aveugle. Frassinetti et collaborateurs (2005) ont ainsi testé l'hypothèse qu'une stimulation audiovisuelle de l'hémi-champ anopsique pourrait améliorer la perception des événements visuels dans l'hémi champ aveugle des patients anopsiques. Cette hypothèse est basée sur les règles spatiales et temporelles qui gouvernent l'intégration multi-sensorielle des neurones du CS (Stein, Wallace & Meredith, 1995). Le principe spatial dépend de l'organisation spatiale des champs récepteurs des neurones multi-sensoriels. En effet, chaque neurone multi-sensoriel du CS a de multiples champs récepteurs correspondant à chaque modalité pour lesquelles ce neurone est activé. Par conséquent, des entrées auditives et visuelles provenant de la même localisation spatiale, pourraient activer un neurone multi-sensoriel du CS. La réponse engendrée par cette combinaison de stimuli sera significativement plus importante que celle provoquée par un stimulus seul. Les résultats obtenus par les auteurs suggèrent que la modalité auditive affecte la modalité visuelle déficitaire. Ainsi, la capacité des patients à percevoir consciemment la présence d'un stimulus visuel dans le champ aveugle, augmente avec l'association entre stimulus visuel et auditif. Il faut noter que cette augmentation de la détection n'est observée que lorsque la cible visuelle et le son proviennent de la même position spatiale.

2.2.1.3 Au-delà de la conception modulaire des systèmes sensoriels

Les illusions perceptives illustrent ainsi bien le fait que l'intégration perceptive ne résulte pas seulement d'une facilitation de la détection par amplification du signal sensoriel, mais combine ces signaux pour former une nouvelle représentation multimodale de la source de l'objet (O'Hare, 1991). Les données issues des illusions perceptives remettent ainsi en cause le cloisonnement des systèmes sensoriels tel qu'il a été défini par Fodor (1983), et montrent que les systèmes sensoriels pourraient s'influencer mutuellement dès les premières étapes de

traitement de l'information. D'autres études menées sur la plasticité cérébrale montrent qu'il existe une plasticité intermodale chez des personnes souffrant de privation sensorielle précoce. Les résultats dans ce type d'études montrent que la privation sensorielle dans une modalité au début du développement a pour conséquence une réorganisation cérébrale du cortex normalement utilisé pour cette modalité, ce cortex pouvant être utilisé par d'autres modalités (Sadato *et al.*, 1996). Toutes ces données considérées ensemble apportent des arguments contre la notion de modularité stricte des modalités sensorielles et les travaux présentés ici montrent qu'il existe des interactions multimodales dans des régions considérées traditionnellement comme unimodales. Ces résultats indiquent que l'interaction entre modalités sensorielles représente la règle plus que l'exception (Shimojo & Shams, 2001). De plus, il est considéré traditionnellement que l'intégration au niveau cortical des informations provenant des différents sens est dédiée à des aires cérébrales associatives de haut niveau. Toutefois des données issues de la neurobiologie (pour revue, voir Ghazanfar & Schroeder, 2006), suggèrent que l'ensemble du néocortex est multi-sensoriel. Ces données nous obligent à reconsidérer la validité de la conception strictement modulaire du traitement multimodal mais aussi de donner plus d'importance au rôle que peuvent jouer les aires modalités-spécifiques dans l'intégration multi-sensorielle.

2.2.2 Acquisition de nouvelles associations multi-sensorielles

Nous avons vu que l'on pouvait remettre en cause la conception modulaire des systèmes sensoriels au sens Fodorien du terme. Nous allons voir maintenant que les apprentissages multi-sensoriels permettent de mieux comprendre les relations qui peuvent se former entre des modalités sensorielles qui sont amenées à être combinées au cours d'une expérience donnée.

2.2.2.1 L'apprentissage perceptif multi-sensoriel

Il est généralement admis que les protocoles d'entraînement multi-sensoriel produisent des apprentissages plus efficaces que les apprentissages unisensoriels à condition qu'il existe des relations appropriées entre chaque modalité (Shams & Seitz, 2008). Pour exemple, Seitz, Kim et Shams (2006) ont comparé des conditions d'entraînement audiovisuel et visuel et ont montré que les sujets ayant bénéficié d'un apprentissage multi-sensoriel avaient de meilleures performances dès les phases d'entraînement par rapport aux sujets placés dans les groupes d'apprentissage unisensoriel. Lorsque les auteurs ont testé ultérieurement l'effet de ces apprentissages uniquement sur le matériel visuel utilisé lors de l'apprentissage, ils se sont aperçus que les performances étaient meilleures dans le groupe d'apprentissage audiovisuel.

Selon les auteurs, la récupération d'informations unisensorielles impliquerait les mêmes structures et représentations impliquées lors de l'encodage unisensoriel. La présentation visuelle ultérieure entrainerait une activation plus forte des structures visuelles concernées. Par contre, l'apprentissage multi-sensoriel, impliquerait la modification des connexions entre modalités sensorielles ou la formation de représentations multi-sensorielles, la présentation visuelle ultérieure entraînant alors une activation plus globale de structures impliquées lors de l'apprentissage (voir FIG. 10).

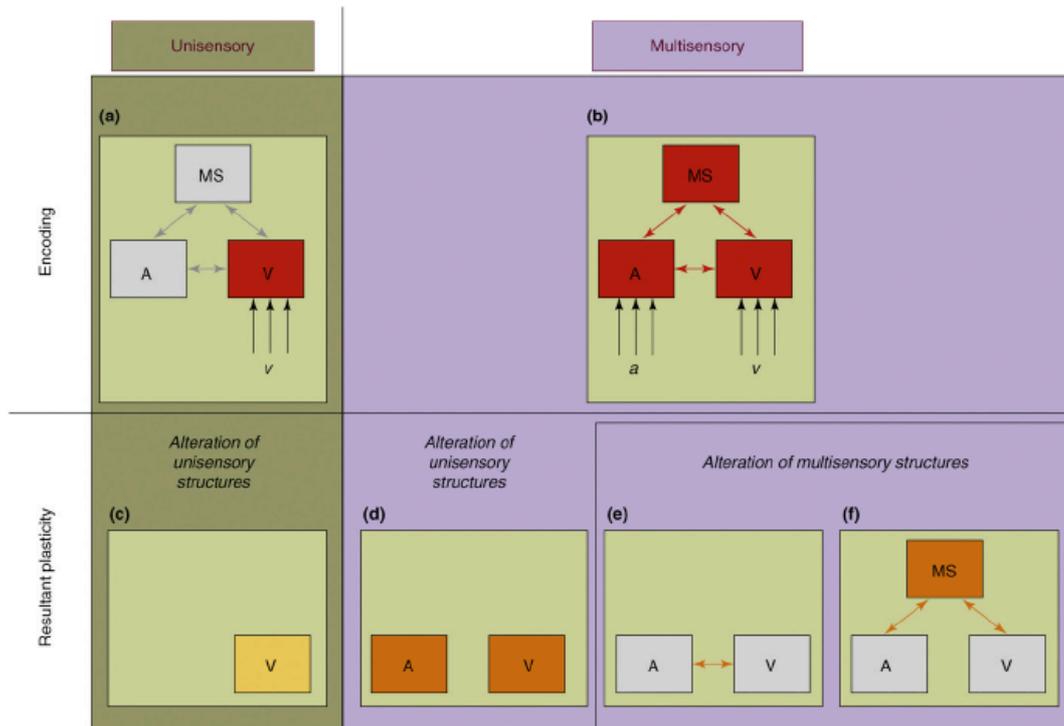


FIG. 10 - Schémas des processus impliqués dans un apprentissage unisensoriel vs multi-sensoriel. Durant l'encodage unisensoriel (a), seules les structures visuelles seraient activées, alors que durant l'entraînement multi-sensoriel (b), une gamme plus importante de structures serait activée (b). Par conséquent l'apprentissage unisensoriel n'entraîne de modification que sur les représentations visuelles (c), alors qu'un apprentissage multi-sensoriel peut modifier les relations entre modalités à différents niveaux : au niveau des structures unisensorielles (d) ou des structures associatives (e), (f).

Ces effets sont observables lorsque les traits auditifs et visuels sont congruents, en d'autre terme, lorsque l'association audiovisuelle ou plus généralement multimodale est redondante. Des auteurs (voir Seitz & Watanabe, 2005) ont montré qu'après la présentation brève d'une série d'items multimodaux, les sujets apprenaient à identifier ces paires audiovisuelles qui montraient des régularités statistiques au sein de la séquence. Peut-on obtenir de tels effets lorsque l'association entre les traits bimodaux est arbitraire ? Il semblerait que la réponse soit moins évidente (Shams & Seitz, 2008). Bien qu'un apprentissage sur une paire de stimuli multi-sensoriels doive induire une meilleure représentation du stimulus unisensoriel, on peut

penser que cet effet sera plus prononcé pour des associations congruentes. On peut définir la congruence comme la relation qui peut exister entre des stimuli à condition que cette relation soit consistante avec les expériences précédentes du sujet. Par exemple, Kim, Seitz et Shams (2008) ont comparé un apprentissage chez trois groupes, un groupe avec un stimulus visuel en mouvement, un groupe avec une congruence audio-visuelle (mouvement dans la même direction), et un groupe audiovisuel non congruent (mouvement dans deux directions opposées). Les auteurs ont montré que la facilitation était spécifique au groupe qui avait reçu un apprentissage audiovisuel congruent et ont conclu que cette facilitation impliquait la présence d'interactions audiovisuelles.

Il existe également des données concernant les effets intermodaux, lorsque ceux-ci impliquent des stimuli de modalités différentes, mais que le sujet ne doit traiter les caractéristiques que d'un seul des deux stimuli. Le trait que le sujet ne doit pas traiter est appelé trait orthogonal. Les effets intermodaux orthogonaux peuvent se définir comme l'influence générée d'un stimulus A non pertinent pour la tâche, sur un stimulus B d'intérêt, lorsque ces deux stimuli sont présentés conjointement. Le sujet devant uniquement traiter les propriétés du stimulus B. Cet effet a été mis en évidence dans une tâche de jugement, où le traitement de certaines caractéristiques dans une modalité a été influencé par une seconde modalité, alors que celle-ci ne fournissait aucune information à propos de la propriété à juger elle-même (Driver & Spence, 2000). Vroomen et De Gelder (2000), en utilisant une tâche de détection, ont également montré qu'un pattern visuel était détecté plus rapidement parmi d'autres patterns visuels lorsque celui-ci était accompagné d'un son qui différait des autres sons d'une séquence. Selon les auteurs, les événements visuels deviendraient groupés avec des événements auditifs concurrents par contiguïté temporelle. Selon d'autres auteurs, dans le domaine de l'apprentissage multi-sensoriel, la façon de concevoir comment l'association de deux modalités sensorielles peut améliorer le traitement d'une des deux modalités, une conception émerge. En effet, l'apprentissage ne pourrait se faire sur les traits de stimuli qui ne sont pas pertinents pour la tâche (Ahissar & Hochstein, 1993). Par la suite, Seitz et collaborateurs (2003) ont complété cette définition en indiquant que l'apprentissage perceptif pouvait intervenir même sur des traits non pertinents, dès lors que ceux-ci seraient proches du seuil de perception. Selon ces mêmes auteurs, les traits non pertinents pour la tâche ne seraient pas « remarqués » par le système cognitif, ils ne seraient donc pas supprimés par le système attentionnel et seraient appris. Récemment, Tsushima et collaborateurs (2008) ont confirmé l'hypothèse que la nature du trait non pertinent (signal fort/faible) déterminait l'apprentissage ultérieur. Ces dernières données offrent un argument pour penser que

l'interaction intermodale est automatique et ne nécessite pas que le sujet porte attention aux deux modalités pour qu'elles soient intégrées, à l'inverse la congruence serait un facteur critique.

2.2.2.2 Corrélats cérébraux des acquisitions de nouvelles associations multi-sensorielles

Des études utilisant diverses tâches, des paradigmes différents ainsi que des espèces différentes, ont apporté l'évidence que des nouvelles associations arbitraires se traduisaient par des modifications dans le cerveau. Certaines expériences ont montré que des primates pouvaient apprendre des associations multimodales, et dans le cas d'associations visuo-haptiques, des aires multimodales telles que le cortex rhinal antérieur seraient nécessaires pour ces types d'apprentissage (Goulet & Murray, 2001). Gibson et Maunsell (1997) ont fait apprendre à des singes macaques des associations arbitraires audiovisuelles. Après cet apprentissage, les auteurs ont mis en évidence une réponse sélective des cellules du cortex inféro-temporale à ces associations multimodales, ce qui indique que l'apprentissage multi-sensoriel modifie l'activité de régions sensorielles considérées comme modalité spécifique. Tanabe et collaborateurs (2005) ont étudié les corrélats neuronaux de l'apprentissage de paires audiovisuelles en utilisant l'IRMf. Ils ont observé que durant l'apprentissage, l'activité du cortex visuel augmentait pendant le délai suivant la présentation du composant auditif de la paire. Réciproquement, les auteurs ont observé une augmentation de l'activité du gyrus temporal supérieur en réponse au composant visuel de la paire. Ces résultats complètent ceux que nous avons cités précédemment, l'interaction multimodale est précoce et les activités des cortex sensoriels modalités spécifiques peuvent être modulées par la présence d'une autre modalité. De plus, il semble que durant un apprentissage multimodal, l'association qui se forme entre deux traits sensoriels se traduit par une modification de l'activité métabolique dans les cortex sensoriels.

2.3 Les modèles de l'intégration multimodale

Nous allons maintenant nous intéresser à la manière dont l'intégration multi-sensorielle peut être modélisée pour rendre compte des effets de facilitation ou non que l'on observe classiquement.

2.3.1 Modèles de compétition, modèle de coactivation et modèle de compétition interactif

La manière dont les informations extraites sont combinées ou intégrées pour activer la réponse appropriée a été sujette à débat. Comment savoir si les informations sont traitées en parallèle ou bien si elles sont intégrées ? Ces deux alternatives ont été confrontées notamment à l'aide du paradigme de détection de la cible redondante. Dans ce type de tâche, le sujet doit appuyer sur une touche lorsqu'une certaine cible visuelle apparaît sur un écran contenant un ou deux stimuli. Lorsque les cibles apparaissent à deux endroits différents, les temps de réponse sont classiquement plus rapides que lorsqu'une seule cible apparaît. Ce résultat est connu, comme nous l'avons vu précédemment sous le nom de *redundant target effect* (voir page 42) ou encore *redundant signals effect* (RSE).

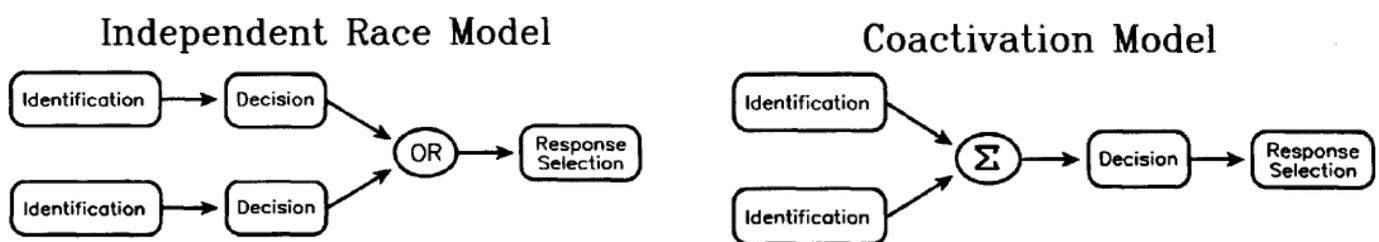


FIG. 11 - Modèle de compétition et modèle de coactivation (D'après Mordkoff & Yantis, 1991)

Deux classes de modèle ont été proposées pour rendre compte du RSE (pour revue, voir Miller, 1982). Le *modèle de compétition* (voir FIG. 11) présuppose que chaque élément présenté à l'écran est évalué indépendamment par un ou plusieurs canaux parallèles (*Independent Race Model*). L'information extraite des différentes sources ne serait jamais combinée, car la réponse du système serait provoquée dès qu'une décision serait prise par l'un ou l'autre canal. Ce type de modèle ne suppose donc pas la présence d'une quelconque intégration entre canaux. D'après le *modèle de compétition*, le RSE serait expliqué par une facilitation statistique (Raab, 1962). Le *modèle de coactivation* (voir FIG. 11) explique le RSE d'une toute autre manière, en effet, l'activation des canaux due à la présence de la cible ne produit pas de compétition, mais cette activation est mise en commun en une seule valeur d'activation. Cette activation totale entraîne une accélération plus importante du traitement que lors de la présence d'une seule cible, ce qui se traduit par des temps de réponse plus rapides. Puisque ces deux classes de modèles permettent d'expliquer le phénomène de RSE, la simple présence d'un RSE ne peut donc pas permettre de les discriminer. Toutefois, ces deux modèles font des prédictions différentes concernant les distributions de temps de réponse des conditions unimodales et multimodales. Par exemple, le modèle de compétition ne peut pas prédire un gain plus important que celui produit par une simple facilitation statistique. A l'inverse, les modèles de coactivation ne sont pas soumis à cette règle, ils

constituent pour cela une alternative intéressante pour expliquer les phénomènes d'intégration multi-sensorielle.

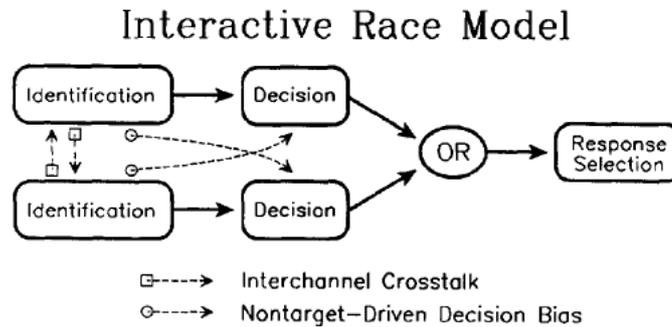


FIG. 12 - Modèle de compétition interactif (D'après Mordkoff & Yantis, 1991)

Toutefois, les modèles de coactivation ne sont pas les seuls modèles alternatifs aux modèles de compétition. Le *modèle de compétition interactif* (voir FIG. 12) partage certains points communs avec les deux modèles présentés ci-dessus, mais se distingue de chacun d'eux. Ce modèle présuppose des processus de décision séparés et une compétition entre les traitements des canaux sensoriels, mais il ajoute que les canaux perceptifs peuvent échanger des informations avant même l'intervention des processus de décision, qui restent spécifiques à chacun des canaux contrairement au modèle de coactivation. Mordkoff et Yantis (1991) décrivent un premier mécanisme qui permet d'apporter de l'information sur l'identité d'un des éléments affichés à l'aide d'un canal perceptif différent. Ce mécanisme est lié à la connaissance plus ou moins implicite des relations de contingences qui peuvent exister entre les stimuli. D'ailleurs, il est établi que la probabilité d'association des paires audiovisuelles influence l'amplitude du RSE : les associations les plus fréquentes induisent un RSE plus important que les associations moins fréquentes (Miller, 1991). Enfin, un second mécanisme concerne le fait qu'un élément distracteur puisse être traité par un système d'identification au sein d'un des canaux et influence le processus de décision de l'autre canal.

2.3.2 Méthodes d'étude de l'intégration multi-sensorielle

2.3.2.1 Le Race Model

La simple présence d'un gain en temps de réaction dans la condition bimodale par rapport à l'une et l'autre des conditions unimodales n'est pas une preuve suffisante de l'existence d'un processus d'intégration, car ce gain peut s'expliquer dans un modèle simple de convergence tardive des voies sensorielles, par un phénomène de facilitation statistique. Le modèle proposé par Miller (1982) permet de mettre en évidence l'existence de réelles interactions

audiovisuelles. En effet ce modèle basé sur la comparaison des distributions de TR dans les conditions auditive, visuelle et audiovisuelle permet de rejeter un modèle d'activations séparées.

Bases mathématiques

Dans le modèle d'activations séparées proposé par Miller (1982), les deux canaux sensoriels auditif et visuel sont parallèles et convergent vers des processus communs. Ce modèle simple lui permet de faire des prédictions sur la distribution des TR bimodaux à partir des distributions de TR unimodaux en faisant un nombre limité d'hypothèses. Dans ce modèle, on part du principe que le TR pour un essai bimodal sera déterminé par le premier des traitements unisensoriels déclenchant les processus communs liés à la réponse, comme dans le modèle d'indépendance de Raab (1962). Dans un essai bimodal donné, le temps de traitement (TT) à l'instant de déclenchement des processus communs est donc le plus petit des deux TT auditif ou visuel. L'ensemble des essais bimodaux correspondent à une distribution bivariée des TT auditifs et visuels, c'est-à-dire une distribution de couples (TT_A, TT_V) . A chaque essai audiovisuel, c'est le plus petit des deux TT qui définit le TT audiovisuel, donc la distribution des TR bimodaux (résultant de la compétition) est égale à la distribution des minima de cette distribution bivariée.

Pour savoir si les données expérimentales sont explicables par le modèle, il faut donc pouvoir estimer indépendamment la distribution des $\min(TT_A, TT_V)$ et la distribution des temps de traitement audiovisuels TT_{AV} à partir de données observables. On peut donc prédire que, dans le modèle d'activations séparées, la distribution des TR bimodaux, qui est observable, est égale à la distribution des minima de la distribution bivariée dont les distributions marginales sont les distributions des TR unimodaux, qui sont toutes deux également observables.

Une façon d'appliquer ce modèle est de partir des moyennes des TR unimodaux, de postuler la normalité et l'égalité des variances de leurs distributions et d'en déduire la distribution des minima et donc leur moyenne. Cette moyenne peut être alors directement comparée au TR moyen obtenu en condition bimodale pour rejeter ou accepter le modèle d'activations séparées. Afin de se passer de l'hypothèse de normalité, Miller (1982) cherche au contraire à estimer cette distribution des minima à partir des distributions effectives des TR unimodaux. Pour cela, il est utile d'utiliser les fonctions de répartition des TR ou des TT.

Soient $p(TRA < t)$, la fonction de répartition des TR auditifs et $p(TRV < t)$, la fonction de répartition des TR visuels. Pour un t donné, $p(TRA < t)$ désigne donc la probabilité qu'un TR

auditif soit inférieur à une certaine valeur t , et $p(TR_V < t)$ la probabilité qu'un TR visuel soit inférieur à t . De même, $p[\min(TR_A, TR_V) < t]$ désigne la fonction de répartition des minima de la distribution bivariée des TR unimodaux en condition bimodale. Dans la condition audiovisuelle, le TR unimodal minimum sera inférieur à une valeur t si le TR auditif est inférieur à t ou si le TR visuel est inférieur à t , ou encore si les deux temps de traitements sont inférieurs à t .

Miller en déduit l'inégalité suivante :

$$p[\min(TR_A, TR_V) < t] \leq p(TR_A < t) + p(TR_V < t), \forall t$$

Cette inégalité est satisfaite par tous les modèles d'activations séparées, quelle que soit la dépendance des distributions de temps de réponse auditifs et visuels. Donc si la distribution des TR audiovisuels observés est telle qu'il existe une valeur de t telle que :

$$p(TR_{AV} < t) > p(TR_A < t) + p(TR_V < t)$$

alors la distribution des TR audiovisuels ne peut s'expliquer par aucun modèle d'activations séparées.

Application de l'inégalité

Il est ainsi facile de vérifier graphiquement si l'inégalité de Miller est respectée ou non. Il suffit de tracer la fonction de répartition des TR audiovisuels et la somme des fonctions de répartition des TR auditifs et visuels. Graphiquement, l'inégalité est falsifiée et le modèle d'activations séparées est rejeté si, à n'importe quel TR t , la fonction de répartition des TR bimodaux se trouve au-dessus de la somme des fonctions de répartition unimodales (voir FIG. 13).

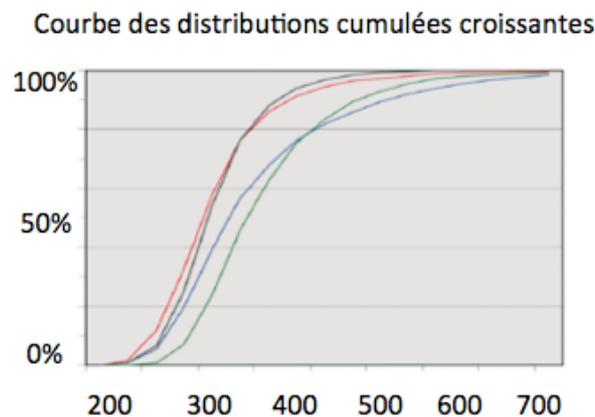


FIG. 13 - Illustration graphique de la falsification des modèles de compétition. De gauche à droite : la première courbe (rouge) correspondant à la fonction de répartition des TR bimodaux est graphiquement supérieure à la seconde courbe (noire) prédite par un modèle d'activations séparées. Les troisième et quatrième courbes correspondent aux fonctions de répartition des TR unimodaux. D'après Molholm et coll. (2002)

Analyse statistique

La falsification de l'inégalité de Miller s'applique en principe pour chaque sujet et est donc difficilement testable pour chaque sujet séparément par un test d'hypothèse. Donc elle ne garantit pas que la facilitation audiovisuelle n'est pas due à un biais d'échantillonnage des TR au niveau d'un sujet. Afin d'évaluer statistiquement la facilitation audiovisuelle, nous pouvons choisir de tester si la falsification de l'inégalité de Miller au niveau du groupe de sujets est attribuable ou non à un biais d'échantillonnage. Nous faisons l'hypothèse que si un effet est significatif au niveau du groupe, c'est qu'il reflète un effet réel au niveau des sujets.

Puisque l'application de l'inégalité de Miller utilise des distributions et non des moyennes de TR, il faut rassembler les distributions des différents sujets sans perdre l'information d'appariement des distributions auditives, visuelles et audiovisuelles de chaque sujet. On ne peut donc se contenter de comparer à chaque valeur t les moyennes des effectifs cumulés des sujets car cela pourrait gommer les violations de l'inégalité dans le cas où les sujets présentent des différences importantes de TR moyens entre eux. Pour obtenir les distributions de groupe, on peut utiliser une technique de regroupement des distributions connue sous le nom de vincentisation, proposée à l'origine par Vincent (1912) et appliquée par Miller (1982) puis Giray et Ulrich (1993) au test statistique de l'inégalité de Miller dans un groupe de sujets. Elle consiste à calculer un fractile donné de la distribution de groupe comme la moyenne de ce fractile à travers les sujets. Cette façon de moyenniser les distributions constitue une sorte de normalisation puisqu'elle permet d'éviter l'injection de variabilité due à des différences de TR absolus entre sujets et de faire ressortir les différences de distributions présentes chez tous les sujets (Ratcliff, 1979).

Pour chaque sujet, il est ainsi nécessaire de calculer la somme de ses distributions de TR unimodales. Ensuite, pour chaque sujet, il faut calculer 19 fractiles d'ordre 20 pour cette distribution et pour la distribution des TR audiovisuels. Afin de tester la significativité de différences entre les fractiles moyennés par sujet, une méthode consiste à réaliser des tests de Student pour vérifier que le fractile audiovisuel est plus faible que le fractile de la somme des distributions de TR unimodaux. Toutefois face à l'accumulation d'erreurs de type I que peut présenter cette méthode, Kiesel, Miller et Ulrich (2007) ont proposé de tester la violation de

l'inégalité seulement entre les percentiles 10% et 25%. En effet les valeurs de t intéressantes sont plutôt les valeurs faibles, puisque l'on s'attend, en cas d'interactions audiovisuelles à une accélération du temps de réaction et dans l'idéal à ce que les plus petits TR bimodaux soient inférieurs aux plus petits TR unimodaux. De plus, ces mêmes auteurs estiment que pour diminuer le biais consistant à produire une violation systématique du modèle, il faut considérer au minimum 20 items par condition.

2.3.2.2 Mesures d'amplitude de la violation

On peut également essayer d'évaluer l'ampleur de la violation, pouvant être définie implicitement comme la surface comprise entre la fonction de répartition audiovisuelle et la somme des fonctions de répartition unimodale, comme une mesure directe de l'effet de facilitation intersensorielle.

Colonus et Diederich (2006) ont ainsi proposé une mesure du gain, qui revient à estimer une aire comprise entre deux courbes. En posant $F_{av}(t) = p(TRA V < t)$, $F_a(t) = p(TRA < t)$, $F_v(t) = p(TRV < t)$, on obtient :

$$E(X) = \int_0^{\infty} R_{av}(t)dt = \int_0^{\infty} \max[1 - F_a(t) - F_v(t), 0] dt - \int_0^{\infty} [1 - F_{av}(t)]dt$$

$$= E[\min(RT_a, RT_v)] - E(RT_{av})$$

En pratique, cette valeur s'obtient en retranchant la moyenne des temps de réponse de la condition bimodale à la moyenne des minima de chaque couple de valeurs unimodales. L'avantage de cette méthode est de pouvoir ensuite inclure cette mesure dans un ensemble de tests statistiques. Toutefois, ce calcul prend en compte l'ensemble de la distribution, par conséquent, s'il existe une surface négative au sein de la violation, elle sera retranchée à la surface positive de la violation, et bien qu'il y ait une violation de l'inégalité à un temps t , la valeur calculée pourra être négative.

Si l'on s'intéresse à une violation de l'inégalité pour des t relativement courts, cette méthode ne semble plus adéquate, car obtenir une valeur négative sur l'ensemble de la distribution ne signifiera pas pour autant une absence d'intégration pour les premiers percentiles d'intérêt par exemple. Une autre méthode consisterait à tracer la courbe correspondant à la surface comprise entre la fonction de répartition audiovisuelle et la somme des fonctions de répartition unimodales afin d'obtenir une approximation moyenne de l'amplitude de la

violation. L'intérêt de cette méthode est que l'on peut ainsi comparer directement des conditions expérimentales au sein d'une même expérience (Mozolic *et al.*, 2008a ; Mozolic *et al.*, 2008b), ou alors comparer l'amplitude du gain dans une condition pour deux groupes distincts (Laurienti, Burdette, Maldjian & Wallace, 2006). Afin de tester la significativité d'une éventuelle différence entre les courbes de répartition, une méthode similaire à celle employée en EEG (Mozolic *et al.*, 2008a; Mozolic *et al.*, 2008b) peut s'appliquer. Elle consiste à réaliser des tests de Student entre chaque courbe pour chaque temps t et afin de diminuer l'erreur générée par les tests multiples, on considère comme significatif uniquement les périodes où 10 t-tests successifs sont inférieurs au seuil $\alpha = 0.05$.

2.3.3 Peut-on différencier intégration et interaction multimodales ?

Le mécanisme d'intégration a d'abord été étudié dans le champ de la perception visuelle et de l'attention. La théorie d'intégration des traits (Treisman & Gelade, 1980) est une théorie de l'intégration sensorielle, qui suggère que nous devenons conscients des objets, de deux façons différentes : soit à l'aide de l'attention focalisée, soit par un traitement de type ascendant (« top-down »). L'étape initiale consiste en une analyse rapide et parallèle des caractéristiques élémentaires de l'objet (couleur, forme, orientation, etc.). L'objet n'existe pas comme entité à cette étape de traitement. L'identification de l'objet et l'extraction des connaissances relatives à l'objet nécessitent l'intégration des caractéristiques élémentaires. Cette intégration est possible seulement si l'attention est focalisée sur une région spécifique du champ visuel. L'attention focalisée rend possible l'intégration de toutes les caractéristiques élémentaires de la région du champ visuel en question. La manière dont est décrite la réémergence des connaissances dans le chapitre précédent n'est pas si différente de celle proposée par Treisman pour obtenir un percept cohérent, la différence majeure provient de l'importance donnée à la composante attentionnelle dans le modèle de Treisman. En effet, dans ce modèle d'intégration, l'attention focalisée est un élément essentiel du liage perceptif des traits sensoriels, alors que selon Versace et collaborateurs (2002), les traits sont avant tout mnésiques.

Dans le modèle de Treisman, on parle d'intégration des composants élémentaires de la scène visuelle, mais on peut généraliser ce mécanisme à des composants présentés dans des modalités différentes. Toutefois, cette considération nous amène à la problématique suivante : doit-on penser qu'une présentation bimodale conduit irrémédiablement à une intégration multi-sensorielle ? C'est ici que nous allons introduire le terme d'interaction multimodale, nous allons essayer de définir par des exemples, en quoi ces deux notions peuvent différer. Le

terme d'interaction peut être utilisé pour signifier l'influence réciproque de composants sensoriels, ou l'influence d'un composant sensoriel sur un autre, mais qui restent néanmoins différenciés (Versace *et al.*, 2009 ; Driver & Noesselt, 2008). A l'inverse, le terme intégration devrait être utilisé lorsque les composants élémentaires de l'expérience ne peuvent plus être différenciés (Goldstone, 2000). Ernst et Bühlhoff (2004) ont proposés deux mécanismes afin de rendre compte de l'émergence d'un percept cohérent et robuste. Ils parlent de mécanisme de *combinaison* (*combination*) et d'*intégration* (*integration*). La *combinaison* serait un traitement non linéaire et permettrait de décrire les interactions entre des signaux sensoriels qui ne seraient pas redondants. Ce mécanisme jouerait notamment un rôle pour lever l'ambiguïté de certaines situations et permettrait la coopération entre modalités sensorielles. Au contraire, l'*intégration*, qui reflèterait un traitement sériel, permettrait de décrire les interactions entre signaux redondants. Ces deux mécanismes interviendraient de manière sérielle, de telle sorte que la *combinaison* soit un traitement subordonné à celui de l'*intégration*. Par exemple, frapper contre un morceau de bois (voir FIG. 14) entraînerait au moins trois estimations sensorielles à propos de la localisation de l'impact (L) qui pourraient être inférées à partir des informations visuelle (V), auditive (A) et proprioceptive (P). Le processus de combinaison sensoriel serait non linéaire car chaque input nécessiterait d'être transformé dans un système de coordonnées communes. A un niveau de traitement plus tardif, les trois signaux L1, L2 et L3 seraient intégrés pour former un percept cohérent de la localisation de l'impact.

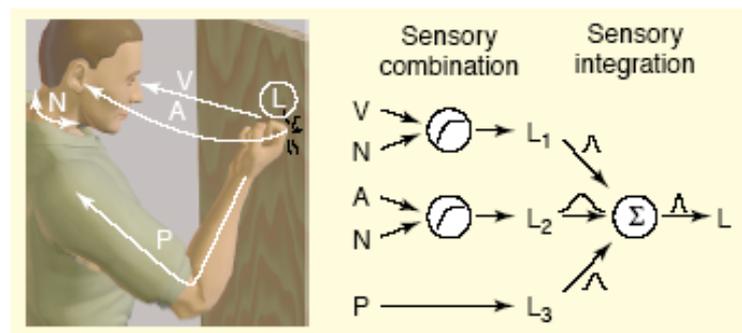


FIG. 14 – Formation d'un percept cohérent faisant appel à deux mécanismes : *combinaison* et *intégration* pour lever l'ambiguïté sensorielle d'une situation donnée, selon Ernst & Bühlhoff (2004)

Prenons par exemple l'étude réalisée par Molholm et collaborateurs (2002) dans laquelle les auteurs étudient les interactions audiovisuelles précoces en EEG. Dans ce paradigme, les sujets devaient détecter des stimuli visuels, auditifs et audiovisuels. Les items visuels pouvaient être présentés à différentes positions sur l'écran, les items auditifs étaient diffusés via un haut-parleur au-dessus de l'écran. Au niveau des temps de réponse, les résultats ont

montré que les items audiovisuels étaient détectés plus rapidement que les items visuels et auditifs. De plus, l'analyse des potentiels évoqués met en évidence des interactions audiovisuelles précoces au niveau des régions pariéto-occipitales droites. La topographie au niveau du scalp, ainsi que la temporalité de cet effet sont consistantes avec l'idée selon laquelle le traitement sensoriel visuel est modulé très précocement par les inputs auditifs. Une interprétation possible est la présence de projections directes des cortex associatifs auditifs vers les aires visuelles précoces, cette interprétation est rendue possible par l'existence de projections des cortex auditifs primaires et des aires associatives auditives vers les aires visuelles V1 et V2 (Falchier, Clavagner, Barone & Kennedy, 2002). Une autre interprétation plus indirecte est possible, si l'on suppose que les aires auditives activent des régions de convergence multi-sensorielle, qui à leur tour activeraient les aires visuelles unisensorielles. En résumé, dans cette étude, des interactions audiovisuelles sont présentes au niveau du scalp sur les régions pariéto-occipitales droites (40-50ms), occipito-temporales (165ms) et centrales (180ms). Les interactions audiovisuelles précoces seraient sous-tendues par les régions du système de traitement du mouvement visuel (MT+), la latéralisation de ces interactions seraient concordantes avec un réseau d'orientation de l'attention exogène classiquement observé dans l'hémisphère droit (Mesulam, 1981). Les interactions plus tardives sont mises en lien avec l'intégration sensori-motrice.

L'étude que nous allons présenter maintenant diffère de la précédente car elle a porté sur la reconnaissance d'associations multimodales préalablement apprises, il y a donc dans cette étude un facteur mnésique qui n'était pas présent dans l'étude de Molholm et collaborateurs (2002). Dans l'étude célèbre de Giard et Perronnet (1999), les auteurs se sont également intéressés à la dynamique d'interaction multi-sensorielle, mais dans un tout autre paradigme. Les sujets devaient effectuer une tâche de catégorisation : pour chaque essai, il fallait indiquer lequel des deux objets A ou B étaient présentés. Chaque objet était défini par des traits auditifs, des traits visuels ou la conjonction des deux. L'objet A était constitué d'un cercle pouvant se déformer en une ellipse horizontale et d'un son de 540Hz, l'objet B était constitué d'un cercle pouvant se déformer en une ellipse verticale et d'un son de 560 Hz. Les sujets répondaient sur une touche pour l'objet A et sur une autre touche pour l'objet B. Les résultats comportementaux ont mis en évidence des temps de catégorisation plus rapides et des pourcentages de bonnes réponses plus élevés lorsque les objets étaient présentés dans deux modalités plutôt que dans une seule. Concernant les corrélats neuronaux de ces effets multimodaux, les auteurs mettent en évidence des interactions précoces (40ms) dans le cortex visuel. De plus, les auteurs mettent en évidence une interaction robuste vers 145-160ms sur

les sites T4-FT4 de l'hémisphère droit. Cette région n'étant pas activée par la présentation des stimuli unimodaux, elle pourrait être le lieu d'interactions polysensorielles ou autrement dit, une région multimodale associative. Cette dernière interaction n'est pas observée dans l'étude de Molholm et collaborateurs (2002). D'un point de vue théorique, la présence de cette différence pourrait représenter la signature électrophysiologique de mécanismes intégratifs versus interactifs. Dans l'étude de Molholm et collaborateurs (2002), la facilitation pour la détection de cibles audiovisuelles pourraient s'expliquer par l'intervention de mécanismes interactifs, alors que dans l'étude de Giard et Perronnet (1999), on est plus à même de penser qu'aux mécanismes interactifs, viennent s'ajouter des mécanismes intégratifs supportés, par exemple, par des aires polymodales.

Alors qu'il semble possible de conceptualiser deux notions distinctes : l'interaction et l'intégration, il reste néanmoins à spécifier le lien qui existe entre elles. Selon certains auteurs, il existerait une relation hiérarchique avec des mécanismes interactifs subordonné à des mécanismes intégratifs (Ernst & Bühlhoff, 2004).

2.3.4 Liens fonctionnels possibles entre interaction et intégration selon Act-in

Versace et collaborateurs (2009) précisent que lorsque les items visuels et auditifs ne sont pas redondants à l'inverse des deux études que nous avons présentées dans la partie précédente, seul le mécanisme d'intégration serait mis à contribution. Par exemple, l'effet McGurk traduit bien le fait que lorsqu'un son et un mouvement des lèvres ne sont pas congruents, le son perçu est différent. Mais peut-on penser le mécanisme d'intégration sans le mécanisme d'interaction? En fait, on peut imaginer que le mécanisme d'interaction précède le mécanisme d'intégration, l'interaction multimodale serait automatique et irrépressible, ce mécanisme interviendrait dans des conditions multimodales ou intramodales. Par exemple, si des items visuels et auditifs sont systématiquement associés, l'interaction entre modalités devrait aboutir à la formation d'un percept cohérent, ces items devraient par conséquent être intégrés. Brunel et collaborateurs (2010) ont réalisé une expérience en deux phases, avec une première phase d'association où les sujets voyaient une forme systématiquement associée à un bruit blanc et une autre forme présentée seule. Dans la tâche d'amorçage qui suivait, il était présenté en amorce une des deux formes géométriques vues dans la première phase et en cible, des images d'objets sonores ou non sonores. Les auteurs ont montré que lorsque la forme avait été précédemment associée à un son, elle facilitait le traitement des items sonores pour un SOA de 500ms. On peut penser qu'il y a eu création d'un percept à partir des interactions multimodales entre une forme géométrique et le bruit blanc dans la phase

d'association. Ces interactions multimodales ont fait que les composants auditifs et visuels ont été intégrés au sein d'une trace. Le concept d'*unitization* (Goldstone, 2000) illustre également l'idée que des composants associés au sein d'une trace deviennent interdépendants. Pour reprendre l'exemple des illusions perceptives et notamment celle de McGurk (McGurk & MacDonald, 1976), la présentation du son /ba/ et du mouvement des lèvres /ga/ entraînerait l'activation de représentations phonologiques associées, car le son /ba/ et sa représentation phonologique sont intégrées au sein d'une trace, tout comme le mouvement des lèvres /ga/ et sa représentation phonologique. La présentation simultanée de ces deux percepts entraînerait une interaction audiovisuelle ainsi qu'une activation conjointe des représentations phonologiques associées et par conséquent un conflit. Le résultat du traitement des deux mécanismes produirait l'illusion perceptive, il existerait donc une compétition entre intégration et interaction en fonction de la nature des associations multimodales et des caractéristiques de celles-ci. Un autre argument issu de travaux menés en électrophysiologie indique qu'il peut exister des mécanismes interactifs même dans des situations de non congruence. Dans un paradigme équivalent à celui utilisé par Giard et Perronet, Fort et collaborateurs (2002) ont présenté trois objets composés chacun de traits visuels, auditifs et audiovisuels. L'objet A était constitué d'un cercle se déformant en ellipse horizontale (V1), d'un son de 530Hz (A1) et de la conjonction des deux (A1V1), l'objet B était constitué d'un cercle se déformant en ellipse verticale (V2), d'un son de 610Hz (A2) et de la conjonction des deux (A2V2). L'objet C'était quant à lui constitué d'un cercle (V3), d'un son de 560Hz (A3), et des conjonctions audiovisuelles A1V2 et A2V1. Les résultats montrent que lors de la présentation bimodale, les objets A et B sont catégorisés plus rapidement que l'objet C, de plus l'objet C est sujet à plus d'erreurs que les objets A et B. Ces résultats comportementaux indiquent bien la présence d'un conflit dans le cas où les items bimodaux sont non redondants. Les résultats obtenus en électrophysiologie sont également très intéressants, car les auteurs observent la présence d'activités neuronales précoces durant la reconnaissance des objets caractérisés par des associations auditives et visuelles non redondantes. Toutefois, les activités intermodales sont moins fortes en amplitude et plus tardives que celles trouvées pour des associations multimodales redondantes, mais elles présentent la même topographie que ces dernières. Les interactions multimodales, si l'on suppose qu'elles interviennent très précocement, seraient présentes également dans des situations de non congruence.

2.4 Conclusion

Nous avons présenté la notion d'intégration multi-sensorielle et nous avons spécifié à partir de celle-ci, deux notions, l'interaction et l'intégration. Cette distinction fait référence à la nature des processus engagés dans la perception ainsi qu'aux relations qu'entretiennent les items présentés dans les différentes modalités. Ainsi, l'interaction serait un processus subordonné au processus d'intégration (Ernst & Bühlhoff, 2004). Il semble que l'interaction puisse qualifier l'influence d'une modalité sur une autre sans que celle-ci ne se traduise par la création d'un percept unifié (Versace *et al.*, 2009 ; Driver & Noesselt, 2008 ; Goldstone, 2000). Nous avons également présenté un modèle mathématique permettant de tester la présence ou non d'une intégration multi-sensorielle. Le Race Model (Miller, 1982) est en effet capable de discriminer entre un simple effet de facilitation sensorielle et entre un processus intégratif multimodal. Toutefois, on peut se demander quelle est la nature des processus que l'on teste réellement avec cet outil. Par exemple, le Race Model pourrait mesurer de l'interaction, de l'intégration ou les deux. De plus, les effets facilitateurs obtenus avec une telle modélisation pourraient également être interprétés au regard du modèle Act-In et des prédictions faites par les mécanismes d'intégration/activation. Il semble donc intéressant de pouvoir discuter des observations faites par ce modèle et des processus mis en jeu dans une tâche quelconque.

2.5 Problématique

2.5.1 Etude de la création d'une trace multi-sensorielle

L'association de plusieurs traits sensoriels aboutit à la formation d'une trace dont un des composants peut réactiver l'ensemble de la trace ultérieurement (Brunel *et al.*, 2009). L'interaction entre des traits sensoriels de différentes modalités devrait aboutir à la création d'une trace et par conséquent à l'intégration de ces traits. Nous allons nous intéresser dans cette série d'expériences à la spécificité des interactions multi-sensorielles par le biais de l'apprentissage. Dans une première expérience, nous allons comparer un apprentissage multi-sensoriel à un apprentissage unisensoriel pour voir si la création d'une trace multimodale produit de meilleures performances qu'un apprentissage unimodal (Expérience 1). Nous testerons ensuite l'aspect spécifique de ces interactions multi-sensorielles en manipulant l'association entre les composants auditifs et visuels (Expérience 2). En modifiant le paradigme d'étude, nous essaierons de quantifier le gain multi-sensoriel à l'aide du Race Model (Expérience 3) dans une condition où il y a eu création d'une trace multi-sensorielle

par rapport à une condition où les associations multi-sensorielles n'ont pas conduit à une telle création.

2.5.2 Intégration/interaction multi-sensorielle et Race Model

Nous avons vu que le Race Model était un outil souvent utilisé pour mesurer un gain intégratif dans des tâches où l'on compare des conditions multimodales à des conditions unimodales. Dans l'étude des relations audiovisuelles, de nombreuses études ont montré que présenter deux stimuli dans deux modalités différentes pouvaient conduire à apporter un gain intégratif par rapport à celui prédit par un modèle de co-activation (Giard & Perronnet, 1999; Molholm *et al.*, 2002; Molholm *et al.*, 2004;). D'autres études se sont également intéressées aux processus intégratifs dans le vieillissement et ont comparé pour cela des groupes de sujets jeunes à des groupes de sujets âgés (Laurienti *et al.*, 2006). L'étude de Laurienti a notamment montré que dans une situation de congruence multimodale sémantique, les sujets âgés présentaient un gain intégratif supérieur à celui présenté par les sujets jeunes. Pour expliquer cet effet surprenant, les auteurs ont proposé que la dégradation présente chez les personnes âgées au niveau du traitement des sens, pouvait être à l'origine de ce qu'ils avaient observé. Laurienti explique que les sujets âgés devraient tirer plus de bénéfices dans les conditions multimodales qu'unimodales, ce qui augmenterait l'amplitude du gain observé par rapport aux sujets jeunes qui gardent un traitement sensoriel optimal.

Dans cet axe de recherche nous allons nous centrer sur les processus intégratifs des sujets jeunes en essayant d'apporter une explication sur le fait que cette population puisse présenter un gain intégratif plus faible que des sujets âgés lorsque celui-ci est mesuré à l'aide du Race Model. Selon nous, cet effet serait en grande partie lié à la méthode de mesure du gain intégratif. Dans une situation de congruence sémantique multimodale, la présentation d'un item visuel active la représentation de l'item auditif et réciproquement (Vallet *et al.*, 2010), par conséquent la présentation des deux items simultanément ne devrait donc pas apporter beaucoup plus de bénéfice que la présentation des items isolément. Cette hypothèse pourrait rendre compte du fait que les sujets jeunes présentent un gain intégratif plus faible que celui des sujets âgés.

Nous allons dans un premier temps utiliser une tâche de catégorisation proche de celle de Laurienti pour reproduire les effets obtenus dans l'étude de 2006. Nous devrions retrouver un gain intégratif plus important chez les sujets âgés que chez les sujets jeunes. Dans un deuxième temps, nous allons utiliser une tâche de détection avec des items qui ne possèdent

pas de liens sémantiques entre eux. Dans cette tâche plus perceptive, on ne devrait pas obtenir de différences entre le groupe de sujets âgés et le groupe de sujets jeunes.

Toutes Les expériences de cette partie ont été conçues et réalisées sur un ordinateur MacBook grâce au logiciel Psyscope (Cohen, MacWhinney, Flatt & Provost, 1993). Les données ont été recueillies à l'aide d'un Mac Book Pro.

La majorité des données brutes obtenues dans cette partie expérimentale sont présentées en Annexes (voir page 241).

2.6 Partie expérimentale

Partie 1

Dans cette partie, nous allons d'abord nous intéresser à la création d'une trace multi-sensorielle à partir d'un apprentissage. Nous testerons également l'aspect spécifique de la création de cette trace et nous essaierons de quantifier le gain multi-sensoriel associé à la création de cette trace. Pour cela, nous allons utiliser une tâche dans laquelle le sujet devra juger de l'orientation d'un trait présenté à l'écran (*i.e.*, oblique ou vertical). Chacun des traits sera spécifiquement associé à un son dans le but de créer une association en mémoire.

Expérience 1

Hypothèses

H1 : L'apprentissage multimodal, dont l'association est constituée d'un trait perceptif d'intérêt et d'un trait perceptif orthogonal, devrait être meilleur qu'un apprentissage unimodal portant seulement sur un trait perceptif d'intérêt.

H2 : La constitution d'une trace en mémoire entre les composants audiovisuels devrait améliorer la détection d'orientation des traits visuels. Les sujets placés dans le contexte multi-sensoriel devraient devenir de plus en plus performants au fil de la tâche comparée aux sujets placés dans le contexte unimodal.

Méthode

Sujets

Vingt sujets ($M = 21.8$; $ET = 3.2$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques ni sensoriels.

Stimuli

Stimuli visuels

Les stimuli visuels pouvaient être de deux sortes : un trait vertical ou un trait oblique (voir Annexes, page 214). Ces traits de couleur noire apparaissaient au centre d'un carré de couleur noire de 7,5 cm de côté, présenté au centre de l'écran. Les traits mesuraient 7 mm de longueur pour 0.07 mm de largeur. Le trait oblique était incliné de -18° par rapport au trait vertical.

Stimuli auditifs

Nous avons utilisé pour cette expérience deux stimuli auditifs, chacun d'une durée de 500 ms. Un des deux stimuli possédait une fréquence de 2.37Hz et pouvait être considéré comme légèrement plus aigu que le second stimulus qui possédait une fréquence de 2Hz. Ces deux stimuli possédaient une fréquence de digitization de 44100Hz et ont été construits à l'aide du logiciel D-SOUND PRO 3.5.1 pour Macintosh. Les stimuli auditifs étaient diffusés via les hauts parleurs internes d'un ordinateur Mac Book PRO.

Procédure et plan expérimental

L'expérience comportait 40 stimuli et les sujets devaient effectuer une tâche de détection d'orientation de traits visuels, qui pouvaient être soit verticaux, soit obliques. Nous avons considéré comme facteur intra-sujet, le BLOC d'apprentissage avec deux modalités : le premier bloc constitué des 20 premiers essais et le second bloc constitué des 20 derniers essais. Nous avons également manipulé le CONTEXTE, facteur inter-sujet à deux modalités : multi-sensoriel ou unisensoriel. Un premier groupe de sujets réalisait la tâche de détection de traits verticaux ou obliques et un deuxième groupe de sujets réalisait cette même tâche alors qu'un son était systématiquement associé à l'apparition des traits visuels. Chaque trait visuel était associé systématiquement avec un seul trait auditif, par exemple le trait visuel vertical était associé au trait auditif aigu, alors que le trait visuel oblique était associé au trait auditif grave. L'association entre le type de trait visuel et le type de trait auditif a été contrebalancée.

Chaque essai débutait par une croix de fixation d'une durée de 500 ms, le stimulus apparaissait ensuite pendant 500ms, enfin un écran blanc d'une durée de 1500 ms apparaissait pendant lequel les sujets pouvaient encore répondre au stimulus. Les sujets devaient appuyer

sur une touche lorsqu'ils détectaient l'orientation verticale et sur une autre touche pour l'orientation oblique. Les touches de réponse étaient « a » et « p », elles ont été contrebalancées. Dans la condition de contexte multi-sensoriel, il n'était donné aucune indication aux sujets concernant les sons. Il était demandé aux sujets d'essayer de répondre le plus rapidement possible sans commettre trop d'erreurs.

Résultats

Un premier filtrage a été fait en supprimant les latences supérieures à 1500 ms. Ensuite seuls les temps de réponse s'écartant de plus ou moins deux écarts types de la moyenne, par condition, ont été conservés (voir TAB. 1). Seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse, moins de 5% des temps de réponse ont été écartés de l'analyse statistique.

		BLOC	
		Bloc 1	Bloc 2
CONTEXTE	multisensoriel	543 22	505 22
	unisensoriel	637 39	539 19

TAB. 1 – Moyennes des temps de réponses et erreurs types en fonction des facteurs BLOC et CONTEXTE

L'analyse de variance conduite sur les latences de réponses montre un effet significatif du facteur BLOC, $F(1,18) = 22.35$; $p < .01$; $\eta^2 = .55$, qui témoigne de l'amélioration des performances au cours de la tâche. Conformément à notre hypothèse, l'analyse de variance révèle un effet tendanciel pour le facteur CONTEXTE, $F(1,18) = 3.32$; $p = 0.08$; $\eta^2 = .16$, qui souligne l'aspect bénéfique du contexte multi-sensoriel de l'apprentissage par rapport au contexte unisensoriel, bien que le trait auditif ne soit pas traité explicitement par les sujets. On note également un effet tendanciel de l'interaction BLOC*CONTEXTE, $F(1,18) = 4.19$; $p = 0.055$; $\eta^2 = .19$ (voir FIG. 15), qui suggère une différence entre l'évolution des performances des deux groupes d'apprentissage. Bien que les temps de réponse diminuent de manière significative entre les deux blocs quelle que soit la condition, il est intéressant de noter que les performances s'améliorent plus pour les sujets placés dans la condition d'apprentissage unisensoriel (98ms) par rapport aux sujets placés dans la condition d'apprentissage multi-sensoriel (39 ms), $t(18) = 2.05$; $p = .05$. Si l'on considère séparément le premier et le deuxième bloc d'apprentissage, il existe une différence significative pour le contexte uniquement pour le premier bloc. On note en effet des temps de réponse plus rapide pour le groupe d'apprentissage sonore (543 ms) comparé au groupe d'apprentissage non sonore (637

ms), $F(1,18) = 4.27$; $p = .027$. On ne retrouve pas de différences significatives dans le second bloc d'apprentissage.

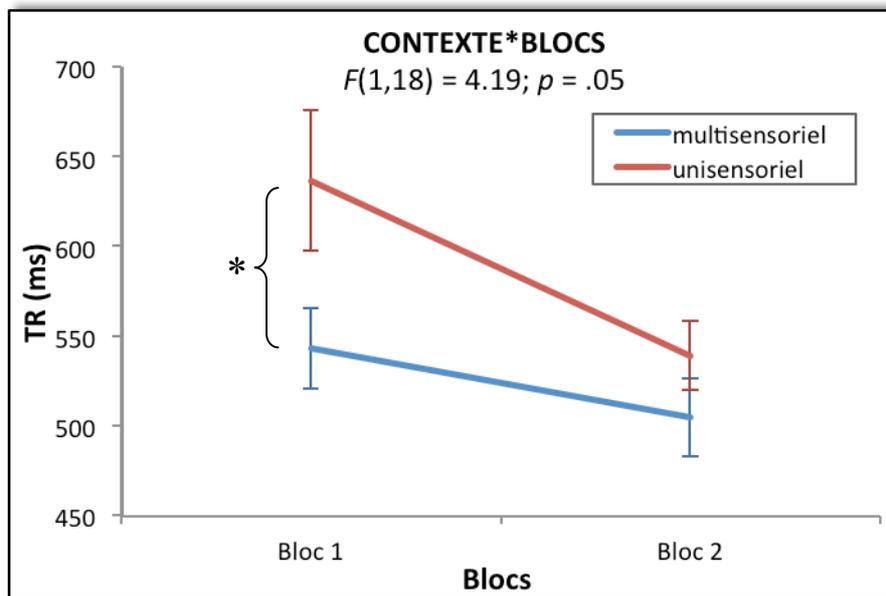


FIG. 15 - Représentation graphique de l'interaction des facteurs CONTEXTE et BLOCS sur les temps de réponse (ms)

Discussion

Les résultats semblent indiquer que les performances des sujets à juger de l'orientation des traits visuels s'améliorent globalement avec l'apprentissage. Bien que les sujets n'aient pas à traiter les sons, nous avons montré que les performances étaient meilleures en multimodal qu'en unimodal mais uniquement pour le BLOC 1. Ainsi, nous avons mis en évidence que la pente d'apprentissage était plus forte dans la condition unimodale. Cet effet peut s'expliquer par le fait que les sujets placés dans un contexte multi-sensoriel bénéficient dès le début de l'apprentissage des effets facilitateurs du trait orthogonal par rapport aux sujets placés dans la condition de contexte unisensoriel. Pour les sujets placés dans le groupe d'apprentissage non sonore, des temps de réponse plus lents dans le premier bloc permettent une amélioration des performances plus importantes dû à l'effet de pratique. Le contexte multi-sensoriel serait donc globalement meilleur que le contexte unisensoriel, avec la présence de performances significativement meilleures et qui s'améliorent au fil de l'apprentissage. On note toutefois que les performances liées au contexte unisensoriel sont équivalentes en fin d'apprentissage, ce qui laisse suggérer que ce type d'apprentissage peut être aussi efficace qu'un entraînement multi-sensoriel mais est plus long à se mettre en place. Seitz (2006) montrent que les données obtenues sur la première session d'entraînement d'un apprentissage sur plusieurs jours suivent

un pattern similaire à celui que nous avons obtenu. Les performances sont meilleures au cours de l'apprentissage pour le groupe multi-sensoriel mais ne se différencient plus du groupe unisensoriel à la fin de la première session d'entraînement. Bien que la procédure expérimentale utilisée ici soit différente de celle de Seitz et collaborateurs, il semble qu'un entraînement unisensoriel parvienne au même niveau de performance qu'un entraînement multi-sensoriel, mais celui-ci reste plus efficace, car plus rapidement performant qu'un entraînement unisensoriel. Selon le modèle Act-in (Versace *et al.*, 2009), l'association systématique de deux composants sensoriels aboutirait à la création d'une trace mnésique/sensorielle même si l'attention n'est pas dirigée intentionnellement sur les deux composants sensoriels, et le processus de création de la trace serait précoce, dès les premiers items et pourrait avoir une influence ultérieure sur une autre tâche (Brunel *et al.*, 2009; Brunel *et al.*, 2010).

Une question semble émerger suite à cette expérience préliminaire, il est en effet intéressant de connaître la nature spécifique ou non de l'intégration multimodale. En effet l'amélioration des performances dans la condition multi-sensorielle dans cette première expérience pourrait être liée à la présence d'un contexte bimodal sans que l'association systématique entre un composant visuel et un composant auditif ne soient prédominante et n'ait entraîné la création d'une trace mnésique/sensorielle. On peut se poser maintenant la question de la spécificité de ces interactions multi-sensorielles sur l'apprentissage, en d'autres termes, l'association systématique d'un composant visuel et d'un composant auditif orthogonal aboutit-elle à la création d'une trace mnésique/sensorielle intégrée ? Pour répondre à cette question nous avons conduit une deuxième expérience dans laquelle un groupe de sujets devait juger de l'orientation de traits visuels associés systématiquement à des sons (barre verticale + son aigu et barre oblique + son grave), alors qu'un autre groupe de sujets réalisait la même tâche mais l'association entre traits visuels et auditifs n'était plus systématique, mais aléatoire.

Expérience 2

L'amélioration des performances dans le contexte multi-sensoriel dépend plus de la relation qu'entretiennent les traits perceptifs entre eux que de la simple présence conjointe de plusieurs modalités sensorielles simultanées. Un apprentissage où des modalités sensorielles sont systématiquement associées entre elles devrait entraîner la création d'une trace mnésique/sensorielle spécifique.

Hypothèses

H1 : Un apprentissage systématique devrait entraîner un gain supérieur à un apprentissage où l'association entre les modalités sensorielles n'est pas systématique.

Méthode

Sujets

Quatre-vingt sujets ($M = 21.9$; $ET = 5.4$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques ni sensoriels.

Stimuli

Stimuli visuels

Les stimuli visuels utilisés dans cette deuxième expérience étaient identiques à ceux utilisés lors de l'expérience 1 (voir page 69).

Stimuli auditifs

Les stimuli auditifs utilisés dans cette deuxième expérience étaient identiques à ceux utilisés lors de l'expérience 1 (voir page 69).

Procédure et plan expérimental

L'expérience comportait 60 stimuli afin d'augmenter le nombre d'items dans chaque bloc par rapport à la première expérience. Nous avons considéré comme facteur intra-sujet, le facteur BLOC d'apprentissage avec deux modalités : un premier bloc constitué des 30 premiers essais et un second bloc constitué des 30 derniers essais. Nous avons également manipulé l'ASSOCIATION, facteur inter-sujet, à deux modalités : association systématique ou association non systématique. Pour un premier groupe de sujets, le trait vertical était associé systématiquement au son aigu et le trait oblique était systématiquement associé au son grave, l'association entre les traits visuels et auditifs a été contrebalancée au sein des sujets du groupe d'association systématique. Pour un autre groupe, le trait vertical et le trait oblique étaient associés dans 50% des essais avec le son aigu et dans 50% des essais avec le son grave. Les touches de réponse ont été contrebalancées au sein des deux groupes.

Les sujets devaient effectuer la même tâche de détection d'orientation de traits que dans l'expérience précédente. La consigne donnée aux sujets des deux groupes ne comportait

aucun élément sur les stimuli auditifs, les sujets devaient uniquement catégoriser les traits visuels. La procédure était rigoureusement identique à celle employée lors de l'expérience 1.

Résultats

Nous avons réalisé une analyse de variance à mesures répétées sur les latences des réponses. Les latences supérieures à 1500 ms n'ont pas été prises en compte, de même que les temps de réponse s'écartant de plus ou moins deux écarts types de la moyenne. Seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse (voir TAB. 2), moins de 5% des temps de réponse ont été écartés de l'analyse statistique. Trois sujets ont été écartés de l'analyse, un pour ne pas avoir respecté les consignes et deux sujets pour avoir totalisé moins de 85% de bonnes réponses sur l'ensemble de la tâche.

		BLOC	
		Bloc 1	Bloc 2
ASSOCIATION	systematique	503 23	448 16
	non systematique	489 18	461 15

TAB. 2 – Moyennes des temps de réponses et erreurs types en fonction des facteurs BLOC et ASSOCIATION

L'analyse de variance conduite sur les latences de réponses montre un effet significatif du facteur BLOC, $F(1,75) = 37.02$; $p < .01$; $\eta^2 = .33$, qui montre l'amélioration des performances au cours de la tâche. On note également un effet significatif de l'interaction BLOC*ASSOCIATION, $F(1,75) = 4$; $p = 0.049$; $\eta^2 = .05$ (voir FIG. 16), qui suggère une différence dans l'amélioration des performances en fonction du groupe d'apprentissage. Conformément à notre hypothèse, il semble que l'amélioration des performances pour le groupe d'association systématique soit meilleure que celle du groupe d'association non systématique. En effet, la comparaison du bénéfice d'apprentissage montre que la pente du groupe d'association systématique (55 ms) est plus importante que celle du groupe d'association non systématique (28ms), $t(75) = 1.97$; $p = 0.026$. L'analyse de variance ne révèle par contre aucun effet du facteur ASSOCIATION, $F(1,75) < 1$.

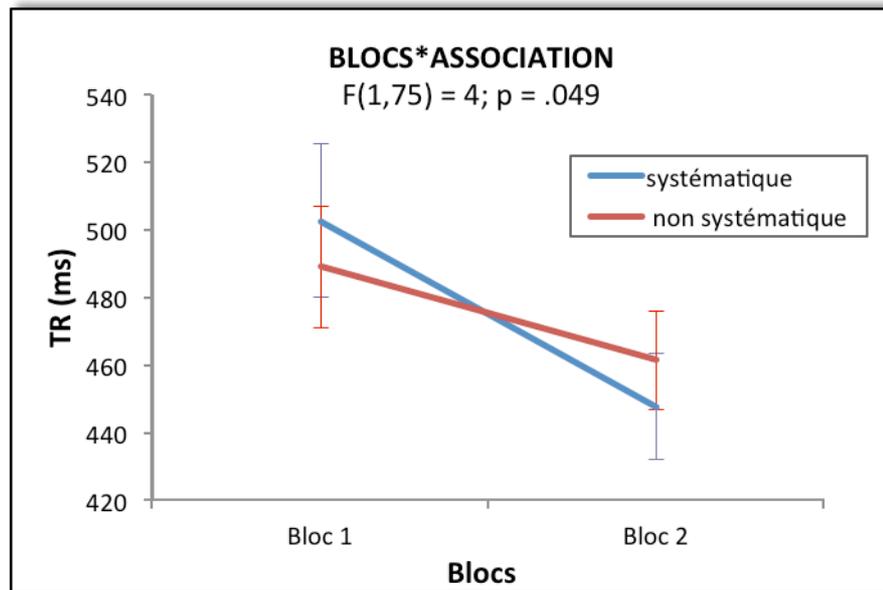


FIG. 16 – Représentation graphique de l'interaction entre les facteur BLOCS et ASSOCIATION sur les temps de réponse (ms)

Discussion

Dans cette deuxième expérience, nous avons voulu tester le caractère spécifique de l'association systématique entre un trait visuel et un trait auditif orthogonal. Les résultats montrent que l'association systématique entre deux composants perceptifs, même si l'un des deux ne doit pas être traité, est meilleure qu'une association non systématique. Toutefois il est important de préciser que la taille de l'effet que nous recherchions est faible ($\eta^2 = .05$), ce résultat doit donc être considéré avec beaucoup de prudence. Nous avons pris pour postulat que la présentation simultanée de deux traits perceptifs entraînerait la création d'une trace spécifique au sein de laquelle les stimuli visuels et auditifs seraient associés. Il serait peut être un ambitieux d'affirmer que l'association systématique a entraîné la création d'une trace sensorielle/mnésique. Toutefois, on peut attester avec prudence qu'un contexte multi-sensoriel spécifique augmente l'amélioration des performances par rapport à un contexte multi-sensoriel non spécifique. Il ne semble pas contradictoire de penser que la congruence entre les traits sensoriels visuels et auditifs aient créé un contexte favorable pour un apprentissage plus performant. On note également que le gain est plus important dans l'expérience 2 que dans l'expérience 1 pour le groupe ayant bénéficié d'un apprentissage multimodal systématique. Il se peut que le nombre différent de sujets impliqués dans les deux expériences puisse expliquer cette variabilité, il faut également noter que le nombre de stimuli n'est pas le même entre la première et la deuxième expérience, ce qui peut également expliquer cette différence.

Dans la littérature sur l'intégration multi-sensorielle, certains auteurs tendent à dire que se baser sur les moyennes des temps de réponse pour attester de la présence ou non d'une intégration multi-sensorielle est insuffisant (Raab, 1962). Il serait donc nécessaire pour aller plus loin que les présents résultats, d'utiliser une méthode permettant de vérifier la présence d'un gain multi-sensoriel. Le Race Model (Miller, 1982) est un outil mathématique qui permet justement de vérifier que le gain obtenu ne soit pas simplement le résultat d'une facilitation statistique. Nous avons donc cherché dans une troisième expérience à valider les résultats obtenus en utilisant le Race Model, en modulant l'association entre les stimuli visuels et auditifs. Nous avons conservé deux groupes, un avec association systématique entre une cible auditive et une cible visuelle et un groupe sans association. L'intérêt d'utiliser le Race Model est double : quantifier l'intégration multi-sensorielle ou le gain produit par chacune de ces associations audiovisuelles, mais également comparer ces deux groupes. En effet l'intérêt est de pouvoir savoir s'il existe dans un groupe, un gain plus important que dans l'autre groupe.

S'il se crée une trace dans une condition et non dans l'autre, on est en mesure de penser que cela va avoir un impact sur les paramètres du Race Model. On peut donc attendre : une différence entre les deux groupes au niveau de la significativité de l'intégration multi-sensorielle (1), on peut également observer une différence dans l'amplitude du gain (2) ou alors entre les courbes de violation de l'inégalité de Miller (3). On pourrait éventuellement poser l'hypothèse que la création d'une trace mnésique/sensorielle pourrait se traduire par un bénéfice plus important que lorsque le contexte est seulement multi-sensoriel. Ce bénéfice devrait s'observer selon un des trois critères définis précédemment.

Toutefois l'utilisation du Race Model nous force à modifier le paradigme utilisé jusque là. En effet, ce modèle nécessite des temps de réponse dans les conditions visuelles, auditives et audiovisuelles, or jusqu'à présent les sujets ne devaient traiter que le composant visuel. Ce qui nous intéresse est avant tout la constitution d'une trace audiovisuelle, quelle que soit la manière de traiter les différentes composantes de la trace. Nous avons montré dans cette expérience le caractère automatique de l'intégration multi-sensorielle et la présence d'un trait orthogonal favorisait la création d'une trace mnésique/sensorielle car la performance des sujets était meilleure dans le cas d'une association systématique que non systématique. Nous ne modifions dans l'expérience suivante que le traitement des composants et non l'association entre eux.

Expérience 3

Hypothèses

H1: Le gain intégratif mesuré par le Race Model devrait être supérieur pour les sujets placés dans la condition d'association systématique comparé aux sujets placés dans la condition d'association non systématique.

Méthode

Sujets

Vingt-quatre sujets ($M = 22.3$; $ET = 4.7$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques et possédaient une vision normale ou corrigée.

Stimuli

Stimuli visuels

Nous avons utilisé deux formes, un cercle gris de 1,88cm de rayon et un carré de même couleur de 3,5cm de côté, de telle sorte qu'ils aient environ la même surface (voir Annexes, page 214). Ces formes géométriques étaient présentées pendant 500ms et pouvaient apparaître à gauche ou à droite de l'écran. Ces stimuli visuels sont repris de l'expérience de Brunel et collaborateurs (2009), nous avons divisé leur taille par 2, comme la présentation était latéralisée, les items visuels paraissaient ainsi moins saillants.

Stimuli auditifs

Les stimuli auditifs étaient identiques à ceux utilisés lors de la première expérience (voir page 69). La durée des sons était de 500ms et ils pouvaient être diffusés soit dans l'oreille gauche soit dans l'oreille droite du sujet via un casque SENNHEISER HD 457.

Stimuli audiovisuels

Les items audiovisuels étaient composés des stimuli visuels et auditifs présentés simultanément. La durée de présentation de l'image et du son était toutes les deux de 500 ms. Le mode d'apparition des stimuli audiovisuels était toujours congruent, si l'image était présentée du côté gauche de l'écran, alors le son était diffusé dans l'oreille gauche du sujet et vice versa.

Procédure et plan expérimental

L'expérience comportait 192 stimuli (voir TAB. 3) répartis en quatre blocs de 48 stimuli chacun. La liste de présentation des items au sein de chaque bloc était la même et l'ordre des items dans la liste était déterminé de manière semi-aléatoire afin d'éviter que plusieurs items d'une même modalité se suivent. Dans cette expérience, Nous avons manipulé en intra-sujet, le facteur MODALITE, avec trois conditions : visuelle, auditive et audiovisuelle. Selon les conditions, seuls les items visuels, auditifs ou audiovisuels apparaissaient, dans 50% des cas à droite et dans 50% des cas à gauche de la croix de fixation. Nous avons également manipulé, en inter-sujet, le facteur ASSOCIATION, avec deux conditions : avec association et sans association. Dans le groupe avec association, les sujets réalisaient la tâche où un type de stimulus visuel était systématiquement associé à un type de stimulus auditif (cercle+son aigu et carré+son grave). L'association entre un type de stimulus visuel (carré ou cercle) et le type de son (aigu ou grave) était contrebalancée entre les sujets. Dans le groupe sans association, les sujets réalisaient la tâche dans laquelle un stimulus visuel était associé dans 50% des cas avec un son et dans 50% des cas avec l'autre son. Les touches de réponse étaient contrebalancées dans les deux groupes.

ASSOCIATION	Avec association/Sans association					
	Visuelle		Auditive		Audiovisuelle	
MODALITE	Gauche	Droite	Gauche	Droite	Gauche	Droite
LATERALITE						
Nombre d'items	32	32	32	32	32	32

TAB. 3 – Répartition des items de l'expérience en fonction des facteurs ASSOCIATION et MODALITE et de la position d'apparition

La présentation latéralisée des items était en lien avec la tâche du sujet, en effet la consigne était de détecter le plus vite possible sans faire trop d'erreurs la position (gauche ou droite) des items. Comme nous n'avons pas d'hypothèses concernant un effet d'intégration multi-sensorielle en lien avec la présentation latéralisée à gauche ou à droite des stimuli, nous n'avons pas considéré ce facteur dans l'analyse.

Chaque essai débutait par une croix de fixation d'une durée de 500ms, puis le stimulus apparaissait pendant 500ms suivi d'un écran blanc d'une durée de 1500ms pendant laquelle le sujet pouvait encore répondre au stimulus. La consigne donnée aux sujets était de détecter le plus rapidement le côté d'apparition du stimulus à l'aide de deux touches du clavier « a » et « p », une sur la gauche et une sur la droite du sujet. Il était indiqué au sujet d'essayer d'aller le plus rapidement possible et de répondre à tous les items visuels, auditifs et audiovisuels.

Résultats

Analyse des temps de réponses et des pourcentages de bonnes réponses

Nous avons réalisé une analyse de variance à mesures répétées sur les latences des réponses. Les latences supérieures à 1500 ms n'ont pas été prises en compte, de même que les temps de réponse s'écartant de plus ou moins deux écarts types de la moyenne. Seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse (TAB. 4), moins de 5% des temps de réponse ont été écartés de l'analyse statistique. Toutes les données ont été retenues pour le calcul des pourcentages de bonnes réponses (voir TAB. 4).

		Modalité (TR)			Modalité (%BR)		
		V	A	AV	%V	%A	%AV
ASSOCIATION	Systématique	392	379	323	0,98	0,95	0,99
		20	22	18	0,01	0,01	0,01
	Non systématique	383	387	323	0,96	0,94	0,96
		13	24	11	0,02	0,02	0,02

TAB. 4 – Moyennes des temps de réponses et des pourcentages de réponses correctes et erreurs types en fonction des facteurs MODALITE et ASSOCIATION

L'analyse de variance conduite sur les latences de réponse met en évidence un effet du facteur MODALITE, $F(2,44) = 57.45$; $p < .01$, en effet les temps de réponse associés à la modalité audiovisuelle sont nettement plus rapide que ceux associés à la modalité auditive, $F(1,22) = 63.62$; $p < .01$, ou visuelle, $F(1,22) = 297.78$; $p < .01$. En revanche, l'analyse de variance ne montre aucun effet du facteur ASSOCIATION, $F(1,22) < 1$, ni d'interaction ASSOCIATION*MODALITE, $F(2,44) < 1$. Ces données suggèrent donc qu'il n'existe pas de différences au niveau du traitement des modalités entre les groupes d'association et de non association.

L'analyse de variance conduite sur les pourcentages de bonnes réponses révèle un effet du facteur MODALITE, $F(2,44) = 9.04$; $p < .01$, mais pas d'effet du facteur ASSOCIATION, $F(1,22) = 1.06$; $p = .31$, ni d'interaction ASSOCIATION*MODALITE, $F(2,44) < 1$.

Analyse des données du Race Model

Seules les données correspondant aux bonnes réponses données par les sujets ont été conservées, de plus les temps de réponse s'écartant de plus ou moins deux écart-type de la moyenne de chaque condition ont été exclus de l'analyse. Les courbes des fonctions de répartition visuelle, auditive et audiovisuelle ont été calculées ainsi que la courbe correspondant à la somme des fonctions de répartition unimodale pour chaque sujet. Ensuite 19 fractiles d'ordre 20 ont été calculés par condition pour chaque sujet, les fractiles ont ensuite été moyennés par groupe. Nous avons considéré ensuite la valeur de l'amplitude de la

violation de l'inégalité de Miller (Colonius & Diederich, 2006), ainsi que les courbes de violation pour chaque groupe (Laurienti, 2006). Tous les traitements ont été réalisés à l'aide du logiciel MATLAB 2007b. Pour cela, plusieurs programmes ont été créés afin de transformer les données comportementales et de tracer les courbes du Race Model (voir Annexes, pages 220 et 229). Une autre fonction de ce programme a été de calculer chaque vincentile pour les distributions audiovisuelles empiriques (AV) et théoriques (A+V). Un autre programme a enfin permis de calculer le coefficient d'intégration tel qu'il a été décrit par Colonius et Diederich (voir Annexes, page 217).

Les courbes obtenues montrent qu'il existe une violation de l'inégalité de Miller. En effet quel que soit le groupe, la fonction de répartition des TR audiovisuels est graphiquement supérieure à la fonction de répartition de la somme des TR unimodaux (voir FIG. 17). Les tests de student réalisés sur les six premiers percentiles pour chaque groupe indiquent qu'il existe une tendance à 20% pour le groupe association, $t(11) = 2.09$, $p = .06$ et pour le groupe sans association, $t(11) = 1.91$; $p = .08$. Cette différence devient significative pour les percentiles 25%-30% dans les deux groupes (voir TAB. 5).

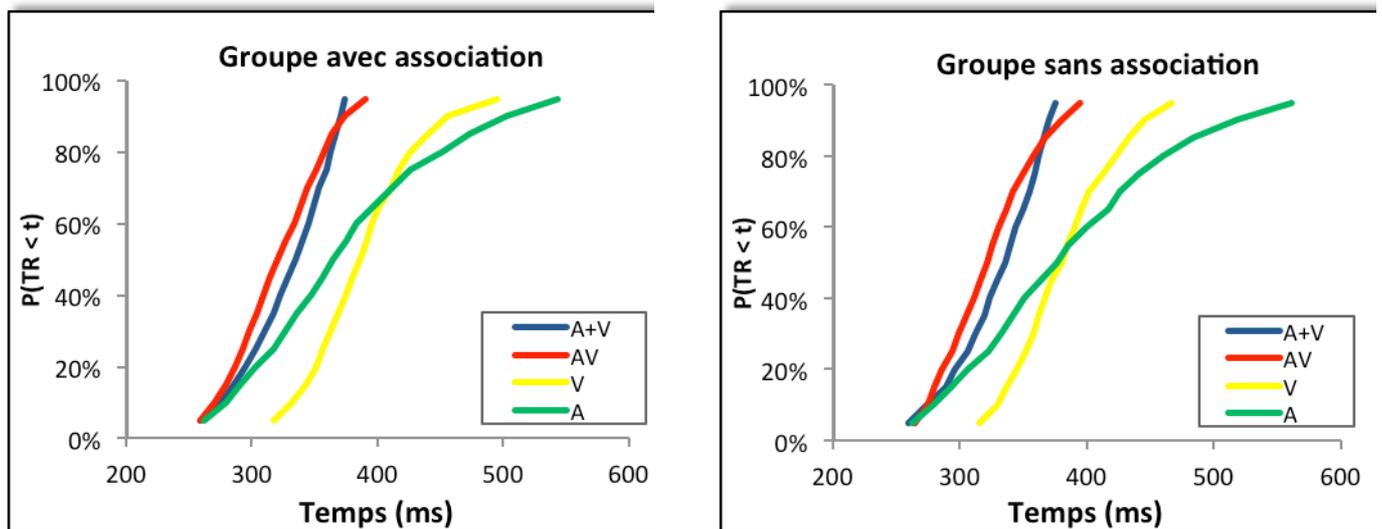


FIG. 17 – Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V) pour les sujets placés dans la condition « avec association » (à droite) et « sans association » (à gauche)

Percentile	t_{11}	p
5%	.71	.5
10%	1.26	.23
15%	1.55	.15
20%	2.09	.06
25%	3.11	.01
30%	3.62	.004

Percentile	t_{11}	p
5%	1.54	.15
10%	0.82	.43
15%	1.80	.1
20%	1.91	.08
25%	2.37	.04
30%	2.44	.03

TAB. 5 - Tableau de significativité des 6 premiers percentiles pour les groupes avec association systématique et sans association

Le calcul des valeurs d'approximation d'amplitude sous la courbe (Colonus & Diederich, 2006) ne montre aucune différence entre le groupe d'association systématique (5.46) et le groupe sans association (5.02), $p = .49$ (Mann-Whitney).

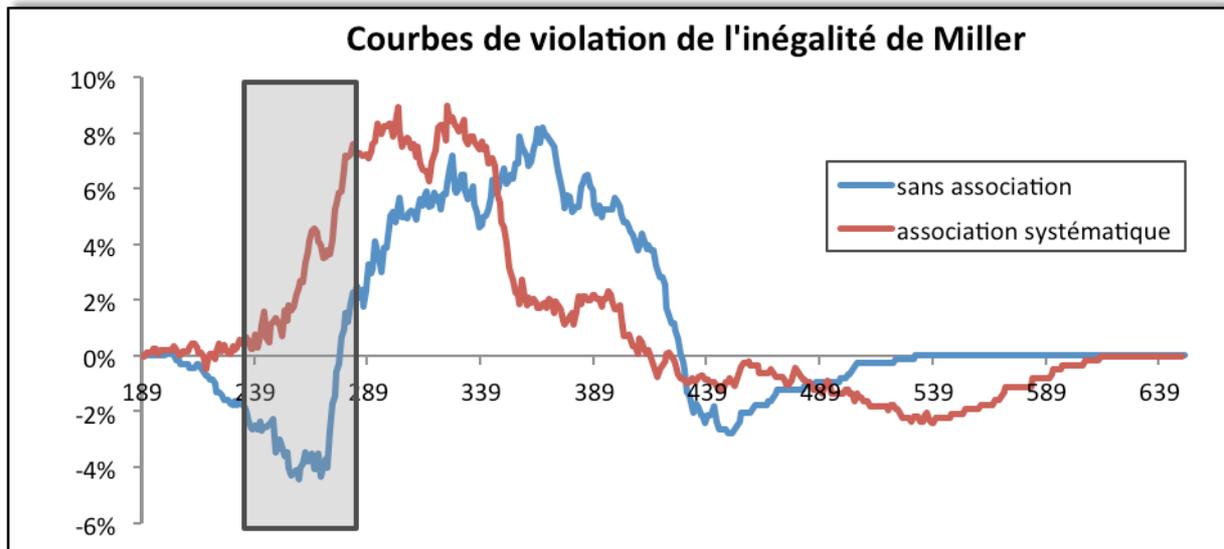


FIG. 18 - Représentation graphique des courbes de violation ($AV - (A+V)$) des sujets placés dans la condition « sans association » et des sujets placés dans la condition « avec association ». Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un $\alpha < .05$.

Cet indicateur ne nous permet donc pas de différencier les deux groupes de sujets en fonction de l'amplitude de la violation de l'inégalité de Miller.

L'analyse conduite sur la comparaison des fonctions de répartition du groupe d'association systématique et du groupe sans association montre une différence significative entre 235ms et 283ms après présentation du stimulus. Les deux groupes diffèrent donc sur une gamme de temps comprenant les deux premiers percentiles de chacune des distributions. Toutefois, il est important de noter que même si la courbe de violation du groupe d'association systématique est supérieure à celle du groupe sans association entre 235ms et 283ms, cette période ne correspond pas à la présence d'une intégration multi-sensorielle significative au regard de la vincentisation (voir FIG. 18).

Discussion

Cette expérience avait pour but, en modifiant le paradigme utilisé dans l'expérience 2 et en utilisant le Race Model, de mettre en évidence une différence entre les groupes d'association

systematique et le groupe sans association. Nous avons également émis l'hypothèse que la constitution d'une trace dans le groupe d'association systématique se caractériserait par une augmentation du gain intégratif par rapport au groupe sans association.

L'analyse des temps de réponse et des pourcentages de bonnes réponses ont montré des résultats classiquement observés dans la littérature, c'est-à-dire un avantage de la condition audiovisuelle sur les conditions unimodales (Giard & Perronnet, 1999 ; Molholm *et al.*, 2002 ; Fort *et al.*, 2002). Toutefois nous cherchions à faire un parallèle entre les résultats de l'expérience 2 et les résultats de cette expérience, et nous n'avons pas relevé d'effet du facteur ASSOCIATION, ni d'interaction entre les facteurs ASSOCIATION et MODALITE. Il semble donc, au regard des latences de réponses et des pourcentages de bonnes réponses que nous n'ayons pas retrouvé les résultats escomptés. Néanmoins, la modification du paradigme dans cette expérience pourrait en partie expliquer ces résultats. Une différence essentielle porte sur le fait que dans cette expérience la position du son était pertinente pour réaliser la tâche alors que le son avait moins d'utilité dans l'expérience 2.

Concernant les résultats du Race Model, nous avons observé une intégration multi-sensorielle significative (percentile 20%-30%) au sein des deux groupes expérimentaux. Ensuite, les deux groupes ont présenté un gain intégratif similaire sur l'ensemble de la distribution, en effet, le groupe d'association systématique (5.46) et le groupe sans association (5.02) ne diffèrent pas sur cette dimension. Enfin, en comparant les courbes de violation de l'inégalité de Miller, nous avons mis en évidence une différence entre les deux groupes, avec un avantage significatif du groupe d'association systématique sur le groupe sans association entre 235ms et 283ms. Cet avantage pourrait ainsi être le marqueur qui différencierait les deux conditions expérimentales et celui-ci se caractériserait sur les temps les plus courts. Cet avantage ne serait donc pas lié à un gain multi-sensoriel plus important, mais à des processus intégratifs plus précoces lorsque l'association est systématique. Nous avons utilisé une tâche de détection et il est possible que l'association entre les traits visuels et auditifs soit moins pertinente pour réaliser la tâche. On peut imaginer que la composante sémantique aurait un rôle plus important dans une tâche de catégorisation que dans une tâche de détection.

Il faut cependant relativiser ce résultat en rapprochant les données obtenues avec la vincentisation et celles obtenues avec les courbes de violation de l'inégalité de Miller. En effet la période de temps où la courbe du groupe d'association systématique est supérieure à celle du groupe sans association ne correspond pas à un percentile significatif au niveau de la

vincentization. Cette période recouvre en fait les deux premiers percentiles de chacune des deux distributions. Or, Kiesel (2007) met en évidence les risques de se focaliser sur le premier percentile (5%) et surtout de ne considérer qu'un seul percentile significatif pour certifier la présence d'une intégration multi-sensorielle. Suivant ces indications, il serait donc trop ambitieux de tirer des conclusions sur ces seuls percentiles bien que certains auteurs ne se réfèrent pas forcément à la vincentization pour attester de la présence d'intégration multi-sensorielle (Mozolic *et al.*, 2008a). Il nous semble en effet plus rigoureux, même si le Race Model est une méthode d'analyse déjà conservatrice, de prendre en compte différents indicateurs disponibles à partir de cette méthode pour conclure sur la nature des processus intégratifs mis en jeu dans telle ou telle tâche.

Enfin, la procédure expérimentale utilisée dans cette troisième expérience est peut être trop différente de celle des expériences précédentes pour pouvoir répondre au questionnement soulevé par les expériences 1 et 2. Les items visuels et auditifs présentés seuls ont peut être rendu l'association systématique moins robuste qu'elle ne pouvait l'être dans l'expérience 2, où pour chaque essai, un item visuel était renforcé par un item auditif.

Conclusion

Dans cette première partie, nous nous sommes intéressés à la façon dont une trace pouvait se créer à partir d'associations multi-sensorielles. Dans une première expérience, nous avons vu qu'un apprentissage multi-sensoriel incident était meilleur qu'un apprentissage unisensoriel mais le bénéfice de la multimodalité semblait marqué surtout en début d'apprentissage (Seitz, 2006). La présentation conjointe de modalités sensorielles distinctes facilitent le traitement d'une seule des modalités même si l'autre trait sensoriel est orthogonal (Vroomen & De Gelder, 2000). Afin de s'assurer de la spécificité de l'association audiovisuelle de l'expérience 1, nous avons conduit une deuxième expérience, dans laquelle nous avons comparé deux apprentissages audiovisuels. Un apprentissage était constitué d'associations systématiques entre un son et une forme alors que pour l'autre apprentissage, les formes et les sons n'étaient pas systématiquement associés. Bien que notre résultat soit à considérer avec beaucoup de prudence, il semble que lorsque l'association est systématique, le traitement soit d'autant plus facilité que la simple présence d'un contexte multimodal. La co-occurrence systématique d'un son et d'une forme semble faciliter l'apprentissage même lorsque l'un des traits est orthogonal et non proche du seuil de détection. Il est cependant difficile d'attester avec certitude que cette association a entraînée la création d'une trace en mémoire. Il aurait

été intéressant de réaliser une tâche d'amorçage à la fin de l'apprentissage pour voir si une véritable association avait été créée entre le trait visuel et le trait auditif. La troisième expérience que nous avons conduite avait pour but de mesurer le gain intégratif associé à la condition d'association systématique. Toutefois, nous n'avons pas mis en évidence de différences significatives entre nos deux groupes. Le paradigme utilisé est certainement trop éloigné de celui utilisé dans l'expérience 2 pour pouvoir faire directement l'analogie avec l'expérience 3. En effet si l'on regarde bien, ces deux expériences ne se ressemblent que très peu. Une détection spatiale était requise dans l'expérience 3 alors qu'une détection d'orientation d'items placés au centre était nécessaire dans l'expérience 2. Cependant, il semble que dans ce type de paradigme, il est difficile de mettre en place une association entre traits visuels et traits auditifs lorsque dans la même phase les items auditifs et visuels sont présentés seuls fréquemment. En effet, cela peut perturber la création d'une unité en mémoire où chaque trait sensoriel est indissociable des autres (Goldstone, 2000).

Tout en conservant comme outil d'analyse le Race Model dans la partie 2, nous allons étudier des associations sémantiques déjà créées en mémoire, ce qui permettra de s'abstraire d'un apprentissage préalable.

Partie 2

Dans cette partie, nous allons utiliser des stimuli signifiant pour étudier le lien entre modalités sensorielles en mémoire lorsque celles-ci font référence au même concept. Nous allons utiliser un outil mathématique, le Race Model (Miller, 1982) qui permet de savoir s'il existe un gain multi-sensoriel dans une condition et si ce gain observé est significatif. Nous utiliserons entre autre le paradigme de l'expérience 3 de la partie 1 et nous ferons varier plusieurs facteurs d'intérêt. Tout d'abord nous allons voir si la nature de la tâche influence la présence ou non de gain multi-sensoriel. Dans un premier temps, nous allons utiliser une tâche de catégorisation (expérience 1) dans laquelle nous allons comparer les performances d'un groupe de sujets jeunes et de sujets âgés. Nous allons essayer de comprendre pourquoi les sujets âgés présentent un gain intégratif plus important que les sujets jeunes dans ce type de tâche (Laurienti *et al.*, 2006). Ensuite nous comparerons les performances de sujets jeunes et de sujets âgés dans une tâche de détection (expérience 2), qui n'implique pas de processus mnésiques. Nous pensons que dans ce type de tâche, il ne devrait plus y avoir de différences entre ces deux groupes de sujets. Nous observerons ensuite l'effet de congruence sémantique entre items de modalités différentes (expérience 3). En effet, nous regarderons si dans une

tâche de détection, la congruence module le gain intégratif multi-sensoriel. Enfin, nous testerons l'hypothèse selon laquelle l'introduction d'une composante attentionnelle peut augmenter le gain multi-sensoriel dans une tâche de catégorisation (expérience 4)

Expérience 1 Tâche de catégorisation multimodale

Nous allons essayer de répliquer les résultats de Laurienti et collaborateurs (2006) dans cette première expérience. Nous avons comparé un groupe de sujets jeune à un groupe de sujets âgés dans une tâche de catégorisation, à l'aide du Race Model.

Hypothèses

H1 : Les temps de réponse devraient être plus rapides chez le groupe de sujets jeunes et l'on devrait obtenir des temps de réponse plus rapide en modalité audiovisuelle quelle que soit le groupe de sujets.

H2 : Le groupe de sujets âgés devrait présenter un gain intégratif mesuré par le Race Model plus important que celui observé chez les sujets jeunes.

Méthode

Sujets

Vingt quatre sujets ont participé à cette expérience. Douze sujets jeunes ($M = 22$; $ET = 2.7$), étudiants à l'Université Lyon II, ont participé à cette expérience. Douze sujets âgés ($M = 78$; $ET = 4.2$) ont participé à cette expérience, seuls les participants âgés ayant obtenu un MMSE (Folstein, Folstein & McHugh, 1975) supérieur à 28 ont été inclus dans l'étude. Tous les sujets de cette expérience étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

Les stimuli visuels

Les stimuli utilisés dans cette expérience étaient les mêmes que ceux utilisés par Brunel et collaborateurs (2009). Les items visuels étaient constitués de deux formes géométriques présentées en nuances de gris : un carré et un cercle. Le carré mesurait 7cm de côté et le cercle mesurait 3,66cm de rayon, ils étaient construits de telle sorte que chacun ait environ la même surface (voir Annexes, page 215). Ces deux formes étaient présentées alternativement

au centre de l'écran pendant 500 ms. Les stimuli visuels ont été construits à l'aide du logiciel Power Point pour Mac.

Stimuli auditifs

Nous avons utilisé deux stimuli auditifs : la verbalisation du mot « carré » et la verbalisation du mot « cercle ». Chaque son possédait la même durée de 500ms. Les stimuli auditifs ont été construits à partir du logiciel Audacity pour Mac.

Stimuli audiovisuels

Les stimuli audiovisuels étaient constitués de la combinaison des items visuels et auditifs, de telle sorte que chaque combinaison soit toujours congruente. Les associations audiovisuelles étaient présentées pendant 500ms.

Procédure et plan expérimental

L'expérience comportait 192 stimuli répartis en quatre blocs de 48 stimuli chacun. La liste de présentation des items au sein de chaque bloc était la même et l'ordre des items dans la liste était déterminée de manière semi-aléatoire afin d'éviter que plusieurs items d'une même modalité se suivent. Nous avons manipulé en intra-sujet le facteur MODALITE, avec trois conditions : visuelle, auditive et audiovisuelle. Tous les items apparaissaient au centre de l'écran. Nous avons manipulé le facteur GROUPE avec deux conditions : un groupe de sujets jeunes et un groupe de sujets âgés.

Chaque essai débutait par une croix de fixation d'une durée de 500ms, puis le stimulus apparaissait pendant 500ms suivi d'un écran blanc d'une durée de 1500ms pendant laquelle le sujet pouvait encore répondre au stimulus. La consigne donnée aux sujets était de catégoriser le plus rapidement le stimulus, en répondant « cercle » ou « carré », à l'aide de deux touches du clavier « a » et « p ». Les touches de réponse étaient contrebalancées chez tous les sujets.

Résultats

Analyse des temps de réponses et des pourcentages de bonnes réponses

Les latences s'écartant de plus ou moins deux écarts types de la moyenne par condition n'ont pas été considérées. De plus, seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse (voir TAB. 6), moins de 5% des temps de réponse ont été écartés de l'analyse statistique. Toutes les données ont été retenues pour le calcul des pourcentages de bonnes réponses (voir TAB. 6).

		Modalité (TR)			Modalité (%BR)		
		V	A	AV	%V	%A	%AV
GROUPE	Agés	674 27	685 19	598 20	0,94 0,02	0,96 0,01	0,95 0,01
	Jeunes	552 32	598 42	523 31	0,96 0,01	0,99 0,001	0,98 0,01

TAB. 6 – Moyennes des temps de réponses et des pourcentages de réponses correctes et erreurs types en fonction des facteurs GROUPE et MODALITE

L'analyse de variance sur les latences de réponse met en évidence un effet du facteur GROUPE, $F(1,22) = 23.4$; $p < .01$; $\eta^2 = .52$, qui témoigne du ralentissement général présent chez les sujets âgés. On note également un effet du facteur MODALITE, $F(2,44) = 45.4$; $p < .01$; $\eta^2 = .67$. Les items audiovisuels sont catégorisés plus rapidement que les items visuels, $F(1,22) = 57.32$; $p < .01$ et plus rapidement que les items auditifs, $F(1,22) = 94.5$; $p < .01$. Il n'existe pas d'interaction GROUPE*MODALITE, $F(2,44) = 2.2$; $p = .12$.

Concernant les taux de réponses correctes, l'analyse de variance met en évidence un effet du facteur GROUPE, $F(1,22) = 5.2$; $p = .03$; $\eta^2 = .19$, ce qui atteste de la plus grande précision des sujets jeunes quelle que soit la modalité de réponse. On note également un effet du facteur MODALITE, $F(2,44) = 4.7$; $p = .014$; $\eta^2 = .18$. Il n'existe pas d'interaction GROUPE*MODALITE, $F(2,44) < 1$.

Analyse des données du Race Model

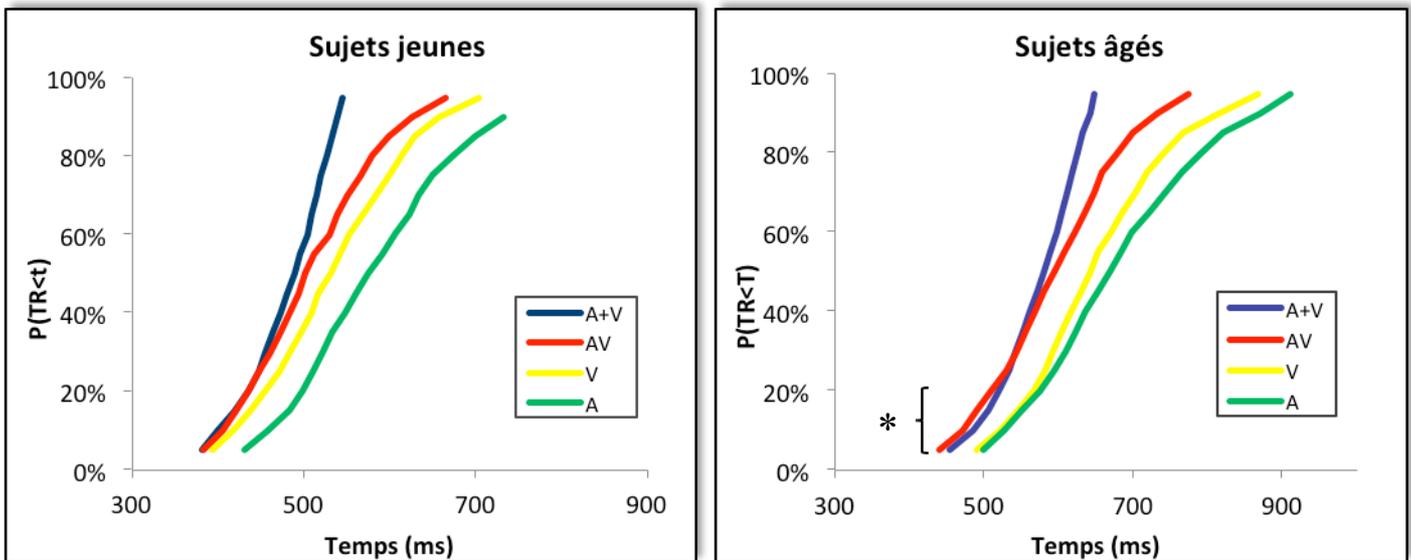


FIG. 19 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V) pour les sujets jeunes et les sujets âgés

Percentile	t_{11}	p
5%	1.04	.32

Percentile	t_{11}	p
5%	2.64	.02

10%	1.38	.2	10%	2.16	.05
15%	.87	.41	15%	2.6	.02
20%	.71	.49	20%	2.25	.04
25%	.9	.39	25%	1.22	.25
30%	1.54	.15	30%	.73	.48

TAB. 7 - Test de Significativité des six premiers percentiles pour la distribution des sujets jeunes et celle des sujets âgés. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués italique.

L'analyse des données du Race Model montre qu'il existe une violation de l'inégalité de Miller uniquement chez les sujets âgés (voir FIG. 19). Cette violation est significative pour les quatre premiers percentiles de la distribution, ce qui témoigne de la présence d'une intégration multi-sensorielle dans cette tâche de catégorisation (voir TAB. 7). A l'inverse, les sujets jeunes ne présentent aucun gain intégratif dans cette tâche, il n'existe en effet à aucun moment de violation de l'inégalité de Miller pour ce groupe de sujets.

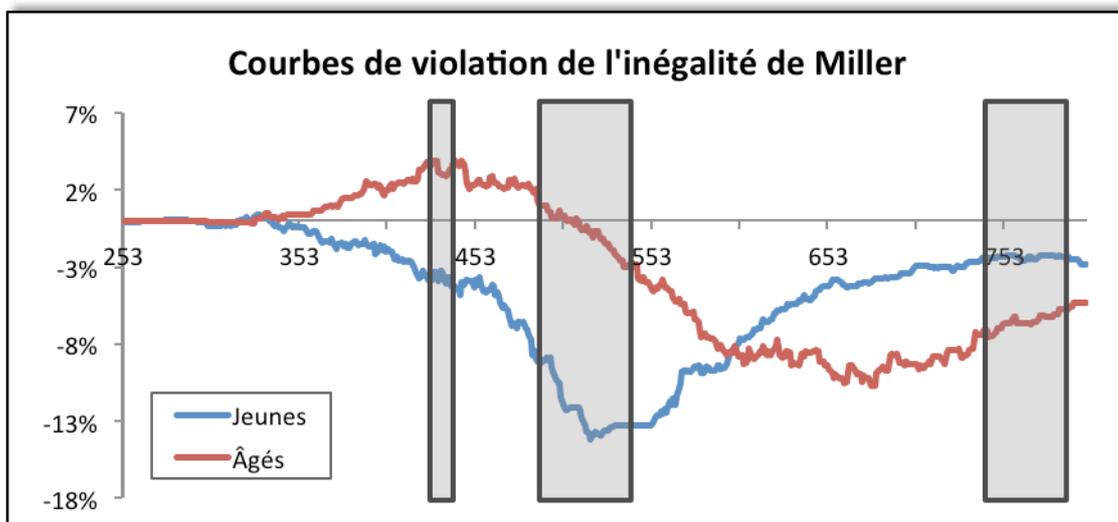


FIG. 20 - Représentation graphique des courbes de violation (AV-A+V) des sujets jeunes et des sujets âgés. Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un alpha < .05.

L'étude de la différence d'amplitude des courbes de violation chez les sujets jeunes et âgés montre que le gain intégratif est significativement supérieur pour les personnes âgées pour trois fenêtres temporelles (voir FIG. 20). La présence de valeurs négatives pour les courbes de violation signifie que la courbe audiovisuelle (AV) était inférieure à la courbe audiovisuelle théorique (A+V). Par conséquent, la moyenne des différences de ces deux courbes produit des valeurs négatives.

Discussion

Cette expérience avait pour but de répliquer les résultats de l'expérience de Laurienti et collaborateurs (2006) dans la condition de congruence sémantique multimodale. Nous avons répliqué en partie les résultats attendus car nous ne retrouvons pas dans notre expérience de gain intégratif chez les sujets jeunes. Toutefois les sujets âgés présentent un gain intégratif non nul et supérieur à celui des sujets jeunes comme attendu.

Il est toutefois étonnant que les sujets jeunes n'intègrent pas dans cette tâche de catégorisation. Selon nos hypothèses, une tâche de catégorisation fait intervenir des processus mnésiques qui peuvent difficilement mettre en évidence un gain tel qu'il est mesuré par le Race Model. En effet la catégorisation d'un item sur la base d'un composant sensoriel (e.g., carré visuel) doit activer les composants sensoriels de la trace. Par conséquent, la présentation d'un composant audiovisuel devrait entraîner peu de bénéfices en plus que le composant visuel ou auditif seul. A l'inverse l'utilisation d'une tâche de détection ne nécessite pas de traiter le composant dans sa globalité pour réaliser la tâche, on peut donc penser qu'elle est moins coûteuse en processus mnésique. Dans cette condition, la présentation d'items auditifs et visuels devrait faire apparaître un gain multi-sensoriel chez les sujets jeunes. En effet, c'est le mécanisme d'interaction multimodale qui devrait intervenir dans cette condition et nous pensons que le Race Model est plus sensible à ce type de processus.

Dans l'expérience 2, nous allons comparer le gain intégratif entre des sujets jeunes et âgés dans une tâche de détection. Nous devrions mesurer à l'aide du Race Model des processus purement perceptifs dans ce type de tâche. Pour cela, nous allons reprendre le paradigme développé dans l'expérience 3 de la partie 1.

Dans cette nouvelle expérience, les sujets jeunes et âgés devraient présenter un gain intégratif significatif. De plus, la différence de gain entre les sujets jeunes et les sujets âgés devrait disparaître.

Expérience 2 Tâche de détection latéralisée sans congruence sémantique

Pour cette expérience, nous avons repris les résultats des sujets jeunes qui avaient réalisé l'expérience 3 de la partie 1 dans la condition « sans association » (voir page 78). Nous avons ensuite ajouté un groupe de sujets âgés sur la même tâche, décrite dans le chapitre précédent.

Hypothèses

H1 : Les temps de réponse devraient être plus rapides chez le groupe de sujets jeunes et l'on devrait obtenir des temps de réponse plus rapide en modalité audiovisuelle quelle que soit le groupe de sujets.

H2 : Contrairement à l'expérience 1, la différence de gain mesurée par le Race Model entre sujets jeunes et sujets âgés devrait disparaître dans cette tâche de détection. Toutefois, on devrait observer un gain significatif dans les deux groupes de sujets.

Méthode

Sujets

Vingt-quatre sujets ont participé à cette expérience. Douze sujets jeunes ($M = 22$; $ET = 3.4$), étudiants à l'Université Lyon II, ont participé à cette expérience. Douze sujets âgés ($M = 77$; $ET = 5.6$) ont participé à cette expérience, seuls les participants âgés ayant obtenu un score supérieur à 28 au MMSE (Folstein, Folstein & McHugh, 1975) ont été inclus dans l'étude. Tous les sujets de cette expérience étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

L'ensemble des stimuli visuels, auditifs et audiovisuels sont en tout point identique à ceux utilisés lors de l'expérience 3 de la partie 1 (voir page 77).

Procédure et plan expérimental

L'expérience comportait 192 stimuli répartis en quatre blocs de 48 stimuli chacun. La liste de présentation des items au sein de chaque bloc était la même et l'ordre des items dans la liste était déterminé de manière semi-aléatoire afin d'éviter que plusieurs items d'une même modalité se suivent. Dans cette expérience, Nous avons manipulé en intra-sujet, le facteur MODALITE, avec trois conditions : visuelle, auditive et audiovisuelle. Selon les conditions, seuls les items visuels, auditifs ou audiovisuels apparaissaient, dans 50% des cas à droite et dans 50% des cas à gauche de la croix de fixation. Nous avons également manipulé, en inter-sujet, le facteur GROUPE, avec deux conditions : un groupe de sujets âgés et un groupe de sujets jeunes. Les sujets réalisaient une tâche dans laquelle le stimulus visuel était associé dans 50% des cas avec un son et dans 50% des cas avec l'autre son. Les touches de réponse étaient contrebalancées dans les deux groupes.

La présentation latéralisée des items était en lien avec la tâche du sujet, en effet la consigne était de détecter le plus vite possible sans faire trop d'erreurs la position (gauche ou droite) des items. Comme nous n'avons pas d'hypothèses concernant un effet d'intégration multi-sensorielle en lien avec la présentation latéralisée à gauche ou à droite des stimuli, nous n'avons pas considéré ce facteur dans l'analyse.

Chaque essai débutait par une croix de fixation d'une durée de 500ms, puis ensuite le stimulus apparaissait pendant 500 ms suivi d'un écran blanc d'une durée de 1500ms pendant laquelle le sujet pouvait encore répondre au stimulus. La consigne donnée aux sujets était de détecter le plus rapidement le côté d'apparition du stimulus à l'aide de deux touches du clavier « a » et « p », une sur la gauche et une sur la droite du sujet. Il était indiqué au sujet d'essayer d'aller le plus rapidement possible et de répondre à tous les items visuels, auditifs et audiovisuels.

Résultats

Analyse des temps de réponses et des pourcentages de bonnes réponses

Nous représentons les résultats comportementaux des sujets jeunes par souci de compréhension pour le lecteur, ce sont les mêmes que ceux qui sont présentés dans le chapitre précédent (voir page 78).

Les latences s'écartant de plus ou moins deux écarts types de la moyenne par condition n'ont pas été considérées. De plus, seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse (voir TAB. 8), moins de 5% des temps de réponse ont été écartés de l'analyse statistique. Toutes les données ont été retenues pour le calcul des pourcentages de bonnes réponses (voir TAB. 8).

		Modalité (TR)			Modalité (%BR)		
		V	A	AV	%V	%A	%AV
GROUPE	Agés	540	488	414	0,98	0,96	0,98
		25	42	18	0,01	0,01	0,01
	Jeunes	383	387	323	0,96	0,94	0,96
		13	24	11	0,02	0,02	0,02

TAB. 8 – Moyennes des temps de réponses et pourcentages de réponses correctes et erreurs types en fonction des facteurs MODALITE et GROUPE

L'analyse de variance sur les latences de réponse met en évidence un effet du facteur GROUPE, $F(1,22) = 17.02$; $p < .01$; $\eta^2 = .44$, qui témoigne du ralentissement général présent chez les sujets âgés. On note également un effet du facteur MODALITE, $F(2,44) = 17.5$; $p < .01$; $\eta^2 = .44$. Les items audiovisuels (369ms) sont détectés plus rapidement que les items

visuels (462ms), $F(1,22) = 85$; $p < .014$, et que les items auditifs (438ms), $F(1,22) = 18.8$; $p < .01$. Il n'existe pas d'interaction GROUPE*MODALITE, $F(2,44) = 2.36$; $p = .11$.

Concernant les taux de réponses correctes, l'analyse de variance ne met pas en évidence d'effet du facteur GROUPE, $F(1,22) < 1$, les sujets âgés et les sujets jeunes présentent un taux de précision équivalent quelle que soit la modalité de réponse. On note un effet du facteur MODALITE, $F(2,44) = 6.5$; $p < .01$; $\eta^2 = .23$. Il n'existe pas d'interaction GROUPE*MODALITE, $F(2,44) < 1$.

Analyse des données du Race Model

Nous représentons les résultats issus de l'analyse faite avec le Race Model des sujets jeunes par souci de compréhension pour le lecteur, ce sont les mêmes que ceux présentés dans le chapitre précédent (voir page 78).

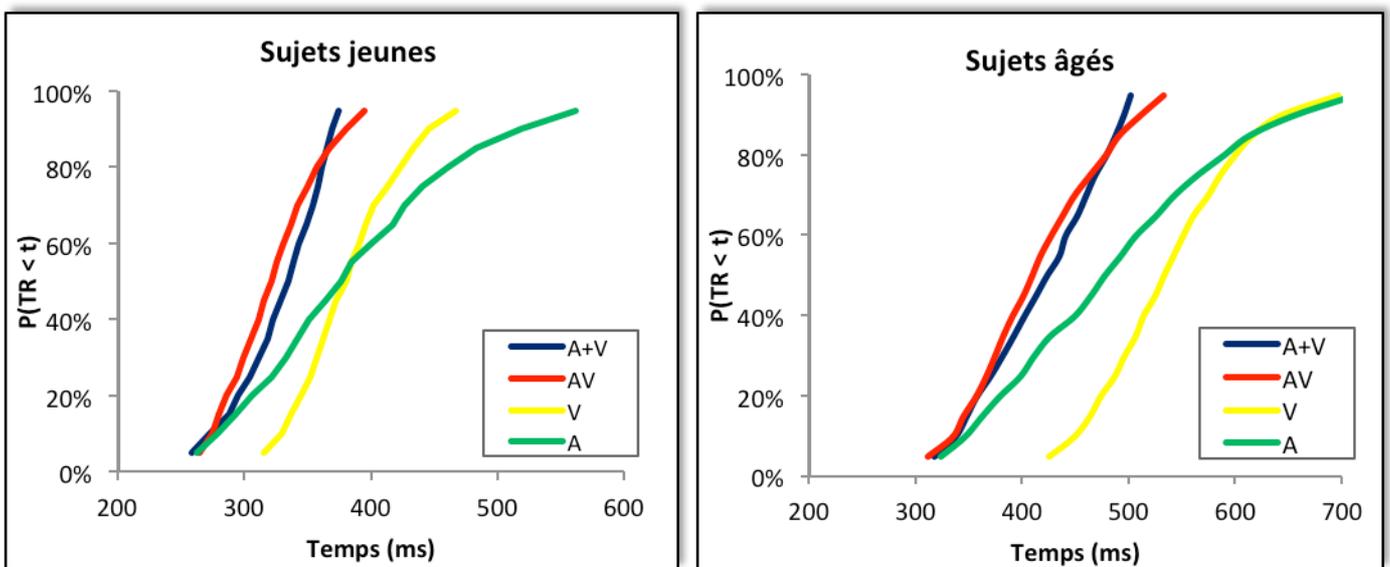


FIG. 21 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V) pour les sujets jeunes et les sujets âgés

Percentile	t_{11}	p	Percentile	t_{11}	p
5%	1.54	.15	35%	1.85	.09
10%	.82	.43	40%	2	.07
15%	1.49	.10	45%	2.14	.056
20%	1.91	.08	50%	2.49	.03
25%	2.37	.037	55%	2.88	.01
30%	2.44	.03	60%	2.21	.049

TAB. 9 - Test de Significativité des six premiers percentiles pour la distribution des sujets jeunes et du septième au douzième percentile pour la distribution des sujets âgés. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués en italique.

L'analyse des résultats obtenus à l'aide du Race Model montre qu'il existe un gain intégratif chez les deux groupes de sujets (voir FIG. 21). Toutefois, les premiers percentiles significatifs chez les sujets âgés sont plus tardifs (50%) que chez les sujets jeunes (25%) (voir TAB. 9).

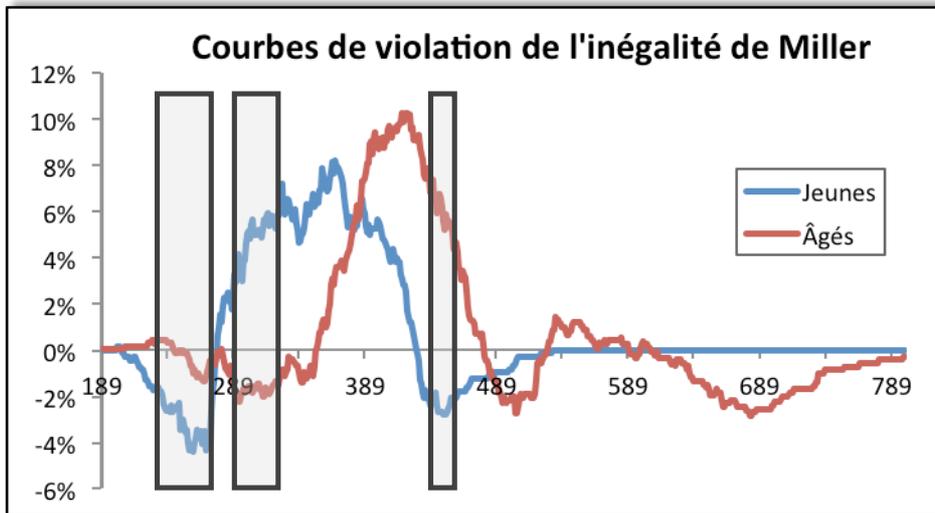


FIG. 22 - Représentation graphique des courbes de violation (AV-A+V) des sujets jeunes et des sujets âgés. Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un $\alpha < .05$.

Lorsque l'on observe les courbes de violation de l'inégalité de Miller, on observe des différences significatives entre sujets jeunes et sujets âgés mais qui semblent être causées par un décalage dans le temps de la distribution de la violation chez les personnes âgées (voir FIG. 22). Ces différences pouvant être liées uniquement au ralentissement des temps de réponse avec le vieillissement, nous avons recalé les courbes (voir FIG. 23) en prenant pour origine le moment où la violation de l'inégalité est effective.

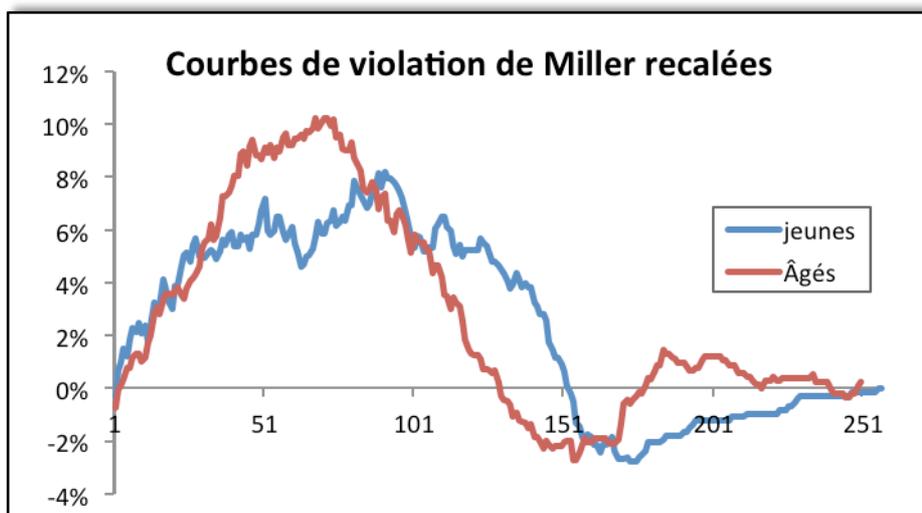


FIG. 23 - Courbes de violation des deux groupes recalées avec pour origine $p(TR < t) > 0$

Lorsque les deux courbes sont recalées sur une origine commune, il n'existe plus de différences entre les gains intégratifs des sujets jeunes et des sujets âgés.

Discussion

Dans cette expérience, nous avons montré que lorsqu'il n'existe plus de lien sémantique entre les items des différentes modalités et que le sujet est soumis à une tâche qui ne mobilise pas de processus mnésiques, alors les performances des sujets jeunes et des sujets âgés sont équivalentes. Nous observons un ralentissement de la distribution des sujets âgés, peut être lié à la tâche qui nécessite que le sujet déplace son attention pour répondre aux stimuli présentés. Ce serait donc un effet attentionnel. Lorsque l'on recalcule les courbes de chaque groupe par rapport à l'origine de la violation, les différences observées disparaissent. L'implication des processus perceptifs/mnésiques mis en jeu dans la tâche pourrait aussi être à l'origine des différences observées entre sujets jeunes et sujets âgés (Laurienti *et al.*, 2006).

Toutefois, cette conclusion doit être considérée avec beaucoup de précaution. En effet nos manipulations expérimentales ne sont pas assez rigoureuses pour conclure définitivement à la suite de ces deux expériences. Entre l'expérience 1 et l'expérience 2 nous avons fait varier trois facteurs, il est donc probable que la différence de tâche ne soit pas la seule explication des effets obtenus ici. Nous avons modifié la nature de la tâche, en réalisant une détection alors qu'il s'agissait d'une catégorisation, nous avons changé le lien sémantique entre les items, de telle sorte qu'il n'y est plus d'association dans l'expérience 2 et enfin nous avons introduit une composante spatiale qu'il n'y avait pas dans l'expérience 1. Tous ces facteurs nécessitent d'être contrôlés.

D'autres auteurs (Peiffer, Mozolic, Hugenschmidt & Laurienti, 2007) ont montré que même dans une tâche de détection simple, des sujets âgés présentaient un gain intégratif plus important que des sujets jeunes tant au niveau des temps de réponse qu'au niveau du gain calculé par le Race Model. Selon ces auteurs, cela suggère que l'augmentation d'intégration multi-sensorielle ne serait pas un épiphénomène du ralentissement cognitif global mais serait bien liée à un changement dans le processus de traitement multi-sensoriel. Ce résultat contraste avec ce que nous avons obtenu dans la présente expérience, bien que le paradigme expérimental soit différent. En effet dans notre tâche, les sujets devaient déplacer leur attention pour répondre, alors que dans l'étude de Peiffer et collaborateurs (2007), les sujets répondaient sur une seule touche à des stimuli apparaissant tous au centre. En fait les résultats dans la littérature sont assez discordants, car dans une étude récente (Stephen, Knoefel, Adair,

Hart & Aine, 2010), les auteurs trouvent un gain intégratif multi-sensoriel plus important chez les sujets jeunes que chez les sujets âgés. Ces auteurs ont utilisé un paradigme de détection original proche du concept de notre expérience. De plus, ils observent une diminution d'amplitude pour les stimuli multi-sensoriels chez les sujets âgés au niveau cérébral, qu'ils interprètent comme une diminution d'intégration corticale chez les sujets âgés.

Dans l'expérience 1 et 2 que nous venons de réaliser, nous avons observé que les sujets jeunes ne présentaient pas de gain intégratif lors d'une tâche de catégorisation (expérience 1) alors qu'ils en présentaient dans une tâche de détection (expérience 2). Toutefois, il est difficile de comparer ces deux expériences en l'état car elles diffèrent sur deux variables : la présentation latéralisée et la congruence sémantique entre les items des modalités auditives et visuelles. Nous allons réaliser maintenant deux expériences exclusivement chez les sujets jeunes pour essayer de contrôler les variations de gains intégratifs observés dans les deux premières expériences.

L'objectif de l'expérience 3 sera de contrôler la congruence entre les modalités auditives et visuelles. Nous allons utiliser la même tâche de détection que dans l'expérience 2 mais les items audiovisuels seront sémantiquement reliés. En comparant les résultats de l'expérience 2 et 3, nous devrions pouvoir contrôler les effets de congruence sémantique pour une tâche de détection, impliquant peu de processus mnésiques.

Dans une quatrième expérience, nous contrôlerons l'effet de la présentation latéralisée. Nous utiliserons une tâche de catégorisation latéralisée, ce qui la rendra distincte de l'expérience 1 uniquement sur ce point. Nous pourrions alors voir l'effet de la composante attentionnelle sur le gain intégratif en comparant les résultats des sujets jeunes entre les expériences 1 et 4.

Expérience 3 Tâche de détection latéralisée avec congruence sémantique

Si ce que mesure le Race Model est plus sensible aux processus mnésiques/perceptifs impliqués dans une tâche plutôt qu'au lien qui unit les différents composants entre eux, alors il ne devrait pas exister de différences entre l'expérience 2 et l'expérience 3 au niveau des courbes du Race Model. En effet, une tâche de détection est peu influencée par des processus mnésiques, par conséquent des conditions modulant l'association sémantique ou non entre les modalités auditives et visuelles ne devrait pas différer entre elles.

Hypothèses

H1: Il ne devrait pas exister de différences de gain intégratif mesuré par le Race Model entre les groupes de sujets jeunes de l'expérience 2 et ceux de l'expérience 3.

Méthode

Sujets

Douze sujets ($M = 20.6$; $ET = 1.2$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

Stimuli visuels

Les stimuli visuels sont les mêmes que ceux utilisés dans l'expérience 3 de la partie 1 (voir page 77)

Stimuli auditifs

Les stimuli auditifs étaient identiques à ceux utilisés lors de l'expérience 1 de la partie 2 (voir page 86). La durée des sons était de 500ms et ils pouvaient être diffusés soit dans l'oreille gauche soit dans l'oreille droite du sujet via un casque SENNHEISER HD 457.

Stimuli audiovisuels

Les items audiovisuels étaient composés des stimuli visuels et auditifs présentés simultanément. La durée de présentation de l'image et du son était toutes les deux de 500 ms. Le mode d'apparition des stimuli audiovisuels était toujours congruent, si l'image était présentée du côté gauche de l'écran, alors le son était diffusé dans l'oreille gauche du sujet et vice versa.

Procédure et plan expérimental

L'expérience comportait 192 stimuli répartis en quatre blocs de 48 stimuli chacun. La liste de présentation des items au sein de chaque bloc était la même et l'ordre des items dans la liste était déterminé de manière semi-aléatoire afin d'éviter que plusieurs items d'une même modalité se suivent. Dans cette expérience, Nous avons manipulé en intra-sujet, le facteur MODALITE, avec trois conditions : visuelle, auditive et audiovisuelle. Selon les conditions, seuls les items visuels, auditifs ou audiovisuels apparaissaient, dans 50% des cas à droite et

dans 50% des cas à gauche de la croix de fixation. Nous avons également manipulé, en inter-sujet, le facteur DETECTION, avec deux conditions : une condition « détection sémantique » et une condition « détection non sémantique » (*i.e.*, condition « sans association » de l'expérience 3, voir page 77). Les sujets réalisaient une tâche dans laquelle le stimulus visuel était correspondait systématiquement au stimulus auditif (*e.g.*, forme d'un carré + verbalisation du mot « carré »).

Chaque essai débutait par une croix de fixation d'une durée de 500ms, puis ensuite le stimulus apparaissait pendant 500ms suivi d'un écran blanc d'une durée de 1500ms pendant laquelle le sujet pouvait encore répondre au stimulus. La consigne donnée aux sujets était de détecter le plus rapidement le côté d'apparition du stimulus à l'aide de deux touches du clavier « a » et « p », une sur la gauche et une sur la droite du sujet. Il était indiqué au sujet d'essayer d'aller le plus rapidement possible et de répondre à tous les items visuels, auditifs et audiovisuels.

Résultats

Analyse des temps de réponses et des pourcentages de bonnes réponses

Les latences s'écartant de plus ou moins deux écarts types de la moyenne par condition n'ont pas été considérées. De plus, seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse (TAB. 10), moins de 5% des temps de réponse ont été écartés de l'analyse statistique. Toutes les données ont été retenues pour le calcul des pourcentages de bonnes réponses (TAB. 10). Nous avons comparé dans cette analyse les données du groupe de « détection non sémantique » (expérience 2) avec celles du groupe de « détection sémantique » (expérience 3).

		Modalité (TR)			Modalité (%BR)		
		V	A	AV	%V	%A	%AV
DETECTION	Sémantique	410 <i>10</i>	446 <i>19</i>	364 <i>11</i>	0,99 <i>0,002</i>	0,97 <i>0,01</i>	1,00 <i>0,002</i>
	Non sémantique	383 <i>13</i>	387 <i>24</i>	323 <i>11</i>	0,96 <i>0,02</i>	0,94 <i>0,02</i>	0,96 <i>0,02</i>

TAB. 10 – Moyennes des temps de réponses et pourcentages de réponses correctes et erreurs types en fonction des facteurs MODALITE et DETECTION

L'analyse de variance réalisée sur les temps de réponse met en évidence un effet du facteur MODALITE, $F(2,44) = 51.9$; $p < .01$; $\eta^2 = .70$. La modalité audiovisuelle (344ms) est détectée plus rapidement que les modalités visuelles (397ms), $F(1,22) = 286$; $p < .01$, ou auditives (417ms), $F(1,22) = 75.2$; $p < .01$. On relève également un effet du facteur DETECTION, $F(1,22) = 4.9$; $p < .05$; $\eta^2 = .18$, les sujets du groupe « détection non

sémantique » détecte plus rapidement les stimuli que les sujets du groupe « détection sémantique ». On ne relève pas d'interaction DETECTION*MODALITE, $F(2,44) = 2.4 ; p = .10 ; \eta^2 = .10$.

L'analyse de variance réalisée sur les taux de bonnes réponses met en évidence un effet du facteur MODALITE, $F(2,44) = 10.5 ; p < .01 ; \eta^2 = .32$. Les modalités visuelles (98%) et audiovisuelles (98%) sont associées à un meilleur taux de bonnes réponses que les items auditifs (96%), $F(1,22) = 13.1 ; p < .01$ et $F(1,22) = 12 ; p < .01$, respectivement. L'analyse de variance met en évidence un effet tendanciel du facteur DETECTION, $F(1,22) = 3.4 ; p = .077 ; \eta^2 = .13$. Les sujets du groupe « détection non sémantique » ont un taux de bonnes réponses plus faibles que les sujets du groupe « Détection sémantique ». On ne relève pas d'interaction DETECTION*MODALITE, $F(2,44) < 1$.

Analyse des données du Race Model

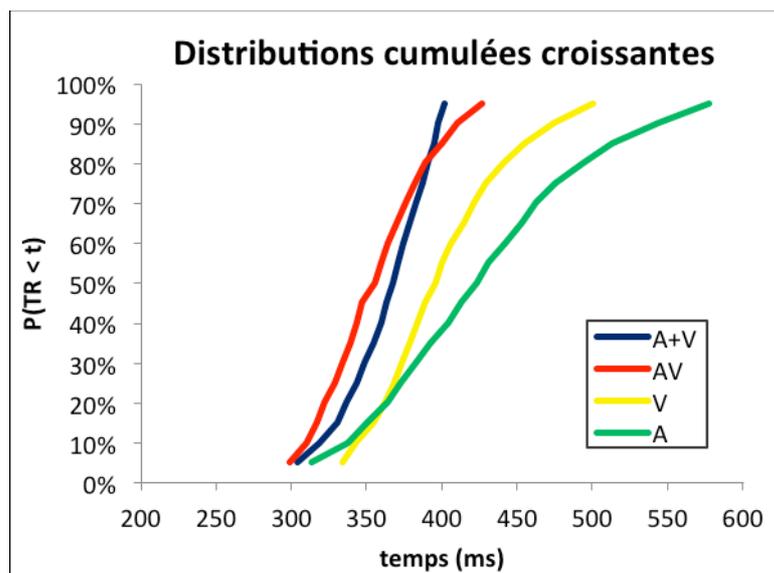


FIG. 24 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V).

Percentile	t_{11}	p
5%	1.53	.15
10%	2.94	<i>.01</i>
15%	4.26	<i>.001</i>
20%	5.08	<i>>.001</i>
25%	8.56	<i>>.001</i>
30%	7.09	<i>>.001</i>

TAB. 11 - Test de Significativité des six premiers percentiles. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués en italique.

L'analyse des données réalisée à l'aide du Race Model montre l'existence d'un gain intégratif multi-sensoriel pour cette tâche de détection où les items ont un lien sémantique entre eux (voir FIG. 24). La significativité de l'intégration est attestée dès le percentile 10% (voir TAB. 11)

Lorsque l'on compare les courbes de violation de l'inégalité de Miller pour les tâches de détection avec congruence sémantique (expérience 3) et celle pour laquelle il n'y a pas de composante sémantique (expérience 2), on s'aperçoit qu'il n'existe aucune différence significative entre ces deux conditions (voir FIG. 25), tant au niveau de la temporalité de la courbe de violation qu'au niveau de l'amplitude.

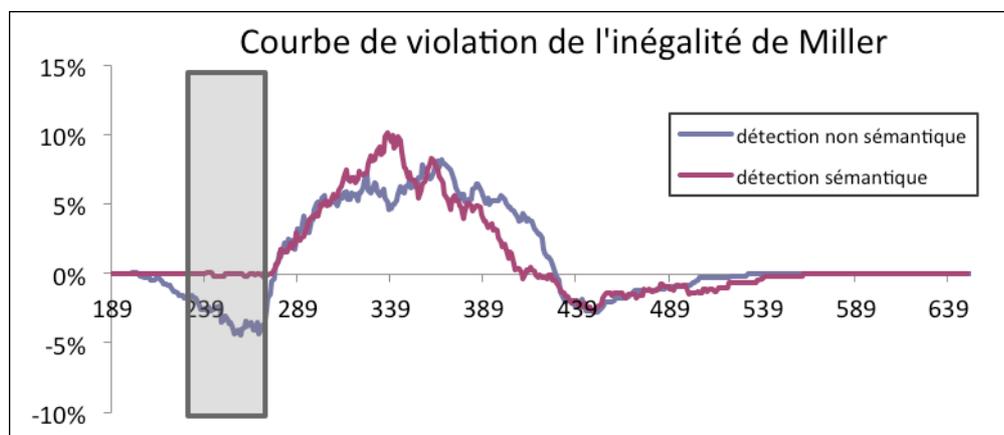


FIG. 25 - Représentation graphique des courbes de violation (AV-A+V) des conditions « détection sémantique » et « détection non sémantique ». Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un $\alpha < .05$.

Discussion

Dans cette expérience, nous avons comparé le gain intégratif et la courbe de violation dans une condition où les items visuels et auditifs étaient sémantiquement reliés (carré + verbalisation du mot carré) et dans une condition où il n'existait pas ce type d'association (carré + son aigu). Dans les deux conditions, les sujets devaient effectuer une tâche de détection. Selon nous, cette tâche est peu coûteuse en processus mnésique et devrait être peu sensible aux relations existant entre les modalités sensorielles. Ce résultat suggère que quel que soit le lien qui existe entre deux items, dans une tâche de détection, le gain intégratif tel qu'il est mesuré par le Race Model sera équivalent. Le Race Model serait donc peu sensible pour discriminer entre des items entretenant des relations en mémoire et des items sans liens sémantiques entre eux.

Lorsque l'on regarde de plus près les résultats comportementaux, on s'aperçoit que les sujets détectent plus rapidement les items dans la condition sans congruence sémantique. Ce résultat peut paraître étonnant et pourrait s'apparenter à un biais mais il est à mettre en parallèle avec les pourcentages de bonnes réponses. En effet les sujets placés dans la condition de non congruence sémantique traite plus rapidement les items et ce quel que soit leur modalité, l'effet n'est donc pas exclusivement porté par le lien qui existe entre les modalités visuelles et auditives. De plus, les sujets placés dans la condition de non congruence sémantique font plus d'erreurs que les autres sujets pour détecter les cibles. Les sujets ne sont donc pas facilités dans le traitement de la condition de non congruence.

Cette expérience nous a permis de contrôler l'influence de la relation sémantique entre les composants sensoriels dans une tâche de détection. Dans une quatrième et dernière expérience, nous allons contrôler l'effet de la position de la cible à détecter par rapport à ce qui a été fait dans l'expérience 1. En effet, on peut se demander si la présentation au centre des items (expérience 1) a créé un effet d'attente et entraîné des performances plafond. Si cette hypothèse est valable, l'introduction d'une composante attentionnelle devrait moduler le gain intégratif.

Expérience 4 Tâche de catégorisation latéralisée

Afin de voir si l'absence de gain intégratif chez les jeunes dans l'expérience 1 pouvait être liée au fait que les items apparaissaient tous au centre de l'écran, nous avons latéralisé la présentation des items comme dans les expériences 2 et 3, les sujets devant ici catégoriser les items présentés. Dans l'expérience 1, les sujets devaient catégoriser des items visuels, auditifs et audiovisuels apparaissant au centre de l'écran ou dans les deux écouteurs du casque. Pour chaque item, les sujets devaient dire si l'item présenté était un carré ou un cercle en appuyant sur une touche ou sur une autre du clavier. Nous suspectons la présence d'un effet plafond dans l'expérience 1 qui aurait par conséquent diminué le gain intégratif. En introduisant un facteur attentionnel, il est possible que cela module l'effet multi-sensoriel en augmentant le gain lié à la présentation bimodale comme cela a déjà pu être montré dans des tâches de détection (Molholm *et al.*, 2002). A l'inverse, si l'introduction d'un facteur attentionnel ne module pas les effets multi-sensoriels dans notre nouvelle expérience, nous serons à même de penser que l'absence de gain intégratif dans l'expérience 1 ne peut s'expliquer par un effet plafond sur les temps de catégorisation.

Hypothèses

H1: On devrait obtenir un gain intégratif supérieur dans cette tâche par rapport aux sujets jeunes de l'expérience 1, seulement et seulement si l'introduction d'un facteur attentionnel est à même de faciliter la tâche de catégorisation multimodale.

Méthode

Sujets

Douze sujets ($M = 20.8$; $ET = 1.9$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

Stimuli visuels

Les stimuli visuels sont les mêmes que ceux utilisés dans l'expérience 3 de la partie 1 (voir page 77).

Stimuli auditifs

Les stimuli auditifs étaient identiques à ceux utilisés lors de l'expérience 1 de la partie 2 (voir page 86).

Stimuli audiovisuels

Les items audiovisuels étaient composés des stimuli visuels et auditifs présentés simultanément. La durée de présentation de l'image et du son était toutes les deux de 500 ms. Le mode d'apparition des stimuli audiovisuels était toujours congruent, si l'image était présentée du côté gauche de l'écran, alors le son était diffusé dans l'oreille gauche du sujet et vice versa.

Procédure et plan expérimental

L'expérience comportait 192 stimuli répartis en quatre blocs de 48 stimuli chacun. La liste de présentation des items au sein de chaque bloc était la même et l'ordre des items dans la liste était déterminé de manière semi-aléatoire afin d'éviter que plusieurs items d'une même modalité se suivent. Dans cette expérience, Nous avons manipulé en intra-sujet, le facteur MODALITE, avec trois conditions : visuelle, auditive et audiovisuelle. Selon les conditions, seuls les items visuels, auditifs ou audiovisuels apparaissaient, dans 50% des cas à droite et

dans 50% des cas à gauche de la croix de fixation. Nous avons également manipulé, en intersujet, le facteur CATEGORISATION, avec deux conditions : une condition « catégorisation latéralisée » et une condition « catégorisation non latéralisée ». Les sujets réalisaient une tâche dans laquelle le stimulus visuel était correspondant systématiquement au stimulus auditif (*e.g.*, forme d'un carré + verbalisation du mot « carré »).

Chaque essai débutait par une croix de fixation d'une durée de 500 ms, puis ensuite le stimulus apparaissait pendant 500 ms suivi d'un écran blanc d'une durée de 1500 ms pendant laquelle le sujet pouvait encore répondre au stimulus. La consigne donnée aux sujets était de catégoriser le plus rapidement le côté d'apparition du stimulus à l'aide de deux touches du clavier. Afin d'éviter un conflit au niveau des réponses données par le sujet, les touches de réponses étaient situées l'une en dessous de l'autre au niveau du clavier et non plus de chaque côté du clavier. Cette manière de répondre diffère clairement de ce qui était demandé aux sujets de réaliser dans l'expérience 1. Toutefois nous avons adopté cette procédure dans le but d'éviter un conflit perceptivo-moteur dans lequel la moitié des items auraient été impliqués. En effet, si nous avons gardé la même procédure que dans l'expérience 1, les sujets auraient du répondre par exemple « carré » avec la touche gauche alors que celui-ci apparaissait à droite. Nous assumons complètement le fait que ces deux procédures diffèrent et ne soient pas complètement comparables mais nous avons estimé que le biais résultant de garder la même procédure aurait été encore plus préjudiciable.

Les sujets répondaient toujours avec l'index et le majeur, les touches de réponse « g » et « t » étaient contrebalancées chez tous les sujets. Il était indiqué au sujet d'essayer d'aller le plus rapidement possible et de répondre à tous les items visuels, auditifs et audiovisuels.

Résultats

Analyse des temps de réponses et des pourcentages de bonnes réponses

Nous avons comparé dans cette analyse les données du groupe de « catégorisation non latéralisée » (expérience 1) avec celles du groupe de « catégorisation latéralisée » (expérience 4).

Les latences s'écartant de plus ou moins deux écarts types de la moyenne par condition n'ont pas été considérées. De plus, seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse (voir TAB. 12), moins de 5% des temps de réponse ont été écartés

de l'analyse statistique. Toutes les données ont été retenues pour le calcul des pourcentages de bonnes réponses (voir TAB. 12).

		Modalité (TR)			Modalité (%BR)		
		V	A	AV	%V	%A	%AV
CATEGORISATION	latéralisée	519	560	491	0,95	0,97	0,97
		32	33	26	0,01	0,01	0,01
	Non latéralisée	552	598	523	0,96	0,99	0,98
		32	42	31	0,01	0,001	0,01

TAB. 12 – Moyennes des temps de réponses et pourcentages de réponses correctes et erreurs types en fonction du facteur MODALITE

L'analyse de variance réalisée sur les temps de réponse met en évidence un effet du facteur MODALITE, $F(2,44) = 39.4$; $p < .01$; $\eta^2 = .64$. La modalité audiovisuelle (507ms) est détectée plus rapidement que les modalités visuelles (535ms), $F(1,22) = 31.3$; $p < .01$, ou auditives (579ms), $F(1,22) = 60.8$; $p < .01$. On ne relève pas d'effet du facteur CATEGORISATION, ni d'interaction MODALITE*CATEGORISATION.

L'analyse de variance réalisée sur les taux de bonnes réponses met en évidence un effet du facteur MODALITE, $F(2,44) = 7.8$; $p < .01$; $\eta^2 = .26$. Les items audiovisuels (96%) et auditifs (97%) sont associées à un meilleur taux de bonnes réponses que les items visuels (96%), $F(1,22) = 7.3$; $p < .05$ et $F(1,22) = 10.3$; $p < .01$ respectivement. On ne relève pas d'effet du facteur CATEGORISATION, ni d'interaction MODALITE*CATEGORISATION, $F(2,44) < 1$.

Analyse des données du Race Model

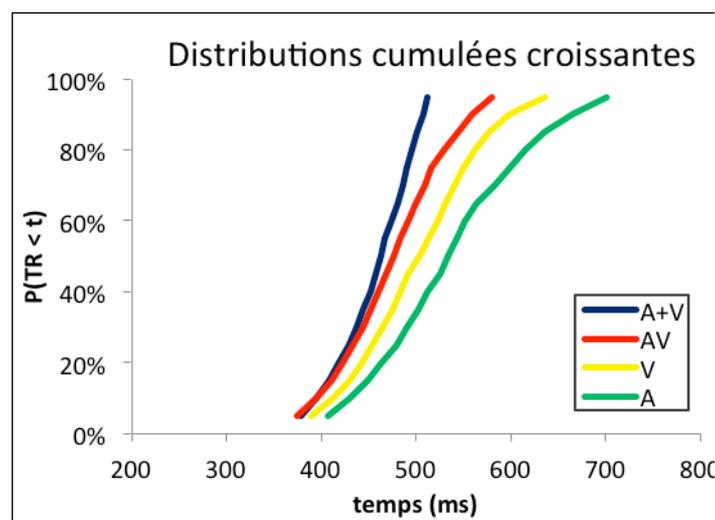


FIG. 26 - Distributions cumulées croissantes visuelles (V), auditives (A), audiovisuelles (AV) et théoriques (A+V).

Percentile	t_{11}	p
5%	1.01	.33
10%	0.76	.47
15%	0.96	.36
20%	1.2	.26
25%	1.2	.26
30%	1.63	.13

TAB. 13 - Test de Significativité des six premiers percentiles. Les percentiles pour lesquels il existe une différence significative entre la courbe AV et la courbe A+V sont indiqués en italique.

Lorsque nous observons les courbes de distributions cumulées croissantes, il s'avère qu'il n'existe pas de violation de l'inégalité de Miller (voir FIG. 26), et ce quel que soit le percentile (voir TAB. 13). Lorsque les items étaient présentés non latéralisés (expérience 1), nous trouvons le même profil de résultat. Ce résultat suggère qu'il n'existe donc pas de gain intégratif multi-sensoriel dans cette tâche de catégorisation que nous introduisions un facteur attentionnel ou non.

Nous avons comparé les courbes de violation de l'inégalité de Miller des sujets placés dans la condition « catégorisation non latéralisée » (expérience 1) et des sujets placés dans la condition « catégorisation latéralisée » (expérience 4) (voir FIG. 27). L'analyse statistique n'a révélé aucune différence entre les deux conditions, l'amplitude de la violation est donc équivalente que les items à catégoriser soit présentés au centre de l'écran ou bien qu'ils soient présentés de façon latéralisés.

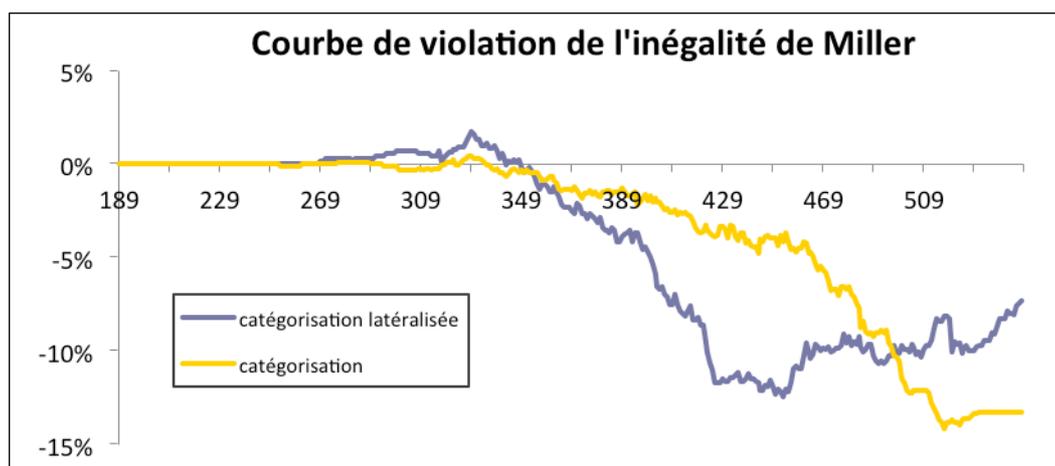


FIG. 27 - Représentation graphique des courbes de violation (AV-A+V) des conditions « catégorisation latéralisée » et « catégorisation non latéralisée ».

Discussion

Cette quatrième expérience avait pour but de contrôler l'effet d'une variable attentionnelle sur le gain intégratif mesuré par le Race Model sur une tâche de catégorisation. Dans l'expérience 1, nous n'avons pas observé de gain intégratif dans une tâche de catégorisation alors que tous les items apparaissaient au centre de l'écran. Dans cette expérience, la présentation latéralisée des mêmes items avait pour but de vérifier l'implication d'une composante attentionnelle sur le traitement intégratif multi-sensoriel. Les résultats obtenus n'ont pas mis en évidence de gain intégratif multi-sensoriel significatif. De plus la courbe de violation en condition « catégorisation latéralisée » ne différait pas de la condition « catégorisation non latéralisée ».

Il semble donc que l'ajout d'une composante attentionnelle ne permette pas d'augmenter le gain intégratif par rapport à une tâche de catégorisation où les items sont présentés au centre de l'écran (expérience 1). Les deux tâches de catégorisation que nous avons réalisé avec ou sans composante attentionnelle ne mettent donc pas en évidence de gain intégratif multi-sensoriel.

Conclusion

Nous allons résumer maintenant l'ensemble des résultats que nous avons obtenu lors de cette deuxième partie. Dans une première expérience, nous voulions répliquer les résultats d'une expérience précédente (Laurient *et al.*, 2006) ayant montré que les sujets jeunes présentaient moins de gain intégratif que des sujets âgés. Bien que nous ayons obtenu le même pattern de résultats, nous avons observé que les sujets jeunes ne présentaient pas du tout de gain multi-sensoriel. Bien que l'hypothèse selon laquelle la dégradation sensorielle chez les sujets âgés puissent expliquer le gain important dans la condition bimodale, il faut préciser pourquoi les sujets jeunes présentent peu ou pas de gain intégratif. Selon nous, dans le cas d'une présentation de composants sémantiquement très reliés, l'activation automatique des autres modalités au sein de la trace peut expliquer pourquoi la présentation bimodale n'apportera pas beaucoup plus de bénéfices qu'une présentation unimodale, déjà efficiente.

L'expérience 2 a montré que sujets jeunes et âgés présentaient un gain intégratif équivalent lorsque les sujets devaient détecter des stimuli plutôt que de les catégoriser. Cette observation montre bien que lorsque l'on change la nature de la tâche ainsi que le type de relation entre items visuels et auditifs, les sujets âgés présentent un gain intégratif équivalent aux sujets jeunes. Ce résultat est en contradiction avec des données de la littérature ayant utilisé une tâche de détection (Peiffer *et al.*, 2007 ; Stephen *et al.*, 2010). Cela montre bien qu'il est difficile de dégager un consensus au niveau du rapport entre processus intégratifs chez les

sujets âgés et chez les sujets jeunes. Il faut noter toutefois que les paradigmes utilisés diffèrent souvent entre eux et que certains peuvent demander plus de ressources attentionnelles (Stephen *et al.*, 2010), ce qui pourrait expliquer les différences observées. Nous avons donc essayé de comprendre les résultats observés chez les sujets jeunes. Il semble que le gain intégratif soit modulé en fonction de la nature de la tâche. Toutefois les expériences 1 et 2 étant différentes sur plusieurs points, nous avons été obligés de conduire deux autres expériences afin de contrôler différents facteurs, tel que la composante attentionnelle et le type de lien entre items visuels et auditifs.

L'expérience 3 avait pour but de contrôler l'aspect sémantique du lien entretenu entre les items visuels et auditifs. Nous avons observé que le gain intégratif était équivalent dans une tâche de détection que le matériel utilisé soit sémantiquement relié ou non. Cela suggère que les liens qu'entretiennent les modalités en mémoire sont moins importants que le type de tâche que les sujets doivent réaliser. Enfin, dans l'expérience 4, nous devions présenter les items de manière latéralisée pour vérifier l'effet de la composante attentionnelle. Il s'est avéré que le type de présentation avait peu d'influences sur le gain intégratif dans une tâche de catégorisation. Lorsque l'on fait une synthèse des résultats obtenus dans les quatre expériences (voir TAB. 14), il s'avère que le gain intégratif varie en fonction du type de tâche et moins en fonction des composantes attentionnelles ou encore du type de lien existant entre les items visuels et auditifs.

	Type de	Composante	Lien sémantique des items	Gain intégratif
Expérience 1	catégorisation	Non	Oui	-
Expérience 2	détection	Oui	Non	+
Expérience 3	détection	Oui	Oui	+
Expérience 4	catégorisation	Oui	Oui	-

TAB. 14 – Tableau récapitulatif des résultats obtenus sur le gain intégratif en fonction des expériences réalisées

Nous avons voulu comparer le gain intégratif en fonction des différentes expériences réalisées (voir FIG. 28). Il semble que le gain intégratif mesuré à l'aide de cette méthode soit plus important pour les tâches de détection que pour les tâches de catégorisation. Pour vérifier cela nous avons réalisé des analyses statistiques (Mozolic *et al.*, 2008a) sur les courbes de violation de l'inégalité de Miller obtenues lors de nos expériences. Toutefois, nous avons restreint nos analyses à la comparaison des expériences 3 et 4, car ce sont les seules de cette série d'expériences à être strictement comparables.

Le résultat de cette analyse indique que la tâche de détection produit un gain intégratif significativement plus important que celui produit dans la tâche de catégorisation (voir FIG. 29). Nous ajouterons que nous n'avons pas réussi à mettre en évidence de gain intégratif dans les tâches de catégorisation bien que d'autres études ont mis en évidence un tel gain (Laurienti *et al.*, 2006). En fonction des processus impliqués dans la tâche, le Race Model ne semble pas aussi sensible dans la mesure du gain intégratif. On peut alors se demander ce qu'il mesure réellement, en effet dans une tâche de détection, peu coûteuse en processus mnésique, il présente un gain intégratif alors que ce n'est pas le cas dans une tâche de catégorisation, plus coûteuse en processus mnésique. Toutefois, on ne peut exclure de nos interprétations, la présence de processus décisionnels dans la tâche de catégorisation qui auraient pu cacher un éventuel gain multi-sensoriel.

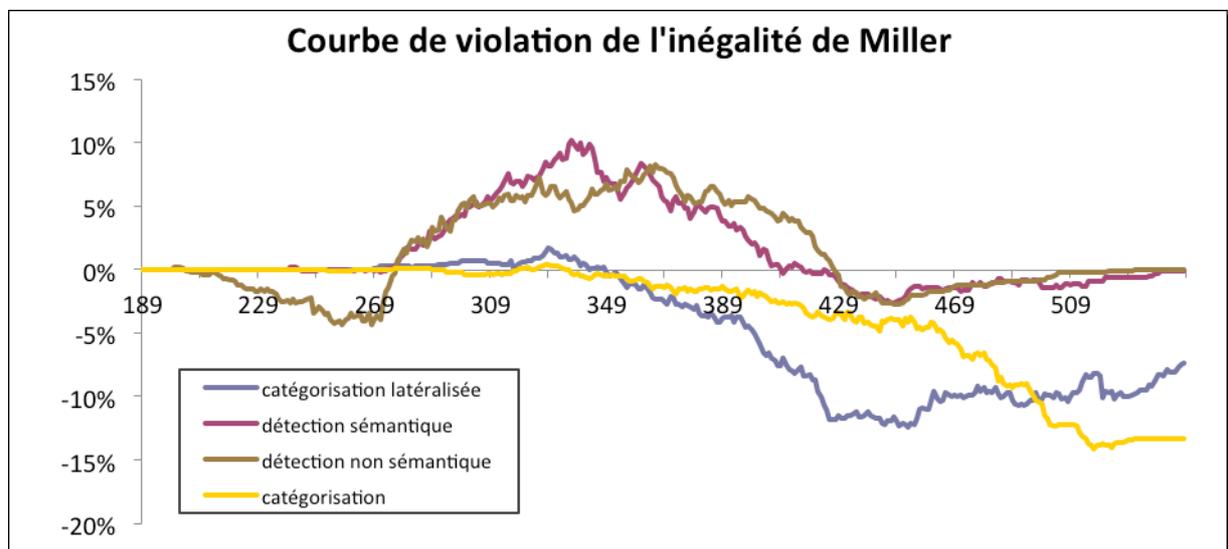


FIG. 28 - Représentation graphique des courbes de violation (AV-A+V) des quatre conditions « catégorisation latéralisée », « catégorisation non latéralisée », « détection sémantique » et « détection non sémantique ».

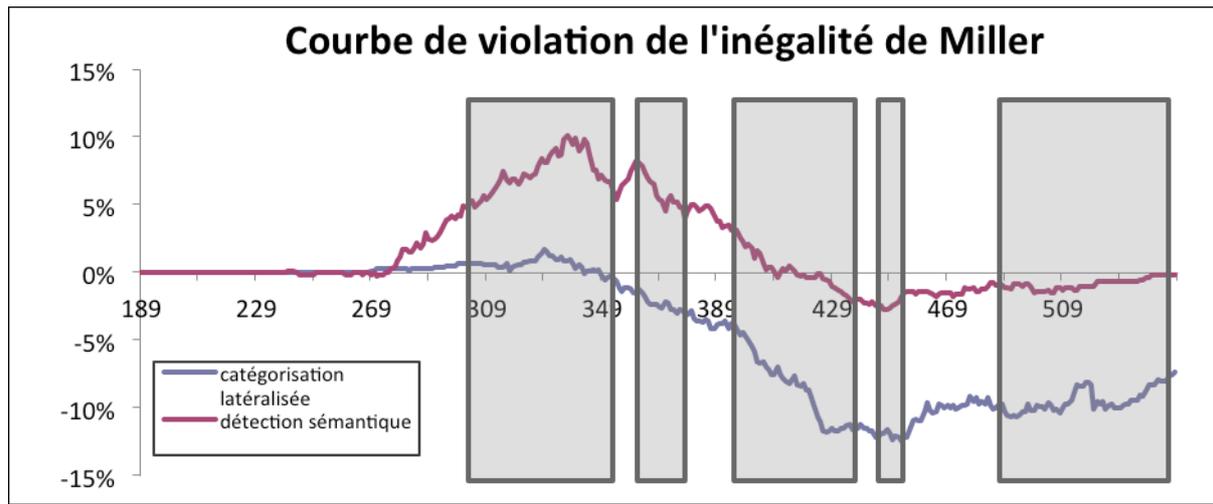


FIG. 29 - Représentation graphique des courbes de violation (AV-A+V) des conditions « détection » et « catégorisation ». Chaque période où les courbes diffèrent significativement est indiquée par une fenêtre grisée, la significativité est considérée pour un $\alpha < .05$.

On peut se questionner quant à la validité de notre hypothèse, selon laquelle le Race Model serait moins sensible pour mesurer des effets multi-sensoriels lorsque les items ont une relation sémantique forte en mémoire ou lorsque la tâche implique des processus mnésiques. Si tel était le cas, alors on devrait observer une distribution cumulée visuelle et auditive plus proche de la distribution audiovisuelle uniquement dans la tâche de détection. Qualitativement, c'est le profil de résultat que nous observons mais il est plus marqué pour la distribution cumulée visuelle (voir FIG. 24 et FIG. 26). Si l'on se réfère au modèle Act-in, le composant auditif serait peut-être moins apte à réactiver le composant visuel dans ces conditions précises. On peut également supposer que cet effet est plus fort en modalité visuelle de par l'effet de dominance visuelle (Colavita, 1974). Il est donc important de rester prudent quant à cette interprétation.

Pour terminer, ces expériences mériteraient d'autres contrôles expérimentaux afin de valider les résultats que nous avons obtenus ici. Le nombre de sujets âgés inclus dans nos expériences est relativement faible comparé à d'autres études (Laurienti *et al.*, 2006). La variabilité de cette population pourrait alors expliquer pourquoi nous avons obtenu des résultats similaires chez les sujets jeunes et chez les sujets âgés dans l'expérience 2. Ensuite, il serait peut-être intéressant de reproduire l'expérience 2 et 3 en intra-sujet afin de s'assurer que la détection d'un matériel sémantique n'entraîne pas plus de gain qu'un matériel non sémantique.

Nous avons donc montré ici que les processus perceptifs et mnésiques impliqués dans une tâche modulent le gain intégratif tel qu'il est mesuré par le Race Model. Le Race Model serait

plus efficace pour mesurer des gains multi-sensoriels en perception plutôt qu'en mémoire, même si cette conclusion est à considérer avec beaucoup de prudence.

Chapitre 3. L'intégration en mémoire à long terme

Nous avons vu dans le premier chapitre que la récupération d'une trace en mémoire, selon le modèle Act-in, était dépendante des relations existant entre les modalités constitutives de la trace. En conséquence, la présentation d'une modalité devrait pouvoir activer une autre modalité non perceptivement présente, mais associée au sein de la trace en mémoire (Brunel *et al.*, 2009). Nous avons vu également que certaines modalités sensorielles étaient plus à même de représenter certains concepts que d'autres modalités, ainsi les êtres vivants seraient représentés majoritairement par des unités visuelles alors que les objets manufacturés seraient représentés majoritairement par des unités fonctionnelles (Farah & McClelland, 1991). Ainsi, en fonction des propriétés des concepts, certaines unités sensorielles auraient plus de poids au sein de la trace pour activer d'autres unités constitutives du concept (Thompson-Schill *et al.*, 1999). Dans ce chapitre, nous allons nous intéresser à la manière dont une trace peut être récupérée en mémoire, en supposant que les unités constitutives de la trace et leurs poids relatifs devraient influencer l'efficacité de la récupération. Selon nous, la multimodalité joue un rôle essentiel dans la récupération en mémoire à long terme. Pour tester cela, nous nous placerons sous l'angle de la *Dual-process Theory*, qui suggère l'existence de deux processus : la familiarité et la recollection pour rendre compte de l'accès à la trace lors de la reconnaissance. Nous nous intéresserons essentiellement à la reconnaissance, car ce type de tâche est à même de rendre compte des mécanismes d'activation multimodale en mémoire à long terme dans la perspective du modèle Act-in. Si l'on présente des items visuels avec un son sémantiquement congruent ou non lors d'une phase d'encodage et que l'on représente uniquement les items visuels lors d'une phase de reconnaissance ultérieure, on pourra voir si la manipulation sensorielle lors de l'encodage a influencé la récupération. Après avoir abordé la notion de processus de familiarité/recollection dans leurs aspects cognitifs et sous l'angle des neurosciences cognitives, nous ferons une revue succincte des travaux ayant manipulé plusieurs modalités sensorielles en mémoire tant dans des tâches nécessitant un accès implicite que dans des tâches nécessitant un accès explicite.

3.1. Modèle de récupération en mémoire à long terme

Il existe plusieurs manières d'accéder à une information en mémoire déclarative, parmi celles-ci, on peut citer le rappel libre ou encore la reconnaissance. En effet, nous avons expliqué dans la partie précédente en quoi une tâche de reconnaissance peut être intéressante pour

étudier l'effet des activations sensorielles en mémoire. Nous allons dans cette partie nous focaliser sur la reconnaissance en essayant de dégager les différents processus qui sous-tendent ce type de récupération mnésique.

3.1.1. Généralité sur les processus de recollection/familiarité

Les résultats provenant des études réalisées en psychologie cognitive, en neuropsychologie ou encore en neuro-imagerie semblent indiquer que la reconnaissance reflèterait deux processus mnésiques distincts : la familiarité et la recollection (Reder, Park & Kieffaber, 2009; Yonelinas, 2002). Cette distinction peut s'illustrer intuitivement lorsque l'on rencontre une personne dans la rue et que l'on a l'impression de la connaître (familiarité), elle nous paraît familière, toutefois nous sommes parfois dans l'incapacité de donner des indications sur le lieu où nous l'avons rencontré précédemment ou même sur son identité (recollection). Cet exemple intuitif suggère tout de même qu'un jugement en mémoire peut être basé soit sur la recollection d'informations sur l'expérience passée soit sur une évaluation de la familiarité du stimulus.

3.1.2. Un ou plusieurs processus pour rendre compte des performances en reconnaissance ?

Dans sa revue de 2003, Yonelinas suggère que s'il est possible d'expliquer les performances en reconnaissance avec un modèle qui n'implique qu'un seul système de mémoire alors ce modèle est préférable. Toutefois, une majorité de recherches a conclu à l'existence de deux processus distincts. Par exemple, dans des tâches de reconnaissance, des études ont montré qu'au moins deux processus mnésiques fonctionnellement distincts étaient nécessaires pour rendre compte des performances (Kelley & Wixted, 2001). Nous allons décrire quelques unes de ces dissociations.

Premièrement, les études qui ont utilisé la vitesse de traitement ont montré que la familiarité était plus rapide que la recollection. Par exemple, dans des épreuves avec une contrainte temporelle, les sujets étaient capables de discriminer plus rapidement des items anciens/nouveaux que de réaliser une discrimination nécessitant de récupérer des informations spécifiques à propos du contexte d'étude (Hintzman, Caulton & Levin, 1998). Yonelinas va distinguer ces deux types de tâche en terme de tâche de reconnaissance d'items et tâche de reconnaissance associative. Deuxièmement, l'analyse des degrés de certitude associés aux réponses a indiqué que la recollection et la familiarité pouvait produire des courbes ROC

distinctes. Si l'on représente graphiquement le degré de certitude en fonction de la proportion de vrais positifs (le sujet a reconnu comme ancien un item étudié précédemment) et de faux positifs (le sujet a reconnu comme ancien un item nouveau), les courbes que l'on observe sont curvilinéaires et leur forme varie selon les conditions expérimentales et les processus mnésiques mis en jeu. De plus, dans les tâches de reconnaissance d'items, tant que les performances sont au-dessus du hasard, les courbes ROC sont toujours curvilinéaires alors que dans des tâches de reconnaissance associative, les courbes ROC sont assez différentes au sens qu'elles ont une forme beaucoup plus linéaire. Différents processus contribueraient aux performances observées dans ces deux types de tâche. Troisièmement, la familiarité et la recollection auraient des corrélats électrophysiologiques distincts. Les travaux ayant étudié les potentiels évoqués ont montré que la reconnaissance d'items sur la base de la familiarité sans recollection engageait des distributions spatiales et temporelles de réponses cérébrales distinctes (Klimesch *et al.*, 2001). Ainsi, dans une expérience, des auteurs ont montré que des items correctement reconnus étaient associés à une onde frontale précoce alors que seuls les items ayant bénéficié d'un encodage profond étaient associés à une onde pariétale gauche plus tardive (Rugg *et al.*, 1998). Ces résultats n'indiquent pas pour autant quelles régions cérébrales supportent les processus de reconnaissance, mais suggèrent qu'il existerait au moins deux processus cérébraux impliqués dans la reconnaissance. Quatrièmement, la recollection est plus sévèrement atteinte que la familiarité lors de certaines atteintes cérébrales. Par exemple, des patients amnésiques présentent de plus grandes difficultés pour des tâches de reconnaissance associative que pour de tâches de reconnaissance d'items. Toutefois, une critique importante provenant de l'observation des patients amnésiques est liée au fait qu'au mieux, on se trouve en présence d'une simple dissociation (Rugg & Yonelinas, 2003). En effet, il n'existe pas de patients qui présentent la dissociation inverse, c'est à dire un processus de recollection intact et un processus de familiarité déficitaire. De plus, d'un point de vue anatomique, le cortex parahippocampique est la structure qui fournit le plus d'inputs à l'hippocampe (Lavenex & Amaral, 2000) et en admettant que l'hypothèse selon laquelle la familiarité est sous tendue par le cortex parahippocampique soit correcte (Yonelinas, 2002), un déficit sélectif de la familiarité sera difficile à mettre en évidence. En effet, des lésions qui perturberaient la familiarité, créeraient une déafférentation suffisamment importante de l'hippocampe, affectant par conséquent la recollection.

Dans la vision classique de la Dual-process theory, la familiarité est un processus continu au sens qu'il s'exprime en terme de degré. Un faible degré de familiarité est associé à un faible

degré de certitude alors qu'un haut degré de familiarité est associé à un haut degré de certitude et des pourcentages de réponses correctes élevés. Au contraire, le processus de recollection est perçu comme catégoriel au sens qu'il intervient ou non. De plus la vision traditionnelle veut que la décision dans une situation de reconnaissance soit basée soit sur un processus soit sur un autre. Le processus de recollection est catégoriel, car lorsqu'il intervient, il confère un degré de certitude élevé quand à l'item à juger, rendant par la même occasion non nécessaire, le recours au processus de familiarité. Lorsque la recollection échoue complètement, le seul recours possible repose sur la familiarité. Une vision alternative repose sur le fait que familiarité et recollection seraient deux processus continus (Wixted, 2007). Les deux processus contribueraient ensemble à évaluer l'aspect ancien ou nouveau d'un item. L'argument issu de l'étude de Rugg et collaborateurs (1998) est alors aussi valable dans ce contexte théorique. En effet, les auteurs avaient montré que des items étudiés lors d'une phase d'encodage, pouvant être reconnus ultérieurement pour certains sur la base de la familiarité et d'autres sur la base de la recollection, produisaient tous deux une réponse électrophysiologique précoce située au niveau des électrodes frontales. Toutefois, seuls les items ayant bénéficié d'un encodage profond et plus susceptible d'être reconnus sur la base de la recollection produisait une réponse électrophysiologique tardive située au niveau des électrodes pariétales gauches. La différence avec les modèles précédents repose sur le fait que la recollection n'est plus perçu comme un processus catégoriel, mais comme un processus continu. Toutefois la recollection et la familiarité resteraient deux processus distincts, la familiarité étant plus lente que la recollection, chaque traitement réalisé par la recollection s'ajouterait au traitement basé sur la familiarité plutôt que de se substituer à lui. Dans une étude récente des auteurs ont montré que le processus recollection pouvait être associé à des degrés de certitude et des pourcentages de réponse corrects faibles autant qu'à des degrés de certitude et des pourcentages de réponses corrects élevés (Mickes, Wais & Wixted, 2009). Ce résultat impliquerait également le processus de recollection dans les souvenirs « plus faibles ».

3.1.3. Les méthodes de mesure

Deux types de méthode ont permis d'étudier la familiarité et la recollection : les méthodes de dissociation de tâche et les méthodes de dissociation des processus.

3.1.3.1. Les méthodes de dissociation de tâche

On trouve dans ces méthodes, celles basées sur la rapidité des réponses et qui ont permis notamment de montrer que les réponses basées sur la familiarité étaient plus rapides que celles données sur la base de la recollection. Ce type de méthode peut donc être efficace pour distinguer la contribution des deux processus. Cette méthode se distingue encore en trois sous procédures. La méthode temps-réponse consiste à observer les performances en reconnaissance en séparant les réponses données rapidement de celles données plus lentement, alors que le sujet a tout le temps de répondre. Les réponses rapides reflètent la familiarité alors que les réponses plus lentes reflètent la recollection (Péris & Tiberghien, 1984).

On trouve également la méthode rappel/reconnaissance qui repose sur l'hypothèse selon laquelle les processus de recollection qui opèrent dans des tâches de reconnaissance sont similaires à ceux qui opèrent dans des tâches de rappel libre, par conséquent les performances en rappel libre peuvent être utilisées comme indice de recollection. En comparant les performances en rappel libre et en reconnaissance qui reflète la familiarité et la recollection, il est alors possible de faire des inférences sur les effets de différentes variables sur la recollection et la familiarité. Par exemple, si une variable a un effet plus important en rappel libre qu'en reconnaissance alors on peut dire qu'elle a un effet plus important en recollection que sur la familiarité. Une limite de cette méthode repose sur le fait que ces deux conditions de test diffèrent en termes d'indices de récupération et du type de réponse requis. Une autre difficulté résulte de la différence des échelles de mesure utilisée dans ces deux types de tâche.

3.1.3.2. Les méthodes de dissociation de processus

La procédure de dissociation des processus (PDP) de Jacoby (1991) mesure la recollection comme la capacité de se souvenir du moment et du lieu où un item a été étudié précédemment. Si un sujet peut récupérer un item sur la base de la recollection alors il peut connaître les détails du contexte d'encodage alors que la familiarité ne peut pas supporter cette assertion. La PDP est consistante avec la plupart des modèles qui postulent l'existence de deux processus et qui admettent que la recollection et la familiarité sont deux processus indépendants. Si ces deux processus sont indépendants, alors il est possible de trouver des manipulations expérimentales qui auront un effet sur un processus mais pas sur l'autre et inversement, par contre si les deux processus sont dépendants alors une telle dissociation est impossible.

Dans la procédure *remember/know*, développée par Tulving (1985), les sujets doivent juger sur en fonction de leur souvenir si un item a été reconnu sur la base du *remembering* (récupération des informations relatives au contexte d'encodage) ou sur la base du *knowing* (l'item est jugé familier en l'absence de récupération des éléments du contexte d'encodage). Les réponses de type *remember* reflètent le processus de recollection mais l'estimation de la familiarité n'est pas aussi simple. Dans des études anciennes, les réponses *know* étaient directement associées au processus de familiarité (*i.e.*, Gardiner, 1988), mais il semblerait qu'elles semblent sous estimées les réponses basées sur la familiarité (Yonelinas & Jacoby, 1995). De plus, selon Tulving, les réponses *know* peuvent être de type implicite ou sémantique, ce qui peut poser problème pour discriminer les processus de familiarité et de recollection en reconnaissance.

3.2. Aspect fonctionnel et structurel de la récupération en MLT

3.2.1. Rôle des structures temporelles médianes

3.2.1.2. Le cortex rhinal

Bien que nous soyons capables de rappeler un très grand nombre d'informations, nos ressources pour encoder par unité de temps semblent plutôt limitées. Certains auteurs postulent que pour fonctionner efficacement, la mémoire doit utiliser des opérations de filtrage qui détecteraient l'aspect familier ou la nouveauté des informations présentées au système. Fernández et Tendolkar (2008) donnent les propriétés nécessaires à une structure devant assurer ce type d'opérations : (1) cette structure doit détecter rapidement l'aspect nouveau/familier des items avant que des opérations d'encodage plus élaborées aient débuté, (2) elle doit coder spécifiquement un grand nombre d'items et il est donc probable qu'elle utilise des représentations conceptuelles ou sémantiques, (3) elle doit être bien intégrée dans les opérations d'encodage afin de pouvoir orienter les ressources de façon appropriée. Les auteurs proposent d'identifier un processus totalement intégré qui permette à lui seul de traiter l'encodage des stimuli et la détection de la familiarité, de telle sorte que plus un item est familier moins la force de l'encodage sera nécessaire et vice et versa (voir FIG. 30). Pour Fernández et Tendolkar, le cortex rhinal se prête tout à fait à ce type de traitement. Par cortex rhinal, les auteurs regroupent les cortex péri-rhinaux et ento-rhinaux qu'il est difficile de distinguer clairement par les méthodes utilisées actuellement.

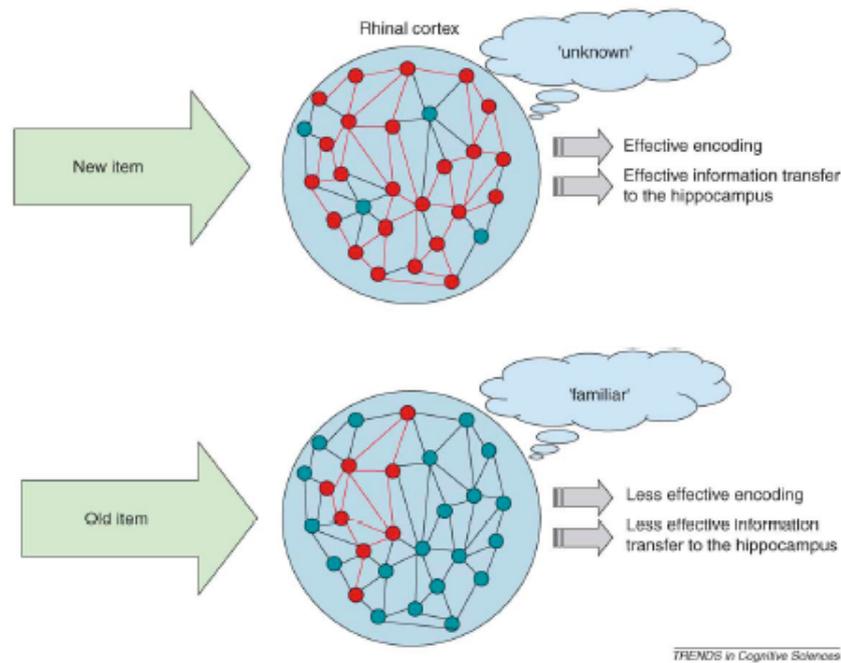


FIG. 30 – Représentation des opérations réalisées par le cortex rhinal, selon Fernández & Tendolkar, 2008

Des études ont montré que le traitement répété d'un item était associé à une réduction de l'activité neuronale, comparé au traitement initial. Cet effet de « suppression de répétition » a déjà été observé dans de nombreuses régions du cerveau notamment dans des paradigmes d'amorçage. Des items présentés visuellement sont souvent associés à une suppression de répétition au niveau du cortex temporal inférieur incluant le cortex rhinal qui fournit les inputs principaux de l'hippocampe. Initialement, les effets de suppression de répétition ont souvent été assimilés au traitement de la nouveauté à l'encodage ou à l'amorçage de répétition, une forme de mémoire implicite. Récemment, des études ont montré des effets de suppression dans le cortex rhinal pour la reconnaissance d'items en dehors de tout contexte d'encodage (Henson *et al.*, 2003). Ce qui signifierait que le cortex rhinal pourrait sous tendre le processus de familiarité (Graham, Barense & Lee, 2010).

3.2.1.3. Lobe temporal médian et « binding »

Il existe de nombreux arguments pour dire que le lobe temporal médian (LTM) joue un rôle important dans la mémoire déclarative (Aggleton & Brown, 2006). Une vision dominante consiste à dire que le LTM est impliqué dans la reconnaissance d'évènements précédemment rencontrés, toutefois il existe un débat concernant les traitements réalisés dans les différentes régions cérébrales qui constituent le LTM. Il a été suggéré récemment que l'hippocampe et les cortex adjacents pourraient supporter différents aspects de la reconnaissance. En effet, il a été montré que les cortex hippocampique et parahippocampique sous-tendraient le processus de

recollection alors que le cortex périrhinal sous-tendrait le processus de familiarité (Fernández & Tendolkar, 2008). Cependant, il existe certaines contradictions dans les études rapportées concernant cette vision (*i.e.*, Opitz, 2010). Pour dépasser ces contradictions, des modèles récents ont proposé que les cortex périrhinaux et parahippocampiques sous tendraient respectivement l'encodage et la récupération d'items spécifiques et d'informations contextuelles alors que les cortex hippocampiques supporteraient le stockage des associations item/contexte. Cette vision est notamment développée suite aux études qui ont montré une augmentation de l'activité du cortex hippocampique dans des tâches de mémoire associative, telles que l'apprentissage de paires associées (Eldridge *et al.*, 2005).

Opitz (2010) propose une conception étendue de la notion de *binding* proposée ci-dessus, en développant l'idée selon laquelle l'hippocampe serait le support d'un « *relational binding* » des représentations corticales des items, actions, etc. et du contexte spatio-temporel qui compose l'unicité de l'expérience. Ces activités de liage peuvent être décrites en terme d'opérations relationnelles qui lient et organisent les éléments individuels de l'expérience. Par exemple, la rencontre de deux personnes dans un restaurant va fournir une information relationnelle à propos de l'identité des individus et du contexte spatio-temporel. Le « *relational binding* » est hautement flexible au sens qu'il permet un réarrangement des éléments de l'expérience individuelle pour faire face aux nouvelles situations. Par exemple, si vous rencontrez dans un restaurant, un collègue que vous avez l'habitude de voir au travail, vous pouvez aisément faire face à cette situation parce qu'il n'existe pas de relations rigides entre votre collègue et le contexte de travail. La flexibilité du « *relational binding* » est complétée par une seconde propriété : la « *compositionality* ». Cela signifie que les éléments d'une relation gardent leur identité quel que soit le liage, la représentation de votre collègue restera la même quelle que soit la situation de rencontre. Troisièmement, le « *relational binding* » est un processus automatique, obligatoire, au sens qu'il est mis en place chaque fois que des informations arbitraires sans relations pré-existantes doivent être réunies au sein d'une expérience. Enfin, les connexions hippocampiques possèdent les caractéristiques anatomiques et computationnelles pour être le support des propriétés du « *relational binding* ». C'est également dans la continuité d'un modèle neuro-computationnel qui suggère que le codage neuronal dans l'hippocampe produit deux phénomènes (O'Reilly & Norman, 2002). Premièrement, un codage distribué va réduire la probabilité que les mêmes neurones dans l'hippocampe soient activés par deux inputs distincts, ce qui va conduire à deux

représentations distinctes et deuxièmement, ce codage distribué va favoriser le liage d'items et de contextes en une représentation cohérente de l'évènement (FIG. 31-A).

A l'inverse, le cortex temporal médian assignerait des représentations similaires à des inputs similaires en utilisant des représentations se chevauchant (O'Reilly & Norman, 2002). Ces représentations deviendraient de plus en plus nettes au cours d'expositions répétées. Par conséquent, la première rencontre avec un évènement va activer faiblement une population large d'unités du cortex temporal médian alors que des évènements répétés et familiers vont entraîner une forte activation d'un petit nombre d'unités (FIG. 31-B). Par ce mécanisme, le cortex temporal médian est capable de supporter le « *representational binding* », c'est-à-dire relier des informations communes partagées par plusieurs évènements. Le cortex temporal médian extrait les régularités présentes dans les expériences récentes. Ces régularités comprennent la fréquence de co-occurrence mais aussi les contingences temporelles. La représentation de ces régularités constitue une connaissance décontextualisée qui est dépendante d'une expérience particulière dans laquelle cette information a été initialement encodée.

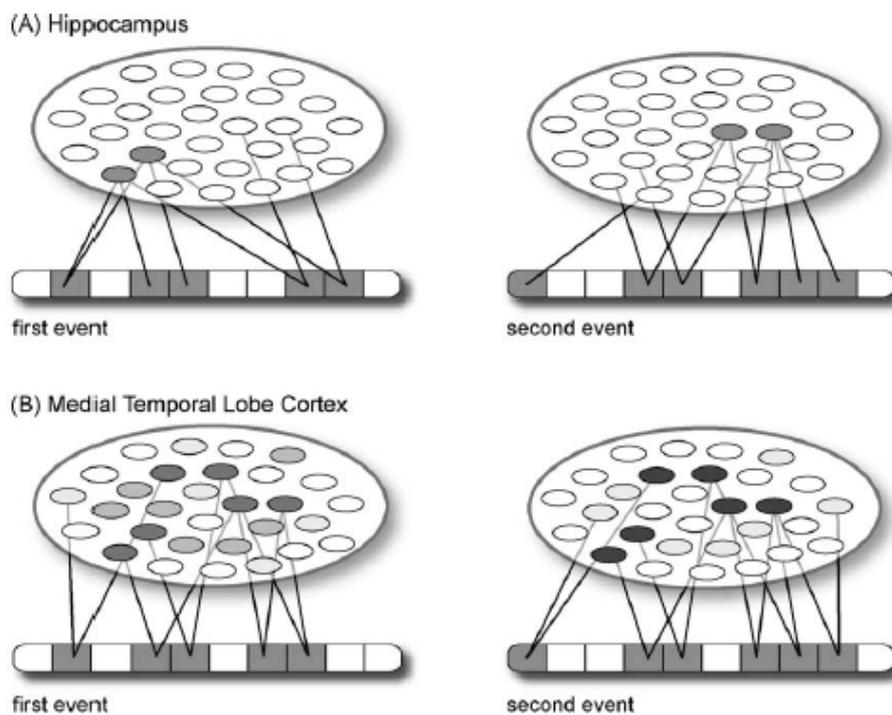


FIG. 31 – Représentation schématique du codage neuronal par deux évènements distincts mais similaires

En résumé, Opitz propose l'implication de l'hippocampe dans le liage relationnel d'éléments d'expériences individuelles grâce à des représentations distribuées, alors que les cortex

périorhinaux et parahippocampiques lieraient des informations représentationnelles à propos des items et du contexte avec des occurrences précédentes similaires.

Si l'on considère maintenant une tâche de mémoire épisodique, l'item présenté durant le test de reconnaissance doit être identifié par rapport à la situation d'encodage. Lors de l'apprentissage, une association doit être faite entre l'item à apprendre et son contexte, plus tard cette association devra être récupérée. Ce type d'association peut être considéré comme un exemple de *relational binding*. Comme nous l'avons vu précédemment à partir d'un modèle neuro-computationnel (Norman & O'Reilly, 2003), le codage neuronal distribué dans l'hippocampe aboutit à la construction de représentations distinctes à partir d'une association arbitraire entre un item et son contexte. L'hippocampe est capable de reconstruire la situation d'encodage, ce qui rend possible la récupération des informations contextuelles. Le *relational binding* d'un item et de son contexte aboutit à la reconnaissance basée sur le processus de recollection. Au contraire, le cortex temporal médian assigne des représentations similaires à des inputs similaires en affinant au cours des présentations le nombre d'unités recrutées. Lors du test, la présentation d'un item précédemment étudié va initier une série de traitement qui consiste en une comparaison entre une représentation du stimulus actuel et la représentation au sein du cortex temporal médian. Ce traitement va établir la similarité qui existe entre l'item étudié et celui stocké au sein du système, cette comparaison est proche de ce que peut décrire Hintzman (1986) avec la notion d'écho. Ce traitement reflèterait donc un processus de familiarité (Opitz, 2010). Par conséquent, le *binding* en mémoire épisodique implique que la recollection et la familiarité émerge du *relational binding* et/ou du *representational binding*.

3.2.2. Mécanismes oscillatoires et processus mnésiques

Depuis de nombreuses années, les neurosciences tentent de comprendre le lien fonctionnel entre les oscillations cérébrales et le comportement humain. Ainsi, les neurosciences cognitives ont rapproché différentes fréquences oscillatoires avec certains processus cognitifs. Le domaine de la mémoire a largement été investigué et de nombreuses hypothèses ont émergé quant aux liens fonctionnels entre mécanismes oscillatoires et processus mnésiques.

3.2.2.1. Introduction

Une oscillation correspond à la fréquence de décharge neuronale que l'on peut observer dans une population de neurones donnée. Plusieurs fréquences ont été observées lors des enregistrements directement réalisés sur le scalp : les ondes delta (autour de 3Hz), les ondes thêta (4-7Hz), les ondes alpha (8-13Hz), les ondes mu (9-15Hz), les ondes beta (15-30Hz) et

les ondes gamma (environ 40Hz). Chacune de ces fréquences a été observée lors de certains phénomènes, par exemple, l'apparition d'un rythme alpha lors du passage yeux fermés/yeux ouverts ou encore le rythme thêta lors de la réalisation de calcul mental. Concernant les ondes gamma, certains auteurs distinguent au moins deux types d'onde gamma (Galambos, 1992) : les ondes gamma évoquées par le stimulus et les ondes gamma induites par le stimulus. Les ondes gamma induites (30-80Hz) consistent en des bouffées oscillatoires dont la latence par rapport au stimulus varie contrairement aux ondes gamma évoquées.

3.2.2.2. Les oscillations gamma

Une théorie suggère que la synchronisation rythmique des décharges neuronales dans la bande de fréquence gamma (environ 40Hz) pourrait fournir le liage spatial et temporel nécessaire permettant ainsi à différentes régions cérébrales de construire un percept cohérent (Singer & Gray, 1995). Ainsi, cette hypothèse rendrait compte de la construction de percepts lors de la perception, mais d'autres auteurs ont élargi cette hypothèse à l'émergence des représentations en mémoire. Selon Tallon-Baudry et Bertrand (1999), les ondes gamma induites déjà impliquées dans les activités de liage perceptif pourraient aussi sous-tendre la notion de représentation des objets. Dans une première étude, Tallon-Baudry et Bertrand (1997) présentaient deux conditions à des sujets. Dans une première condition, les sujets voyaient soit des stimuli constitués de tâches noires sans signification sur un fond gris ou des stimuli dont les tâches noires formaient un chien (dalmatien) ou enfin des stimuli dont les tâches noires étaient distordues. Les sujets devaient simplement compter les stimuli avec les tâches noires. Dans une deuxième condition, les sujets devaient compter le nombre de chiens qui étaient orientés vers la gauche. Il faut préciser que dans cette condition, les sujets étaient entraînés à détecter le stimulus cible. La différence principale entre ces deux conditions portait sur le fait que dans la condition 2, les sujets avaient une représentation du chien qu'il devait détecter. Les auteurs ont observé une augmentation importante de l'activité gamma-induite dans la condition 2 plus importante pour le stimulus cible que pour ce même stimulus présenté dans la condition 1. Les auteurs en ont déduit que cette activité gamma induite pouvait refléter l'activation top-down d'une représentation interne du stimulus. Les mêmes auteurs (Tallon-Baudry, Bertrand, Peronnet & Pernier, 1998) ont réalisé une autre étude qui comportait deux conditions. Dans une condition, les sujets voyaient une forme visuelle sans signification pendant 400 ms puis la forme disparaissait pendant 800 ms. Suite à ce délai, la même forme ou une autre forme apparaissait et les sujets devaient dire si les deux formes présentées étaient identiques. Dans l'autre condition, une forme était présentée pendant 400

ms mais après le délai de 800 ms, les sujets devaient observer si le point de fixation changeait de couleur. La deuxième condition ne faisait pas appel à la mémoire contrairement à la première, où les sujets devaient conserver une représentation de la forme visuelle en mémoire à court terme pour réaliser la tâche. Les auteurs ont montré dans la condition 1 la présence d'une activité gamma associée au délai de rétention de la forme visuelle par rapport à la condition 2 au niveau de l'électrode occipitale O1. L'activité dans la bande gamma semble ainsi liée à l'émergence et au maintien à court terme des représentations en mémoire. De plus, la topographie de cette activité semble indiquer que la trace mnésique d'un stimulus visuel est maintenue dans les aires visuelles, celles-là même qui sont activées lors de la présentation du stimulus (Jensen, Kaiser & Lachaux, 2007 ; Tallon-Baudry & Bertrand, 1998). Cela n'exclut toutefois pas l'intervention d'aires associatives (Driver & Noesselt, 2008).

D'autres études réalisées à l'aide d'électrodes intra-cérébrales chez des patients épileptiques ont montré que l'activité gamma dans l'hippocampe et la synchronisation entre le cortex rhinal et l'hippocampe étaient corrélées avec un encodage réussi en mémoire à long terme (Fell *et al.*, 2001). Des auteurs ont testé chez des sujets sains à l'aide de la MEG la contribution des bandes de fréquence gamma et thêta dans les processus mnésiques (Osipova *et al.*, 2006). Ils ont montré à des sujets des images de lieux ou de bâtiments inconnus lors d'une phase d'encodage. Les mêmes sujets ont participé à la phase de reconnaissance ayant eu lieu 5 à 10 minutes après l'encodage. Les sujets ne devaient répondre aux items seulement s'ils étaient sûrs de les avoir vus ou non. Les auteurs ont montré que l'encodage et la récupération étaient associés à des modulations de l'activité oscillatoire dans la bande fréquence gamma, mais aussi dans la bande de fréquence thêta. Un encodage réussi, pour lequel la reconnaissance ultérieure était correcte, était accompagné d'une augmentation de la magnitude de l'activité gamma dans les régions occipitales (BA18/19). L'activité gamma était plus importante pour les items correctement reconnus par rapport à ceux ayant été oubliés par rapport à la phase d'encodage. Les auteurs ont également montré que la magnitude de l'activité thêta était modulée par la tâche plus précisément dans le lobe temporal droit. L'augmentation d'activité dans la bande gamma durant l'encodage et la récupération en mémoire déclarative reflèterait une augmentation de la synchronisation neuronale alors que, selon les auteurs, les oscillations thêta pourrait jouer un rôle dans la plasticité synaptique et permettrait de faciliter l'encodage.

Ces différentes études démontrent que l'encodage et la récupération en mémoire à long terme et donc l'activation d'une représentation peuvent être associées avec l'augmentation d'activité

dans la bande gamma. Cette activité électrophysiologique pourrait donc constituer le substrat neuronal de l'activation des traces en mémoire et pourrait de plus supporter l'hypothèse d'une vision distribuée de la mémoire du fait des propriétés de ce type d'activité neuronale. Il faut toutefois rester prudent quant à cette interprétation car les activités oscillatoires dans la bande gamma ne sont pas les seules à être mises en évidence dans des tâches de mémoire. En effet la bande thêta est aussi impliquée dans la mémoire à court terme et dans la mémoire à long terme. Il faut donc imaginer des interactions plus complexes entre les différentes fréquences oscillatoires pour rendre compte des différents processus mnésiques plutôt que de penser un isomorphisme pur entre domaine oscillatoire et processus cognitif.

3.2.2.3. Une vision interactive: l'hypothèse du « process binding »

D'autres auteurs (Klimesch, Freunberger & Sauseng, 2010) proposent un type de liage légèrement différent qu'il nomme « *process binding* » ou encore « *large scale integration* ». Selon eux, les oscillations dans les différentes bandes de fréquence refléteraient différents types de processus cognitifs et ces différentes oscillations se synchroniseraient si la tâche requiert une co-activation ou une intégration des différents processus cognitifs. Klimesch et collaborateurs suggèrent que chaque oscillation dans les bandes de fréquence basse (4-14Hz) semble sous-tendre des processus de contrôle dans deux systèmes de mémoire : la mémoire de travail et le système de connaissance. Ce système ne trouve pas de description dans un registre cognitif, mais semble issu de la démarche des neurosciences cognitives, ce serait un système de mémoire déclaratif se rapprochant d'un système de mémoire sémantique comme a pu le décrire Tulving (1991). Ces processus de contrôle permettraient l'accès et la récupération des informations stockées en mémoire. Les oscillations dans la bande thêta (4-6Hz) seraient associées à des processus d'accès et de récupération d'informations nouvelles et d'informations épisodiques alors que les oscillations dans la bande alpha haute seraient liées à des processus d'accès et de récupérations d'informations déjà bien établies en mémoire. Pour ces mêmes auteurs, les synchronisations gamma ne reflètent pas un processus mnésique spécifique au sens que ce type d'oscillations n'est pas lié à un type de mémoire en particulier.

Bien que l'implication des oscillations gamma et thêta ne soit plus à démontrer dans le domaine de la mémoire, leur rôle fonctionnel reste encore à préciser. Certains auteurs (Gruber, Tsivilis, Giabbiconi & Müller, 2008) ont tenté récemment d'expliquer leur rôle dans le cadre des processus de familiarité et de recollection. Dans la première phase de leur expérience, les sujets voyaient des images d'objets pour lesquelles ils devaient dire si les

objets tenaient ou non dans une boîte à chaussure. Lors de la seconde phase, les sujets devaient juger si chaque item était ancien ou nouveau. Lorsque les items étaient catégorisés comme ancien, les sujets devaient se rappeler quelle classification ils avaient fait lors de l'encodage. Les auteurs ont observé une réponse plus forte d'activité gamma induite (210-330 ms) pour les items correctement catégorisés comme anciens par rapport aux nouveaux items (Osipova *et al.*, 2006). Par contre il n'existait pas de différence d'activité gamma parmi les items anciens, pour lesquels les sujets se rappelaient de la catégorisation qu'ils avaient faites lors de l'encodage et ceux pour lesquels ils ne se souvenaient pas. A l'inverse, il existait une différence de réponse dans la bande thêta (600-1200 ms) liée au souvenir de la catégorisation réalisée lors de l'encodage. Gruber et collaborateurs ont proposé suite à cette expérience que les modulations d'activité gamma induite pourraient refléter des processus associatifs précoces liés aux processus de familiarité alors les modulations d'activité thêta induite, plus tardives, reflèteraient des mécanismes de récupération de type recollection. Cette expérience permet d'affiner les relations entre mécanismes oscillatoires cérébraux et processus mnésiques. Les auteurs ont essayé de faire le lien entre processus mnésiques (familiarité/recollection) et activités oscillatoires plutôt qu'entre stocks mnésiques ou accès à un stock et activités oscillatoires. Cette vision fonctionnaliste permet ainsi de mettre à jour des modulations oscillatoires plus fines en lien avec des processus cognitifs.

Nous allons maintenant présenter des travaux qui ont étudié l'effet de la multimodalité en mémoire à l'aide de paradigmes divers. L'objectif étant de montrer qu'une manipulation sensorielle modifie le mode de récupération des connaissances, que l'accès à ces connaissances se fassent de manière implicite ou explicite.

3.3. Interaction multimodale en mémoire

Manipuler l'encodage module la récupération dans des tâches très diverses (Murray *et al.*, 2004; Lehmann & Murray, 2005; Brunel *et al.*, 2009; Brunel *et al.*, 2010; Giard & Perronnet, 1999; Fort *et al.*, 2002; Von Kriegstein & Giraud, 2006). Nous allons faire une brève revue de travaux ayant montré des modulations liées aux manipulations multimodales lors de l'encodage dans des tâches de mémoire à court terme et à long terme, ayant utilisé des accès implicites ou explicites.

3.3.1. Récupération liée au contexte multimodal

Il a déjà été montré qu'un apprentissage multimodal conduit à de meilleures performances qu'un apprentissage unimodal (Seitz & Dinse, 2007 ; Seitz *et al.*, 2006 ; Seitz & Watanabe, 2005 ; Giard & Perronnet, 1999). Le fait qu'un apprentissage multimodal soit plus robuste qu'un apprentissage unimodal, peut s'expliquer par la présence d'un contexte d'encodage plus riche qui serait réactivé lors de la récupération. En effet, la richesse du contexte d'encodage peut faciliter la récupération explicite du contenu d'une expérience (Paivio, 1986).

Dans une étude en tomographie par émission de positrons (Nyberg, Habib, McIntosh & Tulving, 2000), des sujets devaient apprendre des mots qui étaient associés ou non à des sons lors d'une phase d'encodage. Les sujets devaient dans une seconde phase reconnaître visuellement les mots qui leur avaient été présentés lors de l'encodage. Les résultats indiquent que la reconnaissance des mots active des régions du cortex auditifs qui étaient activés lors de l'encodage alors qu'il n'était pas demandé aux sujets de récupérer l'information auditive. Ces observations suggèrent que les composants auditifs de l'évènement multi-sensoriel sont stockés au sein du cortex auditif, et qu'ils pourraient être réactivés lorsque des éléments de la même trace sont activés. Dans une étude similaire réalisée en IRMf (Wheeler, Petersen & Buckner, 2000), des sujets devaient apprendre des images et des sons, ensuite un test de rappel leur était administré. Les résultats indiquent que lorsque les sujets doivent se rappeler d'items visuels et auditifs, une partie des régions activées lors de l'encodage est réactivée lors de la récupération de ces mêmes informations, ce qui corrobore les résultats de l'expérience précédente. Ces données ne suggèrent toutefois pas que les performances sont meilleures lorsque l'encodage est multimodal, cela montre qu'un encodage multimodal s'accompagne d'une réactivation automatique. Ceci est une condition nécessaire pour voir si cette réactivation automatique facilite les performances en récupération.

En utilisant des associations de stimuli sans signification, d'autres études montrent des résultats semblables. En IRMf, Tanabe, Honda et Sadato (2005) montrent que la présentation d'items auditifs préalablement associés à des items visuels pouvait réactiver des régions visuelles et réciproquement. Dans un paradigme de type conditionnement, Meyer et collaborateurs (2007) ont présenté à des sujets des stimuli visuels dans une première phase puis des stimuli auditifs dans une seconde phase. Dans une troisième phase, les auteurs ont présenté les stimuli visuels en même temps que les stimuli auditifs. Enfin durant une quatrième phase, les sujets pouvaient voir la même association audiovisuelle que lors de la troisième phase, soit les stimuli visuels seuls. Lors de la présentation des stimulations visuelles seules, les auteurs ont observé des activations dans le cortex visuel mais également

dans le cortex auditif droit. Ces résultats démontrent l'implication des aires sensorielles auditives dans la perception de stimulations visuelles, cette observation reflète la formation d'associations multi-sensorielles et ne peut être attribuée à une simple sensation auditive.

Lors d'une expérience multi-sensorielle, une association peut se former entre des stimuli de différentes modalités pour créer une trace mnésique. Lorsque l'un des items est présenté dans ultérieurement, cela entraîne entre autre l'activation automatique des régions cérébrales impliquées dans le traitement du stimulus de l'autre modalité (Meyer *et al.*, 2007 ; Baier *et al.*, 2006). Cette activation multimodale automatique peut être dû à l'association préalable existant entre les deux stimuli, on peut même supposer qu'elle facilite la récupération en renforçant le poids de la trace. Nous allons maintenant présenter des études qui ont manipulé des associations multimodales lors de l'encodage et ont observé leur effet lors d'une phase de récupération.

3.3.2. Effet d'un encodage multimodal sur la récupération

Nous allons présenter ici différents travaux comportementaux permettant de comprendre le lien qu'il peut exister entre les modalités sensorielles en mémoire et l'effet de ces liens à la fois sur l'encodage et sur la récupération. Pour cela, nous allons détailler quelques études qui manipulent les effets de congruence sémantique audiovisuelle. L'idée sera de présenter des résultats comportementaux, ainsi que des paradigmes expérimentaux ayant permis d'observer ces effets.

3.3.2.1. Dans les études à court terme

Dans un ensemble d'expériences, Chen et Spence (2010) ont manipulé l'effet de congruence sémantique audiovisuelle. Leur hypothèse était de voir si la présentation d'un son congruent ou non facilitait l'identification d'une présentation visuelle masquée. Dans l'expérience 1, les auteurs présentaient une image à identifier pendant 27 ms qui était immédiatement masquée par la suite. En même temps que l'image, un son congruent ou non, ou un bruit blanc (condition contrôle) était diffusé. Les auteurs ont montré que les sujets identifiaient avec un pourcentage de bonnes réponses significativement plus élevé les images accompagnées d'un son sémantiquement congruent. De plus la présentation d'un son non congruent entraînait une diminution des performances par rapport à la condition contrôle. Les auteurs mettent en évidence un effet facilitateur de la condition de congruence sémantique et un effet perturbateur de la condition de non congruence en utilisant une condition de contrôle multimodale (bruit blanc) et non uniquement une condition de contrôle unimodale. Dans une

deuxième expérience, les auteurs ont montré qu'en utilisant une condition contrôle unimodale, les performances ne différaient plus entre la condition de non congruence et la condition contrôle (Molholm *et al.*, 2004). Cette différence entre les expériences 1 et 2 pourrait s'expliquer par le fait que les sujets soient dans un état d'alerte moins important dans les essais unimodaux par rapport aux essais multimodaux. Il a également été montré que l'interférence intermodale est moins forte que l'interférence intramodale à des stades précoces de traitement (Laurienti *et al.*, 2004). Enfin, dans les expériences 3 et 4, les auteurs ont fait varier le SOA entre l'apparition de l'image et du son ; ils montrent que pour un SOA de 307ms, l'effet facilitateur de la congruence sémantique est toujours présent, alors que pour un SOA de 533ms cet effet n'est plus présent. Il existerait ainsi une fenêtre temporelle où les représentations sensorielles resteraient activées, il est intéressant de noter qu'il n'est pas nécessaire de respecter une contiguïté temporelle stricte pour observer ces phénomènes intermodaux. Chen et Spence interprètent ces résultats selon une conception structuraliste, à l'aide du module de mémoire conceptuelle à court terme (Potter, 1993) afin d'expliquer le traitement sémantique des items visuels. Ce module contiendrait un buffer à court terme qui servirait d'interface entre mémoire et perception. Une fois un item visuel identifié, sa signification serait accessible et maintenue temporairement. Cette représentation sémantique serait alors structurée au sein d'un contexte cohérent et consolidé en mémoire, ou alors déstructurée et rapidement oubliée. La présentation d'un item auditif congruent entraînerait une facilitation de l'accès au sens de l'item visuel alors qu'un item auditif non congruent perturberait l'accès au sens. Cette interprétation est plausible pour expliquer un phénomène à court terme, mais il devient plus difficile d'expliquer des phénomènes inter-sensoriels à long terme (Murray *et al.*, 2004 ; Lehmann & Murray, 2005). On pourrait également envisager que l'activation du composant visuel à identifier soit renforcée par l'activation d'un composant auditif congruent et que ces activations conjointes soient également possibles dans des délais plus longs. Le lien crée entre ces composants permettrait d'expliquer la facilitation inter-sensorielle à long terme pour des connaissances sémantiques (Vallet *et al.*, 2010).

Dans une autre expérience, Schneider et collaborateurs (2008) se sont intéressés aux effets de congruence sémantique en utilisant cette fois un paradigme d'amorçage inter-sensoriel à court terme. Les auteurs ont testé l'influence réciproque de percepts auditifs et visuels d'objets naturels car peu d'études jusqu'ici se sont intéressées aux effets entre modalités sensorielles à l'aide de l'amorçage. Dans une première expérience, les auteurs ont proposé aux sujets de catégoriser des sons cibles après avoir vu des images qui pouvaient ou non être reliées

sémantiquement aux sons. La tâche des sujets était de dire pour chaque item, si celui-ci tenait ou non dans une boîte à chaussure. Les auteurs ont également manipulé la modalité d'amorçage, il y avait en effet une condition unimodale (amorce sonore – cible sonore) et une condition multimodale (amorce visuelle – cible sonore). Les résultats de cette première étude mettent en évidence que les items auditifs cibles sont catégorisés plus rapidement lorsqu'ils sont sémantiquement congruents avec l'amorce, quelle que soit sa modalité. Dans une seconde expérience, les auteurs ont essayé de reproduire les effets obtenus lors de la première étude mais en utilisant cette fois, une cible visuelle et une amorce auditive ou visuelle. Afin de rendre leur tâche comparable, les auteurs ont dégradé les images présentes dans leur base de données, jusqu'à obtenir des pourcentages de réponse correcte comparables à ceux des sons présentés en cible dans la première étude. Deux facteurs étaient alors manipulés, la congruence sémantique et le type d'amorçage, unimodal ou intermodal. Une nouvelle fois, les résultats montrent un effet d'amorçage dû à l'effet du facteur congruence. Les items visuels cibles sont catégorisés plus rapidement lorsqu'il existe une congruence sémantique avec l'amorce, et ce quelle que soit la modalité. Cette expérience confirme les résultats précédents avec un paradigme différent, toutefois elle ne permet pas de préciser à quel niveau les interactions multimodales ont lieu, on peut éventuellement supposer au vu des travaux de la littérature dans le domaine que l'hypothèse d'interactions relativement précoces entre modalités sensorielles est plausible (Giard & Perronnet, 1999 ; Molholm *et al.*, 2004). On peut penser que le traitement perceptif a lieu dans les aires modalité spécifiques, mais l'information entre ces aires se propage à des niveaux de traitement précoce par un possible mécanisme de *binding*, pour aboutir à une représentation multi-sensorielle (Senkowski *et al.*, 2008). Dans une tâche comparable, Greene, Easton et LaShell (2001) ont testé le transfert de la modalité visuelle vers la modalité auditive et réciproquement en utilisant des séquences vidéos et auditives. Dans la phase d'étude, les sujets voyaient des séquences vidéo ou auditives d'items et dans la phase de test, les sujets devaient identifier des items dégradés, pixellisé pour les séquences visuelles et bruités pour les séquences auditives. Si les sujets étudiaient la séquence vidéo dans la première phase, ils devaient identifier la séquence auditive du même item dans la seconde phase, afin de tester le transfert d'une modalité vers une autre. Les résultats ont mis en évidence que les items étudiés dans la modalité auditive amorçaient les items à identifier dans la modalité auditive mais non dans la modalité visuelle, alors que les items étudiés dans la modalité visuelle amorçaient les items à identifier dans la modalité auditive et visuelle. Une partie de ces résultats sont en contradiction avec ceux obtenus par Schneider et collaborateurs (2008), en effet, ici, les items étudiés dans la modalité

auditive n'amorcent pas les items visuels dégradés alors que dans l'étude précédente c'est l'inverse qui est observé. Cette différence peut être expliquée par la quantité de dégradation des images dans l'expérience de Schneider et collaborateurs (2008), en effet selon ces auteurs, le fait que les images soient difficiles à identifier, permettrait à l'information communiquée par le canal auditif de faciliter l'identification des objets.

Noppeney et collaborateurs (2007) ont réalisé une tâche d'amorçage intersensoriel à court terme dans laquelle ils présentaient aux sujets en amorce une image ou un mot écrit d'animaux ou d'outils et en cible un son ou un mot prononcé oralement. Dans cette étude réalisée en IRMf, les auteurs ont voulu étudié les effets d'une présentation visuelle sur la catégorisation de stimuli auditifs. Les sujets devaient catégoriser la cible selon qu'elle représente un objet ou un animal plus ou moins lourd que 4kg. Trois facteurs étaient manipulés dans cette expérience : la congruence entre la cible et l'amorce, la nature de l'amorce (image ou mot écrit) et la nature de la cible (son ou mots prononcés). Les résultats comportementaux ont mis en évidence un effet de congruence et un effet de la nature de la cible. On observe ici une interférence liée à la nature de la relation entre l'amorce et la cible, les sujets étant plus lents lorsque l'amorce et la cible n'étaient pas sémantiquement reliées. Il existe donc un effet d'amorçage visuo-auditif et ce quel que soit le matériel. L'absence d'interaction de la congruence avec le type de matériel suggère que l'image ou le mot écrit présentés en amorce produit le même effet d'amorçage quel que soit le type de cibles. On retrouve ici un résultat observé dans les expériences précédemment décrites avec en plus l'idée selon laquelle il n'existe pas de différences de nature entre matériel verbal et non verbal, en effet il semble qu'un composant lexical ou non est en mesure d'activer un autre composant lexical.

3.3.2.1.2. Dans les études à long terme

Von Kriegstein et Giraud (2006) ont testé l'idée selon laquelle la présentation d'items unimodaux devrait activer des réseaux neuronaux intermodaux ayant été formés lors d'un apprentissage associatif antérieur et augmenter ainsi les performances comportementales. Les auteurs ont utilisé deux groupes de sujets, un groupe apprenait à associer des voix avec des visages, les voix apportaient des informations redondantes à propos du genre, du groupe ethnique et de l'âge. Le second groupe apprenait à associer des voix avec des noms. A l'inverse des visages, les noms sont reliés arbitrairement aux voix. Cet apprentissage correspond à une association congruente, mais arbitraire. De plus, les auteurs ont utilisé des images et des sonneries de téléphone portable comme contrôle, afin de tester l'importance

d'une source physique commune des sons et des images pour la formation d'associations multi-sensorielles précoces et la reconnaissance unimodale auditive ultérieure. Bien qu'une sonnerie et un téléphone portable renvoient à une source multimodale commune, leur association est arbitraire et n'importe quelle sonnerie peut être associée à n'importe quel téléphone portable. Dans la première partie de l'expérience, les sujets écoutaient et identifiaient les stimuli auditifs (voix et sonneries) alors que l'activité neuronale était enregistrée. Dans une seconde partie, les participants apprenaient à associer ces voix et ces sonneries avec les stimuli visuels. Un groupe de sujets apprenaient des associations voix-visages et sonneries-portables alors qu'un autre groupe de sujets apprenaient à associer des voix-noms de personnes et des sonneries-noms de téléphones. Dans la troisième partie de l'expérience, la présentation des items auditifs seuls était répétée. Les auteurs ont analysé les résultats comportementaux et de neuroimagerie correspondant aux sessions de reconnaissance des stimuli auditifs avant et après l'apprentissage. Les résultats mettent en évidence qu'une association entre une voix et un visage lors de la phase d'apprentissage augmente l'activité due à la présentation de la voix seule dans le gyrus fusiforme (FFA), le précuneus, les cortex pariétaux bilatéraux et le cortex préfrontal droit. Ces régions sont impliquées dans le traitement des personnes familières, indépendamment de la modalité considérée. De plus, l'augmentation d'activité dans le FFA durant la reconnaissance des voix intervient uniquement dans le groupe d'apprentissage voix-visage. L'activation des cortex visuels spécifiques en réponse aux inputs auditifs sont donc seulement observés après l'exposition de stimuli multi-sensoriels redondants (visages-voix), mais pas après l'exposition à de l'information multimodale arbitraire (voix-nom, sonnerie-nom) même si la paire de stimulus était congruente au niveau sensoriel (sonnerie-téléphone). Si l'on s'intéresse maintenant à l'influence de l'association visage-voix sur la reconnaissance ultérieure de la voix, les résultats comportementaux montrent une augmentation de la reconnaissance des voix de 14% alors que l'association voix-nom n'augmente que de 5% les performances par rapport aux performances de la première phase.

Dans une étude réalisée en EEG, Murray et collaborateurs (2004) ont utilisé un paradigme de reconnaissance continue (voir FIG. 32), dans lequel les sujets voyaient des images qui pouvaient ou non être accompagnées d'un son. Ces images représentaient des animaux ou des objets et étaient présentées avec le son naturel produit par ceux-ci. Pour chaque image, les sujets devaient dire si oui ou non l'item était présenté pour la première fois. Les auteurs manipulaient dans cette expérience la condition de la première présentation des images. Les

items pouvaient être présentés, lors de leur première apparition, uniquement visuellement, ou accompagné du son leur étant naturellement associé. Pour la seconde présentation, les résultats montrent que les images étaient mieux reconnues lorsqu'elles étaient présentées initialement avec un son que lorsqu'elles étaient présentées uniquement dans la modalité visuelle. De plus, l'amplitude des potentiels évoqués générés par les items présentés dans la condition multimodale la première fois était plus importante que ceux générés par les items présentés dans la condition unimodale. Cette différence serait sous-tendue par le complexe occipito-latéral (COL), habituellement impliqué dans le traitement visuel des objets. Cette étude est digne d'intérêt car elle montre qu'une seule présentation multi-sensorielle suffit à améliorer les performances en reconnaissance. Toutefois, elle ne permet pas de conclure en termes de spécificité d'encodage multi-sensoriel. Des travaux menés par la même équipe (Lehmann & Murray, 2005) ont amené des éléments de réponse.

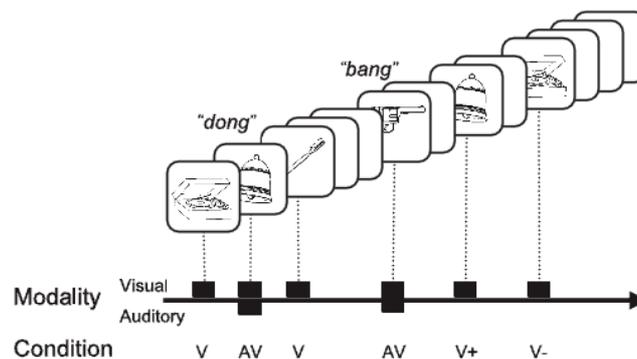


FIG. 32 - Procédure utilisée dans l'expérience de Murray et collaborateurs (2004)

En utilisant le même paradigme qu'en 2004, Lehmann et Murray (2005) ont comparé une condition où la première présentation de l'image était associée à un son pur (1000Hz) et une condition où la première présentation était unimodale. Les résultats de cette première expérience montrent que les pourcentages de bonnes réponses sont plus élevés lorsque les images sont présentées visuellement que lorsqu'elles sont associées à un son pur. Ce résultat indique que le contexte multi-sensoriel n'explique pas à lui seul l'amélioration des performances. En effet, le fait que l'image soit associée à un son n'augmente pas nécessairement la probabilité de récupérer cet item en mémoire, bien au contraire. La probabilité de récupération serait dépendante des attributs sémantiques que partagent les stimuli auditifs et visuels. Selon les auteurs, la présentation répétée d'un son pur avec des images entraîneraient une dégradation des représentations de ces images en mémoire, ce qui expliquerait qu'elles soient moins bien reconnues. Dans une deuxième expérience Lehmann et Murray ont comparé une condition dans laquelle la première présentation de l'image était

unimodale et une condition où la première présentation était associée à une vibration tactile (50Hz). Dans cette expérience, les auteurs n'ont relevé aucune différence entre ces deux conditions, l'ajout d'un contexte multi-sensoriel ne permettrait donc pas d'améliorer les performances du fait du manque de congruence entre ces associations. Dans une seconde série d'expérience, les auteurs ont manipulé le caractère sémantique des associations multi-sensorielles. Dans une condition congruente, la première présentation de l'image pouvait être associée au son naturel, dans une seconde condition non congruente, la première présentation de l'image pouvait être associée à un autre son (image d'un chien + miaulement d'un chat), dans une dernière condition la première présentation était uniquement visuelle. Les performances sont meilleures dans la condition congruente, ce qui conforte l'hypothèse de congruence des attributs sémantiques des différentes modalités sensorielles afin de récupérer plus facilement les items en mémoire. Les résultats de Lehmann et Murray sont particulièrement intéressants même si on peut s'interroger sur le fait que les mêmes images soient présentées plusieurs fois lors de la reconnaissance, ce qui est nécessaire pour pouvoir réaliser une acquisition EEG, mais ce qui peut également biaiser les performances des sujets, en augmentant la familiarité des items déjà reconnus auparavant.

Dans une expérience récente de Vallet et collaborateurs (2010), les auteurs ont présenté des images ou des sons d'animaux et d'objets manufacturés que les sujets devaient catégoriser le plus rapidement possible, lors d'une première phase. La moitié des items étaient masqués à l'aide d'un bruit blanc pour les images et d'un masque visuel pour les sons. Dans la seconde phase, les sujets devaient catégoriser de nouveau les items qui leur étaient présentés mais ils devaient cette fois catégoriser le même concept sous format auditif dans la seconde phase. Un tiers des concepts présentés dans la seconde phase était des concepts nouveaux. Les auteurs ont montré que les items anciens étaient catégorisés plus rapidement que les items nouveaux, ce qui témoigne selon eux d'un transfert de la modalité visuelle vers la modalité auditive entre la phase d'encodage et celle de récupération. De plus, les auteurs avaient utilisé un masque, de telle sorte qu'un tiers des items vus lors de l'encodage. Les résultats obtenus montrent que les items anciens masqués sont catégorisés plus lentement que les items anciens non masqués. La manipulation sensorielle lors de l'encodage a donc perturbé la catégorisation de ces items dans la phase de récupération, ce qui montre que les connaissances en mémoire sont basées sur les modalités sensorielles (Brunel *et al.*, 2010; Versace *et al.*, 2002 ; Versace *et al.*, 2009)

3.4 En résumé

Ces différentes études à court ou à long terme ont montré clairement qu'une association entre un item visuel et un item auditif conduisait à des modulations du traitement, même si cette association n'avait été faite qu'une fois (Murray *et al.*, 2004 ; Lehmann & Murray, 2005). Ces effets intermodaux semblent donc robustes et reproductibles selon les études et les paradigmes expérimentaux utilisés. Peu d'études proposent un cadre théorique explicatif à ces effets multimodaux en mémoire. Mis à part le modèle de Potter (1993) qui tentent de rendre compte des effets de congruence sémantique à l'aide d'un buffer à court terme, donc selon une approche clairement structuraliste, nous n'avons pas à notre disposition de modélisation des interactions multimodales lors de la récupération en mémoire.

3.5 Problématique

Dans cette série d'expériences, nous nous proposons d'étudier le rôle de la multimodalité à l'aide de deux tâches qui reposent sur un accès différent en mémoire : une tâche de reconnaissance, sous tendue par des mécanismes implicites et explicites (Yonelinas, 2002) et une tâche de catégorisation, sous tendue par des mécanismes implicites. Beaucoup de travaux ont montré les effets de la multimodalité en mémoire (Murray *et al.*, 2004 ; Lehmann & Murray, 2005), mais sans jamais faire véritablement référence à un modèle cognitif précis. De plus, aucun des travaux que nous avons présentés dans le Chapitre 1 n'apporte d'explications cohérentes sur les effets de la multimodalité. L'intérêt sera de préciser les relations multimodales en mémoire, selon un modèle cognitif de mémoire à traces multiples, en tenant compte des composantes explicites et implicites de traitement.

Versace et collaborateurs postulent que l'activation des composants de la trace associée en mémoire est automatique et interviendrait dès la présentation d'un des composants de cette trace. De plus, nous avons vu qu'une manipulation sensorielle lors de l'encodage avait des effets lors de la récupération (Murray *et al.*, 2004 ; Brunel *et al.*, 2009 ; Vallet *et al.*, 2010). Ainsi, présenter des items visuellement avec un son congruent va faciliter la récupération lors d'une seconde présentation de l'item visuel seul (Murray *et al.*, 2004). Nous pensons qu'un encodage multimodal congruent devrait faciliter la récupération tant pour le processus de familiarité que pour celui de recollection. A l'inverse, présenter un item visuellement avec un son non congruent, sans signification, devrait perturber la récupération de l'item visuel seul dans une seconde phase par rapport à une condition de base (Lehmann & Murray, 2005). Nous pensons que cet effet d'interférence sensoriel peut s'expliquer par la perturbation de l'activation des autres composantes de la trace et rendre la récupération

ultérieure de la trace plus difficile. De plus, cet effet devrait s'observer pour des étapes de traitement précoce, nous pensons qu'il devrait jouer essentiellement sur le processus de familiarité. Enfin, ce qui pourrait expliquer cet effet d'interférence sensoriel repose sur le fait que lors de l'encodage, une partie des items soit associée avec le même son (bruit blanc). Lors de la récupération, ces items seraient moins discriminables par rapport à des items présentés tous avec des sons différents.

Nous allons donc dans un premier temps mettre au point un paradigme de catégorisation que nous pourrons ensuite appliquer à une tâche de reconnaissance. En effet, ces deux types de tâche font appel à des mécanismes implicites (catégorisation/reconnaissance) et explicite (reconnaissance). Nous allons dans un premier temps tester un paradigme original de catégorisation en deux phases permettant de montrer l'avantage d'une condition multimodale d'encodage incident par rapport à une condition unimodale. Ensuite, nous utiliserons ce paradigme sur une tâche de reconnaissance. L'intérêt d'avoir un seul paradigme pour tester la catégorisation et la reconnaissance est de pouvoir contrôler parfaitement les conditions expérimentales. Nous allons présenter des items visuels ou auditifs dans la condition unimodale et des items auditifs et visuels dans la condition multimodale. Les items auditifs et visuels ne seront pas présentés simultanément dans la condition multimodale afin d'éviter que certains items soient composés de deux modalités et d'autres items que d'une seule. Le format de présentation dans la phase d'encodage sera fait sous forme d'images de concepts pour les items visuels et sous forme de sons naturels de concepts pour les items auditifs. Dans la phase de test, nous pourrons utiliser par exemple les noms des concepts, la catégorisation (vivant/non vivant) se fera donc sur l'étiquette verbale représentant les concepts (Noppeney *et al.*, 2007). L'étiquette verbale est un élément de la trace mnésique, l'activation des dimensions sensorielles devrait permettre l'activation de la composante lexicale du concept. Nous utiliserons également une condition interférente, représentée par un bruit blanc, dont l'association avec un item visuel est censée perturber l'activation des dimensions sensorielles associées en mémoire (Brunel *et al.*, 2010 ; Vallet *et al.*, 2010). En plus du statut de condition contrôle dans certaines expériences (Lehman & Murray, 2005), l'interférence possède un statut particulier dans le modèle Act-in, car l'interprétation du mécanisme sous jacent est encore mal connue.

Selon nos hypothèses, dans une tâche de catégorisation, l'activation des composants de la trace devrait suffire à faciliter le traitement des items placés dans la condition multimodale par rapport aux autres conditions. Concernant la condition unimodale, bien que

les items n'aient pas été présentés dans plusieurs modalités, ils devraient réactiver automatiquement un composant sensoriel associé en mémoire au sein de la trace. A l'inverse, la présentation simultanée d'une interférence avec des items visuels ou auditifs devrait gêner la réactivation du composant associé en mémoire. Nous devrions donc obtenir des temps de réponses plus rapides pour les items placés dans la condition multimodale par rapport à ceux des autres conditions. Les items placés en condition unimodale devraient être catégorisés plus rapidement que les items placés dans la condition interférente.

Dans une tâche de reconnaissance, les items placés dans la condition multimodale, nécessitent l'intégration des différents composants de la trace. La performance devrait donc être meilleure pour ces items que pour ceux présentés dans les conditions unimodales et interférentes. Les items placés dans la condition unimodale devraient tout de même réactiver les composants de la trace, même s'ils n'ont pas été associés à un son lors de la première phase. Cette activation automatique des composants de la trace ne devrait se manifester uniquement dans la composante implicite de la reconnaissance, en augmentant la familiarité des items unimodaux à l'inverse des items présentés avec une interférence, pour qui l'activation des composants associés aurait été perturbée. Nous devrions donc obtenir des pourcentages de réponses correctes plus élevés pour les items placés dans la condition multimodale que pour des items placés dans la condition unimodale ou interférente. Dans la composante implicite de la tâche de reconnaissance, nous devrions observer un pourcentage de réponses correctes plus élevé pour les items placés dans la condition unimodale par rapport à ceux placés dans la condition interférente.

3.6 Partie expérimentale

Toutes les expériences réalisées ont été conçues et réalisées sur un ordinateur MacBook grâce au logiciel Psyscope (Cohen *et al.*, 1993). Les données de cette expérience ont été recueillies sur le même ordinateur au laboratoire EMC.

La majorité des données brutes obtenues lors de cette partie expérimentale est présentée en Annexes (voir page 241).

Expérience 1

Hypothèses

H1 : Un encodage multimodal favorise le traitement ultérieur de l'étiquette verbale du concept en comparaison à un encodage unimodal ou interférent.

H2 : La présentation simultanée d'une interférence sensorielle avec une dimension sensori-motrice relative à un concept perturbe la réactivation de la trace.

Méthode

Sujets

Soixante sujets ($M = 22.1$; $ET = 5.2$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

Stimuli visuels

Les images des concepts

Nous avons sélectionné trente-deux images faisant référence à des animaux et trente-deux images représentant des outils ou des instruments de musique. La plupart des images appartenaient à une base de données d'images de notre équipe, d'autres images ont été ajoutées via les moteurs de recherche internet « Picsearch ». Les images de la base de données possédaient une dimension standard de 10.3x7.7cm, ce qui donne à une distance de 40cm, un

degré d'angle visuel de $14.4 \times 10.9^\circ$. Les images ne correspondant pas à ce format, ont été redimensionnées tout en conservant la proportion initiale, à l'aide de l'application internet « PicNik ». Les images étaient présentées durant 1000 ms au centre de l'écran.

Les noms des concepts

Le nom de chaque concept vivant ou non vivant sélectionné pour cette expérience a été écrit en noire, avec une police Times New Roman et une taille de 54. Chaque nom a été placé au centre d'une diapositive à l'aide du logiciel Power Point. Ces noms ont été choisis sur la base des couples images/sons de notre base de données. Nous avons également réalisé un pré-test dans lequel nous nous sommes assuré que les noms associés aux images leur correspondaient bien.

L'interférence visuelle

Nous avons utilisé un dessin sans signification construit à partir de PHOTOSHOP, représentant des lignes de couleur différentes auxquelles nous avons appliqué une distorsion verticale. Cette image possédait les mêmes dimensions que les images des concepts. Ce masque visuel avait pour objectif de perturber l'activation des dimensions sensorielles à partir de la présentation d'un composant perceptif auditif (Vallet *et al.*, 2010).

Stimuli auditifs

Les sons des concepts

Nous avons sélectionné de la même façon trente-deux sons, seize représentaient des cris d'animaux et seize représentaient des sons émis par des outils ou des instruments de musique. Certains sons ont été directement repris d'une base de données existant dans notre équipe, les autres sons ont été récupérés via le moteur de recherche internet « FindSounds ». Chaque son était édité à l'aide du logiciel AUDACITY pour que leur durée soit égale à 1000 ms. Les sons étaient diffusés à l'aide d'un casque stéréo SENNEISER HD.

L'interférence auditive

Pour l'interférence auditive, nous avons généré un bruit blanc à l'aide du logiciel AUDACITY, d'une durée de 1000 ms. Ce son avait pour objectif de perturber l'activation des dimensions sensorielles à partir de la présentation d'un composant perceptif visuel (Vallet *et al.*, 2010).

Procédure et plan expérimental

L'expérience était composée de deux phases distinctes. Lors d'une première phase d'encodage incident, les sujets devaient catégoriser chacun des stimuli qui leur étaient présentés en fonction de sa catégorie d'appartenance (vivant/non vivant) en répondant sur deux touches distinctes. Lors de la deuxième phase, les sujets devaient de nouveau réaliser une tâche de catégorisation, mais cette fois sur des mots représentant soit des concepts étudiés lors de la première phase soit des concepts nouveaux. Nous avons utilisé dans cette expérience trois groupes expérimentaux, un groupe réalisait un encodage unisensoriel auditif, un groupe réalisait un encodage unisensoriel visuel et un autre groupe réalisait un encodage multi-sensoriel audiovisuel. Chaque encodage était réalisé de telle manière qu'un concept était présenté deux fois à un intervalle de 9 items en moyenne. Un concept apparaissait une première fois visuellement et une seconde fois visuellement dans le groupe d'encodage unisensoriel visuel, il apparaissait une première fois auditivement et une seconde fois auditivement dans le groupe d'encodage unisensoriel auditif et enfin il apparaissait une première fois auditivement puis une seconde fois visuellement dans le groupe d'encodage multi-sensoriel audiovisuel. De plus, chaque concept pouvait être présenté accompagné d'une interférence visuelle pour les items auditifs ou visuelle pour les items visuels. Dans le groupe d'encodage audiovisuel, le premier item auditif était interféré visuellement et le second item visuel était interféré auditivement. Le caractère ancien ou nouveau des concepts a été contrebalancé au sein de chaque groupe, nous avons également contrebalancé les items interférés dans chacun des groupes d'encodage. Cela nous a conduits à créer quatre listes expérimentales par groupe d'encodage, soit douze listes au total. Dans la seconde phase, les sujets catégorisaient 64 noms de concepts, dont 32 noms correspondaient à des concepts que les sujets avaient étudiés dans la première phase sous format visuel, auditif ou audiovisuel et 32 noms qui correspondaient à des concepts nouveaux (voir TAB. 15). Dans cette expérience, nous avons manipulé un facteur inter-sujet GROUPE à trois modalités : Visuel, Auditif ou Audiovisuel et deux facteurs intra-sujets, le facteur CATEGORIE à deux modalités : vivants et non vivants et le facteur ITEM à trois modalités : mot ancien interféré, mot ancien non interféré et mot nouveau.

		ITEM					
		ancien sans interférence		ancien avec interférence		nouveaux	
		vivant	non vivant	vivant	non vivant	vivant	non vivant
GROUPE	Visuel	8	8	8	8	16	16
	Auditif	8	8	8	8	16	16
	Audiovisuel	8	8	8	8	16	16

TAB. 15 – Répartition des stimuli en fonction des facteurs GROUPE et ITEM

Dans la phase d’encodage, les items étaient présentés dans un paradigme de catégorisation continue, constitué d’un écran blanc d’une durée de 500 ms puis d’un point de fixation présenté au centre de l’écran qui durait 500 ms. Ensuite le stimulus visuel ou auditif était présenté pendant une durée de 1000 ms. Lorsque l’item était interféré, le masque visuel ou auditif était diffusé simultanément que le stimulus d’intérêt. Il était demandé aux sujets de répondre le plus rapidement possible mais quelle que soit la réponse du sujet, la durée de présentation de l’item était constante pour favoriser l’encodage des concepts lors de cette phase. Si le sujet ne donnait pas de réponse dans un délai de 2000 ms après la disparition du stimulus, l’expérience passait à l’essai suivant. Il y avait un délai de 5 minutes entre la phase d’encodage incident et la phase de test où le sujet était invité à réaliser une tâche distractive. Dans la phase de test, un essai était constitué comme lors de la phase précédente, d’un écran blanc de 500 ms, puis d’un point de fixation de 500 ms, puis du mot présenté au centre de l’écran. Les mots restaient affichés jusqu’à ce que le sujet réponde. Les sujets devaient catégoriser les mots à l’aide des mêmes touches qu’ils avaient utilisé lors de la première phase. Il était indiqué aux sujets de répondre le plus rapidement possible en faisant le moins d’erreurs. Les sujets répondaient à l’aide des touches « a » et « p », ces touches étaient contrebalancées.

Résultats

Les latences supérieures à 2000 ms n’ont pas été prises en compte, seules les latences correspondant aux bonnes réponses ont été conservées pour l’analyse. Nous avons conduit une analyse de variance à mesures répétées sur les pourcentages de réponses correctes et une autre sur les latences de réponses (TAB. 16). Une analyse par item a été réalisée sur les latences de réponse.

	nouveau		ancien			
	vivant	non vivant	vivant		non vivant	
			interf	sans	inter	sans
AA	674	771	687	695	769	755
	20	31	23	25	34	30
VV	659	737	638	663	716	735
	28	27	27	25	36	38
AV	644	725	627	625	737	723
	20	24	20	19	28	24

TAB. 16 – Moyenne des temps de réponses et erreurs types en fonction des facteurs GROUPE, ITEM et CATEGORIE

L'analyse de variance conduite sur les pourcentages de réponses correctes ne révèle aucun effet significatif. L'analyse de variance conduite sur les latences de réponse révèle un effet du facteur CATEGORIE, par sujet $F_s(1,57) = 102.52$; $p < .01$; $\eta^2 = .64$, les mots représentant des concepts vivants (657 ms) sont catégorisés plus rapidement que les mots représentant des concepts non vivants (741 ms), on retrouve ce résultat par item $F_i(1,62) = 21.68$; $p < .01$; $\eta^2 = .26$. L'analyse ne révèle pas d'effet du GROUPE par sujet, $F_s(2,57) < 1$, mais il existe un effet par item, $F_i(2,124) = 29.88$; $p < .01$; $\eta^2 = .33$. Le groupe qui a suivi l'encodage audiovisuel (687ms) a catégorisé plus rapidement les mots que le groupe visuel (715ms), $F(1,62) = 14.75$; $p < .01$, et plus rapidement que le groupe auditif (751ms), $F(1,62) = 50$; $p < .01$. Nous ne notons pas d'interaction GROUPE*CATEGORIE par sujet, $F_s(2,57) < 1$, ni par item, $F_i(2,124) = 2.3$; $p = .10$. Nous ne relevons pas d'effet du facteur ITEM par sujet, $F_s(2,114) < 1$ mais il existe un effet tendanciel par item (voir FIG. 33), $F_i(2,124) = 2.81$; $p = .06$; $\eta^2 = .04$. Cette tendance se traduit par un avantage sur les temps de catégorisation des mots ayant été vus lors de l'encodage dans la condition interférente (703ms) par rapport aux mots nouveaux (728ms), $F(1,62) = 8.78$; $p < .01$. La différence entre les mots anciens placés dans la condition interférence et les mots anciens placés dans la condition sans interférence n'est pas significative, $F(1,62) = 1.79$; $p = .09$, tout comme la différence entre les mots anciens placés dans la condition sans interférence et les mots nouveaux, $F(1,62) < 1$. L'analyse ne révèle pas d'interaction ITEM*GROUPE, ni par sujet $F_s(4,114) < 1$ ni par item $F_i(4,248) = 1.5$; $p = .2$. Elle ne révèle pas non plus d'interaction CATEGORIE*ITEM par sujet $F_s(2,114) < 1$ et par item non plus $F_i(2,124) < 1$. Enfin, nous n'observons pas de double interaction CATEGORIE*ITEM*GROUPE, par sujet $F_s(4,114) < 1$ et par item $F_i(4,248) = 1.2$; $p = .3$.

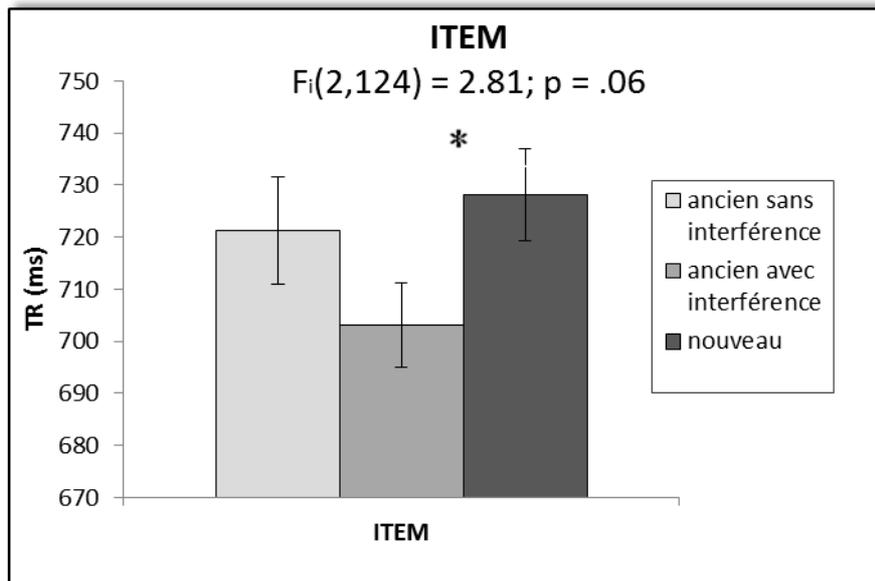


FIG. 33 - Représentation graphique de l'effet du facteur ITEM sur les temps de réponse (ms)

Discussion

Nous pensons mettre en évidence un avantage pour le groupe ayant suivi un encodage multi-sensoriel par rapport aux groupes ayant suivi un encodage unisensoriel. De plus, nous avons émis l'hypothèse qu'une interférence sensorielle présentée en même temps qu'une composante perceptive perturberait l'activation des dimensions sensori-motrices associées à cette composante. Ces hypothèses n'ont pas été confirmées, en effet nous n'avons pas relevé d'interaction entre les facteurs ITEM et GROUPE, tant sur les latences de réponse que sur les pourcentages de bonnes réponses. L'analyse par sujets ne nous a donc pas permis de valider nos hypothèses, toutefois l'analyse par item nous a apporté certains éléments de réponse. Cette analyse a montré une différence entre les groupes d'encodage, les sujets placés dans le groupe d'encodage multi-sensoriel étant globalement plus rapides pour catégoriser les mots dans la seconde phase que les sujets placés dans les groupes d'encodage unisensoriel. Cette différence se manifeste quelle que soit les conditions expérimentales dans lesquelles sont placés les mots. Il est intéressant de noter qu'après avoir réalisé un encodage multimodal, les sujets sont plus rapides pour catégoriser un matériel différent du format d'apprentissage, présenté dans une seule modalité. On peut penser que le contexte multi-sensoriel a favorisé l'amélioration globale des performances des sujets. On note également que l'analyse par item révèle un effet tendanciel du facteur ITEM. On constate (voir FIG. 33) que les items ayant été présentés en condition « concepts avec interférence » sont catégorisés plus rapidement que les items placés dans les conditions « concepts nouveaux », et même s'il n'y a pas de différences significatives entre les conditions « concepts sans interférence » et « concepts avec

interférence », on peut penser que l'interférence n'a pas joué le rôle perturbateur attendu (Vallet *et al.*, 2010 ; Brunel *et al.*, 2010 ; Lehman et Murray, 2005). Bien qu'il faille rester prudent quand aux interprétations, l'effet obtenu par item n'étant que tendanciel, on ne peut pas statuer sur la nature des processus mis en jeu par l'interférence, étant donné les résultats de cette première expérience.

Un des problèmes majeurs de cette expérience est de ne pas avoir obtenu d'effet d'amorçage, de la phase d'encodage vers la phase de test. Bien que les effets multimodaux que nous recherchions ne soient pas présents dans cette première expérience nous pourrions proposer plusieurs explications. Premièrement, l'encodage incident réalisé ici ne permet pas l'activation de traces mnésiques permettant une facilitation de traitement ultérieur. Deuxièmement, l'encodage réalisé ici ne peut pas amorcer des mots, car l'étiquette verbale n'a pu être activée lors de la phase de test par des dimensions sensori-motrices. Enfin l'interférence utilisée dans cette expérience a pu perturber l'encodage des items qui n'étaient pas placés dans cette condition interférence. Pour toutes ces raisons, il nous semble raisonnable de réaliser une nouvelle expérience en modifiant les conditions expérimentales. Nous allons supprimer la condition interférence afin de simplifier le plan expérimental et voir si la simple exposition sensorielle dans la phase d'encodage a une influence sur la catégorisation des mots dans la seconde phase. Nous testerons de nouveau la condition interférence dans une expérience ultérieure.

Expérience 2

Hypothèses

H1 : Un encodage multimodal favorise le traitement ultérieur de l'étiquette verbale du concept en comparaison à un encodage unimodal ou interférent.

Méthode

Sujets

Trente six sujets ($M = 21.6$; $ET = 2.5$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

Stimuli visuels

Les images

Nous avons sélectionné 40 images à partir des images de l'expérience précédente, 20 représentant des concepts vivants et 20 des concepts non vivants. Cette sélection plus sévère des items avait pour but de ne conserver que les couples images/sons pour lesquels le pré-test avait donné les valeurs les plus fortes lorsque les sujets devaient dire si les images et les sons présentés allaient bien ensemble. Les images avaient les mêmes caractéristiques en termes de dimensions ou de durée d'affichage que lors de la première expérience.

Les mots

Seuls les noms des concepts représentés par les couples images/sons conservées pour cette expérience ont été retenus. Ils possédaient les mêmes caractéristiques que lors de la première expérience.

Stimuli auditifs

La sélection des 40 sons s'est faite en parallèle de celle des images, les critères de sélection sont expliqués ci-dessus. Les caractéristiques des sons sont en tout point similaire à celles de l'expérience 1.

Procédure et plan expérimental

L'expérience était composée de deux phases distinctes. Lors d'une première phase d'encodage incident, les sujets devaient catégoriser chacun des stimuli qui leur étaient présentés en fonction de sa catégorie d'appartenance (vivant/non vivant) en répondant sur deux touches distinctes « a » et « p ». Dans une condition expérimentale, un premier groupe de sujets (Visuel) voyait les concepts deux fois sous forme visuelle uniquement, dans l'autre condition expérimentale, un autre groupe de sujets (Audiovisuel) voyait les concepts également deux fois, une première fois sous forme auditive et une deuxième fois sous forme visuelle. Dans chacun des groupes expérimentaux, les listes étaient construites de telle manière qu'il y ait en moyenne 9 items entre la présentation des deux stimuli représentant le même concept. Les sujets devaient catégoriser 40 stimuli lors de l'encodage incident, avec deux stimuli par concept, ce qui représentait un total de 20 concepts. Dans la deuxième phase, les sujets voyaient des noms représentant les concepts qu'ils devaient une nouvelle fois catégoriser (vivant/non vivant) à l'aide des mêmes touches « a » et « p ». Les sujets voyaient 20 noms représentant des concepts qu'ils avaient vus lors de la première phase, et 20 noms

représentaient des concepts nouveaux. Dans cette expérience, nous avons manipulé deux facteurs intra-sujets: le facteur CATEGORIE à deux modalités (vivant/non vivant) et le facteur ITEM à deux modalités (nom ancien/nom nouveau). Nous avons également manipulé un facteur inter-sujet : le facteur MODALITE à deux modalités (Audiovisuel/Visuel). Nous avons créé deux listes dans lesquelles nous avons contrebalancé la nature ancien/nouveau des items sur l'ensemble des sujets. Les touches de réponses étaient également contrebalancées sur l'ensemble des sujets.

Dans la phase d'encodage, les items étaient présentés dans un paradigme de catégorisation continue, constitué d'un écran blanc d'une durée de 500 ms puis d'un point de fixation présenté au centre de l'écran qui durait 500 ms. Ensuite le stimulus visuel ou auditif était présenté pendant une durée de 1000 ms. Il était demandé aux sujets de répondre le plus rapidement possible, mais quelle que soit la réponse du sujet, la durée de présentation de l'item était constante pour favoriser l'encodage des concepts lors de cette phase. Si le sujet ne donnait pas de réponse dans un délai de 2 secondes après la disparition du stimulus, on passait à l'essai suivant. Il n'y avait pas de délai entre la phase d'encodage incident et la phase de test, un écran marquait la fin de la première phase et invitait le sujet à lire les consignes de la deuxième partie de l'expérience. Dans la phase de test, un essai était constitué comme lors de la phase précédente, d'un écran blanc de 500 ms, puis d'un point de fixation de 500 ms, puis du mot présenté au centre de l'écran. Les mots restaient affichés jusqu'à ce que le sujet réponde. Les sujets devaient catégoriser les mots à l'aide des mêmes touches qu'ils avaient utilisé lors de la première phase. Les touches de réponse ont été contrebalancées. Il était indiqué aux sujets de répondre le plus rapidement possible en faisant le moins d'erreurs.

Résultats

Les latences supérieures à 2000 ms n'ont pas été prises en compte, seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse. Nous avons conduit une analyse de variance sur les pourcentages de réponses correctes et une autre sur les latences de réponses (voir TAB. 17). Une analyse par item a été réalisée sur les latences de réponses.

	Temps de réponse (ms)				Pourcentage de réponses correctes (%)			
	Ancien		Nouveau		Ancien		Nouveau	
	vivant	non vivant	vivant	non vivant	vivant	non vivant	vivant	non vivant
VV	575	622	592	637	0,95	0,94	0,97	0,93
AV	18	17	18	16	0,01	0,02	0,01	0,02
	550	603	564	605	0,95	0,97	0,96	0,93

	17	16	18	16	0,01	0,02	0,02	0,02
--	----	----	----	----	------	------	------	------

TAB. 17 – Moyennes des temps de réponses et des pourcentages de réponses correctes et erreurs types en fonction des facteurs ITEM, CATEGORIE et MODALITE

L'analyse de variance conduite sur les pourcentages de réponses correctes n'a révélé aucune différence significative. Concernant les latences de réponse, l'analyse de variance montre un effet du facteur ITEM, par sujet $F_s(1,34) = 8.4; p < .01; \eta^2 = .20$ et par item $F_i(1,38) = 9.1; p < .01; \eta^2 = .28$, en effet les mots faisant référence aux concepts étudiés dans la première phase (587 ms) sont traités plus rapidement que les mots faisant référence à des concepts nouveaux (599 ms). L'analyse de variance révèle également un effet du facteur MODALITE par sujet $F_s(1,34) = 4,8; p < .05; \eta^2 = .12$ et par item $F_i(1,38) = 47.52; p < .01; \eta^2 = .56$, qui suggère une rapidité de traitement significativement plus importante chez les sujets placés dans le groupe Audiovisuel par rapport aux sujets placés dans le groupe Visuel. Il existe également un effet du facteur CATEGORIE par sujet $F_s(1,34) = 46.3; p < .01; \eta^2 = .58$ et par item $F_i(1,38) = 15.06; p < .01; \eta^2 = .28$, les mots se rapportant aux concepts vivants (570 ms) sont traités plus rapidement que ceux se rapportant aux concepts non vivants (617 ms). Contrairement à notre hypothèse initiale, l'interaction entre les facteurs MODALITE*ITEM (voir FIG. 34) n'est significative ni par sujet $F_s(1,34) = 1.4; p = .25; \eta^2 = .04$, ni par item $F_i(1,38) < 1$. On n'observe ni d'interaction CATEGORIE*MODALITE, $F_s(1,34) < 1$, ni de double interaction ITEM*CATEGORIE, $F_s(1,34) < 1$, ni de double interaction ITEM*CATEGORIE*MODALITE, $F_s(1,34) < 1$.

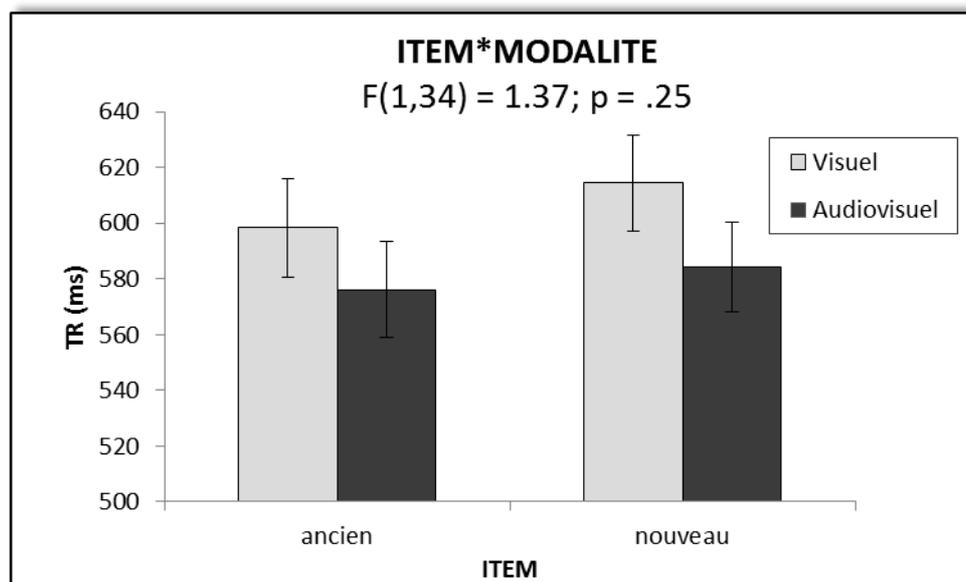


FIG. 34 – Représentation graphique de l'interaction entre les facteurs ITEM et MODALITE sur les temps de réponse (ms)

Discussion

Dans cette expérience, nous cherchions à montrer qu'un encodage multi-sensoriel était meilleur qu'un encodage unisensoriel pour la catégorisation ultérieure de mots faisant référence aux concepts étudiés dans la phase précédente. En effet, le nom d'un concept faisant partie de la trace mnésique, nous pensions que l'activation des dimensions sensorielles faciliterait l'activation de l'étiquette verbale du concept. Par rapport à la première expérience, nous avons retiré la condition d'interférence, ainsi que le groupe expérimental Auditif. Nous avons émis l'hypothèse que le groupe placé dans la condition Audiovisuelle traiterait plus rapidement les mots appartenant à des concepts étudiés lors de la phase d'encodage par rapport au groupe placé dans la condition Visuelle. Or l'interaction ITEM*MODALITE ne s'est pas révélée significative, il semble donc que le traitement des mots faisant référence aux concepts anciens n'est pas plus efficace dans le groupe Audiovisuel. L'activation de l'étiquette verbale n'a peut être pas été assez forte dans le groupe Audiovisuel par rapport au groupe Visuel. Cette absence de différences est peut être aussi liée à un encodage incident peu efficace. Il se peut que la catégorisation de mots faisant référence à un concept ne soit pas influencée par le nombre de modalités sensorielles mises en jeu. Nous notons la présence d'un effet du facteur ITEM qui montre que globalement les items anciens sont catégorisés plus rapidement que les items nouveaux quel que soit le groupe. Il semble qu'il existe dans cette expérience un effet d'amorçage de la première phase vers la seconde entre des concepts présentés dans un format visuel ou sonore et l'étiquette verbale des concepts, toutefois la nature du format ne semble pas avoir d'influence sur cet amorçage. Dans la première expérience, nous n'avions pas observé d'effets d'amorçage, les résultats obtenus ici semblent indiquer que l'activation de la composante lexicale à partir de manipulations sensorielles est possible, mais le nombre de modalités sensorielles impliquées importe peu. Enfin, la présence d'un effet du facteur MODALITE seul nous apparaît comme surprenant, mais concorde avec nos résultats précédents. Les noms de concepts, qu'ils soient anciens ou nouveaux, sont traités plus rapidement dans le groupe Audiovisuel que dans le groupe Visuel. Il se pourrait donc que la présentation multi-sensorielle entraîne une accélération globale des performances par rapport à la présentation unisensorielle des concepts. On ne peut toutefois pas écarter l'hypothèse du biais d'échantillonnage, les sujets placés dans ce groupe ont peut être été plus rapides que ceux placés dans l'autre groupe, indépendamment des conditions expérimentales.

À la vue de ces résultats, il nous paraît important de réaliser une nouvelle expérience où le facteur MODALITE serait manipulé en intra-sujet. Cela va nous permettre de mieux contrôler les différences liées à la variabilité inter-sujet. Nous nous proposons alors de tester cet effet de

contexte bimodal/unimodal d'activation des concepts en présentant lors de la première phase, des concepts en condition Visuelle ainsi que des concepts en condition Audiovisuelle, et de tester ce type d'encodage sur la même phase de catégorisation de mots que dans cette expérience.

Expérience 3

Hypothèses

H1 : Nous nous attendons à observer de meilleures performances en catégorisation de mots pour les items qui ont été encodés de façon multimodale par rapport à ceux qui ont été étudiés en condition unimodale.

Méthode

Sujets

Trente deux sujets ($M = 20.9$; $ET = 0.7$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques, ni de troubles de lecture, et possédaient une vision normale ou corrigée.

Stimuli

Les images, les sons ainsi que les noms utilisés dans cette expérience sont les mêmes que ceux de la deuxième expérience.

Procédure et plan expérimental

L'expérience se déroulait en deux phases de la même manière que lors de l'expérience 2, avec une phase d'encodage incident puis une phase de test. Les sujets devaient effectuer une tâche de catégorisation (vivant/non vivant) sur des images et des sons d'animaux ou d'objets dans la première phase. Dans la seconde phase, ils effectuaient la même tâche de catégorisation avec des noms de concepts.

Encodage (20 concepts)				
CATEGORIE	Vivants		Non vivants	
	10		10	
MODALITE	AV	VV	AV	VV
	5	5	5	5

TAB. 18 – Répartition des stimuli de la phase d'encodage en fonction des facteurs CATEGORIE et MODALITE

La phase d’encodage incident comportait 40 stimuli au total, comme chaque concept pouvait être présenté deux fois, sous forme auditive puis visuelle ou deux fois sous forme visuelle, cette phase comportait 20 concepts (voir TAB. 18). Lors de la phase de test, les sujets voyaient tous les concepts dont la moitié avait été étudiée lors de l’encodage et l’autre moitié constituée de concepts nouveaux. Nous avons constitué deux listes expérimentales dans laquelle le caractère ancien ou nouveau des concepts était contrebalancé sur l’ensemble des sujets. Nous avons également contrebalancé le format de présentation des items au sein de ces deux listes, de telle sorte qu’un item ancien apparaisse soit sous forme audio-visuelle soit visuelle-visuelle, ce qui nous a conduits à créer deux nouvelles sous listes expérimentales pour chacune des listes, soit un total de quatre listes expérimentales. Les listes expérimentales étaient construites de telle sorte que deux présentations d’un même concept soient espacées d’en moyenne 9 items. Dans la phase de test, les sujets voyaient l’ensemble des noms des concepts, ceux qu’ils n’avaient pas vus lors de l’encodage, au nombre de 20 et 20 autres qu’ils avaient étudiés précédemment (voir TAB. 19). Dans cette expérience, nous avons manipulé deux facteurs intra-sujets : le facteur CATEGORIE à deux modalités : vivant et non vivant, et le facteur ITEM à trois modalités : mot ancien présenté deux fois visuellement, mot ancien présenté une fois auditivement et une fois visuellement et mot nouveau.

Phase de test (40 concepts)				
CATEGORIE	vivants		non vivants	
	20		20	
ITEM	mots anciens		mots nouveaux	
	10	10	10	10
	VV	AV	VV	AV
	5	5	5	5

TAB. 19 – Répartition des stimuli lors de la phase de récupération en fonction des facteurs CATEGORIE et ITEM

La procédure utilisée est la même que celle de l’expérience 1 (voir page 136), nous ne la détaillerons donc pas de nouveau. Les sujets devaient répondre à l’aide de deux touches « a » et « p », une pour indiquer vivant et l’autre pour indiquer non vivant, les touches de réponse ont été contrebalancées sur l’ensemble des sujets.

Résultats

Les latences supérieures à 2000 ms n’ont pas été prises en compte, seules les latences correspondant aux bonnes réponses ont été conservées pour l’analyse. Trois sujets ont été exclus de l’analyse, deux sujets pour ne pas avoir respecté les consignes et un autre sujet pour ses latences de réponse très lentes. Nous avons conduit deux analyses de variance, une sur les

latences de réponse (voir TAB. 20) et l'autre sur les pourcentages de réponses correctes (voir TAB. 21).

Temps de réponse (ms)					
Vivant			Non vivant		
Ancien AV	Ancien VV	Nouveau	Ancien AV	Ancien VV	Nouveau
630	626	624	658	678	669
20	16	17	20	20	15

TAB. 20 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et ITEM

Pourcentage de réponses correctes (%)					
Vivant			Non vivant		
Ancien AV	Ancien VV	Nouveau	Ancien AV	Ancien VV	Nouveau
0,99	0,99	0,93	0,99	0,98	0,95
0,01	0,01	0,01	0,01	0,02	0,01

TAB. 21 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et ITEM

L'analyse de variance réalisée sur les latences de réponse a mis en évidence un effet du facteur CATEGORIE, $F(1,28) = 31.9; p < .01; \eta^2 = .53$, qui indique que les mots de concepts vivants (627 ms) sont catégorisés plus rapidement que les mots de concepts non vivants (668 ms). L'analyse ne révèle pas d'effet du facteur ITEM, $F(2,6) < 1$, il n'existe pas d'interaction ITEM*CATEGORIE, $F(2,56) = 1.2; p = .30$.

L'analyse de variance conduite sur les pourcentages de réponses correctes ne montre pas d'effet du facteur CATEGORIE, $F(1,28) < 1$, par contre on observe un effet du facteur ITEM (voir FIG. 35), $F(2,56) = 11.02; p < .01; \eta^2 = .28$. Les sujets ont commis moins d'erreur pour catégoriser les mots dont les concepts avaient été présentés de manière audiovisuelle (AV) (99%) et les mots dont les concepts avaient été présentés de manière visuelle (VV) (98%) par rapport aux mots nouveaux (94%), $F(1,28) = 20.48; p < .01$ et $F(1,28) = 12.06; p < .01$ respectivement. Toutefois, il n'existe pas de différences entre les pourcentages de bonnes réponses des mots anciens AV et des mots anciens VV, $F(1,28) < 1$.

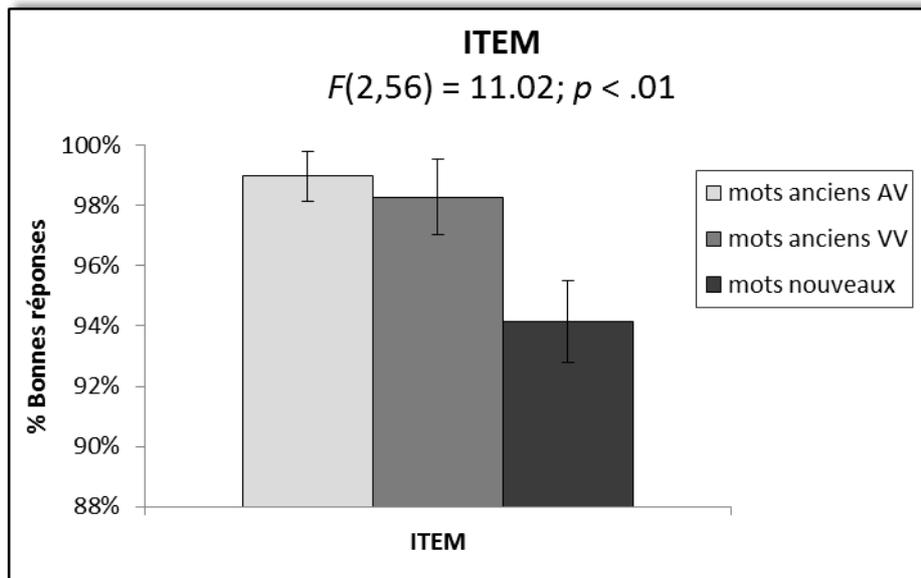


FIG. 35 – Représentation graphique de l'effet du facteur ITEM sur les pourcentages de réponses correctes (%)

Discussion

Dans cette expérience, nous avons cherché à montrer qu'un encodage incident multimodal de concepts vivants ou non vivants était plus efficace pour catégoriser les noms de ces concepts dans une phase ultérieure par rapport à un encodage incident unimodal. Nous n'avons pas réussi à montrer cela lors de la deuxième expérience et nous avons décidé de ne pas garder l'encodage incident multimodal ou unimodal en inter-sujet mais plutôt en intra-sujet afin de mieux contrôler les effets liés aux différences d'échantillonnage des groupes.

La première chose que l'on a pu observer dans cette expérience est l'absence d'effet sur les latences de réponses. Cela dénote avec la première expérience où les principaux effets que nous avons trouvés se situaient sur les latences de réponse. Ici l'absence d'effet du facteur ITEM montre clairement qu'il n'y a aucune différence significative entre les temps de catégorisation, quelle que soit les conditions expérimentales. Nous ne relevons donc pas d'avantage de la condition multimodale sur les autres conditions concernant les latences de réponses, contrairement à nos hypothèses initiales. Il est toutefois intéressant de noter que l'effet du facteur ITEM se manifeste sur les pourcentages de bonnes réponses, on peut voir que les items anciens sont catégorisés avec moins d'erreurs que les items nouveaux. On peut donc affirmer que les meilleures performances en catégorisation pour les items anciens peuvent être attribuées à l'encodage incident. Toutefois, même si cet encodage semble avoir eu une influence sur les performances en catégorisation dans la phase de test, il ne permet pas de différencier la condition multimodale de la condition unimodale. On peut essayer

d'apporter plusieurs explications pour tenter de comprendre cette absence d'effet. Premièrement, le peu d'items par condition expérimentale et par conséquent le calcul de moyennes sur trop peu de temps de réponse a pu empêcher un effet d'être mis en évidence. Deuxièmement, un biais relatifs aux noms des concepts pourrait aussi expliquer l'absence d'effets. Les noms des concepts ont été choisis en fonction des couples images/sons de la base de données de notre équipe et non l'inverse, or on sait que pour la lecture de mots, de nombreuses variables influent sur le temps de lecture, comme le nombre de syllabes, la fréquence des mots dans la langue, ainsi que leur concrétude ou leur valence par exemple. La difficulté de trouver des paires images/sons correspondant à une catégorisation vivant/non vivant a induit un moins bon contrôle du matériel verbal utilisé pour cette expérience.

C'est pour cette raison que nous allons conduire une nouvelle expérience en remplaçant les noms des concepts de la deuxième phase par des images. Afin de mieux contrôler le matériel expérimental, nous allons utiliser des images du même concept mais qui ne seront pas les mêmes (différents exemplaires) que celles présentées dans la phase d'encodage. En effet, représenter les mêmes images dans la phase de test pourrait avantager les items de la condition unimodale qui verraient trois fois la même image contre une fois pour les items présents dans la condition bimodale. De plus, ce procédé s'apparenterait à de l'amorçage de répétition, ce que nous ne souhaitons pas. Nous avons donc choisi de présenter une image différente dans la deuxième phase, qui apparenterait notre tâche à une tâche d'amorçage conceptuelle cross-modale à long terme.

Expérience 4

Hypothèses

H1 : Un encodage multimodal devrait favoriser le traitement ultérieur d'une image distincte de celle présentée lors de l'encodage mais représentant le même concept, en comparaison à un encodage unimodal.

Méthode

Sujets

Trente sujets ($M = 20.4$; $ET = 4.8$), étudiants à l'Université Lyon II, ont participé à cette expérience. Tous les sujets étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient ni de troubles neurologiques ni de troubles sensoriels.

Stimuli

Nous avons ajouté 8 couples images/sons par rapport à l'expérience précédente, ce qui a conduit à un total de 48 concepts, vingt quatre concepts vivants et vingt quatre concepts non vivants. Quarante couples images/sons ont été repris de notre base de données et les 8 nouveaux couples ont été sélectionnés via les moteurs de recherche internet « Picsearch » pour les images et « FindSounds » pour les sons. Les nouvelles images ont été standardisées en terme de dimension à l'aide du logiciel « PicNik » et les sons ont été standardisés en terme de durée à l'aide du logiciel AUDACITY.

Nous avons également ajouté 48 nouvelles images des concepts que nous disposions déjà. Ces images étaient destinées à être présentées uniquement lors de la seconde phase. Vingt quatre de ces images représentaient des concepts non vivants et vingt quatre autres des concepts vivants. Ces images ont toutes été choisies via des moteurs de recherche d'image sur Internet, et elles ont été sélectionnées pour se différencier perceptivement suffisamment des images utilisées pour la phase d'encodage, afin d'éviter les effets d'amorçage de répétition. Nous disposions de triplets constitués de deux exemplaires différents du même concept sous forme imagée et d'un son produit par le concept dans son environnement naturel ou dans des conditions normales d'utilisation.

Procédure et plan expérimental

L'expérience était composée de deux phases distinctes comme précédemment. Lors d'une première phase d'encodage incident, les sujets devaient catégoriser chacun des stimuli qui leur étaient présentés en fonction de sa catégorie d'appartenance (vivant/non vivant) en répondant sur deux touches distinctes « a » et « p ». Les sujets voyaient les concepts deux fois, soit dans le même format (visuel), soit dans deux formats (auditif puis visuel). Les listes d'items étaient construites de telle manière qu'il y ait en moyenne 9 items entre la présentation des deux stimuli représentant le même concept.

Le plan expérimental est un carré latin constitué de 3 listes expérimentales construites chacune de telle sorte qu'elles comportent toutes le même nombre d'oiseaux, de mammifères, d'instruments de musique, d'outils. Chaque liste comportait 16 items, 8 vivants et 8 non vivants. Dans un groupe expérimental, tous les items d'une liste étaient présentés lors de l'apprentissage dans la condition visuelle-visuelle (VV) ou dans la condition audio-visuelle (AV) ou alors les items n'étaient pas présentés car ils faisaient partis des items nouveaux

(Nvx). Nous avons utilisé trois groupes expérimentaux où chaque liste était associée à une condition différente (voir TAB. 22).

	Liste 1	Liste 2	Liste 3
Groupe 1	AV	VV	Nvx
Groupe 2	VV	Nvx	AV
Groupe 3	Nvx	AV	VV

TAB. 22 – Répartition des stimuli dans les listes expérimentales en fonction des groups de sujet

Par exemple, les sujets placés dans le groupe 1 voyaient les items de la liste 1 dans la condition audiovisuelle et ceux de la liste 2 dans la condition visuelle-visuelle. Les items de la liste 3 n'étaient présentés que dans la phase test, ils correspondaient aux items non étudiés. La phase d'encodage incident comportait 64 stimuli au total, soit 32 concepts répétés deux fois. Dans la deuxième phase, les sujets voyaient des images représentant des concepts qu'ils avaient vus ou non dans la phase d'encodage. Toutefois, aucune des images présentées lors de la phase test n'a réellement été vue lors de la première phase, car toutes les images de la seconde phase représentaient un autre exemplaire du même concept. Dans la phase de test, les items de toutes les listes étaient présentés sous forme imagée seulement. Dans cette expérience, nous avons manipulé deux facteurs intra-sujets, le facteur CATEGORIE à deux modalités : vivant ou non vivant et le facteur ITEM à trois modalités images de concepts présentés en condition AV, images de concepts présentés en condition VV ou images de concepts nouveaux. Chaque condition du facteur ITEM comportait 16 stimuli dans la phase de tests, il y avait donc 32 stimuli anciens et 16 stimuli nouveaux.

Dans la phase d'encodage, les items étaient présentés dans un paradigme de catégorisation continue, constitué d'un écran blanc d'une durée de 500 ms puis d'un point de fixation présenté au centre de l'écran qui durait 500 ms. Ensuite le stimulus visuel ou auditif était présenté pendant une durée de 1000 ms. Il était demandé aux sujets de répondre le plus rapidement possible mais quelle que soit la réponse du sujet, la durée de présentation de l'item était constante pour favoriser l'encodage des concepts lors de cette phase. Si le sujet ne donnait pas de réponse dans un délai de 2 secondes après la disparition du stimulus, on passait à l'essai suivant. Il n'y avait pas de délai entre la phase d'encodage incident et la phase de test, un écran marquait la fin de la première phase et invitait le sujet à lire les consignes de la deuxième partie de l'expérience. Dans la phase de test, un essai était constitué comme lors de la phase précédente, d'un écran blanc de 500 ms, puis d'un point de fixation de 500 ms, puis d'une image présentée au centre de l'écran. Les images restaient affichées jusqu'à ce que le sujet réponde. Les sujets devaient catégoriser les images à l'aide des mêmes touches qu'ils

avaient utilisé lors de la première phase. Il était indiqué aux sujets de répondre le plus rapidement possible en faisant le moins d'erreurs. Les touches de réponses étaient contrebalancées sur l'ensemble des sujets, mais pas entre la première et la deuxième phase de test pour un même sujet.

Résultats

Les latences supérieures à 2000 ms n'ont pas été prises en compte, seules les latences correspondant aux bonnes réponses ont été conservées pour l'analyse. Nous avons conduit deux analyses de variances, l'une sur les pourcentages de bonnes réponses (voir TAB. 24) et une autre sur les latences de réponses (TAB. 23). De plus, nous n'avons réalisé d'analyse par item que sur les latences de réponse.

vivant			non vivant		
ancien VV	ancien AV	nouveau	ancien VV	ancien AV	nouveau
511	509	513	517	501	511
9	9	14	10	9	9

TAB. 23 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et ITEM

vivant			non vivant		
ancien VV	ancien AV	nouveau	ancien VV	ancien AV	nouveau
0,97	0,97	0,97	0,97	1,00	0,98
0,012	0,010	0,010	0,012	0,000	0,008

TAB. 24 - Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et ITEM

L'analyse de variance conduite sur les pourcentages de bonnes réponses révèle un effet du facteur CATEGORIE, $F(1,29) = 4.19$; $p = .049$; $\eta^2 = .13$, les images correspondant aux concepts non vivants (98%) sont mieux catégorisés que les images correspondant aux concepts vivants (97%). L'analyse de variance ne révèle en revanche aucun effet du facteur ITEM, $F(2,58) = 1.24$; $p = .30$, ni d'interaction CATEGORIE*ITEM, $F(2,58) < 1$.

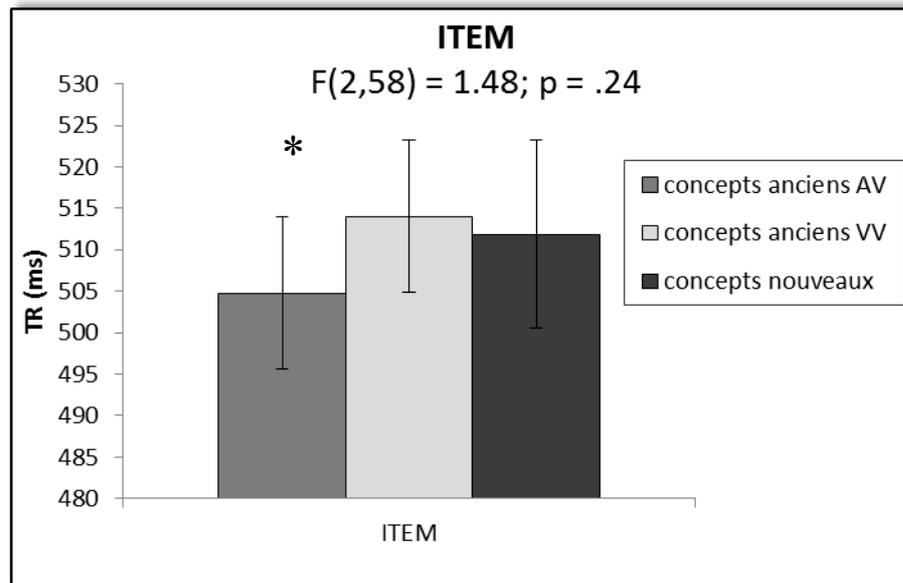


FIG. 36 – Représentation graphique de l'effet du facteur ITEM sur les temps de réponse (ms)

L'analyse de variance conduite sur les latences de réponses ne révèle aucun effet du facteur CATEGORIE, $F_s(1,29) < 1$ et $F_i(1,46) < 1$. Contrairement à nos hypothèses, nous ne notons pas d'effet du facteur ITEM, $F_s(2,58) = 1.48 ; p = .24$ et $F_i(2,92) < 1$, il n'y a donc pas de différence significative entre les images correspondant aux concepts présentés visuellement lors de la première phase (VV) (514ms), les images correspondant aux concepts présentés de manière audiovisuelle (AV) (505ms) et les images nouvelles (512ms). Comme nous avons une hypothèse a priori forte concernant le sens de la différence entre les images anciennes VV et AV, nous avons réalisé une comparaison planifiée bien que l'effet ITEM ne soit pas significatif. Selon cette comparaison, il existe une différence entre les temps de réponse pour les images anciennes AV et les images anciennes VV, $F(1,29) = 4.08 ; p = .026$, il n'existe cependant pas de différences entre les temps de réponse des images anciennes VV et AV et des images nouvelles, $F(1,29) < 1$. Les images anciennes AV semblent donc être catégorisées plus rapidement que les images anciennes VV. L'analyse ne révèle pas d'interaction CATEGORIE*ITEM, $F_s(2,58) < 1$ et $F_i(2,92) < 1$.

Discussion

Dans cette expérience, nous avons modifié la phase de test en changeant les noms de concepts par des images différentes de celles présentées lors de l'encodage, mais représentant les mêmes concepts. Nous avons décidé de mettre des images différentes dans la phase de test pour éviter des effets qu'on aurait pu imputer à de l'amorçage de répétition pour la condition unimodale, dont on sait que les effets sont plus forts que ceux produits par de l'amorçage

conceptuel. Cette manipulation avait pour objectif de montrer que l'encodage incident multimodal était meilleur que l'encodage incident unimodal. Les résultats obtenus n'ont guère été concluants, en effet nous n'avons observé aucun effet, tant sur les pourcentages de bonnes réponses que sur les latences de réponses. Ces résultats contrastent avec ceux obtenus précédemment, puisqu'il existait un effet d'amorçage sur les latences de réponse dans l'expérience 2 et sur les pourcentages de bonnes réponses dans l'expérience 3. Il semble que le changement de format dans la phase de test ait gêné la présence d'un effet d'amorçage, car la condition « images de concepts nouveaux » est traitée aussi rapidement que les autres. Comme nous avons une hypothèse forte, nous avons tout de même analysé la différence entre les latences de réponse des images dont le concept avait été présenté dans le format audiovisuel lors de la première phase et celles dont le concept avait été présenté dans le format visuel-visuel lors de la première phase. Il en est ressorti une différence significative entre ces deux conditions, que nous pourrions interpréter comme une implication des processus d'activation/intégration. Toutefois, l'absence d'effet d'amorçage significatif ne nous permet pas d'interpréter plus ces résultats. Il se peut également que l'encodage incident que nous utilisons ne soit pas adapté pour avoir un impact sur la phase de test. En utilisant un paradigme légèrement différent, Vallet et al. (2010) a mis en évidence un effet d'amorçage inter-sensoriel de la modalité auditive vers la modalité visuelle d'un même concept, et réciproquement. Les auteurs interprètent cet effet comme une facilitation de la modalité amorcée causée par l'activation du concept. Cet effet que nous cherchons à mettre en évidence ne semble pas se manifester dans le paradigme que nous avons choisi ici et peine à valider les hypothèses liées aux mécanismes d'activation/intégration.

Lors de la première expérience nous avons utilisé les noms des concepts dans la seconde phase, et nous avons observé un effet d'amorçage global de l'encodage, ainsi qu'un avantage pour le groupe multimodal par rapport au groupe unimodal. Lors de la seconde expérience, nous avons modifié le plan expérimental pour que la modalité d'encodage soit en intra-sujet. Nous avons montré un effet d'amorçage de la première phase vers la seconde sur les pourcentages de bonnes réponses, mais aucune différence entre les conditions multimodale et unimodale. Dans la troisième expérience, nous avons changé les noms de concepts par des images dans la phase de test, en utilisant des images différentes de celles utilisées lors de l'encodage. Les résultats semblent moins clairs même si une différence semble exister entre les conditions unimodale et multimodale. Il paraîtrait donc nécessaire de modifier le paradigme dans une expérience suivante. Nous avons préféré une phase d'encodage où les

items auditifs et visuels n'étaient pas présentés simultanément dans le but de renforcer l'activation de la trace dans le temps et faciliter le traitement ultérieur, mais il semble que cette présentation des modalités ne soit pas suffisante et même inefficace pour produire des effets à long terme. En effet, on peut imaginer que la présentation « différée » des items auditifs et visuels d'un même concept lors de la phase d'encodage ne permette pas réellement de créer des associations qui soient plus fortes que la présentation de deux items visuels ou auditifs du même concept.

Le paradigme que nous avons utilisé jusqu'à présent ne semble pas adapté pour mettre en évidence les effets que nous recherchions. Il nous semble intéressant à présent de réaliser une tâche de reconnaissance qui va nous permettre d'évaluer les processus de recollection et de familiarité, ce qui n'était pas le cas dans notre tâche de catégorisation. Pour cela, nous allons garder le même type de paradigme en deux phases en modifiant certains éléments de l'expérience. Tout d'abord, lors de la phase d'encodage, les modalités visuelles et auditives seront présentées simultanément pour les items audiovisuels. La présentation d'un composant perceptif est censé activer les autres dimensions associées au sein de la trace mnésique (Versace *et al.*, 2002; Versace *et al.*, 2009), si l'on présente simultanément deux composants perceptifs, l'activation au sein de la trace devrait donc être plus robuste que la présentation d'un seul composant perceptif. La présentation ultérieure d'un seul des composants devraient donc activer l'autre dimension et faciliter son traitement (Brunel *et al.*, 2009). Nous allons inclure également une condition d'interférence (item visuel + bruit blanc) en plus de la condition unimodale et audiovisuelle. Enfin lors de la phase de récupération, tous les items seront présentés visuellement et correspondront à l'image vue lors de l'encodage. De plus, nous allons reprendre notre matériel, en effet certaines images étaient en contexte alors que d'autres images ne l'étaient pas. Afin d'éviter la présence d'effet parasite lié à la nature du matériel expérimental, nous avons repris l'ensemble du matériel, remplacé des images et des sons dans le but que tous les stimuli soient harmonisés.

Expérience 5

Hypothèses

H1 : Les items placés dans la condition multimodale devraient être mieux reconnus que les items placés dans les autres conditions.

H2 : Les items placés dans la condition unimodale devraient être mieux reconnus que les items placés dans la condition interférence uniquement dans la composante implicite de la tâche de reconnaissance (processus de familiarité).

H3 : Le degré de certitude devrait être plus élevé dans la condition multimodale que dans les autres conditions mais aussi plus important dans la condition unimodale que dans la condition interférence.

Méthode

Sujets

Vingt-quatre sujets ($M = 24$; $SD = 8$), tous étudiants à l'Université Lyon II, ont participé à cette expérience. Tous étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient pas de troubles neurologiques et possédaient une vision normale ou corrigée.

Stimuli

Stimuli visuels

Dans cette expérience, nous avons augmenté le nombre d'items, nous avons ajouté 36 concepts (18 vivants et 18 non vivants). Augmenter le nombre de stimuli était réalisé dans le souci d'ajouter plus d'items par condition expérimentale. Des images trouvées sur le site « Fotolia » ont été ajoutées à notre base de données. Elles ont été standardisées comme les images présentes dans les expériences précédentes (voir Annexes, page 233).

Stimuli auditifs

De la même manière, nous avons ajouté autant de sons que d'images soit 36 sons de concepts. Ces sons ont été pour la plupart trouvés à l'aide du site « Findsounds ». Quand ils ne correspondaient pas en termes de durée aux autres sons de la base de données, ils étaient modifiés à l'aide du logiciel AUDACITY.

Masque auditif

Le masque sonore était un bruit blanc d'une durée de 1250ms, généré par le logiciel AUDACITY. Il était présenté en même temps que les items visuels, et sa durée excédait de 250 ms la présentation des images afin de produire une interférence sonore. L'interférence utilisée dans cette expérience était plus longue de 250 ms par rapport à l'interférence utilisée

lors de la première expérience, notre intention était de perturber l'activation des modalités sensorielles liées au concept en mémoire (Vallet *et al.*, 2010).

Procédure et plan expérimental

L'expérience était constituée de deux phases distinctes : une phase d'encodage incident, où les sujets devaient catégoriser (vivant/non vivant) des images, et une phase de test, dans laquelle les sujets devaient reconnaître parmi d'autres les images qui leur avaient été présentées lors de la première phase. Les réponses s'effectuaient à l'aide des touches « a » et « p » du clavier dans les deux phases de l'expérience. Ces deux phases étaient séparées par un délai de 5 mn, temps durant lequel les sujets réalisaient une grille de Sudoku (voir FIG. 37).

Dans la phase d'encodage incident, les sujets pouvaient voir 42 images (24 vivantes et 24 non vivantes) issues de trois conditions expérimentales : une condition unimodale, où l'image était présentée seule, une condition multimodale, où l'image était accompagnée du son produit par le concept en situation écologique, et une condition interférente, dans laquelle l'image était accompagnée d'un bruit blanc. Les sujets avaient pour consigne de ne traiter que les images, il n'était rien précisé concernant les sons. Lors de la phase de test, les sujets voyaient l'ensemble des quatre vingt quatre images : les quarante-deux qu'ils avaient catégorisées dans la première phase et quarante-deux nouvelles, et avaient pour consigne de reconnaître les images précédemment vues. Nous avons donc manipulé dans cette expérience deux facteurs intra-sujets : le facteur CATÉGORIE à deux modalités (vivant/non-vivant) et le facteur MODALITÉ à trois modalités (items ayant été encodés en condition multimodale, unimodale ou interférence). Deux listes ont donc été créées dans lesquelles la moitié des items étaient vus lors de l'encodage et l'autre moitié était présentée seulement pendant la phase de test, ces deux listes ont été contrebalancées sur l'ensemble des sujets. L'attribution des items dans les conditions expérimentales était également contrebalancée sur l'ensemble des sujets au sein de trois sous listes. Chaque sous liste était constituée de 14 items (7 vivants et 7 non vivants), sélectionnées de telle manière que chacune des sous listes soit équilibrée en termes de catégorie (nombre de volatile ou nombre d'instruments de musique) et de la fréquence des items présents à l'intérieur de celle-ci.

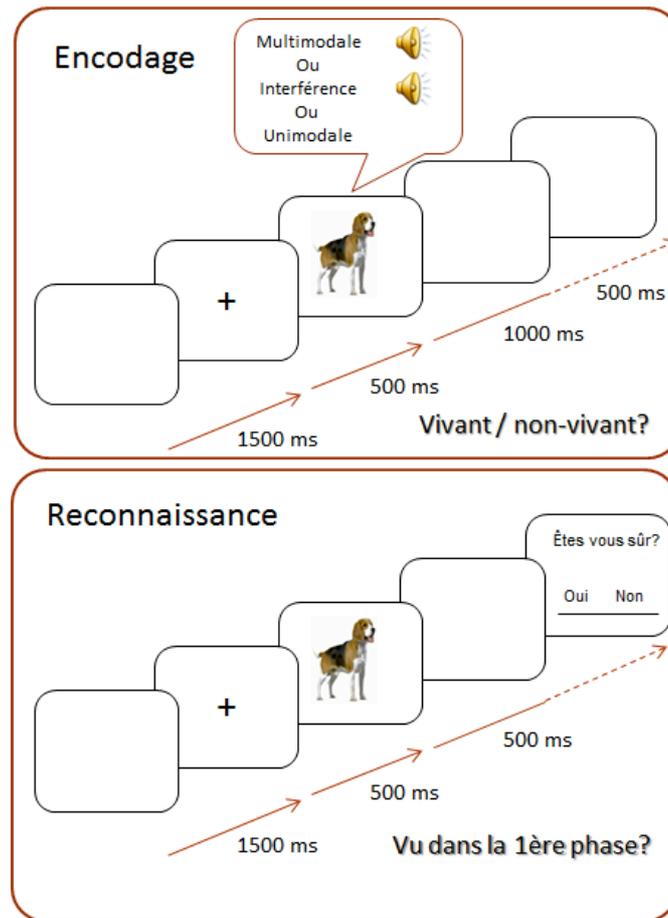


FIG. 37 – Représentation schématique du paradigme expérimental. La phase d’encodage et la phase de récupération étaient séparées par un délai de 5mn pendant lequel les sujets réalisaient un Sudoku.

Chaque essai débutait par un écran blanc d’une durée de 1500ms suivi d’un point de fixation de 500ms. Les sujets devaient focaliser leur attention sur le point de fixation jusqu’à l’apparition du stimulus. Lors de la phase d’encodage incident, les images restaient affichées pendant 1000ms quelle que soit l’action des sujets, qui devaient les catégoriser (vivant/non vivant) le plus rapidement possible. Les sujets avaient 1500ms pour répondre dès l’apparition du stimulus. Si le sujet ne répondait pas durant cette période, l’essai suivant débutait. Lors de la phase du test, les images restaient affichées pendant 500ms afin d’accroître la sensibilité de la reconnaissance. Dès que la réponse était donnée, le sujet était invité à évaluer le degré de certitude de sa réponse sur une échelle continue de type Lickert qui s’affichait à l’écran. Le sujet pouvait cliquer avec la souris sur un continuum allant de: « je ne suis pas du tout sûr de ma réponse » à « je suis sûr de ma réponse ». Les touches de réponse des sujets et les temps de réponse ont été enregistrés lors de cette expérience. Nous avons également enregistré la position de la souris sur l’axe horizontal, lors de la phase de reconnaissance, lorsque les sujets devaient restituer leur degré de certitude.

Résultats

Phase d'encodage

Lors de la phase d'encodage, nous avons filtré les temps de réponse situés au-delà de 2000ms, considérant ces latences comme trop lentes pour une catégorisation vivant/non-vivant, nous avons ensuite calculé les temps de réponse (voir TAB. 25) et les pourcentages de réponses correctes (voir TAB. 26) par condition et par sujets.

vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence
572	607	578	616	630	582
23	25	23	22	24	22

TAB. 25 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et MODALITE

vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence
0,97	0,97	0,93	0,96	0,95	0,96
0,01	0,01	0,02	0,02	0,02	0,02

TAB. 26 - Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE

L'analyse de variances à mesures répétées conduite sur les temps de réponse lors de la phase d'encodage révèle un effet significatif du facteur CATÉGORIE, $F(1,23) = 5.51$; $p = .028$; $\eta_p^2 = .19$, en effet les items vivants sont catégorisés plus rapidement que les items non vivants. L'analyse de variance met également en évidence un effet significatif du facteur MODALITÉ, $F(2,46) = 13.86$; $p < .01$; $\eta_p^2 = .38$, les sujets répondent plus rapidement lorsque les items sont présentés avec un son naturel ($M = 594\text{ms}$) ou avec un bruit blanc ($M = 580\text{ms}$) que lorsqu'ils sont présentés seulement visuellement ($M = 619\text{ms}$), $F(1,23) = 4.47$; $p = .045$ et $F(1,23) = 24.91$; $p < .01$ respectivement. Nous ne relevons pas d'interaction MODALITÉ*CATÉGORIE.

Concernant les pourcentages de bonnes réponses, l'analyse de variance ne révèle aucun effet des facteurs MODALITÉ, CATÉGORIE, ni d'interaction entre ces facteurs. Les performances sont assez hautes quelle que soit la condition de présentation des items ($M = 96\%$).

Phase de test

Concernant la phase de reconnaissance, nous n'avons appliqué aucun filtre sur les temps de réponse étant donné que les sujets avaient pour consigne de privilégier la qualité plutôt que la rapidité de leurs réponses. Nous avons ensuite considéré les pourcentages de réponses correctes par condition et par sujets (voir TAB. 27). Pour les degrés de certitude, les positions de la souris sur l'axe horizontal ont été normalisées en utilisant les déciles de la distribution de chaque sujet. Aucun filtre n'a été appliqué, la moyenne des degrés de certitude a ensuite été calculée par condition et par sujet.

Pourcentage de reconnaissance correcte (%)

vivant				non vivant			
ancien			nouveau	ancien			nouveau
multimodal	unimodal	interférence		multimodal	unimodal	interférence	
0,89	0,85	0,83	0,85	0,85	0,77	0,75	0,90
0,03	0,03	0,04	0,02	0,03	0,05	0,04	0,02

TAB. 27 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE

Nous avons réalisé deux analyses de variance à mesures répétées, une sur les pourcentages de bonnes réponses et une sur les degrés de certitude dans la phase de test. Un sujet a été écarté de l'analyse pour ne pas avoir respecté les consignes.

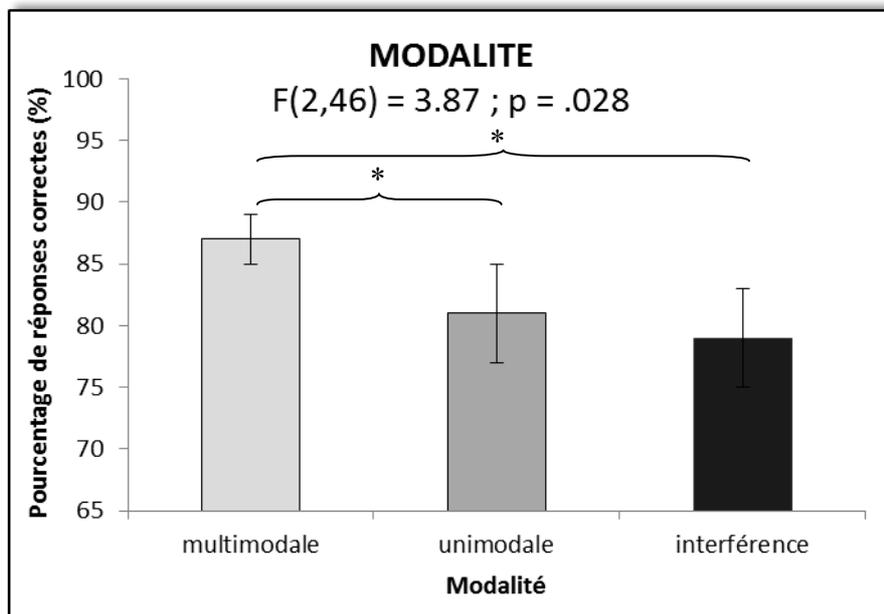


FIG. 38 – Représentation graphique de l'effet du facteur MODALITE sur les pourcentages de réponses correctes (%)

L'analyse de variance conduite sur les pourcentages de bonnes réponses, uniquement pour les items reconnus comme anciens, révèle la présence d'un effet du facteur CATÉGORIE, $F(1,23) = 9.52 ; p < .01 ; \eta_p^2 = .29$. Les items vivants sont mieux reconnus ($M = 86\%$), quelle

que soit la condition expérimentale, que les items non vivants ($M = 79\%$). Nous ne notons aucune interaction MODALITÉ*CATÉGORIE, $F < 1$. L'analyse met en évidence un effet du facteur MODALITÉ, $F(2,46) = 3.87$; $p = .028$; $\eta_p^2 = .14$ (voir FIG. 38). Les items placés dans la condition multimodale ($M = 87$) sont significativement mieux reconnus que les items placés dans la condition unimodale ($M = 81$), $F(1,23) = 3.61$; $p = .035$, ou interférente ($M = 79\%$), $F(1,23) = 6.36$; $p < .01$. On ne note pas de différence significative entre la condition unimodale et interférente, $F(1,23) < 1$.

Afin de voir si la nature de l'encodage avait influencé la familiarité des items lors de la reconnaissance, nous avons réalisé une analyse de détection du signal sur les pourcentages de bonnes réponses. Nous avons calculé les paramètres d' et β pour chacune des trois conditions : multimodales, unimodales et interférentes et nous avons réalisé deux analyses de variance séparées en prenant ces paramètres comme variables dépendantes. L'analyse de variance conduite sur les d' met en évidence un effet du facteur MODALITÉ, $F(2,46) = 3.48$; $p = .039$; $\eta_p^2 = .13$ (voir FIG. 39). Il semble que les items anciens encodés dans la condition multimodale ($M = 2.32$) soient mieux discriminés des items nouveaux que ceux encodés dans la condition unimodale ($M = 2.13$), $F(1,23) = 3.39$; $p = .039$, ou interférence ($M = 2.07$), $F(1,23) = 4.97$; $p = .018$. L'analyse de variance conduite sur les β n'a pas mis en évidence d'effet du facteur MODALITÉ, $F(2,46) = 2.53$; $p = .09$.

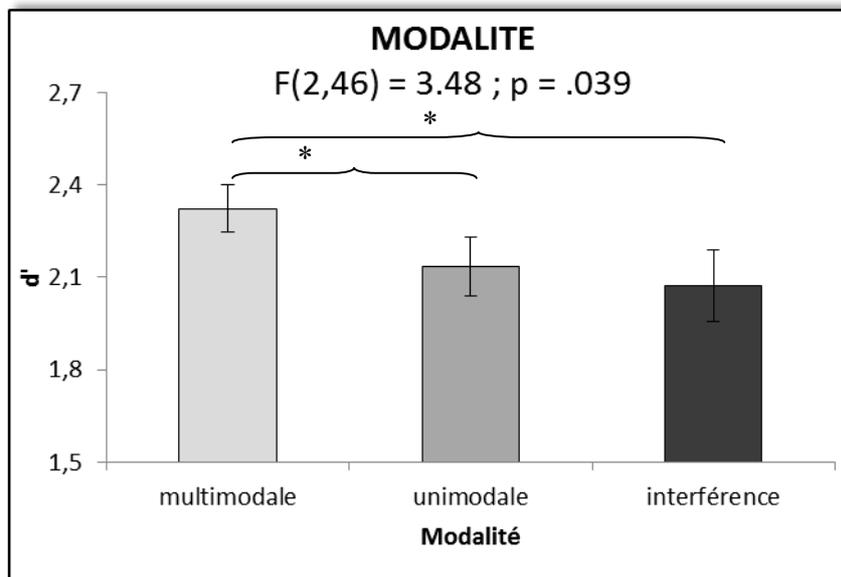


FIG. 39 - Représentation graphique de l'effet du facteur MODALITE sur le d'

Ces résultats semblent indiquer que la manipulation expérimentale a influencé la familiarité des images plutôt que le caractère décisionnel. La familiarité reflétant un processus

automatique, nous avons conduit une nouvelle analyse dans laquelle nous avons considéré séparément les pourcentages de réponses correctes, en fonction des temps de réponse, rapides ou lents, de chaque sujet (Péris & Tiberghien, 1984).

Cette procédure va nous permettre de tester l’hypothèse selon laquelle la condition unimodale devrait être meilleure que la condition interférence dans une récupération de type implicite. En effet, il est établi qu’une tâche de reconnaissance fait intervenir des processus explicites mais également implicites (Yonelinas, 1994). Nous avons donc cherché à voir si la différence entre ces deux conditions pouvait être tributaire de processus distincts dans une tâche explicite. Pour cela nous avons considéré la médiane des temps de réponse par condition expérimentale (multimodale, unimodale et interférence) pour chacun des sujets, et nous avons calculé ensuite les pourcentages de réponses correctes pour chacune des conditions du facteur MODALITÉ, pour les temps de réponse rapides et lents (Péris & Tiberghien, 1984). Ce type d’analyse permet de dissocier un traitement automatique, basé sur la familiarité (réponses rapides), d’un traitement plus contrôlé basé sur la recollection (réponses lentes).

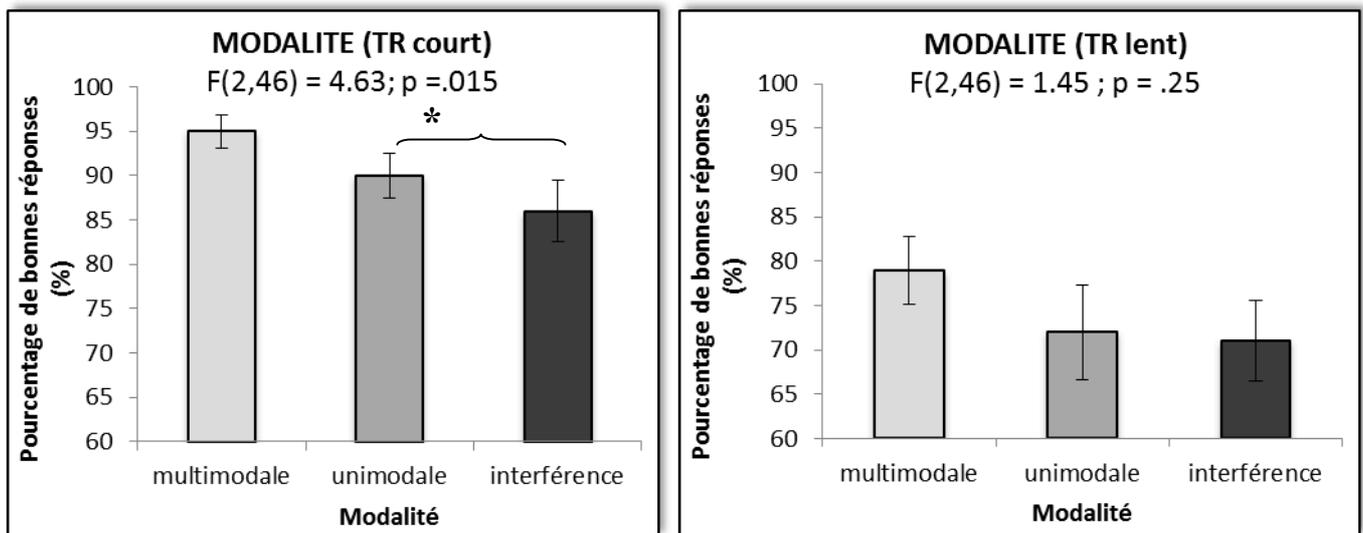


FIG. 40 – Représentation graphique de l’effet du facteur MODALITÉ en considérant séparément les pourcentages de réponses correctes pour les temps de réponses rapides et pour les temps de réponses lents

L’analyse de variance conduite sur les pourcentages de réponses correctes pour les temps de réponses rapides révèle un effet principal du facteur MODALITÉ, $F(2,46) = 4.63; p = .015; \eta_p^2 = .17$, alors que l’on ne note pas d’effet du facteur MODALITÉ pour les TR lents, $F(2,46) = 1.45; p = .25; \eta_p^2 = .06$ (voir FIG. 40). Si l’on s’intéresse à la différence entre les conditions unimodales et interférentes pour les TR rapides, on s’aperçoit qu’il existe une différence entre la condition unimodale ($M = 90\%$) et interférente ($M = 86\%$), $F(1,22) = 3.62; p = .035$.

Nous avons également conduit une analyse de variance sur le degré de certitude associé aux bonnes réponses (voir TAB. 28). Cette analyse révèle uniquement un effet du facteur MODALITÉ, $F(2,46) = 3.59$; $p = .036$; $\eta_p^2 = .13$ (voir FIG. 41). Les sujets semblent plus sûrs de leurs réponses lorsque les items étaient présentés dans la condition interférente ($M = 5,68$) que lorsqu'ils étaient présentés dans la condition unimodale ($M = 5,10$), $F(1,23) = 7.4$; $p = .01$. On ne relève pas de différence entre les degrés de certitude des conditions interférence et multimodale ni de différences entre les degrés de certitude des conditions unimodale et multimodale.

Degré de certitude					
vivant			non vivant		
multimodale	unimodale	interference	multimodale	unimodale	interference
5,19	4,97	5,80	5,57	5,23	5,56
0,25	0,29	0,20	0,22	0,25	0,22

TAB. 28 – Moyennes des degrés de certitude et erreurs types en fonction des facteurs MODALITE et CATEGORIE

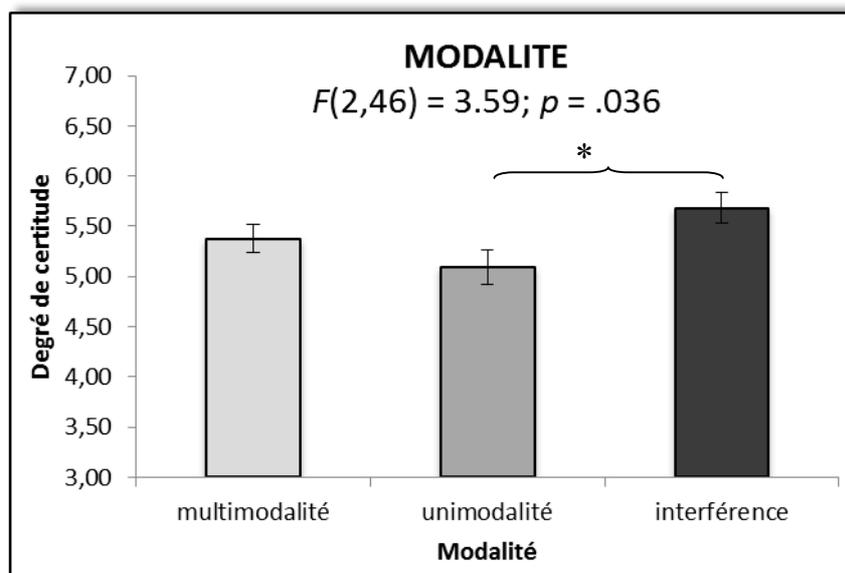


FIG. 41 - Représentation graphique de l'effet du facteur MODALITÉ sur le degré de certitude

De plus, nous relevons une corrélation négative significative entre le pourcentage de réponses correctes et le degré de certitude, $r(22) = -.42$; $p < .05$, moins les items interférés étaient retenus, plus ils étaient associés à un degré de certitude élevé.

Discussion

Les résultats obtenus dans cette expérience confirment notre hypothèse principale, les items vus dans la condition multimodale en phase d'encodage bénéficient de meilleures performances en reconnaissance lors de la phase de test. Avec un paradigme original, nous répliquons des résultats obtenus avec une procédure expérimentale différente (Murray *et al.*,

2004). A notre connaissance, notre étude est la première utilisant ce type de matériel, à inclure dans son plan expérimental des conditions multimodales, interférentes et unimodales. Il est également important de préciser que la condition interférente de cette expérience utilise un bruit blanc, donc un stimulus purement sensoriel, contrairement à l'association souvent utilisée dans la littérature entre des items sémantiquement non congruents, telle que l'image d'un chien présentée en même temps que le miaulement d'un chat (Schneider *et al.*, 2008; Yuval-Greenberg & Deouell, 2007). Ce résultat confirme qu'une interférence sensorielle peut perturber l'activation des connaissances en mémoire (Brunel *et al.*, 2009) et ainsi diminuer significativement les performances en reconnaissance. Selon le modèle de Versace et collaborateurs (2009), cet avantage de la multimodalité pourrait s'expliquer par le fait que la présentation d'un item visuel en phase test pourrait réactiver le composant auditif lui ayant été associé lors de la première phase par un mécanisme d'intégration multimodale. Selon ces auteurs, la présentation d'un seul composant de la trace permettrait également de réactiver automatiquement l'ensemble des composants associés (Versace *et al.*, 2002; Versace *et al.*, 2009). Ainsi la présentation d'un composant visuel seul dans la condition unimodale pourrait réactiver un composant auditif associé au sein de la trace, ce qui ne serait pas le cas dans la condition interférente. Cette différence entre la condition unimodale et interférente, se manifeste uniquement sur les temps de réponses rapides, et serait donc sous tendue par un processus de familiarité, c'est-à-dire par une composante plus implicite de la reconnaissance. Cette explication peut contraster avec certains résultats de la littérature (Murray *et al.*, 2004 ; Lehmann & Murray, 2005) qui montrent un avantage de la condition unimodale sur la condition interférente en reconnaissance. Nous faisons l'hypothèse, selon le paradigme utilisé et des temps de réponses rapides des sujets, que les auteurs ont essentiellement testé une composante implicite de la reconnaissance dans leur expérience.

Une explication de l'avantage de la condition interférence au niveau du degré de certitude, pourraient être liée au fait que les items interférés seraient reconnus essentiellement par le mécanisme de recollection, lequel est associé à un niveau de conscience élevé. Forcément lorsqu'un item est reconnu (ils le sont moins que dans les autres conditions), la reconnaissance s'accompagne d'un fort degré de certitude. Cette explication est sous tendue par la présence d'une corrélation négative entre le degré de certitude et le pourcentage de bonnes réponses dans la condition interférente. Cette observation peut aussi être liée au fait que les items interférés soient tous associés avec le même bruit sans signification. En effet, cela aurait pour conséquence de les rendre moins distinctif en mémoire, toutefois certains

items interférés seraient plus discriminables que les autres. En effet certains items interférés sont peut-être plus fréquents ou encore plus distinctifs de par leurs propriétés intrinsèques. Cette observation est tout à fait compatible avec l'interprétation de Murray et collaborateurs (2004).

Peut-on penser que les performances dans la condition interférente sont moins bonnes que les autres à cause d'un encodage incident moins robuste ? Bien que les temps de réponse associés à la condition interférente, dans la phase d'encodage, soient plus rapides que ceux associés à la condition unimodale, les pourcentages de réponses correctes sont similaires dans toutes les conditions. On peut ainsi exclure une hypothèse de type « speed-accuracy trade-off » pour les items placés dans la condition interférente.

N'ayant pas observé de différences entre les conditions unimodales et interférence pour la composante de familiarité dans la tâche de reconnaissance, nous proposons de réaliser une nouvelle expérience en gardant la même phase d'encodage mais en demandant aux sujets de réaliser un rappel libre lors de la seconde phase. Cette tâche est réputée pour être saturée en processus de recollection. Si l'on propose l'hypothèse selon laquelle l'interférence sensorielle perturbe la qualité de la représentation, mais pas l'accès à la trace, alors on ne devrait pas observer de différences entre la condition interférence et la condition unimodale. De plus, les items présentés dans la condition multimodale devrait être plus rappelés que les items présentés dans les autres conditions, ce qui devrait argumenter en faveur d'une facilitation du traitement des items multimodaux pour le processus de recollection.

Expérience 6

Hypothèses

H1 : Les items placés dans la condition multimodale devraient être mieux rappelés que les items placés dans les autres conditions.

H2 : Les items unimodaux et interférés devraient être rappelés dans les mêmes proportions si l'hypothèse d'une diminution de la qualité de la trace est valide.

Méthode

Sujets

Vingt-quatre sujets ($M = 21$; $ET = 2$), tous étudiants à l'Université Lyon II, ont participé à cette expérience. Tous étaient volontaires et ont rempli un formulaire de consentement. Les

participants ne présentaient pas de troubles neurologiques et possédaient une vision normale ou corrigée.

Stimuli

Les stimuli utilisés lors de cette expérience sont identiques en tout point à ceux de l'expérience 5.

Procédure et plan expérimental

L'expérience était constituée de deux phases distinctes : une phase d'encodage incident, où les sujets devaient catégoriser (vivant/non vivant) des images à l'aide des touches « a » et « p » du clavier, et une phase de test, dans laquelle les sujets devaient rappeler oralement pendant 2 minutes les images qu'ils avaient catégorisées pendant la première phase. Ces deux phases étaient séparées par un délai de 5 mn, temps durant lequel les sujets réalisaient une grille de Sudoku.

La phase d'encodage incident était la même que celles utilisées dans les expériences 6 & 7. Nous avons donc manipulé dans cette expérience deux facteurs intra-sujets : le facteur CATÉGORIE à deux modalités (vivant/non-vivant) et le facteur MODALITÉ à trois modalités (items ayant été encodés en condition multimodale, unimodale ou interférence).

Lors de la phase de rappel libre, les sujets étaient informés qu'ils avaient catégorisés des images dans la première partie de l'expérience et qu'ils allaient devoir à présent essayer de rappeler pendant deux minutes le plus grand nombre d'images possibles. Il était précisé aux sujets qu'il fallait dénommer précisément les images qu'ils avaient vues. Les réponses produites oralement par les sujets étaient écrites sur une feuille de dépouillement. Les touches de réponse des sujets et les temps de réponse ont été enregistrées lors de cette expérience.

Résultats

Phase d'encodage

Lors de la phase d'encodage, nous avons filtré les temps de réponse situés au-delà de 2000ms, considérant ces latences comme trop lentes pour une catégorisation vivant/non-vivant, nous avons ensuite calculé les temps de réponse (voir TAB. 29 et les pourcentages de réponses correctes (voir TAB. 30) par condition et par sujets.

Temps de réponse (ms)	
vivant	non vivant

multimodal	unimodal	interférence	multimodal	unimodal	interférence
559	583	570	604	588	589
21	20	20	21	15	21

TAB. 29 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et MODALITE

Pourcentage de réponses correctes (%)

vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence
0,98	1,00	0,98	0,98	0,99	0,98
0,01	0,00	0,01	0,01	0,01	0,01

TAB. 30 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE

L'analyse de variances à mesures répétées conduite sur les temps de réponse lors de la phase d'encodage révèle un effet significatif du facteur CATÉGORIE, $F(1,23) = 12.14$; $p < .01$; $\eta_p^2 = .35$, en effet les items vivants sont catégorisés plus rapidement que les items non vivants. L'analyse de variance ne révèle pas d'effet du facteur MODALITÉ, $F(2,46) < 1$. Nous ne relevons pas non plus d'interaction MODALITÉ*CATÉGORIE, $F(2,46) = 2.26$; $p = .11$.

Concernant les pourcentages de bonnes réponses, l'analyse de variance ne révèle aucun effet des facteurs MODALITÉ, CATÉGORIE, ni d'interaction entre ces facteurs. Les performances sont assez élevées quelle que soit la condition de présentation des items ($M = 99\%$).

Phase de rappel libre

Nous avons calculé les moyennes d'items rappelés lors de la seconde phase (voir TAB. 31), par condition et par sujet. Nous avons écarté de l'analyse les mots produits par les sujets qui ne se rapportaient à aucune image catégorisée lors de la première phase.

Pourcentage d'items rappelés par condition (%)

vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence
0,18	0,14	0,13	0,15	0,10	0,10
0,02	0,02	0,02	0,01	0,01	0,02

TAB. 31 – Moyennes des pourcentages d'items rappelés et erreurs types en fonction des facteurs CATEGORIE et MODALITE

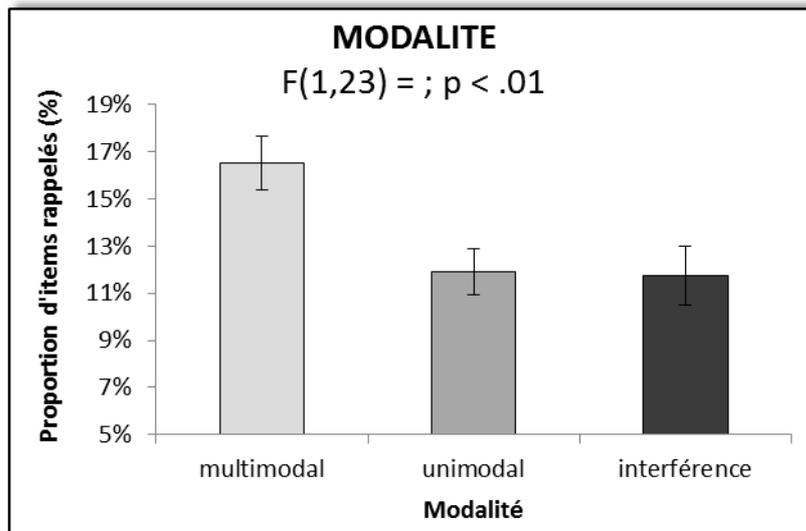


FIG. 42 – Effet du facteur MODALITE sur la proportion totale d'items rappelés par condition

L'analyse de variance conduite sur la proportion d'items corrects rappelés par condition révèle la présence d'un effet du facteur CATÉGORIE, $F(1,23) = 24.17$; $p < .01$; $\eta_p^2 = .51$. Les items vivants sont plus rappelés ($M = 15.08\%$) que les items non vivants ($M = 11.71\%$). Nous ne notons aucune interaction MODALITÉ*CATÉGORIE, $F(2,46) < 1$. L'analyse met en évidence un effet du facteur MODALITÉ, $F(2,46) = 5.33$; $p < .01$; $\eta_p^2 = .19$ (voir FIG. 42). Les items placés dans la condition multimodale ($M = 16.52\%$) sont significativement mieux reconnus que les items placés dans la condition unimodale ($M = 11.9\%$), $F(1,23) = 7.87$; $p < .01$, ou interférente ($M = 11.76\%$), $F(1,23) = 7.04$; $p < .01$. On ne note pas de différence significative entre la condition unimodale et interférente, $F(1,23) < 1$.

Discussion

Dans cette expérience, nous avons émis l'hypothèse selon laquelle les items présentés en condition multimodale devraient être mieux rappelés que ceux présentés dans les autres conditions. En effet, le traitement de la multimodalité en mémoire devrait être amélioré tant pour le processus de familiarité que pour celui de recollection. Notre analyse consistant à séparer les pourcentages de bonnes réponses en considérant les latences de réponses courtes et lentes nous a permis de valider notre hypothèse uniquement sur la composante de familiarité. Dans cette tâche de rappel libre, considérée principalement comme une tâche faisant intervenir le processus de recollection, nous validons notre hypothèse. Dans une expérience plus ancienne (Thompson & Paivio, 1994), des auteurs avaient déjà manipulé la nature de l'encodage en présentant soient des images, des sons ou la présentation simultanée de sons et d'images. Dans une seconde phase de rappel libre, ils avaient montré l'avantage d'une

présentation bimodale sur des présentations unimodales et avaient interprété ces résultats dans le cadre de la théorie du double codage. Dans notre étude, nous n'avons pas présenté de sons seuls durant l'encodage mais nous avons retrouvé l'avantage de la condition multimodale sur la condition unimodale. Nous pouvons donc penser que la multimodalité a une influence sur la récupération mnésique tant au niveau de la composante de familiarité que pour celle de recollection, les résultats étant globalement concordants avec l'expérience 5.

Dans cette expérience, nous n'avons pas trouvé de différences entre la condition interférence et la condition unimodale. En effet, les sujets rappellent autant de mots correspondant aux concepts interférés dans la première phase que de mots correspondant aux concepts présentés uniquement dans la modalité visuelle. Ce résultat indique que dans une tâche sous tendue par des processus de recollection, le rappel n'est pas facilité par une présentation unimodale comparée à la présentation d'une image et d'un bruit blanc. Ce résultat peut paraître étonnant si l'on suppose qu'une perturbation précoce a forcément des conséquences sur des processus plus tardifs. Toutefois, au regard de notre expérience précédente, cela ne semble pas contradictoire, en effet sur les résultats globaux, les items interférés et unimodaux ne différaient pas, le pourcentage d'items correctement rappelés étaient identiques. Toutefois le degré de certitude était plus important pour les items interférés par rapport aux items unimodaux, de plus on relevait une corrélation négative entre pourcentage de réponses correctes et degré de certitude pour les items interférés. Cette dernière observation montre bien qu'il existe un traitement distinct pour les items interférés par rapport aux items unimodaux alors que globalement le pourcentage d'items correctement reconnus est identique dans les deux conditions. En outre, le pourcentage d'items rappelés est significativement plus important dans la condition multimodale que dans la condition interférence, le format bimodal ne permet donc pas de produire à lui seul un codage plus robuste. Une présentation visuelle avec une interférence sensorielle aboutirait à une représentation distincte d'une présentation bimodale congruente, au regard des expériences 5 et 6. Dans notre expérience de rappel libre, ce ne serait donc pas l'accès à la trace qui serait perturbé puisque le taux d'items rappelés dans les deux conditions est sensiblement le même, mais plutôt la qualité de la représentation qui serait dégradée. Pour vérifier cette hypothèse, il faudrait réaliser une tâche pour tester la mémoire de la source, dans laquelle on demanderait aux sujets de se rappeler du jugement qu'ils auraient fait sur l'item lors de l'encodage (Gruber *et al.*, 2008). Ainsi, selon cette hypothèse, pour les items interférés dont les sujets se seraient souvenus, on pourrait imaginer que la remémoration du jugement serait plus difficile que pour les items unimodaux.

Pour conclure cette série d'expérience, nous nous proposons de réaliser une dernière expérience de catégorisation basée sur le paradigme de l'expérience 5. Il est intéressant de revenir à cette tâche de catégorisation avec notre nouveau paradigme et avec l'harmonisation des items que nous avons réalisée entre les expériences 4 et 5. Dans l'expérience 5, nous avons montré que l'interférence sensorielle pouvait perturber le traitement à un niveau précoce, qui pourrait être assimilé à un traitement implicite. Nous allons tester l'hypothèse selon laquelle dans une tâche majoritairement sous tendue par des processus implicites, nous devrions retrouver l'avantage de la condition unimodale sur la condition interférence.

Expérience 7

Hypothèses

H1 : Les items placés dans la condition multimodale devraient être catégorisés plus rapidement que les items placés dans les autres conditions.

H2 : Les items placés dans la condition unimodale devraient être catégorisés plus rapidement que les items placés dans la condition interférence.

Méthode

Sujets

Vingt-quatre sujets ($M = 22$; $ET = 3$), tous étudiants à l'Université Lyon II, ont participé à cette expérience. Tous étaient volontaires et ont rempli un formulaire de consentement. Les participants ne présentaient pas de troubles neurologiques et possédaient une vision normale ou corrigée.

Stimuli

Les stimuli utilisés dans cette expérience sont en tout point identiques à ceux décrits précédemment (voir page 157) dans l'expérience 5.

Procédure et plan expérimental

L'expérience était constituée de deux phases espacées par une tâche distractive d'une durée de 5mn. Dans les deux phases, phase d'encodage et phase de test, les sujets devaient catégoriser (vivant/non vivant) les images qui s'affichaient à l'écran à l'aide des touches « a » et « p » du clavier. La phase d'encodage incident était strictement identique à celle de l'expérience précédente. Dans la phase de test, les sujets devaient catégoriser (vivant/non

vivant) les images qui s'affichaient à l'écran comme lors de la phase d'encodage. Le matériel utilisé, ainsi que l'ordre de présentation des items étaient similaires à ceux de la phase de reconnaissance de la première expérience. Les sujets devaient donc catégoriser, soit des images qu'ils avaient catégorisées selon les conditions unimodale, multimodale et interférente dans la phase d'encodage incident, soit de nouvelles images. Nous avons, dans cette deuxième expérience, manipulé deux facteurs intra-sujets: le facteur ITEM à 4 modalités (item ancien unimodal, item ancien multimodal, item ancien interféré ou item nouveau) et le facteur CATÉGORIE à deux modalités (vivant/non-vivant). De plus, nous avons conservé le facteur intra-sujet MODALITÉ à trois modalités (encodage multimodal, unimodal ou interférent) uniquement pour la phase d'encodage afin de contrôler les effets de la manipulation sensorielle lors de cette phase.

Lors de cette phase de test, chaque essai débutait par un écran blanc d'une durée de 1500ms suivi d'un point de fixation de 500ms. les sujets devaient focaliser leur attention sur le point de fixation jusqu'à l'apparition du stimulus. Les images étaient ensuite présentées pendant seulement 500ms afin d'accroître la difficulté de la catégorisation. L'essai suivant débutait à partir du moment où le sujet avait donné sa réponse. Il était indiqué aux sujets de répondre le plus rapidement possible sans commettre trop d'erreurs. Les touches de réponse étaient contrebalancées sur l'ensemble des sujets mais pas entre la première et la seconde phase pour un même sujet.

Résultats

Nous avons calculé les moyennes des pourcentages de réponses correctes et des temps de réponses, pour les deux phases, par condition et par sujet. Nous avons supprimé les latences supérieures à 2000ms au niveau des données brutes pour chaque sujet, considérant ces temps comme trop lents pour une tâche de catégorisation vivant/non-vivant.

Phase d'encodage

Temps de réponses (ms)					
vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence
565	586	556	597	609	562
18	18	13	19	20	17

TAB. 32 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et MODALITE

Pourcentage de réponses correctes (%)	
vivant	non vivant

multimodal	unimodal	interférence	multimodal	unimodal	interférence
0,96	0,98	0,98	0,94	0,96	0,96
0,02	0,01	0,01	0,02	0,02	0,02

TAB. 33 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et MODALITE

L'analyse de variance réalisée sur les temps de réponse (voir TAB. 32) met en évidence un effet principal du facteur CATÉGORIE, $F(1,23) = 4.59$; $p = .04$; $\eta_p^2 = .17$, ainsi qu'un effet principal du facteur MODALITÉ, $F(2,46) = 7.22$; $p < .01$; $\eta_p^2 = .24$. L'analyse ne révèle pas d'interaction significative CATÉGORIE*MODALITÉ.

Concernant les pourcentages de bonnes réponses (voir TAB. 33), l'analyse ne montre aucun effet des facteurs CATÉGORIE, MODALITÉ ni d'interaction entre ces facteurs. Les performances sont bonnes, quelle que soit la condition de présentation ($M = 96\%$).

Phase de test

Temps de réponses (ms)

vivant				non vivant			
ancien			nouveau	ancien			nouveau
multimodal	unimodal	interférence		multimodal	unimodal	interférence	
486	488	491	496	502	517	516	521
10	8	7	9	9	10	12	9

TAB. 34 – Moyennes des temps de réponses et erreurs types en fonction des facteurs CATEGORIE et ITEM

Pourcentage de réponses correctes (%)

vivant				non vivant			
ancien			nouveau	ancien			nouveau
multimodal	unimodal	interférence		multimodal	unimodal	interférence	
0,94	0,95	0,95	0,97	0,97	0,98	0,95	0,97
0,02	0,02	0,02	0,01	0,01	0,02	0,02	0,01

TAB. 35 – Moyennes des pourcentages de réponses correctes et erreurs types en fonction des facteurs CATEGORIE et ITEM

Nous avons conduit deux analyses de variance, une sur les temps de réponse (voir TAB. 34) et une sur les pourcentages de bonnes réponses (voir TAB. 35). L'analyse de variance conduite sur les pourcentages de bonnes réponses n'a mis en évidence aucun effet des facteurs ITEM et CATÉGORIE, ni d'interaction entre ces facteurs. Les performances atteignent un plafond et ce, quelle que soit la condition ($M = 96\%$). L'analyse de variance conduite sur les temps de réponse révèle un effet du facteur CATÉGORIE, $F(1,23) = 23.07$; $p < .01$; $\eta_p^2 = .50$, et un effet du facteur ITEM, $F(3,69) = 4.4$; $p < .01$; $\eta_p^2 = .16$ (voir FIG. 43). Conformément à notre hypothèse, dans la phase de test, les items placés dans la condition multimodale ($M = 494\text{ms}$) sont catégorisés plus rapidement que les items unimodaux ($M =$

503 ms), $F(1,23) = 5.78$; $p = .01$ et plus rapidement que les items interférés ($M = 504$ ms), $F(1,23) = 4.44$; $p = .02$. De plus, il existe un effet d'amorçage pour les items anciens multimodaux et pour les items anciens unimodaux qui sont caractérisés plus rapidement que les items nouveaux ($M = 508$ ms), $F(1,23) = 16.4$; $p < .01$ et $F(1,23) = 3$; $p = .048$, respectivement. Toutefois on ne note pas de différences entre les items placés dans la condition unimodale et ceux placés dans la condition interférence, $F(1,23) < 1$.

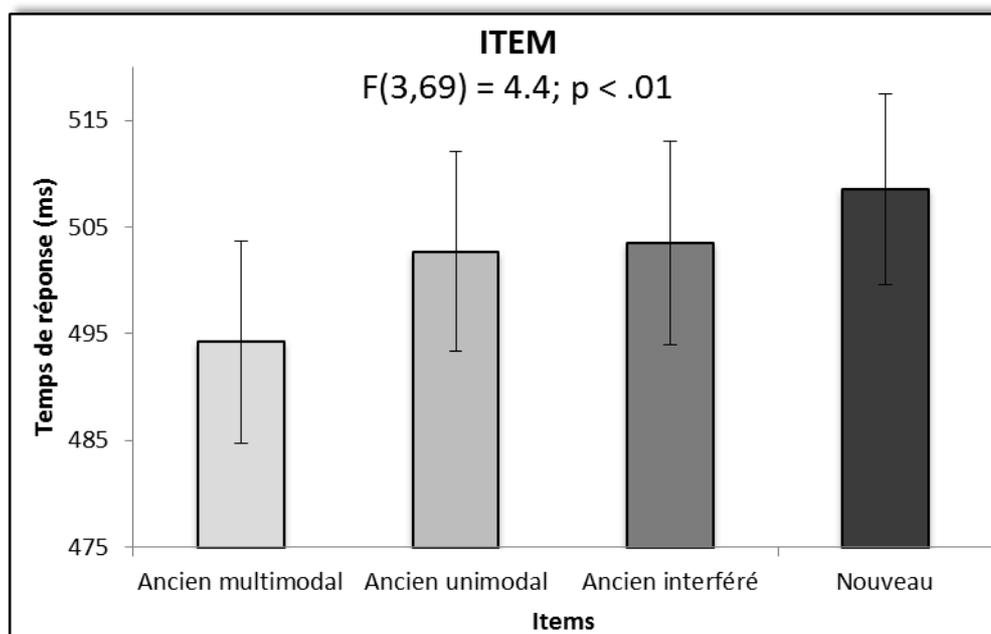


FIG. 43 - Représentation graphique de l'effet du facteur ITEM sur les temps de réponse

Discussion

Dans cette expérience, nous montrons qu'une présentation multimodale permet de faciliter la catégorisation de la présentation ultérieure d'un composant visuel. La présentation de l'item visuel seul dans la phase de récupération permettrait la réactivation du composant auditif associé lors de la première phase. Il semble toutefois que les mécanismes en jeu ne soient pas les mêmes concernant la présentation unimodale. Le fait qu'il n'y ait pas de différence dans la deuxième phase entre les items présentés uniquement visuellement et ceux présentés avec une interférence dans la première phase, contredit des résultats obtenus précédemment dans une tâche de catégorisation (Vallet *et al.*, 2010). Dans cette étude, les sujets devaient catégoriser dans une première phase des items présentés visuellement qui pouvaient être masqués auditivement ou non. Dans une deuxième phase, les sujets devaient catégoriser des items auditifs, la moitié d'entre eux correspondaient aux sons émis par les items, masqués ou non, présentés dans la première phase, l'autre tiers était des items nouveaux. Les sujets catégorisaient plus rapidement les items auditifs lorsque ceux-ci avaient été présentés sous

forme visuelle non masquée dans la première phase. Ce que les auteurs ont interprété comme le résultat d'une activation du composant auditif en mémoire lors de la présentation visuelle, facilitant la catégorisation du composant auditif dans la deuxième phase. Dans notre expérience, l'activation d'une dimension auditive lors de la présentation d'une image lors de la première phase ne semble pas faciliter le traitement de la même image lors de la seconde phase. La présentation de la même image dans la deuxième phase ne permettrait pas à l'activation automatique du composant auditif en mémoire de faciliter la catégorisation, car la présentation visuelle seule suffirait à catégoriser les items présentés (Schneider *et al.*, 2008). Cette explication permettrait d'expliquer les résultats différents obtenus entre les conditions interférentes et unimodales dans cette expérience et celle de Vallet et collaborateurs (2010).

D'autre part, lorsque l'on s'intéresse à la phase d'encodage, on voit que les items multimodaux ne sont pas catégorisés plus rapidement que les items unimodaux, et le pourcentage de bonnes réponses est identique quelles que soient les conditions. Il est donc peu probable que l'on soit en présence d'un biais lié aux conditions d'encodage. Nous pouvons rejeter ici l'hypothèse de l'avantage de la multimodalité basée sur un seul effet attentionnel.

Cette expérience ne nous permet donc pas de mettre en évidence un avantage de la condition unimodale par rapport à la condition interférente. Nous avons déjà émis une première hypothèse explicative consistant en l'avantage de la modalité visuelle sur la réactivation du composant mnésique pour la catégorisation, mais nous ne pouvons rejeter l'idée selon laquelle les mécanismes implicites impliqués dans la tâche de catégorisation et ceux impliqués dans la tâche de reconnaissance ne sont pas identiques.

3.7 Synthèse

Ce chapitre avait plusieurs objectifs, tout d'abord d'élaborer un paradigme expérimental permettant de mettre en évidence des effets multimodaux dans différentes tâches de mémoire. Nous voulions montrer qu'une présentation multimodale congruente facilitait le traitement de familiarité, mais aussi de recollection. Une autre hypothèse reposait sur le fait que l'interférence sensorielle présentée lors de l'encodage pouvait perturber la reconnaissance des items, et ce, pour des traitements précoces. Enfin, nous voulions intégrer nos résultats à la lumière du modèle de mémoire à traces multiples Act-in.

Les expériences 1 à 4 ont eu pour objectif de mettre en place un paradigme original valable dans des tâches de reconnaissance mais aussi de catégorisation, pour pouvoir tester les effets de la multimodalité selon des types de récupération explicite et implicite. Cette série

d'expériences a conduit à des résultats parfois ambigus ou contradictoires et n'a pas réussi à mettre en évidence les effets escomptés. En guise d'interprétation, nous pensons que pour qu'un encodage multimodal soit opérant, il faut que les différentes modalités du même concept soient présentées dans des fenêtres temporelles proches et non à plusieurs items d'intervalle. Il semble d'ailleurs que lorsque les items de différentes modalités sont spatialement et temporellement proches, ils produisent des effets multi-sensoriels optimaux (Lewald & Gusky, 2003; Meredith & Stein, 1993). Fort de ce constant, nous avons adapté un nouveau paradigme en appliquant cette règle lors de la présentation des stimuli dans la phase d'encodage. Nous avons pu observer des effets multimodaux dans les différentes tâches de mémoire que nous avons testées. Concernant la multimodalité, ce nouveau paradigme a mis en évidence des résultats robustes et reproductibles dans plusieurs expériences. En outre, nous avons retrouvé des résultats déjà présents dans la littérature (Murray *et al.*, 2004 ; Lehmann & Murray, 2005). Il semble donc que cet outil expérimental puisse être utilisé dans les prochaines recherches visant à tester l'effet d'un encodage multimodal sur une récupération ultérieure.

Concernant nos hypothèses, nous avons montré qu'un encodage multimodal pouvait favoriser la récupération implicite et explicite. Dans une tâche de reconnaissance, nous avons observé un avantage de la condition multimodale dans la composante de familiarité et dans une tâche de rappel libre, nous avons observé le même avantage de la condition multimodale sur le processus de recollection par rapport aux autres conditions. Il semble donc que les interactions multimodales en mémoire puissent avoir des effets, lors de la récupération, pour des composantes précoces et plus tardives de traitement. Avoir encodé une information multimodale permettrait ainsi d'augmenter ultérieurement la familiarité des items ainsi que la capacité de récupérer l'ensemble des éléments du contexte d'étude (recollection). Lorsque nous avons présenté une image avec un bruit blanc lors de l'encodage, nous avons vu que cela entraînait une dégradation des performances par rapport à la condition multimodale dans toutes les tâches de mémoire. La congruence entre les éléments des diverses modalités lors d'une première présentation est donc essentielle pour favoriser le traitement ultérieur des items (Schneider *et al.*, 2008; Yuval-Greenberg & Deouell, 2007; Lehmann & Murray, 2005 ; Vallet *et al.*, 2010). Nous avons également observé que l'interférence sensorielle perturbait le processus de familiarité dans la tâche de reconnaissance. De plus, nous n'avons pas retrouvé cette perturbation dans la tâche de catégorisation, censée reposer sur des mécanismes implicites tout comme la familiarité, ce qui semble contradictoire. D'autres expériences seront

nécessaires pour explorer cet effet. Nous avons également montré que bien que le pourcentage de bonnes réponses associé à l'interférence ne différait pas globalement de la condition unimodale, celui-ci était corrélé négativement au degré de certitude. Selon nous, cela montre que l'interférence dégrade la représentation de la trace alors que son accès reste possible par rapport à la condition unimodale. Pour valider cette hypothèse, nous pourrions réaliser une expérience nécessitant que les sujets se souviennent des détails de l'item présenté lors de la première phase, comme la couleur d'un cadre entourant les items visuels par exemple. Si la représentation est dégradée avec l'interférence, ce jugement devrait être moins bon pour les items interférés que pour les items unimodaux et multimodaux.

On peut se demander si ces résultats peuvent s'expliquer selon la *Dual-process theory*. Il semble que oui si l'on considère que la recollection et la familiarité sont tous deux des processus continus (Witxed, 2007). Ces deux processus restent distincts, mais la recollection apporte un traitement supplémentaire à celui réalisé par la familiarité. Nous avons vu ici que la multimodalité pouvait augmenter le traitement de familiarité et de recollection. Concernant l'interférence sensorielle, nous pouvons nous demander à quelle étape de traitement elle module la récupération en mémoire. D'après nos résultats, l'interférence perturberait le processus de familiarité sans perturber le processus de recollection. Toutefois, une analyse plus fine, nous a permis de supposer que bien que l'accès à la trace soit possible, la qualité de la représentation de la trace pourrait être perturbée. La familiarité et la recollection pourrait être envisagée comme deux processus interdépendants au sens où le produit de la familiarité a une répercussion sur le traitement réalisé par la recollection.

Qu'en est-il du lien entre mécanisme d'activation postulé par le modèle Act-in et nos résultats ? Selon ce modèle, il semble plausible que la présentation d'un item visuel dans une phase de récupération puisse réactiver un composant associé au sein de la trace et permette ainsi de faciliter le traitement. Nous avons vu que cette facilitation était présente de manière robuste dans différentes tâches de mémoire uniquement lorsque la présentation bimodale lors de l'encodage regroupait des stimuli sémantiquement congruents. Nous avons émis l'hypothèse que l'interférence sensorielle pouvait perturber l'activation des composants de la trace donc à un niveau précoce de traitement, ce que nous avons observé pour la composante implicite de la tâche de reconnaissance. Toutefois, dans la tâche de catégorisation, nous n'avons pas répliqué ce résultat, ce que nous avons tenté d'expliquer par le fait que dans cette tâche, la modalité auditive associée au sein de la trace n'a pas la possibilité de faciliter le traitement, la modalité visuelle étant suffisante pour catégoriser les items. Il serait intéressant

de dégrader les items visuels lors de la seconde phase en catégorisation. En effet, il a été proposé que la difficulté d'accès à l'information visuelle rendrait plus opérant l'information communiquée par la modalité auditive (Schneider *et al.*, 2008).

En résumé, cette partie expérimentale ne répond pas à toutes les interrogations que nous avons posées dans la problématique, il semble d'ailleurs qu'elle apporte d'autres questions il est nécessaire d'apporter des réponses si l'on veut connaître les interactions des modalités sensorielles en mémoire à long terme. Il est donc essentiel de réaliser d'autres expériences afin d'élucider les points de réflexion suivants : l'interférence sensorielle perturbe-t-elle la représentation de la trace tout en laissant possible l'accès à la trace ? Quel est le rôle de la modalité auditive associée au sein de la trace, dans une tâche de catégorisation portant sur la modalité visuelle ?

4. Discussion générale

4.1 Rappel de notre problématique

Les objectifs de ce travail étaient multiples et se sont appuyés sur le modèle théorique Act-in. Ce modèle de mémoire à traces multiples conçoit la mémoire comme un système unique où les traces sont multidimensionnelles et où les expériences vécues par un individu sont représentées au sein de ces traces. Ce modèle est régi par deux mécanismes principaux que sont l'activation et l'intégration et nous avons pour but de mieux comprendre les liens fonctionnels qu'entretiennent ces deux mécanismes. L'intégration et l'activation sont intrinsèquement liées et sont mises en œuvre par le système mnésique dès lors qu'il s'agit de faire émerger une connaissance ou de créer une trace à partir d'une expérience. En d'autres termes, l'activation et l'intégration interviennent tant dans des processus de perception que dans des processus mnésiques. Nous avons donc étudié les relations entre activation et intégration au travers de tâches perceptives et de tâches mnésiques. Nous avons introduit également la notion d'interaction qui selon nous est essentielle pour rendre compte des mécanismes multi-sensoriels perceptifs.

Dans un premier axe de recherche, nous avons essayé d'étudier l'interaction et l'intégration en perception. Nous avons tout d'abord travaillé avec des items de différentes modalités non reliées en mémoire. Selon nous l'association systématique entre deux traits sensoriels doit aboutir à la constitution d'une trace mnésique. La constitution d'une telle trace est progressive et doit se traduire par la création d'une nouvelle entité, dans laquelle les modalités sensorielles préexistantes sont devenues indissociables (Goldstone, 2000). De plus cette association systématique doit se traduire par de meilleures performances que celles observées suite à l'exposition d'un simple contexte multimodal, où les traits sensoriels ne sont pas systématiquement associés. Ces interactions audiovisuelles, lorsqu'elles sont congruentes doivent aboutir à la formation d'une trace mnésique.

Dans un deuxième temps, nous avons utilisé un matériel signifiant dans lequel les items visuels et auditifs renvoyaient aux mêmes concepts afin d'étudier le gain intégratif dans diverses conditions. Nous avons utilisé le Race Model en faisant varier divers paramètres qui nous semblaient important pour observer des modulations de gains multi-sensoriels. Selon nous, cet

outil mesure de l'interaction multimodale plus que de l'intégration. Nous avons donc fait varier la nature de la tâche en utilisant une tâche de détection et une tâche de catégorisation. Ces deux tâches n'impliquent pas les mêmes processus, la tâche de catégorisation demande de traiter une ou des propriétés intrinsèques du stimulus alors que la tâche de détection ne nécessite qu'un traitement partiel de l'item, en ce sens qu'elle demande simplement de détecter la présence de cet item. Le Race Model devrait être plus sensible à des interactions intermodales lorsque les items ne sont pas reliés en mémoire, nous devrions donc observer un gain intégratif multi-sensoriel plus important dans des tâches de détection que dans des tâches de catégorisation. Nous sommes partis d'un résultat de la littérature qui montre que les sujets âgés présentent un gain multi-sensoriel plus important que les sujets jeunes. A partir de cela, nous avons manipulé divers facteurs pour comprendre ce résultat notamment chez les sujets jeunes.

Dans un deuxième axe de recherche, nous avons voulu observer l'effet d'un encodage multimodal sur la récupération à long terme. Il existe des études qui ont montré qu'un encodage multimodal pouvait apporter un bénéfice lors de la récupération (Giard & Perronnet, 1999; Lehmann & Murray, 2005 ; Murray *et al.*, 2004 ; Vallet *et al.*, 2010). Selon le modèle Act-in, les unités sensorielles sont l'élément constitutif des connaissances en mémoire, il est donc logique de penser qu'une manipulation sensorielle lors d'une phase d'encodage pourrait moduler la récupération en mémoire à long terme. Pour cela, nous avons mis au point un paradigme en deux phases permettant de mesurer les effets de l'encodage sur la récupération. Afin de tester les effets sensoriels, nous avons utilisé une condition multimodale congruente (*e.g.*, image du chien et aboiement du chien), une condition multimodale non congruente (*e.g.*, image du chien et bruit blanc) et une condition unimodale (*e.g.*, image du chien). Nous nous attendions à ce que la multimodalité favorise les performances dans diverses tâches de mémoire à long terme, de plus nous pensions qu'une interférence sensorielle perturberait la récupération.

4.2 Rappel des principaux résultats expérimentaux

4.2.1 Expériences portant sur l'intégration multi-sensorielle en perception

Dans cette première partie, nous nous sommes intéressés à la façon dont une trace pouvait se créer à partir d'associations multi-sensorielles. Dans une première expérience, nous avons vu qu'un apprentissage multi-sensoriel incident était meilleur qu'un apprentissage unisensoriel, et le

bénéfice de la multimodalité semblait marqué notamment en début d'apprentissage (Seitz, 2006). Dans une seconde expérience, nous avons comparé deux apprentissages audiovisuels. Un apprentissage était constitué d'associations systématiques entre un son et une forme alors que pour l'autre apprentissage, les formes et les sons n'étaient pas systématiquement associés. Bien que notre résultat soit à considérer avec beaucoup de prudence, il semble que lorsque l'association est systématique, le traitement est plus efficace que lorsqu'il existe un simple contexte multimodal. Une troisième expérience avait pour but de mesurer le gain intégratif associé à la condition d'association systématique par rapport à la condition d'association non systématique. Nous n'avons pas mis en évidence de différences significatives entre nos deux groupes, certainement car le paradigme utilisé était trop différent.

Dans un second temps, nous avons utilisé des stimuli composés de traits sémantiquement redondants. Dans une première expérience, nous voulions répliquer les résultats d'une expérience précédente (Laurienti *et al.*, 2006) ayant montré que les sujets jeunes présentaient moins de gain intégratif que des sujets âgés dans une tâche de catégorisation. Bien que nous ayons obtenu le même pattern de résultats, nous avons observé que les sujets jeunes ne présentaient pas de gain multi-sensoriel. Selon nous, dans le cas d'une présentation de composants sémantiquement très reliés, l'activation automatique des autres modalités au sein de la trace peut expliquer pourquoi la présentation bimodale n'apporte pas beaucoup plus de bénéfices qu'une présentation unimodale, déjà efficiente. L'expérience 2 a montré que sujets jeunes et âgés présentaient un gain intégratif équivalent lorsqu'ils devaient détecter des stimuli plutôt que les catégoriser. Cette observation montre bien que lorsque l'on change la nature de la tâche ainsi que le type de relation entre items visuels et auditifs, les sujets âgés présentent un gain intégratif équivalent aux sujets jeunes. Il semble donc que le gain intégratif soit modulé par la nature de la tâche. Nous avons conduit deux autres expériences où nous avons contrôlé des variables tel que des facteurs attentionnels ou encore la congruence sémantique des traits dans la tâche de détection. Nous avons ainsi pu observer que le gain intégratif était équivalent dans une tâche de détection que le matériel utilisé soit sémantiquement relié ou non. Cela suggère que les liens qu'entretiennent les modalités en mémoire sont moins importants lorsque la tâche ne nécessite pas de traiter l'un des composants sensoriels, mais plutôt de traiter l'objet sur la base d'une propriété extrinsèque (*i.e.*, position spatiale). Enfin dans une dernière expérience, nous avons montré qu'une composante

attentionnelle, souvent impliquée dans le rehaussement du gain intégratif, ne modulait pas ce gain lorsque le sujet doit catégoriser un stimulus bimodal.

Bien que ces expériences méritent un contrôle expérimental plus rigoureux, Nous avons montré que le Race Model était plus sensible pour mesurer des gains multi-sensoriels dans des tâches perceptives que dans des tâches plutôt mnésiques. Il semble que lorsque les items des différentes modalités sont reliés en mémoire, il est plus difficile d'observer un gain intégratif, nous pouvons donc supposer que le Race Model mesure de l'interaction multi-sensorielle plutôt que de l'intégration multi-sensorielle.

4.2.2 Expériences portant sur intégration et récupération en mémoire à long terme

Une grosse partie du travail a consisté à mettre en place un paradigme valide permettant de tester un matériel multimodal selon plusieurs tâches de mémoire. Nous voulions montrer qu'une présentation multimodale congruente facilitait la récupération en mémoire à long terme. Une autre hypothèse reposait sur le fait que l'interférence sensorielle présentée lors de l'encodage pouvait perturber la reconnaissance des items.

Concernant nos hypothèses, nous avons montré qu'un encodage multimodal pouvait favoriser la récupération en mémoire et cela, dans des conditions d'accès implicite et explicite. Dans une tâche de reconnaissance, nous avons observé un avantage de la condition multimodale dans la composante de familiarité et dans une tâche de rappel libre, nous avons observé le même avantage de la condition multimodale sur le processus de recollection par rapport aux autres conditions. Il semble donc que les interactions multimodales en mémoire puissent avoir des effets, lors de la récupération, pour des composantes précoces et plus tardives de traitement. Avoir encodé une information multimodale permettrait ainsi d'augmenter ultérieurement la familiarité des items ainsi que la capacité de récupérer l'ensemble des éléments du contexte d'étude (recollection). Lorsque nous avons présenté une image avec un bruit blanc lors de l'encodage, nous avons vu que cela entraînait une diminution des performances par rapport à la condition multimodale dans toutes les tâches de mémoire. Nous avons également montré que bien que le pourcentage de bonnes réponses associé à l'interférence ne différait pas globalement de la condition unimodale, celui-ci était corrélé négativement au degré de certitude. Selon nous,

cela montre que l'interférence dégrade la représentation de la trace alors que son accès reste possible par rapport à la condition unimodale.

Bien qu'en reconnaissance les items interférés et unimodaux se distinguent en faveur des items unimodaux uniquement sur la composante précoce de traitement (*i.e.*, familiarité), nous n'avons pas retrouvé ce résultat dans une tâche de catégorisation, censée reposer sur des mécanismes implicites. Cela nous a permis de proposer une nouvelle hypothèse concernant le lien entre activation des modalités sensorielles au sein de la trace et mode d'accès à la trace. Nous détaillerons cette hypothèse par la suite.

4.3 Implications théoriques

Les différentes expériences que nous avons conduites dans ce travail de thèse ont porté tant sur la mémoire que sur la perception mais avaient en commun l'aspect multi-sensoriel de la manipulation expérimentale. Nous allons maintenant tenter d'apporter ou de confirmer des éléments théoriques à partir de nos observations, nous tenterons ensuite de spécifier les principales notions que nous avons abordées dans ce travail : l'interaction, l'intégration et l'activation.

4.3.1 Multimodalité et conception distribuée de la mémoire

4.3.1.1 Spécificité de l'association multimodale

Nous avons étudié l'effet de la multimodalité tant lors de l'apprentissage que sur la mémoire à long terme. Lors de notre revue succincte de la littérature, nous avons vu que la multimodalité pouvait améliorer les performances des sujets dans diverses tâches (Brunel *et al.*, 2009 ; Giard & Perronnet, 1999 ; Lehmann & Murray, 2005 ; Molholm *et al.*, 2002 ; Murray *et al.*, 2004 ; Seitz & Dinse, 2007 ; Seitz & Watanabe, 2005 ; Vallet *et al.*, 2010). Ce qui nous a questionnés dans ce travail, c'est la nature de l'association créée lors d'une expérience multi-sensorielle. Les sujets bénéficient-ils seulement d'un contexte multi-sensoriel aspécifique pour améliorer leur performance ou existe-t-il une réelle association entre des traits sensoriels pouvant expliquer les bénéfices observés. En fait, il est peu aisé de répondre à cette question, en tout cas sur la base de notre travail. Nous essaierons d'apporter quelques éléments de réponses.

Sur la base d'un apprentissage multi-sensoriel, nous avons montré qu'un contexte multi-sensoriel apportait de meilleures performances qu'un contexte unisensoriel. Nous avons également montré qu'un contexte multi-sensoriel, lorsque les traits sensoriels sont systématiquement associés, apporte de meilleures performances qu'un contexte multi-sensoriel où les associations entre traits sensoriels sont aléatoires. Bien que ce résultat soit à considérer avec précaution, nous pouvons penser qu'une association multi-sensorielle systématique est plus efficace qu'une association multi-sensorielle non systématique. Toutefois, ce type d'association n'est pas un prérequis à l'existence d'un gain intégratif multi-sensoriel, puisqu'il existe en l'absence de redondance systématique entre traits sensoriels (Molholm *et al.*, 2002). Il semble donc que la présence d'un contexte multi-sensoriel puisse améliorer les performances de façon aspécifique et de façon spécifique lorsque les interactions entre traits sensoriels conduisent à l'intégration des traits au sein d'une trace multi-sensorielle. Nous avons également montré que la nature du traitement conduit sur les stimuli sensoriels jouait un rôle non négligeable au niveau du gain multi-sensoriel que nous pouvions mesurer. En effet, nous avons montré qu'une tâche nécessitant peu l'implication de processus mnésiques était plus à même de mettre en évidence un gain intégratif par rapport à une tâche fortement saturée en processus mnésiques. Il semble donc que le gain intégratif soit modulé en fonction que la tâche porte sur une caractéristique intrinsèque du stimulus (*e.g.*, trait auditif ou visuel) ou sur une caractéristique extrinsèque (*e.g.*, position spatiale).

Nous avons pu établir la spécificité de l'association multi-sensorielle grâce à nos expériences conduites sur des stimuli signifiants. En effet, nous avons montré que l'association entre un stimulus visuel signifiant (*e.g.*, image d'un chien) perturbait la reconnaissance ultérieure de cet item par rapport à l'association de ce même item visuel et du son lui correspondant (*e.g.*, aboiement). Cette dernière observation montre bien qu'un contexte multi-sensoriel n'est pas le seul garant de l'amélioration du traitement. La congruence des traits sensoriels et la nature des processus impliqués dans la tâche semblent essentielles pour expliquer l'amélioration du traitement dans un contexte multi-sensoriel.

Enfin, ces résultats confortent notre position théorique consistant à envisager la mémoire de manière distribuée et multi-sensorielle. En effet, l'association spécifique entre des traits sensoriels congruents permet de rendre compte d'une meilleure récupération en mémoire à long

terme. A l'inverse, une association non congruente est moins bonne pour la reconnaissance ultérieure des items visuels mais c'est surtout la nature du trait auditif qui est intéressant à considérer ici. Car l'interférence, comme nous l'avons appelée, représentée par un simple bruit blanc, suffit à perturber l'émergence des connaissances en mémoire. Cela nous permet de penser que la mémoire humaine est ancrée sur la sensorialité car une manipulation sensorielle lors d'un encodage permet de perturber ou de faciliter l'émergence des connaissances en mémoire à long terme. Nous allons développer plus en détail cet aspect dans la partie suivante.

4.3.1.2 Réinterprétation des effets d'interférence sensorielle en mémoire

Dans les études que nous avons présentées, il y a un résultat intéressant à commenter. Lehmann et Murray en 2005 ont montré que présenter une image d'un objet ou d'un animal en même temps qu'une stimulation tactile ne perturbait pas la reconnaissance ultérieure de ces mêmes images. Toutefois, en 2004, Murray et collaborateurs ont montré en utilisant le même paradigme, que présenter une image d'objet ou d'animal en même temps qu'un son sans signification perturbait la reconnaissance ultérieure de ces mêmes images. Si l'on se place dans une conception structuraliste de la mémoire, il est facile de concevoir qu'une stimulation tactile n'interfère pas avec l'activation d'une représentation en mémoire, il devient plus compliqué d'expliquer pourquoi un son sans signification lors de l'encodage peut perturber la reconnaissance ultérieure des items. On peut supposer à la rigueur qu'une surcharge dans le buffer épisodique (Baddeley, 2000) conduirait à dégrader l'activation des concepts en mémoire mais on ne comprendrait pas pourquoi cette dégradation n'interviendrait que pour l'association de certaines modalités entre-elles. La conception structuraliste qui conçoit les connaissances de manière amodale ne nous semble pas propice à expliquer de tels effets.

Tentons maintenant d'expliquer ces effets sensoriels lors de l'encodage à l'aide d'une vision distribuée de la mémoire. Partons de la présentation d'une image de chien. Cette présentation va activer automatiquement les composants sensoriels associés au sein de la trace (Versace *et al.*, 2002). Ainsi, voir un chien doit réactiver le son associé (aboiement) ou encore la texture (pelage). Par conséquent, si on présente une stimulation auditive ou tactile sans signification en même temps qu'une image lors d'une phase d'encodage, on devrait perturber l'activation des composants sensoriels associés au sein de la trace. Selon cette logique, quelle que soit la modalité

sensorielle de la stimulation sans signification, elle devrait perturber la récupération. Or ce n'est pas ce que l'on observe dans l'expérience de Murray et Lehmann (2005).

On peut alors essayer de compléter ce modèle basé sur l'activation des modalités sensorielles par un facteur de poids du composant sensoriel réactivé. En effet, on peut penser que tous les composants sensoriels n'ont pas les mêmes capacités à réactiver les autres modalités sensorielles (Thompson-Schill *et al.*, 1999). On peut même dire que le poids des modalités sensorielles constitutives des concepts n'est pas le même selon les concepts. Ainsi pour les êtres vivants, les unités visuelles auraient plus d'importance que les autres unités constitutives de la trace (Carbonnel, 2000).

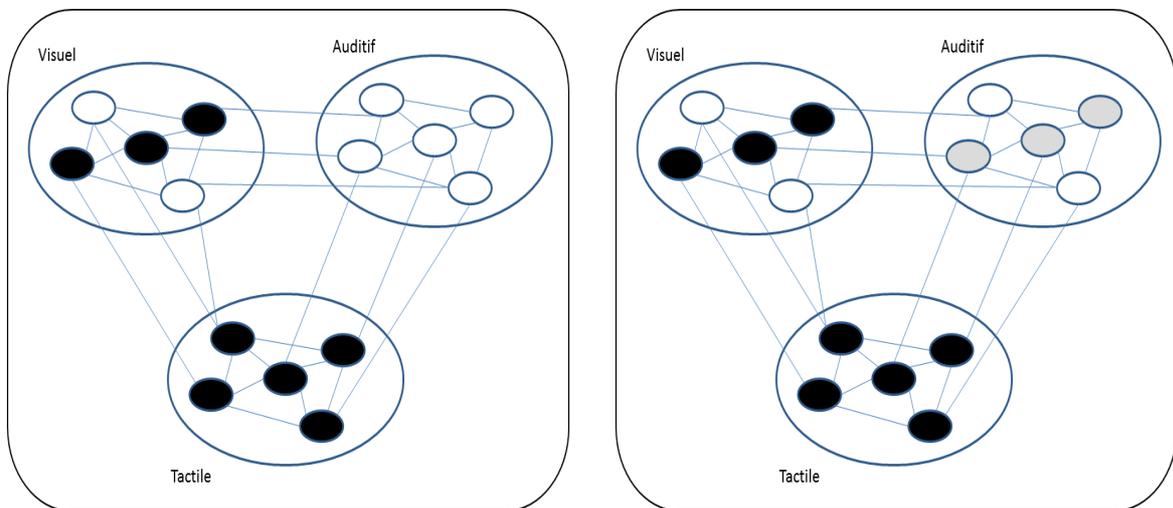


FIG. 44 – Représentation schématisée de l'activation d'une trace mnésique dans un réseau neuronal suite à la présentation d'un stimulus visuel représentant un concept X et d'une stimulation tactile non significative. A -A gauche, la présentation des stimuli a entraîné l'activation de régions dédiées au traitement de ces modalités. B -A droite, l'activation automatique (en gris) des régions auditives suite à la stimulation visuelle.

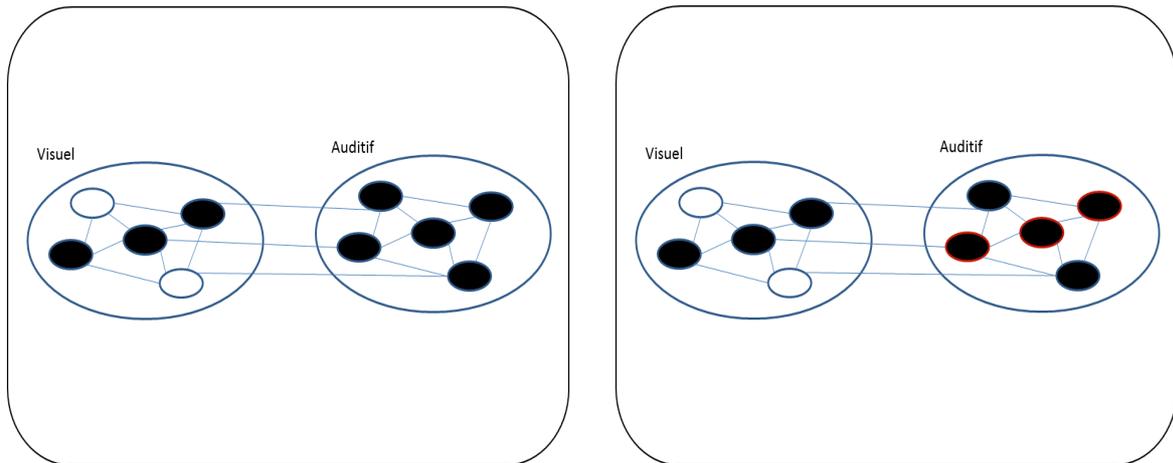


FIG. 45 - Représentation schématique de l'activation d'une trace mnésique dans un réseau neuronal suite à la présentation d'un stimulus visuel représentant le concept X et d'une stimulation auditive non significative. A - A gauche, la présentation des stimuli a entraîné l'activation de régions dédiées au traitement de ces modalités. B - A droite, l'activation automatique (entouré en rouge) des régions auditives suite à la stimulation visuelle est superposée à l'activation par le stimulus auditif non significatif.

Si un concept est peu représenté par des modalités tactiles, alors on peut imaginer que, présenter une stimulation tactile sans signification sera moins apte à perturber l'activation des composants sensoriels de la trace. A l'inverse, si un concept est représenté majoritairement par des unités visuelles et auditives, alors présenter un son sans signification en même temps que l'image d'un chien aura plus de chance de perturber l'activation automatique des composants sensoriels en mémoire. Il est également important de ne pas confondre les propriétés tactiles et les propriétés fonctionnelles. En effet, on sait que les propriétés fonctionnelles peuvent définir majoritairement les objets (Farah & McClelland, 1991 ; Carbonnel, 2000), on pourrait alors penser qu'une stimulation tactile pourrait perturber le traitement des objets en particuliers. Or par unités fonctionnelles, il faut plutôt envisager les composants moteurs, au sens où ils définissent le lien qui unit l'objet à l'action qu'il est possible de réaliser.

Prenons comme exemple, le concept X , défini essentiellement par des propriétés auditives et visuelles. Lors d'un encodage, on présente le concept X avec une stimulation tactile. La présentation du concept X sous forme imagée et la stimulation tactile non significative vont entraîner des activations sensorielles liées aux modalités stimulées (voir FIG. 44–A). Puis le composant visuel va activer les traits auditifs associés au sein de la trace (voir FIG. 44–B) (Versace *et al.*, 2002 ; Versace *et al.*, 2009). Maintenant imaginons que lors de l'encodage, nous

présentons en même temps que la forme imagée de notre concept X , une stimulation auditive non signifiante. Comme lors de la première situation, la présentation visuelle et auditive vont entraîner des activations sensorielles liées aux modalités stimulées (voir FIG. 45–A). Le composant visuel va activer ensuite les traits auditifs associés au sein de la trace, mais cette fois il devrait y avoir une interférence entre les traits perceptivement activés (stimulation auditive effective) et les traits réactivés par le composant visuel (voir FIG. 45–B). Cette interférence lors de l’encodage peut perturber la reconnaissance de différentes façons : soit en interférant avec l’activation des composants sensoriels associés au sein de la trace, soit en diminuant la distinctivité de la trace. Ces deux explications ne sont pas exclusives, on peut même supposer que la perturbation de l’activation des composants sensoriels de la trace lors de l’encodage entraîne une diminution de la distinctivité de cette trace. On serait alors dans une relation de cause à effet, quoi qu’il en soit d’autres études seront nécessaires pour vérifier ces hypothèses.

Une autre hypothèse a été suggérée pour expliquer cet effet d’interférence lors d’une association entre un item visuel et un son sans signification. Selon certains auteurs, la présentation de plusieurs items visuels avec un seul son à travers les essais aurait pour conséquence d’introduire du « bruit » dans la formation de traces multi-sensorielles distinctes (Lehmann & Murray, 2005). Cette hypothèse est tout à fait en accord avec nos résultats et notre interprétation. En effet cette notion de « bruit » dans la trace pourrait tout à fait entraîner une diminution de la distinctivité des traces, comme nous l’avons déjà évoquée. De telle sorte qu’associer toujours le même son à différents items visuels devraient perturber la récupération des items visuels. On suppose également que cet effet pourrait être modulé en fonction des propriétés des items, telle que la fréquence par exemple. En effet, un item peu fréquent serait plus à même d’être sujet à cet effet d’interférence, qu’un item plus fréquemment rencontré, qui ne subirait pas ou peu l’effet d’interférence sensoriel.

4.3.2 Intégration interaction et activation

Nous allons à présent discuter du lien que nous pouvons maintenant établir entre les mécanismes postulés par le modèle Act-In. Nous parlerons du mécanisme d’intégration et d’activation, nous aborderons également le mécanisme d’interaction.

Selon le modèle Act-in, la présentation d'un composant sensoriel va entraîner l'activation automatique des éléments associés au sein de la trace qui vont s'intégrer progressivement pour aboutir à l'émergence des connaissances (Versace *et al.*, 2009). On peut également penser que c'est parce que les éléments d'une trace sont intégrés entre eux que les composants sensoriels peuvent activer d'autres composants d'une même trace. Plus les composants seront intégrés entre eux et plus il sera aisé pour un élément d'activer un autre élément de la trace. L'activation des composants sensoriels par d'autres composants sensoriels serait alors le marqueur d'une intégration des traits. Des études ont montré que suite à l'association de traits sensoriels simples, l'un pouvait réactiver l'autre dans une tâche d'amorçage (Brunel *et al.*, 2009). D'autres études ont montré qu'un composant sensoriel pouvait réactiver un autre composant sensoriel pour des concepts cette fois-ci (Vallet *et al.*, 2010). Dans le chapitre 3, nous avons également montré que des items visuels précédemment associés à des items auditifs congruents contenaient plus d'informations que des items visuels précédemment présentés seuls ou avec un bruit blanc. Ceci signifie, qu'au moment de la récupération, le contexte d'encodage multimodal est facilitateur (Nyberg *et al.*, 2000 ; Wheeler *et al.*, 2000), mais nous pensons que c'est plutôt la réactivation des composants sensoriels associés au sein de la trace qui vont être facilitateurs (Vallet *et al.*, 2010).

Selon cette conception, les mécanismes d'activation et d'intégration seraient intrinsèquement liés puisque de l'un dépendrait l'autre. On peut toutefois se demander si cette vision n'est pas trop proche d'une simple mémoire associative, au sens où toute récupération peut s'expliquer par une activation d'un composant à partir d'un autre. Même si l'activation est nécessaire, elle n'est selon nous pas suffisante pour rendre compte de l'émergence des connaissances. Il est nécessaire que les réseaux activés, s'intègrent en synchronisant leur activité par exemple pour faire émerger un ensemble de traces. Pour s'activer, des éléments doivent être intégrés, mais pour que la connaissance émerge, les composants activés doivent s'intégrer progressivement.

Le mécanisme d'interaction multimodale peut s'intégrer dans notre modèle, pour décrire la manière dont les éléments se combinent entre eux pour former une trace. Dans le chapitre 2, nous avons vu que toute interaction multimodale ne produit pas les mêmes performances lors de l'apprentissage selon l'association systématique des traits entre eux ou non. En fonction de

différents facteurs telle que la contigüité spatiale et temporelle par exemple, les interactions multimodales vont plus ou moins lier les composants entre eux.

4.4 Ouverture et perspectives

Ce travail de thèse a apporté quelques réponses en lien avec la problématique que nous avons posée au départ. Toutefois, les réponses apportées sont à l'origine de nombreuses autres questions qu'il sera intéressant de considérer dans des travaux futurs. Nous allons proposer quelques expériences dont le but sera de répondre aux interrogations que nous avons soulevées dans ce travail. Enfin, le modèle Act-in sur lequel repose l'assise théorique de ce travail de thèse nous permet aussi de repenser les liens entre mémoire et perception. Nous aborderons une conception alternative à la vision classique *top-down* des influences de la mémoire sur la perception. Pour cela nous nous appuyerons sur des arguments issus des neurosciences, ainsi que sur des travaux menés dans notre équipe de recherche.

4.1 Des hypothèses supplémentaires

Comme nous l'avons évoqué plus haut, le paradigme que nous avons mis en place dans le chapitre 3 semble à même de constituer le socle expérimental aux expériences que nous allons décrire.

Une première expérience pourrait avoir pour but d'éclairer l'effet de l'interférence sensorielle lors de l'encodage sur la qualité de la trace. Nous avons émis l'hypothèse selon laquelle cette interférence pouvait diminuer la distinctivité de la trace, nous allons proposer une expérience pour évaluer cette hypothèse. La phase d'apprentissage sera la même que celle décrite dans l'expérience 5 du chapitre 3 avec trois conditions : unimodale, multimodale et interférente. Les sujets devront catégoriser des items visuels faisant référence à des êtres vivants ou non. Nous ajouterons autour des items un cadre de couleur. Lors de la phase de reconnaissance, si les sujets reconnaissent un item, ils devront en plus essayer de se souvenir de la couleur du cadre qui entourait l'item. Si l'interférence sensorielle dégrade la qualité de la trace, alors il devrait y avoir plus de bonnes réponses dans la condition multimodale par rapport à la condition interférence.

Dans une deuxième expérience, nous remplacerons le son servant dans la condition interférence par autant de sons sans signification différents qu'il n'y a d'items visuels dans cette condition.

Dans une deuxième phase, les sujets devront reconnaître les items comme ils l'ont fait dans l'expérience 5 du chapitre 3. Cette manipulation a pour but de tester l'hypothèse avancée par Lehmann et Murray (2005). Selon nous, la simple présentation d'un item non congruent avec un item visuel signifiant devrait entraîner des performances moins bonnes lors de la seconde phase. Toutefois, il se pourrait que les items dans cette nouvelle condition d'interférence soient mieux reconnus que ceux dans la condition d'interférence utilisée dans nos expériences car la distinctivité de la trace pourrait être moins atténuée avec cette manipulation.

Une troisième expérience portera sur la tâche de catégorisation que nous avons réalisée dans l'expérience 7 du chapitre 3. L'absence de différences significatives entre les conditions unimodales et interférentes nous avaient interpellée. L'hypothèse avancée était que pour catégoriser visuellement des items, le son présenté lors de l'encodage a un rôle très mince. En effet la réactivation du composant auditif, lorsque la tâche ne porte que sur un trait, peut difficilement moduler le traitement requis (Schneider *et al.*, 2008). Afin de tester cette hypothèse, nous allons utiliser dans la seconde phase de notre paradigme, des items visuels dégradés qui apparaitront progressivement aux sujets (*e.g.*, Schneider *et al.*, 2008). En rendant plus difficile l'accès à la trace lors de la récupération, on laisserait la possibilité au composant auditif réactivé de faciliter le traitement des items unimodaux par rapport aux items interférés.

Enfin dans une perspective plus large, il serait intéressant d'observer les corrélats neuronaux de notre tâche de reconnaissance de l'expérience 5 du chapitre 3. Nous avons montré qu'une association multimodale congruente était meilleure qu'une association multimodale non congruente, constituée d'un item visuel signifiant et d'un son sans signification (Lehmann & Murray, 2005). Des études ont déjà observé les corrélats d'associations multimodales non congruentes, mais en associant toujours un item visuel signifiant avec un son signifiant non relié (*e.g.*, image du chien et miaulement du chat) (Yuvall-Greenberg & Deouell, 2007). Par exemple, il a été observé une augmentation d'activité dans la bande gamma lors de la présentation d'items sémantiquement congruents par rapport à des items non congruents dans des tâches d'amorçage à court terme (Schneider *et al.*, 2008). A notre connaissance, il n'existe pas encore de travaux en imagerie qui ont réalisé une association multimodale non congruente lors de l'encodage comme la notre ou celle de Lehmann et Murray (2005). Il serait intéressant d'étudier les oscillations cérébrales sous jacentes aux conditions interférence et multimodale, notamment dans la bande

gamma, lors de la phase de reconnaissance. En effet, on peut se demander si une augmentation d'activité gamma dans les régions visuelles (Tallon-Baudry *et al.*, 1998) pourrait être mise en évidence et si une activité plus importante reflèterait le recrutement d'un réseau distribué dans la condition où les items ont été encodés de manière multimodale par rapport à un encodage unimodal ou interférent.

4.2 Mémoire et perception, vers une nouvelle conception

4.2.1 Une nouvelle représentation du lobe temporal médian

Revenons maintenant sur la description que nous avons brièvement faite sur le rôle du lobe temporal médian (LTM) en mémoire notamment dans le chapitre 3 de ce travail de thèse. Dans la vision modulaire du cerveau, c'est à dire pour laquelle chaque région est spécialisée dans une fonction psychologique, le lobe temporal médian occupe une place particulière pour la mémoire dite déclarative. Le LTM inclut l'hippocampe, le cortex entorhinal, le cortex parahippocampique et le cortex péri-rhinal. On oppose à ce système, un système de représentation perceptive (Tulving & Schacter, 1990) dans la voie visuelle ventrale, qui joue un rôle essentiel pour l'apprentissage perceptif et la mémoire. Dans le champ des neurosciences cognitives de la mémoire, la conception en systèmes multiples de mémoire est le paradigme dominant.

Récemment, des auteurs ont posé la question de l'utilité d'une vision modulaire et ont proposé de comprendre le rôle des régions du LTM en lien avec le traitement qui y est réalisé et les représentations qui y sont contenues (pour revue, voir Bussey & Saksida, 2007). En fait ces auteurs proposent que le LTM soit considéré comme l'extension de la hiérarchie des représentations du système visuel ventral. Plus précisément, le cortex péri-rhinal, situé à l'interface entre les systèmes 'mnésiques' et 'perceptifs' dans le paradigme des systèmes de mémoire multiple est une structure adéquate pour tester l'hypothèse selon laquelle le cerveau est divisé en systèmes dédiés à la mémoire et à la perception. Si une structure, tel que le cortex péri-rhinal, est impliquée tant dans des processus mnésiques que perceptifs, ceci pourrait être un argument en défaveur de la conception modulaire. Le cortex péri-rhinal peut être considéré comme faisant partie de la voie visuelle ventrale plutôt que du LTM (Murray & Bussey, 1999). Ce système est supposé contenir une organisation hiérarchique des représentations des stimuli visuels. Bussey et Saksida (2002) ont créé un modèle connexionniste simplifié de ce modèle (voir

FIG. 46). La couche *feature*, correspondant aux régions caudales au cortex péri-rhinal, contiendrait des représentations des traits simples des objets. La couche *feature conjunction*, correspondant au cortex péri-rhinal, contiendraient quant à elle une représentation complexe des conjonctions entre les traits simples. Ce modèle nommé *perceptuel-mnemonic/feature conjunction* (Bussey & Saksida, 2002) propose que le cortex péri-rhinal possède des fonctions perceptives et mnésiques.

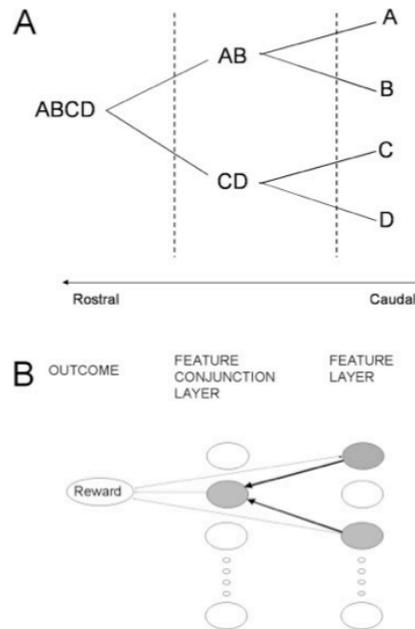


FIG. 46 – Représentation du modèle d’organisations des représentations visuelles dans la voie visuelle ventrale (VVV) imaginé par Bussey et Saksida (2007). A – Les traits simples (A, B, C et D) sont représentés dans les régions caudales de la VVV alors que des représentations plus complexes sont stockées dans des régions plus rostrales. B – Diagramme du modèle connexionniste proposé par Bussey et Saksida (2002).

Si ce modèle est un bon prédicteur des relations entre les processus perceptifs et le cortex péri-rhinal, alors une lésion sélective de ce cortex devrait entraîner une perturbation dans des tâches nécessitant des traitements perceptifs de haut niveau. Dans une expérience, Bussey, Saksida et Murray (2003), créent à partir d’images et à l’aide d’une technique de morphing, des paires de stimuli, dont le nombre de traits en commun augmente progressivement. Ainsi, il existe des paires relativement similaires perceptivement et des paires dissimilaires perceptivement. Les auteurs ont montré chez le singe que des lésions du cortex péri-rhinal perturbaient l’acquisition des discriminations de ces paires, seulement lorsque les paires étaient proches perceptivement. Deuxièmement, les auteurs ont étudié l’effet d’augmentation de similarité perceptive pour des discriminations perceptivement distinctes précédemment apprises. Les auteurs ont d’abord

entraîné les singes jusqu'à ce qu'un critère de réussite soit atteint pour discriminer entre deux stimuli non transformés. Ensuite, lors d'une session de test, les auteurs ont mesuré les performances des singes selon trois conditions : la paire entraînée et deux autres paires de stimuli transformés à partir de la paire entraînée. Les singes avec une lésion du cortex péri-rhinal étaient encore déficitaires lorsque la demande perceptive était importante mais ne présentaient pas de perturbation lorsqu'elle était faible. Les auteurs ont ensuite testé l'implication du cortex péri-rhinal dans des fonctions perceptives pour la reconnaissance d'objet (Bartko *et al.*, 2007). Les auteurs ont montré des objets à des rats puis immédiatement, un autre objet pouvant entretenir avec le premier un haut degré ou un faible degré de similarité perceptive. Les auteurs ont observé que les rats avec lésion péri-rhinale possédaient de faibles performances dans la condition de reconnaissance d'objet avec une haute similarité perceptive. A l'inverse, ils présentaient les mêmes performances que les rats contrôles dans la condition de faible similarité perceptive.

Ces résultats suggèrent l'implication du cortex péri-rhinal dans des tâches de discrimination visuelle. Le cortex péri-rhinal contiendrait des représentations de conjonctions complexes (voir), et ces représentations seraient essentielles pour résoudre des tâches où des discriminations visuelles d'un haut niveau de complexité sont requises. De plus, la lésion du cortex péri-rhinal semble entraîner les mêmes perturbations des fonctions perceptives dans la reconnaissance d'objets. Il existe de plus en plus de données qui démontrent la contribution des structures du LTM dans la perception et dans la mémoire d'objets, de scènes (pour revue, voir Graham *et al.*, 2010). Il a longtemps été considéré que la mémoire était une fonction distincte des autres fonctions telle que la perception (Squire, 2009), il semble aujourd'hui nécessaire qu'une conception plus dynamique et interactive des relations entre mémoire et perception s'enracine dans la pensée scientifique. De plus, cette intrication des processus devrait permettre de revoir les modélisations classiques de la mémoire humaine afin de mieux expliquer et de mieux prédire les nombreuses dissociations observées dans le champ de la neuropsychologie (Graham *et al.*, 2010).

Nous allons terminer cette discussion générale en proposant une conception des relations entre mémoire et perception à partir des travaux réalisés au sein de notre équipe de recherche.

4.2.2 D'une vision top-down vers une vision horizontale des relations perception/mémoire

Nous avons montré qu'une manipulation sensorielle lors de l'encodage modifiait tant la construction des traces que la récupération en mémoire. Ceci est envisageable si l'on suppose que la mémoire est basée essentiellement sur les modalités sensorielles. Ceci est tout à fait compatible avec ce qui est prédit par la *grounded cognition* (Barsalou, 1999, 2003 ; Barsalou *et al.*, 2003). Selon cette théorie, la mémoire est basée sur des traces sensori-motrices qui peuvent être activées quand une partie du stimulus encodé ou de la situation est traitée. Par conséquent, notre mémoire serait le résultat de notre propre perception (pour revue, voir Barsalou, 2008). De nombreuses études ont montré que des régions cérébrales dites modalités spécifiques, qui ne sont pas impliquées dans le traitement perceptif d'un stimulus, vont s'activer quand les sujets réalisent une tâche de mémoire (Meyer *et al.*, 2007 ; Nyberg *et al.*, 2002 ; Wheeler *et al.*, 2002). En d'autre terme, chaque forme de connaissance (*i.e.*, épisodique ou conceptuelle) dépend des propriétés sensori-motrices. Pour résumer, notre mémoire est créée durant nos expériences perceptuelles et est constituée de traces sensori-motrices.

Des études récentes se sont intéressées au lien entre perception et mémoire et ont montré qu'il existait un recouvrement entre processus mnésiques et processus perceptifs. Vermeulen, Corneille et Niedenthal (2008) ont montré par exemple que la mémoire pouvait dépendre des mêmes systèmes sensori-moteurs que ceux impliqués dans la perception du stimulus. Par exemple, dans une condition où la charge sensorielle était importante, trois items (visuels ou auditifs) étaient présentés avant une tâche de vérification de propriétés (visuelle ou auditive). Après que les sujets aient réalisé la tâche de vérification de propriétés, les mêmes items ou trois autres étaient présentés. Dans cette phase, les sujets devaient décider si les trois items étaient les mêmes ou étaient différents à ceux de la première présentation. Les auteurs ont montré que les temps de réponse étaient plus longs pour la tâche de vérification de propriétés et la tâche de reconnaissance quand ces deux tâches impliquaient la même modalité que des modalités différentes. Van Dantzig, Pecher, Zeelenberg et Barsalou (2008) ont également montré que la perception pouvait influencer les processus mnésiques. Dans leur étude, les sujets devaient décider si un stimulus perceptif, présenté dans trois modalités possibles (visuelle, auditive, ou tactile), apparaissait à droite ou à gauche de l'écran. Immédiatement après, les sujets réalisaient une tâche de vérification de propriétés. Les résultats ont montré que les temps de réponse étaient plus longs quand la modalité du stimulus perceptif et la modalité impliquée dans la tâche de vérification de propriétés étaient différentes. Les auteurs ont conclu que les systèmes de

traitement conceptuel et perceptuel sont basés sur le même système distribué. En d'autres termes, les activations en mémoire semblent impliquées les mêmes régions cérébrales que celles impliquées durant le traitement perceptif (pour revue, voir Slotnick *et al.*, 2006). On peut donc penser que si une dimension sensorielle est perçue ou activée en mémoire alors elle va pouvoir avoir une influence dans une tâche de mémoire ou une tâche perceptive. Une expérience menée au sein de notre équipe et que nous avons déjà évoquée, a testé l'hypothèse selon laquelle les dimensions sensori-motrices réactivées puissent influencer une tâche perceptive. Ainsi Brunel et collaborateurs (2009) ont montré que la réactivation d'une dimension sensorielle pouvait moduler le traitement d'une tâche perceptive. Dans une première phase, les auteurs créaient une association systématique entre une forme et un son (*i.e.*, un carré et un bruit blanc) et ils présentaient une autre forme seule (*i.e.*, un cercle). Dans un paradigme d'amorçage à court terme, les auteurs présentaient en amorce la forme visuelle seule (*i.e.*, le carré ou le cercle) et en cible un son que les sujets devaient catégoriser selon qu'il était aigu ou grave. Les auteurs ont montré que seulement la forme associée à un son lors de la première phase, amorçait le son dans la seconde phase. L'activation du composant auditif lors de la présentation de l'amorce entraîne une facilitation du traitement du composant cible. Mais ici, l'activation du composant mnésique se faisait avant le traitement perceptif (Van Dantzig *et al.*, 2008).

Dans une étude récente menée au sein de notre équipe (Riou, Lesourd, Brunel & Versace, 2011), nous avons voulu tester l'influence de l'activation de la dimension visuelle en mémoire sur une tâche perceptive. Nous avons voulu montrer que la manipulation de différence de taille typique, non perceptivement présente, entre une cible et des distracteurs pouvait influencer le pourcentage de bonnes réponses dans une tâche perceptive. Nous avons pour cela utilisé un paradigme de recherche visuelle (Treisman & Gormican, 1988) dans lequel les cibles et les distracteurs diffèrent sur une dimension physique, comme leur taille par exemple. Dans notre expérience, nous avons demandé aux sujets d'identifier une image (représentant un objet familier) perceptivement plus grande ou plus petite que les autres images présentées à l'écran. En d'autres termes, la cible qui était toujours l'image d'un objet de grande taille (taille typique) pouvait varier en terme de taille perceptive à l'écran. Les distracteurs qui avaient toujours la même taille perceptive à l'écran représentaient soit des objets de petite taille typique (*e.g.*, un sifflet) ou de grande taille typique (*e.g.*, une armoire). Deux conditions étaient alors envisageables : dans une première expérience, la différence de taille typique était congruente avec la différence de taille à

l'écran. Dans une deuxième expérience, la différence de taille typique n'était pas congruente avec celle présente à l'écran. Comme nous pensons que mémoire et perception partagent les mêmes ressources ou sont basés sur des systèmes communs, ces deux fonctions devraient s'influencer réciproquement. Notre hypothèse principale portait sur le fait que la taille typique des objets est automatiquement simulée pendant la perception (Barsalou, 1999, 2003). Dans ces deux conditions, nous devrions observer deux effets contraires, dépendant de la congruence ou non de la différence de taille typique. La taille typique des cibles était choisie de telle manière que ces objets soient de grande dimension. Par conséquent, quand les distracteurs ont une grande taille typique, il n'existe pas de différence de taille typique. Par contre, quand les distracteurs ont une petite taille typique, il existe une différence de taille typique. Si la taille typique des objets est activée lors de leur perception, dans l'expérience 1, la congruence entre la différence de taille typique et perceptive devrait conduire à des temps de recherche visuelle de la cible plus rapides. A l'inverse, dans l'expérience 2, la non congruence entre la différence de taille typique et perceptive devrait conduire à un ralentissement des temps de réponse dans la tâche de recherche visuelle. Les résultats obtenus dans l'expérience 1 ont montré que lorsque la différence de taille physique était congruente avec la différence de taille typique, les temps de réponse étaient plus courts. Ceci suggère qu'ajouter une différence de taille typique à une différence de taille perceptive existante améliorerait les performances dans une tâche de recherche visuelle. A l'inverse, les résultats obtenus dans l'expérience 2 ont montré que les temps de réponse étaient plus lents lorsque la différence de taille typique n'était pas congruente avec la différence de taille perceptive. Ajouter une différence de taille typique non congruente à une différence de taille perceptive perturbe le traitement dans cette tâche de recherche visuelle. Ces résultats confirment nos hypothèses selon lesquelles les dimensions mnésiques sont réactivées durant la perception.

L'ensemble de ces travaux insiste sur l'interaction qui existe entre les dimensions perceptives et les dimensions mnésiques. Il constitue ainsi un argument supplémentaire en faveur du lien entre les mécanismes perceptifs et les mécanismes mnésiques. On peut concevoir à présent que les interactions entre mémoire et perception peuvent être portées par l'activation de composants sensori-moteurs perceptivement absents et par leur influence sur une tâche en cours. Cette conception pourrait se modéliser par des interactions de type horizontale et non plus comme l'activation de représentations de haut niveau sur des composants perceptifs de bas niveau.

4.5 Quelques mots pour terminer

Nous nous sommes intéressés tout au long de ce travail de thèse, aux mécanismes sous-jacents à la formation des traces multimodales en mémoire et aux liens entretenus par les modalités sensorielles en mémoire lors de la récupération. Nous avons vu qu'il était possible d'expliquer les effets d'interférence sensorielle observées tant dans ce travail de thèse que dans la littérature à partir d'une conception distribuée de la mémoire. Les relations entre modalités perceptivement présentes et modalités perceptivement absentes tel que nous avons pu le voir, peuvent nous questionner sur les liens plus généraux qui existent entre perception et mémoire. Nous avons montré qu'une association multi-sensorielle lorsqu'elle était systématique, conduisait à de meilleures performances qu'une association non systématique au cours d'un apprentissage. Nous avons aussi montré que l'association d'un composant visuel signifiant (*e.g.*, chien) et d'un composant auditif non signifiant (*e.g.*, bruit blanc) pouvait perturber la reconnaissance ultérieure du composant visuel présenté seul. À l'inverse, la présentation d'une association multi-sensorielle congruente lors de l'encodage (*e.g.*, image de chien et aboiement de chien) facilitait la reconnaissance ultérieure du même composant visuel. Ceci montre que le composant perceptivement absent influence le traitement du composant perceptivement présent (Brunel *et al.*, 2009 ; Riou *et al.*, 2010). L'activation d'un composant en mémoire et son influence sur la perception nous a permis également de penser les liens entre mémoire et perception d'une manière alternative. Selon cette conception, les liens entre mémoire et perception se définiraient plutôt de manière horizontale que verticale selon la vision traditionnelle *top-down*.

Bibliographie

- Aggleton, J.P., & Brown, M.W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences*, 22(3), 425–444.
- Ahissar, M., Hochstein, S. (1993). Attentional control of early perceptual learning. *Proc Natl Acad Sci USA*, 90,5718-5722.
- Baddeley, A. (2000). The episodic buffer : a new component of working memory. *Trends In Cognitive Sciences*, 4(11), 417-423.
- Baier, B., Kleinschmidt, A., & Müller, N.G. (2006). Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *The Journal of Neuroscience*, 26, 12260-12265.
- Baron-Cohen, S., Harrison, J., Goldstein, L.H. & Wyke, M. (1993). Coloured speech perception: is synaesthesia what happens when modularity breaks down. *Perception*, 22, 419-426.
- Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Barsalou, L.W. (2003). Situated simulation in the human conceptual system. *Language and cognitive Processes*, 18, 513-562.
- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, 22, 577-660.
- Barsalou, L.W., Simmons, W.K., Barbey, A.K., & Wilson, C.D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7, 84-91.
- Bartko, S.J., Winters, B.D., Cowell, R.A., Saksida, L.M., & Bussey, T.J. (2007). Perceptual functions of perirhinal cortex in rats: Zero-delay object recognition and simultaneous oddity discriminations. *Journal of Neuroscience*, 27, 2548–2559.
- Brunel, L., Labeye, E., Lesourd, M. & Versace, R. (2009). The sensory nature of episodic memory: sensory priming effects due to memory trace activation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35(4), 1081-1088.
- Brunel, L., Lesourd, M., Labeye, E., & Versace, R. (2010). The sensory nature of knowledge: Sensory priming effects in semantic categorization. *Quarterly Journal of Experimental Psychology*, 23, 1-10.

- Busse, L., Roberts, K.C., Crist, R.E., Weissman, D.H., & Woldorff, M.G. (2005). The spread of attention across modalities and space in a multisensory object. *PNAS*, *102*, 18751-18756.
- Bussey, T.J., & Saksida, L.M. (2007). Memory, Perception, and the Ventral Visual-Perirhinal-Hippocampal Stream: Thinking Outside of the Boxes. *Hippocampus*, *17*, 898-908.
- Bussey, T.J., Saksida, L.M., Murray, E.A. (2003). Impairments in visual discrimination after perirhinal cortex lesions: Testing “declarative” versus “perceptual-mnemonic” views of perirhinal cortex function. *European Journal of Neuroscience*, *17*, 649–660.
- Bussey, T.J., Saksida, L.M., & Murray, E.A. (2002). The role of perirhinal cortex in memory and perception: Conjunctive representations for object identification. In M.P., Witter & F.G., Wouterlood (Eds.), *The Parahippocampal Region: Organization and Role in Cognitive Functions*. (pp 239–254).Oxford: Oxford University Press.
- Carbonnel, S. (2000). Les conceptions “systems unique” de la mémoire: Applications à la neuropsychologie. *Revue de Neuropsychologie*, *10*, 53-76.
- Carbonnel, S., Charnallet, A., David, D., & Pellat, J. (1997). One or several semantic system(s): Maybe none. Evidence from a case study of modality and category-specific “semantic” impairment. *Cortex*, *33*, 391-417.
- Charnallet, A., Rousset, S., Carbonnel, S., & Pellat, J. (1996). A case study of a strong perceptual deficit without agnosia: Evidence against sequential perception and memory. *Brain and Cognition*, *32*, 115-117.
- Charnallet, A., & Carbonnel, S. (1995). Agnosie et déficits catégorie-spécifiques : confrontation des modèles symboliques et épisodiques. *Revue de Neuropsychologie*, *5*, 463-488.
- Chen, Y.-C., & Spence, C. (2010). When hearing the bark helps to identify the dog : semantically-congruent sounds modulate the identification of masked picture. *Cognition*, *114*, 389-404.
- Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257-271.
- Colavita, F.B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409-412.

- Colonius, H., & Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*, 148-154.
- Damasio, A.R. (1989). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25-62.
- Delbeuck, X., Colette, F., & Van der Linden, M. (2007). Is Alzheimer's disease a disconnection syndrome? Evidence from a crossmodal audio-visual illusory experiment. *Neuropsychologia*, *45*, 3315-3323.
- Delbeuck, X., Van der Linden, M., & Collette, F. (2003). Alzheimer's disease as a disconnection syndrome? *Neuropsychological Review*, *13*, 79-92.
- Driver, J., & Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on 'Sensory-Specific' Brain Regions, Neural Responses, and Judgments. *Neuron*, *57*, 11-23.
- Driver, J., & Spence, C. (2000). Multisensory perception: beyond molarity and convergence. *Current Biology*, *10*, R731-R735.
- Eggert, G.H. (1977). Wernicke's Works on Aphasia: A Sourcebook and Review (Vol.1). The
- Eldridge, L.L., Engel, S.A., Zeineh, M.M., Bookheimer, S.Y., & Knowlton, B.J. (2005). A dissociation of encoding and retrieval processes in the human hippocampus. *Journal of Neuroscience*, *25*, 3280-3286.
- Eimer, M., & Schröger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology*, *35*, 313-327.
- Ernst, M.O., & Bühlhoff, H.H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*, 162-169.
- Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience*, *22*, 5749-5759.
- Farah, M.J., Hammond, K.M., Mehta, Z., & Ratcliff, G. (1989). Category-specificity and modality-specificity in semantic memory. *Neuropsychologia*, *27*, 193-200.
- Farah, M.J., & McClelland, J.L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*, 339-357.

- Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C.E., & Fernández, G. (2001). Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nature Neuroscience*, *4*, 1259-1264.
- Fernández, G., & Tendolkar, I. (2008). The rhinal cortex: 'gatekeeper' of the declarative memory system. *Trends in Cognitive Sciences*, *10*, 358-362.
- Fodor, J.A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge: The MIT Press.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "'Mini-mental state". A practical method for grading the cognitive state of patients for the clinician". *Journal of psychiatric research*, *12*, 189-98.
- Fort, A., Delpuech, C., Pernier, J., Giard, M.H. (2002). Early auditory-visual interactions in human cortex during nonredundant target identification. *Cognitive Brain Research*, *14*, 20-30.
- Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., & Làvadas, E. (2005). Audiovisual integration in patients with visual deficits. *Journal of Cognitive Neurosciences*, *17*, 1442-1452.
- Galambos, R. (1992). In E. Basar and T.H. Bulcock (Eds). *Induced Rhythms in the Brain*. (pp. 201-216) Birkhauser.
- Gardiner, J.M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309-313.
- Ghazanfar, A.A., & Schroeder, C.E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*, 278-285.
- Giard M.H., & Peronnet F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473-490.
- Gibson, J.R., & Maunsell, J.H. (1997). Sensory modality specificity of neural activity related to memory in visual cortex. *Journal of Neurophysiology*, *78*, 1263-1275.

- Giray, M., & Ulrich, R. (1993). Motor coactivation revealed by response force in divided and focused attention. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1278-1291.
- Gloor, P. (1997). *The temporal Lobe and the Limbic System*. Oxford, UK: Oxford University Press.
- Goldstone, R.L. (2000). Unitization during category learning. *Journal of Experimental Psychology: Human perception and performance*, *26*, 86-112.
- Goulet, S., & Murray, E.A. (2001). Neural substrates of crossmodal association memory in monkeys: the amygdale versus the anterior rhinal cortex. *Behavioural Neuroscience*, *115*, 271-284.
- Graham, K.S., Barense, M.D., & Lee, A.C.H. (2010). Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia*, *48*, 831-853.
- Greene, A.J., Easton, R.D., & LaShell, L.S.R. (2001). Visual-auditory events: cross-modal perceptual priming and recognition memory. *Consciousness and Cognition*, *10*, 425-435.
- Grossenbacher, P.G. & Lovelace, C.T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in Cognitive Sciences*, *5*, 36-41.
- Grüber, T., Tsvivilis, D., Giabbiconi, C.M., & Müller, M.M. (2008). Induced electroencephalogram oscillations during source memory: familiarity is reflected in the gamma band, recollection in the theta band. *Journal of Cognitive Neurosciences*, *20*, 1043-1053.
- Hamilton, W. (1859). *Lectures on Metaphysics and logic*. Vol. 1. Boston: Gould & Lincoln.
- Henson, R.N., Cansino, S., Herron, J.E., Robb, W.G., & Rugg, M.D. (2003). A familiarity signal in human anterior medial temporal cortex? *Hippocampus*, *13*, 301-304.
- Hintzman, D.L. (1986). « Schema abstraction » in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.
- Hintzman, D.L., Caulton, D.A., & Levitin, D.J. (1998). Retrieval dynamics in recognition and list discrimination: further evidence of separate processes of familiarity and recall. *Memory & Cognition*, *26*, 449-462.

- Horowitz, L.M., & Prytulak, L.S. (1969). Redintegrative memory. *Psychological Review*, 76, 519-531.
- Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513–541.
- James, W. (1890). *The Principles of Psychology*. New York: Dover.
- Jensen, O., Kaiser, J., & Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*, 30, 317-324.
- Kelley, R., & Wixted, J.T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 701-722.
- Kiesel, A., Miller, J., & Ulrich, R. (2007). Systematic biases and Type I error accumulation in tests of the race model inequality. *Behavior Research Methods*, 39, 539-551.
- Kim, R.S., Seitz, A.R., & Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One*, 3: e1532.
- Klimesch, W., Doppelmayr, M., Yonelinas, A., Kroll, N.E., Lazzara, M., Röhms, D., & Gruber, T. (2001). Theta synchronization during episodic retrieval: neural correlates of conscious awareness. *Cognitive Brain Research*, 12, 33-38.
- Klimesch, W., Freunberger, R., & Sauseng, P. (2010). Oscillatory mechanisms of process binding in memory. *Neuroscience & Behavioral Reviews*, 34, 1002-1014.
- Labeye, E., Oker, A., Badard, G., & Versace, R. (2008). Activation and integration of motor components in a short-term priming paradigm. *Acta Psychologica*, 129, 108-111.
- Lambon Ralph, M.A., & Patterson, K. (2008). Generalization and Differentiation in Semantic Memory: Insights from Semantic Dementia. *Annals of New York Academy of Sciences*, 1124, 61-76.
- Laurienti, P.J., Burdette, J.H., Maldjian, J.A., & Wallace, M.T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, 27, 1155-1163.
- Laurienti, P.J., Kraft, R.A., Maldjian, J.A., Burdette, J.H. & Wallace, M.T. (2004). Semantic congruence is a critical factor in multisensory behavioural performance. *Experimental Brain Research*, 158, 405-414.

- Laurienti, P.J., Burdette, J.H., Wallace, M.T., Yen, Y.F., Field, A.S., & Stein, B.E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, *14*(3), 420-429.
- Lavenex, P., & Amaral, D.G. (2000). Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus*, *10*(4), 420-430.
- Lehmann, S., & Murray, M.M. (2005). The role of multisensory memories in unisensory object discrimination. *Cognitive Brain Research*, *24*, 326-334.
- Lewald, J. & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research*, *16*, 468-478.
- Lockhart, R.S. (1969). Retrieval asymmetry in the recall of adjectives and nouns. *Journal of Experimental Psychology*, *79*, 12-17.
- Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., & Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, *270*, 5233.
- Martin, A., Wiggs, C.L., Ungerleider, L.G., & Haxby, J.V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 6566.
- McClelland, J.L., McNaughton, B.L., O'Reilly, R.C. (1995). Why are there complementary learning systems in the hippocampus and neocortex: insights from the success and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419-457.
- McDonald, J.J., Teder-Sälejärvi, W.A., & Hillyard, S.A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*, 906-908.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748.
- Mesulam, M.M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, *10*, 309-325.
- Meyer, M., Baumann, S., Marchina, S., & Jancke, L. (2007). Hemodynamic responses in human multisensory and auditory association cortex to purely visual stimulation. *BMC Neuroscience*, *8*, 14.
- Mickes, L., Wais, P.E., & Wixted, J.T. (2009). Recollection is a continuous process: implications for dual-process theories of recognition memory. *Psychological Sciences*, *20*, 509-515.

- Miller, J.O. (1991). Channel interaction and the redundant targets effect in bimodal divided attention. *Journal of Experimental Psychology : Human Perception and Performance*, *17*, 160-169.
- Miller, J.O. (1982). Divided attention : Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247-279.
- Miller, L.M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *Journal of Neuroscience*, *25*, 5884-5893.
- Molholm, S., Martinez, A., Shpaner, M., & Foxe, J.J. (2007). Object-based attention is multisensory: co-activation of an object's representations in ignored sensory modalities. *European Journal of Neuroscience*, *26*, 499-509.
- Molholm, S., Ritter, W., Javitt, D.C., & Foxe, J.J. (2004). Multisensory visual-auditory object recognition in humans: a high density electrical mapping study. *Cerebral Cortex*, *14*, 452-465.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., & Foxe, J.J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans : a high-density electrical mapping study. *Cognitive Brain Research*, *14*, 115-128.
- Mordkoff, J.T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 520-538.
- Moscovitch, M., & Nadel, L. (1998). Consolidation and the hippocampal complex revisited: in defense of the multiple-trace model. *Current Opinion in Neurobiology*, *8*, 297-300.
- Moscovitch, M., & Nadel, L. (1999). Multiple-trace theory and semantic dementia: Response to K.S. Graham (1999). *Trends in Cognitive Sciences*, *3*, 87-89.
- Mozolic, J.L., Hugenschmidt, C.E., Peiffer, A.M., Laurienti, P.J. (2008a). Modality-specific attention attenuates multisensory integration. *Experimental Brain Research*, *184*, 39-52.
- Mozolic, J.L., Joyner, D., Hugenschmidt, C.E., Peiffer, A.M., Kraft, R.A., Maldjian, J.A., & Laurienti, P.J. (2008b). Cross-modal deactivations during modality-specific selective attention. *BMC Neurology*, *8*.
- Murray, E.A., & Bussey, T.J. (1999). Perceptual-mnemonic functions of peri- rhinal cortex. *Trends In Cognitive Sciences*, *3*, 142-151.

- Murray, M.M., Michel, C.M., De Peralta, R.G., Ortigue, S., Brunet, D., Andino, S.G., & Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *NeuroImage*, *21*, 125-135.
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*, 217-227.
- Nadel, L., & Moscovitch, M. (1998). Hippocampal contributions to cortical plasticity. *Neuropharmacology*, *37*, 431-439.
- Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus*, *10*, 352-368.
- Noppeney, U., Josephs, O., Hocking, J., Price, C.J., & Friston, K.J. (2007). The effect of prior visual information recognition of speech and sounds. *Cerebral Cortex*, *18*(3), 598-609.
- Norman, K.A., O'Reilly, R.C., 2003. Modeling hippocampal and neocortical contributions to recognition memory: a complementary learning systems approach. *Psychological Review*, *110*, 611-646.
- Nyberg, L., Habib, R., McIntosh, A.R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *PNAS*, *97*(20), 11120-11124.
- O'Hare, J.J. (1991). Perceptual integration. *Journal of the Washington Academy of Sciences*, *81*, 44-59.
- O'Reilly, R.C., & Norman, K.A. (2002). Hippocampal and neocortical contributions to memory: advances in the complementary learning systems approach. *Trends in Cognitive Sciences*, *6*, 505-510.
- Opitz, B. (2010). Neural binding mechanisms in learning and memory. *Neuroscience and Biobehavioral Reviews*, *34*, 1036-1046.
- Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *The Journal of Neuroscience*, *26*, 7523-7531.

- Paivio, A (1986). *Mental representations: a dual coding approach*. Oxford. England: Oxford University Press.
- Patterson, K., & Hodges, J.R. (2000). Semantic dementia: one window on the structure and organization of semantic memory. In L. Cermak (Ed.), *Handbook of Neuropsychology, 2nd Edition, Vol. 2: Memory and Its Disorders* (pp. 313-333). Amsterdam/London: Elsevier.
- Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations : Modality-specific effects of prior activation. *Psychonomic Bulletin & Review, 11(1)*, 164-167.
- Peiffer, A.M., Mozolic, J.L., Hugenschmidt, C.E., & Laurienti, P.J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport, 18*, 1077-1081.
- Péris, J-L., & Tiberghien, G. (1984). Effets de contexte et recherche conditionnelle dans la reconnaissance de visages non familiers. *C P C: European Bulletin of Cognitive Psychology, 4*, 323-334.
- Potter, M.C. (1993). Very short-term conceptual memory. *Memory & Cognition, 21*, 156-161.
- Raab, D. (1962). Statistical facilitation of simple reaction time. *Transactions of the New York Academy of Sciences, 43*, 574-590.
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446-461.
- Reder, L.M., Park, H., & Kieffaber, P.D. (2009). Memory systems do not divide on consciousness: reinterpreting memory in terms of activation and binding. *Psychological Bulletin, 135*, 23-49.
- Riou, B., Lesourd, M., Brunel, L., & Versace, R. (2011). Visual memory and visual perception: when visual perception improves visual search. *Memory & Cognition*, doi: 10.3758/s13421-011-0075-2.
- Rogers, T.T., Lambon Ralph, M.A., Garrard, P., Bozeat, S., McClelland, J.L., Hodges, J.R., & Patterson, K. (2004). The structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological Review, 111*, 205-235.

- Rousset, S. (2000). Les conceptions “système unique” de la mémoire: aspects théoriques. *Revue de Neuropsychologie*, *10*, 27-51.
- Rousset, S., Charnallet, A., Diederichs, H., Schreiber, A.C., Carbonnel, S., & Pellat, J. (1996). Can semantic impairment be a consequence of autobiographical memory loss? Second International Conference on Memory, Abano Terme, Italie.
- Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595-598.
- Rugg, M.D., & Yonelinas, A.P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitives Sciences*, *7*, 313-319.
- Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M.P., Dold, G., & Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, *380*, 526–528.
- Schneider, T.R., Debener, S., Oostenveld, R., & Engel, A.K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage*, *42*, 1244-1254.
- Schneider, T.R., Engel, A.K., & Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Experimental Psychology*, *55*, 121-132.
- Seitz, A.R., & Dinse, H.R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology*, *17*, 148-153.
- Seitz, A.R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology*, *16*, 1422-1427.
- Seitz, A., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, *9*, 329–334.
- Seitz, A.R., & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature*, *422*, 36.
- Senkowski, D., Schneider, T.R., Foxe, J.J., & Engel, A.K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends in Neurosciences*, *31*, 401-409.

- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788.
- Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *NeuroReport*, *12*, 3849-3852.
- Shams, L., & Seitz, A.R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411-417.
- Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, *11*, 505-509.
- Slotnick, S.D., & Schacter, D.L. (2006). The nature of memory related activity in early visual areas. *Neuropsychologia*, *44*, 2874–2886.
- Simmons, W.K., & Barsalou, L.W. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*, 451-486.
- Singer, W., & Gray, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, *18*, 555-586.
- Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, *128*, 153-161.
- Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science & Technology*, *28*, 61-70.
- Snowden, J., Griffiths, H., & Neary, D. (1995). Autobiographical experience and word meaning. *Memory*, *3*, 225-246.
- Squire, L.R., Cohen, N.J., Nadel, L. (1984). The medial temporal region and memory consolidation: a new hypothesis. In H. Weingartner & E. Parker (Eds.), *Memory consolidation* (pp. 185-210). Hillsdale, NJ: Erlbaum.
- Squire, L. R. (2009). The legacy of patient H.M. for neuroscience. *Neuron*, *61*, 6–9.
- Squire, L.R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*, 1380-1386.
- Stein, B.E., & Meredith, M.A. (1993). *The merging of the senses*. Cambridge: The MIT Press

- Stein, B.E., Wallace, M.T., & Meredith, M.A. (1995). Neural Mechanisms mediating attention to multisensory cues. In M.S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 683-702). Cambridge: MIT Press.
- Stephen, J.M., Knoefel, J.E., Adair, J., Hart, B., & Aine, C.J. (2010). Aging-related changes in auditory and visual integration measured with MEG. *Neuroscience Letters*, *484*, 76-80.
- Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*, 151-162.
- 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*, 722-734.
- Tallon-Baudry, C., & Bertrand, O., Perronnet, F., & Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short-term memory task in humans. *Journal of Neuroscience*, *18*, 4244-4254.
- Talsma, D., Doty, T.J., & Woldorff, M.G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex*, *17*, 679-690.
- Tanabe, H.C., Honda, M., & Sadato, N. (2005). Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. *The Journal of Neuroscience*, *25*(27), 6409-6418.
- Teder-Sälejärvi, W.A., Di Russo, F., McDonald, J.J., & Hillyard, S.A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*, 1396-1409.
- Thompson, V.A., & Paivio, A. (1994). Memory for pictures and sounds: independence of auditory and visual codes. *Canadian Journal of Experimental Psychology*, *48*(3), 380-398.
- Thompson-Schill, S.L., Aguirre, G.K., D'Esposito, M., & Farah, M.J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia*, *37*, 671-676.
- Treisman, A. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*, 97-136.
- Treisman, A.M., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search

asymmetries. *Psychological Review*, 95, 15–48.

Tsushima, Y., Seitz, A.R., & Watanabe, T. (2008). Task-irrelevant learning occurs only when the irrelevant feature is weak. *Current Biology*, 18, 516-517.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology : Human Perception and Performance*, 24, 830-846.

Tulving, E. (1995). *Organization of memory: Quo Vadis?* In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 839-847). Cambridge, Mass: MIT Press.

Tulving, E. (1991). Concepts of human memory. In L.R. Squire, G. Lynch, & J. McLaugh (Eds.), *Memory: organization and locus of change* (pp. 3-32). New York, Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26, 1-12.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory*,. London: Academic Press.

Tulving, E., Schacter, D.L. (1990). Priming and human memory systems. *Science*, 247, 301–306.

Vallet, G., Brunel, L., & Versace, R. (2010). The perceptual nature of the cross-modal priming effect. *Experimental Psychology*, 6, 1-7.

Van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, 32, 579–590.

Van den Bergh, O., Vrana, S., & Eelen, P. (1990). Letters from the heart : affective categorization of letter combinations in typists and nontypists. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 16(6), 1153-1161.

Vermeulen, N., Corneille, O., & Niedenthal, P. M. (2008). Sensory load incurs conceptual processing costs. *Cognition*, 109, 287–294.

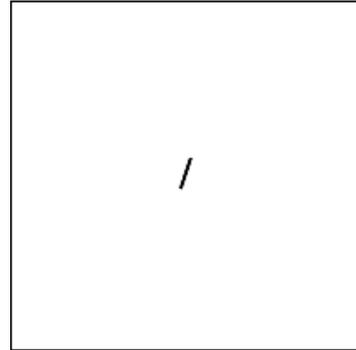
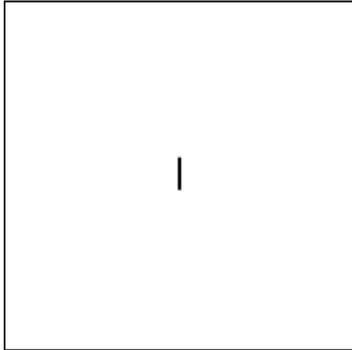
Versace, R., Nevers, B., & Padovan, C. (2002). *La mémoire dans tous ses états*. Marseille: Solal.

Versace, R., Labeye, E., Badard, G., & Rose, M. (2009). The Contents of Long-Term Memory and the Emergence of Knowledge. *The European Journal of Cognitive psychology*, 21, 522-560.

- Vincent, S.B. (1912). The function of the vibrissae in the behaviour of the white rat. *Behavior Monographs*, 1(5).
- Von Kriegstein, K., & Giraud, A.L. (2006). Implicit multisensory associations influence voice recognition. *PLOS Biology*, 4(10): e326. DOI:10.1371/journal.pbio.0040326.
- Vroomen, J., & De Gelder, B. (2000). Sound enhances visual perception : cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1583-1590.
- Warrington, E.K., & McCarthy, R.A. (1987). (1987). Categories of knowledge. Further fractionations and an attempted integration. *Brain*, 110, 1273-1296.
- Wheeler, M.E., Petersen, S.E., & Buckner, R.L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. *PNAS*, 97(20), 11125-11129.
- Wixted, J.T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114, 152-176.
- Yonelinas, A.P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *Journal of Memory and Language*, 46, 441-517.
- Yonelinas, A.P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 1341-1354.
- Yonelinas, A.P., & Jacoby, L.L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, 34, 622-643.
- Yuvall-Greenberg, S., & Deouell, L.Y. (2007). What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. *The Journal of Neuroscience*, 27(5), 1090-1096.
- Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, 28, 611-619.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168-171.

A. Annexes – Matériel experimental

Annexe 1 – Stimuli visuels utilisés lors des expériences 1 et 2 du chapitre 2 – Partie 1

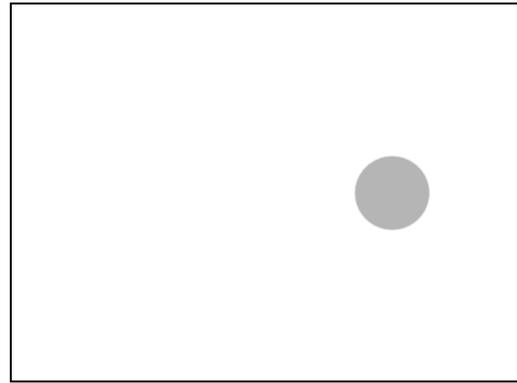
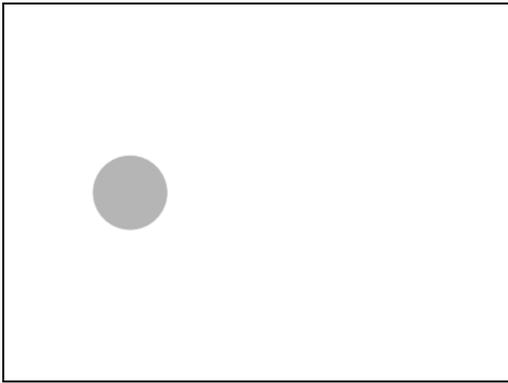


Annexe 2 – Stimuli auditifs utilisés lors de l'expérience 1 et 2 du chapitre 2 – Partie 1

Aigu.aiff	Grave.aiff
-----------	------------

Annexe 3 – Stimuli visuels utilisés lors de l'expérience 3 du chapitre 2 – Partie 1 et lors des expériences 2, 3 et 4 du chapitre 2 – Partie 2

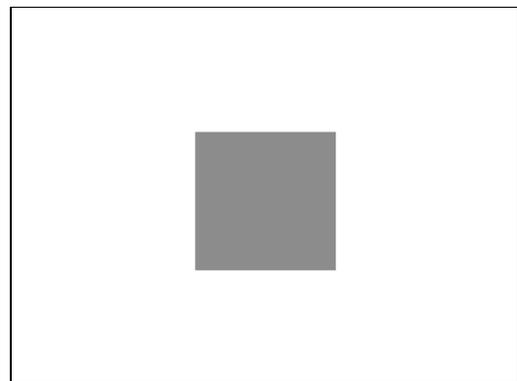
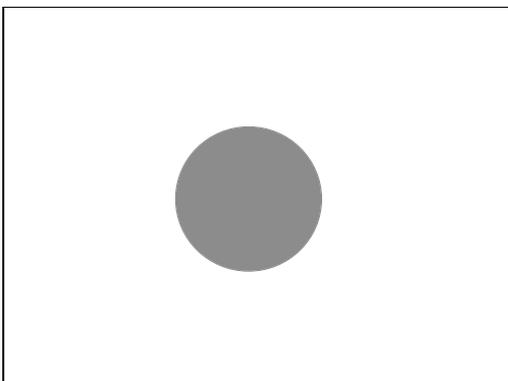




Annexe 4 - Stimuli auditifs utilisés lors de l'expérience 3 du chapitre 2 - Partie 1 et lors de l'expérience 2 du chapitre 2 - Partie 2

AiguG.aiff	aiguD.aiff
graveG.aiff	graveD.aiff
Silence.aiff	

Annexe 5 - Stimuli visuels utilisés lors de l'expérience 1 du chapitre 2 - Partie 2



Annexe 6 - Stimuli auditifs utilisés lors de l'expérience 1 du chapitre 2 - Partie 2

rond.aiff	carre.aiff
silence.aiff	

Annexe 7 - Stimuli auditifs utilisés lors de l'expérience 3 et 4 du chapitre 2 - Partie 2

rondG.aiff	rondD.aiff
carreG.aiff	carreD.aiff
silence.aiff	

Annexe 8 – Fichier matlab Colonius.m permettant de calculer un indice d'intégration multi-sensoriel (Colonius & Diederich, 2006)

```
clear

[file, path] = uigetfile('*.xls', '');

%%%récupère une matrice excel avec des lignes vides et la condense%%%%%%%%
b = xlsread(file);

n = length(b);

%remplissage direct de la matrice B

j = 0;

k = 0;

l = 0;

for i = 1:n;

    %premiere colonne

    a = isnan(b(i,1));

    if (a == 0)

        j = j+ 1;

        B(j,1)= b(i,1);

    end

    %deuxieme colonne

    d = isnan(b(i,2));

    if (d == 0)

        k = k+ 1;

        B(k,2)= b(i,2);

    end

    %troisieme colonne

    c = isnan(b(i,3));

    if (c == 0)
```

```

        l = l+ 1;

        B(l,3)= b(i,3);

    end

    i = i+1;
end

%%%%%connaitre le nb de valeurs dans la colonne%%%%%%%%%%

nbV = 0;

nbA = 0;

nbAV = 0;

sumAV = 0;

i=1;

while (B(i,3)> 0)

    nbAV = nbAV+1;

    sumAV = sumAV+B(i,3);

    i=i+1;

end

meanAV = sumAV/nbAV;

i=1;

while (B(i,1)> 0)

    nbV = nbV+1;

    i=i+1;

end

for i =1:nbV

C(i) = B(i,1);

end

x=C(:);

x = sort(x);

```

```
clear C;

i=1;
while (B(i,2)> 0)

    nbA = nbA+1;

    i=i+1;
end

for i =1:nbA
C(i) = B(i,2);
end

y=C(:);

y = sort(y);

k=nbA+1;

for i=1:nbA

    E(k-i)=y(i);
end

y=E(:);

m = min(nbV,nbA);

M = max(nbV,nbA);

for i = 1:m

matcomp(i)=min(x(i),y(i));

end

%if nbV>nbA

%    for i=m+1:M

%        matcomp(i)=x(i);

%    end

%else

%    for i=m+1:M
```

```
%      matcomp(i)=y(i);

%      end

%end

matcomp=matcomp';

A=mean(matcomp);

C = A - meanAV;

fprintf(1,'Estimation aire sous la courbe pour sujet %s: %f\n',file,C);

%S=input('\nnom du fichier de sortie ?','s');

%xlswrite(strcat(S, '.xls'),matcomp);
```

Annexe 9 – Fichier matlab RaceModel_ind.m permettant de calculer pour chaque sujet les 19 fractiles pour les distributions auditives, visuelles, audiovisuelles et pour la courbe théorique audiovisuelle. Il permet également de tracer les distributions, il donne un aperçu de la violation de l'inégalité de Miller.

```
%RaceModel_ind largement inspiré de cumulposav2d (J.BESLE, U821): trace et
enregistre les distributions cumulees des temps de

%reaction a partir d'un fichier excel

%      auteur : Mathieu Lesourd, EMC. 16/01/2008

%nouvelle version permet de lire directement un fichier sans que celui-ci
%ait été préalablement nettoyé

%      auteur : Mathieu Lesourd, EMC. 30/01/2009

%nouvelle version permettant de lire des fichier sans intégration

%      auteur : Maud PERRIN (étudiante Licence 3 MISASHS), 20/02/2009

clear

hold off
```

```
limiteinf=50;

[file, path] = uigetfile('*.xls', '');

%récupère une matrice excel avec des lignes vides et la condense

b = xlsread(file);

n = length(b);

%remplissage direct de la matrice B

j = 0;

k = 0;

l = 0;

for i =1:n;

    %premiere colonne

    a = isnan(b(i,1));

    if (a == 0)

        j = j+ 1;

        B(j,1)= b(i,1);

    end

    %deuxieme colonne

    d = isnan(b(i,2));

    if (d == 0)

        k = k+ 1;

        B(k,2)= b(i,2);

    end

    %troisieme colonne

    c = isnan(b(i,3));

    if (c == 0)

        l = l+ 1;
```

```
        B(1,3)= b(i,3);

    end

    i = i+1;
end

nV = j;

nA = k;

nAV = 1;

M = max(B); % renvoie les maximum de chaque colonne
N = max(M); % renvoie la valeur maximum de M donc de la matrice B

densite= zeros(N,4);

fprintf(1,'calcul des fonctions de frequences cumulees dans le fichier %s
par cumul.m\n', file);

for i = 1:nV;

% pour les TR visuels

    if B(i,1)>limiteinf

        densite(B(i,1),1) = densite(B(i,1),1)+1;

    end

end

for i = 1:nA;

% pour les TR auditifs
```

```
    if B(i,2)>limiteinf
        densite(B(i,2),2) = densite(B(i,2),2)+1;
    end
end

for i = 1:nAV;

% pour les TR audiovisuel

    if B(i,3)>limiteinf
        densite(B(i,3),3) = densite(B(i,3),3)+1;
    end
end

TRv = mean(B(:,1));
TRa = mean(B(:,2));
TRav = mean(B(:,3));

sdTRv = std(B(:,1));
sdTRa = std(B(:,2));
sdTRav = std(B(:,3));

%autre methode pour les calculer
%TRv = sum(B(:,1))/nV; TRa = sum(B(:,2))/nA; TRav = sum(B(:,3))/nAV;
%sv=0; sa=0; sav=0;
%for i=1:nV;
```

```
%      sv = sv + ((B(i,1)-TRv)*(B(i,1)-TRv));

%end

%for i=1:nA;

%      sa = sa + ((B(i,2)-TRa)*(B(i,2)-TRa));

%end

%for i=1:nAV;

%      sav = sav + ((B(i,3)-TRav)*(B(i,3)-TRav));

%end

%sdTRv = sqrtm(sv/(nV-1));

%sdTRa = sqrtm(sa/(nA-1));

%sdTRav = sqrtm(sav/(nAV-1));

fprintf(1,'V:      %f(%f)                A:      %f(%f)                AV:
%f(%f)\n',TRv,sdTRv,TRa,sdTRa,TRav,sdTRav );

%fprintf(1,'V: %f   A: %f   AV: %f',nV,nA,nAV);

for i=1:3 % conversion en fréquences

    densite(:,i) = densite(:,i)/sum(densite(:,i));

end

freqcumul = cumsum(densite); % calcul des freq cumulées croissantes pour
les 3 modalités

somme = densite(:,1) + densite(:,2); % calcul la somme des fréquences A et
V
```

```
cumulsomme = cumsum(somme); % distribution cumulée correspondant à A+V

for i =1:N

    if cumulsomme(i)>1

        cumulsomme(i) = 1;

    end

end

violation = freqcumul(:,3) - cumulsomme;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%calcul de la magnitude du
RSE%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

m=1;

taille = length(freqcumul(:,3));

while (freqcumul(m,3) < 1) & (m < taille) %size(freqcumul))

    m = m+1; %determine la limite m pour l'affichage des valeurs en
abscisse

end % et pour le calcul de la
magnitude du RSE

n=m-1;

while (violation(n) <= 0 & n >2)

    n = n-1; %determine la limite m ou A+V passe definitivement au dessus
de AV

    %si n arrive à 0 il n'y a pas d'integration

end % pour le calcul de la magnitude du RSE

deltaT=0;
```

```
for i=1:m

    if violation(i)>0

        violationpositif(i)=violation (i);

        %calcul de la latence d'integration

        deltaT=deltaT+1;

    else

        violationpositif(i)=0;

    end

end

end

%Matrice de violation de l'inégalité de Miller avec seulement les valeurs
%positives

j=1;

for i = 1:m

    if violationpositif(i)>1e-04;

        alpha(j) = violationpositif(i);

        j=j+1;

    end

end

end

magnitudo= sum(violationpositif);

%/sum(freqcumul(1:n,3))

fprintf(1,'magnitudo = %f intervalle = %f',magnitudo, deltaT);

%%%% Tracé des distributions %%%%%

figure(1)
```

```
%subplot (2,1,1);

subplot('Position',[0.15 0.4 0.75 0.5]);

plot(freqcumul(:,1),'r-','LineWidth',1);

hold on;

plot(freqcumul(:,2),'g-','LineWidth',1);

plot(cumulsomme,'b-','LineWidth',1);

plot(freqcumul(:,3),'y','LineWidth',2);

legend('V','A','A+V','AV',4);

axis([0 m*1.1 0 1]);

ylabel('Prob. ');

xlabel('TR (msec)');

S = sprintf('Fonction de densite des frequences cumulees : P(X <=
TR)\n%s',file);

title(S);

hold off;

subplot('Position',[0.15 0.1 0.75 0.2]);

plot(violation);

plot(violationpositif,'g');

axis([0 m*1.1 -0.3 0.3]);

title('Violation du modele de competition');

xlabel('TR (msec)');

hold off;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%vincentization%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

j=1;
```

```
k=1;

l=1;

n=1;

for i=2:N

    if freqcumul(i,1)>l*0.05 & l<20

        vincentize(l,1)=i;

        l=l+1;

    end

    if freqcumul(i,2)>n*0.05 & n<20

        vincentize(n,2)=i;

        n=n+1;

    end

    if freqcumul(i,3)>j*0.05 & j<20

        vincentize(j,3)=i;

        j=j+1;

    end

    if cumulsomme(i)>k*0.05 & k<20

        vincentize(k,4)=i;

        k=k+1;

    end

end

figure(2)
```

```

plot(vincentize(:,1),[1:19]','r-','LineWidth',1)

hold on

plot(vincentize(:,2),[1:19]','g','LineWidth',1)

plot(vincentize(:,4),[1:19]','b','LineWidth',1)

plot(vincentize(:,3),[1:19]','y','LineWidth',2)

legend('A','V','A+V','AV',4)

ylabel('Prob.')

xlabel('TR (msec)')

title('Vincentization');

%set(gca,'XLim',[0 m*1.1]);

xlim([0 m*1.1]);

hold off

fprintf('\nProcessus de vincentization')

S=input('\nnom du fichier de sortie ?','s');

save(strcat(S,'.mat'),'vincentize','-ASCII');

fprintf('\ncreation de la distribution individuelle de violation du
model?');

T=input('\(O)ui ou (N)on','s');

if T == 'o'

xlswrite(strcat(S,'RMI.xls'),violation);

end

```

Annexe 10 – Fichier matlab cumul_intersujets.m cree un fichier par fractile regroupant les valeurs pour ce fractile pour chaque sujet.

```
%cree a partir de fichiers de fractiles de deux distributions (un fichier
```

```
par sujet)

%      un fichier de distribution des valeurs entr les sujets pour chaque
fractile (un fichier par fractile)

%      auteur : Julien Besle, U821. 28/02/2005 – modifié par Mathieu
Lesourd, labo EMC (2008)

clear all

hold off

nbsuj=input('nombre de sujets ?');

for i=1:nbsuj

    [filename, pathname] = uigetfile('*.mat', 'Fichier de donnees du
sujet');

    vincentize(:, :, i)=load(filename, '-ASCII');

end

nbfractile = size(vincentize,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%fichier de
sortie%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

S=input('nom de l''analyse ?','s');

for i=1:nbfractile

    ficout=fopen(['', strcat(S,int2str(i))], 'w');

    fprintf(ficout,'variable 1\tsujets\ttemps de reaction\n');

    for j=1:nbsuj

        fprintf(ficout,'a+v\t%d\t%f\n',j,vincentize(i,4,j));
```

```

end

for j=1:nbsuj

    fprintf(ficout,'av\t%d\t%f\n',j,vincentize(i,3,j));

end

fclose(ficout);

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%tracage    de    la    distribution
vincentisee%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

vincentized=zeros(nbfractile,4);

for i=1:nbsuj

    vincentized=vincentized+vincentize(:, :, i);

end

vincentized=vincentized/nbsuj;

plot(vincentized(:,1),[1:19]','r-','LineWidth',2)

hold on

plot(vincentized(:,2),[1:19]','g','LineWidth',2,'Marker','^','MarkerEdgeColor','k','MarkerFaceColor','g','MarkerSize',2)

plot(vincentized(:,4),[1:19]','b','LineWidth',2,'Marker','o','MarkerEdgeColor','k','MarkerFaceColor','b','MarkerSize',2)

plot(vincentized(:,3),[1:19]','y','LineWidth',2,'Marker','s','MarkerEdgeColor','k','MarkerFaceColor','y','MarkerSize',2)

legend('V','A','A+V','AV',2)

ylabel('P(TR < t) (%)')

xlabel('t (msec)')

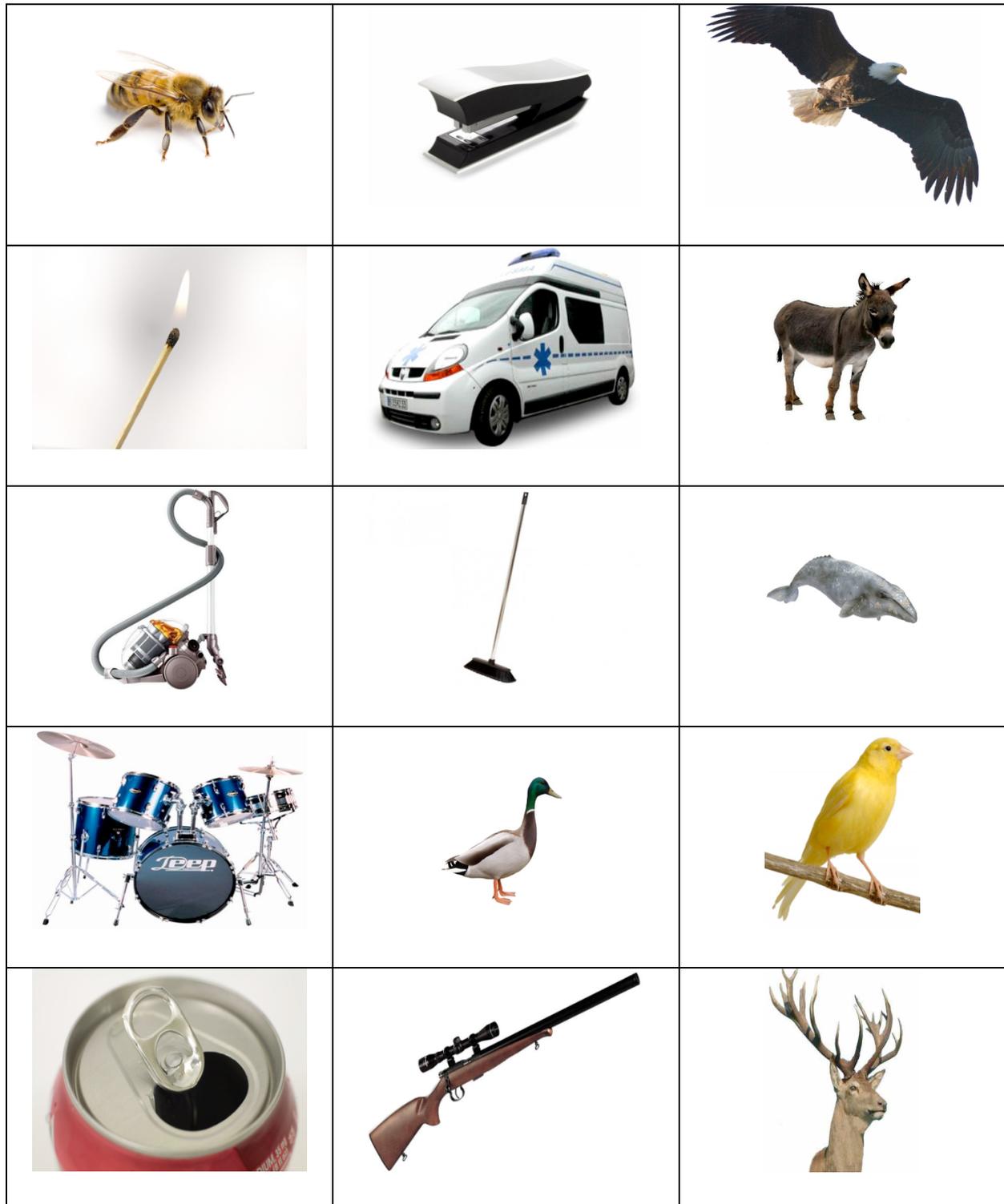
xlim([0 max(vincentized(nbfractile,2),vincentized(nbfractile,1))+100]);

```

Annexes

```
title(strcat('Vincentisation de ',int2str(nbsuj),' sujets'));  
hold off;
```

Annexe 11 – Liste des stimuli visuels utilisés pour les expériences 5 à 7 du chapitre 3



Annexes

Annexe 12 – Liste des stimuli auditifs utilisés pour les expériences 5 à 7 du chapitre 3

Abeille.aiff	Hippopotame.aiff
Agrafeuse.aiff	Imprimante.aiff
Aigle.aiff	Lion.aiff
Allumette.aiff	Loup.aiff
Ambulance.aiff	Machinecrire.aiff Machinelaver.aiff
Ane.aiff	Maracas.aiff
Aspirateur.aiff	Marteau.aiff
Balai.aiff	Marteaupiqueur.aiff
Baleine.aiff	Moto.aiff
Batterie.aiff	Mouche.aiff
Bblanc.aiff	Mouette.aiff
Bblanc2.aiff	Moustique.aiff
Canard.aiff	Mouton.aiff
Canari.aiff	Otarie.aiff
Canette.aiff	Ours.aiff
Carabine.aiff	Panda.aiff
Cerf.aiff	Perceuse.aiff
Chameau.aiff	Perroquet.aiff
Chat.aiff	Piano.aiff
Cheval.aiff	Pigeon.aiff
Chevre.aiff	Police.aiff

Chien.aiff	Poule.aiff
Chimpanze.aiff	Poussin.aiff
Chouette.aiff	Rat.aiff
Ciseau.aiff	Reveil.aiff
Cloche.aiff	Saxophone.aiff
Cochon.aiff	Scie.aiff
Cochondinde.aiff	Sechecheveux.aiff
Coq.aiff	Serpent.aiff
Corbeau.aiff	Sifflet.aiff
Cornemuse.aiff	Silence.aiff
Crocodile.aiff	Silence2.aiff
Dauphin.aiff	Telephone.aiff
Dinde.aiff	Tigre.aiff
Djembe.aiff	Tondeuse.aiff
Elephant.aiff	Train.aiff
Flute.aiff	Triangle.aiff
Gong.aiff	Trompette.aiff
Grenouille.aiff	Tronconneuse.aiff
Grillon.aiff	Vache.aiff
Guitare.aiff	Ventilateur.aiff
Guitarelec.aiff	Violon.aiff
Harmonica.aiff	Xylophone.aiff

Annexes

Harpe.aiff	
Helicoptere.aiff	

B. Annexes – Données brutes**Chapitre 2**

Partie 1

Expérience 1

		% bonnes réponses	Bloc 1	Bloc 2
G1	sujet 1	95	572	530
	sujet 2	98	588	537
	sujet 3	95	522	557
	sujet 4	95	539	483
	sujet 5	98	447	383
	sujet 6	93	523	483
	sujet 7	95	485	459
	sujet 8	93	496	512
	sujet 9	93	558	464
	sujet 10	100	704	640
G2	sujet 11	100	699	596
	sujet 12	100	660	537
	sujet 13	100	625	542
	sujet 14	100	660	597
	sujet 15	100	926	616
	sujet 16	93	440	427
	sujet 17	100	550	488
	sujet 18	100	573	533
	sujet 19	95	634	578
	sujet 20	100	603	480

Expérience 2

G1	% bonnes	Bloc 1	Bloc 2	G2	% bonnes	Bloc 1	Bloc 2
sujet 1	0,92	445	400	sujet 41	0,93	420	406
sujet 2	0,97	335	340	sujet 42	0,98	407	376
sujet 3	0,95	512	480	sujet 43	0,95	527	429
sujet 4	0,98	430	430	sujet 44	0,93	386	428
sujet 5	0,95	402	400	sujet 45	0,98	499	488
sujet 6	1,00	802	605	sujet 46	0,95	382	392
sujet 7	0,98	635	547	sujet 47	0,98	422	387
sujet 8	0,93	634	365	sujet 48	0,95	529	458
sujet 9	0,98	766	605	sujet 49	0,97	421	398
sujet 10	0,98	458	381	sujet 50	0,93	766	590
sujet 11	0,97	321	297	sujet 51	0,97	450	454
sujet 12	1,00	435	421	sujet 52	0,92	421	422
sujet 13	0,98	533	489	sujet 53	0,88	341	317
sujet 14	0,92	390	392	sujet 54	0,98	707	662
sujet 15	0,97	488	466	sujet 55	0,93	470	584
sujet 16	0,93	787	725	sujet 56	0,97	415	416
sujet 17	0,93	614	559	sujet 57	1,00	480	452
sujet 18	0,92	411	458	sujet 58	0,98	499	433
sujet 19	0,95	843	638	sujet 59	0,92	535	497
sujet 20	0,92	405	391	sujet 60	0,93	422	410
sujet 21	1,00	412	385	sujet 61	0,87	378	406
sujet 22	1,00	565	443	sujet 62	0,90	364	379
sujet 23	0,98	461	417	sujet 63	0,98	429	349
sujet 24	0,93	772	619	sujet 64	0,93	483	410
sujet 25	0,97	399	372	sujet 65	0,95	524	500
sujet 26	1,00	480	413	sujet 66	0,98	619	566
sujet 27	0,95	377	357	sujet 67	1,00	747	691
sujet 28	0,98	431	419	sujet 68	0,95	391	359
sujet 29	0,92	326	324	sujet 69	0,97	525	574
sujet 30	1,00	583	567	sujet 70	0,95	402	386
sujet 31	1,00	700	553	sujet 71	0,97	406	411
sujet 32	0,98	391	335	sujet 72	0,93	703	611
sujet 33	1,00	447	426	sujet 73	1,00	672	581
sujet 34	0,95	455	438	sujet 74	0,98	530	503
sujet 35	1,00	375	415	sujet 75	1,00	422	410
sujet 36	0,95	374	348	sujet 76	0,97	385	378
sujet 37	0,98	482	401	sujet 77	0,97	557	508
sujet 38	0,98	472	468	sujet 78	1,00	555	512
sujet 39	0,92	454	367				

Expérience 3

groupe	V	A	AV	TxV	TxA	TxAV
SYST						
sujet 1	345	391	300	1,00	0,98	1,00
sujet 2	393	378	321	0,97	0,91	0,98
sujet 3	443	444	366	0,97	0,98	0,95
sujet 4	347	283	267	0,92	0,89	0,98
sujet 5	447	361	331	1,00	1,00	1,00
sujet 6	332	295	261	1,00	0,92	1,00
sujet 7	346	386	293	1,00	0,97	1,00
sujet 8	319	319	282	1,00	0,97	1,00
sujet 9	442	456	372	1,00	1,00	1,00
sujet 10	390	357	311	0,95	0,86	0,97
sujet 11	557	558	491	1,00	0,97	0,94
sujet 12	339	319	279	0,98	0,97	1,00
NON SYST						
sujet 13	379	375	336	0,89	0,78	0,81
sujet 14	339	327	297	0,98	0,98	0,97
sujet 15	341	326	300	1,00	0,94	1,00
sujet 16	341	286	275	1,00	0,97	0,97
sujet 17	413	483	347	0,97	0,95	0,98
sujet 18	362	300	292	0,98	0,98	0,98
sujet 19	363	381	314	1,00	1,00	0,98
sujet 20	434	453	340	0,92	0,89	0,92
sujet 21	353	329	297	0,86	0,80	0,86
sujet 22	448	467	376	1,00	1,00	0,98
sujet 23	442	489	393	1,00	1,00	1,00
sujet 24	386	430	315	0,97	1,00	1,00

Partie 2

Expérience 1

groupe	V	A	AV	%V	%A	%AV
Jeunes	484	510	440	92,5	98,75	98,75
Jeunes	435	438	418	98,75	98,75	97,5
Jeunes	642	730	616	96,25	100	97,5
Jeunes	594	632	520	95	96,25	97,5
Jeunes	554	540	510	95	98,75	93,75
Jeunes	485	547	452	97,5	95	95
Jeunes	499	546	484	98,75	100	98,75
Jeunes	519	595	494	97,5	97,5	95
Jeunes	548	555	492	97,5	100	98,75
Jeunes	509	538	477	97,5	100	98,75
Jeunes	491	492	472	100	97,5	97,5
Jeunes	546	581	510	95	100	97,5
Agés	518	626	516	100	98,75	97,5
Agés	669	639	622	81,25	87,5	87,5
Agés	708	726	642	93,75	93,75	90
Agés	755	830	682	90	97,5	98,75
Agés	632	652	552	91,25	88,75	90,00
Agés	776	705	613	86,25	90,00	96,25
Agés	821	767	710	100	98,75	100
Agés	718	679	607	81,25	96,25	90
Agés	524	607	469	96,25	97,5	93,75
Agés	674	688	634	96,25	97,5	98,75
Agés	638	648	610	100	98,75	100
Agés	626	627	589	96,25	100	96,25

Expérience 2

groupe	V	A	AV	%V	%A	%AV
Jeunes	379	375	336	379	375	336
Jeunes	339	327	297	339	327	297
Jeunes	341	326	300	341	326	300
Jeunes	341	286	275	341	286	275
Jeunes	413	483	347	413	483	347
Jeunes	362	300	292	362	300	292
Jeunes	363	381	314	363	381	314
Jeunes	434	453	340	434	453	340
Jeunes	353	329	297	353	329	297
Jeunes	448	467	376	448	467	376
Jeunes	442	489	393	442	489	393
Jeunes	386	430	315	386	430	315
Agés	518	626	516	100	98,75	97,5
Agés	669	639	622	81,25	87,5	87,5
Agés	708	726	642	93,75	93,75	90
Agés	755	830	682	90	97,5	98,75
Agés	632	652	552	91,25	88,75	90,00
Agés	776	705	613	86,25	90,00	96,25
Agés	821	767	710	100	98,75	100
Agés	718	679	607	81,25	96,25	90
Agés	524	607	469	96,25	97,5	93,75
Agés	674	688	634	96,25	97,5	98,75
Agés	638	648	610	100	98,75	100
Agés	626	627	589	96,25	100	96,25

Expérience 3

détection sémantique						
	V	A	AV	TxV	TxA	TxAV
sujet 1	437	419	381	0,98	0,92	0,98
sujet 2	386	348	318	0,98	0,98	1,00
sujet 3	377	486	350	1,00	0,95	1,00
sujet 4	410	451	352	1,00	0,95	1,00
sujet 5	406	389	339	1,00	1,00	1,00
sujet 6	385	410	339	1,00	1,00	0,98
sujet 7	483	577	450	0,98	0,97	1,00
sujet 8	435	531	403	1,00	0,94	1,00
sujet 9	445	463	390	0,98	0,95	1,00
sujet 10	389	434	350	1,00	1,00	1,00
sujet 11	411	483	379	1,00	0,97	1,00
sujet 12	357	361	316	1,00	1,00	0,98

Expérience 4

catégorisation latéralisée						
	V	A	AV	TxV	TxA	TxAV
sujet 1	412	453	399	0,89	0,98	0,95
sujet 2	433	507	452	0,95	1,00	0,98
sujet 3	434	435	410	0,95	0,97	1,00
sujet 4	459	483	437	0,95	0,97	0,97
sujet 5	711	725	638	0,91	0,98	0,98
sujet 6	543	696	514	0,92	0,92	0,95
sujet 7	508	595	510	0,92	1,00	0,94
sujet 8	528	523	466	0,95	0,98	0,97
sujet 9	528	605	475	0,95	0,95	0,95
sujet 10	528	553	515	1,00	0,97	0,98
sujet 11	434	447	421	0,98	0,98	0,97
sujet 12	707	700	658	1,00	0,97	0,98

Chapitre 3

Expérience 1

	nouveau		ancien				
	vivant	non vivant	interf	sans	inter	sans	
Groupe AA	sujet 1	859	1151	778	968	1212	1127
	sujet 2	751	814	864	658	768	725
	sujet 3	642	717	622	579	601	680
	sujet 4	619	593	614	615	571	638
	sujet 5	663	812	674	719	908	882
	sujet 6	748	910	740	770	713	907
	sujet 7	761	822	836	745	864	765
	sujet 8	754	857	870	873	881	754
	sujet 9	578	640	610	595	598	653
	sujet 10	633	719	662	641	715	630
	sujet 11	695	749	650	733	703	748
	sujet 12	700	858	796	706	848	845
	sujet 13	621	695	669	574	725	698
	sujet 14	684	824	732	689	854	761
	sujet 15	468	524	479	535	552	517
	sujet 16	648	669	651	708	703	674
	sujet 17	528	655	543	555	676	659
	sujet 18	696	691	687	784	737	709
	sujet 19	761	829	650	808	875	845
	sujet 20	668	899	612	643	873	883
Groupe VV	sujet 21	668	671	631	733	650	776
	sujet 22	585	747	635	680	638	732
	sujet 23	1109	1163	1056	1014	1288	1413
	sujet 24	771	835	679	644	766	813
	sujet 25	662	657	610	686	631	682
	sujet 26	556	667	573	553	601	628
	sujet 27	610	638	525	533	679	610
	sujet 28	544	656	488	486	533	594
	sujet 29	698	791	771	694	825	796
	sujet 30	792	797	703	765	707	757
	sujet 31	640	655	632	597	674	644
	sujet 32	581	701	552	571	657	701
	sujet 33	582	764	640	588	757	710
	sujet 34	636	779	631	680	804	715
	sujet 35	565	652	552	618	602	701
	sujet 36	683	796	667	793	688	726
	sujet 37	663	762	691	710	888	708
	sujet 38	656	649	599	694	624	684
	sujet 39	566	711	541	629	659	696
	sujet 40	622	649	591	598	653	615

		Nouveau	
		vivant	non vivant
	sujet 41	748	813
	sujet 42	566	615
	sujet 43	831	783
	sujet 44	526	617
	sujet 45	557	637
	sujet 46	654	750
Groupe AV	sujet 47	645	661
	sujet 48	587	706
	sujet 49	812	964
	sujet 50	609	704
	sujet 51	558	654
	sujet 52	573	584
	sujet 53	652	776
	sujet 54	707	724
	sujet 55	539	613
	sujet 56	686	721
	sujet 57	587	795
	sujet 58	660	718
	sujet 59	612	679
	sujet 60	782	980

	Ancien				Non vivant			
	AVi	vivant AVs	VAi	VAs	AVi	AVs	VAi	VAs
721	675	739	675	689	813	875	805	
573	514	502	514	612	717	565	675	
844	609	695	609	748	801	1041	784	
548	522	461	522	544	700	696	596	
584	569	541	569	607	696	740	597	
691	654	654	654	576	747	796	867	
670	596	569	596	574	581	553	573	
619	575	643	575	637	751	642	746	
761	836	891	836	1013	898	987	839	
618	616	633	616	620	665	701	610	
524	541	629	431	684	617	682	707	
528	560	583	513	598	625	625	618	
635	631	665	657	921	650	763	825	
604	605	705	748	725	683	845	906	
503	565	518	493	589	616	593	645	
563	673	591	673	741	635	813	642	
603	713	552	713	879	697	829	738	
602	649	589	649	759	724	775	772	
589	652	498	652	704	639	722	612	
954	772	684	772	885	1236	1151	895	

Expérience 2

	Temps de réponse				% réponses correctes			
	ancien		nouveau		ancien		nouveau	
	vivant	non	vivant	non	vivant	non	vivant	non
	487	545	517	558	1,00	1,00	1,00	0,80
sujet 1	487	545	517	558	1,00	1,00	1,00	0,80
sujet 2	566	743	619	623	0,90	1,00	1,00	1,00
sujet 3	545	664	514	676	1,00	1,00	0,90	0,90
sujet 4	576	634	670	624	1,00	1,00	0,90	0,90
sujet 5	533	592	530	574	0,90	1,00	0,90	1,00
sujet 6	541	576	538	601	0,90	0,90	0,90	1,00
sujet 7	636	725	711	701	1,00	0,90	0,90	1,00
sujet 8	468	502	523	525	0,80	0,90	1,00	0,90
sujet 9	605	738	633	672	1,00	1,00	1,00	1,00
sujet 10	539	578	588	606	1,00	0,90	1,00	1,00
sujet 11	581	633	667	700	1,00	0,90	1,00	0,90
VV sujet 12	591	590	607	607	0,90	1,00	1,00	1,00
sujet 13	618	596	657	684	0,90	1,00	1,00	0,90
sujet 14	662	622	748	721	1,00	0,80	1,00	1,00
sujet 15	688	699	600	680	1,00	1,00	1,00	0,90
sujet 16	481	573	493	675	0,90	0,80	1,00	0,70
sujet 17	523	567	512	638	1,00	0,80	1,00	0,70
sujet 18	645	644	635	691	0,90	1,00	1,00	1,00
sujet 19	497	675	540	621	1,00	1,00	0,90	0,80
sujet 20	576	624	581	635	1,00	0,90	1,00	1,00
sujet 21	576	611	548	594	1,00	1,00	1,00	1,00
sujet 22	507	571	572	560	1,00	1,00	1,00	0,90
sujet 23	585	634	551	701	0,90	0,80	1,00	1,00
sujet 24	404	520	419	489	1,00	1,00	0,90	1,00
sujet 25	499	538	467	529	1,00	1,00	1,00	0,90
sujet 26	641	655	653	637	0,90	1,00	0,90	1,00
AV sujet 27	539	646	527	623	0,90	1,00	1,00	1,00
sujet 28	496	481	487	510	1,00	0,80	0,90	0,80
sujet 29	553	521	576	619	0,90	1,00	0,80	0,80
sujet 30	547	622	583	620	0,90	1,00	1,00	1,00
sujet 31	556	633	673	667	1,00	1,00	0,90	0,80
sujet 32	554	639	612	666	0,80	0,90	1,00	1,00
sujet 33	526	580	502	555	0,80	0,90	1,00	1,00
sujet 34	605	600	623	589	1,00	1,00	0,90	0,90
sujet 35	446	565	454	541	0,90	1,00	0,90	0,90
sujet 36	544	553	552	592	1,00	1,00	1,00	1,00

Expérience 3

	Temps de réponse					
	non vivant			vivant		
	nouveaux	ancien VV	ancien AV	nouveaux	ancien VV	ancien AV
sujet 1	693	657	624	580	632	596
sujet 2	617	532	529	474	512	474
sujet 3	870	807	792	800	774	732
sujet 4	621	707	604	611	696	577
sujet 5	703	678	621	660	645	647
sujet 6	693	668	634	610	589	628
sujet 7	764	698	661	708	698	716
sujet 8	771	782	626	668	687	579
sujet 9	652	575	569	549	494	517
sujet 10	525	455	530	514	508	445
sujet 11	602	493	541	519	552	554
sujet 12	583	576	559	579	545	564
sujet 13	769	856	794	761	745	846
sujet 14	616	615	628	543	547	584
sujet 15	700	747	801	658	711	601
sujet 16	767	821	849	679	706	821
sujet 17	770	831	774	650	682	769
sujet 18	664	731	776	668	596	635
sujet 19	601	679	530	505	494	517
sujet 20	643	684	669	714	709	641
sujet 21	587	598	569	533	546	602
sujet 22	650	821	629	678	712	664
sujet 23	604	709	626	660	661	799
sujet 24	570	642	569	534	529	570
sujet 25	626	622	701	597	596	534
sujet 26	596	481	517	544	562	523
sujet 27	783	840	920	831	746	803
sujet 28	652	669	655	604	581	611
sujet 29	712	694	777	674	701	733

% de réponses correctes					
nouveaux	non vivant		nouveaux	vivant	
	ancien VV	ancien AV		ancien VV	ancien AV
1,00	1,00	1,00	1,00	1,00	1,00
0,80	0,60	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
0,90	1,00	1,00	1,00	1,00	1,00
0,90	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	0,90	1,00	0,80
0,90	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	0,80	1,00	0,90	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	0,90	1,00	1,00
0,80	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
0,90	1,00	1,00	1,00	0,80	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	0,90	1,00	1,00
1,00	1,00	1,00	0,90	0,80	1,00
0,80	1,00	1,00	0,90	1,00	1,00
1,00	1,00	0,80	0,90	1,00	1,00
1,00	1,00	1,00	0,80	1,00	1,00
0,90	1,00	1,00	0,90	1,00	1,00
0,80	1,00	1,00	0,90	1,00	1,00
0,90	1,00	1,00	0,90	1,00	1,00
1,00	1,00	1,00	0,70	1,00	1,00
0,90	1,00	1,00	0,90	1,00	1,00
1,00	1,00	0,80	0,90	1,00	1,00
1,00	1,00	1,00	0,90	1,00	1,00
1,00	1,00	1,00	0,90	1,00	1,00

Expérience 4

	Temps de réponse					
	vivant			non vivant		
	ancien VV	ancien AV	nouveau	ancien VV	ancien AV	nouveau
sujet 1	497	547	538	535	496	541
sujet 2	623	566	585	565	580	630
sujet 3	526	564	441	479	482	542
sujet 4	457	508	455	470	467	512
sujet 5	520	530	568	494	515	524
sujet 6	441	466	417	434	452	450
sujet 7	528	606	460	524	488	519
sujet 8	529	585	692	622	560	580
sujet 9	470	456	430	415	416	427
sujet 10	461	432	441	515	487	476
sujet 11	523	507	559	545	543	541
sujet 12	600	616	574	631	631	630
sujet 13	492	487	497	469	432	471
sujet 14	530	531	516	561	515	476
sujet 15	584	555	714	581	576	552
sujet 16	524	483	475	465	478	518
sujet 17	536	493	464	552	559	483
sujet 18	561	520	546	530	547	522
sujet 19	587	440	644	527	465	507
sujet 20	523	497	526	590	520	533
sujet 21	470	458	446	463	421	430
sujet 22	462	450	458	452	438	441
sujet 23	444	437	466	480	472	445
sujet 24	476	473	535	504	476	515
sujet 25	470	482	475	491	494	513
sujet 26	525	527	511	492	484	481
sujet 27	525	541	531	560	529	541
sujet 28	531	463	463	520	538	520
sujet 29	489	521	484	521	511	484
sujet 30	436	516	475	513	456	523

% réponses correctes					
ancien VV	vivant		ancien VV	non vivant	
	ancien AV	nouveau		ancien AV	nouveau
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	0,88	0,88	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	0,88	1,00	1,00	1,00	1,00
1,00	0,88	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	0,88	0,88	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
0,75	0,88	0,88	1,00	1,00	0,88
0,88	1,00	1,00	1,00	1,00	1,00
0,88	0,88	1,00	1,00	1,00	1,00
1,00	1,00	1,00	0,88	1,00	1,00
1,00	0,88	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	0,88
1,00	1,00	0,88	1,00	1,00	1,00
1,00	0,88	1,00	0,88	1,00	1,00
0,88	1,00	1,00	1,00	1,00	0,88
1,00	1,00	0,88	0,75	1,00	1,00
1,00	1,00	0,88	1,00	1,00	1,00
1,00	1,00	1,00	0,88	1,00	0,88
1,00	1,00	1,00	1,00	1,00	1,00
0,88	1,00	1,00	1,00	1,00	1,00
1,00	1,00	0,88	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
0,88	1,00	1,00	1,00	1,00	1,00
1,00	1,00	0,88	1,00	1,00	1,00
1,00	0,88	1,00	0,88	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
0,88	1,00	1,00	1,00	1,00	1,00

Expérience 5

Encodage

	Temps de réponse					
	vivant			non vivant		
	multimodal	unimodal	interférence	multimodal	unimodal	interférence
sujet 1	458	451	472	465	516	433
sujet 2	668	798	730	791	808	714
sujet 3	547	636	598	660	628	590
sujet 4	511	592	447	550	594	564
sujet 5	496	508	509	533	569	538
sujet 6	511	576	560	513	542	457
sujet 7	511	576	497	499	517	508
sujet 8	563	584	534	634	547	530
sujet 9	460	462	452	519	469	455
sujet 10	543	568	584	592	548	538
sujet 11	455	453	482	580	560	498
sujet 12	903	938	751	856	882	854
sujet 13	663	606	686	634	656	624
sujet 14	743	686	725	705	667	606
sujet 15	564	624	538	682	682	627
sujet 16	555	652	563	596	610	536
sujet 17	620	746	735	728	728	659
sujet 18	396	405	393	429	431	445
sujet 19	714	697	617	664	772	712
sujet 20	582	573	570	583	513	503
sujet 21	619	702	785	769	777	727
sujet 22	481	546	507	547	579	526
sujet 23	531	509	465	522	810	601
sujet 24	626	687	678	730	721	726

% réponses correctes						
	vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence	
1,00	1,00	1,00	0,86	1,00	1,00	1,00
0,86	0,86	0,71	1,00	1,00	0,86	0,86
1,00	1,00	0,86	1,00	1,00	1,00	1,00
0,71	0,86	0,71	1,00	0,86	0,57	0,57
1	1	1	1,00	1,00	0,86	0,86
1,00	0,86	1,00	0,86	1,00	1,00	1,00
1,00	1,00	0,86	1,00	1,00	0,86	0,86
1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	0,86	1,00	1,00
1,00	1,00	1,00	0,86	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	0,86	1,00	0,71	0,86	0,86	0,86
1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	0,86	1,00	1,00	1,00	1,00
1,00	1,00	0,86	1,00	0,71	1,00	1,00
0,86	1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00	1,00
0,86	1,00	0,71	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	0,86	1,00	0,86	0,86	1,00	1,00
1,00	1,00	0,86	0,86	0,86	1,00	1,00
1,00	1,00	1,00	1,00	0,86	1,00	1,00

Reconnaissance

	vivant			
	ancien multimodal	Ancien unimodal	ancien interféré	nouveau
sujet 1	1,00	1,00	0,71	0,81
sujet 2	1,00	1,00	0,86	0,90
sujet 3	0,86	0,86	0,71	0,95
sujet 4	0,86	0,86	1,00	0,90
sujet 5	0,43	0,57	0,29	0,86
sujet 6	0,71	0,86	1,00	0,86
sujet 7	0,71	1,00	1,00	0,90
sujet 8	1,00	1,00	1,00	0,67
sujet 9	1,00	1,00	1,00	0,86
sujet 10	1,00	0,57	0,57	0,81
sujet 11	0,86	1,00	1,00	0,95
sujet 12	0,86	0,86	0,71	0,86
sujet 13	1,00	0,57	0,71	0,71
sujet 14	0,86	1,00	0,86	0,62
sujet 15	1,00	0,86	1,00	0,86
sujet 16	1,00	1,00	0,86	0,71
sujet 17	0,71	0,57	0,57	0,95
sujet 18	1,00	1,00	1,00	0,90
sujet 19	1,00	0,71	1,00	0,81
sujet 20	0,86	0,71	0,71	0,90
sujet 21	0,71	0,86	0,86	0,95
sujet 22	0,86	0,57	0,86	0,90
sujet 23	1,00	1,00	0,71	0,81
sujet 24	1,00	1,00	1,00	0,86

	non vivant			
	ancien multimodal	ancien unimodal	ancien interféré	nouveau
sujet 1	1,00	0,71	0,71	0,76
sujet 2	0,57	0,71	0,71	0,95
sujet 3	0,71	0,86	0,86	0,95
sujet 4	0,86	0,86	0,86	0,90
sujet 5	0,57	0,29	0,29	1,00
sujet 6	0,86	0,71	0,71	0,76
sujet 7	1,00	0,86	1,00	0,81
sujet 8	1,00	1,00	0,86	0,86
sujet 9	0,71	1,00	0,86	0,76
sujet 10	0,86	0,43	0,43	1,00
sujet 11	0,86	0,71	0,71	0,81
sujet 12	0,86	0,57	0,86	0,95
sujet 13	0,86	1,00	0,57	0,76
sujet 14	0,86	0,71	0,71	0,95

Annexes

sujet 15	1,00	0,86	1,00	0,90
sujet 16	0,71	0,71	0,71	1,00
sujet 17	0,71	0,14	0,71	1,00
sujet 18	1,00	1,00	1,00	0,86
sujet 19	0,86	1,00	0,29	0,86
sujet 20	0,86	0,71	0,43	1,00
sujet 21	0,86	1,00	1,00	0,81
sujet 22	1,00	0,71	0,86	0,95
sujet 23	0,86	1,00	0,86	1,00
sujet 24	1,00	1,00	1,00	0,95

Expérience 6

Encodage

	Temps de réponse					
	vivant			non vivant		
	multimodal	unimodal	interérence	multimodal	unimodal	interérence
sujet 1	681	713	799	869	690	719
sujet 2	512	487	607	514	589	525
sujet 3	433	461	427	461	501	504
sujet 4	785	672	625	658	629	797
sujet 5	557	622	649	579	542	524
sujet 6	524	530	499	625	621	514
sujet 7	536	616	528	549	597	579
sujet 8	560	489	460	525	545	568
sujet 9	547	553	505	538	601	558
sujet 10	520	494	477	521	509	502
sujet 11	672	629	603	689	559	643
sujet 12	551	668	552	602	601	609
sujet 13	579	583	546	621	576	566
sujet 14	718	818	749	798	726	723
sujet 15	542	630	605	680	598	589
sujet 16	474	549	566	525	556	615
sujet 17	521	572	590	617	561	532
sujet 18	524	571	582	555	560	584
sujet 19	797	791	763	811	784	853
sujet 20	437	489	442	503	449	457
sujet 21	422	494	488	491	492	434
sujet 22	475	513	506	578	643	597
sujet 23	545	541	587	640	633	631
sujet 24	497	508	518	547	551	520

% réponses correctes

	vivant			non vivant		
multimodal	unimodal	interférence	multimodal	unimodal	interférence	
1,00	1,00	1,00	0,86	0,86	1,00	
1,00	1,00	0,86	0,86	1,00	0,86	
0,86	1,00	1,00	0,86	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
0,86	1,00	1,00	1,00	1,00	0,86	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	0,86	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
0,86	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	0,86	1,00	1,00	1,00	
1,00	1,00	1,00	1,00	1,00	1,00	
1,00	1,00	1,00	0,86	1,00	0,86	
1,00	1,00	1,00	1,00	1,00	1,00	

Rappel

% items rappelés par condition						
	vivant			non vivant		
	multimodal	unimodal	interférence	multimodal	unimodal	interférence
sujet 1	0,43	0,00	0,07	0,14	0,21	0,00
sujet 2	0,14	0,07	0,14	0,07	0,07	0,07
sujet 3	0,14	0,07	0,14	0,21	0,07	0,14
sujet 4	0,29	0,07	0,21	0,14	0,21	0,14
sujet 5	0,21	0,14	0,00	0,07	0,07	0,21
sujet 6	0,21	0,07	0,21	0,07	0,14	0,29
sujet 7	0,14	0,00	0,21	0,14	0,00	0,00
sujet 8	0,07	0,21	0,21	0,21	0,07	0,00
sujet 9	0,07	0,07	0,21	0,29	0,00	0,00
sujet 10	0,21	0,14	0,14	0,07	0,14	0,07
sujet 11	0,29	0,14	0,00	0,29	0,14	0,07
sujet 12	0,14	0,29	0,14	0,14	0,00	0,21
sujet 13	0,14	0,07	0,21	0,14	0,07	0,07
sujet 14	0,21	0,14	0,07	0,14	0,07	0,07
sujet 15	0,07	0,36	0,14	0,14	0,14	0,29
sujet 16	0,14	0,14	0,14	0,07	0,14	0,07
sujet 17	0,14	0,29	0,07	0,21	0,00	0,07
sujet 18	0,07	0,14	0,21	0,21	0,14	0,07
sujet 19	0,07	0,29	0,07	0,21	0,00	0,00
sujet 20	0,14	0,14	0,07	0,14	0,14	0,00
sujet 21	0,07	0,14	0,14	0,07	0,14	0,00
sujet 22	0,36	0,07	0,14	0,21	0,07	0,21
sujet 23	0,21	0,14	0,00	0,14	0,14	0,14
sujet 24	0,29	0,14	0,21	0,07	0,14	0,21

Annexes

Expérience 7

Encodage

	Temps de réponse					
	vivant			non vivant		
	multimodal	unimodal	interférence	multimodal	unimodal	interférence
sujet 1	478	527	487	505	551	515
sujet 2	512	570	497	541	543	570
sujet 3	713	652	562	707	585	530
sujet 4	752	759	645	800	727	666
sujet 5	546	528	486	546	509	492
sujet 6	496	557	539	500	540	502
sujet 7	513	609	509	507	486	492
sujet 8	495	479	530	640	596	601
sujet 9	585	668	555	671	654	447
sujet 10	493	551	556	546	553	514
sujet 11	604	585	591	676	720	591
sujet 12	492	549	617	545	561	524
sujet 13	624	642	607	631	616	641
sujet 14	528	564	573	617	631	658
sujet 15	712	679	656	792	838	735
sujet 16	563	593	551	626	630	671
sujet 17	511	604	532	505	619	569
sujet 18	659	668	722	554	848	506
sujet 19	510	566	547	568	574	541
sujet 20	522	467	478	482	448	476
sujet 21	731	788	608	702	657	715
sujet 22	444	437	447	527	508	425
sujet 23	575	518	537	644	651	589
sujet 24	511	515	503	494	574	520

% réponses correctes					
multimodal	vivant			non vivant	
	unimodal	interférence	multimodal	unimodal	interférence
1,00	1,00	1,00	1,00	1,00	1,00
0,86	0,86	0,86	0,71	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
0,71	1,00	1,00	1,00	0,71	1,00
1,00	0,86	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	0,86	0,86	1,00	1,00
1,00	1,00	1,00	0,86	1,00	0,71
1,00	1,00	1,00	0,86	1,00	1,00
1,00	1,00	1,00	0,86	0,86	0,57
1,00	1,00	1,00	0,86	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	0,86	1,00
1,00	0,86	1,00	1,00	0,71	1,00
0,86	1,00	1,00	0,71	1,00	1,00
1,00	0,86	1,00	1,00	1,00	1,00
0,86	1,00	0,86	0,86	0,86	1,00
1,00	1,00	1,00	1,00	1,00	0,86
1,00	1,00	1,00	1,00	1,00	1,00
0,86	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00	1,00

Catégorisation

	Temps de réponse							
	vivant				non vivant			
	ancien		interférence	nouveau	ancien		interférence	nouveau
multimodal	unimodal	multimodal			unimodal			
sujet 1	452	469	462	484	512	512	512	527
sujet 2	479	479	486	475	480	572	587	529
sujet 3	541	532	523	576	599	599	556	563
sujet 4	500	493	511	472	515	518	568	526
sujet 5	413	410	440	432	409	435	418	467
sujet 6	502	463	508	479	519	559	525	542
sujet 7	502	526	523	475	570	538	505	542
sujet 8	418	391	403	385	409	392	384	433
sujet 9	502	514	506	527	507	523	527	524
sujet 10	446	477	439	458	478	486	491	519
sujet 11	463	464	492	492	497	478	478	499
sujet 12	427	470	460	465	469	500	511	490
sujet 13	509	510	515	507	501	516	567	526
sujet 14	458	435	452	447	514	547	538	525
sujet 15	403	543	457	496	503	468	451	480
sujet 16	518	487	537	537	509	538	550	515
sujet 17	553	549	511	551	528	564	543	552
sujet 18	478	483	463	515	502	487	471	482
sujet 19	484	467	487	469	449	452	495	466
sujet 20	470	480	516	490	475	503	461	506
sujet 21	504	497	507	537	493	528	488	542
sujet 22	475	491	509	520	512	513	539	524
sujet 23	600	532	546	542	523	598	579	576
sujet 24	568	552	521	582	582	589	651	642

	% réponses correctes							
	vivant				non vivant			
	ancien			nouveau	ancien			nouveau
	multimodal	unimodal	interférence		multimodal	unimodal	interférence	
sujet 1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
sujet 2	1,00	1,00	0,86	1,00	1,00	1,00	1,00	0,95
sujet 3	0,86	1,00	1,00	0,95	1,00	1,00	1,00	1,00
sujet 4	1,00	0,86	0,86	0,95	0,71	1,00	0,71	1,00
sujet 5	0,71	0,86	0,86	0,76	0,86	1,00	1,00	0,86
sujet 6	1,00	0,86	1,00	0,95	1,00	1,00	1,00	0,95
sujet 7	1,00	0,86	1,00	1,00	1,00	1,00	1,00	1,00
sujet 8	0,71	1,00	1,00	0,95	0,86	1,00	1,00	0,86
sujet 9	1,00	1,00	1,00	0,95	1,00	1,00	0,86	1,00
sujet 10	0,71	1,00	0,86	1,00	0,86	1,00	0,86	1,00
sujet 11	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,95
sujet 12	1,00	1,00	1,00	0,95	1,00	1,00	1,00	1,00
sujet 13	0,86	1,00	1,00	0,95	1,00	1,00	1,00	0,95
sujet 14	0,86	0,57	0,86	0,86	1,00	0,57	0,86	0,90
sujet 15	0,86	1,00	1,00	1,00	1,00	0,86	1,00	0,86
sujet 16	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
sujet 17	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
sujet 18	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
sujet 19	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
sujet 20	1,00	0,86	0,86	1,00	1,00	1,00	0,71	1,00
sujet 21	1,00	1,00	0,71	0,95	1,00	1,00	1,00	1,00
sujet 22	1,00	1,00	1,00	1,00	1,00	1,00	0,71	0,95
sujet 23	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
sujet 24	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,95