

THÈSE DE DOCTORAT DE L'UNIVERSITÉ LUMIÈRE LYON 2

ECOLE DOCTORALE DE SCIENCES COGNITIVES

Présentée par Julien Besle

Pour obtenir le grade de Docteur de l'Université Lyon 2

Spécialité : Sciences Cognitives - Mention : Neurosciences

Interactions audiovisuelles dans le cortex auditif chez l'homme

Approches électrophysiologique et comportementale

Soutenance publique le 22 mai 2007 devant le jury composé de :

M^r Pascal Barone (Examineur)

M^{me} Nicole Bruneau (Rapporteur)

M^r Jean-Luc Schwartz (Rapporteur)

M^{me} Marie-Hélène Steiner-Giard (Directrice de thèse)

M^r Rémy Versace (Examineur)

Table des matières

I	Revue de la littérature	3
1	Convergence audiovisuelle en neurophysiologie	5
1.1	Aires associatives corticales	5
1.1.1	Études électrocorticographique (ECoG) de la convergence multisensorielle	5
1.1.2	Convergence audiovisuelle au niveau du neurone unitaire	8
1.1.3	Aires de convergence dans le cortex frontal	9
1.1.4	Effet de l'anesthésie sur les interactions multisensorielles	9
1.2	Convergence audiovisuelle dans le cortex visuel	10
1.3	Convergence corticale chez l'homme	11
1.4	Convergence sous-corticale	12
1.4.1	Colliculus Supérieur / Tectum optique	13
1.4.2	Autres structures sous-corticales	16
1.5	Études anatomiques de la convergence multisensorielle	17
1.6	Conclusion	19
2	Interactions Audiovisuelles en psychologie	21
2.1	Effets intersensoriels sur les capacités perceptives	22
2.1.1	Effets dynamogéniques	22
2.1.2	Modèles explicatifs de l'effet dynamogénique	22
2.1.3	Effet dynamogénique et théorie de la détection du signal	24
2.1.4	Modèles de détection d'un stimulus bimodal au seuil	24
2.2	Correspondance des dimensions synesthésiques	25
2.2.1	Établissement des dimensions synesthésiques	26
2.2.2	Réalité des correspondances synesthésiques	27
2.2.3	Correspondance des intensités	29
2.2.4	Résumé	30
2.3	Temps de réaction audiovisuels	31
2.3.1	Premières études	31
2.3.2	Paradigme du stimulus accessoire	33
2.3.3	Paradigme d'attention partagée	36
2.4	Conflit des indices spatiaux auditifs et visuels	42
2.4.1	Ventriloquie	43
2.4.2	Facteurs influençant l'effet de ventriloquie	45
2.4.3	Niveau des interactions dans l'effet de la ventriloquie	46

2.5	Conflit des indices temporels	47
2.6	Conclusion	48
3	Perception audiovisuelle de la parole	49
3.1	Contribution visuelle à l'intelligibilité	49
3.1.1	Complémentarité des informations auditives et visuelles de parole	50
3.1.2	Redondance des informations auditives et visuelles de parole	51
3.1.3	Facteurs liés à la connaissance de la langue	51
3.2	Effet McGurk	52
3.2.1	L'hypothèse VPAM	53
3.2.2	Intégration audiovisuelle pré-phonologique	54
3.2.3	Influence des facteurs linguistiques et cognitifs	55
3.3	Facteurs spatiaux et temporels	56
3.4	Modèles de perception de la parole audiovisuelle	58
3.4.1	Modèles post-catégoriels	58
3.4.2	Modèles pré-catégoriels	60
3.5	Conclusion	61
4	Intégration AV en neurosciences cognitives	63
4.1	Comportements d'orientation	63
4.1.1	Orientation vers un stimulus audiovisuel chez l'animal	64
4.1.2	Saccades oculaires vers un stimulus audiovisuel, chez l'homme	65
4.1.3	Expériences chez l'animal alerte et actif	66
4.2	Effet du stimulus redondant	67
4.2.1	Premières études	67
4.2.2	Tâches de discrimination	67
4.2.3	Tâche de détection	68
4.3	Perception des émotions	69
4.4	Objets écologiques audiovisuels	70
4.5	Conditions limites de l'intégration AV	71
4.6	Illusions audiovisuelles	72
4.6.1	Intégration audiovisuelle pré-attentive	72
4.6.2	Application du modèle additif	73
4.6.3	Activités corrélées à une illusion audiovisuelle	74
4.7	Perception audiovisuelle de la parole	74
4.8	Conclusion	77
5	Problématique générale	79
II	Méthodes	81
6	Approches électrophysiologiques	83
6.1	Bases physiologiques des mesures (s)EEG/MEG	83
6.2	ElectroEncéphaloGraphie (EEG)	84

6.2.1	Enregistrement	84
6.2.2	Analyse des potentiels évoqués (PE)	86
6.3	MagnétoEncéphaloGraphie (MEG)	90
6.3.1	Champs magnétiques cérébraux	90
6.3.2	Procédure d'enregistrement	91
6.4	StéréoElectroEncéphaloGraphie (sEEG)	92
6.4.1	Localisation des électrodes	92
6.4.2	Procédure d'enregistrement	93
6.4.3	Calcul du PE et rejet d'artéfacts	94
6.4.4	Résolution spatiale et représentation spatiotemporelle	94
6.4.5	Étude de groupe et normalisation anatomique	95
7	Approche méthodologique de l'intégration AV	99
7.1	Falsification de l'inégalité de Miller	99
7.1.1	Bases mathématiques et postulats	99
7.1.2	Application de l'inégalité	102
7.1.3	Biais potentiels	104
7.1.4	Analyse statistique de groupe	105
7.2	Modèle additif	106
7.2.1	Falsification du modèle additif en EEG/MEG	107
7.2.2	Interprétation des violations de l'additivité en EEG/MEG	109
7.2.3	Comparaison avec le critère d'additivité en IRM fonctionnelle	109
8	Méthodes statistiques en (s)EEG/MEG	111
8.1	Tests multiples	111
8.2	Tests Statistiques sur les données individuelles	113
8.2.1	Tests sur les essais élémentaires	113
8.2.2	Test du modèle additif par randomisation pour des données non ap- pariées	114
8.2.3	Remarques	115
 III Interactions audiovisuelles dans la perception de la parole		
9	Étude en EEG et comportement	119
9.1	Rappel de la problématique	119
9.2	Méthodes	120
9.2.1	Sujets	120
9.2.2	Stimuli	120
9.2.3	Procédure	121
9.2.4	Expérience comportementale complémentaire	122
9.2.5	Analyse des résultats	122
9.3	Résultats	123
9.3.1	Résultats comportementaux	123
9.3.2	Résultats électrophysiologiques	123

9.4	Discussion	125
9.4.1	Comportement	125
9.4.2	Résultats électrophysiologiques	127
10	Étude en sEEG	131
10.1	Introduction	131
10.2	Méthodes	134
10.2.1	Patients	134
10.2.2	Stimuli et procédure	134
10.2.3	Calcul des potentiels évoqués	134
10.2.4	Analyses statistiques	135
10.3	Résultats	136
10.3.1	Données comportementales	136
10.3.2	Réponses évoquées auditives	136
10.3.3	Réponses évoquées visuelles	138
10.3.4	Violations du modèle additif	141
10.3.5	Relations entre réponses auditives, visuelles et interactions audiovisuelles	144
10.4	Discussion	145
10.4.1	Activité du cortex auditif en réponse aux indices visuels de parole	146
10.4.2	Interactions audiovisuelles	149
10.4.3	Comparaison avec l'expérience EEG de surface	151
11	Effet d'indigage temporel	153
11.1	Introduction	153
11.2	Expérience comportementale 1	155
11.2.1	Méthodes	156
11.2.2	Résultats	159
11.2.3	Discussion	162
11.3	Expérience comportementale 2	163
11.3.1	Méthodes	164
11.3.2	Résultats	166
11.3.3	Discussion	169
11.4	Discussion générale	170
IV	Interactions audiovisuelles en mémoire sensorielle	173
12	Introduction générale	175
12.1	MMN Auditive	175
12.2	Rappel de la problématique	176
13	Étude comportementale	179
13.1	Introduction	179
13.2	Méthodes	180

13.2.1	Sujets	180
13.2.2	Stimuli	180
13.2.3	Procédure	181
13.2.4	Analyses	182
13.3	Résultats	182
13.4	Discussion	183
14	Additivité des MMNs auditives et visuelles	185
14.1	Introduction	185
14.2	Méthodes	187
14.2.1	Sujets	187
14.2.2	Stimuli	187
14.2.3	Procédure	187
14.2.4	Analyses	188
14.3	Résultats	188
14.4	Discussion	191
15	Représentation auditive d'une régularité AV	195
15.1	Introduction	195
15.2	Méthodes	196
15.2.1	Sujets	196
15.2.2	Stimuli	196
15.2.3	Procédure	197
15.2.4	Analyses	197
15.3	Résultats	198
15.4	Discussion	201
16	MMN à la conjonction audiovisuelle	205
16.1	Introduction	205
16.2	Méthodes	207
16.2.1	Sujets	207
16.2.2	Stimuli	207
16.2.3	Procédure	207
16.2.4	Analyses	208
16.3	Résultats	208
16.4	Expérience comportementale complémentaire	210
16.5	Discussion	211
V	Discussion générale	215
17	Discussion générale	217
17.1	Interactions audiovisuelles précoces dans la perception de la parole	217
17.2	Représentation d'un évènement audiovisuel en mémoire sensorielle auditive	218
17.3	Interactions audiovisuelles dans le cortex auditif	219

A Données individuelles des patients	223
B Articles	239
Bibliographie	287

Quatrième partie

Interactions audiovisuelles en mémoire sensorielle

Chapitre 12

Introduction générale

12.1 MMN Auditive

La négativité de discordance (Mismatch Negativity, MMN) est une onde des potentiels évoqués auditifs, observée en réponse à tout changement sonore dans un environnement de stimulation répétitive. On peut l'observer dans un protocole dit *oddball* : on présente au sujet une suite de sons identiques ("standards") dans lesquels on introduit occasionnellement des sons "déviants" (Näätänen, Gaillard & Mantysalo, 1978). La MMN est observée quelle que soit la nature du trait acoustique déviant par rapport aux standards (la hauteur tonale, la durée, l'intensité, la localisation, etc...), aussi bien lorsque le sujet prête attention aux stimuli que lorsque son attention est dirigée vers une autre tâche ou une autre modalité sensorielle. La détection par le cerveau d'un changement dans l'environnement implique la conservation d'une trace physiologique du stimuli précédents. La MMN reflèterait donc un processus automatique de discordance neuronale entre cette trace mnésique des stimuli passé et l'entrée d'un nouveau stimulus implique. La MMN est en partie générée dans le cortex auditif secondaire (par exemple Kropotov et coll., 2000).

Plusieurs autres interprétations non mnésiques de la MMN ont été exclues, par exemple que la différence de traitement du son standard et du son déviant provienne de la différence physique entre les stimuli et donc de l'activation de populations de neurones partiellement différentes. Cette interprétation peut facilement être rejetée en comparant le potentiel évoqué par le même stimulus dans un contexte où il est standard et dans un contexte où il est déviant : la différence entre ces deux conditions révèle toujours l'existence d'une MMN.

Une autre hypothèse qui n'implique pas l'existence d'une trace mnésique est que la MMN reflèterait la différence de fréquence d'apparition des stimuli standards et déviants. Ainsi, si la population de neurones répondant au stimulus répond d'autant moins que le stimulus est présenté souvent, en raison par exemple de l'existence d'une période réfractaire, la moyenne des réponses au stimulus déviant devrait être différente de la moyenne des réponses au stimulus standard, même si ces sons sont identiques. Cette hypothèse de *refractoriness* peut être rejetée en comparant la réponse au même son, dans le cas où il est déviant parmi des sons standards et dans une condition appelée équiprobable dans laquelle il est présenté, avec la même probabilité, parmi plusieurs stimuli différents ayant la même fréquence de présentation (Schröger & Wolff, 1996). Dans ce cas, on continue à observer

une MMN. Donc le même stimulus, présenté avec la même fréquence d'apparition, mais dans un cas où il brise une régularité (lorsqu'il est présenté dans une suite de standards) et dans un cas où il ne brise aucune régularité (la condition équiprobable) donne lieu à des traitements différents qui ne peuvent être attribués qu'à l'effet de l'organisation des autres stimuli de la séquence, en l'occurrence la répétition des sons standards.

Ainsi contrôlée, l'observation d'une MMN implique donc l'existence d'une représentation mnésique du son standard à laquelle le son déviant est comparé. Cette représentation mnésique est souvent assimilée à la mémoire sensorielle ou échoïque, mise en évidence de façon comportementale dans l'effet de récence lors d'une tâche de rappel ou l'effet de masquage auditif (Hawkins & Presson, 1986) et des tentatives ont été faites de lier la représentation indexée par la MMN et la mémoire échoïque (Cowan, Winkler, Teder & Näätänen, 1993 ; Winkler, Reinikainen & Näätänen, 1993). Il existe toutefois d'autres candidats électrophysiologiques à la corrélation avec la mémoire échoïque. Certains auteurs proposent ainsi que l'existence de périodes réfractaires, en particulier dans le cas de l'onde N1, peut être interprété comme un phénomène mnésique et sous-tendre la mémoire échoïque (Lu, Williamson & Kaufman, 1992b, 1992a ; McEvoy, Levänen & Loveless, 1997). Cette question étant loin d'être tranchée, on utilisera donc le terme de mémoire sensorielle auditive au sens de "ce qui est indexé par la MMN", sans faire d'hypothèse sur une correspondance avec la mémoire sensorielle mise en évidence avec des techniques comportementales.

12.2 Rappel de la problématique

Récemment, l'interprétation de la nature des représentations mnésiques reflétées par la MMN a été révisée par certains auteurs. En effet, de nombreuses études ont montré que la MMN n'est pas générée uniquement lorsque standards et déviants diffèrent sur un ou plusieurs traits acoustiques élémentaires, mais également lors de violations de régularités acoustiques plus complexes impliquant des relations entre plusieurs stimuli auditifs (par exemple : Horvath, Czigler, Sussman & Winkler, 2001 ; Korzyukov, Winkler, Gumenyuk & Alho, 2003 ; Tervaniemi, Maury & Näätänen, 1994) ou plusieurs traits élémentaires d'un même stimulus (violation d'une conjonction de 2 traits, voir la partie 16.1 page 205 ; Paavilainen, Simola, Jaramillo, Näätänen & Winkler, 2001). Ces données, entres autres, ont mené à l'idée que la mémoire sensorielle auditive indexée par la MMN a pour fonction de représenter toute régularité dans un environnement sonore complexe. Le rôle fonctionnel de cette représentation serait de détecter n'importe quelle anomalie de cet environnement sonore pouvant représenter une menace ou intérêt pour l'organisme (Winkler, Karmos & Näätänen, 1996).

Dans l'introduction de cette thèse, nous avons avancé divers résultats neuro-anatomiques, comportementaux et neurophysiologiques suggérant que les informations visuelles pouvaient influencer des traitements spécifiques à la modalité auditive. De même, dans la première partie expérimentale de cette thèse, nous avons montré que la vision pouvait moduler l'activité auditive à des étapes relativement précoces du traitement dans le cortex auditif, dans le cas particulier, il est vrai, de la perception de la parole. Puisque des traitements, censés être purement auditifs, sont en réalité influencés par la vision, cette influence pourrait avoir des répercussions sur la représentation en mémoire sensorielle de

l'environnement sonore et en particulier de ses régularités. La question que nous posons dans cette deuxième partie expérimentale est la suivante : les régularités audiovisuelles sont-elles représentées en mémoire sensorielle auditive ? Autrement dit, si un stimulus auditif est constamment associé à un stimulus visuel, cette composante visuelle va-t-elle être incluse dans la représentation du son en mémoire sensorielle auditive ?

Il existe plusieurs façons d'aborder cette question : notre première approche sera comportementale et exploitera le lien qui existe entre la mémoire sensorielle et la détection de la déviance. Dans les trois expériences suivantes, nous étudierons la question de la représentation d'une régularité audiovisuelle en étudiant diverses influences visuelles possibles sur le marqueur électrophysiologique de la mémoire sensorielle auditive : la MMN.

Chapitre 13

Détection d'une déviance audiovisuelle : étude comportementale

13.1 Introduction

Il a été montré à plusieurs reprises que les performances dans une tâche de détection d'un stimulus déviant présenté parmi des stimuli standards, étaient corrélées aux caractéristiques de la MMN automatiquement évoquée par ces déviants lorsque les sujets n'y prêtent pas attention. Ainsi, Tiitinen, May, Reinikainen et Näätänen (1994) ont montré, d'une part, que la latence de la MMN à une déviance fréquentielle et le temps de détection des mêmes déviants décroissaient avec l'amplitude de la déviance de manière identique et, d'autre part, étaient fortement corrélés. Par ailleurs, Novitski, Tervaniemi, Huotilainen et Näätänen (2004) ont montré que l'amplitude et la latence de la MMN à une déviation fréquentielle étaient corrélées à la fois au temps de détection et au taux de détection de cette déviation, la MMN étant d'autant plus grande que les performances sont bonnes. Ces résultats suggèrent que les performances comportementales dans la détection d'un stimulus déviant présenté parmi des stimuli distracteurs standards sont directement liées aux processus indexés par la MMN (voir Schröger, 1997, pour une revue). Une façon d'étudier si ces processus, qui, on l'a vu, mettent en jeu la mémoire sensorielle auditive, peut être influencé par des informations visuelles, est de comparer les temps de détection d'une déviation auditive et d'une déviation audiovisuelle d'un événement audiovisuel standard. Si le temps de détection d'une déviation audiovisuelle est plus rapide que celui d'une déviation auditive, c'est que la dimension visuelle du stimulus entre en compte dans le processus de comparaison aboutissant à la détection de la déviance.

Deux études ont montré que la déviance occasionnelle d'un stimulus bimodal sur ses deux dimensions auditive et visuelle simultanément, était détectée plus rapidement qu'une déviance uniquement sur sa dimension auditive ou sur sa dimension visuelle (Squires et coll., 1977 ; Teder-Sälejärvi et coll., 2002). Cependant ce résultat pourrait s'expliquer, tout comme l'effet du stimulus redondant, par un phénomène de facilitation statistique dans un modèle d'activations séparées : si le temps de détection du premier processus de détection auditif ou visuel arrivé à son terme détermine le temps de détection d'un essai donné, alors le temps de détection de deux déviations simultanées sera en moyenne inférieur au temps

de détection d'une seule déviance sans que l'on n'ait besoin de postuler d'interactions entre les processus auditifs et visuels de détection de déviance (voir la partie 7.1 page 99).

Pour exclure cette possibilité, il faut tester l'inégalité de Miller sur la distribution des temps de détection des déviances auditive, visuelle et audiovisuelle d'un événement audiovisuel standard. Si cette inégalité est falsifiée et les modèles d'activations séparées rejetés, alors on pourra supposer que les processus auditif et visuel de détection de la déviance ont interagi. Dans la mesure où le processus de détection de la déviance auditive est lié à la comparaison du déviant auditif avec la représentation présente en mémoire sensorielle auditive, ce résultat serait compatible avec la mise en jeu de la dimension visuelle dans cette comparaison.

Mais d'autres explications (non exclusives) sont possibles puisque la diminution du TR pourrait aussi bien refléter une influence des informations auditives dans le processus analogue de détection de la déviance visuelle. Par ailleurs les interactions audiovisuelles pourraient concerner des étapes de traitement en aval de la comparaison à la trace mnésique, comme, par exemple, ceux impliqués dans la détection consciente de la déviance ou dans la réponse motrice.

L'inégalité de Miller a été testée par Schröger et Widmann (1998) dans une telle tâche de détection de stimuli audiovisuels déviants sur leur localisation spatiale, soit dans la dimension visuelle, soit dans la dimension auditive, soit dans les deux dimensions : les temps de détection des déviants audiovisuels étaient significativement plus rapides que ceux prédits par les modèles d'activations séparées. Il semble donc que les processus auditifs et visuels de détection de la déviance interagissent. Nous avons voulu tester l'inégalité de Miller avec d'autres types de stimuli standards et déviants. Ces stimuli sont ceux qui seront utilisés dans les expériences suivantes. Une violation de l'inégalité de Miller permettrait d'établir que les stimuli utilisés sont susceptibles de donner lieu à des interactions audiovisuelles au niveau de la mémoire sensorielle auditive.

13.2 Méthodes

13.2.1 Sujets

Quinze sujets droitiers (dont 8 de sexe féminin) âgés en moyenne de 23,1 ans ont passé cette expérience. Aucun sujet ne souffrait de troubles neurologiques. Ils avaient tous une audition normale et une vision normale ou corrigée.

13.2.2 Stimuli

Les stimuli utilisés étaient inspirés de ceux utilisés par notre équipe dans des expériences précédentes et qui ont permis de mettre en évidence des interactions audiovisuelles précoces (revue dans Fort & Giard, 2004, et dans la partie discrimination PE). Nous avons utilisé 4 types de stimuli audiovisuels A_1V_1 , A_1V_2 , A_2V_1 et A_2V_2 , représentés dans la figure 13.1 page suivante.

Les composantes visuelles de ces stimuli consistaient en une déformation horizontale (V_1) ou verticale (V_2) transitoire d'un cercle jaune sur fond noir, ayant un diamètre de 2°

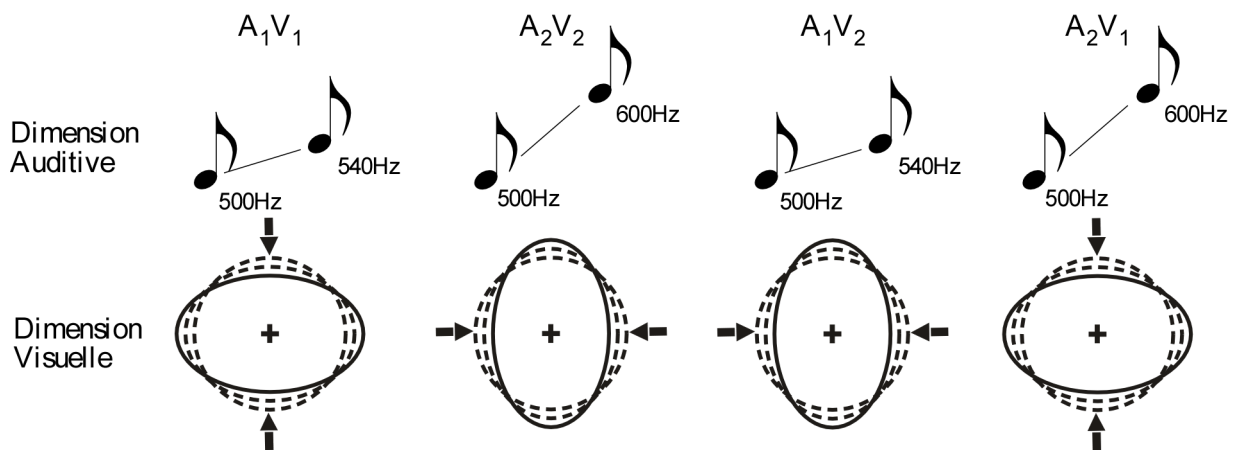


FIG. 13.1 – Stimuli audiovisuels utilisés dans les expériences comportementales et électrophysiologiques sur la mémoire sensorielle. Chaque stimulus était constitué d’une composante auditive A_1 ou A_2 et d’une composante visuelle V_1 ou V_2 .

d’angle visuel. La déformation avait une durée totale de 140 ms incluant le retour du cercle à son état initial. L’amplitude de la déformation du cercle à son maximum représentait 33% du diamètre du cercle de départ.

Les composantes auditives des stimuli consistaient en un son pur enrichi des deux premières harmoniques paires dont la fréquence fondamentale variait linéairement soit de 500Hz à 540Hz (A_1), soit de 500 Hz à 600 Hz (A_2) sur une durée de 140 ms (montée/descente : 14 ms).

La taille des déviations auditives et visuelles a été choisie de façon à ce que, sur un groupe de sujets, le TR pour discriminer le stimulus A_1 du stimulus A_2 soit équivalent au TR pour discriminer le stimulus V_1 du stimulus V_2 . Nous avons choisi d’équilibrer la discriminabilité des composantes auditives et visuelles car plusieurs études ont montré que la diminution du TR en condition audiovisuelle est maximale dans ces conditions (par exemple Squires et coll., 1977).

Dans la moitié des blocs expérimentaux, le stimulus A_1V_1 était présenté avec une probabilité de 76% (standard) et les stimuli A_1V_2 , A_2V_1 et A_2V_2 (respectivement déviants visuel, auditif et audiovisuel) étaient présentés avec une probabilité de 8% chacun. Dans l’autre moitié des blocs, le stimulus A_2V_2 était standard et les stimuli A_2V_1 , A_1V_2 et A_1V_1 étaient déviants. Les stimuli audiovisuels standards seront désormais notés AV, et les déviants auditifs, visuels et audiovisuels $A'V$, AV' et $A'V'$ respectivement.

13.2.3 Procédure

Un bloc de stimuli commençait par la présentation du cercle sur l’écran, qui restait présent pendant toute la durée d’un bloc.

Les stimuli standards et déviants étaient présentés de façon pseudo-aléatoire avec, pour contrainte, qu’un bloc commençait nécessairement par au moins 3 stimuli standards et que deux stimuli déviants étaient séparés par au moins 1 stimulus standard.

La tâche du sujet consistait à fixer la croix de fixation (centre du cercle) et à cliquer le

plus vite possible avec l'index sur le bouton gauche de la souris à chaque apparition d'un stimulus déviant, que la composante déviante soit auditive, visuelle ou audiovisuelle.

Un total de 1000 stimulations (dont 80 déviants de chaque type) a été présenté en 4 blocs expérimentaux d'une durée approximative de 2 minutes 20 secondes chacun. L'intervalle interstimulus était de 560 ms.

Les blocs ayant pour standard les stimuli A1V1 et A2V2 étaient présentés dans un ordre aléatoire et différent pour chacun des sujets.

13.2.4 Analyses

Seuls les TR supérieurs à 150 ms et inférieurs à 1500 ms étaient pris en compte, les autres étant considérés comme des fausses alarmes. Les temps de détection ont été analysés conformément aux méthodes exposées dans la partie 9.2.5 page 122.

13.3 Résultats

Les temps moyens pour détecter les déviants auditifs, visuels et audiovisuels étaient respectivement 446, 429 et 356 ms (écarts-types : 52, 51 et 38 ms). Les taux d'erreurs (cibles manquées) étaient de 6,67%, 23,30%, et 3,33 % dans les conditions auditive, visuelle et audiovisuelle respectivement (écarts-types : 4,67%, 15,01% et 3,36%). La figure 13.2 montre les fonctions de répartition des temps de détection pour les déviants auditifs, visuels et audiovisuels ainsi que la somme des fonctions de répartition auditive et visuelle.

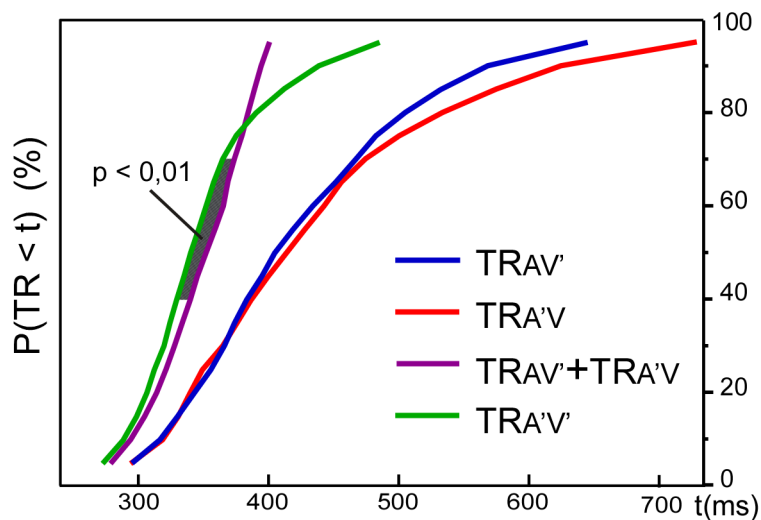


FIG. 13.2 – Application de l'inégalité de Miller. $TR_{AV'}$: fonction de répartition des temps de détection du déviant visuel ; $TR_{A'V}$: fonction de répartition des temps de détection du déviant auditif ; $TR_{AV'} + TR_{A'V}$: somme des 2 fonctions de répartition unimodales ; $TR_{AV'}$: fonction de répartition des temps de de détection du déviant audiovisuel. La partie hachurée désigne les zones où les fractiles correspondants des deux fonctions de répartition sont significativement différents.

Comme on peut le constater, les temps de détection des déviants audiovisuels étaient plus courts que ceux prédits par les modèles d'activations séparées. Cette différence était

significative au seuil $p < 0,05$ pour les fractiles $t(0,2)$ à $t(0,7)$ et significative au seuil $p < 0,01$ pour les fractiles $t(0,4)$ à $t(0,7)$.

13.4 Discussion

Commençons par noter que le taux d'erreur était plus élevé pour les déviants visuels que pour les déviants auditifs ou audiovisuels. Il est probable que les sujets ont sacrifié l'exactitude pour garder des TR équivalents pour les déviants auditif et visuel. Nos résultats confirment ceux de Schröger et Widmann (1998) : la déviance audiovisuelle d'un évènement audiovisuel est détectée plus rapidement qu'une déviance sur une des deux dimensions sensorielles seulement. Le rejet des modèles d'activations séparées a, depuis lors, été répliqué par Teder-Sälejärvi, Di Russo, McDonald et Hillyard (2005) pour des déviances auditives et visuelles sur l'intensité (sonie et brillance), sous l'hypothèse d'indépendance des distributions de TR unimodales, puis par Gondan et coll. (2005) pour des déviants qui consistaient en une répétition du stimulus standard, dans le cas général des modèles d'activations séparées. Ces résultats et les nôtres suggèrent une coactivation entre les processus auditif et visuel de détection de la déviance. Comme nous l'avons souligné dans l'introduction, cela ne garantit pas l'existence d'une dimension visuelle de la représentation de l'évènement audiovisuel en mémoire sensorielle auditive, mais cela montre l'existence d'interactions audiovisuelles dans un processus mettant vraisemblablement en jeu cette mémoire sensorielle.

Dans l'étude de Schröger et Widmann (1998), les potentiels évoqués par les trois déviants et par le stimulus standard avaient été enregistrés. Lorsque l'on calcule la différence entre déviants et standards auditifs dans ce type de protocole, où le sujet a pour tâche de détecter les déviants, on observe en plus de la MMN, des ondes plus tardives telles que la N2b et la P3, vraisemblablement associées au traitement conscient de la déviance. En appliquant le modèle additif, ils ont pu établir que des interactions audiovisuelles prenaient place à partir de 180 ms au niveau de l'onde N2b et de l'onde P3, mais pas au niveau de la MMN. Leur conclusion était donc que ce qui expliquait le gain de temps de réaction était attribuable à une co-activation au niveau des processus conscients de détection de la déviance plutôt qu'à la comparaison automatique des traces en mémoire sensorielle.

Toutefois, la MMN et la N2b sont deux ondes qui se recouvrent partiellement, et il est possible que des interactions audiovisuelles prennent place vers la fin du processus indexé par la MMN et soient superposées à des interactions au niveau de l'onde N2b. Dans l'expérience suivante, nous allons donc appliquer le modèle additif aux différences entre déviants et standards dans une situation où le sujet ignore les stimuli et où les processus étudiés (indexés par la MMN) sont automatiques.

Chapitre 14

Additivité des MMNs auditives et visuelles

Cette étude ayant fait l'objet d'une publication (Besle, Fort & Giard, 2005), elle ne sera que brièvement présentée ici. Les détails en sont décrits dans l'article, intégrée au manuscrit en annexe (page 257).

14.1 Introduction

Pour les raisons exposées dans l'introduction générale de cette partie, nous pensons que la trace en mémoire sensorielle est susceptible d'incorporer des régularités visuelles lorsqu'elles sont associées à des régularités auditives puisque des interactions audiovisuelles ont probablement lieu avant les processus responsables de la construction de la trace (Giard & Peronnet, 1999). Une incorporation de la sorte devrait nécessairement se traduire par une différence entre les MMN générées par une déviance auditive et une double déviance auditive et visuelle d'un événement audiovisuel. En effet plusieurs expériences suggèrent que les MMN générées par des déviants différant d'un son standard sur différents traits acoustiques sont générées dans différentes parties du cortex auditif (Giard et coll., 1995 ; Rosburg, 2003). Par ailleurs, l'amplitude de la MMN générée par une déviance sur le même trait acoustique augmente avec l'amplitude de la déviance (Novitski et coll., 2004 ; Tiitinen et coll., 1994). Si les caractéristiques visuelles d'un événement audiovisuel sont intégrées à la trace en mémoire sensorielle auditive, alors des déviations différentes (auditives et audiovisuelles) d'un événement audiovisuel devraient générer des MMN différentes.

Cependant, on ne peut se contenter de comparer la MMN générée par un déviant auditif et un déviant audiovisuel car on doit tenir compte de l'éventuelle existence de processus de détection automatique de la déviance visuelle. Des études récentes ont mis en évidence une onde analogue à la MMN dans la modalité visuelle (revue dans Pazo-Alvarez et coll., 2003) et ont montré que cette MMN visuelle possède certaines des caractéristiques d'un marqueur des processus de comparaison automatique à une trace mnésique (en mémoire sensorielle visuelle) : indépendance à l'attention (Heslenfeld, 2003), exclusion de l'hypothèse de *refractoriness* (Czigler, Balazs & Winkler, 2002 ; Pazo-Alvarez, Amenedo & Cadaveira,

2004, voir cependant Kenemans, Jong & Verbaten, 2003 pour une autre hypothèse non mnésique). La MMN visuelle semble être générée dans les aires occipitales (Berti & Schröger, 2004), mais certaines études ont décrit une composante additionnelle plus antérieure dans la MMN visuelle (Czigler et coll., 2002 ; Heslenfeld, 2003). Pour comparer la MMN auditive à une déviance audiovisuelle à celle générée par une déviance audiovisuelle, il faudra donc corriger pour l'existence éventuelle de la MMN visuelle, ce qui revient à tester l'additivité des ondes générées par des déviations auditives, visuelles et audiovisuelles d'un même standard audiovisuel.

L'étude de Schröger et Widmann (1998) semble indiquer que la violation d'additivité concerne les processus en aval de ceux indexés par la MMN, mais il n'était pas possible de séparer dans cette étude les violations d'additivité dues à la MMN de celles dues à l'onde N2b. Une autre étude plus ancienne avait testé l'additivité des MMN auditives et visuelles (Nyman et coll., 1990) et n'était pas parvenue à mettre en évidence des processus visuels de détection automatique de la déviance. Les auteurs avaient donc conclu à la spécificité auditive de la MMN auditive. La violation de l'additivité n'avait cependant pas été testée statistiquement dans cette étude.

Par ailleurs, plusieurs études, déjà mentionnées, ont rapporté l'existence d'une MMN auditive évoquée par une déviance visuelle d'un événement audiovisuel, par exemple dans le cas de l'illusion McGurk (Colin et coll., 2004 ; Colin, Radeau, Soquet, Demolin et coll., 2002 ; Möttönen et coll., 2002 ; Sams et coll., 1991), de l'illusion de ventriloquie (Colin, Radeau, Soquet, Dachy & Deltenre, 2002 ; Stekelenburg et coll., 2004), ainsi que dans le cas d'un biais visuel dans la perception d'émotions portées par une voix (de Gelder, Bocker, Tuomainen, Hensen & Vroomen, 1999). Dans tous les cas, sauf le dernier, l'existence d'une illusion irrépressible préservait la possibilité que les informations visuelles aient été converties sous forme auditive et que le processus de comparaison des traces aboutissant à la MMN ait été indépendant de toute interaction audiovisuelle. Quoiqu'il en soit, aucune de ces études n'a envisagé que la MMN enregistrée dans ces conditions, c'est-à-dire la différence entre les réponses à l'événement audiovisuel standard et à l'événement audiovisuel déviant sur sa composante visuelle, pouvait refléter en réalité un processus visuel de comparaison du stimulus à une trace (indexant une supposée mémoire sensorielle visuelle) ou tout autre processus visuel automatique dû à la présence d'une déviance visuelle.

Tester l'additivité des MMN auditive et visuelle peut donc permettre de répondre à plusieurs questions : Observe-t-on une MMN à une déviance visuelle d'un événement audiovisuel ? Et si oui, cette MMN reflète-t-elle une influence visuelle sur un processus auditif, même en l'absence d'une illusion audiovisuelle, ou un processus visuel automatique de détection de la déviance (la MMN visuelle) ? Observe-t-on une modulation de la MMN auditive par la présence d'une déviance visuelle qui pourrait refléter le fait que le processus auditif de comparaison du déviant à une trace en mémoire sensorielle auditive est influencé par les informations visuelles ? Pour répondre à ces différentes questions, il est important d'étudier la topographie de la violation de l'additivité, le cas échéant : une topographie auditive peut signifier soit que la déviance visuelle d'un événement audiovisuel provoque une MMN auditive, comme dans le cas des illusions (McGurk ou ventriloquie), soit que la

MMN auditive évoquée par le déviant audiovisuel a été influencée par la présence d'informations visuelles. En revanche, une topographie visuelle suggérerait que c'est le processus de détection automatique de la déviance visuelle qui est influencé par les informations auditives.

Dans ce dernier cas cependant une ambiguïté peut provenir du fait qu'on connaît mal la topographie de la MMN visuelle et de l'existence possible d'une composante antérieure de la MMN visuelle. Afin de résoudre cette ambiguïté, le cas échéant, nous avons également enregistré la MMN visuelle évoquée par nos stimulations dans une condition visuelle seule.

14.2 Méthodes

14.2.1 Sujets

Les sujets étaient les mêmes que ceux ayant participé à l'expérience précédente. En réalité l'expérience électrophysiologique a été réalisée avant l'expérience comportementale, le même jour.

14.2.2 Stimuli

Les stimuli étaient identiques à ceux utilisés dans l'expérience comportementale.

14.2.3 Procédure

Puisqu'il s'agissait de mesurer dans cette expérience des processus automatiques, il fallait s'assurer que les sujets portent leur attention ailleurs que sur les événements audiovisuels. À cette fin, la tâche du sujet était, dans cette expérience, de répondre le plus rapidement possible lorsque la croix de fixation disparaissait. Cette disparition avait une durée de 120 ms et avait une probabilité d'occurrence de 13%. Elle était cependant désynchronisée par rapport aux événements audiovisuels et ne pouvait se produire que pendant un essai standard (pour éviter de rejeter trop d'essais déviants dans le calcul des potentiels évoqués), et n'avait jamais lieu dans un essai précédent un déviant (pour éviter que les potentiels évoqués par les déviants ne soient contaminés par des processus liés à la réponse, étant donné l'intervalle inter-stimulus relativement faible). Ainsi, le sujet devait regarder l'écran sur lequel étaient présentés les événements audiovisuels, avec son attention dirigée vers une autre tâche. Il avait, de plus, pour consigne d'ignorer le cercle et les sons.

Pour le test de l'additivité des MMNs, un total de 3200 stimulations ont été présentées (dont 8%, c'est-à-dire 256 déviants, de chaque type). Les stimuli étaient mélangés aléatoirement avec des contraintes identiques à l'expérience comportementale et répartis en 12 blocs d'une durée approximative de 2 minutes 30. Dans la moitié de ces blocs, le stimulus A_1V_1 était le standard, dans la deuxième moitié, c'était le stimulus A_2V_2 .

Pour la condition visuelle seule, les séquences de stimuli étaient du même type que dans les conditions audiovisuelles, excepté qu'aucun son n'était présenté : la probabilité d'occurrence d'un déviant visuel (V') était donc de 16% et celle d'un stimulus standard

(V) de 84%. Un total de 1600 stimulations (dont 512 déviants) a été présenté, réparties en 6 blocs. Dans la moitié de ces blocs, le stimulus V1 était le standard.

Les blocs audiovisuels et visuels seuls étaient présentés dans un ordre aléatoire, différent d'un sujet à un autre.

14.2.4 Analyses

Pour le calcul des PE standards moyens, tous les essais ayant inclu une cible, ainsi que les essais suivant immédiatement un déviant ont été exclus. Après rejet des artéfacts d'enregistrement, le nombre moyen d'essais par sujet pour le calcul des PE étaient de 1299, 649 et 204 respectivement pour les standards audiovisuels, les standards visuels et chacun des 4 types de déviants (auditif, visuel, audiovisuel et visuel seul). La ligne de base était prise entre 100 ms et 0 ms avant la stimulation.

Pour le paradigme audiovisuel, les MMN auditive ($MMN_{A'V}$), visuelle ($MMN_{AV'}$) et audiovisuelle ($MMN_{A'V'}$) ont été calculées respectivement comme la différence, point par point, entre les potentiels évoqués par les déviants A'V, AV' et A'V' et le potentiel évoqué par le standard AV. Chaque PE déviant ou standard était donc lui-même une moyenne des potentiels évoqués par deux stimuli différents dans un rôle particulier (le potentiel évoqué déviant audiovisuel était par exemple la moyenne du potentiel évoqué par le stimulus A_1V_1 dans son rôle de déviant et la moyenne du potentiel évoqué par le stimulus A_2V_2 dans son rôle de déviant).

Dans le paradigme visuel seul, la MMN visuelle ($MMN_{V'}$) a été calculée comme la différence entre les potentiels évoqués par le déviant V' et le standard V.

Tous les tests statistiques étaient des tests de Student appariés. Pour éviter le problème des tests multiples, nous n'avons effectué chaque test qu'à un échantillon correspondant au pic maximum de la MMN concernée, sur une valeur moyennée sur une fenêtre de 40 ms autour de la latence de ce pic.

Pour le test de la violation du modèle additif, nous avons arbitrairement choisi la latence du maximum d'amplitude de la $MMN_{A'V}$ car l'objectif premier était de montrer une modulation visuelle de la MMN auditive.

14.3 Résultats

Les TR pour la tâche distractive dans les blocs audiovisuels et visuels étaient de 404 et 409 ms respectivement (écart-types : 51 et 52 ms). Les taux de cibles manquées étaient respectivement de 3,51 et 3,24% (écarts-types : 3,13 et 3,11%). Aucune des deux mesures n'était significativement différente entre les deux conditions.

Les MMN A'V, AV' et A'V' du paradigme audiovisuel sont illustrées dans les figures 14.1 page suivante et 14.2 page 190.

La $MMN_{A'V}$ (courbe rouge sur la figure 14.1) a son pic vers 198 ms et présente la topographie fronto-centrale habituelle (figure 14.2.A) avec inversion de polarité aux mastoïdes, typique des activités générées dans le cortex auditif. Les tests statistiques à la latence du maximum sont très significatifs sur l'ensemble du scalp.

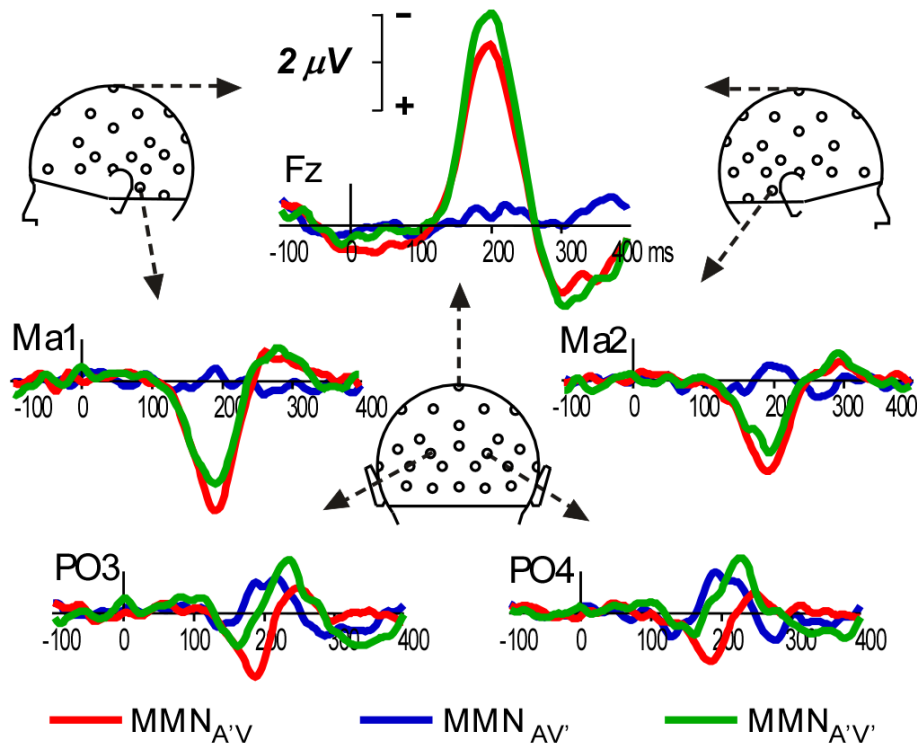


FIG. 14.1 – MMN générées par les déviants A'V, AV' et A'V' dans le paradigme audiovisuel, sur un sous-ensemble d'électrodes.

La $MMN_{AV'}$ présente une topographie bilatérale occipitale (figure 14.2.B), avec deux pics, l'un vers 192 ms et le suivant vers 215 ms (courbe bleue sur la figure 14.1). Sa topographie postérieure suggère qu'elle est générée dans le cortex visuel. On n'a pas observé d'activité plus antérieure ou typique d'activations du cortex auditif. Les tests de Student menés à la latence du premier pic indiquent des potentiels significativement différents de 0 sur un grand nombre d'électrodes occipitales.

La $MMN_{A'V'}$ (courbe verte sur la figure 14.1) ressemble fort à la $MMN_{A'V}$, avec un pic d'amplitude à la même latence (199 ms). Si on regarde cependant plus attentivement les électrodes occipitales PO3 et PO4, on constate qu'elle se rapproche de la $MMN_{AV'}$. Au niveau de la topographie des potentiels (figure 14.2.C), il est très difficile de la distinguer de celle de la $MMN_{A'V}$. Mais la topographie des densités radiales de courant permet de distinguer clairement des générateurs temporaux, identiques à ceux de la MMN auditive, et des générateurs occipitaux.

Nous avons comparé l'amplitude de la $MMN_{A'V'}$ à la somme des amplitudes des $MMN_{A'V}$ et $MMN_{AV'}$, à la latence du pic de la MMN auditive : l'additivité est significativement violée sur 12 électrodes situées pour la plupart sur l'hémiscalep gauche (figure 14.3.A page 191). La topographie de la violation du modèle additif est centrée sur une zone pariéto-occipitale gauche et ne ressemble ni à la topographie auditive, ni à la topographie visuelle.

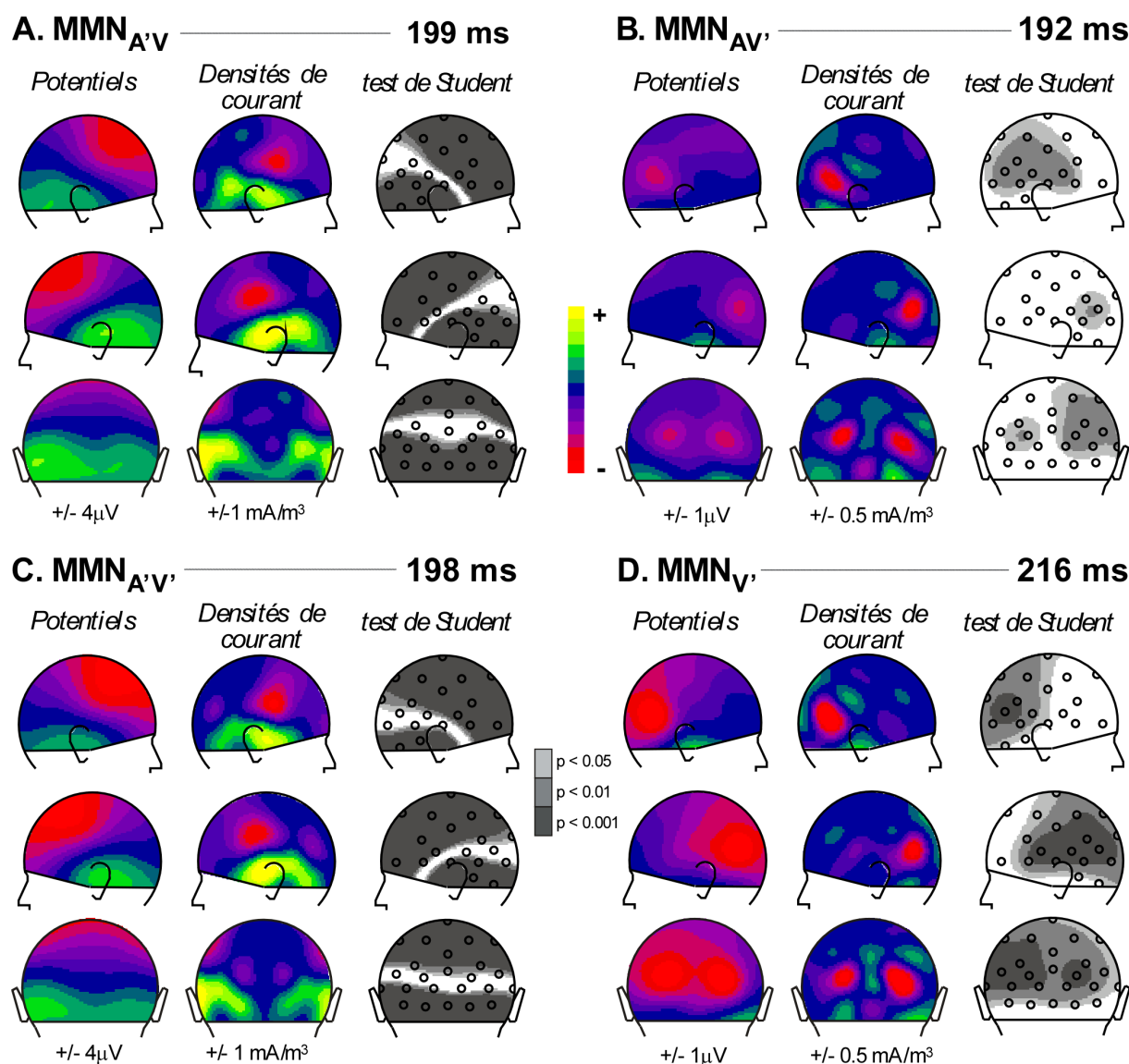


FIG. 14.2 – Topographies des MMN générées par les déviants A'V (A), AV' (B) et A'V' (C) dans le paradigme audiovisuel, et par le déviant V' dans le paradigme visuel seul (D), à la latence de leurs pics d'amplitude respectifs. Le maximum de l'échelle de couleur est indiqué sous chaque ensemble de cartes. Les cartes des tests de Student en niveaux de gris indiquent la significativité des amplitudes par rapport à la ligne de base.

La MMN visuelle générée en contexte unimodal ($MMN_{V'}$) est illustrée dans la figure 14.4 page 192 et comparée à la MMN visuelle générée en contexte bimodal ($MMN_{AV'}$). Les deux MMNs sont très ressemblantes, comme on peut le constater également sur la topographies des potentiels et des densités radiales de courant (figures 14.2.B et 14.2.D de la présente page). La $MMN_{V'}$ semble cependant ne posséder qu'un pic d'amplitude vers 216 ms. La différence d'amplitude entre les MMN visuelles générées dans les deux contextes à cette latence est significative sur 8 électrodes (figure 14.3.B page ci-contre). La topographie de la différence est difficile à interpréter mais suggère que cette différence d'amplitude n'est

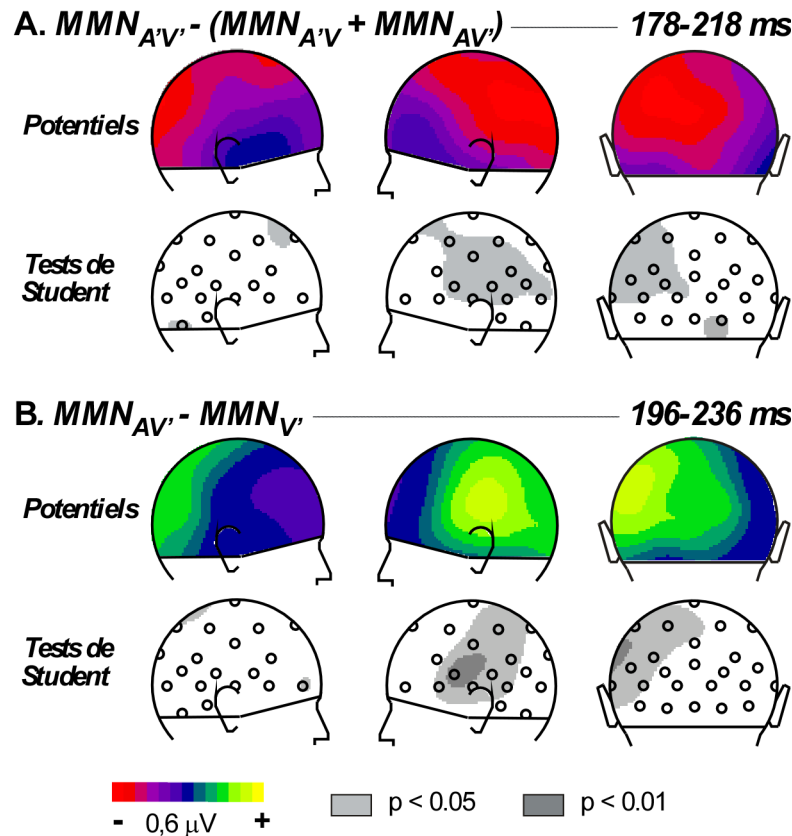


FIG. 14.3 – A. Topographie de la violation du modèle additif [$MMN_{A'V'} - (MMN_{A'V} + MMN_{AV'})$] à la latence du maximum des MMNs auditives (198 ms), dans le paradigme audiovisuel. B. Topographie de la différence entre les MMN visuelles en contexte unimodal ($MMN_{V'}$) et bimodal ($MMN_{AV'}$) à la latence du second pic de la $MMN_{AV'}$. L'échelle est commune à toutes les cartes de potentiels. Les cartes de Student indiquent la significativité des différences.

pas due à une modulation d'amplitude des générateurs de la MMN visuelle.

14.4 Discussion

Une MMN générée par une déviation audiovisuelle d'un événement audiovisuel présente donc les deux caractéristiques suivantes : elle est composée d'un générateur supra-temporal et d'un générateur occipital, ce qui indique qu'elle met en jeu à la fois les aires sensorielles auditives et les aires sensorielles visuelles ; mais elle n'est pas strictement égale à la somme des MMN générées d'une part par une déviance auditive et d'autre part par une déviance visuelle du même événement audiovisuel. Les processus indexés par les MMN visuelle et auditive semblent donc n'être pas totalement indépendants. Contrairement aux conclusions de Schröger et Widmann (1998), la coactivation qui facilite le temps de détection des déviants audiovisuels semble commencer dès l'étape de détection automatique de la déviance, qui repose sur l'existence d'une représentation des sons (et des images) standards en mémoire sensorielle.

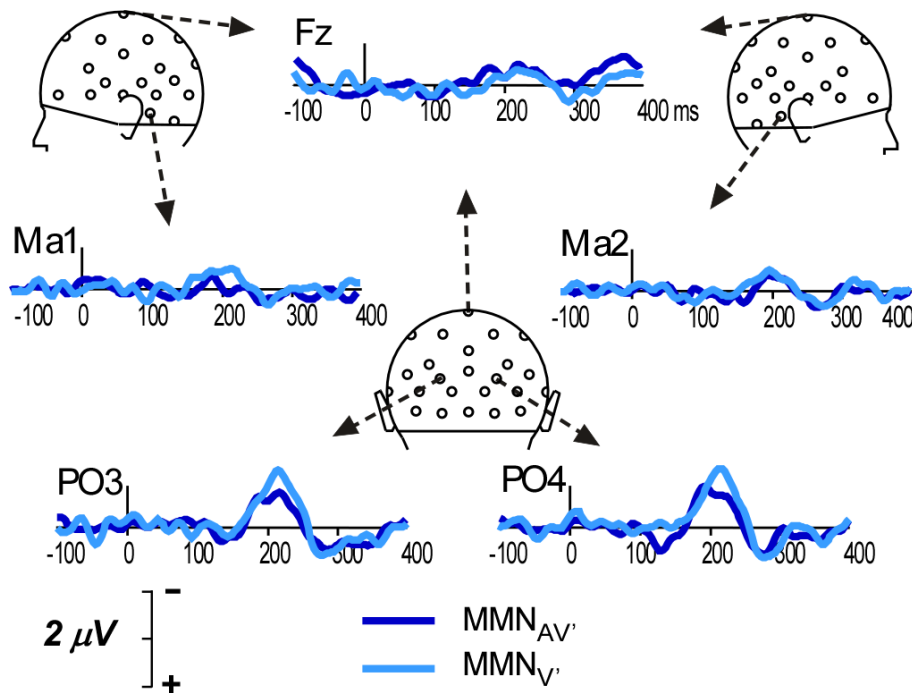


FIG. 14.4 – MMNs visuelles générées en contexte unimodal ($MMN_{V'}$) et bimodal ($MMN_{AV'}$), à un sous-ensemble d'électrodes.

La violation de l'additivité semble corroborer notre hypothèse de l'existence d'une trace audiovisuelle en mémoire sensorielle auditive. Il existe cependant d'autres explications de la violation de l'additivité à considérer.

D'abord, la non-additivité pourrait avoir été provoquée par la présence d'une MMN auditive générée par la déviance visuelle, comme dans le cas de l'illusion McGurk (Colin et coll., 2004 ; Colin, Radeau, Soquet, Demolin et coll., 2002 ; Möttönen et coll., 2002 ; Sams et coll., 1991) et de la ventriloquie (Colin, Radeau, Soquet, Dachy & Deltenre, 2002 ; Stekelenburg et coll., 2004). Cette explication semble cependant ici peu plausible car il est peu probable que la présentation d'un déviant visuel dans notre protocole ait pu modifier la perception auditive du standard auditif comme dans le cas des illusions. Récemment, Saint-Amour, De Sanctis, Molholm, Ritter et Foxe (2007) ont montré que la MMN aux syllabes McGurk déviant sur leur dimension visuelle provenait effectivement du cortex auditif et non d'une détection de la déviance visuelle.

D'autres études récentes ont toutefois montré que le même phénomène était observable lorsque les stimuli auditifs et visuels entretenaient des liens étroits sans pour autant créer une illusion audiovisuelle. Ainsi des stimuli audiovisuels écologiques tels que l'action d'un marteau sur un clou peut provoquer une activité auditive ressemblant à une MMN lorsque sa dimension visuelle est déviant (Ullsperger, Erdmann, Freude & Dehoff, 2006). Il en est de même pour des associations audiovisuelles arbitraires stockées en mémoire à long terme telles que les associations graphème/phonème (Yumoto et coll., 2005) et pour des associations apprises pour les besoins de l'expérimentation (associations symboliques arbi-

traires : Widmann, Kujala, Tervaniemi, Kujala & Schröger, 2004, ou physique : Aoyama, Endo, Honda & Takeda, 2006). Notons que dans tous ces études, les informations visuelles étaient disponibles avant le stimulus auditif (dans l'étude de Yumoto et coll., 2005, l'effet n'était plus observé lorsque le délai était trop réduit) si bien qu'il est possible que la MMN auditive ait pu être générée parce que le son présenté violait une attente créée par les informations visuelles. Dans notre étude, au contraire, les informations auditives et visuelles étaient disponibles au même moment, étaient associées de manière arbitraire sans être apprises avant l'expérience.

Quoiqu'il en soit, la violation observée dans notre expérience ne présente pas la topographie typique des activités générées dans le cortex auditif, ce qui rend peu probable cette explication.

D'autres explications de la non additivité semblent plus plausibles. Par exemple, dans la mesure où l'on observait un MMN visuelle d'origine occipitale en réponse à un déviant visuel, il est possible que la trace visuelle indexée par cette MMN ait été modifiée par la présence d'informations auditives, à l'inverse de notre hypothèse de départ. La topographie de la violation du modèle additif ne nous permet pas de conclure en faveur de l'une ou l'autre des hypothèses car elle ne présente ni les caractéristiques d'une activité générée dans le cortex visuel, ni celles d'une activité générée dans le cortex auditif.

La différence inattendue entre les MMN visuelles générées en contexte audiovisuel et en contexte visuel suggère néanmoins que la trace en mémoire sensorielle visuelle (sous l'hypothèse que la MMN visuelle a une origine mnésique, voir Czigler, sous presse, pour une revue) a intégré des informations sur la régularité auditive. En effet, la seule différence entre les deux protocoles était que dans le cas audiovisuel, les stimuli visuels étaient toujours associés à un stimulus auditif, et en particulier, que le standard visuel était associé au standard auditif dans 76% des essais. Étant donné que la déviance qui génère la MMN visuelle était la même dans les deux conditions, et que les traitements associés au stimulus auditif doivent disparaître dans la différence entre les PE standards et déviants, une interprétation tentante est que la trace en mémoire sensorielle visuelle a enregistré l'association régulière des standards auditifs et visuels. Cette explication n'est bien sûr pas incompatible avec notre hypothèse initiale : les deux processus de détection automatique de la déviance pourraient être influencés chacun par les informations de l'autre modalité sensorielle.

Une autre hypothèse à considérer pour expliquer la violation de l'additivité est que les informations auditives et visuelles n'ont interagi que dans le traitement des déviants, sans que les traces auditive et visuelle n'aient elles-mêmes été influencées par les informations de l'autre modalité sensorielle. Nos résultats montrent sans ambiguïté que les traitements des déviants auditifs et visuels ont interagi avant 200 ms de traitement. En effet, si tel n'était pas le cas, les MMN auditive et visuelle auraient dû être additives, même si les traces auditives et/ou visuelles intègrent des informations intersensorielles. Par contre la violation de l'additivité pourrait s'expliquer uniquement par une coactivation en aval du processus de comparaison, tout en préservant le caractère modalité-spécifique des traces mnésiques.

Dans le domaine auditif, par exemple, plusieurs études ont montré une violation de

l'additivité des MMN à la déviance simultanée sur deux traits acoustiques (Czigler & Winkler, 1996 ; Winkler, Czigler, Jaramillo, Paavilainen & Näätänen, 1998). Dans les deux cas, la MMN à la double déviance avait une amplitude inférieure à la somme des MMN aux déviances simples, comme si la détection d'une des deux déviances diminuait l'importance de l'autre déviance, suggérant l'existence de processus communs déclenchés par les deux déviances. De la même façon il est possible que la détection d'une déviance dans une modalité ait diminué le traitement de l'autre déviance, provoquant une violation de l'additivité. Cette explication est néanmoins insuffisante au moins pour la MMN visuelle puisqu'on trouvait une différence entre les MMN visuelles dans les contextes audiovisuel et visuel, qui peut difficilement s'expliquer par une différence de traitement des déviances.

En résumé, cette expérience ne nous a pas permis de vérifier sans ambiguïté notre hypothèse de départ, à savoir que la représentation de l'évènement en mémoire sensorielle auditive inclut des informations sur la régularité visuelle.

Ajoutons que nos données confirment l'origine occipitale de la MMN visuelle, dont la seule représentation topographique disponible était jusqu'à présent celle de Berti et Schröger (2004) dans une étude où l'attention des sujets était portée sur les stimuli, mais où la dimension de la déviance n'était pas pertinente pour la tâche à réaliser. Nos données ne suggèrent en revanche pas l'existence d'un composante antérieure de la MMN visuelle. Notons que nous n'avons pas contrôlé l'hypothèse de *refractoriness* dans notre expérience et qu'on ne peut donc formellement conclure que notre MMN visuelle est le marqueur d'une mémoire sensorielle visuelle. Plus généralement, nous ne pouvons exclure que la violation de l'additivité des MMN auditive et visuelle, résulte d'un phénomène de *refractoriness*. Cela impliquerait cependant l'existence de populations neuronales sensibles à l'association de stimuli auditifs et visuels particuliers.

Chapitre 15

Représentation d'une régularité audiovisuelle en mémoire sensorielle auditive

15.1 Introduction

L'expérience précédente n'a pas permis de montrer formellement qu'une régularité audiovisuelle est codée en mémoire sensorielle auditive. En revanche, nos données suggèrent que la représentation en mémoire sensorielle visuelle, si elle existe, inclut des informations auditives puisque la MMN visuelle unimodale était différente de celle générée par une déviance visuelle d'un événement audiovisuel. Pour montrer que la mémoire sensorielle auditive inclut des éléments visuels, il nous faut donc montrer, réciproquement, que la MMN générée par la déviance auditive d'un événement audiovisuel est différente de celle générée par la même déviance en contexte unimodal auditif.

Pour cela nous allons présenter dans un bloc expérimental unimodale un son standard pouvant dévier occasionnellement sur sa fréquence et dans un autre bloc audiovisuel, les mêmes sons standards et déviants mais associés à un stimulus visuel standard. Ainsi, dans le bloc audiovisuel, les traitements évoqués par les stimuli visuels devraient disparaître dans le calcul de la MMN. Si une différence subsiste entre les MMNs évoquées dans les deux blocs, elle devrait être due à des différences dans les processus de détection automatique de la déviance auditive. Mais puisque les déviations sont identiques dans les deux blocs, la différence devrait provenir de la différence existant dans la mémoire sensorielle auditive entre la trace d'un événement standard auditif et la trace d'un événement standard audiovisuel.

Nous allons comparer la même MMN auditive générée dans deux contextes différents : un contexte audiovisuel et un contexte auditif seul. Il reste donc toujours la possibilité que la simple présence d'informations visuelles module la MMN. Il a été montré par exemple que la MMN auditive peut être influencée par la présence de stimuli visuels émotionnels (Surakka, Tenhunen-Eskelinen, Hietanen & Sams, 1998), par la charge attentionnelle visuelle (Otten, Alain & Picton, 2000 ; Valtonen, May, Makinen & Tiitinen, 2003 ; Zhang, Chen, Yuan, Zhang & He, 2006) ou la direction de l'attention sélective vers la modalité visuelle ou

auditive (Alho, 1992 ; Dittmann-Balcar, Thienel & Schall, 1999 ; Muller-Gass, Stelmack & Campbell, 2006 ; Woods, Alho & Algazi, 1992).

Le fait que la tâche distractive visuelle soit la même dans les blocs auditif et audiovisuel devrait être un contrôle suffisant pour exclure ces effets attentionnels dans la mesure où elle devrait équilibrer l'attention visuelle soutenue de la même façon dans les deux blocs. Cependant, des stimuli visuels distracteurs peuvent provoquer des déplacements involontaires de l'attention visuelle (spatiale) et avoir une influence sur l'amplitude de la MMN (Mathiak, Hertrich, Zvyagintsev, Lutzenberger & Ackermann, 2005). De plus il est difficile de dire si la simple présence d'un stimulus visuel, même hors du focus attentionnel, pourrait influencer de manière non spécifique la MMN auditive, car cela n'a jamais été testé.

Un meilleur contrôle serait donc de montrer que l'effet des informations visuelles sur la MMN auditive a lieu lorsque les stimuli audiovisuels constituent une véritable régularité audiovisuelle, c'est-à-dire lorsque les mêmes événements auditifs et visuels sont associés de manière régulière, mais pas lorsque l'association audiovisuelle standard varie d'un essai à l'autre. On pourrait ainsi séparer l'effet non spécifique de la présence de stimuli visuels sur la MMN auditive de la construction d'une véritable représentation de l'évènement audiovisuel régulier en mémoire sensorielle auditive. Nous avons donc ajouté une condition de stimulation que nous avons appelé "audiovisuelle équiprobable" dans laquelle des sons standards et déviants identiques à ceux des autres conditions pouvaient être associés de manière équiprobable à quatre stimuli visuels différents.

Notre hypothèse est donc que la MMN générée par une même déviance auditive devrait être différente dans le cas où elle dévie par rapport à une régularité auditive (condition auditive unimodale), une régularité audiovisuelle (condition audiovisuelle) ou une régularité auditive accompagnée d'informations visuelles ne constituant pas une régularité (condition audiovisuelle équiprobable). En particulier, nous prédisons que la MMN auditive dans la condition audiovisuelle devrait se différencier à la fois de la MMN auditive dans la condition unimodale et de celle générée dans la condition audiovisuelle équiprobable, ces deux dernières devant être identiques, si la simple présence d'informations visuelles n'a pas d'effet sur la MMN auditive.

15.2 Méthodes

15.2.1 Sujets

Seize sujets droitiers (dont 9 de sexe féminin) âgés en moyenne de 24 ans (écart-type : 2,5 ans) ont participé à cette expérience. Aucun sujet ne souffrait de troubles neurologiques. Ils avaient tous une audition normale et une vision normale ou corrigée.

15.2.2 Stimuli

Les stimuli utilisés étaient identiques à ceux des deux expériences précédentes excepté dans la condition audiovisuelle équiprobable, où deux types de composantes visuelles supplémentaires ont été ajoutés. Il s'agissait de déformations du cercle dans deux directions

obliques (V_3 et V_4), montrés dans la figure 15.1.

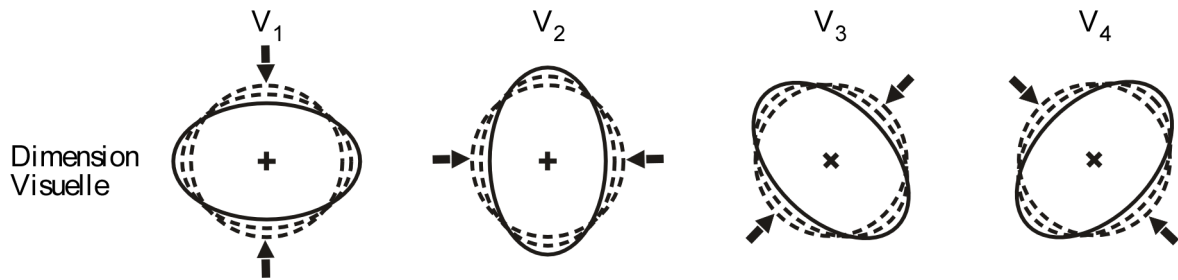


FIG. 15.1 – Composantes visuelles des stimuli audiovisuels utilisées dans la condition audiovisuelle équiprobable

15.2.3 Procédure

Chacune des trois conditions de stimulation comprenait 1600 événements, dont 320 (20%) étaient des déviants auditifs, répartis en 8 blocs de 200 stimuli d'une durée d'environ 1 minutes 50.

Dans la condition auditive unimodale, le stimulus standard (A) était le stimulus A_1 et le déviant (A') était le stimulus A_2 , dans la moitié des blocs. Dans l'autre moitié, les rôles de A_1 et A_2 étaient inversés.

Dans la condition audiovisuelle, les stimuli auditifs standard A et déviant A' d'un même bloc étaient toujours accompagnés du même stimulus visuel V. Dans la moitié des blocs, le stimulus standard (AV) était le stimulus A_1V_1 et le stimulus déviant (A'V) était le stimulus A_2V_1 . Dans l'autre moitié, les stimuli standard et déviant étaient les stimuli A_2V_2 et A_1V_2 .

Dans la condition audiovisuelle équiprobable, les sons standard et déviant pouvaient indifféremment être associés à l'un des quatre stimuli visuels. Dans la moitié des blocs, il y avait donc 4 standards audiovisuels A_1V_1 , A_1V_2 , A_1V_3 et A_1V_4 , présentés chacun dans 20% des essais et 4 déviants audiovisuels A_2V_1 , A_2V_2 , A_2V_3 et A_2V_4 présentés chacun dans 5% des essais. Dans l'autre moitié des blocs, les probabilités d'occurrence étaient inversées entre les stimuli constitués du son A_1 et ceux constitués du son A_2 . Les standards et les déviants dans cette condition seront désormais nommés AVEq et AV'eq.

Les 24 blocs de stimulations étaient présentés dans un ordre aléatoire, différent d'un sujet à l'autre. La tâche distractive et la probabilité d'occurrence de la disparition du point de fixation étaient identiques à celles utilisées dans l'expérience précédente.

Les contraintes de succession appliquées aux déviants et aux standards étaient identiques à celles des deux expériences précédentes.

15.2.4 Analyses

Après rejet des artefacts d'enregistrement et l'exclusion des essais standards ayant contenu une cible ou suivant un déviant, le nombre moyen d'essais par sujet pour le calcul des potentiels évoqués moyens étaient de 741, 755, 826, 279, 284 et 317 respectivement

pour les standards A, AV, AVeq et les déviants A', A'V et A'Veq. La ligne de base était prise entre 100 ms et 0 ms avant la stimulation.

Les MMN auditives $MMN_{A'}$, $MMN_{A'V}$ et $MMN_{A'Veq}$ dans chacune des 3 conditions étaient calculées comme la différence, point par point, entre les potentiels évoqués par chacun des déviants et les potentiels évoqués par chacun des standards.

Pour les tests statistiques, nous avons voulu limiter les hypothèses faites sur la latence des effets, tout en limitant le risque de première espèce global à 5%. Nous avons donc testé la différence entre les MMN à toutes les latences dans une fenêtre 150-250 ms (correspondant à la latence de la MMN auditive dans l'expérience précédente) grâce à un test bilatéral de permutation des conditions appariées (Efron & Tibshirani, 1993, p212; $2^{16} = 65536$ permutations). Les tests multiples ont été pris en compte au niveau de chaque électrode par la méthode du minimum d'échantillons significatifs successifs, avec un risque local $\alpha = 0,05$ et un risque global $\alpha_{global} = 0,05$ (voir la partie 8.1 page 112).

15.3 Résultats

Les TR moyens dans la tâche distractive étaient respectivement de 334, 345 et 348 ms dans les conditions auditive, audiovisuelle et audiovisuelle équiprobable (écarts-types : 52, 53 et 50 ms). Les TR dans les trois conditions étaient significativement différents ($p < 0,00007$) et cette différence était due au fait que le TR dans la condition auditive était plus rapide que dans les deux conditions audiovisuelles (auditive contre audiovisuel : $p < 0,003$; auditive contre équiprobable : $p < 0,0001$). Les taux de cibles manquées pour les 3 conditions étaient respectivement de 1,17%, 1,00% et 1,25% (écarts-types : 1,25%, 1,12% et 1,64%). Ils n'étaient pas significativement différents.

Les figures 15.2 page suivante et 15.3 page 200 présentent les potentiels évoqués par les événements standards et déviants dans les 3 conditions de présentation. Dans la condition auditive (figure 15.2.A), les sons standards et déviants évoquaient une série d'ondes fronto-centrales, caractéristiques du traitement d'un stimulus auditif, visibles notamment sur Cz : une P50 avec un pic d'amplitude vers 60 ms et une inversion de polarité dont le maximum se situe aux mastoïdes, puis une minuscule N100 avec un pic d'amplitude à 100 ms, dont la faible amplitude est probablement due à l'intervalle inter-stimulus relativement rapide. À partir d'environ 120 ms, les potentiels évoqués par les standards et les déviants se séparent et les déviants évoquent une onde d'amplitude importante (la MMN) dont le pic négatif se situe vers 200 ms sur les électrode fronto-centrales et qui présente une inversion de polarité aux mastoïdes.

Dans les deux conditions audiovisuelles (figures 15.2.B et 15.2.C), les potentiels évoqués par les standards sont des agrégats complexes de réponses sensorielles auditives et visuelles. Concernant la modalité visuelle, on peut remarquer sur les électrodes occipitales (O1 et O2 sur la figure), superposées aux réponses auditives, d'abord une onde positive avec un pic d'amplitude vers 130 ms et une onde négative dont le pic d'amplitude se trouve vers 170 ms. Ces ondes sont aussi bien évoquées par les standards que par les déviants. Comme dans la condition auditive, les réponses évoquées par les standards et déviants commencent à différer vers 120 ms.

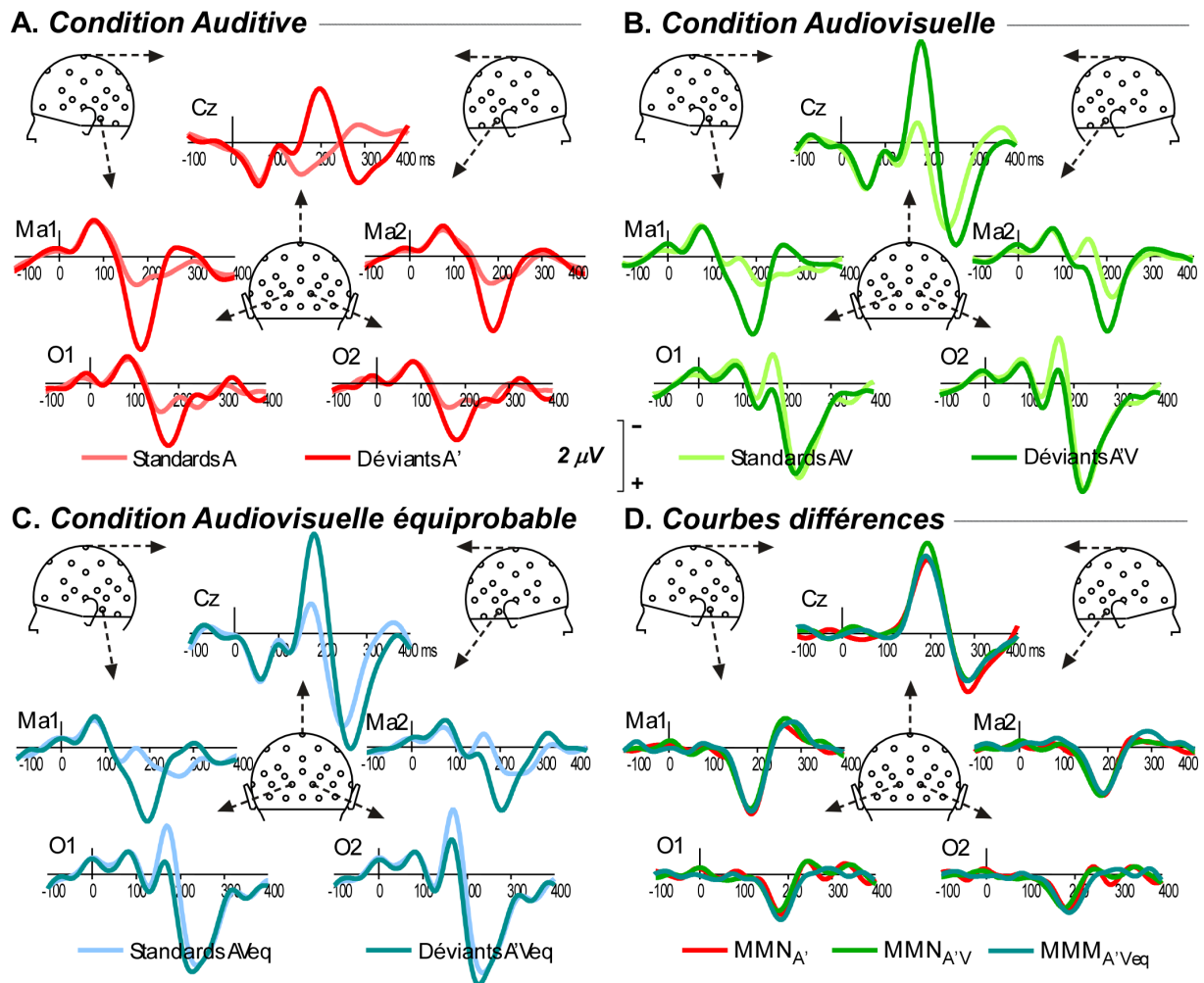


FIG. 15.2 – PE standards et déviants générés dans les conditions auditive unimodale (A), audiovisuelle (B) et audiovisuelles équiprobable (C). D. MMN auditives générées dans les trois conditions.

Lorsque l'on calcule la différence entre les réponses aux déviants et standards, on obtient des courbes très similaires dans les trois conditions (figure 15.2.D), ce qui correspond au fait que la déviance était identique dans ces conditions. La MMN_{A'} avait son pic sur l'électrode Fz à 192 ms ($-2,764\mu V$), la MMN_{A'V} à 194 ms ($-3,021\mu V$) et MMN_{A'Veq} à 192 ms ($-2,714\mu V$).

Comme on peut le constater sur la figure 15.3 page suivante, les topographies des 3 MMN sont très similaires, aussi bien au niveau des potentiels que des densités radiales de courant. L'amplitude du pic négatif de la MMN_{A'V} semble cependant plus importante. Les tests de permutation de la différence entre les MMN_{A'} et MMN_{A'V} (figure 15.4 page suivante) montrent en effet que l'amplitude des deux MMN est significativement différente sur plusieurs électrodes pariéto-centrales entre 180 et 205 ms. Seule la différence sur l'électrode CP1 subsiste lorsque les test multiples sont pris en compte.

Concernant le test de la comparaison entre la condition audiovisuelle et la condition audiovisuelle équiprobable (figure 15.5 page 201), la différence d'amplitude entre les deux

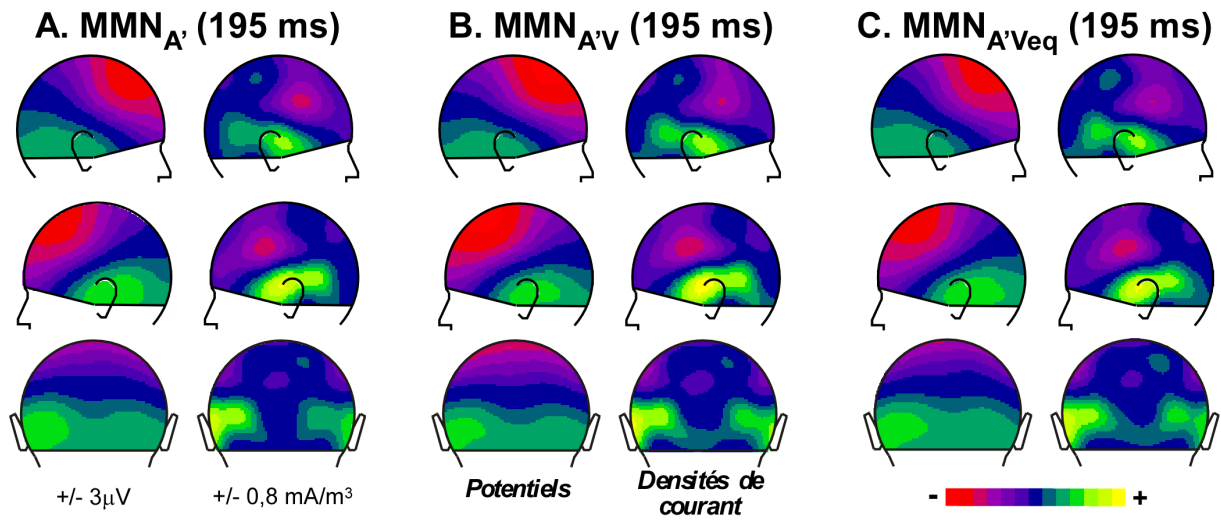


FIG. 15.3 – Topographies des MMN auditives générées dans les conditions auditive unimodale (A), audiovisuelle (B) et audiovisuelles équiprobable (C) à 195 ms.

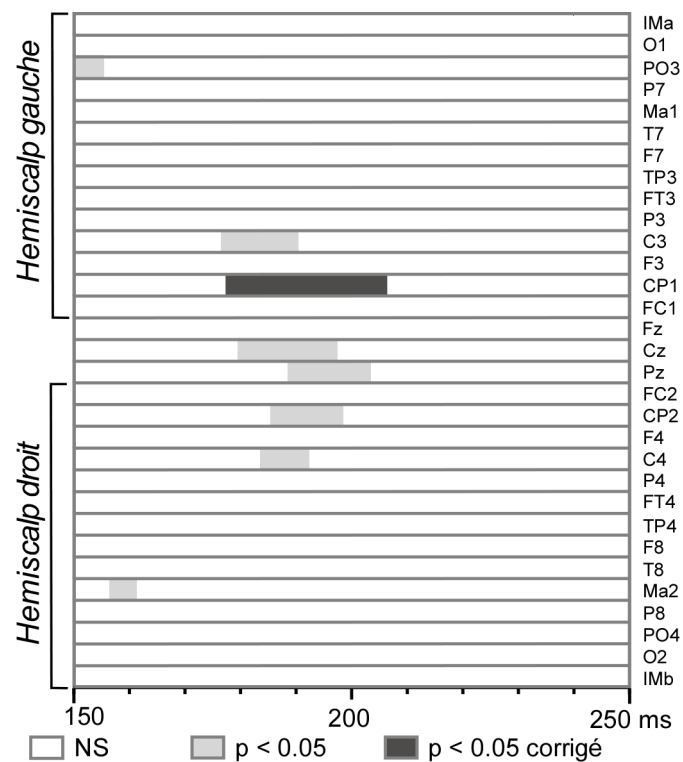


FIG. 15.4 – Résultats des tests de permutation de la différence entre les $MMN_{A'}$ et $MMN_{A'V}$, entre 150 et 250 ms. Le niveau de gris indique la significativité. $p < 0,05$ corrigé : le nombre de tests significatifs successifs dépasse le nombre d'échantillons minimal nécessaire pour limiter le risque global à 0,05.

MMN est significative également, mais seulement sur l'électrode fronto-centrale FC2. De plus, cette différence ne subsiste pas lorsque les test multiples sont pris en compte. La même

comparaison génère d'autres tests significatifs (ne résistant pas plus aux corrections) à une latence plus tardive (entre 215 et 245) ms sur plusieurs électrode pariéto-occipitales à gauche et à droite.

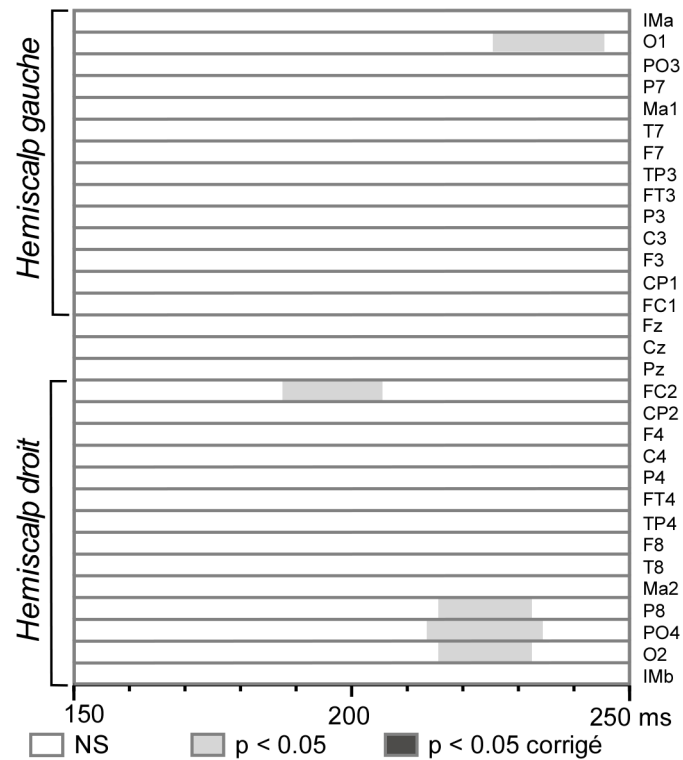


FIG. 15.5 – Résultats des tests de permutation de la différence entre les $MMN_{A'V}$ et $MMN_{A'Veq}$, entre 150 et 250 ms. Le niveau de gris indique la significativité. $p < 0,05$ corrigé : le nombre de tests significatifs successifs dépasse le nombre d'échantillons minimal nécessaire pour limiter le risque global à 0,05.

Le résultat des deux tests statistiques suggère que les MMN différaient à deux latences et en deux zones différentes du scalp. La figure 15.6 page suivante montre la topographie des deux différences testées, au cours du temps. Alors que la topographie de la différence entre les $MMN_{A'}$ et $MMN_{A'V}$ présente un pôle unique commençant sur les électrode centrales et se terminant sur les électrodes frontales, celle de la comparaison des deux conditions audiovisuelles semble être une superposition de la même différence et d'une seconde différence plus tardive et clairement occipitale.

Les densités radiales de courants correspondant à ces différences ne présentaient pas de topographie suffisamment stable (il s'agit d'une différence de différences) pour aider à cette interprétation.

15.4 Discussion

Les résultats vont dans le sens de nos hypothèses puisque la MMN générée par la déviance auditive d'un évènement audiovisuel standard diffère de la MMN unimodale générée par la même déviance auditive. La différence est faible, mais néanmoins significative, même

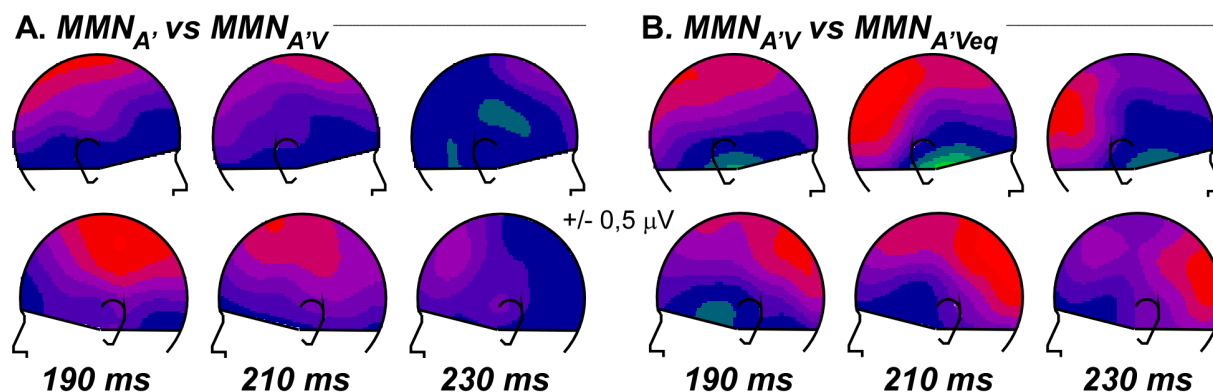


FIG. 15.6 – A. Topographies des différences entre les MMN auditives dans la condition auditive ($MMN_{A'}$) et la condition audiovisuelle ($MMN_{A'V}$), entre 190 et 230 ms. B. Topographies des différences entre les MMN auditives dans la condition audiovisuelle ($MMN_{A'V}$) et la condition audiovisuelle équiprobable ($MMN_{A'Veq}$), entre 190 et 230 ms.

en tenant compte des nombreux tests effectués. De plus, ce résultat ne peut s'expliquer simplement par un effet non spécifique de la présence de stimuli visuels, puisque, dans la condition audiovisuelle équiprobable, lorsque les stimuli auditifs étaient associés avec 4 stimuli visuels différents, il semble que les informations visuelles n'aient pas modifié la MMN auditive par rapport à la condition auditive seule.

Notre interprétation est donc que la représentation d'un événement audiovisuel en mémoire sensorielle est différente de celle d'un événement purement auditif, et que l'incorporation de l'élément visuel de la régularité ne peut se faire que si les stimuli auditifs et visuels sont associés de façon consistante au cours des essais. La mémoire sensorielle auditive, telle qu'indexée par la MMN auditive, semble donc stocker des régularités visuelles lorsque celles-ci sont associées à des régularités auditives.

Il semble que la MMN générée dans la condition audiovisuelle équiprobable présente des caractéristiques qui lui sont propres, puisqu'elle présentait une seconde différence par rapport à la condition audiovisuelle, vers la fin de la MMN et sur des électrodes plus occipitales. On peut supposer que la MMN auditive générée dans la condition audiovisuelle équiprobable différait de celle générée dans la condition auditive seule. Il serait hasardeux de s'aventurer à en donner une interprétation, étant donné, d'une part, que nous n'avions aucune hypothèse a priori sur cette différence, et d'autre part, que la significativité de cette différence peut s'expliquer par le nombre de tests effectués.

Il semble que la présence d'informations visuelles ait eu pour effet de ralentir le temps de réaction des sujets dans leur tâche distractive puisque celui-ci était plus rapide d'environ 10 millisecondes dans la condition auditive seule par rapport aux deux conditions audiovisuelles. Les sujets ont donc pu être légèrement distraits par le mouvement du cercle dans leur traitement de la disparition du point de fixation. Mais ils l'étaient tout autant dans la condition audiovisuelle et dans la condition audiovisuelle équiprobable. Cette distraction ne peut donc expliquer ni les effets spécifiques des informations visuelles régulières sur la MMN auditive, qui n'existaient pas dans la condition audiovisuelle équiprobable, ni les effets spécifiques à la condition audiovisuelle équiprobable.

Même si la déviance, c'est-à-dire la différence entre déviants et standards, était la même dans chacune des trois conditions, les déviants utilisés n'étaient pas, à proprement parler, identiques, puisqu'ils étaient purement auditifs dans une condition, audiovisuels dans les deux autres. Il reste donc toujours la possibilité que la différence entre les MMN proviennent simplement de la différence de traitement des déviants lorsqu'ils sont auditifs ou audiovisuels. Cette explication ne dispense pas de l'existence d'interactions entre les traitements auditifs et visuels (en effet, s'il n'y avait aucune interaction, les traitements visuels devraient purement et simplement s'éliminer dans le calcul de la MMN auditive et il n'y aurait aucune différence entre les MMN), mais compromettrait notre interprétation en termes de mémoire sensorielle auditive. Il paraît difficilement soutenable, cependant, que la partie visuelle du déviant modifie son traitement dans le cortex auditif, sans qu'il en soit de même pour les standards et que donc la représentation de l'évènement standard en mémoire sensorielle auditive soit affectée par la présence d'informations visuelles.

Il semble donc que l'association régulière d'un stimulus auditif donné, avec un stimulus visuel donné, finisse par générer la perception d'un objet audiovisuel à part entière. La représentation sensorielle de cet objet pourrait être stockée en mémoire sensorielle auditive et en mémoire sensorielle visuelle (si l'on en croit les résultats de la comparaison des MMN_{AV} et MMN_V de l'expérience précédente). Le stockage de cette représentation audiovisuelle intégrée dans ces deux mémoires sensorielles pourrait être à l'origine de la facilitation pour la détection d'un déviant audiovisuel, mis en évidence dans la première expérience.

Chapitre 16

MMN à la conjonction audiovisuelle

Cette étude ayant été acceptée pour publication (Besle et coll., sous presse), elle ne sera que brièvement présentée ici. Les détails en sont décrits dans la publication, intégrée au manuscrit, intégrée au manuscrit en annexe (page 267). Cette expérience a été réalisée au centre MEG du CERMEP, à Lyon. Les données ont été acquises par Romaine Mayet, en DEA sous la direction de Dominique Morlet et analysées en collaboration avec Anne Caclin et Dominique Morlet.

16.1 Introduction

Nos expériences précédentes suggèrent qu'une régularité audiovisuelle est représentée en mémoire sensorielle auditive et peut-être en mémoire sensorielle visuelle. Cependant, cela n'a été montré qu'assez indirectement, en étudiant l'influence d'une régularité visuelle sur la représentation d'une régularité auditive et vice-versa. Dans l'expérience suivante, nous avons tenté de savoir si la représentation de cette régularité audiovisuelle peut être à l'origine d'une activité de type MMN lorsque la régularité est violée, autrement dit s'il existe une représentation mnésique sensorielle à part entière d'une association particulière et régulière entre un trait auditif et un trait visuel. Pour cela, nous avons présenté des stimuli audiovisuels déviants, ne différant de la régularité audiovisuelle que sur la façon dont les traits auditifs et visuels sont combinés (conjonction de traits), chaque trait auditif ou visuel pris isolément ne constituant pas la violation d'une régularité auditive ou visuelle.

Ces déviants à la conjonction de deux traits ont déjà été utilisés dans des études sur la MMN auditive pour montrer que la mémoire sensorielle auditive ne stocke pas uniquement des représentations indépendantes des traits acoustiques élémentaires, mais également des représentations de leurs combinaisons particulières (Gomes, Bernstein, Ritter, Vaughan & Miller, 1997 ; Sussman, Gomes, Nousak, Ritter & Vaughan, 1998 ; Takegata, Paavilainen, Näätänen & Winkler, 1999 ; Takegata, Huotilainen, Rinne, Näätänen & Winkler, 2001 ; Winkler, Czigler, Sussman, Horvath & Balazs, 2005). Dans ces expériences, plusieurs sons standards différant sur deux traits acoustiques (par exemple un son fort et aigu et un son faible et grave) sont présentés avec une probabilité équivalente. Les sons déviants occasionnels ont un trait identique à l'un des standards sur une dimension et un trait identique à un autre standard sur l'autre dimension (par exemple un son fort et grave).

Ainsi, les deux traits acoustiques du déviant pris séparément appartiennent à une régularité acoustique et sont donc représentés en mémoire sensorielle auditive. Un tel déviant génère une MMN auditive qui ne peut être attribuée à aucun des deux traits élémentaires de déviance, et on peut en conclure que la conjonction des deux traits elle-même est représentée en mémoire sensorielle auditive. Un tel résultat a également été rapporté récemment dans la modalité visuelle (Winkler et coll., 2005).

Avec un tel protocole appliqué au cas audiovisuel, on peut donc tester si la conjonction audiovisuelle en tant que telle est représentée en mémoire sensorielle, et si oui dans quelle modalité : auditive, visuelle ou les deux. Nous avons donc présenté des événements audiovisuels standards équiprobables (A_1V_1 et A_2V_2) et des déviants audiovisuels (A_1V_2 et A_2V_1), dont les composantes auditives et visuelles sont présentes dans les standards mais dont la conjonction est inédite par rapport aux standards. Si une représentation mnésique sensorielle de la régularité audiovisuelle existe en tant que telle, on devrait observer une différence dans le traitement des standards et des déviants, bien que les traits auditifs et visuels appartiennent chacun à une régularité unisensorielle. Nos expériences précédentes suggèrent que la régularité audiovisuelle est codée à la fois en mémoire sensorielle auditive et en mémoire sensorielle visuelle. Nous prédisons donc que cette MMN à la conjonction audiovisuelle devrait présenter à la fois des générateurs auditifs et visuels.

Pour cette expérience, deux contrôles importants doivent être réalisés pour éviter de confondre la MMN à la conjonction de traits audiovisuels avec d'autres processus. D'une part, la MMN ne doit pas être due à une différence de caractéristiques physiques entre déviants, ce qui n'a pas toujours été contrôlé dans les études de MMN à la conjonction auditive (voir par exemple : Gomes et coll., 1997 ; Sussman et coll., 1998). Dans notre expérience, 2 standards et 2 déviants étaient utilisés, les 2 déviants présentant les mêmes traits auditifs et visuels que les 2 standards, si bien que les traitements propres aux différents traits auditifs et visuels disparaissaient dans le calcul de la MMN à la conjonction.

D'autre part, une MMN ne doit pas être générée par la détection d'une déviance locale dans une seule modalité. Il a en effet été montré, dans la modalité auditive, que des représentations des traits élémentaires et de la conjonction de traits pouvaient coexister en mémoire sensorielle auditive (Takegata et coll., 2001, 1999). De la même façon, une représentation de la conjonction audiovisuelle coexiste sans doute avec les représentations des parties unimodales de la régularité.

Or, pour éviter que l'alternance des deux standards ne soit elle-même une régularité et que la MMN soit provoquée par la violation de cette régularité, on doit présenter aléatoirement les 2 standards. Dans ce cas, un déviant à la conjonction peut être précédé du même standard présenté plusieurs fois. Comme, d'une part, il suffit de trois standards pour qu'une trace se constitue (Cowan et coll., 1993), et même moins dans le cas où le stimulus a déjà été présenté précédemment (Nousak, Deacon, Ritter & Vaughan, 1996 ; Winkler, Cowan, Csepe, Czigler & Näätänen, 1996), et comme, d'autre part, un déviant à la conjonction (par exemple A_2V_1) diffère d'un standard (par exemple A_2V_2) sur un des deux traits, il est possible que la MMN générée contienne une composante unisensorielle (visuelle dans notre exemple), générée par la probabilité locale dans la série présentée.

De la même façon, un standard donné diffère de l'autre standard ou d'un déviant sur

au moins un trait. Une MMN unisensorielle pouvait donc être générée également par un standard si les stimuli précédents présentaient plusieurs fois de suite le même trait. Pour contrôler de tels effets indésirables, nous nous sommes assuré que les pourcentages de stimuli présentant un trait auditif ou un trait visuel donné, précédés par un nombre donné de traits identiques dans l'autre modalité, étaient équivalents pour les standards et les déviants à la conjonction sur l'ensemble de l'expérience. Ainsi, toute MMN unisensorielle, qu'elle soit auditive ou visuelle, devrait disparaître dans le calcul de la MMN à la conjonction.

Cette expérience a été réalisée en MEG. Puisque nous n'avions pas, au laboratoire, l'expérience de ce qu'est une MMN auditive en MEG (souvent appelée MMF pour *Mismatch Field*¹), les sujets ont de plus participé à une expérience purement auditive dans laquelle des sons standards et déviants étaient présentés.

16.2 Méthodes

16.2.1 Sujets

Dix sujets droitiers (dont 5 de sexe féminin) âgés en moyenne de 29 ans (écart-type : 7 ans) ont participé à cette expérience. Aucun sujet ne souffrait de troubles neurologiques. Ils avaient tous une audition normale et une vision normale ou corrigée.

16.2.2 Stimuli

Les stimuli étaient identiques à ceux utilisés dans les expériences précédentes, excepté quelques détails : les mouvements des stimuli visuels étaient constitués de 5 trames d'une durée de 33 ms chacune. La durée des sons était de 167 ms (dont 10 ms de montée/descente)

16.2.3 Procédure

L'intervalle interstimulus était de 583 ms. 2600 stimuli audiovisuels (dont 312 déviants) ont été présentés, répartis dans 10 blocs d'une durée de 2 minutes 30 environ chacun.

Dans tous les blocs, les stimuli A_1V_1 et A_2V_2 étaient utilisés comme standards, avec une probabilité d'occurrence de 44% chacun. Les déviants étaient les stimuli A_1V_2 et A_2V_1 et avaient une probabilité d'occurrence de 6% chacun.

La tâche distractive était identique à celle des expériences précédentes, excepté que la disparition du point de fixation avait une probabilité d'occurrence de 10%.

Dans le paradigme auditif unimodal, les stimuli A_1 et A_2 jouaient, tour à tour, les rôles de standards et déviants avec des probabilités d'occurrence respectives de 88 et 12%. 1700 stimuli (dont 204 déviants) ont été présentés, répartis en 4 blocs de 4 minutes 10. Le sujet devait lire une livre de son choix et ignorer les sons.

L'expérience auditive unimodale était toujours réalisée à la suite de l'expérience audiovisuelle. Dans tous les cas, chaque bloc de stimuli commençait par au moins trois standards, et un déviant était toujours précédé d'au moins 3 standards.

¹Pour plus de clarté, nous garderons la dénomination MMN bien que la négativité n'ait pas le même sens en MEG qu'en EEG.

16.2.4 Analyses

Les champs magnétiques évoqués (CME) de chaque sujet ont été calculés en excluant les 3 essais standards de début de chaque bloc, les essais standards suivant un déviant ainsi que ceux suivant une cible. Contrairement aux analyses EEG, le seuil de rejet des artéfacts était choisi pour chaque sujet de manière à ne pas rejeter plus de 85% des essais. La ligne de base était prise entre 100 ms et 0 ms avant la stimulation.

La MMN à la conjonction audiovisuelle était calculée comme la différence entre les CME aux stimuli déviants A_1V_2 et A_2V_1 et les CME aux stimuli standards A_1V_1 et A_2V_2 .

Nous avons testé la différence entre les CME aux déviants et aux standards dans une fenêtre 140-300 ms par des tests bilatéraux de permutation des conditions appariées (Efron & Tibshirani, 1993, p212; $2^{10} = 1024$ permutations). Les tests multiples ont été pris en compte au niveau de chaque électrode par la méthode du minimum d'échantillons significatifs successifs, avec un risque local $\alpha = 0,05$ et un risque global $\alpha_{global} = 0,05$ (voir la partie 8.1 page 112).

Contrairement à l'enregistrement EEG, dans lequel les électrodes sont disposées par rapport à des repères anatomiques propres à chaque sujet, les capteurs MEG sont disposés de façon rigide les uns par rapport aux autres, et sans rapport précis avec l'anatomie des sujets. Selon la taille de la tête et sa position dans le casque MEG, les capteurs peuvent donc enregistrer des signaux de provenances légèrement différentes selon les sujets. Cela introduit une variabilité non négligeable lors du moyennage des données de plusieurs sujet et limite la puissance des tests statistiques de groupes. Notre étude a donc été complétée par des analyses statistiques sur les données individuelles de chaque sujet (voir la partie 8.2.1 page 113). Pour chaque essai, les champs magnétiques des essais élémentaires standards et déviants ont été comparés par des tests de randomisation pour groupes indépendants. Les tests multiples ont été pris en compte de la même façon que pour les tests de groupe.

Pour avoir une idée de la variabilité de la position de la tête des sujets dans le casque, nous avons mesuré l'écart moyen de position de la tête entre les sujets pris deux à deux, et pour un même sujet entre les deux parties de l'expérience, grâce à trois bobines électromagnétiques placées sur la tête du sujet dans le casque MEG.

16.3 Résultats

Le temps de détection des sujets dans la tâche distractive était de 418 ms (écart-type : 50 ms) et le taux de cibles manquées inférieur à 1%.

Les mesures de la position relative de la tête dans le casque MEG montrent que celle-ci variait en moyenne de $4,6 \text{ mm} \pm 0,6 \text{ mm}$ au cours de l'expérience. La différence inter-sujet pouvait atteindre 40 mm.

La figure 16.1 page ci-contre montre les MMN auditive (16.1.A) et à la conjonction audiovisuelle (16.1.B). Les CME standards et déviants auditifs diffèrent significativement sur un grand nombre de capteurs entre 155 et 250 ms. La topographie de la différence

consiste en une inversion de polarité au niveau des capteurs temporaux correspondant à une activité vraisemblablement générée dans le plan supratemporal².

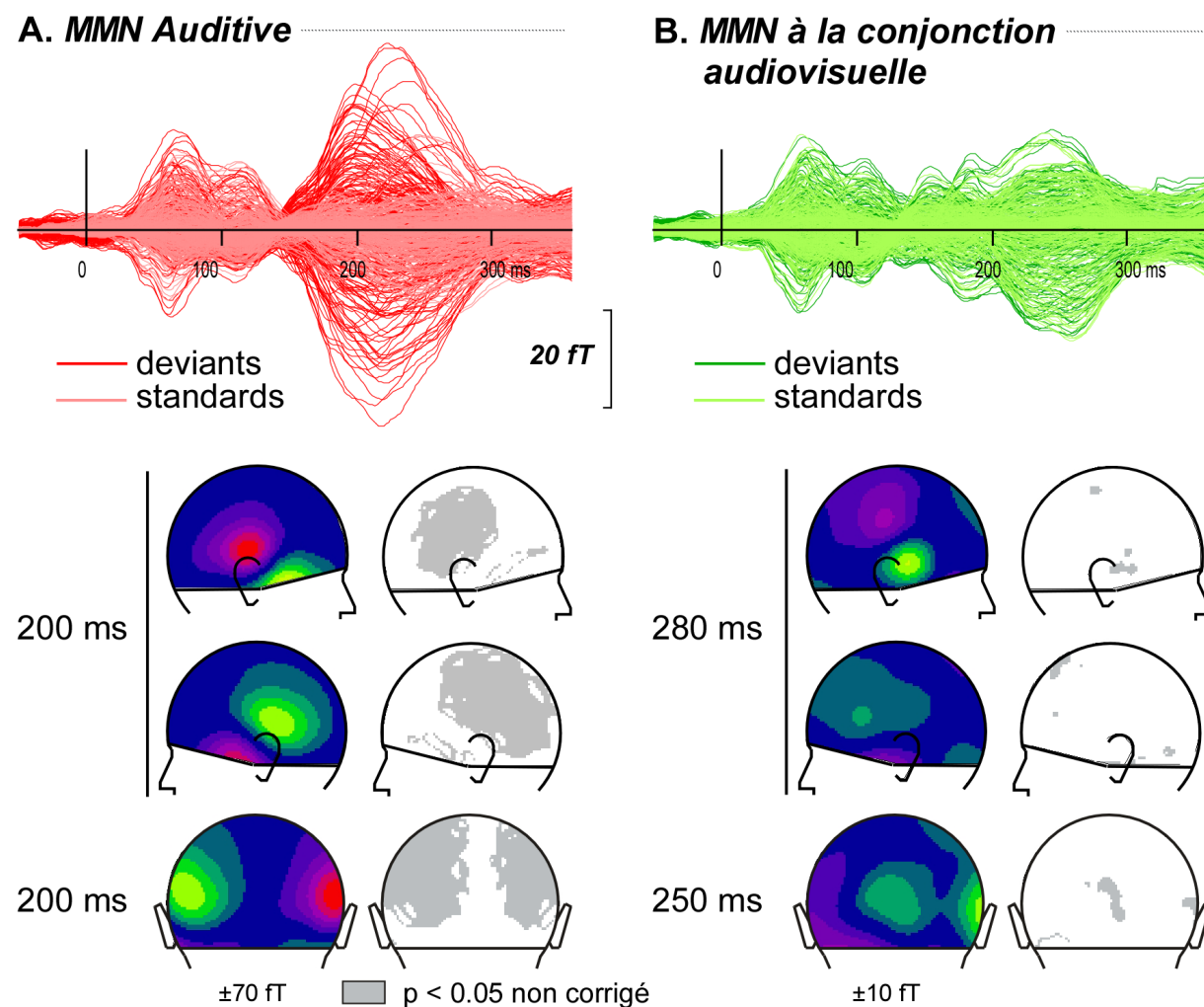


FIG. 16.1 – MMN auditive (A) et à la conjonction audiovisuelle (B). Les courbes sont les CME standards et déviants, enregistrés sur l'ensemble des capteurs et superposés. Les cartes en niveaux de gris indiquent les résultats au test de permutation de la différence entre CME standard et déviant à la latence de la topographie. Noter la différence d'échelle entre les cartes des deux MMN.

Concernant la MMN à la violation d'une conjonction audiovisuelle, les courbes évoquées par les déviants et les standards différaient à peine. Leurs amplitudes étaient cependant significativement différentes sur quelques capteurs occipitaux entre 235 et 265 ms et sur quelques capteurs temporaux gauches vers 280 ms, mais la différence ne subsistait pas à la prise en compte des tests multiples. La topographie de la différence est illustrée dans la figure 16.1.B à ces deux latences. Autour de 280 ms, elle présente une inversion de polarité

²Une inversion de polarité des signaux MEG signant une activité générée dans le cortex auditif doit, en principe, apparaître perpendiculaire au plan supra-temporal. Rappelons toutefois que la topographie MEG représente les activités enregistrées au niveau du casque, et non du scalp du sujet comme en EEG, et dépend donc de l'orientation de la tête des sujets par rapport au casque.

Sujets	MMN auditive		MMN conjonction audiovisuelle		
	Temporal Gauche	Temporal Droit	Temporal Gauche	Temporal Droit	Occipital
S1	150-240	150-250	190-265	195-265	215-235
S2	190-260	200-275	-	285-330 ?	-
S3	160-230	160-230	-	-	-
S4	170-230	160-230	245-260	245-295	245-265
S5	170-250	170-240	265-275 ?	270-300 ?	280-305 ?
S6	180-260	160-270	-	-	-
S7	180-250	180-270	220-270 ?	175-205 ?	245-255 ?
S8	170-240	180-230	265-295 ?	-	-
S9	205-215	200-250	-	-	230-270
S10	150-230	170-230	275-295	275-295	245-255

TAB. 16.1 – Latences (en ms) de début et de fin des réponses significatives pour la MMN auditive et pour la MMN à la conjonction audiovisuelle, chez chacun des sujets. Le point d’interrogation désigne les réponses dont la topographie est instable.

similaire à celle de la MMN auditive et qui peut donc refléter des activités dans le cortex auditif. Par ailleurs, vers 250 ms, la topographie de la MMN à la conjonction audiovisuelle présente une composante postérieure sur les aires occipitales, qui n’est pas présente dans la MMN auditive unimodale

La table 16.1 donne, pour chaque sujet, les fenêtre de latence dans lesquelles les CME aux standards et aux déviants différaient significativement sur des capteurs temporaux ou occipitaux (seuil non corrigé). Alors que tous les sujets montraient une MMN auditive, seuls 3 sujets sur 10 montraient clairement une MMN à la violation de la conjonction audiovisuelle en regard des aires occipitales et des aires temporales de façon bilatérale, 5 sujets sur 10 ne montraient qu’une différence temporale unilatérale ou occipitale, seulement marginalement significative ou instable, et 2 sujets ne montraient aucune différence significative. La latence de ces différences était assez variable d’un sujet à l’autre, contrairement à la latence de la MMN auditive.

La figure 16.2 page ci-contre illustre la topographie de la MMN auditive et de la MMN à la conjonction audiovisuelle pour un sujet particulier (10)

16.4 Expérience comportementale complémentaire

Étant donnée la faiblesse de la MMN audiovisuelle, nous avons voulu savoir si les sujets étaient capables de détecter comportementalement une déviance à la conjonction audiovisuelle et comparer leurs performances à la détection d’une déviance auditive. Six sujets âgés de 30 ans (écart-type : 7 ans) ont participé à cette expérience complémentaire. Trois de ces sujets avaient participé à l’expérience MEG. Les stimuli auditifs et visuels étaient identiques à ceux de l’expérience MEG. La tâche des sujets consistait à cliquer le plus rapidement possible lors de la présentation d’un stimulus déviant à la conjonction audiovisuelle

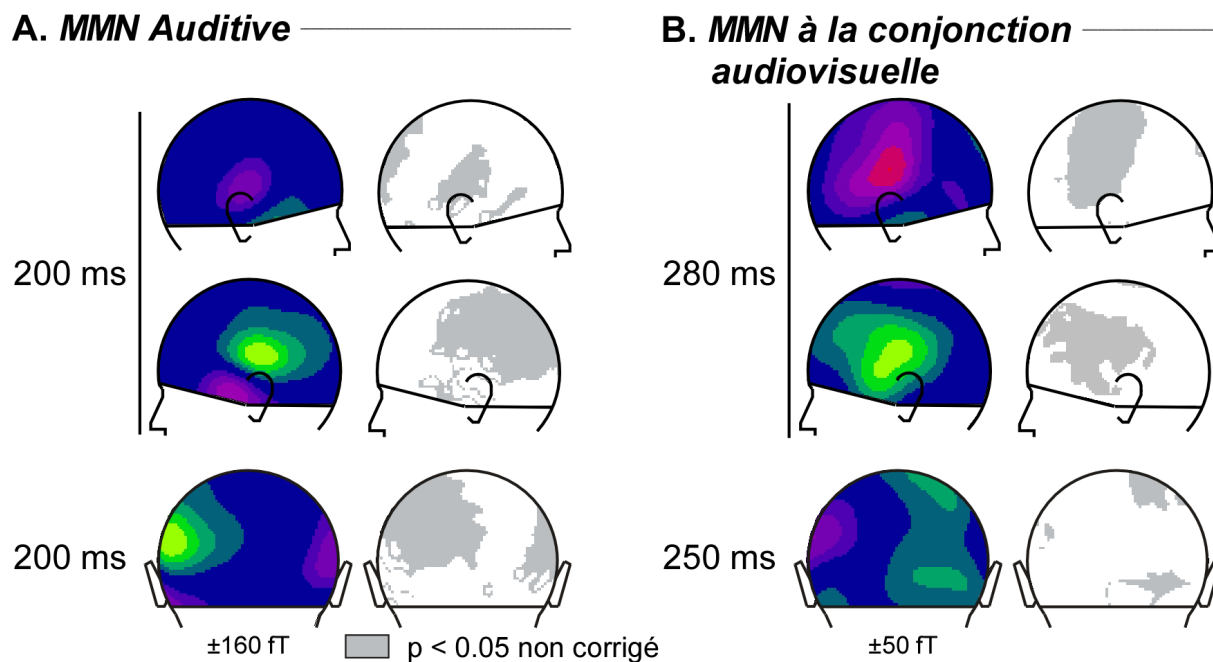


FIG. 16.2 – Topographie des MMN auditive (A) et à la conjonction audiovisuelle (B) chez le sujet S10

dans un bloc audiovisuel, et d'un stimulus déviant auditif dans un bloc contrôle auditif unimodal. Pour chacune des deux conditions, 600 stimuli (72 déviants) étaient présentés, répartis sur 6 blocs d'une durée approximative d'une minute. Les sujets étaient placés dans une situation la plus proche possible de celle des enregistrements MEG.

Les résultats de l'expérience comportementale sont rapportés dans le tableau 16.2 page suivante. Comparées à la détection auditive, les performances dans la détection d'une déviance à la conjonction audiovisuelle étaient assez médiocres puisqu'en moyenne les sujets ne détectaient que 67 % des cibles, avec un temps de détection relativement long. On peut également remarquer que, parmi les 3 sujets ayant participé aux deux expériences (S1, S9 et S10), les sujets S1 et S10, qui montraient une MMN à la conjonction, étaient également ceux dont les performances comportementales, dans la tâche de détection de la conjonction déviante, étaient les meilleures, comparées aux performances du sujet S9, qui ne montrait pas de MMN dans cette condition.

16.5 Discussion

Dans cette expérience, nous avons tenté de mettre en évidence l'existence d'une représentation intégrée de l'association particulière d'un trait auditif et d'un trait visuel en mémoire sensorielle. Les résultats sont moins clairs que ceux des expériences précédentes, dans lesquelles nous cherchions à mettre en évidence l'influence d'une régularité audiovisuelle sur la représentation d'une régularité auditive en mémoire sensorielle.

Sujets	Cibles manquées (%)		TR (ms)	
	Auditif	Audiovisuel	Auditif	Audiovisuel
S'1	2	47	406	750
S'2 (=S1)	0	2	407	673
S'3 (=S10)	2	16	425	628
S'4	2	50	378	778
S'5	5	23	390	771
S'6 (=S9)	0	56	461	739
Moyenne (\pm écart-type)	2 \pm 2	33 \pm 22	411 \pm 29	723 \pm 60

TAB. 16.2 – Performances et TR de détection dans l'expérience comportementale complémentaire.

Plusieurs facteurs peuvent expliquer la faiblesse des effets observés. Une première raison pourrait tenir à la technique d'enregistrement utilisée : les activités magnétiques générées par des dipôles radiaux sont relativement invisibles à la MEG. Or, lorsque l'on considère la topographie des différences de réponse obtenues dans les deux expériences précédentes, elles évoquent une origine plutôt radiale que tangentielle au scalp. La faible MMN à la conjonction observée dans la présente expérience, pourrait n'être constituée que de la composante tangentielle d'une activité principalement radiale.

Par ailleurs, il a été montré que l'amplitude de la MMN dépend de la force de trace mnésique, c'est-à-dire de la régularité et de la fréquence avec lesquelles le ou les standards sont présentés : ainsi l'amplitude de la MMN auditive est plus faible lorsque deux standards plutôt qu'un seul sont présentés (Winkler, Paavilainen & Näätänen, 1992) ou lorsque le standard varie légèrement sur un trait (Winkler et coll., 1990). Comme la mise en évidence d'une MMN à la conjonction de traits nécessitait l'utilisation de deux standards différents, on ne pouvait espérer obtenir une MMN d'amplitude importante.

Enfin, dans le domaine auditif, il a été montré que l'amplitude et la latence de la MMN sont corrélées à la capacité du sujet à détecter explicitement la déviance (Tiitinen et coll., 1994). Or notre expérience comportementale complémentaire montre que la plupart des sujets avaient beaucoup de difficulté à détecter la déviance à la conjonction audiovisuelle, tant en termes de performances qu'en termes de temps de traitement. De plus il semble exister un lien entre la force de la trace et les performances puisque les sujets les plus performants dans la tâche de détection étaient également ceux qui montraient la MMN à la conjonction la plus robuste. Dans ce cas, la détection d'une violation de la conjonction audiovisuelle pourrait être un processus automatique basé sur l'existence d'une représentation de la régularité audiovisuelle, indexée par la MMN à la conjonction d'amplitude assez faible que nous observons. Le nombre de sujets est cependant insuffisant pour conclure sur cette corrélation.

Toutefois, il se pourrait également que la tâche à réaliser pour détecter la conjonction de deux traits auditif et visuel repose sur des processus différents de la détection d'un trait simple dans une modalité et non sur l'existence d'une trace mnésique de la régularité audiovisuelle. En cela, la détection de la déviance à la conjonction audiovisuelle se distinguerait du cas purement auditif puisque que les déviations de conjonctions de traits

acoustiques donnent lieu à des MMN relativement robustes (Gomes et coll., 1997 ; Sussman et coll., 1998 ; Takegata et coll., 2001, 1999 ; Winkler et coll., 2005), ce qui suggère qu'il existe des processus automatiques de détection de la violation d'une conjonction analogues à ceux de la détection d'un trait dans le cas purement auditif. Cependant, à ma connaissance, le détection d'une déviation à une conjonction auditive n'a pas été testé comportementalement.

