

# ETUDE DE LA COMPREHENSION DE PHRASES: APPORT DES POTENTIELS EVOQUES CEREBRAUX

- THESE -

Sous la direction du Professeur  
Ulrich FRAUENFELDER

présentée à la  
Faculté de Psychologie et des Sciences de l'Education  
de l'Université de Genève  
pour l'obtention du grade de Docteur en Psychologie  
par

Elisabeth FONTENEAU  
de France

THESE N° 316  
Genève, 2003



# Remerciements

En premier lieu, je tiens à remercier le Prof. Uli Frauenfelder, mein "Doktor Vater", sans lequel ce travail n'aurait jamais vu le jour. Après m'avoir aiguillé sur le bon chemin il m'a laissé libre dans la réalisation de ma recherche. Le laboratoire de psycholinguistique expérimental s'est avéré un endroit productif et stimulant grâce à des échanges scientifiques multiples.

J'aimerais remercier les membres de mon jury de thèse: P. Bovet, J-F. Démonet et C. Michel pour avoir accepté de prendre le temps de lire ce manuscrit, et de le commenter.

Ce travail est le fruit de collaboration d'une part, avec le Laboratoire de Cartographie de fonctions cérébrales des Hopitaux Universitaires de Genève (Prof. Christoph Michel) et d'autre part, avec le département de linguistique (Prof. Luigi Rizzi).

Cette thèse a vu le jour sur le centre de Recherche Batelle à Carouge, rien que d'écrire ce nom me transporte dans une autre dimension ou potentiels évoqués rimaient avec "bricolage" et questions incessantes. La mise en place du laboratoire EEG de la Faculté de Psychologie et des Sciences et de l'Education ainsi que l'utilisation de la technique des PEs a nécessité la "synchronie" temporelle de billions (?) de neurones issus de deux cerveaux. Merci à Roberto Caldara pour avoir suivi ce chemin avec moi que se soit dans un box "à cage à poule" ou encore à une table des Fleurs de Lys.

La création des stimuli (chap 5 & 6) si bien contrôlées a été méticuleusement élaborée par Julie Palix. Le temps que tu as passé sur ce travail a permis que des liens d'amitiés s'installent entre nous.

Mon éloignement de Genève dans la partie finale de ce travail a conduit à de nombreux échanges par email et fax. Julie Franck a toujours répondu présente à cette surcharge de travail qu'était de relire mon manuscrit. Fréda, Odile, et Céline ont participé activement aux relectures finales, et je leur en suis très reconnaissante.

Afin de ne pas oublier les collègues et amis qui ont gravité autour de moi ainsi que de mon travail je n'en citerais aucun. Mais, ils m'ont permis tour à tour d'éclairer ma recherche, ou tout simplement d'égayer mes journées. Merci pour toute cette énergie positive qui a gravité autour de moi.

Je remercie le Prof. Heather van der Lely qui m'a recueillie dans son centre alors que j'étais encore en pleine période de rédaction. Toujours à l'écoute de mon dur labeur, et de la progression, elle m'a laissé le temps nécessaire pour que ce travail aboutisse.

Pour finir, j'aimerais remercier tous les participants ayant volontairement subi des shampoings après mes expériences.

Cette recherche a bénéficié du soutien du programme plurifacultaire "langage et communication". Le projet "carrière académique" ainsi que la Société Académique de Genève ont soutenu financièrement cette thèse dans sa période finale.

Ce travail est dédié à mes parents et à ma famille.

*"A yes, the word!  
The word is as central to psycholinguistics as the cell is to biologists."*

(Balota, 1994)

## Résumé

Les considérations théoriques ainsi que les données empiriques issues d'études psycholinguistiques suggèrent que la classe d'un mot peut permettre de structurer les représentations et/ou les processus d'accès au lexique. Nous explorons ce thème à l'aide d'une technique d'imagerie cérébrale, les potentiels évoqués ou PEs. Si des représentations et/ou des processus d'accès séparés sont activés lors de la reconnaissance des deux classes, alors des patterns électrophysiologiques distincts sur le scalp doivent émerger de l'activité cérébrale. Une première expérience (chapitre 4) indique qu'une composante N280 - négativité maximale sur les régions antérieures gauches - est différemment répartie sur le scalp, 280 ms après la présentation des mots de classe fermée par rapport à celle enregistrée pour les mots de classe ouverte. De plus, la latence de cette composante s'est avérée corrélée avec la fréquence et la longueur mais aussi avec la classe de mots. Puisque les mots de classe fermée sont plus fréquents et plus courts que ceux de classe ouverte, les interprétations issues de cette expérimentation restent irrésolues. L'hypothèse selon laquelle la composante N280 reflète l'activation spécifique de représentations et/ou des processus d'accès lexical séparés pour les mots de classe fermée est testée dans deux nouvelles expériences à l'aide d'un corpus de mots dont les propriétés lexicales distributionnelle (fréquence d'occurrence) et physique (longueur) sont à la fois contrôlées et manipulées. Alors que pour des présentations en isolation (chapitre 6), les différences de topographie de la composante N280 émergent entre les deux classes de mots pour des basses fréquences (< 400 occurrences/million), à l'opposé, en contexte de phrase (chapitre 5) celles-ci sont observées pour les mots de fréquences élevées (> 400 occurrences/million). L'absence d'effet de topographie précoce pour les basses fréquences en contexte de phrase (chapitre 5) est reliée aux résultats d'un test de *Cloze Procedure* (chapitre 7) indiquant que le contexte employé avant nos mots critiques (chapitre 5) ajoutent des contraintes supplémentaires sur la classe ainsi que l'identité des mots de classe ouverte de basse fréquence. En outre, quel que soit le mode de présentation (isolation ou contexte) les décalages de latence de la composante N280 sont dus à l'effet de classe pour des mots de haute fréquence suggérant des processus d'accès plus rapide pour les mots de classe fermée que pour ceux de classe ouverte de haute fréquence.

De manière plus générale, une imbrication étroite entre les propriétés lexicales (fréquence et longueur) et la classe de mot est observée: la dissociation de leurs effets suggère que les réseaux neuronaux impliqués dans les processus d'identification sont différemment sensibles aux types d'informations apportés par la classe de mots.

# Sommaire

INTRODUCTION .....	1
--------------------	---

## CHAPITRE 1

MODELES PSYCHOLINGUISTIQUES ET CLASSE DE MOTS .....	7
---	---

1 Différentes variables influençant la reconnaissance d'un mot .....	8
2 Modèles d'accès au lexique.....	12
2.1 Modèle à deux lexiques de Glanzer .....	13
2.2 Modèles à un lexique mental .....	14
2.2.1 Modèle sériel de Forster.....	14
2.2.2 Modèle parallèle de Morton .....	16
2.2.3 Nature des processeurs .....	18
2.2.3.1 Autonomie de Forster .....	18
2.2.3.2 Interactivité de Morton .....	20
2.3 Conclusions .....	21
3 La classe de mot: une distinction catégorielle? .....	22
3.1 Terminologie.....	22
3.2 Des propriétés linguistiques distinctives.....	24
4 Un lexique unique ou deux lexiques distincts? .....	27
4.1 Arguments en faveur d'une séparation en deux lexiques .....	28
4.1.1 Travaux précurseurs sur la production .....	29
4.1.1.1 Comparaison adultes sains et patients aphasiques.....	30
4.1.1.2 Hypothèse de Garrett (1975).....	32
4.1.2 Travaux en perception .....	33
4.1.2.1 Détection de lettres .....	33
4.1.2.2 Recherche de mots.....	34
4.1.2.3 Acquisition et classe de mots.....	36
4.1.2.4 Hypothèse de Friederici (1980) .....	37
4.1.3 Travaux de Bradley et collaborateurs .....	40
4.1.3.1 Interférence visuelle .....	40
4.1.3.2 Présentation tachistoscopique .....	41
4.1.3.3 Hypothèse de Bradley: classe dépendante (1978) .....	42
4.2 Arguments en faveur d'une architecture commune: un seul lexique .....	44
4.2.1 Problème de réplcation des résultats de Bradley et collaborateurs.....	44
4.2.1.1 Effet de fréquence sur les mots de classe fermée .....	45
4.2.1.2 Hypothèse de Gordon & Carramazza (1982).....	47
4.2.2 Hypothèses alternatives .....	48
4.2.2.1 Hypothèse de Shillcock & Bard (1993) .....	48
4.2.2.2 Hypothèse de Herron & Bates (1997) .....	49
5 Conclusions.....	50

## CHAPITRE 2

POTENTIELS EVOQUES ET CLASSE DE MOTS .....	54
1 Bases neuronales de l'activité électroencéphalographique .....	55
1.1 Notion de dipôle unitaire .....	56
1.2 Dipôles équivalents et propriétés des neurones .....	57
2 Les Potentiels Evoqués (PEs): notions de base .....	59
2.1 Définition .....	59
2.2 Aspects méthodologiques .....	59
2.3 Contribution à l'étude des processus psycholinguistiques.....	61
2.3.1 Avantages.....	62
2.3.2 Limites .....	62
2.3.3 PEs et violations de règles linguistiques .....	64
3 Apport des PEs dans l'étude du traitement des deux classes de mots .....	65
3.1 Les travaux de Kutas & Hillyard (1983).....	65
3.2 Variabilité des résultats chez les sujets sains.....	67
3.2.1 Aspect méthodologique: la fenêtre précoce et le nombre d'électrodes.....	69
3.2.2 La nature des stimuli: rôle de la fréquence et longueur.....	71
3.2.2.1 La fréquence et la longueur affectent-elles la latence de la N280? .....	71
3.2.2.2 Et si la fréquence et la longueur étaient contrôlées? .....	78
3.2.3 Isolation vs. contexte? .....	79
3.3 Déficits et développement des patterns PEs normaux.....	82
3.3.1 Description des PEs de patients aphasiques de Broca .....	82
3.3.2 Avant 20 mois, la composante N280 n'est pas latéralisée .....	84
3.3.3 Importance de l'âge d'acquisition du langage .....	86
4 Conclusions.....	86
5 Problématique .....	88
5.1 Caractériser les représentations/processus pour les deux classes de mots.....	90
5.2 Etudier l'effet de classe de mots en fonction de la bande de fréquence d'occurrence .....	92
5.3 Etudier l'effet de contexte sur les deux classes de mots .....	93
5.4 Apport méthodologique: les analyses spatio-temporelles .....	94

## CHAPITRE 3

METHODES D'ANALYSE DES PEs: APPLICATION A LA N400 .....	96
1 Introduction.....	97
1.1 Limites des analyses de tracés .....	98
1.1.1 Dimension temporelle: définir la fenêtre d'analyse.....	98
1.1.2 Dimension spatiale .....	100
1.2 Quelles solutions adopter? .....	101
1.2.1 Analyses spatio-temporelles basées sur la cartographie des PEs .....	101
1.2.1.1 Le <i>Global Field Power</i> : GFP .....	102
1.2.1.2 La <i>Global Dissimilarity</i> : GD .....	104
1.2.2 Micro-états fonctionnels et segmentation.....	105

1.3	Segmentation et composante N400 .....	107
1.4	Objectifs .....	108
2	Méthode.....	110
2.1	Sujets .....	110
2.2	Stimuli.....	110
2.3	Procédure .....	111
2.4	Enregistrement de l'EEG.....	111
3	Analyses des PEs .....	112
3.1	Analyse des tracés .....	113
3.2	Analyse de segmentation adaptée.....	115
4	Résultats.....	117
4.1	Résultats comportementaux .....	117
4.2	Analyse des tracés .....	118
4.2.1	Visualisation des tracés.....	118
4.2.2	Analyse topographique par région d'intérêt .....	119
4.2.3	Analyse de latence de pics .....	120
4.3	Analyse spatio-temporelle .....	121
4.3.1	Détermination des micro-états.....	121
4.3.2	Détermination des micro-états spécifiques .....	123
4.3.3	Analyse topographique par région d'intérêt .....	124
4.3.4	Evolution temporelle des micro-états .....	126
5	Discussion.....	126
5.1	La N400 composée d'un ou de deux processus ?.....	127
5.2	Topographie des deux segments .....	128
5.3	Interprétation fonctionnelle de ces segments.....	129
5.4	Amplitude de l'effet N400 .....	131
5.5	Contributions méthodologiques .....	132
5.5.1	Référence moyenne.....	132
5.5.2	La segmentation .....	133
6	Conclusions.....	134

## CHAPITRE 4

### EFFET DE CLASSE ET PROPRIETES LEXICALES NON CONTROLEES ..... 135

1	Introduction.....	136
2	Méthode.....	139
2.1	Sujets .....	139
2.2	Stimuli.....	139
2.3	Procédure .....	140
2.4	Enregistrement de l'EEG.....	140
3	Résultats.....	141
3.1	PEs et classe de mots.....	141
3.1.1	Analyse des tracés .....	141
3.1.1.1	Visualisation des tracés.....	141
3.1.1.2	Analyse topographique par région d'intérêt .....	142
3.1.1.3	Analyse de latence de la N280 .....	146
3.1.2	Analyse spatio-temporelle .....	147

3.1.2.1 Détermination des micro-états .....	147
3.1.2.2 Détermination des micro-états spécifiques .....	148
3.1.2.3 Analyse topographique par région d'intérêt .....	149
3.1.2.4 Evolution temporelle des micro-états .....	154
3.2 PEs et effet de la catégorie grammaticale .....	154
3.2.1 Topographie du segment 7 .....	155
3.2.2 Latence du segment 7 .....	157
4 Discussion.....	159
4.1 Tracés et effet de classe de mots .....	159
4.1.1 Effet précoce visuel: la P2 .....	159
4.1.2 Effet précoce: la N280.....	160
4.1.3 Effet tardif: la N400 .....	161
4.1.4 Effet tardif: la LAN .....	162
4.2 Apport de la segmentation .....	162
4.3 Interprétation de la latence de la N280 et des propriétés lexicales .....	165
4.4 Les prépositions, des mots de classe fermée ? .....	166
5 Conclusions.....	168

## CHAPITRE 5

### EFFET DE CLASSE ET PROPRIETES LEXICALES CONTROLEES EN CONTEXTE..... 169

1 Introduction.....	170
2 Méthode.....	174
2.1 Sujets .....	174
2.2 Stimuli.....	174
2.3 Procédure .....	177
2.4 Enregistrement de l'EEG.....	177
3 Résultats.....	178
3.1 Comparaison CF vs. CO.....	178
3.1.1 Analyse des tracés .....	178
3.1.1.1 Visualisation des tracés.....	178
3.1.1.2 Analyse de topographie par région d'intérêt .....	179
3.1.1.3 Analyse de latence de la N280 .....	180
3.1.2 Analyse spatio-temporelle .....	181
3.1.2.1 Détermination des micro-états .....	181
3.1.2.2 Détermination des micro-états spécifiques .....	182
3.1.2.3 Analyse topographique par région d'intérêt .....	183
3.1.2.4 Evolution temporelle des micro-états .....	185
3.1.3 Synthèse CF vs. CO .....	185
3.2 Comparaison CF vs. CO haute fréquence .....	186
3.2.1 Analyses des tracés .....	186
3.2.1.1 Visualisation des tracés.....	186
3.2.1.2 Analyse de topographie par région d'intérêt .....	187
3.2.1.3 Analyse de latence de la N280 .....	188
3.2.2 Analyse spatio-temporelle .....	188
3.2.2.1 Détermination des micro-états .....	188
3.2.2.2 Détermination des micro-états spécifiques .....	189



3.2.2.3	Analyse topographique par région d'intérêt .....	189
3.2.2.4	Evolution temporelle des micro-états .....	191
3.2.3	Synthèse CF-H vs. CO-H .....	192
3.3	Comparaison CF vs. CO basse fréquence.....	193
3.3.1	Analyse des tracés .....	193
3.3.1.1	Visualisation des tracés.....	193
3.3.1.2	Analyse de topographie par région d'intérêt .....	194
3.3.1.3	Analyse de latence de la N280 .....	194
3.3.2	Analyse spatio-temporelle .....	195
3.3.2.1	Détermination des micro-états .....	195
3.3.2.2	Détermination des micro-états spécifiques .....	195
3.3.2.3	Analyse topographique par région d'intérêt .....	196
3.3.2.4	Evolution temporelle des micro-états .....	197
3.3.3	Synthèse CF-B vs. CO-B .....	198
4	Discussion générale.....	198
4.1	L'effet de classe ne repose pas uniquement sur les propriétés lexicales.....	198
4.1.1	Les analyses des tracés .....	198
4.1.2	L'analyse de segmentation adaptée.....	200
4.2	Hypothèses quant à l'interprétation fonctionnelle des effets .....	203
4.3	L'effet de classe varie avec la fréquence .....	205
5	Conclusions.....	206

## CHAPITRE 6

EFFET DE CLASSE ET PROPRIETES LEXICALES CONTROLEES EN ISOLATION.....		207
1	Introduction.....	208
2	Méthode.....	210
2.1	Sujets .....	210
2.2	Matériel.....	210
2.3	Procédure .....	210
2.4	Enregistrement de l'EEG.....	211
3	Résultats.....	212
3.1	Comparaison CF vs. CO.....	212
3.1.1	Analyses des tracés .....	212
3.1.1.1	Visualisation des tracés.....	212
3.1.1.2	Analyse topographique par région d'intérêt .....	213
3.1.1.3	Analyse de latence .....	215
3.1.2	Analyse de segmentation adaptée.....	217
3.1.2.1	Détermination des micro-états .....	217
3.1.2.2	Détermination des micro-états spécifiques .....	218
3.1.2.3	Analyse topographique par région d'intérêt .....	219
3.1.2.4	Evolution temporelle des micro-états .....	222
3.1.3	Synthèse CF vs. CO .....	222
3.2	Comparaison CF vs. CO haute fréquence .....	223
3.2.1	Analyse des tracés .....	223
3.2.1.1	Visualisation des tracés.....	223
3.2.1.2	Analyse topographique par région d'intérêt .....	224
3.2.1.3	Analyse de latence .....	225

3.2.2 Analyse de segmentation adaptée.....	226
3.2.2.1 Détermination des micro-états .....	226
3.2.2.2 Détermination des micro-états spécifiques .....	226
3.2.2.3 Analyse topographique par région d'intérêt .....	227
3.2.2.4 Evolution temporelle des micro-états .....	228
3.2.3 Synthèse CF vs. CO haute fréquence.....	228
3.3 Comparaison CF vs. CO basse fréquence.....	228
3.3.1 Analyse des tracés .....	228
3.3.1.1 Visualisation des tracés.....	229
3.3.1.2 Analyse topographique par région d'intérêt .....	230
3.3.1.3 Analyse de latence .....	233
3.3.2 Analyse de segmentation adaptée.....	234
3.3.2.1 Détermination des micro-états .....	234
3.3.2.2 Détermination des micro-états spécifiques .....	234
3.3.2.3 Analyse topographique par région d'intérêt .....	235
3.3.2.4 Evolution temporelle des micro-états .....	238
3.3.3 Synthèse CF vs. CO basse fréquence .....	239
4 Discussion.....	239
4.1 Effet de classe en isolation: apport de la segmentation adaptée.....	239
4.2 Interprétation fonctionnelle des effets .....	241
4.3 Effet de classe en fonction du rang de fréquence .....	243
5 Conclusions.....	244

## CHAPITRE 7

LE CONTEXTE AFFECTE-T-IL IDENTIQUEMENT LES DEUX CLASSES DE MOTS? .....	245
1 Introduction.....	246
2 Test de complétion ( <i>Cloze Procedure</i> , CP) .....	248
2.1 Méthode .....	248
2.2 Résultats.....	248
3 Données électroencéphalographiques .....	250
3.1 Analyses.....	250
3.2 Résultats.....	252
3.2.1 CF vs. CO.....	252
3.2.2 CF vs. CO haute fréquence .....	255
3.2.3 CF vs. CO basse fréquence.....	255
4 Discussion.....	256
4.1 Test de complétion.....	256
4.2 Observation des effets de classe pour les deux expériences .....	258
4.3 Effet général du contexte sur les PEs .....	259
4.4 Effet spécifique sur la composante P100: 0-150 ms.....	260
4.5 Effet spécifique sur la composante N280: 250-350 ms.....	261
5 Conclusions.....	263

## CHAPITRE 8

DISCUSSION GENERALE ET CONCLUSIONS .....	264
1 Récapitulatif.....	265
2 Modèle neurocognitif de Friederici 2002 .....	266
3 Analyse de nos résultats à la lumière du modèle de Friederici.....	268
3.1 Mise en correspondance des segments et des phases .....	268
3.2 Accès lexical sans apport de contexte .....	270
3.3 Accès lexical avec apport de contexte .....	272
4 Contribution de nos recherches.....	273
4.1 Adaptation du modèle neurocognitif de Friederici .....	273
4.2 La segmentation adaptée.....	276
5 Perspectives.....	276
6 Conclusions.....	278
BIBLIOGRAPHIE .....	279
ANNEXES .....	292

## INTRODUCTION

---

Comment lisons-nous? Les mots et les phrases lus sont compris au fur et à mesure de leur perception, c'est à dire en temps réel. Cette capacité est possible car la reconnaissance des mots donne directement accès à un riche ensemble d'informations telles que les propriétés orthographiques, phonologiques, morphologiques, syntaxiques et sémantiques. L'exploitation adéquate de ces informations va permettre au lecteur d'aboutir très rapidement à une représentation mentale de l'énoncé. Ainsi, accéder à la reconnaissance d'un mot semble constituer une étape primordiale dans l'activité de compréhension du langage écrit (Segui, 1989).

Pour expliquer de telles performances et les mécanismes impliqués dans une telle activité, les psycholinguistes se réfèrent à la métaphore du "lexique mental". Ce concept, originellement introduit par Oldfield (1966), désigne une trace en mémoire à long terme où sont stockées les diverses connaissances linguistiques relatives à chacun des mots (informations orthographiques, phonologiques, morphologiques, syntaxiques et sémantiques). Les processus permettant d'accéder à ces informations sont considérés comme automatiques, en ce sens que le sujet n'est ni conscient des différentes étapes qui les constituent, ni du résultat final. En outre, la rapidité de leur mise en œuvre est exceptionnelle. En effet, un lecteur moyen confronté à l'identification d'environ 150 à 200 mots par minute (Ferrand, 2001) parvient à retrouver les informations pertinentes à l'intérieur d'un ensemble de plus de 60000 entrées lexicales en moins d'une seconde (Pinker, 1999). En d'autres termes, un sujet doit donc consulter son lexique mental deux ou trois fois par seconde lors de la lecture d'une phrase.

La problématique développée dans cette thèse s'apparente à ce thème en questionnant à la fois la nature des représentations ainsi que celle des processus d'accès lexicaux. Cette thèse s'inscrit dans la mouvance des neurosciences cognitives, et a pour but d'étudier l'organisation et le fonctionnement des propriétés neuronales lors de la compréhension du langage. La rapidité de ces processus nécessite d'une part de postuler l'existence de plusieurs principes organisateurs efficaces concourant à la structuration de ce grand "dictionnaire", et d'autre part d'utiliser une mesure permettant de suivre le décours temporel de l'activation des différentes étapes.

## Principes organisateurs du lexique mental

L'étude des mécanismes psychologiques impliqués dans la reconnaissance des mots constitue l'un des principaux pôles de recherche en psycholinguistique, l'objectif principal étant de caractériser les procédures et les représentations internes impliquées dans la reconnaissance des mots. Visant à clarifier l'activité de mise en relation du stimulus perceptif visuel et des représentations internes stockées dans le lexique mental, les études psycholinguistiques tentent de répondre aux questions portant sur la nature des représentations mentales et sur les mécanismes impliqués dans l'accès à ces représentations. Dans cette perspective, l'influence des différentes propriétés linguistiques sur la rapidité de reconnaissance des mots a été largement étudiée. Sont incluses les propriétés physiques (longueur, composition segmentale), les propriétés distributionnelles (la fréquence d'occurrence du mot et des parties qui le composent, pour une définition voir chapitre 1 section 1), et les propriétés linguistiques (orthographe, phonologie, morphologie, forme syntaxique, sémantique, voir définition chapitre 1 section 3.2). En particulier, l'effet de la fréquence d'occurrence, constitue l'un des résultats les plus robustes de la littérature. De nombreux travaux corroborent l'hypothèse selon laquelle plus un mot est fréquent, plus il est reconnu ou identifié rapidement (Glanzer & Ehrenreich, 1979; Howes & Solomon, 1951; Jacoby & Dallas, 1981; Rao & Proctor, 1984). Par conséquent, l'implémentation de cette propriété distributionnelle dans les contraintes architecturales des modèles est inévitable. Cependant, les effets de cette propriété distributionnelle sur la nature des représentations ou encore sur les processus d'accès au lexique, ne sont pas appréhendés de manière consensuelle par les différentes approches théoriques. Par exemple, si Glanzer & Ehrenreich (1979) proposent que la fréquence d'occurrence d'un mot affecte la nature de la représentation - les mots de haute fréquence étant stockés dans un lexique mental séparé de celui des mots de basses fréquences - pour Forster (1976) en revanche, la fréquence d'occurrence affecte la manière dont les représentations mentales vont être vérifiées lors des processus d'accès au lexique - tous les mots étant stockés dans le même lexique mental, mais examinés en deux temps, les mots de haute fréquence étant traités en premier lieu (pour une discussion détaillée voir chapitre 1 section 2.1).

Les propriétés linguistiques ont aussi été implémentées dans les modèles de compréhension du langage. Dès 1975, Bolinger suppose l'existence éventuelle d'une dichotomie dans la

classification linguistique et la fonctionnalité des mots: alors que certains mots participent à l'élaboration de la structure syntaxique - les mots de classe fermée - d'autres apportent les informations sémantiques - les mots de classe ouverte <sup>1</sup> (Bolinger, 1975). Comme pour les hypothèses relatives aux propriétés distributionnelles, les modèles implémentant ces propriétés linguistiques s'opposent. Pour certains, les deux classes de mots sont stockées dans deux lexiques mentaux séparés (Bradley, 1978; Friederici & Schoenle, 1980; Garrett, 1975; voir chapitre 1 section 4.1) alors que pour d'autres (Gordon & Caramazza, 1982; Herron & Bates, 1997; Shillcock & Bard, 1993; voir chapitre 1 section 4.2), la classe de mot affecte les processus d'accès au lexique, et non la nature des représentations. L'étude des bases de données (Content, Mousty, & Radeau, 1990; Francis & Kucera, 1982) indique que les diverses propriétés d'un mot sont fortement interconnectées. Par exemple, si l'on s'intéresse au cas particulier de la classe de mot, il apparaît que les mots de classe fermée sont aussi plus fréquents et plus courts que les mots de classe ouverte dans de nombreuses langues différentes. Ainsi, les propriétés physiques (la longueur) et distributionnelles (la fréquence) sont fortement liées à la classe de mot. Malgré cette interconnexion entre les propriétés d'un mot (classe de mots, longueur et fréquence, par exemple), très peu d'études ont analysé leurs effets simultanés. En effet, les modèles psycholinguistiques actuels s'intéressent à une, voire deux de ces propriétés, mais rarement plus.

Dans cette thèse, nous avons décidé de retenir une propriété linguistique - la classe des mots - comme éventuel principe organisateur du lexique mental. Comme nous venons de l'évoquer, il semble difficile d'étudier les effets de la classe de mot en délaissant les autres propriétés lexicales qui lui sont liées. Ainsi, nous proposons de tester les effets de la classe de mot sur les processus d'accès au lexique mental, tout en contrôlant et manipulant les propriétés physiques (longueur) et distributionnelles (fréquence d'occurrence dans la langue). Dans un souci de clarté, nous allons dans la suite de cette thèse regrouper sous le terme de **propriétés lexicales**, les propriétés physique (longueur) et distributionnelle (fréquence d'occurrence) des mots.

---

<sup>1</sup> Une définition précise des deux classes de mots est décrite dans le chapitre 1 section 3

Plus spécifiquement, les questions adressées dans cette thèse sont les suivantes:

- Possédons-nous plusieurs lexiques mentaux? Si oui, les mots de classes fermée et ouverte sont-ils représentés dans deux lexiques séparés?
- La classe de mots peut-elle influencer l'accès au lexique mental indépendamment des propriétés lexicales? Si oui, comment?

Trois différents axes vont être développés dans cette recherche. Le principal objectif de cette thèse est de dissocier les effets de la classe de mots de ceux liés aux propriétés lexicales (telles que la fréquence et la longueur). L'activité électrique des deux classes de mots sera analysée lorsque d'une part, les propriétés lexicales ne sont pas contrôlées et d'autre part, lorsque ces dernières le sont. En désambiguïsant les effets des propriétés lexicales et de classe de mots, nous allons pouvoir spécifier si des représentations et/ou des processus d'accès lexical différents sont activés lors de la reconnaissance des deux classes de mots. Le second objectif est d'analyser la reconnaissance des mots de classe ouverte et fermée en présence et en absence de contexte. Ainsi les mots contrôlés en fréquence et en longueur vont être introduit soit dans un contexte de phrase soit dans une liste de mot. En comparant l'effet de classe dans ces deux conditions expérimentales, nous allons pouvoir cibler plus précisément l'influence, si elle existe, du contexte linguistique sur la reconnaissance des deux classes de mots. Enfin, nous proposons d'apporter une pierre constructive dans les analyses statistiques portant sur les PEs. Alors que l'accroissement constant du nombre de capteurs sur le scalp apporte de nombreux problèmes, nous tenterons d'élucider ceux-ci par l'apport d'une nouvelle méthode d'analyse: la segmentation adaptée (Lehmann & Skrandeis, 1980).

## Pourquoi étudier le langage écrit plutôt que la parole?

La communication nécessite la conversion d'une idée ou concept en un code accessible à l'interlocuteur. L'un de ces codes est le langage, et ses manifestations principales sont la parole et le langage écrit. La lecture et l'écriture sont dérivées de notre compétence à comprendre et produire la parole: l'assignation des sons à des symboles conventionnels se fait en accord avec une communauté linguistique qui utilise ce système particulier de lecture et d'écriture. Malgré le fait que la parole soit le premier mode commun des processus langagiers, la psycholinguistique s'est plutôt focalisée sur le traitement du langage écrit.

Il existe des similarités entre la parole et le langage écrit. En effet, la compréhension à l'oral comme à l'écrit vise un seul et même but: l'élaboration d'une représentation sémantique de l'énoncé (du mot, de la phrase, du discours). Tandis qu'à l'oral le système linguistique opère sur des signaux acoustiques codés sous formes de représentations phonologiques, le système de compréhension de l'écrit opère sur des signaux visuels codés sous la forme d'une représentation orthographique. Bien qu'il semble relativement pertinent de postuler que les systèmes responsables du traitement phonologique et orthographique soient différents, les composantes des traitements ultérieurs, c'est à dire syntaxique et sémantique, sont supposées identiques pour les deux modalités de présentation (Levelt, Roelofs, & Meyer, 1999).

Dans le vaste domaine de la psycholinguistique, nous avons vu l'intérêt de focaliser nos recherches sur le domaine du langage écrit. D'une part, la technique usitée dans cette thèse - les potentiels évoqués (PEs) - nécessite une synchronisation de l'enregistrement avec la présentation des stimuli. Ainsi, la recherche d'une frontière d'un mot en milieu de phrase dans un signal acoustique de parole continue aurait pu engendrer des problèmes dans les résultats. D'autre part, alors que la lecture d'un mot peut s'effectuer sur un seul point de fixation (si celui-ci n'est pas trop long), l'écoute d'un stimulus auditif nécessite un traitement de gauche à droite défini par la longueur de présentation du stimulus lui-même. Les différences temporelles entre les mots peuvent, par conséquent, induire des modulations dans les caractéristiques des composantes (latence et/ou distribution sur le scalp) qui ne sont pas forcément liées aux conditions expérimentales manipulées. Bien que techniquement possible (Friederici, Pfeifer, & Hahne, 1993; Holcomb & Neville, 1990; Kutas & van Petten, 1994) nous sommes tout à fait conscients du manque d'études se penchant sur l'épineux problème de la reconnaissance des mots parlés à l'aide des PEs.

## Structure de la thèse

Dans le **chapitre 1**, nous présentons succinctement trois modèles d'accès au lexique opposés quant à leurs approches relatives à la nature des représentations activées (un ou plusieurs lexiques mentaux), mais aussi à la nature de la recherche lexicale (active/sérielle versus passive/parallèle). Ces modèles s'opposent aussi sur la manière dont les informations des multiples niveaux d'analyses (orthographique, morphologique, sémantique, syntaxique) peuvent être ou non utilisées lors de la reconnaissance d'un mot (modulaire versus interactif).



Dans un second temps, une revue de la littérature de la psychologie cognitive en chronométrie mentale nous permet de rendre compte des hypothèses majeures liées à la dissociation des classes de mots. Le **chapitre 2** est axé sur l'apport d'une technique d'imagerie cérébrale à l'étude de la distinction entre classes de mots. Ainsi, après avoir présenté la technique des potentiels évoqués (ou PEs), nous exposons les différentes données issues des recherches qui l'utilisent, ainsi que les limites qu'elles présentent. La dernière partie de ce chapitre relate les trois principaux axes de recherche que nous proposons dans ce travail. Les quatre chapitres suivants exposent les parties expérimentales réalisées. Le **chapitre 3** décrit les deux méthodes d'analyse utilisées (l'analyse de tracés ainsi que l'analyse de segmentation adaptée proposée). Nous avons délibérément appliqué celles-ci à un paradigme expérimental connu, celui de la présentation de violations sémantiques. Alors que ce type d'erreur suscite l'occurrence d'une composante N400 désormais bien définie, l'application en parallèle des deux méthodes sur ces données nous permettra de mettre en évidence les intérêts et les faiblesses respectifs. Le **chapitre 4** analyse l'effet de classe de mots sur les PEs lorsque ceux-ci sont intégrés dans un contexte de phrase. De plus, une analyse séparée (noms, verbes, adverbes, prépositions, quantificateurs, non-quantificateurs) sera effectuée pour chaque catégorie grammaticale, afin de déterminer si les composantes obtenues pour chaque classe de mots sont bien homogènes, tant au niveau des distributions sur le scalp, que pour les latences d'apparition. Le **chapitre 5** permet une généralisation des résultats obtenus dans le chapitre 4 en contrôlant les propriétés lexicales (fréquence et longueur) pour les deux classes de mots. Ainsi, nous utiliserons des mots de classe fermée et ouverte présentant des propriétés lexicales équivalentes dans un contexte de phrase. Les composantes distinctives obtenues dans le chapitre 4 pourront alors être affiliées soit à des effets de classe soit à des effets liés aux propriétés lexicales. Le **chapitre 6** présente les mots utilisés dans le chapitre 5, sans contexte de phrase. Ainsi, nous allons pouvoir mettre en évidence s'il existe des composantes spécifiques à la reconnaissance des mots hors contexte pour chacune des catégories observées. Le **chapitre 7** examine les PEs des deux classes de mots obtenus en contexte de phrase ainsi qu'en isolation (chapitre 5 et 6). Puisque les mêmes mots sont utilisés dans les deux expériences, nous pourrions caractériser précisément si les PEs des mots de classe fermée et ouverte sont affectés différemment par le contexte linguistique. Enfin, le **chapitre 8** résume les différents résultats de notre travail que nous tentons de discuter à la lumière des résultats de la littérature.

# CHAPITRE 1

## MODELES PSYCHOLINGUISTIQUES ET CLASSE DE MOTS

---

### *Résumé*

Dans ce chapitre nous commençons par exposer succinctement les différents facteurs influençant la reconnaissance d'un mot. Puisque la propriété distributionnelle qu'est la fréquence d'occurrence du mot semble affecter la rapidité ainsi que la justesse de la reconnaissance d'un mot, nous présentons trois principaux modèles psycholinguistiques s'opposant sur la manière dont l'effet de fréquence est expliqué. La distinction peut s'effectuer sur la nature des représentations, les mots de haute fréquence et de basse fréquence étant stockés dans des lexiques mentaux séparés, ou encore sur les processus d'identification d'un mot, les mots étant alors organisés selon leur fréquence dans un même et unique lexique. Cependant, il semble intéressant de noter que de nombreux facteurs sont directement liés à la fréquence, et par conséquent ces premiers modèles peuvent être quelque peu incomplets quant à l'implication des autres propriétés lors la reconnaissance d'un mot. Ainsi, la classe de mot est une propriété linguistique du mot qui est corrélée avec la fréquence et la longueur. Alors que l'étude de ces 3 propriétés n'est pas systématique, comment lever l'ambiguïté entre les effets de fréquence, de celui de classe de mot, ou encore de longueur?

Après une définition précise de la classe de mot, nous examinons les arguments empiriques issus de la psychologie cognitive sur l'étude de la distinction entre les deux classes de mots. Différentes hypothèses sont émises se différenciant sur la nature des représentations et sur la manière dont les processeurs utilisent les informations d'un mot. Dans un premier temps, nous présentons les différents résultats en faveur d'une séparation dans le stockage des représentations des deux classes de mots, ainsi que les différentes hypothèses émanant de ces résultats. Par la suite, les arguments inconsistants avec un stockage séparé seront présentés ainsi que les hypothèses qui en découlent.

## 1 Différentes variables influençant la reconnaissance d'un mot

Les recherches portant sur l'étude des processus de compréhension du langage ont très souvent utilisé les termes "identification lexicale", "accès au lexique" ou encore "reconnaissance des mots", sans aucun consensus terminologique (Tanenhaus & Lucas, 1987). Dans cette introduction, nous voulons préciser les termes employés et indiquer les étapes de traitement référencées pour chacun. Le terme **identification lexicale** sera réservé aux différentes étapes depuis le décodage de l'entrée sensorielle jusqu'à l'appariement avec les représentations internes stockées dans le lexique mental. La seconde opération, **accès au lexique**, impliquera l'activation des différentes informations (orthographiques, phonologiques, sémantiques, syntaxiques) associées avec cette entrée. Le dernier terme, **la reconnaissance d'un mot**, fera référence aux processus qui englobent l'identification lexicale ainsi que l'accès au lexique.

L'un des principaux objectifs de la psycholinguistique est de déterminer les facteurs pouvant influencer la rapidité ou encore la précision de l'identification d'un mot. Dès 1951, Howes & Solomon observent que les mots les plus usités, par conséquent de haute fréquence, sont reconnus plus facilement et plus rapidement que les mots de basse fréquence (Howes & Solomon, 1951). Depuis cette expérience princeps, l'effet de la fréquence sur l'identification d'un mot n'est plus remis en question et la plupart des modèles actuels font intervenir ce facteur dans la reconnaissance d'un mot. Le débat s'est donc déplacé vers la détermination du locus de l'effet de fréquence, et de la manière dont cette variable affecte les processus d'identification.

Alors que la fréquence d'occurrence semble être un facteur primordial pour la reconnaissance d'un mot, d'autres facteurs peuvent aussi influencer la rapidité d'un tel mécanisme. Rubin (1980) passe en revue quelques 51 propriétés des mots dont certaines sont étroitement liées ou corrélées avec la fréquence d'occurrence d'un mot.

Après un bref descriptif de l'effet de fréquence, nous allons passer en revue quelques propriétés d'un mot qui nous semblent pertinentes pour la suite de notre exposé.

- **La fréquence d'occurrence**

Rubin (1980) note qu'un des meilleurs prédicteurs pour les performances dans un paradigme expérimental est la fréquence d'occurrence. La **fréquence d'occurrence** est une mesure objective permettant d'estimer le nombre de fois qu'un lecteur a rencontré une chaîne de caractères représentant un mot donné. Depuis l'expérience princeps de Howes & Solomon (1951), des résultats concordants ont été publiés à partir de différentes tâches telles que la décision lexicale (Forster & Chambers, 1973; Rubenstein, Garfield, & Millikan, 1970; Scarborough, Gerard, & Cortese, 1979; Taft, 1979), la nomination (Forster & Chambers, 1973) et la comparaison de mots (Chambers & Forster, 1975). La fréquence d'occurrence d'un mot semble être un facteur important pour la rapidité et la justesse de l'identification d'un mot: plus la fréquence est élevée et plus la reconnaissance du mot est rapide et précise. L'influence persistante de cette variable sur les processus d'accès au lexique a fait que les principales recherches ont étudié celle-ci. Nous verrons dans la section suivante comment les théories psycholinguistiques ont réussi à implémenter cet effet dans les contraintes des modèles.

- **La longueur**

Un mot de haute fréquence tend à être court: la **longueur** d'un mot est, par conséquent, fortement corrélée avec sa fréquence d'occurrence. Gough (1972) argumente que les lettres sont traitées une à une pendant 15 millisecondes dans le *buffer* visuel (ce rythme est plus lent pour les moins bons lecteurs). C'est pourquoi il semble normal que les mots les plus longs soient plus durs à identifier que les mots courts. L'une des complications possible pour ce résultat est qu'il existe plusieurs manières de calculer la longueur d'un mot: le nombre de lettres, le nombre de syllabes, le nombre de phonèmes. En fonction de la modalité de présentation des stimuli, le contrôle de la longueur s'effectue différemment, et par conséquent peut influencer les résultats.

- **La familiarité**

La **familiarité** d'un mot (ou fréquence subjective) est aussi hautement corrélée avec sa fréquence d'occurrence (ou fréquence objective). La familiarité peut être perçue comme la fréquence personnelle. Il est possible qu'un mot présente une basse fréquence objective mais que le sujet soit habitué et familier avec ce mot. Par exemple depuis le début de cette thèse, le lecteur aura lu plusieurs fois le mot "psycholinguistique". Le lecteur aura donc une familiarité, ou encore fréquence subjective, pour cet item qui a considérablement augmenté depuis le début, alors que ce mot conserve une même et basse fréquence objective. Selon Gernsbacher

(1984), la familiarité correspond au nombre de fois qu'un mot est écrit, produit, lu ou entendu. Cet auteur indique que la familiarité d'un mot est fortement corrélée avec la fréquence objective pour les mots de haute fréquence. En revanche, pour les mots de basse fréquence, cette corrélation diminue considérablement.

- **Le voisinage orthographique**

Le terme de **voisinage orthographique** peut soit être défini par la taille, soit par la fréquence des mots. La taille du voisinage orthographique est définie comme le nombre de mots partageant toutes les lettres, sauf une, à la même position avec le mot cible (Coltheart, Davelaar, Jonasson, & Besner, 1977). La fréquence du voisinage orthographique est calculée par la présence ou non d'un voisin orthographique de fréquence plus élevée que le mot cible (Grainger, 1990). Alors que, Andrews (1989) trouve un effet facilitateur de la taille du voisinage orthographique, d'autres études n'ont pas répliqué ce résultat (Carreiras, Perea, & Grainger, 1997; Grainger, 1990). Grainger (1990) suggère que la fréquence des voisins orthographiques est un indice plus important que le nombre de ceux-ci: il observe que la performance des sujets diminue si le mot cible a un voisin orthographique avec une fréquence d'occurrence similaire, voire plus élevée que la sienne. Aucun effet de la taille du voisinage orthographique n'est enregistré dans les recherches de Grainger (1990). Depuis ces résultats contradictoires sur l'étude des effets de voisinage (fréquence et taille) de nombreuses recherches ont essayé de résoudre cette question. Les résultats montrent que l'effet de voisinage orthographique dépend: 1) de la tâche utilisée, 2) de l'indice étudié: la fréquence ou la taille du voisinage orthographique, 3) de la langue. Ferrand (2001) a réalisé une revue complète de la littérature sur ce sujet.

- **La répétition**

L'effet de **répétition** est considéré comme une facilitation dans le traitement d'un mot lorsque celui-ci est présenté une deuxième fois (Scarborough, Cortese, & Scarborough, 1977). Cet effet peut persister plus de 24 heures. Dans un nombre conséquent d'études, il a été reporté que la fréquence et l'effet de répétition pouvaient interagir, de telle sorte que les mots de basse fréquence bénéficient plus de la répétition que les mots de haute fréquence (Duchek & Neely, 1989; Forster & Davis, 1984; Norris, 1984). Une explication possible pour cette interaction est que les effets de fréquence et de répétition agissent au même niveau dans le traitement d'un mot (Monsell, 1985). Ainsi, exposer un sujet à une deuxième présentation d'un mot va temporairement augmenter sa fréquence effective par rapport à celle observée avant la répétition. De plus, selon Monsell (1985), l'effet de répétition affectera plus les mots

de basse fréquence. Jacoby & Dallas (1981) suggèrent une alternative à cette hypothèse lexicale des effets de répétition. Ces auteurs indiquent que l'effet de répétition peut-être de type épisodique: le bénéfice obtenu par la répétition d'un mot ne provient pas du résultat de la modification de la représentation lexicale en mémoire, mais plutôt de l'existence d'une mémoire qui est sensible au contexte avec lequel le mot est présenté. Ainsi, selon ces auteurs, l'interaction entre répétition et fréquence de mots s'explique par le fait que les mots de haute fréquence co-occurrent avec un nombre plus important de contextes linguistiques différents que les mots de basse fréquence, et, par conséquent, vont moins bénéficier d'une nouvelle présentation. Ces différentes hypothèses sont explicitées plus amplement dans une revue récente (Bowers, 2000).

- **Le contexte linguistique**

Le dernier facteur que nous voulons évoquer ici est l'influence du **contexte linguistique** sur la reconnaissance des mots. Meyer & Schvaneveldt (1971) démontrent que le contexte sémantique dans lequel un mot est présenté affecte les temps de réponse des sujets pour réaliser une décision lexicale. Par exemple, le mot "*butter*" entraîne une décision lexicale plus rapide lorsqu'il est précédé d'un mot relié sémantiquement tel que "*bread*" comparé à un mot non relié "*dog*". De plus, Becker (1979) présente l'évidence d'une interaction entre fréquence d'occurrence et l'effet du contexte (sémantique): les mots de basse fréquence profitent plus de la présentation d'un contexte comparé aux mots de haute fréquence. Goodman, McClelland, & Gibbs (1981) reportent l'influence du contexte syntaxique sur la reconnaissance des mots. En effet, il montre que les sujets sont plus rapides pour répondre au mot "*oven*" lorsqu'il est précédé d'une amorce syntaxique appropriée "*my*", comparé à une amorce qui est syntaxiquement inappropriée telle que "*he*". L'influence du contexte (sémantique et/ou syntaxique) n'est plus à remettre en question aujourd'hui; seul le problème du locus de cet effet est encore discuté.

Nous venons de voir que la fréquence d'occurrence semble être corrélée à beaucoup d'autres variables telles que la familiarité ou encore la longueur. De plus, elle interagit avec le voisinage orthographique, la répétition et le contexte linguistique. Malgré cette confusion possible entre les différentes variables pouvant expliquer l'effet de fréquence, les modèles psycholinguistiques ont mis en place des hypothèses précises afin d'expliquer cet effet.

## 2 Modèles d'accès au lexique

Les investigations expérimentales portant sur la reconnaissance des mots ont permis l'élaboration des modèles contemporains de l'accès au lexique. Deux types de questions ont principalement été analysés. La première est de comprendre l'organisation des informations lexicales qui sont inhérentes au lexique mental lui-même. Le deuxième type de question est de déterminer les mécanismes à partir desquels ces informations sont utilisées. Les modèles actuels de l'identification des mots diffèrent principalement sur les points suivants :

- **La nature des représentations lexicales:** certains modèles (Glanzer & Ehrenreich, 1979) postulent une séparation du lexique mental en fonction de la fréquence d'occurrence (propriété distributionnelle), alors que pour d'autres (Forster, 1976; Morton, 1969), cette propriété distributionnelle va permettre d'organiser les entrées dans un même lexique mental.
  
- **La nature des processeurs:**
  - Ces deux derniers modèles (Forster, 1976; Morton, 1969) se différencient sur la manière dont le code sensoriel est comparé aux représentations lexicales lors des procédures d'accès: la recherche dans la liste de candidats potentiels peut s'effectuer soit en série (accès séquentiel, chaque candidat est comparé l'un après l'autre avec l'entrée sensorielle; Forster, 1976) soit de manière simultanée (accès en parallèle, les candidats sont comparés en même temps à l'entrée sensorielle; Morton, 1969).
  - l'accès au lexique peut s'effectuer soit sous l'influence des informations sensorielles seules (modèle autonome ou modulaire), soit sous l'influence d'informations contextuelles multiples (modèle interactif ou non modulaire). Ainsi, les modèles autonomes postulent que l'accès au lexique ne bénéficie que de l'extraction des informations sensorielles du signal (traitement de bas en haut), alors que les modèles de type interactif privilégient d'une part le traitement de bas en haut, mais aussi de haut en bas (interaction des informations contextuelles avec celles issues du traitement sensoriel du signal).

D'après l'explication donnée sur le locus de l'effet de fréquence, deux principales classes de modèles émergent: les modèles mettant en place deux lexiques mentaux pour les mots de haute fréquence et de basse fréquence, et les modèles implémentant la fréquence d'occurrence

comme organisateur des entrées lexicales dans le lexique mental. Nous allons à présent expliquer les mécanismes de fonctionnement de ces différents types de modèles.

## 2.1 Modèle à deux lexiques de Glanzer

Glanzer & Ehrenreich (1979) montrent, pour une tâche de décision lexicale, que l'existence de l'effet classique de fréquence sur la reconnaissance de mots dépend des propriétés fréquentielles des items qui précèdent le mot cible. Ces auteurs trouvent un effet classique de fréquence lorsque les mots sont présentés dans des listes mixtes où des mots de haute, moyenne et basse fréquences sont mélangés. En revanche, lorsque les mots sont présentés dans des listes composées soit de mots de haute fréquence, soit uniquement de basse fréquence, alors le pattern de résultat est modifié selon les 4 formes suivantes: 1) les mots de haute fréquence présentent un temps de décision plus rapide dans des listes pures composées uniquement de mots de haute fréquence par rapport à des listes mixtes; 2) les temps de décision pour les mots de basse fréquence sont plus rapides dans des listes pures par rapport à des listes mixtes; 3) il n'existe pas de différence de temps de réaction entre les mots de moyenne et de basses fréquences dans les listes pures; 4) les participants répondent plus rapidement aux non-mots dans les listes pures composées de mots de haute fréquence par rapport à des listes mixtes ou encore à des listes pures de mots de moyennes fréquences. Glanzer & Ehrenreich (1979) interprètent leurs résultats en divisant la recherche lexicale en deux voies parallèles. Une première recherche tient place dans un lexique plus concis, où seuls les mots de haute fréquence sont répertoriés. La seconde voie d'accès prend place dans un deuxième lexique constitué de la plupart des autres mots. Aucune des deux voies lexicales ne présente d'ordre fréquentiel de recherche. Glanzer & Ehrenreich (1979) suggèrent que pour des listes mixtes, le lexique concis sera d'abord consulté, et si aucun candidat ne correspond, alors le deuxième lexique sera vérifié. Cette interprétation des résultats conduit à aboutir à un effet graduel de la fréquence lors d'une tâche de décision lexicale, puisque les mots de haute fréquence sont stockés dans le lexique concis (le premier) alors que les mots de basse fréquence ne sont jamais répertoriés dans celui-ci. Les mots de moyenne fréquence occupent une place intermédiaire: ces mots peuvent occasionnellement appartenir au premier lexique comme être listés dans les deux lexiques. Cependant, pour des effets de listes pures, la stratégie employée par les processeurs est totalement différente. Glanzer & Ehrenreich (1979) postulent que la probabilité de recherche dans le lexique concis est une fonction dépendante du nombre d'items de haute fréquence précédant le mot cible et n'est pas obligatoire: si tous les mots précédents sont de haute fréquence alors la recherche s'effectue dans le lexique



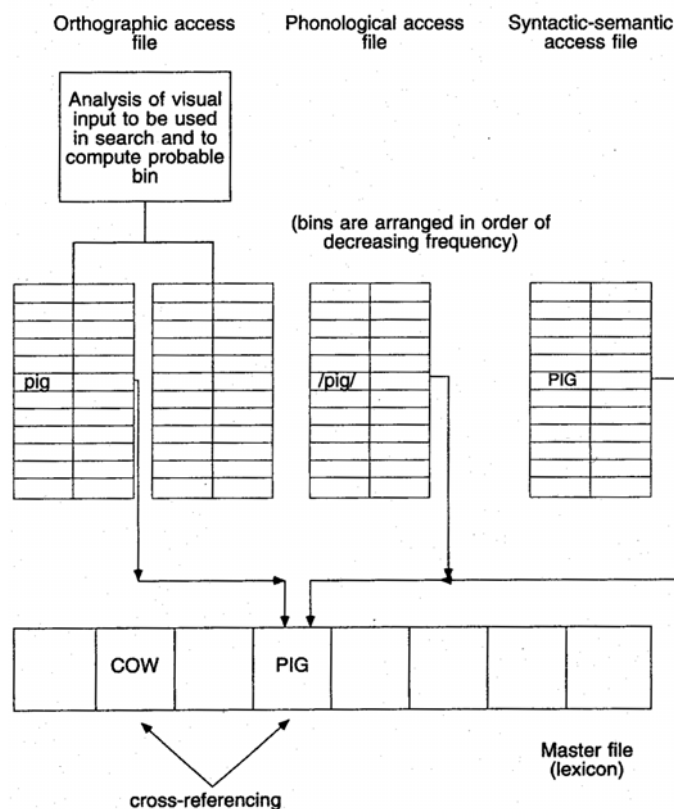
concis; si les mots précédant le mot cible sont tous de basse fréquence, alors cette recherche est annulée et celle-ci débute automatiquement par le deuxième lexique. Par conséquent, les mots de basse fréquence seront identifiés plus rapidement dans des listes pures de mots de basse fréquence que dans des listes mixtes puisque la recherche dans le premier lexique sera évitée. Le temps de décision pour les non-mots sera aussi plus rapide lorsqu'ils sont présentés dans des listes pures de mots de basse fréquence pour les mêmes raisons. Pour finir, Glanzer & Ehrenreich (1979) indiquent que la stratégie normale lorsque les listes sont mixtes est de tester en premier le lexique concis où sont stockés les mots de haute fréquence.

## 2.2 Modèles à un lexique mental

A l'opposé, certains auteurs (Forster, 1976; Morton, 1969) refusent l'idée que cette propriété distributionnelle (fréquence d'occurrence) des mots puisse être un critère de division dans des lexiques mentaux séparés. Puisque cette propriété modifie la rapidité de la reconnaissance des mots, la fréquence d'occurrence apparaît alors comme organisateur des entrées dans un même et unique lexique mental. C'est à partir de ce dernier point que les modèles de Morton (1969) et Forster (1976) vont différer.

### 2.2.1 Modèle sériel de Forster

Forster (Forster & Olbrei, 1973; Forster & Ryder, 1971) propose un modèle dans lequel les entrées sensorielles sont comparées de manière sérielle aux représentations dans le lexique mental. La comparaison de l'entrée sensorielle avec les représentations internes s'effectue de manière active. Il distingue les voies ou fichiers d'accès (*access file*) du lexique mental (*master file*). Le lexique central regroupe toutes les informations des mots telles que la phonologie, l'orthographe, la sémantique et la syntaxe. Ce lexique mental ne peut être interrogé directement, il est questionné par des pointeurs à partir des fichiers d'accès périphériques (*access file*). Pour cet auteur, le premier principe organisateur du lexique mental réside dans la constitution des fichiers d'accès. Ils sont soit organisés sur des critères orthographiques pour la modalité visuelle, phonologiques pour la modalité auditive, et syntaxico-sémantique pour la production de la parole (voir figure 1).



**Figure 1:** Modèle d'accès lexical par recherche sérielle. Les casiers (*bins*) sont ordonnés par ordre de fréquence et les différentes voies d'accès (*access files*) sont indépendantes, tandis que le lexique mental (*master file*) est amodal. D'après Forster (1976).

Ainsi, une suite de lettres sera traitée par le fichier d'accès orthographique. En revanche, le lexique mental (*master file*) ne conserve que des informations amodales, c'est à dire qui ne sont pas liées à la modalité de présentation. Les voies d'accès sont structurées en *bins* ou répertoires, de manière à ce que la recherche dans cette encyclopédie soit rapide et efficace. Les entrées ou fiches sont organisées sur des bases de similitudes dans les *bins* ou répertoires. Pour la voie d'accès orthographique, la similitude sera les premières lettres des entrées sensorielles. Les mots commençant par les mêmes lettres sont alors regroupés dans le même répertoire ou *bin*, et par conséquent la recherche lexicale se limitera à la consultation des entrées lexicales qui y sont stockées.

Le deuxième principe organisateur du lexique mental, est une structuration des *bins* ou répertoires dans les fichiers d'accès en fonction de la fréquence d'occurrence. Ce modèle est caractérisé par le fait que les *bins* présentent une hiérarchie fréquentielle: les items les plus fréquents seront consultés en premier. Après la sélection du *bin* où les candidats sont stockés, la recherche s'effectue par ordre décroissant de fréquence jusqu'à ce que le code sensoriel présente des similitudes avec la représentation. Les différences de traitement entre mots de

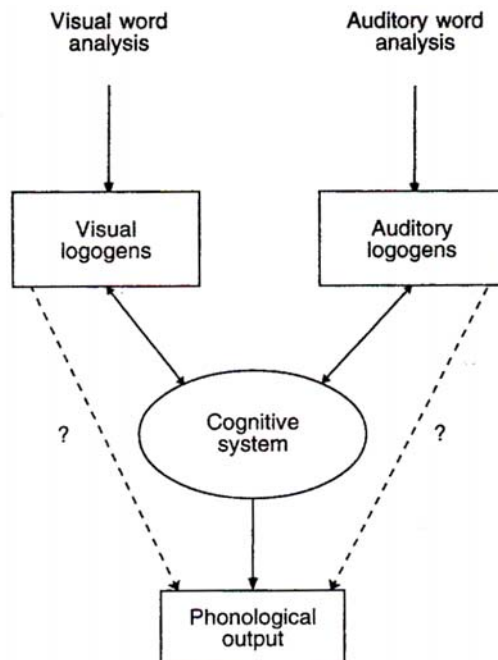
haute fréquence et mots basse fréquence sont expliquées par une recherche préférentielle sur les mots de haute fréquence; puis, si l'entrée sensorielle ne correspond pas, le processeur accède aux fiches des mots de basse fréquence. L'identification lexicale consiste en une recherche séquentielle à travers les fiches dans les répertoires. Lorsque le code sensoriel correspond à une fiche dans le répertoire du fichier d'accès, alors l'information est transmise au lexique mental (*master file*) et le mot peut alors être identifié: les informations qui lui sont associées sont alors activées. Cette dernière étape constituant l'accès au lexique.

Concrètement, l'effet classique de fréquence est obtenu lorsqu'un mot de haute fréquence produit un temps de décision plus court comparativement à des mots de fréquence intermédiaire. Ces derniers mots sont à leur tour identifiés plus rapidement que les mots de basse fréquence. Un effet de graduation est alors enregistré et les temps de décision les plus longs sont obtenus pour des non-mots orthographiquement légaux. Pour la plupart des modèles se basant sur un principe de recherche séquentielle, l'effet graduel de fréquence est le résultat de deux hypothèses: a) la recherche dans le lexique mental commence toujours par les mots de haute fréquence; b) la recherche peut s'arrêter d'elle-même si aucune correspondance n'est trouvée. Les temps les plus longs obtenus pour les non-mots résultent du fait qu'une recherche exhaustive dans la liste de mots est nécessaire pour établir qu'il n'existe pas d'entrée correspondant à l'item présenté. Puisque dans les deux cas la totalité du répertoire doit être scanné, le temps de décision lexicale pour les non-mots tend à approcher celui des mots de basses fréquences.

### 2.2.2 Modèle parallèle de Morton

Dès la fin des années 1960, Morton postule l'existence d'une activation en parallèle lors de l'accès au lexique (Morton, 1969; Morton, 1979). Depuis, d'autres modèles se sont inspirés de ce postulat (Gordon, 1983; McClelland & Rumelhart, 1982). Le modèle de Morton est principalement caractérisé par la description qu'il donne des représentations lexicales en mémoire. Chaque entrée lexicale est caractérisée par un "logogène" ou "détecteur de mots" qui est sensible à différents indices pertinents. Le logogène possède un niveau d'activation qui varie en fonction de l'accumulation de ces indices. Cette accumulation est dite passive et en parallèle, c'est à dire que le logogène accumule les indices pertinents jusqu'à ce que le niveau d'activation dépasse un seuil au-dessus duquel l'identification est possible. Les informations provenant de l'analyse sensorielle sont alors comparées en simultané à tous les logogènes stockés dans le lexique mental. Le niveau d'activation d'un logogène varie en fonction du

nombre d'indices coïncidant avec les informations sensorielles. De plus, le seuil du logogène doit être dépassé pour qu'il y ait identification du mot. Selon Morton, l'accès au lexique est donc un mécanisme qui compare en parallèle l'entrée sensorielle avec toutes les représentations stockées dans le lexique mental. Le modèle original de Morton (Morton, 1969) conceptualisait un seul logogène pour chaque mot quelque soit la modalité testée: le même logogène est utilisé lors de la reconnaissance auditive et visuelle. Cet arrangement est appelé amodal. Ce type d'arrangement prédit clairement que la modalité de présentation des stimuli n'est pas importante. Par conséquent, lors d'une tâche d'amorçage, si l'amorce est présentée dans la modalité auditive, alors que la cible l'est en modalité visuelle, la facilitation observée dans le traitement devrait être équivalente à celle obtenue lorsque l'amorce et la cible sont présentées en mode visuel. L'étude de Winnick & Daniel (1970) montre des résultats à l'encontre de cette prédiction: lorsque l'amorce et la cible ne sont pas présentées dans la même modalité, la facilitation est plus faible que lorsqu'une seule et même modalité est utilisée. Ainsi, ces résultats contredisent-ils la nature amodale des logogènes. Par conséquent, Morton (1979) a révisé son modèle en divisant les logogènes en deux groupes pour les différentes modalités: pour chaque mot désormais deux logogènes peuvent être activés en fonction de la modalité de présentation (voir figure 2).



**Figure 2:** Modèle d'accès au lexique par recherche parallèle. D'après (Morton, 1969), la version présentée dans cette figure est la version révisée de Morton (1979).

Morton suggère que le seuil d'activation des mots varie en fonction de leur fréquence d'occurrence. Ainsi, plus un mot est de haute fréquence et plus son seuil sera bas, en revanche, plus un mot est de basse fréquence et plus son seuil est élevé. Le modèle de Morton permet, par ce postulat, d'expliquer l'effet de fréquence obtenu sur les mots. Ainsi, un mot très fréquent dans la langue, va présenter un seuil d'activation assez bas (peu d'information seront nécessaires pour qu'il soit identifié), alors qu'un mot rare disposera d'un seuil plus élevé (nécessité d'une quantité d'information plus importante). Alors que Forster explique l'effet de fréquence par une propriété structurelle du lexique mental, Morton interprète l'effet de fréquence par les propriétés mêmes des logogènes.

### 2.2.3 Nature des processeurs

Les modèles de Morton (1969) et Forster (1976) s'opposent aussi sur la manière dont les informations d'un mot sont employées afin d'aboutir à une interprétation du mot, de la phrase ou du discours.

#### 2.2.3.1 Autonomie de Forster

D'après le modèle de Forster, l'accès lexical ne se réalise que sur la base des informations décodées à partir de l'entrée sensorielle même. Forster postule l'autonomie de fonctionnement des processus d'accès au lexique (Forster & Olbrei, 1973; Forster & Ryder, 1971). Ce principe fait partie d'une théorie plus générale développée par Fodor (1983) et selon laquelle l'activité mentale humaine est caractérisée en "modules". La notion d'autonomie de fonctionnement découle directement de l'idée que les sous-systèmes ou "modules" impliqués dans la compréhension du langage fonctionnent de manière encapsulée du point de vue informationnel, c'est à dire, les informations circulent de bas en haut et chaque niveau d'analyse n'utilise que les informations du niveau qui le précède directement. Selon Forster, les procédures d'accès au lexique sont déterminées exclusivement par les informations issues des niveaux infra-lexicaux de traitement, et par les principes d'organisation du lexique lui-même. Une conséquence de l'hypothèse de l'autonomie est qu'elle conduit à postuler des restrictions importantes sur les possibilités d'échange entre les différentes sources d'informations. Ainsi, ces échanges ne peuvent prendre place qu'à la sortie respective de chaque processeur. En d'autres termes, les informations provenant des niveaux supérieurs de traitement (circulation des informations de haut en bas) ne peuvent affecter les procédures de traitement des niveaux inférieurs.

Dès le début des années soixante-dix, Forster et collaborateurs présentent des données expérimentales en faveur de la théorie autonomiste d'un traitement syntaxique (Forster & Ryder, 1971) avec un paradigme de présentation visuelle séquentielle et rapide (*rapid sequential visual presentation* ou RSVP). Les mots d'une phrase sont présentés un à un et ce, avec un rythme très rapide de 16 mots par secondes. Le sujet doit alors restituer le plus grand nombre de mots possible. Forster & Ryder (1971) montrent alors que la complexité de la structure syntaxique de la phrase influence la performance de restitution des mots, alors que la congruence sémantique ne l'affecte pas. Ainsi, les mots sont mieux restitués lorsque la phrase présente une structure syntaxique acceptable. Ces auteurs postulent que l'ordre logique d'enchaînement des mots au sein d'une phrase ne peut être déterminé que par les propriétés syntaxiques des stimuli. Ce traitement initial de la catégorie syntaxique des mots est appelé "processeur syntaxique". Ce processeur syntaxique n'a pas pour rôle de déterminer si le verbe est utilisé correctement d'un point de vue sémantique. Il va spécifier si un verbe ou un nom doit être utilisé à cet endroit précis dans l'énoncé, indiquant ainsi l'ordre d'apparition des mots, et par conséquent la structure syntaxique de la phrase. De ce fait, ce processeur initial syntaxique traite les informations de structure, et non les informations sémantiques des items spécifiques utilisés dans la phrase. Le produit de ce traitement initial est ensuite proposé à un second processeur qui poursuit l'analyse de l'énoncé en combinant les informations d'ordre lexical, sémantique et pragmatique. Ainsi, le second processeur a pour tâche de sélectionner les arguments de la phrase sur les bases de la représentation syntaxique donnée par le premier processeur. Selon cette interprétation, la construction de la représentation syntaxique de la phrase s'effectue de manière autonome, indépendante de toute influence par les informations sémantiques.

L'évaluation empirique de ce modèle a conduit Forster et collaborateurs de reformuler plus précisément son fonctionnement. En effet, ce modèle postule que les informations contextuelles, quelle que soit leur nature, ne peuvent influencer la rapidité de l'accès au lexique. Forster explique l'effet facilitateur du contexte sur la reconnaissance d'un mot comme non pas une influence sur les processus d'accès au lexique, mais plutôt sur les processus d'intégration du mot avec le contexte précédent (Forster, 1981b). Depuis sa formulation, cette interprétation a reçu des validations empiriques, notamment, à partir d'études portant sur les ambiguïtés lexicales. Swinney (1982) montre qu'à partir de mots ambigus ou homographes (par exemple: jumelle, avocat), les différentes représentations sont activées, et ce n'est que plus tard que le contexte sélectionne la bonne interprétation du mot.

Forster interprète donc l'effet du contexte comme un effet post-lexical: après une première phase non sélective d'accès au lexique où toutes les interprétations sémantiques et syntaxiques du mot sont activées, un processus d'acceptation contextuelle prend place et va conduire à ne conserver que l'interprétation du mot ambigu qui est appropriée. Lorsque le contexte est réduit à un seul mot, Forster explique l'effet facilitateur (les participants sont plus rapides pour identifier les mots qui sont précédés par un mot qui leur sont reliés sémantiquement comparé à lorsque les deux mots n'ont aucun lien) par une activation intra-lexicale. La présentation du premier mot a activé non seulement sa propre représentation, mais encore celle des autres mots auxquels il est relié dans le lexique mental. Ainsi, c'est cette propagation aux mots reliés qui serait à l'origine de la facilitation observée. Ce type d'activation est compatible avec l'hypothèse de l'autonomie, car les effets de contexte sont alors considérés comme automatiques et prennent place à l'intérieur du lexique lui-même. Ainsi deux types d'activation peuvent être mis en place dans ce modèle. D'une part, une activation par la localisation de la fiche appropriée dans un répertoire (effet dû aux informations de l'entrée sensorielle même) et d'autre part, une activation directe via les relations sémantiques entre les mots (effet dû aux propriétés structurelles même du lexique mental).

### 2.2.3.2 Interactivité de Morton

Morton (Morton, 1969; Morton, 1979) postule que les processus de reconnaissance des mots sont "interactifs". Le seuil de chaque logogène peut être modulé à partir d'informations sensorielles, mais aussi à partir des informations provenant des niveaux d'analyses supérieurs comme les informations sémantiques et/ou syntaxiques provenant du contexte.

Afin de rendre compte des effets de contexte, Morton avance l'idée que la présentation d'un contexte tel que "le chat mange la .." va activer le logogène du mot "souris", avant même la présentation physique du stimulus en question. Ainsi l'accumulation des indices provenant de l'analyse perceptive, ainsi que des informations contextuelles vont permettre d'atteindre rapidement le seuil d'activation du logogène "souris" comparé par exemple, au logogène "table". Dans le cadre du modèle interactif de Morton, l'état d'activation du logogène peut par conséquent, être modifié par les informations provenant du contexte linguistique. Ainsi, les informations sémantiques et syntaxiques peuvent directement agir sur le niveau d'activation des logogènes. Cette interprétation de l'effet du contexte est en complète contradiction avec l'hypothèse de la modularité des processus d'accès développée par Forster et ses

collaborateurs. Selon Morton (1969; 1979), les logogènes se situent à l'interface entre le système sensoriel et le système cognitif (voir figure 2). C'est pourquoi leur seuil peut être affecté par l'influence des deux sources en même temps. Le traitement d'un mot est par conséquent, déterminé par les informations phonologiques, lexicales, sémantiques et syntaxiques dont dispose le locuteur à ce moment précis. En d'autres termes, les informations provenant des multiples niveaux d'analyses sont utilisables par le locuteur afin de réaliser la représentation de la phrase. Morton (1969; 1979) postule l'utilisation de plusieurs types d'informations lors de l'interprétation de la phrase. Selon lui, la compréhension d'une phrase ne passe plus par un niveau de représentation purement syntaxique. Les informations d'ordre sémantique et pragmatique influencent directement le traitement syntaxique. La construction de l'une des représentations peut-être influencée, plus ou moins rapidement, par la construction de l'autre représentation. Ce type de modèle implique que l'analyse initiale de la structure syntaxique soit modulée par des informations de provenances multiples (lexicale, thématique, etc.). Plus récemment, Tanenhaus, Garnsey, & Boland (1990) proposent un modèle dans lequel les connaissances sémantiques du lexique guideraient les processus syntaxiques. Ainsi, ces informations lexicales permettraient d'influencer les décisions syntaxiques ultérieures.

### 2.3 Conclusions

Nous venons de voir trois explications différentes du rôle de la fréquence sur la rapidité d'identification d'un mot. A l'heure actuelle, aucun consensus n'est obtenu, et ces trois modèles peuvent expliquer un aussi grand nombre de résultats. Un seul point commun entre eux est que la fréquence d'occurrence agit à un niveau pré-lexical, permettant ainsi d'accélérer l'identification d'un mot.

Dès 1978, Bradley apporte des données problématiques pour ces trois classes de modèles: l'effet de fréquence peut ne pas être uniforme pour tous les mots (Bradley, 1978). Pour certaines catégories grammaticales (articles, pronoms, prépositions, etc.), l'effet de fréquence est absent: le fait qu'un mot soit de fréquence élevée n'affecte pas sa vitesse de traitement. Ces différentes catégories grammaticales regroupent des mots qui sont en général très fréquents et très courts. A l'opposé, en regroupant les noms, verbes, adjectifs, Bradley montre que cette deuxième macro-classe présente un effet de fréquence lors d'une tâche de décision lexicale. Dans cette deuxième classe, les mots sont moins fréquents et plus longs que ceux appartenant à la première classe. La séparation entre ces deux classes de mots n'est ni



aléatoire, ni basée sur des critères de regroupement par fréquence d'occurrence ou encore de longueur. En réalité ces deux macro-classes disposent de propriétés linguistiques différentes. Nous verrons dans la section 3.2 la description précise de celles-ci. Face à ces résultats, Bradley (1978) propose donc un modèle en faveur d'une architecture séparée en deux lexiques mentaux pour ces deux classes de mots. Elle indique que ce n'est plus la fréquence d'occurrence d'un mot qui accélère son processus d'identification mais plutôt sa classe.

Puisque les facteurs fréquence et longueur semblent fortement liés à la classe de mots (tout au moins dans les stimuli utilisés par Bradley), les modèles actuels de reconnaissance visuelle de mots peuvent avoir omis d'implémenter une variable primordiale lors de l'accès au lexique. Dans la section suivante nous allons préciser la terminologie utilisée pour distinguer les deux classes de mots.

### 3 La classe de mot: une distinction catégorielle?

#### 3.1 Terminologie

Dans la tradition linguistique, il existe une tendance à diviser les catégories grammaticales en deux classes: les éléments grammaticaux (encore appelés mot de fonction) et ceux de contenu (Bolinger, 1975). La première classe est constituée d'unités composées de mots ou morphèmes (déterminants, morphèmes de temps ou d'accord) qui servent principalement les fonctions grammaticales, alors que la seconde classe fournit les informations sémantiques dans la phrase (noms, verbes, adjectifs). Depuis quelques années et grâce aux développements de la grammaire transformationnelle, cette distinction a refait surface sous les termes de classe fonctionnelle et classe lexicale (Chomsky, 1995) et s'applique autant à des morphèmes qu'à des mots entiers. L'importance de cette distinction a reçu de fortes contributions psycholinguistiques depuis quelques années. Nous verrons dans la section 4, que les arguments empiriques proviennent de l'étude de différentes populations telles que les adultes sains, les patients aphasiques, ou encore les enfants. Les différences obtenues dans les processus de traitement peuvent refléter des variations dans les propriétés syntaxiques des deux macro-classes, et alors fournir un fondement psychologique aux récents développements des théories linguistiques. Cependant, malgré l'importance des catégories fonctionnelles pour les théories linguistiques, et malgré la robustesse des arguments issus de la psycholinguistique, il reste encore d'importantes questions non résolues comme par exemple: à partir de quels critères sont constituées les catégories fonctionnelles?

Typiquement, la grammaire transformationnelle inclut dans les catégories fonctionnelles les morphèmes de temps et d'accord, qui n'appartiennent pas à la catégorie majeure constituée de noms, verbes, d'adjectifs, d'adverbes et des prépositions. Cependant, à travers la littérature psycholinguistique, les expressions telles que "ici", "là", etc., avec des critères linguistiques appartenant à la catégorie dite majeure ou encore des marqueurs de discours tels que "par conséquent" sont souvent amalgamés avec les éléments de la classe des mots de fonction. À l'inverse, certaines expressions habituellement considérées comme appartenant à la classe de mots de fonction (certains quantificateurs: "quelques", "plusieurs") sont catégorisées comme des mots de la classe de contenu. Les relations entre les arguments empiriques et les distinctions théoriques sont alors beaucoup plus problématiques qu'il n'y apparaît. Ceci est particulièrement vrai lorsque l'on essaye de déterminer si la classe des mots de fonction est catégorielle, comme l'ont suggéré des travaux portant sur l'acquisition du premier langage (Morgan, Shi, & Alopenna, 1996). Si tel est le cas, alors la nature des critères permettant cette division catégorielle et la manière dont les deux classes interagissent deviennent des questions importantes.

La distinction entre éléments fonctionnels et lexicaux est souvent confondue avec la distinction entre classe ouverte et fermée (Quirk, Greenbaum, Leech, & Svartvik, 1972). Par définition, les mots de classe fermée (articles, verbes auxiliaires, pronoms, prépositions, conjonctions, déterminants, quantificateurs, etc.) sont constitués d'un nombre limité et fixe d'items (environ 150-250 mots en anglais; Petocz & Oliphant, 1988) et n'acceptent que très rarement de nouveaux candidats. Leur nombre total dans une langue est inférieur au nombre d'items appartenant à la classe ouverte. À l'opposé, les mots de classe ouverte, ou lexicaux, contiennent un nombre plus important de candidats (noms, verbes, adjectifs, et la plupart des adverbes) et est en constante augmentation par l'addition de nouveaux membres soit par simple création, soit par dérivation d'items déjà existants. Par exemple, suite à l'extraordinaire expansion de l'utilisation des outils informatiques, les mots tels que micro-ordinateur, courrier électronique, tableur, ont été créés, mais aussi des mots comme souris ont pris des sens différents.

Comme les mots de classe fermée sont en petit nombre dans le lexique mental, ils sont forcément répétés plus souvent comparés aux mots de classe ouverte. Ainsi, le petit nombre des mots de classe fermée est contrebalancé par leur fréquence d'occurrence: les mots de classe fermée sont généralement plus fréquents que les mots de classe ouverte. En plus des différences de caractéristiques numéraires et de fréquence d'occurrence, les mots de classe

fermée présentent d'autres propriétés linguistiques qui peuvent les différencier des mots de classe ouverte. Nous avons décidé de ne pas employer la distinction mots de fonction/contenu ou encore mots grammaticaux/lexicaux, mais plutôt classe fermée/ouverte, car cette dernière est la plus utilisée dans le domaine de la psycholinguistique et des neurosciences cognitives. Ainsi pour la suite de notre exposé, nous n'utiliserons plus que les termes mots de classe ouverte et mots de classe fermée qui respectivement feront référence aux unités lexicales (ou encore de contenues) et fonctionnelles (ou encore grammaticales).

### 3.2 Des propriétés linguistiques distinctives

Ces deux classes de mots (ouverte et fermée) présentent d'autres distinctions plus fondamentales que leur nombre effectif.

- **Phonologie et morphologie**

Pollack & Pickett (1964) présentent l'évidence que les mots de classe fermée sont moins "saillants" que les mots de classe ouverte: il semble plus difficile de les reconnaître sans l'apport du contexte. En Anglais, Cutler & Norris (1988) indiquent que les mots de classe fermée ne suivent pas les règles générales d'accentuation métrique: leurs voyelles tendent à être réduites. Cette différence phonologique peut aussi se retrouver dans l'absence d'une syllabe initiale accentuée pour les mots de classe fermée (9,5 %) comparés aux mots de classe ouverte (90 %; Cutler & Carter, 1987). La pauvreté du statut phonologique des mots de classe fermée se reflète aussi dans leur structure morphologique. En Anglais, les mots de classe fermée peuvent être encodés comme des morphèmes ou encore comme des clitiques: par exemple, *I'll work it out* (pour *will*); *we've arrived* (pour *have*). Cependant les réductions phonologiques peuvent faire intervenir les mots de classe ouverte dans certaines conditions contextuelles: par exemple *wanna* pour *want to*, ou bien encore *gonna* pour *going to*. De plus, dans certaines conditions, les mots de classe fermée peuvent être accentués afin d'augmenter le niveau d'attention. Alors que les réductions phonologiques et morphologiques sont indicatives du statut fonctionnel d'un mot, elles n'en ressortent pas comme étant un critère spécifique.

Pour finir, les mots de classe fermée ne prennent pas (ou que très rarement) de marque de dérivation ou encore de flexion. Par conséquent, cette classe présente moins de variance que celle des mots de classe ouverte. Il est certain que les livres portant sur la morphologie discutent principalement de cet effet sur les mots de classe ouverte, mais il existe des

exceptions où la morphologie dérivationnelle peut s'appliquer aux mots de classe fermée. Notons que certains mots de classe ouverte ne présentent pas non plus les propriétés dérivationnelles classiques. Ainsi, l'absence de morphologie dérivationnelle sur les mots de classe fermée ne semble pas être un critère suffisant pour classer un item dans la classe fermée.

- **Syntaxe**

Les mots de classe fermée apparaissent dans des contextes syntaxiques restreints comparés aux mots de classe ouverte. Par exemple, les articles font intervenir des constructions où un nom suit (le chat, le chien) ou encore un adjectif (le gros chat, le gros chien). De même, les quantificateurs (plusieurs, certains) peuvent apparaître indépendants ou bien dans des constructions faisant intervenir un nom (plusieurs chats, certains chiens) ou encore avec l'article "de" (plusieurs de ces chats, certains de ces chiens), etc. En revanche, pour les mots de classe ouverte, le contexte syntaxique varie plus facilement, et n'est pas défini par des sous-classes de mots. Par exemple, les verbes peuvent être utilisés avec ou sans complément d'objet direct (il mange, il mange du pain), les noms peuvent présenter un syntagme prépositionnel qualificatif ou pas (le cartable, le cartable de Pierre), etc. Ainsi, définir qu'un mot est un verbe ou encore un nom, ne donne que très peu d'informations quant aux mots qui vont suivre. A l'opposé, déterminer qu'un mot appartient à la catégorie grammaticale des quantificateurs augmente considérablement les chances de prédiction pour les mots qui suivent (Cann, 2000).

Un autre aspect s'appliquant aux mots de classe fermée est qu'il n'y a pas de processus qui altère leurs propriétés de sélection comme pour les mots de classe ouverte. Par exemple, un verbe transitif pourra être modifié par l'ajout d'un complément d'objet (il travaille la terre). Ainsi, il n'existe pas de processus s'appliquant aux mots de classe fermée permettant de modifier les arguments sémantiques ou encore leur "valence" (il travaille correctement *vs.* Il travaille mal) .

- **Sémantique**

La différence la plus citée dans la littérature est certainement que les mots de classe fermée ont une interprétation grammaticale, alors que celle des mots de classe ouverte est indicative de sens. En d'autres termes, les mots de classe ouverte sont les catégories grammaticales assignant le sens ou encore la référence extralinguistique, alors que les mots de classe fermée sont responsables des informations structurales dans la phrase et indiquent les relations

grammaticales (Levelt, 1989; Schmauder, 1996). Ainsi, les noms sont des mots qui typiquement réfèrent à des personnes, des places ou des choses, les verbes sont des mots qui précisent des actions, des processus, des états, etc.. Les mots de classe fermée typiquement contraignent les relations entre les mots de classe ouverte, et indiquent les positions où les informations sémantiques, associées à un mot de classe ouverte, sont présentes.

Suite à cette distinction basique, on trouve aussi des différences dans le comportement de ces deux classes de mots. Les mots de classe ouverte sont susceptibles de subir des changements sémantiques variés (comme par exemple des réductions, dégradations ou encore des extensions) qui peuvent par conséquent affecter leur sens (McMahon, 1994); leur interprétation est aussi favorable à des phénomènes socioculturels. Au contraire, les mots de classe fermée sont plus résistants face à ces altérations et semblent moins affectés par les pressions culturelles ou encore sociales.

Nous venons de voir que les deux classes de mots (ouverte et fermée) semblent présenter de nombreuses propriétés distinctes et ce, dans de nombreux aspects linguistiques tels que la phonologie, la morphologie, la syntaxe ou encore la sémantique. Les mots de classe fermée tendent à être phonologiquement et morphologiquement réduits, sont susceptibles d'apparaître dans des contextes syntaxiques restreints, et leur sens peut être restreint dans certains contextes linguistiques. A l'opposé, les mots de classe ouverte ne présentent pas les mêmes propriétés, semblent être plus libres face à la morphologie et apparaissent dans des contextes plus larges. Malheureusement aucune de ces propriétés linguistiques n'est individuellement suffisante ou nécessaire pour différencier les deux classes de mots. Prenons l'exemple des prépositions. En fonction des critères distinctifs utilisés, ce type de mots peut changer d'une classe à l'autre. Sur des bases phonologiques, Kean (1979) catégorise les prépositions en fonction du nombre de leurs syllabes: les prépositions à une syllabe sont considérées comme des éléments appartenant à la classe fermée alors que les prépositions à deux syllabes comme des items de la classe ouverte. En utilisant une description plus syntaxique, la distinction s'effectue sur la base qu'un élément puisse être utilisé en tête de phrase/syntagme ou pas. Les prépositions sont alors considérées comme des éléments appartenant à la classe ouverte (Neeleman & Weerman, 1999). Un dernier critère pouvant être utilisé s'appuie sur des différences sémantiques (Klosek, 1979). Ainsi, pour cette classification, les prépositions perçues comme porteuses de sens sont problématiques car catégorisées comme des éléments de classe ouverte, alors que certains noms totalement dépourvus de sens (plutôt abstraits) sont déterminés comme des éléments de la classe fermée. Ces différences de catégorisation des

prépositions mais plus largement des deux classes de mots rendent les résultats empiriques peu clairs et parfois, ne permettent pas aisément des interprétations.

#### 4 Un lexique unique ou deux lexiques distincts?

Puisque les mots de classes ouverte et fermée semblent présenter autant de propriétés linguistiques distinctes, ils peuvent être de bons candidats pour mettre à l'épreuve l'élaboration des modèles théoriques sur la reconnaissance des mots. Dans cette perspective, de nombreuses recherches en psycholinguistique ont essayé de transposer les diverses propriétés linguistiques de ces deux classes de mots à différentes représentations psychologiques ou/et à des processus d'accès lexicaux distincts. Rechercher un mot particulier dans notre lexique mental est une procédure complexe. Une possibilité de réduire cette complexité est de posséder une organisation efficace des mots en différents sous-lexiques mentaux. Les candidats plausibles pour cette organisation peuvent être les mots de classe ouverte et fermée.

Afin de définir la distinction dans le traitement et l'identification des mots de classe fermée et ouverte, deux caractéristiques sont suggérées:

- La première propriété porte sur la **nature des représentations**: une première hypothèse repose sur un accès à des lexiques séparés pour les deux classes de mots. Les opposants à cette hypothèse admettent qu'un même et unique lexique mental soit activé. Pour ces derniers, les différentes propriétés intrinsèques à chaque classe de mots peuvent être utiles afin d'organiser les mots dans un même lexique mental.
- La deuxième propriété caractérise la manière dont les **processeurs** fonctionnent: les partisans de l'hypothèse autonomiste suggèrent que le contexte ne peut avoir aucune influence sur l'identification des mots de classe fermée et/ou ouverte (Forster & Chambers, 1973; Forster & Ryder, 1971). A l'opposé, l'hypothèse d'une interaction postule que les informations de type haut en bas (*top-down*) peuvent affecter la rapidité ainsi que la qualité de reconnaissance des mots lors de l'accès au lexique (Morton, 1969; Morton, 1979).

## 4.1 Arguments en faveur d'une séparation en deux lexiques

L'un des débats les plus enflammés dans ce domaine d'étude est focalisé sur la manière dont les deux classes de mots sont stockés/représentés dans le lexique mental. Ainsi, deux hypothèses émergent quant à l'étude de la nature des représentations dans le lexique mental. Une première admet qu'un seul lexique mental est utilisé pour les deux classes de mots, alors que la seconde postule deux lexiques mentaux séparés pour chaque classe de mots. Bien qu'un stockage séparé de ces deux classes de mots soit une idée séduisante, les partisans de cette hypothèse assumant ce type de distinction doivent répondre à certaines questions. Par exemple, quels sont les indices dans le signal auditif ou visuel permettant de sélectionner le lexique voulu?

Si deux lexiques séparés existent pour les deux classes de mots, alors il semble nécessaire qu'un ou plusieurs indices indiquent aux processeurs quel lexique mental il lui faut activer pour un mot donné. Les indices pertinents permettant cette distinction peuvent provenir de deux types de flux d'informations. De bas en haut, seules les différences physiques dans la réalisation des mots de classe fermée et ouverte seront prises en compte et le modèle retenu appartient alors à la classe des modèles autonomes (Forster, 1979; Forster & Ryder, 1971). De haut en bas, les informations de type contextuel sont aussi utilisées et le modèle est alors de type interactif (Morton, 1969; Morton, 1979). La mise en place de mécanismes séparés pour l'accès au lexique des deux classes de mots ne présente aucun problème architectural pour un modèle autonome (où seules les informations perceptives du mot sont utilisées) si les caractéristiques sub-lexicales des mots de classe ouverte et fermée sont clairement distincts. En effet, des indices pertinents permettant de distinguer les deux classes de mots sont nécessaires pour que les processeurs activent le lexique voulu. Selon cette idée, Sorensen, Cooper, & Paccia (1978) démontrent que les membres de la classe fermée, lorsqu'ils sont prononcés dans des phrases, ont des durées plus courtes que les mots de classe ouverte. De plus, ces deux classes de mots semblent présenter des caractéristiques phonologiques différentes, tout au moins dans la langue anglaise (Cutler, 1993; Cutler & Norris, 1988). Les mots de classe ouverte commencent, ou sont constitués, par une syllabe métrique forte dite accentuée (c'est à dire une syllabe incluant une voyelle pleine). Les mots de classe fermée semblent, eux, être composés d'une simple syllabe faible ou non accentuée (Cutler, 1993). Ces auteurs argumentent que la distinction entre syllabes accentuées et non accentuées est cruciale et permet de catégoriser les mots dans l'un ou l'autre des deux classes. Bien qu'il existe une tendance générale à pouvoir distinguer les deux classes de mots sur la base

d'indices métriques, il semble que cette hypothèse ne soit pas totalement efficace. En effet, selon Shillcock, Bard, & Spensley (1988), seuls 66 % des mots de classe fermée, en anglais, présentent une syllabe non accentuée contre 87 % de syllabes accentuées pour les mots de classe ouverte. De plus, en hollandais, la correspondance entre classe de mots et accentuation n'est pas si flagrante (Haveman, 1997). Un mot sur quatre en Anglais, et une proportion plus importante en hollandais, devrait être mal classés par cette technique d'analyse, rendant cette hypothèse métrique difficilement responsable de la distinction entre classe de mots.

Bien que d'un point de vue théorique, de vastes différences existent entre hypothèses à un lexique et deux lexiques, il est en pratique difficile de les dissocier empiriquement. Néanmoins, afin de choisir entre ces deux hypothèses, il est utile d'étudier les différents résultats et d'examiner comment chacun des modèles permet d'expliquer ces données. Ainsi, dans le pire des cas ou aucun des modèles ne peut-être rejeté, il semblera tout au moins utile de les affiner.

#### 4.1.1 Travaux précurseurs sur la production

Les premières recherches en aphasiologie ont typiquement centré leurs études sur les déficits en production des patients. La comparaison des résultats obtenus à partir de populations adultes saines et de patients aphasiques semble des plus pertinentes pour la confrontation des contraintes que présentent les modèles portant sur la compréhension de mots. En effet, les études portant sur la structure du lexique mental chez les patients aphasiques permettent de développer les théories du langage où différentes composantes du traitement peuvent être sélectivement atteintes et défectueuses. Une lésion focale de l'hémisphère dominant (l'hémisphère gauche chez une majorité de droitiers) n'entraîne pas simplement une réduction générale des performances langagières. Au contraire, il apparaît que différentes régions du cerveau semblent concourir à la réalisation de différentes fonctions linguistiques (Caramazza & Berndt, 1978; Geschwind, 1972). Pour les psychologues, l'intérêt majeur des recherches sur les patients aphasiques est d'identifier différents types d'aphasies qui permettront de prédire l'ensemble des symptômes du déficit. L'occurrence de tels troubles sous-jacents semblerait refléter les divisions naturelles de l'organisation du langage chez les sujets sains (Caramazza, 1996; Caramazza & Shelton, 1998). Il paraît primordial à la vue de cette analyse que les modèles de compréhension du langage ne doivent pas être en désaccord avec les données des patients aphasiques. De plus, comme il existe des arguments en faveur de connexions multiples entre les processus psychologiques et le cerveau chez ces patients, il est



possible de trouver une altération très fine rendant compte de processus fonctionnant indépendamment chez les adultes sains. Par conséquent, l'étude des patients aphasiques va permettre d'affiner les modèles élaborés sur les adultes sains en implémentant les contraintes de ce type de données. C'est pourquoi nous analyserons la littérature des sujets sains et des patients aphasiques en parallèle afin de mieux cerner les déficits de ces derniers.

#### 4.1.1.1 Comparaison adultes sains et patients aphasiques

Lors de la production spontanée chez des adultes sains, Garrett (1975; 1982) indique des patterns d'erreurs différents pour les deux classes de mots. Les transpositions, ou échanges de mots, tels que *the sunshine flourish in marigold* au lieu de *the marigold flourish in sunshine*, semblent impliquer les mots de classe ouverte alors que les mots de classe fermée participent à des erreurs où un seul item change de position. Dans cette même idée, Pechmann & Zerbst (2002) utilisent un paradigme d'interférence de mots dans la dénomination d'images. Ils présentent à des adultes sains une phrase, puis immédiatement une image qui peut être soit précédée, soit suivie (variation du *Stimulus Onset Asynchrony*, SOA: intervalle de temps compris entre le début de deux stimuli successifs) par un distracteur. Le mot distracteur peut être de trois types de relation avec l'image présentée: identique, de la même classe (ouverte, nom) ou d'une autre classe (fermée). Ces auteurs ont alors montré que l'effet d'interférence est sensible à la classe de mot: les distracteurs de même classe suscitent des temps de dénomination plus élevés par rapport à ceux de la classe fermée. Alors que ces 3 études (Garrett, 1975; Garrett, 1982; Pechmann & Zerbst, 2002) semblent indiquer des processus séparés pour les deux classes de mots en production, d'autres études ont remis en question ce résultat: à partir d'erreurs phonologiques de production (Dell, 1990), ou encore à partir d'erreurs de type structurale (Stemberger, 1984).

Les travaux en aphasiologie montrent clairement que différentes aires corticales sont nécessaires pour la production des mots de classe ouverte et fermée. Par exemple, le discours de patients aphasiques de Broca (lésion focale en région antérieure gauche) est relativement riche en mots de classe ouverte avec des mots de classe fermée manquants, voir mal utilisés (Grodzinski, 1984). Gardner & Zurif (1975) comparent les performances des patients aphasiques de Broca lors de la lecture à voie haute d'une part, pour des particules grammaticales (*or, at*) et d'autre part, pour des noms courts imaginables (*bee, web*). Alors que les particules grammaticales présentent 27 % d'erreurs de production, les noms imaginables ne sont mal prononcés qu'à 15 %. En français, Andreewsky & Seron (1975) sélectionnent des

mots qui sont délibérément bivalents (homographe), tels que les mots "or", "mais", "est", etc.. Ces mots peuvent être utilisés soit comme des mots de classe ouverte ou fermée en fonction du contexte linguistique dans lequel ils apparaissent. Même dans un contexte désambiguïsant l'utilisation de ces mots comme étant des mots de classe fermée, les patients agrammatiques semblent lire ces mots comme appartenant à la classe ouverte. Par exemple, le mot "mais" est lu avec deux syllabes (maï̃s, graine comestible) plutôt qu'avec une (mais, conjonction); le mot "est" est prononcé avec 3 phonèmes (point cardinal) plutôt qu'avec un seul (auxiliaire). Friederici & Schoenle (1980) répliquent les résultats de Andreewsky & Seron (1975) en anglais: le patient JL lit correctement 71 % des mots de classe ouverte contre seulement 18 % des mots de classe fermée. De plus, ce patient est incapable de lire les homographes grammaticalement bivalents. Ces résultats suggèrent que le patient de Andreewsky & Seron (1975) assimile les mots homographes aux mots de classe ouverte alors que le contexte sélectionne l'autre classe. Au contraire, le patient de Friederici & Schoenle (1980) les identifie comme des mots de classe fermée. Dans les deux cas, les résultats suggèrent une distinction computationnelle entre les deux classes de mots et un déficit de la voie spécifique des mots de classe fermée pour ces deux patients.

Les patients agrammatiques de Nespoulous, *et al.* (1988) omettent ou substituent plus facilement les mots de classe fermée par rapport à ceux de la classe ouverte dans un contexte de phrase. En revanche, la production de mots isolés de classe fermée semble être fonctionnelle. Biassou, Obler, Nespoulous, Dordain, & Harris (1997) indiquent des résultats quelque peu différents de ceux de Nespoulous, *et al.* (1988). Les patients agrammatiques de Biassou, *et al.* (1997) présentent un nombre plus important d'erreurs phonologiques (addition, substitution, omission de phonèmes) et de distorsions phonétiques (durée des segments produits) pour les mots de classe fermée. Contrairement à Nespoulous, *et al.* (1988) ces résultats sont stables à la fois, lorsque les mots sont produits en isolation et au sein de phrases. Pour Biassou, *et al.* (1997) les informations de catégorie grammaticale sont encore disponibles chez les patients, ce qui leur permet de distinguer et d'identifier différemment les deux classes de mots en isolation et dans le contexte de phrase. Au contraire, les patients de Nespoulous, *et al.* (1988) ne semblent plus disposer de cette information. Les résultats portant sur les patients semblent donc montrer des variations dans les déficits observés. Alors que Biassou, *et al.* (1997) indiquent la disponibilité d'informations de la catégorie grammaticale, Nespoulous, *et al.* (1988) suggèrent le contraire.

#### 4.1.1.2 Hypothèse de Garrett (1975)

Garrett (1975) argumente que la distinction entre classes de mots est d'une importance fondamentale lors des processus de production et de compréhension. Cet auteur suggère des représentations séparées pour les deux classes de mots ainsi que des traitements à différents niveaux d'analyse. Dans leur modèle de production, deux niveaux de traitement sont activés: le niveau fonctionnel et le niveau positionnel. Le résultat de l'activation du niveau fonctionnel consiste en une suite non ordonnée de mots de classe ouverte avec leurs relations entre eux non encore définis. Au niveau positionnel, la structure syntaxique est ensuite définie: les mots de classe fermée servent de support à la structure syntaxique afin que les mots de classe ouverte soient insérés dans un ordre correct. En accord avec ce modèle, les deux classes de mots sont identifiées à partir de deux niveaux différents dans les processus de production. Ce type de distinction n'est pas seulement fondamental pour les processus de production, mais il s'avère qu'en compréhension, différents mécanismes ont aussi pu être mis en évidence. Ainsi, cette hypothèse computationnelle a été par la suite élargie au domaine de la compréhension du langage. Pour la perception de phrases, un modèle à deux niveaux peut aussi être élaboré: cette fois-ci, les mots de classe fermée sont analysés au niveau positionnel en premier avant les mots de classe ouverte qui eux sont traités au niveau fonctionnel (Garrett, 1975). Garrett (1975) suggère que les mots de classe fermée tiennent un rôle prépondérant dans la construction de la structure syntaxique de la phrase. En particulier, ils fournissent une base à partir de laquelle les mots de classe ouverte sont répartis dans la phrase. Par conséquent, pour la perception de phrase, Garrett (1975) propose que les mots de classe fermée doivent recevoir un traitement différent et qui précède celui des mots de classe ouverte. Pour finir, l'architecture retenue par Garrett est de type autonome (Forster, 1976). Chaque composante du processus de traitement d'un mot est sensible à des informations spécifiques provenant des niveaux d'analyses inférieurs (bas en haut) et par conséquent est aveugle face aux informations provenant des niveaux supérieurs (haut en bas). Pour cet auteur, l'identification des mots de classe fermée et ouverte est strictement insensible aux informations contextuelles, seules les informations provenant des niveaux de traitement perceptif de la cible peuvent être prises en compte.

### 4.1.2 Travaux en perception

Dès 1977, Rayner présente des données chez des adultes sains indiquant que le temps de fixation, lors de la lecture silencieuse, est plus court pour les mots de classe fermée que pour ceux de classe ouverte. Depuis, ces résultats ont été répliqués par Carpenter & Just (1983) suggérant que ces deux classes de mots impliquent des processus de traitement différents. Nous avons vu qu'en production spontanée ou provoquée les patients agrammatiques présentent des déficits spécifiques sur les mots de classe fermée. Les études portant sur la perception du langage montrent que ces patients ont aussi des déficits spécifiques pour comprendre les mots de classe fermée. Friederici & Schoenle (1980) montrent une dissociation dans les performances de lecture de patients aphasiques pour des lésions focales dans les aires de Broca (lésion antérieure gauche) et de Wernicke (lésion postérieure gauche). Les patients aphasiques de Wernicke font plus d'erreurs sur les mots de classe ouverte, alors que pour les patients aphasiques de Broca, le pattern inverse est observé (augmentation d'erreurs sur les mots de classe fermée). De plus, les deux types de patients utilisent les homographes (*I can/the can, I will/the will*) comme appartenant à la classe de mot fermée même si le contexte biaise l'utilisation de la classe de mot ouverte. S'il est correct que les homographes sont représentés par plus d'une représentation mentale, alors l'une de celle-ci peut être spécifiquement détériorée en fonction de la localisation de la lésion. Friederici & Schoenle (1980) interprètent ces résultats comme une évidence pour une architecture séparée pour les deux classes de mots.

#### 4.1.2.1 Détection de lettres

A l'aide d'un paradigme de détection de lettres dans des passages de texte écrit, Drewnowski & Healy (1977) observent un nombre plus élevé d'erreurs de détection de lettres (t) lorsque les mots cibles sont des mots de classe fermée (*the*) par rapport à des mots de classe ouverte (*weather*). La proportion d'erreurs sur les mots *the* et *and* est réduite quand ceux-ci sont intégrés dans des passages dans lesquels la structure syntaxique des phrases est perturbée. En particulier, lorsque la structure locale syntaxique est conservée (par exemple lorsque le mot *the* est suivi d'un nom), la détection de lettre reste difficile dans ces mots à haute fréquence. En revanche, si le contexte syntaxique local est perturbé (par exemple lorsque le mot *the* est suivi d'un verbe) alors le nombre d'erreurs de détection de lettres est réduit. Ces effets n'ont

pas été mis en évidence lorsque la cible était "ant" plutôt que "and". Par conséquent, Drewnowski & Healy (1977) montrent une différence de qualité dans le traitement de ces deux classes de mots: les mots *and* et *the*, appartenant à la classe fermée, ne sont pas identifiés de manière isolée, mais sont plutôt lus dans des unités plus larges tels que de courtes phrases syntaxiques. Depuis cette expérience princeps, ces résultats ont été répliqués à partir de différentes expériences en anglais (Healy & Drewnowski, 1983; Rosenberg, Zurif, Brownell, Garrett, & Bradley, 1985), en français (Saint-Aubin & Poirier, 1997) et en hébreu (Koriat, Greenberg, & Goldshmid, 1991). Afin d'expliquer ces résultats, (Drewnowski & Healy, 1977; Healy & Drewnowski, 1983) émettent l'hypothèse suivante: les sujets engagent en parallèle différents traitements incluant celui des traits, des lettres, des syllabes, des mots et des niveaux supra-mots. Dès qu'une unité est reconnue à un certain niveau d'analyse, alors les traitements aux niveaux inférieurs sont interrompus, même s'ils ne sont pas terminés. Ainsi selon l'hypothèse de Healy & Drewnowski (1983), l'omission de lettres proviendrait d'une reconnaissance des unités de mots alors que le traitement des lettres n'est pas encore achevé.

Des conclusions similaires proviennent des études de Rosenberg, *et al.* (1985). Les patients aphasiques de Broca ne présentent pas le pattern normal (que les patients aphasiques de Wernicke présentent) lors de la détection de lettres dans des mots. Les patients aphasiques de Broca identifient les deux classes de mots de manière similaire, alors que les sujets normaux présentent des erreurs plus nombreuses lorsque le mot cible est un mot de classe fermée (Drewnowski & Healy, 1977; Healy & Drewnowski, 1983; Koriat, *et al.*, 1991; Saint-Aubin & Poirier, 1997).

#### 4.1.2.2 Recherche de mots

Bradley, Garrett, & Zurif (1980) et Friederici (1985) utilisent une tâche de recherche de mots, afin de tester la différence entre les deux classes de mots. Dans une telle tâche, les participants écoutent une phrase et doivent chercher une cible qui est spécifiée avant la présentation de la phrase. Les sujets doivent alors donner leur réponse en pressant aussi rapidement qu'ils le peuvent une clef de réponse. Les temps de réaction ainsi que les taux d'erreurs sont enregistrés. Bradley, *et al.* (1980) et Friederici (1985) observent que des patients agrammatiques reconnaissent plus rapidement les mots de classe ouverte que ceux de classe fermée. Friederici (1986) montre que les patients avec une aphasie de Wernicke (lésion en région postérieure gauche) ont un pattern similaire (temps de réponses plus courts pour les mots de classe fermée que pour ceux de classe ouverte) mais sont plus lents que les sujets

normaux. En revanche, les patients aphasiques de Broca (lésion en région antérieure gauche) présentent des temps de reconnaissance pour les mots de classe ouverte plus rapides que pour ceux de classe fermée (pattern inverse de celui des adultes sains et des patients aphasiques contrôles) alors que leur taux d'erreur ne varie pas significativement. Il existe d'autres considérations dans le domaine de la neuropsychologie montrant que ces patients sont également déficients dans la récupération et l'utilisation des informations structurales apportées par un mot de classe fermée. Par exemple, Zurif & Caramazza (1976) fournissent indirectement une autre preuve de différence de traitement entre les deux classes de mots. Lorsqu'ils demandent à des patients aphasiques de Broca de construire la structure hiérarchique de la phrase (ils demandent aux patients quels mots vont ensemble), ceux-ci regroupent les articles et autres mots de classe fermée ensemble, contrairement aux sujets normaux qui les regroupent plutôt avec un mot de classe ouverte. Les patients agrammatiques de Broca semblent donc présenter des difficultés pour reconnaître les mots de classe fermée, alors que, généralement, ils montrent peu de difficultés pour identifier les mots de classe ouverte. Le pattern observé chez ces patients est une diminution significative de la compréhension des mots de classe fermée. Ainsi, ces patients ne peuvent plus utiliser la voie pré-lexicale qui est spécifique, rapide et automatique pour les mots de classe fermée (Friederici, 1985; Bradley, 1980; Friederici, 1986). A l'opposé, les patients dits anomiques, ou aphasiques de Wernicke, ont des problèmes pour identifier les mots de classe ouverte (Benson, 1979). Ces observations sur les patients aphasiques sont aussi retrouvées chez des patients dyslexiques. En effet, les patients présentant une dyslexie "profonde" ou bien encore une dyslexie phonologique ont des problèmes pour lire les mots de classe fermée alors que les mots de classe ouverte ne présentent que peu de perturbations (Coltheart, 1980; Marshall & Newcombe, 1966; Morton & Patterson, 1980). Les patients dyslexiques "de surface" présentent des processus pour les mots de classe ouverte qui sont perturbés (Fromkin, 1987). Nous pouvons noter qu'une double dissociation existe chez les patients aphasiques et dyslexiques: alors que les patients aphasiques de Broca et dyslexiques "profonds" présentent un déficit pour les mots de classe fermée, les patients aphasiques de Wernicke et dyslexiques "de surface" présentent des troubles pour le traitement des mots de classe ouverte. Différentes aires corticales semblent nécessaires pour identifier ces deux classes de mots. Comme les adultes sains, les patients aphasiques de Broca semblent avoir des difficultés lors de l'identification des mots de classe fermée (par rapport aux mots de classe ouverte): l'aire de Broca localisée dans l'hémisphère gauche en région antérieure semble donc cruciale pour l'activation d'une voie spécifique pour les mots de classe fermée.

#### 4.1.2.3 Acquisition et classe de mots

Etudier les propriétés et les principes qui gouvernent le développement de systèmes cognitifs peut guider l'élaboration des modèles théoriques sur le fonctionnement à l'âge adulte. En effet, si les structures cérébrales réagissent différemment lors de la maturation, alors il semble possible d'avancer l'hypothèse d'un degré de différenciation dans des sous-systèmes fonctionnels. Comme Fodor (1983) le suggère, l'un des critères pour définir un "module mental" est d'observer des décours temporels différents dans la maturation des zones corticales.

Il existe une distinction certaine entre la production et la perception des deux classes de mots chez les enfants. L'acquisition des connaissances morpho-syntaxiques doit être étroitement connectée à l'utilisation des mots de classe fermée. Cette hypothèse suggère que les mots de classe fermée doivent être analysés correctement très précocement et que certains indices pertinents doivent aider les nouveau-nés à classifier/distinguer les deux classes de mots. Puisque les mots de différentes classes ne sont qu'en partie définis par leur position relative par rapport aux autres éléments dans la phrase, un enfant ne peut sur les seules bases des statistiques de co-occurrence, catégoriser correctement les deux classes de mots. Ainsi, un enfant a besoin d'utiliser d'autres indices indépendants afin de distinguer les deux classes de mots. Shi, Morgan, & Allopenna (1998) pour les langues telles que le mandarin, le chinois et le turque, et Strömquist, Ragnarsdottir, & Richthoff (1984) pour le scandinave indiquent que les mots de classe fermée présentent des indices acoustiques et phonologiques différents (voyelle plus courte, amplitude plus faible, structure syllabique simplifiée) par rapport aux mots de classe ouverte. Le fait que ces indices existent n'indique pas forcément que les enfants puissent les utiliser afin de catégoriser correctement les deux classes de mots. Shi, Werker, & Morgan (1999) indiquent grâce à un paradigme de succion que les nouveau-nés (1 à 3 jours) utilisent les indices acoustico-phonologiques afin de discriminer les deux classes de mots. De plus, Shi & Werker (2001) montrent que pour l'anglais et dès 6 mois, les enfants préfèrent écouter les mots de classe ouverte. A l'opposé Höhle & Weissenborn (sous presse) signalent que pour une langue germanique cette préférence n'est présente qu'à partir de 7 mois. Suite à ces résultats, Morgan, Shi, & Allopenna (1996) suggèrent que les enfants utilisent les propriétés phonologiques pour regrouper correctement les mots dans les deux classes. Ainsi, les enfants focaliseraient leur attention sur les mots possédant des caractéristiques phonologiques spécifiques aux mots contenant des informations de sens, soit

les mots de classe ouverte. Pour ces auteurs, les enfants possèdent une certaine conception des mots de classe fermée lors de l'acquisition du langage.

Bloom (1970), Flores d'Arcais (1981) et Radford (1990) montrent que les enfants produisent plus tardivement certains mots de classe fermée par rapport aux mots de classe ouverte. Gleitman & Wanner (1982) suggèrent que cette absence dans la production des jeunes enfants provient du fait que ces mots sont réduits phonologiquement et, par conséquent, difficiles à reproduire. Friederici (1983; 1986) complète les résultats de (Bloom, 1970; Flores d'Arcais, 1981; Radford, 1990) à l'aide d'une expérience de recherche de mot. Dans une telle tâche, le participant reçoit une amorce avant la présentation de la phrase, et doit répondre le plus rapidement sur une clef de réponse dès qu'il entend la cible. Friederici (1983; 1986) montre que les enfants de moins de 8-9 ans détectent moins rapidement les mots de classe fermée que ceux de classe ouverte; ces enfants font aussi beaucoup plus d'erreurs pour ces items. Avant 11 ans, les informations sémantiques apportées par le contexte linguistique influencent le traitement des deux classes: les mots sont reconnus plus rapidement lorsque l'amorce et la cible sont reliés sémantiquement. Vers 11 ans, les enfants présentent des temps de réponses plus lents que ceux enregistrés auprès d'adultes, mais le pattern est semblable: 1) les mots de classe fermée sont reconnus plus rapidement que ceux de classe ouverte, 2) l'effet facilitateur d'un contexte lié sémantiquement est observé sur les mots de classe ouverte seulement, 3) aucun effet contextuel n'est observé sur la reconnaissance des mots de classe fermée. Ainsi, il semble, d'une part, que lors du développement du langage chez les enfants, les temps de réponses pour les mots de classe fermée diminuent progressivement et que, d'autre part, ce processus devient de plus en plus indépendant des contraintes sémantiques. Ces résultats suggèrent donc que les enfants mettent en place des processus de plus en plus automatiques et qu'en même temps, ils deviennent de plus en plus autonomes face à l'influence des informations sémantiques de type haut en bas. Le décours temporel du développement de l'utilisation des mots de classe fermée étant plus lent que celui de ceux de classe ouverte, Friederici (1983; 1986) conclut que ces mots sont traités par des structures cérébrales différentes.

#### 4.1.2.4 Hypothèse de Friederici (1980)

C'est à partir du modèle de Garrett (1975) que Friederici & Schoenle (1980) suggère un modèle similaire permettant d'expliquer les résultats obtenus à partir de différents paradigmes expérimentaux mais aussi à partir de différentes populations. Ainsi ces auteurs postulent que



deux lexiques mentaux différents sont activés lors de l'identification des mots de classe ouverte et fermée. Sur la base du modèle de (Forster, 1979; Forster & Ryder, 1971), Friederici & Schoenle suggèrent que la recherche dans le lexique mental s'effectue selon des mécanismes séquentiels. De plus, cette recherche s'effectuerait dans deux lexiques mentaux différents pour les deux classes de mots. Chaque lexique présente une organisation hiérarchique fréquentielle. Comme les mots de classe fermée sont plus fréquents et que le lexique qui les regroupe est plus petit (en nombre) que celui des mots de classe ouverte, alors l'identification des premiers sera plus rapide.

Le deuxième postulat de (Forster, 1979; Forster & Ryder, 1971) portant sur l'autonomie de l'accès au lexique est aussi retenu, mais d'une manière moins catégorique. En effet, dès 1986, Friederici montre avec une tâche de recherche de mots que le contexte affecte différemment les deux classes de mots. Les informations sémantiques apportées par le contexte linguistique accélèrent la reconnaissance des mots de classe ouverte, alors qu'aucune influence sur l'identification des mots de classe fermée n'est observée. Lorsque le contexte est neutre (sans charge sémantique) les deux classes de mots sont reconnues aussi rapidement. Ces résultats corroborent l'hypothèse d'un effet différencié lors de l'identification des mots de classe fermée dans un contexte de phrase par rapport à ceux de la classe ouverte: le contexte sémantique semble affecter l'identification des mots de classe ouverte de façon plus prononcée. Friederici (1986) utilise le concept suivant afin d'expliquer leurs résultats: les mots de classe fermée étant d'un nombre limité, l'espace correspondant à ce lexique est restreint, et la procédure de recherche dans ce lexique devrait prendre moins de temps que celle des mots de classe ouverte. De plus, lorsque suffisamment d'informations sémantiques sont présentes dans le contexte, le nombre de candidats de la classe ouverte peut alors diminuer, et les deux classes de mots seront reconnues aussi rapidement. Friederici conclut que l'identification des mots de classe fermée s'effectue de manière isolée et en complète autonomie (Forster, 1979; Forster & Olbrei, 1973) face aux informations sémantiques. Ainsi, elle suggère une certaine autonomie du traitement syntaxique. En d'autres termes, Friederici (1983) propose que les informations structurales provenant du contexte peuvent permettre d'influencer l'activation du lexique spécifique des mots de classe fermée. L'influence de ces informations dans les processus d'identification des mots de classe fermée ne viole pas l'hypothèse de l'autonomie car celles-ci n'affectent que les processeurs des mots de classe fermée. Forster (Forster, 1979; Forster & Olbrei, 1973) réfute l'idée selon laquelle le contexte (sémantique et/ou syntaxique) puisse intervenir lors de l'accès au lexique. Alors qu'une interprétation post-lexicale des effets de contexte sous-tend une autonomie stricte des

processus de traitement (Forster, 1979; Forster & Olbrei, 1973), l'hypothèse proposée par Friederici maintient ce postulat grâce à une activation en parallèle des deux types de traitement. Les informations contextuelles sémantiques influencent l'identification des mots de classe ouverte, alors que les informations contextuelles syntaxiques affectent l'identification des mots de classe fermée, aucune influence réciproque (le contexte sémantique ne peut affecter l'identification des mots de classe fermée et vice versa) n'est possible. Etant donné ces interprétations, Friederici suggère qu'il peut être possible de montrer d'une part, des dysfonctionnements de la voie spécifique des mots de classe fermée chez les patients dits agrammatiques aphasiques de Broca et d'autre part, que les enfants en bas âge, disposant d'informations limitées sur la structure des phrases, doivent présenter des patterns différents de ceux observés chez les adultes.

En effet, les résultats sur les adultes sains et les patients agrammatiques supportent l'hypothèse d'une distinction computationnelle de ces deux classes de mots. Alors que les patients avec une aphasie de Wernicke (lésion en région postérieure gauche) présentent un pattern similaire mais plus lent que les adultes sains (temps de réponses plus courts pour l'identification des mots de classe fermée que pour ceux de classe ouverte, ainsi qu'une influence du contexte sémantique sur la reconnaissance des mots de classe ouverte seulement), les patients aphasiques de Broca montrent un effet du contexte sémantique sur la reconnaissance des mots de classe ouverte uniquement (comme pour les adultes sains et les patients aphasiques de Wernicke) mais présentent des réponses plus lentes pour les mots de classe fermée que pour ceux de classe ouverte (Friederici, 1986). Ainsi, comme les adultes sains, les patients aphasiques de Broca semblent identifier les mots de classe fermée indépendamment des informations sémantiques, mais la voie spécifique de traitement pour ces mots est déficiente. En effet, ces auteurs proposent que le mécanisme nécessaire pour accéder rapidement aux mots de classe fermée est désorganisé par les dommages neurologiques de ces patients et que, par conséquent, un mécanisme unique de reconnaissance est utilisé pour les deux classes de mots. Ces résultats suggèrent que chez les patients aphasiques de Broca d'une part, la voie automatique d'accès aux mots de classe fermée est déficiente mais d'autre part, que cette voie fonctionne encore en complète autonomie (seules les informations contextuelles syntaxiques sont utilisées lors de la reconnaissance des mots de classe fermée).

### 4.1.3 Travaux de Bradley et collaborateurs

#### 4.1.3.1 Interférence visuelle

Bradley (1978) étudie l'effet d'interférence en modalité visuelle. Dans ce paradigme, les mots des deux classes sont accolés/enchâssés en initiale à des non-mots. Les participants doivent alors réaliser une tâche de décision lexicale sur ces items hybrides. Le point primordial de ce paradigme expérimental est que la portion initiale de lettres est suffisante pour initier l'accès à une représentation dans le lexique mental. Ainsi Taft & Forster (1976) trouvent que le temps de réaction obtenu pour les non-mots contenant des mots réels comme amorce initiale (*footmilge*) est plus long par rapport à des non-mots complets (*mowmilge*). Ces auteurs proposent que l'accès au lexique du mot *foot* a créé une interférence avec la tâche du participant qui était de rejeter le non-mot *footmilge*. Le système de reconnaissance des mots est temporairement induit en erreur, impliquant une augmentation dans le coût du traitement. Si le mot enchâssé est un mot de classe fermée (*suchmilge*) alors aucune augmentation du temps de décision lexicale n'est observée (Bradley, 1978). Sur la base de ces données, Bradley conclut que les mots de classe fermée ne semblent pas participer à la reconnaissance des mots comme ceux de classe ouverte. De plus, elle insiste sur le fait que ces deux classes de mots semblent montrer des processus d'accès lexicaux différents. Alors que Kolk & Blomert (1985) et Shapiro & Jensen (1986) répliquent cet effet d'interférence en modalité visuelle, et Matthei & Kean (1989) en modalité auditive, ces auteurs infèrent que les effets enregistrés à partir de ce paradigme reflètent des processus post-lexicaux plutôt que d'accès lexical pur.

Alors que pour les adultes sains, l'effet d'interférence semble être présent uniquement pour les mots de classe ouverte, Bradley, *et al.* (1980) montrent que chez les patients aphasiques de Broca celui-ci est présent pour les deux classes de mots. Ces résultats ont été répliqués par Kolk & Blomert (1985) et Matthei & Kean (1989) et par conséquent suggèrent que la voie spécifique de traitement des mots de classe fermée est endommagée chez ces patients.

#### 4.1.3.2 Présentation tachistoscopique

Quelques années plus tard, Bradley & Garrett (1983) étudient la distinction entre classes de mots, grâce à un paradigme en présentation tachistoscopique. Les stimuli sont présentés unilatéralement dans chaque hémichamp visuel. Un stimulus présenté à un hémichamp visuel est initialement projeté dans le cortex controlatéral. Avec une tâche de dénomination de mot, Bradley & Garrett (1983) rapportent différents taux de reconnaissance pour les deux classes de mots, uniquement lorsque ceux-ci sont présentés dans l'hémichamp visuel droit (HVD; ou hémisphère gauche, HG). Alors que les deux classes de mots sont aussi bien nommées dans l'hémichamp visuel gauche (HVG; ou encore hémisphère droit, HD), une asymétrie est donc présente dans l'hémichamp visuel droit (HVD-HG): le taux de reconnaissance pour les mots de classe ouverte est plus élevé que ceux de classe fermée. Ainsi, ces auteurs concluent que ces résultats reflètent des différences qualitatives dans l'identification des mots de classe ouverte et fermée entre les deux hémisphères: l'hémisphère gauche (HG) met en jeu l'activation d'un mécanisme spécifique de reconnaissance pour les mots de classe fermée. C'est à dire, les mots de classe fermée présentés dans l'hémichamp visuel droit (HVD-HG) sont identifiés avec une voie différente de celle qui opère les mots de classe ouverte pour ce même champ visuel. En revanche, dans l'hémichamp visuel gauche (HVG-HD) les stimuli de classe fermée et ouverte sont identifiés par un mécanisme similaire. Une remarque importante est à noter dès à présent et va ternir cette interprétation. Si l'hémisphère gauche (HG) utilise un processus spécifique d'accès au lexique pour les mots de classe fermée alors, il semble quelque peu étonnant de trouver que la reconnaissance des mots de classe fermée soit moins exacte que celle des mots de la classe ouverte. En effet, si une voie est spécialisée dans l'identification de l'une des classes de mots, l'on s'attend à ce que cette voie travaille de manière efficace et rapide. Suite à cette contradiction, Chiarello & Nuding (1987) présentent différentes expériences en présentation tachistoscopique à l'aide des tâches telles que la dénomination de mot, ou encore de décision lexicale. Leurs résultats montrent que les mots de classe fermée sont reconnus plus lentement que les mots de classe ouverte lorsque les stimuli sont présentés dans l'hémichamp visuel gauche (HVG-HD). Aucune différence n'est observée lorsque la présentation est réalisée dans l'hémichamp visuel droit (HVD-HG). Ce résultat indique que l'hémisphère droit (HVG-HD) semble avoir des difficultés pour traiter tout au moins certains mots de classe fermée. Ces résultats sont répliqués par Mohr, Pulvermuller, & Zaidel (1994). Alors que les études de Bradley & Garrett (1983) d'une part, et de Chiarello & Nuding (1987), Mohr, *et al.* (1994) d'autre part, semblent quelque peu contradictoires, il en émerge un résultat commun: les mots de classe fermée semblent être mieux reconnus (plus

justement et plus rapidement) par l'hémisphère gauche (HVD-HG) que par l'hémisphère droit (HVG-HD). Aucune latéralisation, ou seulement très faible, est observée pour les mots de classe ouverte. Ainsi, ces résultats convergent-ils vers l'idée que chacun des deux hémisphères dispose d'un lexique mental pour les mots de classe ouverte, alors que le lexique mental lié aux mots de classe fermée semblerait être restreint à l'hémisphère gauche (HVD-HG). Une entrée lexicale de classe fermée, si elle est présentée dans l'HVG-HD, doit alors transférer ces informations à l'hémisphère gauche HVD-HG, entraînant par conséquent un ralentissement dans les processus de traitement ainsi qu'un nombre plus important d'erreurs. Shapiro & Jensen (1986) utilisent le paradigme de présentation en tachistoscope avec un design de mots (de l'une des deux classes) enchâssés avec des non-mots. Ces auteurs montrent que seuls les non-mots constitués avec une initiale de mots de classe ouverte présentent un effet d'interférence (c'est à dire un temps de réponse plus long). De plus, les différences d'interférence entre les deux classes de mots sont significatives uniquement dans l'hémichamp visuel droit (HVD-HG). Ainsi ces auteurs concluent que le processus de reconnaissance de mots distinguant ces deux classes est mis en place uniquement lorsque les items sont présentés initialement à l'hémisphère gauche (HVD-HG). Ces résultats répliquent ceux de Kolk & Blomert (1985) et suggèrent l'activation d'une voie spécifique pour les mots de classe fermée uniquement lorsque ceux-ci sont présentés dans l'hémisphère gauche HVD-HG.

#### 4.1.3.3 Hypothèse de Bradley: classe dépendante (1978)

Bradley (1978) suggère que les deux classes de mots sont d'une part, stockées dans deux lexiques mentaux différents et d'autre part, que les processus d'accès aux représentations sont distincts. Grâce aux études en présentation tachistoscopique (Bradley & Garrett, 1983; Chiarello & Nuding, 1987; Mohr, *et al.*, 1994; Shapiro & Jensen, 1986), il semble possible de conclure à des différences qualitatives lorsque les deux classes de mots sont présentées sélectivement à l'hémisphère gauche (HG). Puisque les mots de classe fermée sont identifiés plus rapidement lorsque présentés à l'hémisphère gauche (HG), Bradley, *et al.* (1980) proposent que le lexique mental regroupant ces mots est localisé spécifiquement dans cet hémisphère (HG). En revanche, l'absence de latéralisation pour les mots de classe ouverte suggère que le lexique mental regroupant ces mots est plus diffus et est représenté dans les deux hémisphères.

Les résultats les plus influents dans l'étude de la distinction entre classes de mots et de la fréquence proviennent de Bradley et collaborateurs. Bradley (1978) et Bradley, *et al.* (1980) rapportent dans une tâche de décision lexicale visuelle que les mots de classe ouverte montrent un effet de fréquence: les sujets répondent plus rapidement pour les mots de haute fréquence. Cet effet n'est pas observé pour les mots de classe fermée. Afin d'expliquer ces différences, Bradley et ses collègues proposent une voie spécifique et indépendante de la fréquence pour l'accès au lexique des mots de classe fermée. Sur ce même paradigme, les patients aphasiques de Broca et de Wernicke présentent des résultats différents. Alors que les patients aphasiques de Wernicke manifestent des performances similaires aux adultes sains, les patients aphasiques de Broca montrent un effet de fréquence sur les mots de classe fermée qui est comparable à celui observé à partir des mots de classe ouverte. Les patients aphasiques de Wernicke identifient les mots de classe fermée de manière identique aux adultes sains (la vitesse ou la précision d'un tel processus ne sont pas affectées par la fréquence d'occurrence du mot). A l'opposé, les patients aphasiques de Broca identifient les mots de classe fermée comme s'ils appartenaient à la classe ouverte (la fréquence d'occurrence modifie la rapidité et la précision de l'identification). Bradley, *et al.* (1980) suggèrent l'existence de deux mécanismes distincts d'accès pour les deux classes de mots. Ils proposent qu'une voie est spécifique aux mots de classe ouverte et est dépendante de la fréquence. La seconde voie d'accès est réservée pour les mots de classe fermée et n'est pas sensible à la fréquence d'occurrence. Bradley, *et al.* (1980) indiquent que les patients aphasiques de Broca ont perdu la possibilité d'utiliser la voie spécifique aux mots de classe fermée, et qu'ils utilisent par conséquent la première voie d'accès sans distinction. L'architecture proposée par Bradley et collaborateurs ne diffère de celle de Friederici que par un seul trait: les mots de classe ouverte présentent une activation qui est proportionnelle à leur fréquence d'occurrence, alors que les mots de classe fermée sont insensibles à la fréquence. A l'opposé, Friederici admet que les deux classes de mots sont sensibles à la fréquence d'occurrence. Par conséquent, l'hypothèse de Bradley positionne la distinction entre les mots de classe ouverte et fermée à un niveau pré-lexical. Elle propose des processus spécifiques pour chaque classe: leur modèle est alors dépendant de la classe de mot.

Bradley suggère une architecture séparée pour ces deux classes de mots ainsi que la mise en place d'un processus automatique de la syntaxe: le contexte a une importance moindre car les informations lexicales sont encapsulées (Forster, 1979; Forster & Ryder, 1971). De plus, dans le cas où la structure métrique ne permet pas d'activer une classe de mots spécifique (Cutler, 1993; Cutler & Norris, 1988), les deux voies d'accès au lexique vont être activées en parallèle

et une compétition inter-classes va se mettre en place. Puisque le lexique mental des mots de classe ouverte regroupe un nombre plus conséquent d'items à vérifier, l'hypothèse de Bradley apporte une solution élégante qui permet d'expliquer comment les mots de classe ouverte peuvent être reconnus au moins aussi rapidement que ceux de la classe fermée. Comme la fréquence n'affecte que les mots de classe ouverte, les mots de classe fermée ne vont jamais atteindre ou dépasser le niveau d'activation des mots de classe ouverte.

## 4.2 Arguments en faveur d'une architecture commune: un seul lexique

Le fait de suggérer une architecture commune pour les deux classes de mots n'invalide pas l'importance d'une discussion sur la distinction entre ces deux classes de mots. En effet, la majorité des auteurs sont d'accord pour indiquer qu'il existe des différences psychologiques dans le traitement de ces deux classes de mot. Le problème est en réalité de déterminer le locus de ces différences: le niveau de distinction est-il lexical ou plus tardif, c'est à dire post-lexical? Les auteurs postulant un seul lexique de stockage pour les deux classes de mots suggèrent que la différence de traitement s'opère après la récupération et l'identification du mot, soit à un niveau post-lexical. Les modèles d'architecture commune pour les deux classes de mots ont émergé suite aux problèmes de réplication des données de Bradley et collaborateurs (Bradley, 1978; Bradley & Garrett, 1983; Bradley, *et al.*, 1980). Ainsi, des résultats contradictoires sur l'indépendance de fréquence pour les mots de classe fermée sont obtenus à partir de différents paradigmes expérimentaux, ou encore de différentes populations (Gordon & Caramazza, 1982; Gordon & Caramazza, 1983; Gordon & Caramazza, 1985; Matthei & Kean, 1989; Segui, Frauenfelder, Lainé, & Mehler, 1987; Segui, Mehler, Frauenfelder, & Morton, 1982). Ces résultats contradictoires sont une cause directe de l'émergence de nouvelles interprétations de la distinction entre les deux classes de mots.

### 4.2.1 Problème de réplication des résultats de Bradley et collaborateurs

Plusieurs réserves peuvent être émises face au type d'organisation proposé par Bradley et collaborateurs. Par exemple, le modèle de Forster (1979; Forster & Ryder, 1971) propose une recherche dans le lexique de manière fréquentielle et séquentielle. A l'opposé, Bradley suggère une indépendance fréquentielle pour le lexique contenant les mots de classe fermée. Kolk & Blomert (1985) indiquent que les hypothèses de Bradley sont en contradiction certaine avec un modèle de recherche fréquentielle. S'il existe un lexique mental spécial pour les mots de classe fermée, alors, vu leur nombre restreint, celui-ci doit être de taille réduite

comparé au lexique général contenant les mots de classe ouverte. Cette différence de taille de lexique devrait influencer l'effet de la fréquence. Selon l'hypothèse de Forster (1979; Forster & Ryder, 1971), l'effet de fréquence (le gain dans l'identification des mots de haute par rapport à ceux de basse fréquence) obtenu à partir d'une petite liste doit être moins important que celui observé dans une liste plus grande. Ainsi, si un lexique spécifique existe pour les mots de classe fermée, et si l'hypothèse d'une recherche séquentielle en fréquence est retenue, alors on doit trouver un effet de fréquence pour les mots de classe fermée qui est réduit par rapport à celui observé pour les mots de classe ouverte. Des résultats réfutant une telle hypothèse ont été obtenus à partir de différents paradigmes expérimentaux.

#### 4.2.1.1 Effet de fréquence sur les mots de classe fermée

Alors que Bradley (1978) montre un effet de la classe de mot avec une tâche d'interférence (où les mots de classe fermée associés en initiale de non-mots sont rejetés plus rapidement que ceux de classe ouverte), Taft & Forster (1976) indiquent, avec un paradigme similaire, que les mots de haute fréquence associés en initiale de non-mots sont plus rapidement rejetés que ceux de basse fréquence. Malheureusement, les mots utilisés par Bradley ne sont pas contrôlés en fréquence, et, par conséquent, nous ne pouvons conclure si leurs résultats reflètent une différence de traitement des mots de classes fermée et ouverte, ou bien plutôt un effet de la fréquence entre les deux classes comme le suggèrent les résultats de Taft & Forster (1976).

De plus, un certain nombre d'investigations n'ont pas répliqué les résultats de Bradley (1978), Bradley & Garrett (1983), Bradley, *et al.* (1980). Ainsi, à partir d'une tâche de décision lexicale, un effet de fréquence est trouvé sur les mots de classe fermée, en anglais (Besner, 1988; Garnsey, 1985; Gordon, 1983; Gordon, 1985, Gordon, 1982; Matthei & Kean, 1989), en français (Segui, *et al.*, 1982), et en hollandais (Kolk & Blomert, 1985). Par exemple, Matthei & Kean (1989) montrent que les temps de décision lexicale sont corrélés avec le log de fréquence mais aussi avec la durée des stimuli pour les deux classes de mots. Ces résultats, obtenus en modalité auditive, contrastent avec ceux de Bradley (1978), Bradley, *et al.* (1980), mais corroborent ceux de Besner (1988), de Garnsey (1985), de Gordon & Caramazza (1982; 1983; 1985), de Segui, *et al.* (1982), et de Kolk, & Blomert (1985).

A partir d'une tâche de décision lexicale, Gordon & Caramazza (1985) indiquent un effet de fréquence pour les deux classes de mots lorsque les fréquences sont inférieures à 400 occurrences/million (2,6 log de fréquence). Pour des mots de très haute fréquence ( $> 400$



occurrences/million,  $2,6 \log$  de fréquence), ils observent une différence d'effet de fréquence entre les deux classes de mots: les mots de classe fermée ne présentent plus d'effet de fréquence, contrairement à ceux de classe ouverte. L'asymétrie obtenue pour des basses fréquences rejoint les résultats de Bradley (1980) et de Bradley, *et al.* (1980). Gordon et Caramazza suggèrent que l'absence d'effet de fréquence sur les mots de classe fermée obtenue par Bradley peut être expliquée par d'une part, un nombre restreint d'items utilisés et d'autre part, des propriétés des mots non contrôlées. En effet, Bradley et collaborateurs utilisent des mots de classe fermée de fréquence supérieure à 400 occurrence/millions (pas d'effet de fréquence) alors que les mots de classe ouverte présentaient des fréquences inférieures à 400 occurrence/millions (présence d'un effet de fréquence).

A partir d'autres paradigmes, l'évidence d'une interaction entre fréquence et classe de mots semble aussi quelque peu confuse: Schmauder, Morris, & Poynor (2000) montrent que l'effet de fréquence influence le temps de lecture (temps de fixation) initial, pour les deux classes de mots. De plus, lorsque fréquence et longueur sont contrôlées, il ne semble pas y avoir de différence notable dans l'effet de fréquence entre les deux classes. En production, Segalowitz & Lane (2000) indiquent des temps de lecture plus rapide pour les mots de classe fermée comparés aux mots de classe ouverte. De plus, la fréquence influence différemment les deux classes de mots: l'effet de fréquence est plus important dans les basses par rapport aux hautes fréquences pour les mots de classe fermée seulement. Segalowitz & Lane (2000) montrent qu'en dessous de 300 occurrences par million ( $< 2.47 \log$  de fréquence), il existe un effet de fréquence robuste pour les deux classes de mots, résultat répliquant celui de Gordon & Caramazza (1982). En revanche, pour les mots de très haute fréquence, soit supérieure à 300 occurrences par million ( $> 2.47 \log$  de fréquence), seuls les mots de classe ouverte présentent l'effet de fréquence, résultat répliquant celui de Bradley (1978). Il semble donc exister une "frontière" au dessus de laquelle les mots de classe fermée ne réagissent plus face à leur fréquence. Sur la base de cette observation, Mohr, *et al.* (1994) présentent les mots de classe ouverte et fermée en présentation tachistoscopique avec deux groupes de fréquences différentes mais appariés pour les deux classes de mots: de très haute fréquence ( $> 400$  occurrences/million) ou bien de haute fréquence ( $< 400$  million, Francis & Kucera, 1982). Ces auteurs montrent tout d'abord un effet de la fréquence: les mots de très haute fréquence sont reconnus plus rapidement et avec plus d'exactitude que les mots de basse fréquence. En revanche, aucune interaction entre les facteurs fréquence et classe de mot n'est observée. L'avantage de l'hémisphère gauche (HVD-HG) pour les mots de classe fermée est observé quelle que soit leur fréquence, qu'ils soient de très haute fréquence ou bien de haute

fréquence. Les mots de classe ouverte de très haute ou de haute fréquence ne présentent aucune asymétrie. Alors que Segalowitz & Lane (2000) montrent qu'une sorte de plateau existe pour les mots de classe fermée au-dessus de 300 occurrences par million dans la rapidité de traitement (tâche de décision lexicale), Mohr, *et al.* (1994) indiquent que la fréquence d'occurrence n'affecte pas la qualité de traitement des mots de classe fermée.

Grâce aux études portant sur les adultes sains, nous avons mis en évidence que seuls Bradley et collaborateurs ont réussi à montrer dans une tâche de décision lexicale l'insensibilité des mots de classe fermée face à la fréquence d'occurrence (Bradley, 1978; Bradley & Garrett, 1983). A l'opposé, de nombreuses études n'ont pu répliquer ce résultat, et ce dans différents paradigmes expérimentaux tel que la décision lexicale (Besner, 1988; Garnsey, 1985; Gordon & Caramazza, 1982; Gordon & Caramazza, 1983; Gordon & Caramazza, 1985; Kolk & Blomert, 1985; Matthei & Kean, 1989; Segui, *et al.*, 1982) et en temps de lecture (Schmauder, *et al.*, 2000). Il est à noter que certaines études n'ont pu répliquer les résultats de Bradley (1978) et de Bradley, *et al.* (1980) sur des patients aphasiques de Broca (Gordon & Caramazza, 1983).

#### 4.2.1.2 Hypothèse de Gordon & Caramazza (1982)

Gordon et collaborateurs ont développé un modèle (Gordon, 1983) présentant des propriétés d'accès au lexique en parallèle (Morton, 1969; Morton, 1979). Dans une série d'expériences (Gordon & Caramazza, 1982; Gordon & Caramazza, 1983; Gordon & Caramazza, 1985), ces auteurs montrent que les deux classes de mots peuvent être sensibles à la fréquence d'occurrence chez des patients aphasiques de Broca, mais aussi des adultes sains. Selon eux, chaque mot présente une sensibilité à la fréquence d'occurrence et les mots de classe ouverte et fermée sont représentés dans un même lexique mental. Selon Gordon (1983), tous les mots présentent un niveau d'activation en fonction de leur fréquence d'occurrence, impliquant que l'on généralise aux mots de classe fermée les mécanismes utilisés pour les mots de classe ouverte. Cette hypothèse d'architecture commune pour les deux classes de mots s'oppose à celle de Bradley et collaborateurs (Bradley, 1978; Bradley, *et al.*, 1980) dans laquelle les processus de traitement des mots de classe fermée seraient indépendants de la fréquence d'occurrence. Selon cette hypothèse, les résultats des patients aphasiques de Broca mettent en évidence une perte des processus syntaxiques plutôt que lexicaux. Des résultats en faveur de ce type d'organisation proviennent de ceux obtenus par Forster (1981a). Ces auteurs comparent des latences de dénomination, pour des mots de haute fréquence présentés à partir

soit d'une liste pure (où seuls des mots de haute fréquence sont présentés), soit d'une liste mixte (où des mots de haute mais aussi de moyenne et de basse fréquence sont présents). Ils montrent des latences équivalentes pour les deux listes. L'argument sous-jacent est que si le lexique mental est organisé en fréquence d'occurrence (deux lexiques différents: l'un pour les hautes et l'autre pour les basses fréquences), alors exécuter une recherche dans un seul des lexiques devrait prendre moins de temps que rechercher dans deux lexiques différents. Bien qu'un résultat positif (c'est à dire une différence de rapidité entre les deux types de listes, voir section 2.1 de ce chapitre) ait été obtenu par Glanzer & Ehrenreich (1979) dans une tâche de décision lexicale, Forster (1981a) conclut que la recherche lexicale s'effectue dans un seul lexique mental débutant par ordre de fréquence, où les entrées lexicales de haute fréquence sont traitées en premier.

Gordon & Carramazza (1982) aboutissent aux mêmes conclusions que Forster (1981a) et indiquent que les deux classes de mots sont représentées dans un même lexique et que les processeurs sont identiques. Ces hypothèses n'impliquent pas forcément que le produit de ces processeurs soit identique. En effet, comme les logogènes possèdent des seuils différents en fonction de leur fréquence, alors les mots de classe fermée, qui sont plus fréquents, seront identifiés plus rapidement que les mots de classe ouverte.

## 4.2.2 Hypothèses alternatives

### 4.2.2.1 Hypothèse de Shillcock & Bard (1993)

Shillcock & Bard (1993) suggèrent que les deux classes de mots sont stockés dans un même lexique mental, mais que les informations de type haut en bas, telles que les informations contextuelles, peuvent influencer l'accès lexical des mots de classe fermée. A partir d'un paradigme de décision lexicale bi-modal, Shillcock & Bard (1993) utilisent des phrases dans lesquelles sont incluses un homophone avec deux significations possibles, soit un mot de classe ouverte, soit de classe fermée (*would/wood*). Immédiatement après la présentation auditive de ces phrases, les sujets visualisent un mot cible qui est synonyme du mot de la classe ouverte (*timber*). Les résultats révèlent que le temps pour réaliser une décision lexicale sur ce synonyme (*timber*) est plus court lorsque le contexte induit le mot de classe ouverte (*wood*) que lorsqu'il induit celui de classe fermée (*would*). Puisque aucune facilitation n'est observée lorsque le contexte biaise l'utilisation d'un mot de classe fermée, les auteurs concluent que les sujets sont capables de prédire l'apparition d'un mot de classe fermée à

partir du contexte, alors que ce résultat n'est pas possible pour les mots de classe ouverte. Shillcock & Bard (1993) suggèrent que le contexte influence différemment la reconnaissance des deux classes de mots. Pour ces auteurs, ces résultats avancent des arguments en faveur d'un processus autonome pour le traitement des mots de classe ouverte alors que le traitement des mots de classe fermée semble être perméable/interactif tout au moins aux informations syntaxiques. Shillcock & Bard (1993) indiquent que si les informations syntaxiques locales favorisent la reconnaissance des mots de classe fermée, il semble moins primordial de déterminer si les deux classes de mots sont stockées dans deux lexiques séparés ou non. En effet, l'influence du contexte sur la reconnaissance des mots de classe fermée va engendrer un processus plus rapide que l'accès au lexique des mots de classe ouverte.

Haveman (1997) réplique l'expérience de Shillcock & Bard (1993) mais présente en plus en fin de phrase un synonyme de l'homophone dont le contexte biaise le sens du mot de classe fermée (condition absente du paradigme de Shillcock & Bard, 1993). Avec un ISI (intervalle inter-stimulus) de 400 ms, elle observe un effet du contexte sur chaque classe de mots ouverte et fermée. Le contexte sélectionne donc le sens approprié de l'homophone qu'il soit de classe ouverte ou fermée. En revanche, en diminuant le ISI entre l'amorce et la cible à zéro ms, Haveman (1997) montre que les deux utilisations de l'homophone sont activées. Ainsi, elle conclut que le contexte n'agit pas à un niveau pré-lexical mais plutôt lors de processus tardifs d'intégration des mots dans le contexte. L'hypothèse selon laquelle le contexte syntaxique peut aider à dissocier les deux classes de mot est exclue par Haveman (1997).

#### 4.2.2.2 Hypothèse de Herron & Bates (1997)

Une alternative aux deux théories précédentes Gordon & Caramazza (1983) et Shillcock & Bard (1993) est proposée par Herron & Bates (1997) et Herron (1998). Dans ce modèle, les processus de compréhension du langage ne présentent pas de division lors de l'identification des mots de classe fermée et ouverte à quelque niveau que ce soit. Au lieu de postuler des processus séparés ou encore des lexiques distincts, ces auteurs suggèrent que les deux classes de mots sont stockées dans un même lexique et que les informations de type phonologique, mais aussi contextuel vont pouvoir influencer simultanément l'accès au lexique. Les informations de type bas en haut et haut en bas vont permettre une intégration rapide et sont utilisées en même temps. Par conséquent, les différences de traitement (résultats) sont les conséquences des propriétés des deux classes de mots (telles leur fréquence, leur salience, leur imagibilité, etc.), et ne proviennent pas des processeurs distincts qui vont contrôler l'accès au

lexique (Bates & Wulfeck, 1989). Herron (1998) et Herron & Bates (1997) postulent donc une complète interactivité pour l'accès au lexique où toutes les informations (phonologiques et contextuelles) peuvent être utilisées simultanément.

## 5 Conclusions

Dans ce chapitre, nous avons examiné les principaux processus qui sont impliqués dans la reconnaissance visuelle des mots et considéré quelques facteurs qui peuvent faciliter ou non cette reconnaissance. Incontestablement, la fréquence d'occurrence d'un mot est un facteur essentiel pour déterminer sa vitesse de traitement, et donc, est une variable implémentée dans la plupart des modèles psycholinguistiques portant sur la reconnaissance des mots. La dichotomie entre les mots de classes ouverte et fermée indique que deux variables distinctives sont aussi importantes, la fréquence d'occurrence et la longueur: les mots de classe fermée sont plus fréquents et plus courts que ceux de classe ouverte.

Les mots de classe fermée sont généralement identifiés plus rapidement par rapport aux mots de classe ouverte. Un modèle postulant deux formes de stockage ou deux voies d'accès au lexique différentes peut par conséquent être proposé (Bradley, 1978). En revanche, lorsque les mots de classe fermée sont incorporés dans des phrases, Gordon & Caramazza (1982) indiquent que ce type de modèle peut être critiqué. En effet, ces auteurs montrent que l'effet de fréquence est variable selon la fréquence d'occurrence étudiée (supérieure ou inférieure à 400 occurrences/million). Pour les hautes fréquences, seul les mots de classe ouverte présentent un effet de fréquence robuste alors que pour des fréquences inférieures à 400 occurrences/million, cet effet est présent pour les deux classes de mots. La prise en compte des fréquences d'occurrence lors de l'étude de la distinction entre classes de mots semble être primordiale afin d'éclairer le débat. Les résultats comportementaux obtenus à partir d'items non contrôlés en fréquence peuvent ne pas refléter des différences qualitatives dans le traitement de ces deux classes de mots, mais plutôt des différences quantitatives dans les propriétés lexicales physique et/ou distributionnelle (longueur et fréquence).

Le contexte linguistique dans lequel est présenté un mot influence lui aussi les processus de reconnaissance de ce dernier. Ainsi, une question pertinente lors de l'élaboration de ces modèles est de déterminer au niveau de quelle étape cette influence est effective. Selon le type de modèle développé, le contexte affecte soit les processus perceptifs lors du traitement du stimulus et, par conséquent, son accès au lexique (modèles interactifs, Morton, 1969), soit il

affecte les processus post-lexicaux, comme les mécanismes d'attente, d'intégration ou autres (modèles autonomes, Forster, 1976). A partir de ces deux types de modèles, six différentes hypothèses ont été émises afin de déterminer plus précisément l(es) étape(s) de la reconnaissance spécifique(s) à chaque classe de mot. Celles-ci sont résumées dans le tableau 1 et s'opposent sur 3 différents critères. (1) La nature des représentations: existe-t-il un ou plusieurs lexiques pour les deux classes de mots? (2) La nature des processeurs activés: les mots de classe fermée sont-ils accédés différemment de ceux de la classe ouverte? (3) La nature des effets de contexte: le contexte influence-t-il l'identification des deux classes de mots?

Auteurs	lexiques différents	processeurs différents	effet du contexte sur CF	effet du contexte sur CO	autonomie
Garrett (1975)	+	+	-	-	+
Friederici (1980)	+	+	+	-	+
Bradley (1978)	+	+	-	-	+
Gordon & Carramazza (1982)	-	-	?	?	?
Shillcock & Bard (1993)	-	-	+	-	+
Herron & Bates (1997)	-	-	+	+	-

**Tableau 1:** Résumé des différentes hypothèses portant sur la distinction entre classes de mots.

Alors que trois hypothèses, Bradley (1978), Friederici & Schoenle (1980) et Garrett (1975) indiquent que les deux classes de mots sont stockées séparément dans le cerveau, Gordon & Carramazza (1982), Herron & Bates (1997) et Shillcock & Bard (1993) postulent l'existence d'un même et unique lexique mental pour les deux classes de mots. Les hypothèses revendiquant un stockage dans deux lexiques séparés pour les deux classes de mots admettent en général que les processeurs analysant ces mots sont aussi différents. En effet, il semble difficile à partir de résultats comportementaux de conclure plus précisément sur cette distinction entre stockage et processeurs.

Bradley (1978) et Garrett (1975) suggèrent une stricte autonomie de l'accès au lexique pour les deux classes de mots: le contexte n'influence en aucune sorte l'accès au lexique. Bien que les hypothèses de Friederici & Schoenle (1980) et de Shillcock & Bard (1993) s'opposent sur le nombre de lexiques connectés pour les deux classes de mots, elles s'accordent pour postuler que le contexte syntaxique peut influencer l'accès au lexique des mots de classe fermée uniquement. Le contexte ne peut affecter la reconnaissance des mots de classe ouverte. Ainsi, une certaine autonomie est avancée car les informations contextuelles syntaxiques ne peuvent influencer que la reconnaissance des mots de classe fermée. Gordon & Carramazza (1982)

n'indiquent par leur position quant à l'effet du contexte sur la reconnaissance des deux classes de mots. Grâce à un fonctionnement interactif, Herron & Bates (1997) postulent que le contexte ainsi que les informations provenant du décodage perceptif des mots peuvent affecter la qualité ainsi que la vitesse d'identification des deux classes de mots.

Afin de décrire le décours temporel d'utilisation des différentes informations apportées par l'entrée lexicale, il semble primordial d'enregistrer des mesures qui reflètent la rapidité des effets enregistrés. Comme l'indique Forster (1989), les mesures de temps de réaction enregistrent une sommation d'événements ou processus difficiles à cerner et à définir.

*"We have by now (one hopes) abandoned the notion that our battery of reaction-time tasks provides us with straightforward measures of the time taken to perform certain elementary mental operations, and we recognize that even the simplest task may involve a highly complex orchestration of component processes."*

Malheureusement, l'interprétation des résultats provenant du domaine de recherche portant sur la compréhension du langage est particulièrement confuse en raison des différentes tâches utilisées. En effet, en demandant à des sujets de réaliser une tâche de décision lexicale ou encore une tâche de dénomination de mots ou d'images, on étudie non seulement les processus de compréhension du langage, mais aussi les processus liés au traitement demandé. Les tâches semblent interférer avec les processus en question. Les mesures en temps réel sont des dimensions qui font intervenir des informations de sources multiples, linguistiques comme non-linguistiques. Ainsi, des processus post-lexicaux peuvent influencer la réponse observée, qui normalement serait sensée refléter le fonctionnement de sous-systèmes automatiques et, par conséquent, qui ne devraient pas être affectés par ce type de sources. Par exemple, les résultats obtenus à partir de tâches de décision lexicale sont particulièrement critiquables car cette tâche semble sensible aux mécanismes post-lexicaux. La stratégie, la motivation, ou encore l'attention que portent le sujet aux stimuli influencent cette mesure (Balota & Chumbley, 1984; Seidenberg, Waters, Sanders, & Langer, 1984). Afin d'étudier ces sous-systèmes, il semble, par conséquent, nécessaire d'élaborer des techniques strictes testant directement les hypothèses, tout en neutralisant les sources d'informations étrangères au traitement en question.

Grâce à l'apport d'une technique d'imagerie cérébrale nous allons étudier la distinction entre classes de mots. Deux caractéristiques de cette technique vont être mis en avant dans notre

recherche. D'une part, grâce à son excellente résolution temporelle, la technique des potentiels évoqués employée va permettre de suivre le déroulement des différentes étapes activées lors de la reconnaissance des deux classes de mots. De plus, puisque cette technique peut être utilisée sans tâche additionnelle, nous allons pouvoir mettre en évidence, s'il existe une différence dans la reconnaissance des deux classes de mots.



# CHAPITRE 2

## POTENTIELS EVOQUES ET CLASSE DE MOTS

---

### *Résumé*

L'un des arguments en faveur d'une organisation neuronale distincte (et par extension cognitive) pour les deux classes de mots est l'observation de patterns électrophysiologiques différents. Après un bref rappel sur la technique de l'électroencéphalogramme (EEG) et des potentiels évoqués (PEs), nous décortiquerons les résultats confus de cette littérature. Le succès de l'utilisation des méthodes psychophysiologiques pour étudier les représentations et processus lors de la reconnaissance des deux classes de mots requiert de nombreux contrôles dans l'expérimentation. En effet, l'étude de l'expérience princeps de Kutas & Hillyard (1983) soulève différents problèmes: méthodologique (fenêtre d'analyse large et nombre de capteurs réduits), nature des stimuli (propriétés confondues avec la classe de mots) et la tâche employée (lecture de phrase présentant des erreurs). La revue de la littérature sera donc réalisée suivant ces 3 points cruciaux et permettra d'éclaircir si certains corrélats électrophysiologiques peuvent être définis comme étant spécifiques à l'une ou l'autre classe de mots. Neville, Mills, & Lawson (1992), en identifiant une composante précoce la N280 (négativité maximale en région antérieure gauche vers 280 ms) comme spécifique lors du traitement des mots de classe fermée, ont ouvert un nouveau champ de recherche au sein du débat concernant le statut biologique des deux classes de mots. La composante N280 est-elle un indice spécifique reflétant les caractéristiques des mots de classe fermée, ou bien reflète-t-elle les modulations des propriétés lexicales physique (longueur) et distributionnelle (fréquence d'occurrence) qui existent entre les deux classes de mots?

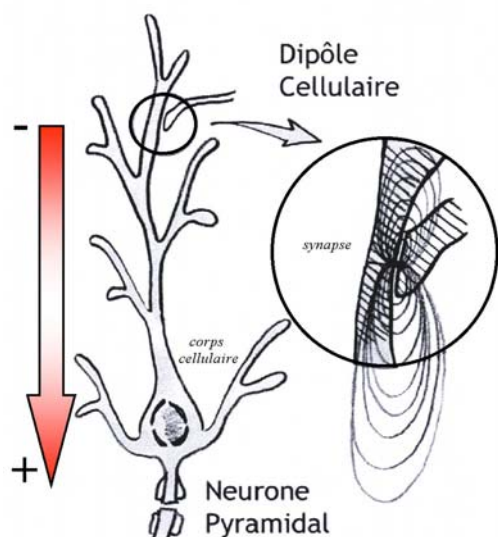
## 1 Bases neuronales de l'activité électroencéphalographique

La manifestation d'un processus neuronal est caractérisée par l'activité électrique des neurones dans les zones corticales et sous corticales concernées. L'information est transférée à travers les neurones, et intégrée à l'intérieur de ceux-ci, via les courants qui circulent à travers les synapses cérébrales activées. Les flux de courants aboutissent à des courants extracellulaires, aussi appelés courants primaires. La sommation des différents courants primaires (provenant de neurones distincts) provoque la circulation de courants plus importants dit de volume, ou de conduction. Les courants de conduction, comme l'indique leur nom, peuvent se propager dans tout le cerveau à travers le liquide céphalo-rachidien, les os du crâne et le scalp. Ainsi, la diffusion de courants de volume induit des différences de potentiel (ddp) qui pourront être enregistrées par des électrodes positionnées sur le scalp. Suite à ce phénomène électrique, les neurones activés ont dépensé de l'énergie. Ainsi, ces cellules vont devoir reformer leur stock de molécules énergétiques (ATP, adénosine tri-phosphate) grâce à l'apport par les vaisseaux sanguins environnant de substances nutritives. Les ressources nécessaires à la poursuite de l'activité vont être apportées grâce à une augmentation d'une part, du débit sanguin local dans les artérioles et d'autre part, d'oxyhémoglobine dans les capillaires. Ces deux derniers événements représentent les réponses hémodynamiques, corrélats indirects et secondaires de l'activité neuronale que vont pouvoir enregistrer la TEP (Tomographie par Emission de Positons) et l'IRMf (Imagerie par Résonance Magnétique Fonctionnelle).

Lorsque deux électrodes sont positionnées sur le cuir chevelu et connectées à un amplificateur différentiel, la sortie de l'amplificateur révèle une fluctuation de voltage au cours du temps. Ce pattern électrique est connu sous le nom d'électroencéphalogramme ou EEG. L'amplitude normale d'un EEG peut varier entre  $\pm 100 \mu\text{V}$ , et sa fréquence jusqu'à 40 Hz et plus (Berger, 1929). L'activité EEG est la manifestation directe à distance des courants primaires: elle est le résultat d'une propagation électrique passive et instantanée émanant de synapses cérébrales actives et aboutissant à une électrode d'enregistrement située au niveau du cuir chevelu.

## 1.1 Notion de dipôle unitaire

Un potentiel d'action (PA) arrivant à une synapse libère des neurotransmetteurs qui ouvrent, au niveau des membranes neuronales post-synaptiques, des canaux ioniques spécifiques. Les ions afflueront dans ces orifices, ce particulièrement si la cellule activée est excitatrice; dans ce cas, le courant transmembranaire sera porté par l'entrée des ions  $\text{Na}^+$  dans la cellule. Une fois dans la cellule, la répartition des ions reflètera la géométrie de celle-ci; par conséquent un flux asymétrique d'ions positifs va migrer vers la partie centrale (corps cellulaire) du neurone, créant un courant électrique intra-cellulaire. Puisqu'un mouvement d'ions vient de se réaliser à l'intérieur du neurone post-synaptique, en partie distale, un puits de courant est constitué. Ce "mini" courant intra-cellulaire est accompagné par l'émission d'un champ magnétique perpendiculaire à celui-ci. En même temps, à l'extérieur du neurone post-synaptique, au niveau de la synapse, il s'est créé une dépression d'ions  $\text{Na}^+$ <sup>1</sup> suscitant un autre courant local extra-cellulaire, appelé courant primaire. A l'opposé, la membrane de ce neurone (partie centrale, vers le soma) va se comporter comme une source de courant<sup>2</sup> (Halgren, Clarke, & Hervé, 1992). La résultante de ces courants (figure 3) peut être schématisée par un dipôle unitaire orienté du puits vers la source: la base de la flèche étant localisée au niveau du puits (charge négative) et la pointe, à la source (charge positive). Si la synapse est inhibitrice, alors les ions rentrant dans la cellule post-synaptique sont des ions négatifs, et l'orientation du dipôle sera opposée.



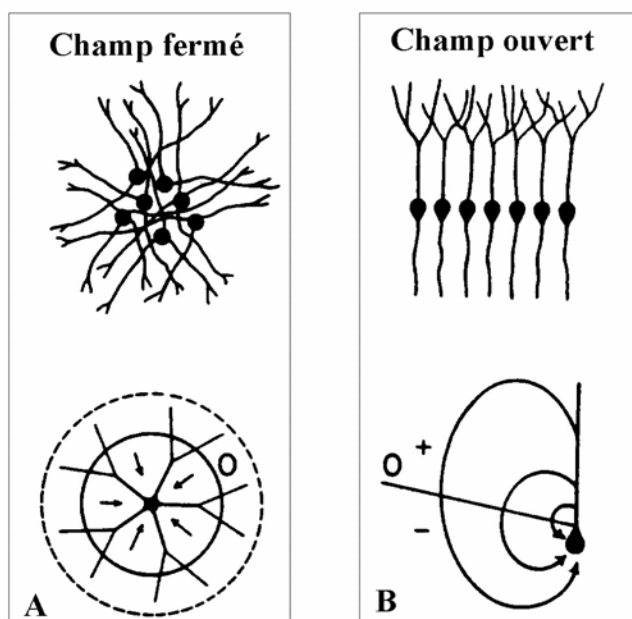
**Figure 3:** Modélisation dipolaire de l'activation d'un neurone. Lors de l'activation, un puits de courant apparaît au niveau de la synapse, et une source de courant au niveau du corps cellulaire. L'accumulation à une distance faible de ces charges opposées conduit à générer un dipôle unitaire (flèche rouge). D'après Thierry (2000).

<sup>1</sup> Comme des ions  $\text{Na}^+$  viennent de rentrer dans le neurone post-synaptique, leur concentration extracellulaire diminue.

<sup>2</sup> La concentration en  $\text{Na}^+$  à l'extérieur du neurone est légèrement diminuée, attirant par conséquent de nouveaux ions  $\text{Na}^+$  dans l'environnement proche. Ces ions vont provenir de la zone extra-cellulaire (partie centrale).

## 1.2 Dipôles équivalents et propriétés des neurones

Un courant de conduction est le résultat de la sommation des différents dipôles unitaires qui sont enregistrés en synchronie et peuvent être orientés dans toutes les directions de l'espace. Au niveau du scalp, le courant de conduction est représenté par un unique dipôle résultant de la sommation totale de l'activité électrique au temps X. Les propriétés des courants de conduction dépendent de la configuration des neurones activés, de leur orientation, de leur emplacement, mais aussi des conductivités des différents tissus cérébraux traversés. Nous voyons directement sur la figure 3, que le caractère dipolaire d'un neurone dépend de sa configuration spatiale. Ainsi, si l'arbre dendritique est symétrique (cas des neurones associatifs, non représenté ici), le dipôle unitaire correspondant sera de faible intensité, contrairement au cas du neurone pyramidal. Selon Lorente de No (1947; cité par Coles, 1990), si les neurones sont organisés en structure nodulaire concentrique, leur activité peut induire des courants locaux constituant un champ de potentiel fermé ou co-activé, non enregistrable en surface. Ainsi le dipôle équivalent sera de faible amplitude (figure 4 A).

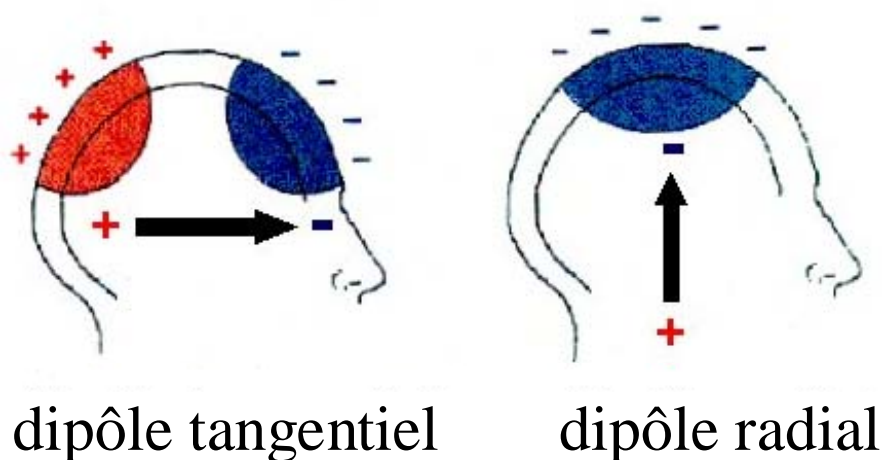


**Figure 4:** A "champ fermé": organisation neuronale en nodules engendrant un dipôle équivalent de faible amplitude, donc non enregistrable à la surface du cuir chevelu. B "champ ouvert": organisation neuronale en couches engendrant un dipôle équivalent enregistrable à distance. D'après Lorente de No (1947).

En revanche, l'organisation neuronale en couches typique du cortex cérébral, avec les corps cellulaires et les dendrites d'une couche orientés de la même façon, est propice à la genèse de champs ouverts (figure 4 B). De tels champs s'étendent au-delà des limites de la population neuronale active et produisent une différence de potentiel entre la surface et la profondeur du cortex (courant de conduction). Les configurations spatiales en champs ouverts sont observées dans la surface corticale où de nombreux neurones pyramidaux sont organisés en macro-

colonnes (environ 3 mm sur 3 mm) perpendiculaires à la surface du cortex. Ainsi, l'EEG permet d'enregistrer l'activation des neurones pyramidaux présents dans cette surface cérébrale.

Les caractéristiques des dipôles équivalents vont dépendre de l'orientation spatiale des différents dipôles unitaires. Ainsi, un dipôle équivalent parallèle (tangential) à la surface du scalp donnera lieu en surface à un potentiel constitué d'un pôle positif et négatif (figure 5). En revanche, si le dipôle équivalent est perpendiculaire (radial) à la surface du scalp, alors seule son extrémité négative ou positive sera observable.



**Figure 5:** Distribution à la surface du scalp des champs de potentiels qui sont produits par deux dipôles différents: tangential et radial. Dans le cas d'un dipôle tangential, les extrémités positive (représentée en rouge) et négative (représentée en bleu) du dipôle équivalent sont détectées par les électrodes de surface. A l'opposé, pour un dipôle équivalent radial, seule l'extrémité en surface peut être détectée, le champ émis par l'autre extrémité se situe en profondeur.

Ce dipôle équivalent, ou courant conducteur se propage depuis la source (générateur) jusqu'à la surface corticale et, par conséquent, est totalement déterminé par la conductance du tissu qu'il traverse. Dès à présent, nous pouvons noter que les enregistrements magnéto-encéphalographique (MEG) ne souffrent pas de distorsion du signal comme l'EEG, car cette technique enregistre une source différente de celle de l'EEG. En effet, consécutif à l'émission des courants intra-cellulaires, nous avons vu qu'un courant magnétique perpendiculaire est émis. C'est ce type d'activité magnétique que la technique du MEG enregistre. Contrairement au dipôle unitaire que l'EEG capte, les courants magnétiques ne sont pas distordus.

Finalement, nous avons vu que les dipôles équivalents représentent la sommation de dipôles unitaires qui, eux-mêmes, sont le reflet de courants post-synaptiques plutôt que de potentiels d'action axonaux (PA; Allison, Wood, & McCarthy, 1986). Afin que ce principe de sommation s'applique à un large ensemble de neurones, il ne suffit pas que leur organisation géométrique soit satisfaite, il faut aussi que les neurones soient activés en synchronie.

## 2 Les Potentiels Evoqués (PEs): notions de base

### 2.1 Définition

Un enregistrement EEG est le reflet de l'activité électrique spontanée du cerveau. Maintenant, supposons que nous présentions à un sujet humain un stimulus spécifique (visuel, auditif, ou autre) et que nous enregistrons son EEG. Nous pouvons définir une fenêtre dans cet EEG qui est synchronisée avec le début de la présentation du stimulus. Dans cette fenêtre temporelle, nous allons pouvoir mettre en évidence des fluctuations de voltage qui sont directement reliées à la présentation du stimulus. Ce sont ces fluctuations de voltages (dépendant du stimulus) que constituent les ERPs, provenant de l'anglais *Event-Related Potentials*, ou encore Potentiels Evoqués (PEs). L'hypothèse de base est que ces PEs sont la manifestation des processus psychologiques spécifiquement mis en jeu lors du traitement des événements/stimuli qui les déclenchent.

### 2.2 Aspects méthodologiques

L'activité EEG est recueillie sur le scalp au moyen d'électrodes placées à différents sites et reliées à des amplificateurs. Pour chaque essai, l'EEG est synchronisé avec l'apparition d'un événement (stimulus ou acte moteur). Les données enregistrées au cours des essais d'une même condition expérimentale sont ensuite sommées puis divisées par le nombre d'essais. Cette technique de moyennage améliore le rapport signal (PEs) sur bruit (activité EEG spontanée) car les PEs sont de faible amplitude (quelques  $\mu\text{V}$ ) par rapport à l'activité EEG spontanée (plusieurs dizaines de  $\mu\text{V}$ ). Ainsi deux principes ressortent de la technique de moyennage. D'une part, plus nombreux seront les essais, plus le rapport signal/bruit sera bon: c'est donc le caractère reproductible de la réponse évoquée qui permet son extraction de l'activité spontanée. D'autre part, seuls les changements de potentiels temporellement corrélés

avec l'événement émergent du bruit de fond constitué par l'électrogénèse corticale (Coles & Rugg, 1995). Nous verrons dans le chapitre 3 les spécificités techniques d'un enregistrement PEs.

Les variations de potentiels corticaux, associées à un événement donné, représentent l'activation de populations neuronales successives qui impliquent l'activation de processus psychologiques pendant le traitement de l'événement. La décomposition des PEs en "composantes", repose sur le principe d'une inférence psychophysiological constituant à attribuer à chaque composante un rôle fonctionnel particulier, en référence aux processus sous-jacents à une tâche ou encore à un stimulus particulier. Une composante peut être définie en termes descriptifs, comme une variation de potentiel caractérisée par sa polarité (par rapport à un niveau de référence), sa latence (par rapport au début de la présentation du stimulus), son amplitude (par rapport à un niveau de base) et sa topographie (répartition sur le scalp). La nomenclature la plus utilisée repose sur cette définition descriptive d'une composante. Une composante peut être nommée par une lettre indiquant la polarité (P pour positif ou N pour négatif), par sa latence moyenne en millisecondes (par exemple: N100, P300), ou enfin, par son rang d'apparition dans la succession des composantes de même polarité (ex: P1, première composante positive, etc.). Dès 1978, Donchin & Heffley (1978) utilisent le terme de composante non plus pour désigner des phénomènes identifiés par inspection visuelle du tracé (approche descriptive), mais plutôt comme un concept théorique. En effet, on peut déduire la fonctionnalité d'une composante qui n'est enregistrée qu'en présence d'une des deux conditions expérimentales. Sutton, Braren, Zubin, & John (1965) différencient deux types de PEs, en fonction de leur latence, et des manipulations expérimentales auxquelles ils sont sensibles. Les PEs précoces ou exogènes (de latence inférieure à 200 ms) sont associés à l'analyse des caractéristiques physiques du stimulus (par exemple la durée ou l'intensité du son). Ils sont facilement reproductibles d'une session à une autre chez un même sujet et sont relativement insensibles aux changements de l'état psychologique du sujet au cours de l'expérience. En revanche les PEs endogènes (de latence supérieure à 200 ms) reflètent des étapes plus tardives du traitement de l'information. Des facteurs tels que la nature des instructions données au sujet, les décisions que celui-ci prend, ou encore la probabilité d'occurrence du stimulus, influencent la latence et/ou l'amplitude des composantes endogènes. En réalité la différenciation entre composantes exogènes et endogènes n'est pas toujours aussi aisée.

## 2.3 Contribution à l'étude des processus psycholinguistiques

Le but de la psychologie cognitive est d'identifier les processus cognitifs qui s'échelonnent entre un stimulus quelconque et une réponse comportementale. Pour cela, les chercheurs se doivent de définir les représentations sur lesquelles les processeurs opèrent, la manière dont elles interagissent (ou travaillent indépendamment) et leurs propriétés temporelles. Les méthodes conventionnelles de la psychologie cognitive ne permettent pas d'observer aussi finement les processus cognitifs ainsi que leurs représentations. Etant donné que les processus cognitifs sont mis en œuvre par le cerveau, il semble intéressant d'explorer la possibilité d'obtenir des réponses sur leur nature grâce aux mesures de l'activité cérébrale. Les mesures de PEs semblent être de bons candidats: c'est grâce à leur excellente résolution temporelle qu'ils vont permettre de suivre et de définir le déroulement temporel de tels processus.

Les mesures des PEs peuvent être utilisées selon deux approches différentes. Une première approche est d'adopter un modèle pour un processus cognitif spécifique, et d'utiliser les mesures des PEs comme test du modèle en question. Le problème de cette approche est qu'il est difficile de déterminer qu'une mesure d'une composante spécifique reflète un processus cognitif précis. Dans le cas où un processus cognitif et sa manifestation dans les mesures des PEs peuvent être définis précisément, des prédictions peuvent être émises sur les variations de la composante en question (amplitude, latence, polarité ou topographie). La deuxième approche possible est de systématiquement manipuler une variable psychologique et d'observer le résultat sur les PEs. Cette approche ne nécessite pas que l'expérimentateur sache sur quelle composante l'effet sera enregistré. Le succès de cette dernière approche est évident, car un nombre conséquent de processus psychologiques suscitent l'occurrence de composantes PEs distinctes sur le scalp sans pour autant que ces dernières soient définies précisément. Nous allons utiliser cette dernière approche en manipulant la classe de mots tout en contrôlant différentes variables qui peuvent affecter les résultats. Le résultat obtenu à partir des PEs pourra alors nous permettre d'une part, de caractériser les processus de reconnaissance pour les deux classes de mots mais d'autre part, de définir plus précisément les composantes distinctives. Dans la section suivante, nous présentons les avantages et les inconvénients de la méthode des PEs (Coles & Rugg, 1995; Kutas & van Petten, 1994; Rugg & Coles, 1995).



### 2.3.1 Avantages

Un avantage considérable de cette méthode est qu'elle offre la possibilité d'examiner en temps réel les différentes étapes du traitement de l'information. En effet, les PEs reflétant directement l'activité électrique post-synaptique, la résolution temporelle qui est de l'ordre de la milliseconde, constitue un avantage majeur de cette méthode. Les PEs vont donc pouvoir mettre en évidence le déroulement temporel de l'utilisation des différentes informations lors de l'activation de processus cognitifs. A l'opposé, la TEP ou l'IRMf mesurent l'activité cérébrale de second ordre (métabolique ou hémodynamique) par rapport à l'activité électrique neuronale, et sont utilisées afin de visualiser l'activation des différentes aires cérébrales sans pouvoir distinguer l'ordre d'activation de celles-ci. La précision des PEs comme indice de la chronométrie des opérations mentales est apparue comme un progrès notable dans le champ de recherche des neurosciences cognitives. Contrairement à d'autres techniques d'imagerie cérébrale, l'enregistrement des PEs ne nécessite pas l'injection de produit radioactif (TEP), même en faible quantité, ni l'exposition à des champs magnétiques (IRMf). Cette méthode peut donc être considérée comme totalement non invasive et peut, par conséquent, être facilement utilisée sur des patients cérébrolésés ainsi que des enfants, voir même des femmes enceintes <sup>3</sup>. Un autre avantage de cette méthode est de rendre possible l'étude d'un processus cognitif sans l'addition d'une tâche supplémentaire dont la nature pourrait influencer ce processus. En effet, l'enregistrement des PEs permet par exemple de mettre en évidence les processus de compréhension d'un mot, sans qu'il soit demandé au sujet une tâche supplémentaire, comme une décision lexicale, mise à part la lecture (de même une réponse motrice n'est pas nécessaire).

### 2.3.2 Limites

En 1988, Meyer, Osman, Irwin, & Yantis démontrent que l'utilisation des PEs comme outil d'investigation du déroulement temporel des processus mentaux présente certaines faiblesses théoriques et méthodologiques. L'une des techniques couramment utilisées pour analyser les enregistrements PEs est de calculer l'amplitude moyenne (en  $\mu\text{V}$ ) et la latence (en millisecondes) d'une composante, dans le cas où une composante est définie comme une variation de potentiel de durée limitée.

<sup>3</sup> Sauf cas pathologiques, les techniques du TEP et de l'IRMf ne le permettent pas, tout au moins en France.

L'amplitude est mesurée par rapport à un niveau de référence ou ligne de base (qui diffère selon les études: soit une période précédant la présentation du stimulus, soit la mesure moyenne de l'activité durant tout l'enregistrement). La mesure de la latence correspond au temps compris entre le début de la présentation du stimulus et l'instant où la composante atteint son maximum d'amplitude (latence du pic) ou encore l'instant où le tracé diverge de la ligne de base (latence du début de la composante). Meyer, *et al.* (1988) soulignent l'absence d'argument physiologique pour affirmer que la latence du maximum d'amplitude d'une composante est directement liée à la fin d'une étape de traitement, et remarquent également la difficulté d'identifier les générateurs intra-corticaux des composantes.

La résolution spatiale et temporelle des techniques d'imagerie cérébrale dépend non seulement de la puissance de résolution intrinsèque (c'est à dire de la résolution individuelle des détecteurs) de chaque système, mais surtout du type de mesure enregistrée. Puisque les PEs mesurent des fluctuations de l'activité électrique, la résolution de cette technique est équivalente aux variations de champs électriques, soit de l'ordre de la milliseconde. A l'opposé, la TEP (Tomographie à Emission de Positons) enregistrant des variations de débit sanguin dispose d'une résolution plus lente car ces fluctuations sont consécutives à l'activité électrique et ne se révèlent qu'après 1 seconde en moyenne. Ainsi, la technique des PEs dispose depuis les premières expérimentations d'une résolution temporelle inégalable (de l'ordre de la milliseconde). En revanche, par comparaison avec d'autres méthodes d'imageries cérébrales telles que la TEP (Tomographie à Emission de Positons) ou l'IRMf (Imagerie par Résonance Magnétique Fonctionnelle), sa résolution spatiale est mauvaise. Nous pouvons noter que des progrès énormes ont été réalisés depuis quelques années afin de pallier cet inconvénient. Ainsi, les nouveaux développements, tant du point de vue de l'enregistrement (augmentation du nombre de capteurs de quelques électrodes à un panel de 64, voir jusqu'à 256 capteurs sur le cuir chevelu), que des méthodes d'analyses (spatio-temporelles)<sup>4</sup>, permettent aux enregistrements PEs modernes d'appartenir à part entière aux techniques d'imagerie fonctionnelle cérébrale, au même titre que la TEP ou encore de l'IRMf.

---

<sup>4</sup> Une définition plus précise des analyses spatio-temporelles sera abordée dans le chapitre 3.

### 2.3.3 PEs et violations de règles linguistiques

La fécondité de l'utilisation des PEs dans le domaine de recherche des processus de compréhension du langage est largement liée à la découverte par Kutas & Hillyard (1980) d'une composante: la N400. Cette composante s'est avérée être un indice sensible à la manipulation de facteurs sémantiques multiples. Dans leur expérience princeps, Kutas & Hillyard (1980) utilisent des phrases présentées visuellement, et terminées soit par des mots congrus (la pizza est trop chaude pour la *manger*) soit par des mots suscitant une violation d'ordre sémantique (la pizza est trop chaude pour la *crier*). Ces auteurs font l'hypothèse que la différence entre ces deux conditions expérimentales est reflétée par une variation de la latence de la composante P300. Cette hypothèse repose sur des études précédentes, montrant que plus l'évaluation du stimulus est complexe, plus la latence de la P300 est longue (Kutas, McCarthy, & Donchin, 1977). A leur surprise, les résultats ne montrent pas de variation de latence de la composante P300 entre mots congrus et incongrus, mais mettent en évidence une nouvelle composante endogène, plus tardive que la P300 (400 millisecondes), de polarité négative et de distribution maximale sur les électrodes centro-pariétales pour les violations sémantiques. Les caractéristiques de cette composante suggèrent à Kutas & Hillyard (1980) de lui donner le nom de composante N400. Toute une série de travaux confirment que cette différence entre les PEs associés aux mots incongrus et congrus, aussi appelée "effet N400", est extrêmement reproductible et indépendant de la langue testée, ainsi que de la modalité de présentation des phrases, visuelle ou auditive (pour une revue, Kutas & van Petten, 1994).

Depuis ces premières expériences utilisant la technique des PEs, les recherches montrent que la présentation de violations dans un contexte de phrase est un bon moyen pour engendrer des effets robustes sur les PEs. Par exemple, les recherches portant sur les violations d'ordre syntaxique indiquent une certaine indépendance du traitement syntaxique face aux informations sémantiques. En 1992, Osterhout & Holcomb observent que la présentation de violations de la structure d'une phrase, ou encore des propriétés de sous-catégorisation de mots, suscite l'occurrence d'une positivité plus tardive que la N400, ayant un maximum d'amplitude vers 600 millisecondes avec une distribution topographique maximale sur les sites postérieurs. Ils appelèrent P600 cette nouvelle composante. Ce genre de résultats ajoute un argument en faveur des modèles théoriques supposant une indépendance des processus de traitement sémantique et syntaxique (Forster, 1976; Forster, 1979; Friederici, 2002).

Les recherches concernant ces deux composantes se sont développées suivant deux grands axes. Il s'agit d'une part, de déterminer si les effets N400 et P600 sont spécifiquement associés à un traitement linguistique et d'autre part, de préciser quelle(s) étape(s) de traitement es(son)t reflétée(s) par ces composantes pour en proposer une interprétation fonctionnelle. Le premier objectif a été poursuivi en recherchant si des violations spécifiquement syntaxique (accord, temps, structure de la phrase, mouvement lié aux interrogatives, etc.) engendrent les mêmes effets (N400 ou P600). Le second objectif a été poursuivi en identifiant les facteurs psycholinguistiques modifiant les caractéristiques (amplitude et latence) des composantes N400 et P600. N'étant pas le but de cette thèse, nous ne développerons pas plus ces résultats, mais nous tenons à préciser l'existence de revues de littérature sur ce domaine (Kutas, van Petten, & Besson, 1988; Kutas & van Petten, 1994). De plus, une courte revue plus spécifique à la composante N400 sera réalisée dans le chapitre 3.

### 3 Apport des PEs dans l'étude du traitement des deux classes de mots

Les recherches portant sur la distinction entre classes de mots en utilisant la technique des PEs sont assez récentes (moins de 15 ans) et sont très peu homogènes face à leurs résultats. Nous allons maintenant détailler ceux-ci, ainsi que les paradigmes expérimentaux qui, selon nous, peuvent influencer considérablement les corrélats électrophysiologiques.

#### 3.1 Les travaux de Kutas & Hillyard (1983)

Suite à la découverte de la composante N400 (Kutas & Hillyard, 1980), Kutas & Hillyard (1983) étudient les PEs des deux classes de mots à travers un texte comportant différents types d'erreurs. Des erreurs d'accord en nombre portant sur les noms ou les verbes (*as a turtle grows its shell **grow** too; some shells **is** even soft*), des erreurs de temps (*when clouds are cooled to a very low temperature snow may form. Ice begins to **grew** around invisible specks of dust that always float in the air*) et des erreurs sémantiques (*Turtles are smarter than most reptiles but not as smart as mammals such as dogs or **socks***) sont présentées. Les sujets doivent simplement lire ces phrases sans qu'une tâche additionnelle ne leur soit demandée. Les auteurs prennent en compte tous les mots qui ne présentent pas d'erreur et forment deux PEs différents en fonction de la classe de mots étudiée. Ils montrent alors des différences entre les PEs des deux classes de mots dès 200 ms et ce, jusqu'à 700 ms. Les mots de classe

fermée présentent des PEs plus négatifs entre 400 et 700 ms, effet maximal sur les sites antérieurs gauches.

Différentes critiques émergent de cette expérience princeps avec le recul, que nous pouvons émettre à ce jour. Tout d'abord, les auteurs analysent les PEs sur une seule fenêtre temporelle et ceci objectivement en visualisant les tracés. Cette fenêtre est comprise entre 200 et 700 ms, et semble englober différents processus: il est fort probable qu'en divisant la fenêtre comprise entre 200 et 700 ms, l'on puisse mettre en évidence des effets plus précis. Un deuxième problème est que les stimuli ne sont pas contrôlés en longueur ni en fréquence: les mots de classe fermée sont donc plus fréquents et plus courts que ceux employés dans la classe ouverte. Nous avons noté à plusieurs reprises que ces deux propriétés, fréquence et longueur, semblent affecter les processus de traitement des mots. Par conséquent, il s'avère difficile, comme le souligne Kutas & Hillyard (1983), de conclure si les effets observés sur le scalp sont purement des effets de classe de mots, ou plus simplement reflètent des différences dans les propriétés lexicales, telles que la fréquence et/ou la longueur. Kutas & Hillyard (1983) utilisent un nombre conséquent d'items pour chaque classe de mots (2031 items pour la classe fermée et 1562 items pour la classe ouverte). Vu la récurrence des mots de classe fermée dans une langue (*of, the, and, but*) ceux-ci ont été forcément répétés et ce, un nombre de fois plus important que les mots de classe ouverte. Un certain nombre d'études constatent que les effets de répétition peuvent être importants et interagir fortement avec d'autres facteurs tels que la fréquence d'occurrence ou encore l'incongruité sémantique (Besson, Kutas, & van Petten, 1992; Rugg, 1990). En outre, certains mots utilisés dans les moyennages apparaissent dans la phrase après la présentation de l'une des différentes erreurs. De nombreuses études indiquent que suite à des erreurs, les PEs des mots suivants sont perturbés jusqu'au mot final de la phrase (Osterhout, 1997). Malgré le fait qu'aucune tâche additionnelle ne soit demandée aux participants, le fait de présenter des erreurs dans des phrases a certainement dû influencer leur stratégie. Enfin, ces auteurs enregistrent les PEs à partir de 8 électrodes, et il s'avère difficile à la lueur des études ultérieures de pouvoir conclure sur les différences de topographie entre mots de classe ouverte et fermée, en raison du manque de précision dans cette étude princeps. Face à l'émergence d'une composante spécifique pour les mots de classe fermée et malgré de nombreux problèmes méthodologiques que nous venons d'énumérer, Kutas & Hillyard (1983) avancent l'hypothèse que les mots de classe ouverte et fermée sont traités différemment dans le cerveau.

### 3.2 Variabilité des résultats chez les sujets sains

Par la suite, les études ont essayé de pallier les différents problèmes énumérés et ce, d'une manière plus ou moins adéquate. Nous pouvons observer dans le tableau 2 les différentes expériences à partir d'adultes sains qui portent sur l'étude de la distinction entre mots de classe fermée et ouverte jusqu'à nos jours. Ces études et leurs résultats sont difficiles à présenter et à commenter car, ces derniers sont parfois contradictoires. Par exemple, certaines études montrent des PEs similaires (négativités antérieures précoces avec un maximum d'activité vers 280 ms, la N280) pour les deux classes de mots (Brown, Hagoort, & ter Keurs, 1999; King & Kutas, 1998; Münte, Wieringa, Weyerts, Szentkuti, Matzke, & Johannes, 2001) alors que d'autres ne mettent en évidence cette négativité (N280) que pour les mots de classe fermée (Keurs ter, Brown, Hagoort, & Stegeman, 1999; Neville, *et al.*, 1992; Pulvermüller, Lutzenberger, & Birbaumer, 1995). Ces résultats contradictoires peuvent être dus en partie dans ces expériences à une large variabilité dans le choix des stimuli présentés ainsi que dans la tâche demandée aux participants. Les stimuli peuvent être soit des morceaux de texte (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; Kutas & Hillyard, 1983; Osterhout, Bersick, & McKinnon, 1997), soit des phrases sans liens (King & Kutas, 1998; Kutas, *et al.*, 1988; Münte, *et al.*, 2001; Neville, Coffey, Holcomb, & Tallal, 1993; Neville, *et al.*, 1992; van Petten & Kutas, 1991), soit des listes de mots non reliés (Garnsey, 1985; Münte, *et al.*, 2001; Nobre & McCarthy, 1994; Osterhout, *et al.*, 1997; Pulvermüller, *et al.*, 1995). Les tâches incluent la lecture simple (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; Kutas & Hillyard, 1983; Kutas, *et al.*, 1988; Münte, *et al.*, 2001), un jugement d'acceptabilité (Neville, *et al.*, 1993; Neville, *et al.*, 1992), ou encore une tâche de décision lexicale (Garnsey, 1985; Münte, *et al.*, 2001; Pulvermüller, *et al.*, 1995).

Deuxièmement, l'interprétation des résultats portant sur la distinction entre classe de mots est sujette à caution car il semble difficile de déterminer quels aspects dans le vocabulaire sont réellement responsables des effets PEs (longueur, fréquence d'occurrence, répétition, contraintes contextuelles, degré d'abstraction, fonction grammaticale, etc.). En particulier, il semble crucial de ne pas confondre les différences quantitatives des propriétés lexicales des stimuli (telles que la fréquence et la longueur) avec les différences qualitatives des propriétés linguistiques (telle que la classe de mot).

Etudes	Langue	corpus		procédure				méthode	résultats PRECOCE (250-350 msec)		résultats TARDIF (350-700 msec)		Corrélation	
		Long	Freq	stimuli	Rep	Durée	SOA		Tâche	El	N280 pour CF	N280 pour CO		Neg pour CF
1	Anglais	oui	oui	mot	non	20	4500	Dec Lex	6	?	?	N350-600, partout sauf broca	NON	
2	Anglais	oui	non	mot	non	500	2100-2200	Det Cat	50	N288 Ant G	N316 Cent	N400 réduite	N364 cent-occ Bilat	
3	Allemand	oui	oui	mot	non	100	2500-4500	Dec Lex	17	N160 & N260 G	N160 & N260 Bilat		N340-500 partout	
4a	Anglais	non	non	mot	oui	300	700	Quest comp	13	non	non	N400-700 post	N400 post	corrélés
5a	Allemand	non	non	mot	non	300	2100	Dec Lex	29	N280	N280	pas de N400-700, N400 réduit	N400 large	non corrélés
6	Hollandais	non	non	mot	oui	400	800	Lect	27	N260 Ant G	N260 Bilat	N400-700 Ant	N450 D	non corrélés
7	Anglais	non	non	phrase	oui	150-200-132	500-900-700	Lect	8	?	?	N400 réduit	N400 post D	
8	Anglais	non	non	phrase	oui	200	600	Rec Cible	13	?	?	N400 réduit	N400 post D	
9	Anglais	non	non	phrase	oui	200	700	Accep sem	12	N280 Ant G	non	N400-700 Ant G	N400 post D	
10	Anglais	non	non	phrase	oui	300	ISI 700	Accep sem	14	?	?	N400-700 Ant G	N400 post D	
11	Anglais	non	non	phrase	oui	200	500	vrai/faux	26	N280 Ant G	N315 Ant G	N400-700 Ant G	N400 post D	corrélés
5b	Allemand	non	non	phrase	non	400	700	Lect	29	N280	N280	N400 réduit, N400-700 pour HF-CF	N400 large	corrélés
12	Anglais	non	non	prose	oui	200	640-760	Lect avec erreur	8	?	?	N400-700 Bilat	P400-700 G	
4b	Anglais	non	non	prose	oui	300	700	Quest comp	13	N350 temp	non	N400-700 post	N400 post	corrélés
13	Hollandais	non	non	prose	oui	400	800	Lect	29	N263 Ant G	N237-N317 Bilat	N400-700 Ant G	P400-700 Ant	non corrélés
14	Hollandais	non	non	prose	oui	400	800	Lect	29	N289 bilat Ant	N312 Bilat Ant	N350-500 Ant G		non corrélés
15	Anglais	non	non	prose	oui	300	700	Lect - accep sem/synt	13	N280 bilat Ant	N319 ou 329 Bilat			corrélés

**Tableau 2:** Tableau récapitulatif des études portant sur la distinction entre classes de mots avec des PEs. Référence pour les différentes expériences: 1, Garnsey (1985); 2, Nobre & McCarthy (1994); 3, Pulvermüller, Lutzenberger, & Birbaumer (1995); 4 (a et b), Osterhout, Bersick, & McKinnon (1997); 5 (a et b), Münte, et al. (2001); 6, Keurs ter, Brown, & Hagoort (2002); 7, Kutas, van Petten, & Besson (1988); 8, van Petten & Kutas (1991); 9, Neville, Mills, & Lawson (1992); 10, Neville, Coffey, Holcomb, & Tallal (1993); 11, King & Kutas (1998); 12, Kutas & Hillyard (1983); 13, Keurs ter, Brown, Hagoort, & Stegeman (1999); 14, Brown, Hagoort, & ter Keurs (1999); 15, Osterhout, Allen, & McLaughlin (2002). Abréviations. Long: longueur, Freq: fréquence d'occurrence des mots, Rep: répétition, El: nombre d'électrodes, Dec lex: décision lexicale, Det Cat: Détection de catégorie, Quest Comp: Question de compréhension, Rec cible: reconnaissance de cible, Lect: Lecture, Accep sem: jugement d'acceptabilité sémantique, accep sem/synt: jugement d'acceptabilité sémantique et syntaxique, vrai/faux: décision de véracité, Ant: antérieure, G: gauche, D: droite, Bilat: bilatérale, Cent: central, Occ: occipital, Post: postérieure, Prop. lex: Propriétés lexicales des stimuli (longueur et fréquence), HF-CF: mot de classe fermée de haute fréquence. Précoce/tardif voir détails dans la section 3.2.1 de ce chapitre.

C'est donc à partir de trois grands thèmes que nous allons approfondir les résultats de la littérature. La première dimension retenue est **méthodologique**. En effet, alors que la première étude de Kutas & Hillyard (1983) utilisent une large fenêtre d'analyse (de 200 à 700 ms), les études suivantes réduisent cette fenêtre afin d'affiner les résultats. De plus, l'augmentation du nombre d'électrodes permet de mieux cerner la répartition topographique de la composante N280 et d'une possible distinction entre les deux classes de mots. La deuxième dimension portera sur le **corpus** et la nature des stimuli employés: nous verrons que certaines études mettent en évidence une possible corrélation entre la latence de la composante N280 et la fréquence des mots utilisés (alors que d'autres ne répliquent pas ce résultat). D'autres études contrôlent la fréquence des deux classes de mots afin d'obtenir un effet de classe "pure" sans contamination par les propriétés lexicales des mots. La dernière dimension sera le **paradigme** expérimental (type de tâche). Ce dernier thème nous semble crucial lorsque l'on étudie la distinction entre classes de mots. En effet, il a été émis comme hypothèse que la distinction entre classes de mots n'est effective que dans un contexte de

phrase, et qu'en isolation, et par conséquent lors de l'accès lexical, les étapes de traitement sont identiques (Osterhout & Bersick, 1999).

### 3.2.1 Aspect méthodologique: la fenêtre précoce et le nombre d'électrodes

Les résultats de Kutas & Hillyard, (1983) sont affinés, grâce surtout à un apport méthodologique: la réduction de la fenêtre temporelle dans les analyses. Le traitement des mots de classe ouverte et fermée peut grossièrement se différencier sur deux fenêtres temporelles: une précoce (entre 200 et 350 ms) et une plus tardive (entre 350 et 700 ms).

Dans la première fenêtre temporelle (entre 200 et 350 ms) Neville, *et al.*(1992) montrent une composante spécifique aux mots de classe fermée. Dans cette étude, les sujets devaient lire des phrases qui étaient soit correctes, soit terminées par un mot incongru sémantiquement, et décider si celles-ci étaient correctes ou non. Ces auteurs rapportent pour les mots de classe fermée une composante négative largement distribuée sur les sites antérieurs gauches. Le maximum d'amplitude pour cette composante est enregistré à 280 ms, d'où son nom: la N280. Dans la fenêtre temporelle tardive (350-700 ms), Neville, *et al.*(1992) montrent que les mots de classe fermée suscitent l'occurrence d'une composante négative largement distribuée sur les sites antérieurs gauches, la LAN (pour *Left Anterior Negativity*) ou encore la N400-700, alors que les mots de classe ouverte suscitent l'occurrence d'une composante N400 localisée sur les sites postérieurs. Neville, *et al.*(1992) examinent l'effet de la fréquence et de la longueur en formant des PEs séparés pour des mots de haute et de basse fréquence, ainsi que des mots courts et longs. Alors qu'ils n'ont pu réaliser la même étude pour les mots de classe fermée, ils montrent que l'effet de fréquence pour les mots de classe ouverte est d'altérer l'amplitude de la composante N400: plus la fréquence du mot est basse et plus l'amplitude de la composante N400 est grande. En revanche, ils notent que l'effet de longueur est reflété par l'augmentation d'une positivité frontale symétrique dès 200 ms jusqu'à la fin de la fenêtre critique: plus un mot est long et plus la positivité est importante. Les mots de classe ouverte courts et fréquents ne suscitent donc pas l'occurrence d'une composante N280 ou encore d'une N400-700. Puisque les mots de classe ouverte de haute fréquence et courts ne présentent pas le pattern des mots de classe fermée (pas de N280 sur les régions antérieures gauches), les auteurs concluent que la N280 est un "marqueur " spécifique de la classe fermée et que cette composante ne reflète aucunement les variations des propriétés lexicales entre les stimuli. De plus, cette composante N280 est distincte de la composante N400 car elle ne



présente pas les mêmes latence et topographie sur le scalp. Ces auteurs interprètent donc fonctionnellement la composante N280 comme étant liée aux processus de traitement des mots de classe fermée. La précocité de la N280, ainsi que les études comportementales, suggèrent que ces mots doivent être traités de manière automatique et rapide. De plus, la distribution de la N280 sur les aires antérieures gauches est compatible avec la large littérature clinique indiquant que des perturbations de la compréhension ainsi que de la production de structures grammaticales sont consécutives à des lésions spécifiques dans ces régions cérébrales. Alors que la composante N400 est enregistrée pour les mots de classe ouverte seulement, ces auteurs rejoignent l'interprétation que cette composante est un marqueur des processus d'intégration sémantique lors de la compréhension du langage (Kutas & Hillyard, 1980). Ainsi, comme deux composantes sont mises en évidence lors du traitement des deux classes de mots, Neville, *et al.* (1992) affirment que deux systèmes neuronaux différents sont responsables du traitement de ces mots.

D'après ces résultats, il semblerait que dans un contexte de phrase les deux classes de mots suscitent l'occurrence de PEs précoces distincts. Il s'avère que si l'on regarde les expériences étudiant cette distinction dans un contexte de phrase, alors, Neville, *et al.* (1992) sont les seuls à obtenir ces résultats (voir tableau 2). En effet, alors que certaines études (Kutas & Hillyard, 1983; Kutas, *et al.*, 1988; Neville, *et al.*, 1993; van Petten & Kutas, 1991) ne nous renseignent pas sur la possibilité d'une distinction dans le rang de latence de la composante N280 (les analyses sont réalisées entre 200, 350 ou 400 et 700 ms), d'autres études (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001) montrent que les deux classes de mots suscitent l'occurrence d'une composante N280 présentant une topographie similaire ou presque sur le scalp. Ces dernières études observent les PEs à partir de 26 voir, 29 électrodes, alors que Neville, *et al.* (1992) n'en n'ont enregistré que 12. L'augmentation du nombre d'électrodes dans certaines études (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001) et par conséquent, l'affinement de l'analyse topographique de la composante N280, ajoute un argument en défaveur d'une distinction précoce dans les PEs entre les deux classes de mots. Ainsi ces résultats invalident-ils partiellement l'hypothèse selon laquelle les mots de classe fermée et ouverte seraient identifiés précocement de manière différente dans le cerveau.

Nous avons vu que, selon Neville, *et al.* (1992), la fréquence ou la longueur n'affectent pas les PEs dans le rang de latence de la composante N280. Des résultats contradictoires sont suggérés par King & Kutas (1998) et Osterhout, *et al.* (1997): la latence de la composante N280 varierait avec les propriétés lexicales des stimuli telles que la fréquence d'utilisation du mot ou sa longueur. Nous allons donc maintenant passer en revue les études portant sur l'étude des propriétés lexicales des stimuli.

### 3.2.2 La nature des stimuli: rôle de la fréquence et longueur

#### 3.2.2.1 La fréquence et la longueur affectent-elles la latence de la N280?

Quelques recherches se sont intéressées à la modulation des caractéristiques de la composante N280 par les propriétés lexicales des stimuli. Très peu d'expériences ont réussi à contrôler la fréquence d'utilisation ainsi que la longueur des mots de classe fermée et ouverte. Ainsi, certaines expériences ont essayé de vérifier si les modulations de fréquence et longueur des mots pouvaient affecter la latence de la composante N280.

Nobre & McCarthy (1994) présentent des listes de mots dans lesquelles les sujets devaient détecter des mots de catégories spécifiques (végétal). Les stimuli comprenaient des mots de classe fermée (prépositions, conjonctions, articles), des mots de classe ouverte (noms concrets), des pseudo-mots (ex: *glubbalt*), des non-mots (ex: *klghdr*) qui étaient appareillés en longueur mais pas en fréquence (CF: 5.8 nombre de lettres, CO: 5.3 nombre de lettres)<sup>5</sup>. Ils mettent alors en évidence une composante N280 latéralisée sur les sites antérieurs gauches pour les mots de classe fermée (latence de 288 ms), alors que les mots de classe ouverte montrent une augmentation de négativité qui est plus symétrique et plus tardive (la N330, latence de 316 ms). De plus, ils trouvent que la composante N400 est significativement réduite pour les mots de classe fermée par rapport aux mots de classe ouverte.

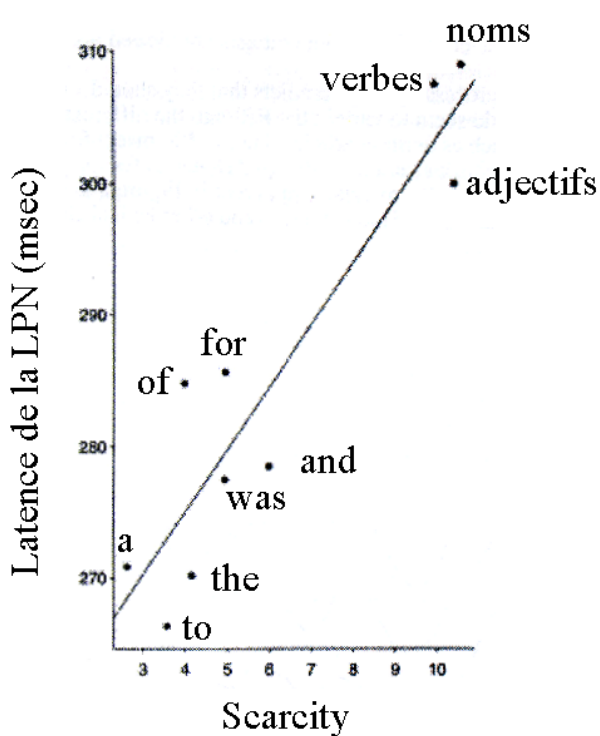
---

<sup>5</sup> La fréquence des stimuli n'est pas indiquée dans cette étude, les auteurs indiquent que les mots de classe fermée sont plus fréquents que les mots de classe ouverte.

Alors que les deux classes de mots suscitent l'occurrence de composantes négatives différemment latéralisées entre 200-350 ms (la N280 pour les mots de classe fermée et la N330 pour ceux de classe ouverte) Nobre & McCarthy (1994) ne concluent pas que ces composantes précoces reflètent des processus différents, et invalident donc les résultats de Neville, *et al.* (1992). En revanche, Nobre & McCarthy (1994) interprètent la diminution d'amplitude de la composante N400 comme un marqueur distinctif entre les deux classes de mots. De plus, comme la longueur des stimuli est contrôlé, ils infèrent que leur résultats ne reflètent tout au moins pas des effets de longueur.

Osterhout, *et al.* (1997) ajoutent une pièce supplémentaire au puzzle qu'est l'étude de la distinction entre classes de mots et leurs propriétés lexicales. Dans une expérience présentant un texte normal, ces auteurs rapportent que les mots de classe ouverte suscitent l'occurrence d'une négativité à 400 ms (N400) et d'une positivité ultérieure (LPC pour *Late Positive Component*). Les mots de classe fermée en revanche présentent deux pics négatifs (N350 et N450) ainsi qu'une autre négativité plus tardive (N400-700). Les mots de classe fermée ne produisent pas de N280, et les négativités N350 et N450 enregistrées ont des distributions similaires sur le scalp à celle de la composante N400 suscitée par les mots de classe ouverte (négativités postérieures). De plus, ces auteurs sous-divisent chaque classe en différentes catégories grammaticales: les mots de classe fermée en articles, prépositions, pronoms et auxiliaires, et les mots de classe ouverte en noms et verbes. Ils montrent que chaque catégorie grammaticale provoque l'occurrence d'une négativité dont la latence fluctue entre 280 et 400 ms: N280 pour les articles, N320 pour les prépositions, N350 pour les pronoms, N360 pour les auxiliaires, N400 pour les noms et verbes. Ainsi, ils concluent que les deux composantes négatives des mots de classe fermée (N350 et N450) ainsi que la N400 enregistrée pour les mots de classe ouverte, sont une seule et même composante. Sa latence fluctue avec les propriétés lexicales des mots telles que la fréquence et la longueur. Afin de vérifier leur hypothèse, ces auteurs effectuent des analyses de régression afin d'observer si les indices longueur et fréquence des mots peuvent être utilisés pour prédire la latence de cette composante négative. Leurs résultats montrent que la latence de ces négativités est hautement corrélée aux propriétés lexicales des mots: leur fréquence ( $R = .46, p < .0001$ ) et leur longueur ( $R = .46, p < .0001$ ). Ils concluent que les variations de latence de ces négativités correspondent plutôt à des différences quantitatives des propriétés lexicales des mots, plutôt qu'à des différences qualitatives de classe de mots.

Un an plus tard, King & Kutas (1998) corroborent les résultats d'Osterhout, *et al.* (1997). Ils présentent des phrases sans lien entre elles et demandent aux sujets de réaliser une tâche de lecture. Après la moitié des phrases, une question de compréhension suscitant une réponse oui/non était posée aux participants. Ils enregistrent alors des composantes négatives précoces pour les mots de classe fermée (N280) ainsi que les mots de classe ouverte (N315). Face à la ressemblance de leur topographie (antérieure gauche) ils estiment que ces négativités N280 et N315 représentent la manifestation d'un même et unique processus et que, ces deux négativités sont en quelque sorte des modulations d'une même composante. Ils prénomment cette composante LPN pour *Lexical Processing Negativity*. Ces auteurs émettent alors l'hypothèse que cette LPN indexe des processus d'accès lexical qui sont affectés par les propriétés lexicales des mots (Osterhout, *et al.*, 1997). De même qu'Osterhout, *et al.* (1997), ces auteurs réalisent des analyses de régression entre la latence de la LPN et la classe de mots, en utilisant l'indice *scarcity* comme facteur indicatif de la fréquence des mots. Ils définissent l'indice lexical, la *scarcity* comme correspondant au logarithme de la fréquence de chaque mot soustrait au logarithme maximum pouvant être obtenu. Le logarithme maximum correspondant à la taille du corpus (soit dans ce corpus, le logarithme maximum est de 6 sur 1 million d'occurrences, voir figure 6).



**Figure 6:** Ligne de régression obtenue sur la grande moyenne (21 sujets) pour la latence de la composante LPN (*lexical processing negativity*). Les points sont obtenus à partir de groupes équivalents en logarithme de fréquence d'items (environ 190 projections d'items, excepté le point correspondant à "the" où 377 essais par sujets sont utilisés). Les dix points correspondent à chaque groupe de mots présentant un intervalle de *scarcity* équivalent entre chaque. Les points sont représentés avec un exemple de mot représentatif de chaque fréquence lexicale. Les résultats de cette analyse de régression montrent que la *scarcity* est un bon prédicteur de la latence de la composante LPN. Notons que les mots de basse *scarcity* sont tous des mots de classe fermée, alors que les mots présentant une *scarcity* élevée appartiennent à la classe ouverte. D'après King & Kutas (1998).

Par conséquent, la *scarcity* est une variable qui est corrélée de manière positive (comme la longueur) avec le temps de réaction dans de nombreuses tâches. Comme les variables longueur et *scarcity* sont deux variables fortement corrélées (à 93 %), King & Kutas (1998) n'utilisent qu'un seul indice pour les deux soit la *scarcity*. Les résultats révèlent que 92 % de la variance de la latence de la LPN est expliquée par l'indice *scarcity*. De plus, ces auteurs montrent qu'une augmentation de 1 logarithme de fréquence correspond à un délai de 5 ms dans la LPN. Lorsque l'on regarde les courbes de régression plus précisément (figure 6), nous pouvons noter que dans les catégories de *scarcity* élevée (supérieure à 8 logarithme) se trouvent des verbes, noms, et adjectifs, appartenant tous à la classe ouverte. En revanche, pour les mots présentant une *scarcity* basse (inférieure à 6 logarithme), on ne trouve quasiment que des mots de la classe fermée (*of, for, was, and, the, to*). Ainsi, le problème de leur analyse de régression est que l'effet de *scarcity* est inséparable de celui de la classe de mot: les mots de classe fermée sont courts et fréquents, et inversement pour les mots de classe ouverte. Leur analyse de régression révèle qu'à l'intérieur de chaque classe de mot, la latence de la LPN est aussi largement affectée par les propriétés de fréquence des mots. Par exemple, les articles montrent une latence d'environ 270 ms alors que les prépositions (qui sont moins fréquentes et moins courtes que les articles) présentent une latence de 300 ms. Cependant dans ce travail, les différences de latence de la LPN entre chaque type de mots au sein des classes ouverte et fermée ne sont pas plus amplement détaillées et analysées statistiquement. King & Kutas (1998) aboutissent à l'idée que la négativité précoce qu'ils nomment LPN est enregistrée par la présentation des mots de classe ouverte comme fermée. Sa latence variant en fonction des propriétés lexicales des stimuli telles que la fréquence, ils concluent que cette composante ne reflète pas une différence de classe de mots à l'inverse des conclusions de Neville *et al.* (1992). Pourtant l'examen de la figure 6 suggère qu'il existe un clivage entre les deux classes de mots formant 2 groupes séparés. De plus, ils notent que la *scarcity* est aussi corrélée avec des temps de fixation de l'œil sur les mots (Thibadeau, Just, & Carpenter, 1983). Ainsi, ces auteurs émettent l'hypothèse que la LPN reflète le contrôle des processus inhibiteurs des saccades oculaires. En effet, pour réaliser un PE, il est nécessaire de présenter des phrases mot par mot au centre de l'écran, cette exigence empêchant le sujet de réaliser des saccades de mot à mot, mais aussi de placer l'œil au niveau d'un point de fixation préférentiel bien précis dans un mot.

Des résultats contradictoires sont obtenus un an plus tard par Brown, *et al.* (1999) ainsi que Keurs ter, *et al.* (1999). Pour Brown, *et al.* (1999), les PEs sont enregistrés à partir de mots

présentés dans une histoire. Deux types de différences sont mis en évidence. Dans une première fenêtre précoce (entre 230 et 350 ms), une composante négative est enregistrée pour les deux classes de mots avec un maximum d'amplitude sur les sites antérieurs bilatéraux. La latence de cette composante n'est pas affectée par la fréquence ou la longueur des mots, mais elle est plus précoce pour les mots de classe fermée (289 ms) que pour ceux de classe ouverte (312 ms). Dans une fenêtre plus tardive (350-500 ms), seuls les mots de classe fermée suscitent une composante négative localisée sur les sites frontaux gauches. Ainsi, ces auteurs se rallient aux conclusions de King & Kutas (1998) et d'Osterhout, *et al.* (1997) pour avancer l'idée que la négativité précoce qui est enregistrée pour les deux classes de mots ne peut pas être définie comme un marqueur distinctif de la classe. En revanche, ne notant pas d'effet des propriétés lexicales telles que la longueur et fréquence sur la latence de la composante précoce, ils rejettent les arguments de King & Kutas (1998) et d'Osterhout, *et al.* (1997) sur la possibilité que la latence de cette composante reflète les variations des propriétés lexicales des mots. De plus, en sous-divisant chaque classe de mots en différentes catégories grammaticales ils montrent d'une part, que seules les catégories appartenant à la classe fermée présentent des variations dans la latence de la composante négative précoce. Alors que les articles et les conjonctions engendrent des latences identiques (277 et 278 ms respectivement) les prépositions produisent une composante négative plus tardive (298 ms). D'autre part, ces auteurs notent que pour les trois catégories de mots de classe fermée les latences sont plus précoces de 16 ms en moyenne sur la partie gauche du scalp que celle située à droite. Les résultats de Brown, *et al.* (1999) contredisent par conséquent ceux de King & Kutas (1998) où ces derniers auteurs trouvent que la latence de la LPN est fortement corrélée à la *scarcity* des mots non seulement de classe fermée mais aussi de classe ouverte.

Récemment, les données de Brown, *et al.* (1999) ont été étudiées avec une autre analyse: les oscillations par bande de fréquence dans l'EEG. Cette analyse calcule la puissance spectrale (en  $\mu\text{V}$ ) pour des bandes de fréquences spécifiques. Le principe de ce type d'analyse est issu des différences théoriques et fonctionnelles entre réponses EEG évoquées et induites. Les réponses EEG induites coïncident avec la présence d'un stimulus (*time locked*) mais ne sont pas en phase avec celui-ci (*phase locked*). A l'opposé, les réponses EEG évoquées (PEs) sont à la fois *time locked* et *phase locked* et de ce fait émergent lors du moyennage. Les réponses induites pouvant apparaître à des moments différents par rapport au stimulus critique seront donc invisibles pour la technique des moyennages utilisés avec les PEs. Cette technique, les oscillations par bande de fréquence, a permis de mettre en évidence des augmentations de

puissance dans les rangs de fréquence thêta (4-7 Hz) durant des tâches mnésiques (Klimesch, 1999), ou encore lors de la reconnaissance de mots (Bastiaansen, van Berkum, & Hagoort, 2002). Bastiaansen & Hagoort (2000) indiquent que pour des mots suscitant des violations d'ordre "accord grammatical", une augmentation spécifique dans le rang de fréquence thêta est aussi observée, avec un effet maximal entre 300 et 500 ms. Ces résultats de Bastiaansen & Hagoort (2000) et de Bastiaansen, *et al.* (2002) semblent spécifiques au rang de fréquence étudiée (absence d'effet dans le rang de fréquence correspondant au rythme alpha, 8-12 Hz). Sur les données de Brown, *et al.* (1999), Bastiaansen & Hagoort (2001) indiquent une augmentation dans le rang de fréquence thêta (4-7 Hz) entre 300 et 500 ms après le début de chaque mot. De plus, ils montrent une différence dans la puissance de fréquence thêta: les mots de classe ouverte présentent une puissance de spectre pour les fréquences thêta plus importante sur l'hémisphère gauche que les mots de classe fermée (pas d'effet sur l'hémisphère droit). Le pic de latence d'apparition de ces augmentations dans le rang de fréquence thêta est aussi différent pour les deux classes de mots: 265 et 345 ms pour les mots de classes fermée et ouverte, respectivement. De même que lors des analyses sur les PEs, ces auteurs réalisent des régressions sur la latence d'augmentation dans la fréquence thêta, et n'indiquent aucun effet significatif. Ainsi, l'augmentation de puissance dans le rang de fréquence thêta semble être indépendant des propriétés lexicales du mot, et directement reliée à la classe de mot, confirmant les résultats obtenus à partir des PEs (Brown, *et al.*, 1999). Actuellement, l'interprétation fonctionnelle de cet effet reste à trouver, mais cette analyse corrobore les résultats de Brown, *et al.* (1999).

En 2001, Münte, *et al.* réalisent une étude dans laquelle les stimuli sont constitués de mots appartenant aux deux classes de mots, avec une variation de la fréquence et de la longueur dans chaque classe. Ainsi, des mots de fréquence moyenne pour les deux classes de mot (fermée: moyenne de 6.6 lettres et 2.83 log de fréquence; ouverte: moyenne de 6.85 lettres et 2.73 log de fréquence), ainsi que des mots de très haute fréquence pour la classe fermée sont présentés (moyenne de 3.17 lettres et 4.34 log de fréquence). Tous les mots critiques sont placés en cinquième position dans des phrases déclaratives. La tâche des participants est de lire ces phrases afin de répondre à des questions après la présentation de celles-ci. Les résultats montrent qu'entre 250 et 350 ms les mots de classe fermée de très haute fréquence se différencient des deux autres conditions expérimentales par leur topographie (amplitude plus importante en région antérieure gauche), mais aussi par la latence de la composante N280 (CF très haute fréquence: 335 ms, CF moyenne fréquence: 344 ms, CO moyenne fréquence: 353

ms). Dans la fenêtre temporelle tardive, la composante N400-700 est d'amplitude plus importante pour les mots de très haute fréquence de classe fermée que pour ceux des deux autres catégories. La mesure de la latence de la composante N280 présente une sensibilité à la fréquence (pour les mots de très haute fréquence), résultat qui confirme ceux de King & Kutas (1998) et d'Osterhout, *et al.* (1997) et vont à l'encontre de ceux de Brown, *et al.* (1999), Keuster, *et al.* (1999) et de Neville, *et al.* (1992).

Dernièrement, Osterhout, *et al.* (2002) présentent une étude dans laquelle ils répliquent leurs résultats de 1997, mais avancent aussi des résultats en faveur d'une modulation des PEs en fonction de la longueur des mots. Dans une expérience où des phrases avec des erreurs de sens ou de syntaxe sont présentées, les analyses portant sur la fréquence ne sont que succinctement expliquées. Ils indiquent que la latence de la composante négative précoce varie entre 298 et 329 ms pour les différentes catégories grammaticales (articles: 298, pronoms: 306, auxiliaires: 310, prépositions: 313, verbes: 319, noms: 329 ms). De plus, ils signalent que la latence fluctue entre les différentes catégories grammaticales et correspond "approximativement" à l'ordre des fréquences moyennes et des longueurs pour chaque catégorie. Malheureusement, les résultats ne sont présentés qu'avec une analyse de Variance ( $F(5,90) = 8.33, p < .001$ ). Alors que dans leur étude d'Osterhout, *et al.* (1997), les analyses de régression étaient présentées, ici, nous ne pouvons déterminer s'il existe un réel lien entre la fréquence et la latence de la composante négative précoce. En effet, il suffit qu'une seule catégorie grammaticale suscite une latence différente des autres pour que l'analyse présentée soit significative. De plus, ces auteurs réalisent des regroupements pour différentes longueurs (2, 3, 4, 5, 6, 7, 8 lettres). Ils observent que la longueur affecte les PEs sur toute la fenêtre temporelle: les PEs des mots courts sont plus négatifs que les PEs associés aux mots longs, entre 150 et 500 ms, sur les sites antérieurs. Entre 500 et 700 le maximum de différence est enregistré à la fois sur les sites antérieurs mais aussi postérieurs, résultats similaires à ceux de Neville, *et al.* (1992). De plus, ces auteurs indiquent que l'amplitude et la latence de deux composantes (à savoir P2, entre 150 et 300 ms, et N300, entre 280 et 320 ms) sont corrélées à plus de 80 % avec la longueur. Bien que la longueur et la fréquence soient corrélées à plus de 99%, ces auteurs concluent que la longueur est la propriété lexicale qui influence le plus les caractéristiques de la composante N280 (par comparaison avec la fréquence d'occurrence).

Les études examinant les effets de la fréquence et de la longueur sur la latence de la composante N280 semblent mitigées quant à leurs conclusions. Alors qu'il existe un



consensus pour conclure que la composante N280 ne reflète pas un marqueur distinctif entre les deux classes de mots (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997), certains (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) indiquent que la fréquence et la longueur influencent significativement la latence, contrairement aux conclusions d'autres (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; Neville, *et al.*, 1992).

### 3.2.2.2 Et si la fréquence et la longueur étaient contrôlées?

Afin de répondre à la question si les caractéristiques (topographie et latence) de la composante N280 sont modulées par la classe de mots ou bien par les propriétés lexicales des stimuli, il est nécessaire de contrôler la fréquence et la longueur des stimuli utilisés. Très peu d'études ont réussi à construire un corpus équilibré en fréquence et longueur avec un nombre suffisant d'items pour chaque classe de mots et qui permette de réaliser une expérience PEs dans de bonnes conditions (rapport signal sur bruit satisfaisant).

La première étude dans ce sens est celle de Garnsey (1985) présentant des groupes de mots de classe fermée et ouverte dont les fréquences et longueurs sont en moyennes équivalentes (2.33 log de fréquence et 5.72 nombre de lettres). Malgré le peu d'électrodes utilisées (6), une analyse en composantes principales montre un pattern général. Il semblerait que, lorsque le sujet réalise une tâche de décision lexicale sur des mots présentés en isolation, les PEs des mots de classe fermée soient plus négatifs que ceux associés aux mots de classe ouverte sur 5 des 6 sites (CPz, Oz, Broca droit, Wernicke gauche et droit) pour une large fenêtre temporelle allant de 375 à 600 ms. Malheureusement, ces auteurs ne différencient pas effet précoce et effet tardif dans leurs analyses.

En allemand, Pulvermüller, Lutzenberger, & Birbaumer (1995) appariaient aussi leurs stimuli en fréquence et en longueur (CO: 3.62 log de fréquence, CF: 3.66 log de fréquence)<sup>6</sup> et ont demandé aux sujets de réaliser une tâche de décision lexicale. Ils montrent alors que les mots de classe fermée suscitent une composante négative précoce enregistrée dès 160 ms avec un maximum d'activité sur les sites antérieurs gauches, alors que les mots de classe ouverte présentent la même composante mais bilatérale. De plus, dans le rang de latence 150-300 ms, Pulvermüller, *et al.* (1995) indiquent que les mots de classe fermée présentent une composante négative qui est plus latéralisée à gauche que celle des mots de classe ouverte. Ces auteurs ne montrent pas de différence de latence dans les composantes. Ainsi, d'après eux, seule la latéralisation (topographie) de la composante négative semble être caractéristique de l'effet de classe de mots; en revanche, la différence de latence elle disparaît lorsque la fréquence et la longueur des mots sont appariées.

Bien que seulement deux études aient contrôlé la fréquence et la longueur, les résultats semblent indiquer que seules des différences de topographie sont le reflet d'une différence dans le traitement des deux classes de mots. De plus, le résultat d'une latence plus précoce pour les mots de classe fermée n'est retrouvé que pour l'étude de Nobre & McCarthy (1994) alors que la fréquence d'occurrence des mots n'est pas contrôlée. Ainsi, on ne peut conclure sur la possibilité que la latence de la composante N280 reflète des différences de propriétés lexicales des mots plutôt qu'une distinction des propriétés linguistiques. Enfin, il semble intéressant de noter que ces deux études - ainsi que l'étude de Nobre & McCarthy (1994) où seule la fréquence est contrôlée - utilisent des mots en isolation avec une tâche quasiment similaire (détection lexicale ou catégorielle). Comme nous l'avons indiqué précédemment, la tâche, mais aussi le contexte, peuvent induire des processus spécifiques pour chaque classe de mots.

### 3.2.3 Isolation vs. contexte?

Trois séries d'études comparent plus précisément les résultats obtenus en isolation et dans un contexte de phrase (Osterhout, *et al.*, 1997), (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999) et (Münte, *et al.*, 2001).

<sup>6</sup> Ces auteurs indiquent que les deux classes de mots présentent des valeurs équivalentes en nombre de lettres, sans pour autant indiquer les valeurs respectives.

Osterhout, *et al.* (1997) comparent deux expériences. Dans l'une, l'ordre des mots dans le texte a été changé de sorte que seul l'aspect sémantique du texte soit conservé: la structure syntaxique est donc inexistante. Dans l'autre, le contexte est préservé. Pour les deux expériences, les mots de classe fermée suscitent l'occurrence de négativités quasi identiques en topographie avec des latences de N350 et N450 ms. En revanche, pour les mots de classe ouverte, la négativité enregistrée varie en fonction du contexte utilisé: lorsque le contexte est un texte normal, la négativité est identique à celle de la composante N400 (postérieure). Par contre, lorsque le contexte est constitué d'une liste de mots sans suite, la négativité présente une latéralisation sur l'hémisphère gauche. Pour les deux types de contextes, chaque catégorie grammaticale (pronoms, articles, auxiliaires, verbes, noms, etc.) présente un seul pic négatif, dont plus de 75 % de la variance est expliquée par les propriétés lexicales des mots (longueur et fréquence). Comme la plupart de ces résultats sont communs aux deux expériences, les auteurs en concluent que ces négativités N350 et N450 reflètent des processus mis en jeu lors de l'identification du mot, mais ne sont pas liés aux processus d'analyse syntaxique mis en place lors de la compréhension de phrase. La comparaison des deux expériences conduit au résultat suivant: la composante N400 est réduite dans le cas où le texte est normal. Ce résultat est cohérent avec l'interprétation selon laquelle l'amplitude de la composante N400 reflète la facilité d'intégration d'un mot avec un contexte (Kutas & Hillyard, 1980): plus le mot est difficilement intégrable et plus cette composante est ample. Osterhout, *et al.* (1997) montrent aussi que la composante N400-700 des mots de classe fermée est réduite dans le cas d'un texte mélangé par rapport à un texte normal (Garnsey, 1985; van Petten & Kutas, 1991). Ainsi, ces observations suggèrent que la composante N400-700 reflète la manière dont les deux classes de mots sont utilisées dans une phrase (ou un texte), plutôt que des différences dans les représentations ou l'accès au lexique mental. En résumé, Osterhout, *et al.* (1997) indiquent que des différences quantitatives dans la fréquence et la longueur des mots sont fortement corrélées avec les variations de la latence de la négativité précoce. Ces corrélations sont retrouvées que le contexte soit un texte ou une suite de mots.

Münte, *et al.* (2001) comparent des PEs obtenus dans des listes de mots et à partir de contexte de phrase. Pour les deux expériences, ces auteurs ne montrent pas de différence significative entre les deux classes de mots dans la fenêtre 160-200 ms - effet de Pulvermüller, *et al.* (1995) - comme entre 280-370 ms (effet N280). De plus, la composante N280 est largement plus visible dans l'expérience où les items sont présentés dans des phrases. Ils observent que la N280 est présente à la fois pour les mots de classe ouverte et fermée. En isolation, les résultats

montrent que la fréquence, ainsi que la classe, ne jouent aucun rôle dans la détermination de la latence de la composante N280 (pour les mots de classe ouverte: basse fréquence, 356 ms; moyenne fréquence, 352 ms; haute fréquence, 340 ms - pour les mots de classe fermée: basse fréquence, 342 ms; moyenne fréquence, 337 ms; haute fréquence, 337 ms; très haute fréquence, 339 ms). A l'opposé, lors de la présentation d'un contexte, la latence pour les mots de très haute fréquence de classe fermée est plus courte que celle des mots de moyenne fréquence (de classe ouverte et fermée). Ces auteurs montrent aussi des différences de topographies: les PEs des mots de classe fermée de très haute fréquence sont plus négatifs sur le scalp entre 400 et 700 ms sur les sites antérieurs et pariétaux par rapport à ceux associés aux autres mots (classe ouverte et classe fermée de basse fréquence). Alors que cette composante est largement distribuée dans l'expérience où les mots sont présentés en isolation, lorsque des phrases sont utilisées celle-ci est plus restreinte sur le scalp sur les régions antérieures gauches. Si l'on regarde les caractéristiques des stimuli, l'on peut noter que la longueur des mots de classe fermée de très haute fréquence (3,17 nombre de lettres) est inférieure aux deux autres catégories (classe fermée moyenne fréquence: 6.66 et classe ouverte moyenne fréquence: 6.85 nombre de lettres). Alors que Neville, *et al.* (1992) et Osterhout, *et al.* (2002) suggèrent que la longueur est un effet qui est enregistré dès 200 ms<sup>7</sup>, il semble possible que l'effet de Münte, *et al.* (2001) soit lié à une différence de longueur dans leur stimuli.

Dans deux séries d'expériences, ter Keurs et collaborateurs étudient aussi l'effet du contexte (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999). Ainsi, ils comparent des PEs obtenus dans un texte (Keurs ter, *et al.*, 1999) et dans une liste de mots (Keurs ter, *et al.*, 2002). Pour les deux expériences, les résultats montrent qu'entre 210 et 325 ms les PEs des mots de classe fermée sont plus négatifs que ceux des mots de classe ouverte (effet N280 classique). Dans le cas où les mots sont présentés en isolation, cette différence de topographie entre les deux classes de mots se restreint aux sites localisés sur les régions antérieures gauches. Pour la fenêtre d'analyse plus tardive, les deux expériences montrent que les mots de classe fermée suscitent l'occurrence d'une composante N400-700 négative, maximale sur les sites antérieurs gauches.

---

<sup>7</sup> Plus le mot est court et plus les PEs sont négatifs sur les sites antérieurs gauches.

Ainsi l'interprétation selon laquelle la composante N400-700 refléterait des processus d'attente liés au contexte est à exclure (Osterhout, *et al.*, 1997). En effet, cette longue négativité enregistrée à partir des mots de classe fermée a été généralement assimilée à la famille de la contingente négative (CNV; Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999). En quelque sorte, cette composante refléterait le fait que le lecteur soit engagé dans des stratégies de prédiction des mots à venir. Cette interprétation concorde avec le rôle syntaxique accordé aux mots de classe fermée dans une phrase. Malheureusement, les résultats de Keurs ter, *et al.* (2002) et de Münte, *et al.* (2001) invalident cette interprétation car ces deux études montrent une composante N400-700 pour les mots de classe fermée dans une liste de mots sans suite et où, par conséquent, le rôle des mots de classe fermée est restreint, voir aboli.

Les études de Osterhout, *et al.* (1997), de Keurs ter, *et al.* (2002), Keurs ter, *et al.* (1999) et de Münte, *et al.* (2001) montrent qu'en isolation (ou en présence d'une liste de mots sans lien) les différences de topographies entre les deux classes de mots sont moins importantes et se restreignent à des distributions sur les sites antérieurs gauches par rapport à lorsque les mots sont présentés dans un contexte de phrase ou de texte. De plus, alors qu'Osterhout, *et al.* (1997) indiquent une corrélation entre la fréquence du mot et la latence de la N280 avec ou sans contexte, Münte, *et al.* (2001) observent celle-ci uniquement dans un contexte de phrase. Pour finir, nous avons aussi vu que les résultats sur l'obtention d'une composante N400-700 dans un contexte de phrase sont aussi controversés par les études de Keurs ter, *et al.* (2002) et de Münte, *et al.* (2001), rendant difficile l'interprétation fonctionnelle de cette composante.

### 3.3 Déficits et développement des patterns PEs normaux

#### 3.3.1 Description des PEs de patients aphasiques de Broca

Keurs ter, *et al.* (1999) conduisent une étude afin de déterminer si les patients atteints d'une aphasie de Broca (lésion fronto-temporal gauche) présentant un déficit grammatical lors de la compréhension de phrases (voir chapitre 1 section 4), présentent également une altération des PEs associés aux mots de classe fermée. Ainsi, ils étudient trois groupes de sujets: des patients présentant des lésions dans l'hémisphère gauche au niveau de l'aire de Broca (fronto-temporal gauche, BR), des patients présentant des lésions dans la région homologue à l'aire de Broca dans l'hémisphère droit (RH), ainsi qu'un groupe de sujets adultes contrôle ne présentant

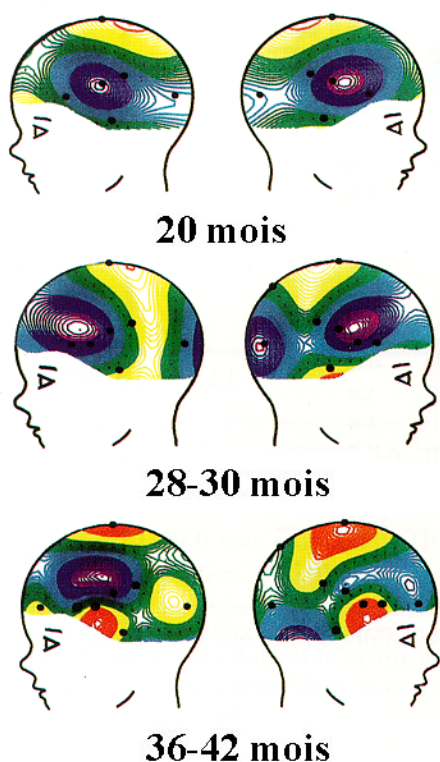
aucune pathologie. Les PEs sont enregistrés à partir d'un texte présenté visuellement que les sujets devaient lire et comprendre. Ces auteurs montrent que les adultes sains présentent des PEs différents pour les deux classes de mots et dans deux rangs de latence. Les mots de classe fermée induisent une négativité sur les sites antérieurs gauches avec un pic à 263 ms (la N280) ainsi qu'entre 400 et 700 ms (la N400-700). Les mots de classe ouverte présentent deux pics négatifs N237 et N317 (entre 210 et 375 ms) avec un maximum d'activité sur les électrodes frontales bilatérales. Aucune différence dans la latence de la composante négative précoce n'est observée entre les deux classes de mots. Les auteurs caractérisent l'effet N280 comme étant un marqueur précoce des processus lexicaux (Neville, *et al.*, 1992): cette composante serait reliée à un processus très précoce comme l'accès au lexique. Sur la base de leurs résultats Keurs ter, *et al.* (1999) interprètent cette N280 comme étant la première manifestation de l'effet de classe lorsque les informations de catégorie lexicale sont activées dans le lexique mental. Alors que pour les patients RH les différences entre les deux classes de mots sont identiques aux effets enregistrés chez les adultes sains pour les deux rangs de latences, précoce et tardif, les patients BR, en revanche, ne présentent pas l'effet de classe dans le rang de latence précoce. En effet, les PEs pour les deux classes de mots sont similaires dans cette fenêtre. Pour le rang tardif, les patients BR présentent des effets de topographies significatifs: une composante N400-700 (négativité antérieure) pour les mots de classe fermée, et une composante N400 (négativité postérieure) pour les mots de classe ouverte. Puisque Keurs ter, *et al.* (1999) suggèrent que la composante précoce reflète l'activation des informations de catégories lexicales, l'absence de différence pour les patients BR dans la fenêtre temporelle précoce suggère que ce type d'information n'est pas encore, ou n'est que partiellement, disponible à ce moment précis (Keurs ter, *et al.*, 1999): le déficit des ces patients se manifesterait par conséquent au niveau lexical du traitement des classes de mots.

Sans considération précise sur l'étape de traitement que représente la composante N280 lors de l'accès au lexique, le déficit des patients BR porte aussi sur les processus de compréhension de phrases, reflété par des perturbations dans les PEs tardifs. En effet, ces patients montrent l'occurrence d'une composante N400 pour les mots de classe ouverte, alors qu'elle est absente chez les adultes sains et les patients RH. Le délai, ou l'absence, d'activation des informations de classes a donc induit chez ces patients des problèmes d'intégration sémantique reflétés par la présence d'une composante N400. L'accès aux informations de catégorie lexicale, ainsi que leurs disponibilités dans le temps, sont nécessaires afin d'aboutir à une compréhension de la phrase cohérente. Puisqu'une même

composante est enregistrée pour les deux classes de mots, les auteurs n'ont pu conclure si le traitement anormal des informations de catégories lexicales chez les patients BR provenait du déficit de traitement spécifique de l'une ou l'autre de classe de mots. En revanche, la présence d'une composante N400 pour les mots de classe ouverte indique que les patients BR ont accès aux informations de nature lexico-sémantique. Cependant, l'amplitude conséquente de cette composante N400 pour les mots de classe ouverte suggère aussi que ces mots sont difficilement intégrés dans le contexte linguistique. En conclusion, on ne peut affirmer si les patients BR présentent un déficit spécifique dans le traitement de l'une ou l'autre des classes de mots. En revanche, un retard évident dans la disponibilité des informations de catégorie lexicale induit probablement leurs problèmes de compréhension.

### 3.3.2 Avant 20 mois, la composante N280 n'est pas latéralisée

L'étude développementale de la distinction entre classe de mots sous-tend l'hypothèse suivante: si deux systèmes neuronaux différents participent à l'élaboration de traitement différents, alors comme les études sur le système visuel le suggèrent, ces systèmes doivent présenter une maturation temporelle et des périodes critiques de développement différentes. Neville & Mills (1997) observent l'évolution de l'acquisition des deux classes de mots chez les enfants âgés entre 20 et 42 mois. Ces auteurs se sont assurés que tous les enfants comprenaient et produisaient les mots qu'ils ont utilisé dans cette expérience. A 20 mois, les PEs des mots de classe ouverte et fermée ne sont pas différents, mais présentent une topographie différente de celle suscitée par les mots que les enfants ne connaissaient pas (voir figure 7). Ces résultats suggèrent que dans les premiers stades du développement du langage, lorsque les enfants s'expriment typiquement avec des mots isolés (ou qu'ils commencent à combiner deux mots), les mots de classe ouverte et fermée sont traités de la même manière dans le cerveau. Entre 28 et 30 mois, lorsque les enfants commencent à parler avec de courtes phrases, les PEs des mots de classe ouverte et fermée présentent des différences topographiques sur le scalp. Cependant, la latéralisation gauche de la composante négative N280 n'est pas encore apparente pour les mots de classe fermée. Entre 36 et 42 mois (3 ans) la plupart des enfants s'expriment par des phrases, et utilisent les mots de classe fermée pour spécifier des relations grammaticales dans la phrase. Alors, comme chez les adultes, une composante négative maximale sur les sites antérieurs gauches, la N280, est enregistrée après la présentation de mots de classe fermée.



**Figure 7:** Analyse topographique des sources de densité de courant (CSD) de l'activité cérébrale pour les mots de classe fermée à la latence de 200 ms. La colonne de gauche représente les densités de courant distribué sur l'hémisphère gauche. Les puits (courant rentrant à l'intérieur de la tête, couleur bleue) et les sources (courant sortant de la tête, couleur rouge) des CSD sont présentés pour 3 tranches d'âges différentes. En haut, à 20 mois les CSD dévoilent des puits de courants (bleus) dans l'hémisphère gauche et droit. Pour des enfants âgés entre 28 et 30 mois (figures du milieu), les puits (bleus) de CSD sont bilatéraux, avec une légère tendance plus prononcée sur l'hémisphère droit. En bas, les CSD pour les enfants âgés de 36 et 42 mois montrent un puits (bleu) présent sur l'hémisphère gauche en région antérieure uniquement. Ce dernier pattern correspond à celui observé chez l'adulte. D'après Neville & Mills (1997).

Les résultats de ces 3 groupes d'enfants sont en accord avec l'hypothèse selon laquelle, initialement, les mots de classe ouverte et fermée sont traités de la même manière dans le cerveau, et ce n'est que progressivement, et avec l'augmentation de l'expérience et la maturation du système langagier, que le traitement va se spécialiser.

De plus, les auteurs montrent que pour les enfants de 20 mois qui disposent d'un vocabulaire de taille réduite (toute classe confondue), les PEs des mots de classe ouverte et fermée ne présentent pas de différence significative. En revanche, dans ce même groupe d'âge, les enfants qui possèdent un vocabulaire de taille plus conséquente présentent des patterns électrophysiologiques similaires à ceux du groupe d'âge supérieur (28-30 mois). Ces données suggèrent que l'activité électrique du cerveau est liée aux capacités langagières plutôt qu'à l'âge chronologique de l'enfant.



### 3.3.3 Importance de l'âge d'acquisition du langage

Neville, *et al.* (1992) étudient les PEs pour les mots de classe ouverte et fermée chez des patients sourds de naissance (surdité congénitale). Ces patients ont appris la langue des signes à l'âge normal d'acquisition du langage alors qu'ils n'ont appris la langue anglaise que bien plus tard (après 8 ans) et ne l'utilisent que dans la modalité visuelle. Les phrases présentées comporte 50 % d'erreurs sémantiques en position finale des phrases. Ces patients sont testés dans la langue anglaise. Ces auteurs montrent que les patients présentent des PEs identiques à ceux des adultes sains pour les mots de classe ouverte (noms) ainsi que pour les phrases dans lesquelles des violations sémantiques sont présentes (présence d'une composante N400). En revanche, les patterns électrophysiologiques sont aberrants pour les mots de classe fermée. Plus spécifiquement, ces patients sourds de naissance ne présentent pas les composantes négatives localisées sur les sites antérieurs gauches pour les deux rangs de latences: la N280 et la N400-700. Ainsi, les patients ayant appris l'anglais tardivement et les adultes de langue maternelle anglaise ne présentent pas de différences électrophysiologiques qualitative ou quantitative pour le traitement des mots de classe ouverte ou encore celui des erreurs sémantiques. Ces résultats montrent que les processus lexico-sémantiques ne sont pas affectés par un retard dans l'expérience/apprentissage et par conséquent sont robustes. A l'opposé de ces résultats, les mots de classe fermée qui véhiculent les informations lexico-syntaxiques suscitent l'occurrence de PEs qui sont qualitativement différents chez les patients sourds par rapport à ceux observés chez les adultes entendants. Le système mis en place lors du traitement des informations grammaticales semble donc plus vulnérable face à l'altération et/ou au retard dans l'acquisition du langage. Le développement des processus grammaticaux ainsi que la maturation des aires corticales qui lui sont associées serait donc plus sensible à un retard dans l'acquisition que ne l'est le développement des processus lexico-sémantiques.

## 4 Conclusions

Les résultats portant sur l'acquisition du langage chez des enfants sains (Neville & Mills, 1997) ainsi que chez des enfants sourds (Neville, *et al.*, 1992) montrent que les deux classes de mots sont différemment sensibles d'une part, à l'exposition de la langue et d'autre part, à l'âge d'acquisition. La différence de maturation des structures corticales permettant la

reconnaissance des deux classes de mots sous-tendrait une différence computationnelle au sein de l'organisation architecturale. De plus, les résultats obtenus sur les patients aphasiques de Broca suggèrent des conclusions similaires. Chez ces patients, une réduction de l'utilisation des mots de classe fermée est accompagnée de patterns électrophysiologiques déviants comparés à ceux des adultes contrôles.

Si la distinction entre les deux classes de mots est effective lors de l'accès au lexique, alors l'on s'attend à enregistrer les différences après les PEs qui sont représentatifs du traitement visuel précoce (encodage du stimulus), tels que le complexe N1-P2. De plus, ces effets doivent précéder les PEs connus pour refléter des processus post-lexicaux, tels que la composante N400 (Chwilla, Brown, & Hagoort, 1995; Chwilla, Hagoort, & Brown, 1998; Holcomb, 1993; Kutas & van Petten, 1994). Etant donné que la composante N280 présente ces caractéristiques (enregistrée après le complexe N1-P2 et avant la composante N400), elle a pu représenter un candidat potentiel permettant de distinguer l'accès au lexique des deux classes de mots. Suite à la revue de la littérature sur les adultes sains, il n'existe aucun consensus sur une possibilité de distinction entre les mots de classe ouverte et fermée grâce à cette composante négative précoce. En effet, un certain nombre d'études montrent que la composante précoce négative est aussi bien enregistrée pour les mots de classe fermée qu'ouverte (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001). De plus, de nombreux résultats s'accordent pour indiquer que la latence de cette composante varie avec les propriétés lexicales (fréquence et longueur) des mots (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997).

Si des corrélats électrophysiologiques distincts ont été mis en évidence, aucune conclusion claire n'est possible quant à déterminer si ces effets proviennent de différences qualitatives des propriétés lexicales entre les deux classes de mots ou bien de différences réelles dans leur rôle respectif au sein d'une phrase. Alors que seulement deux études Garnsey (1985) et Pulvermüller, *et al.* (1995) contrôlent la fréquence et la longueur des items utilisés, celles-ci montrent des différences de topographies précoces entre les deux classes de mots mais pas d'effet de latence. Les résultats de ces deux expériences suggèrent donc un traitement précoce différent entre les deux classes de mots, tout au moins lorsque les mots sont lus en isolation. Ces résultats ne nous permettent pas d'être catégorique quant à l'étape de traitement reflétée par la composante N280. En effet, afin de désambiguïser les effets obtenus par Garnsey (1985) et Pulvermüller, *et al.* (1995), il aurait fallu présenter les mots au sein de phrase. Si

une même différence de PEs est enregistrée entre les deux classes de mots à la fois dans un contexte de phrase et en isolation, alors l'effet ne pourra pas être lié à des étapes d'intégration au sein du contexte, mais aura lieu directement lors de l'accès au lexique (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). En effet, la similarité des effets N280 (différence entre mots de classe ouverte et fermée) obtenus en isolation et dans un contexte de phrase (ou de texte) indique que les différences entre les deux classes de mots sont précoces et correspondent à des étapes prenant part lors de l'accès au lexique. Malheureusement dans ces dernières expériences, les propriétés lexicales ne sont pas contrôlées.

Si certaines études contrôlent les propriétés lexicales, alors l'influence du contexte sur la composante N280 est négligée. A l'opposé, si l'effet du contexte est étudié, l'influence des propriétés lexicales n'est observée qu'à partir de régressions et non contrôlée. Le cœur de notre travail portera donc sur la distinction entre classe de mots lorsque les propriétés lexicales ne rentrent plus en compte, puisqu'elles seront contrôlées. Grâce à la présentation des stimuli dans un contexte de phrase et en isolation, nous pourrons déduire si les différences de traitement entre les deux classes de mots ont lieu à des étapes incluses lors de l'accès au lexique - effet mis en évidence dans les deux types de présentation - ou bien à des étapes plus tardives comme l'intégration des informations lexicales au sein du contexte - effet enregistré seulement lorsque les mots sont présentés en contexte de phrase.

## 5 Problématique

Cette thèse évalue le statut de la distinction entre les mots de classe ouverte et fermée au sein des modèles psycholinguistiques actuels. Comme nous l'avons vu dans le chapitre 1, la psychologie cognitive a permis l'élaboration des modèles de fonctionnement cognitif, tout en précisant la nature des représentations, ainsi que les processeurs mis en jeu. La reconnaissance d'un mot est une fonction cognitive automatique, irrépressible et inconsciente (voir introduction). Questionner la nature d'un tel processus nous a semblé plus évident avec une technique d'imagerie cérébrale fonctionnelle, la technique des potentiels évoqués ou PEs. Deux principaux atouts ont retenu notre choix pour cette technique: d'une part sa **résolution temporelle** et d'autre part, un **rapprochement des conditions expérimentales** avec les situations réelles.

La résolution temporelle de la technique des PEs est de l'ordre de la milliseconde, et par conséquent avoisine la rapidité des processus de compréhension du langage. La technique des PEs permet de recueillir des indices plus détaillés par rapport à un temps de réaction. En effet cette mesure permet de déterminer d'une part, quelles informations influencent l'activité cérébrale, et donc le processus cognitif en question mais d'autre part, de suivre le déroulement temporel de cette influence. Les potentiels évoqués permettent donc d'étudier la dynamique d'un processus cognitif. L'enregistrement des PEs est complémentaire aux autres méthodes d'imageries cérébrales. En effet, la précision spatiale ne peut être obtenue qu'avec la TEP ou encore l'IRMf. Dans cette thèse nous avons donc décidé d'axer nos recherches sur l'étude de la dynamique des fonctions cognitives au détriment de la possibilité de localiser celles-ci. Les mesures des PEs peuvent être recueillies sans qu'une tâche additionnelle ne soit demandée aux participants. Ce dernier point nous semble crucial afin d'étudier les représentations des deux classes de mots. En effet, prenons l'exemple d'une tâche de décision lexicale dans laquelle un participant doit détecter les mots parmi des non-mots. Ce type d'analyse ne survient que très rarement dans la vie courante. Les participants ont donc certainement déployé une stratégie propre qui a pu interférer avec le processus en question. Nous nous efforcerons dans la plupart de nos expériences d'éliminer ce genre d'artefact et d'utiliser des protocoles se rapprochant le plus de la lecture.

Les quatre principaux objectifs de notre travail, explicités dans cette section, sont développés tour à tour dans les chapitres 3, 4, 5, 6 et 7. Nous tenons à préciser que l'explication dans cette section de nos différents objectifs ne suit pas le déroulement chronologique des chapitres de la thèse; mais plutôt notre intérêt théorique portant sur la modélisation de la reconnaissance des deux classes de mots. C'est ainsi que le dernier objectif, méthodologique, prend part dans le premier chapitre expérimental, le chapitre 3.

## 5.1 Caractériser les représentations/processus pour les deux classes de mots

La revue de littérature nous indique que bons nombres d'études montrent des corrélats électrophysiologiques différents pour les deux classes de mots (Neville, *et al.*, 1993; Neville, *et al.*, 1992; Nobre & McCarthy, 1994; Pulvermüller, *et al.*, 1995). Ces différences peuvent être des variations de topographie dans une fenêtre temporelle précoce comprise en général entre 250 et 350 ms, et/ou plus tardive entre 400 et 700 ms, mais aussi de latence sur la composante précoce N280 enregistrée entre 250 et 350 ms. Ces résultats soutiennent l'hypothèse selon laquelle il existe des différences qualitatives dans l'activation précoce des représentations des deux classes de mots. Ces conclusions sont toutefois à considérer avec précaution. En effet, les études contrastant les PEs des deux classes de mots ont produit une quantité importante de données contradictoires (pour un résumé voir chapitre 2, section 3 et tableau 2). Toutes ne sont pas en accord avec celles rapportées par Neville *et al.* (1992). Le problème de la majorité de ces études est qu'un grand nombre de facteurs confondus avec la classe de mots n'ont que rarement été manipulés, ou ne fut-ce que contrôlés dans les analyses. Déterminer précisément quelles propriétés (longueur, fréquence, classe, répétition, contexte, etc.) suscitent des différences de traitement, et donc affectent les PEs des deux classes de mots, est une question des plus importantes en psycholinguistique.

Depuis quelques années, un certain nombre d'études indiquent que l'effet précoce que reflète la N280, (variation de topographie et de latence), auparavant interprété par des différences de traitement entre les deux classes de mots, pourrait être en réalité lié à des variations de propriétés lexicales entre les deux classes de mots. Ainsi, différentes études (King & Kutas, 1998; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) montrent que les variations de latence de la composante N280 sont directement liées à la longueur et la fréquence des mots: plus un mot est court et fréquent, plus la latence de la composante négative est précoce. A l'opposé d'autres études (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999) observent une latence plus courte pour les mots de classe fermée par rapport à celle des mots de classe ouverte sans pour autant que les propriétés lexicales (longueur et fréquence) n'affectent celle-ci. Alors que seulement deux études manipulent des items contrôlés en fréquence et en longueur (Garnsey, 1985; Pulvermüller, *et al.*, 1995), aucune ne rapportent de différences de latence de la composante N280. Seules des différences

de topographie sont à noter, suggérant que des représentations différentes, c'est à dire, localisées à différents endroits du cerveau, sont activées pour ces deux classes de mots.

Notre travail apporte une contribution expérimentale visant à expliquer les contradictions dans les résultats précédents. Ainsi, une question centrale est de déterminer si les différences électrophysiologiques observées sur le scalp résultent de différences dans les propriétés linguistiques des deux classes de mots ou bien plutôt de différences dans leurs propriétés physique (longueur) et/ou distributionnelle (fréquence). Cette question est au cœur d'un débat actuel concernant la statut biologique des deux classes de mots: la classe ouverte et la classe fermée.

Dans une première expérience (chapitre 4), les propriétés lexicales des stimuli utilisés ne sont pas contrôlés: les mots de classe fermée sont donc plus fréquents et plus courts que ceux de classe ouverte. Cette première étude vise tout d'abord à obtenir une base électrophysiologique pour le traitement de ces deux classes de mots en langue française. En effet, il semble intéressant de noter qu'aucune expérience sur l'étude de la distinction entre les deux classes de mots (voir tableau 2) n'a actuellement enregistré les PEs sur des stimuli en français. Si la distinction entre ces deux classes de mots est liée à des différences fondamentales quant à leur rôle linguistique, alors nous devrions trouver cet effet dans l'ensemble des langues, et par conséquent en français. D'autre part, des analyses portant sur les différentes catégories grammaticales (noms, verbes, adverbes, prépositions, quantificateurs, non-quantificateurs) sont présentées afin de définir précisément si, à l'intérieur de chaque classe de mots, la configuration électrique sur le scalp pour la fenêtre d'analyse de la composante N280 est uniforme ou non. Enfin, des analyses de régression (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; King & Kutas, 1998) sur les latences des composantes précoces négatives sont réalisées afin déterminer un éventuel effet des propriétés lexicales sur la précocité de ces composantes négatives.

L'étude de l'effet de classe de mots nécessite de spécifier les autres propriétés qui peuvent influencer et interférer avec les résultats. Nous avons vu que deux études (Garnsey, 1985; Pulvermüller, *et al.*, 1995) ont contrôlé la fréquence et la longueur des mots. Comme les théories linguistiques s'accordent pour attribuer aux mots de classe fermée un rôle prépondérant dans la construction de la structure syntaxique de la phrase, l'on s'attend à enregistrer des différences entre les PEs des deux classes de mots qui soit maximal dans un

contexte de phrase, et moindre lorsque les mots sont présentés en isolation. Malheureusement, aucune de ces deux expériences (Garnsey, 1985; Pulvermüller, *et al.*, 1995) ne présentent les mots dans un contexte linguistique. L'expérience du chapitre 5 utilise des stimuli contrôlés pour les deux classes de mots dans un contexte de phrase. Dans le chapitre 5, nous avons créé des listes de mots pour les deux classes présentant des fréquences et longueurs similaires. Chacun des mots cibles est présenté en position fixe dans la phrase (4<sup>ème</sup> position) afin de limiter un éventuel effet du contexte, aucun mot cible n'est répété, et une simple tâche de compréhension est demandée aux participants. En comparant les résultats des expériences du chapitre 4 et 5, nous pourrions mettre en évidence les corrélats électrophysiologiques directement liés à des différences fonctionnelles entre les deux classes de mots de ceux liés aux différences entre les propriétés lexicales (physique et/ou distributionnelle).

## 5.2 Etudier l'effet de classe de mots en fonction de la bande de fréquence d'occurrence

Contrôler la fréquence d'occurrence des mots semble être primordial afin de spécifier plus précisément les effets sur le scalp. Malheureusement, la littérature issue de la psychologie cognitive indique un autre problème à surmonter. En effet, Gordon & Caramazza (1985) ont montré que les effets de fréquence (c'est à dire qu'un mot de haute fréquence est reconnu plus rapidement qu'un mot de basse fréquence) liés à la classe de mots varient en fonction du rang de fréquence étudié. Si ce rang est inférieur à  $2,6 \log$  de fréquence (soit 400 occurrences/million), alors l'effet de fréquence est observé pour les deux classes de mots, au dessous de cette limite, seuls les mots de classe ouverte répondent positivement à la fréquence d'occurrence. Les résultats de Bradley (1978), indiquant que les mots de classe fermée sont insensibles à la fréquence d'occurrence, peuvent donc s'expliquer par le fait que ces auteurs n'ont pas apparié la fréquence d'occurrence entre les deux sets de stimuli. Les items de la classe fermée ont été choisis parmi un "pool" typique de mots de classe fermée, soit, de très haute fréquence. Ainsi, les résultats de Gordon & Caramazza (1985) suggèrent que la reconnaissance des mots montre un effet plafond: au dessus de 400 occurrences/million, soit  $2,6 \log$  de fréquence, les mots (toute classes confondues) ne présentent plus l'effet de fréquence.

Le deuxième but de cette thèse est d'étudier la distinction entre classe de mots pour les deux rangs de fréquence prédéfinis par Gordon & Caramazza (1985). C'est pourquoi dans l'étude du chapitre 5, nos stimuli pour chaque classe de mots sont sous-divisés en groupes de mots dont la fréquence est supérieure ou inférieure à  $2,6 \log$  de fréquence (Gordon & Caramazza, 1985). La comparaison entre l'effet de classe pour des mots de haute fréquence ( $> 2,6 \log$  de fréquence) et pour ceux de basse fréquence ( $< 2,6 \log$  de fréquence) apporte des indices pertinents quant à l'influence de la fréquence d'occurrence sur la reconnaissance des deux classes de mots.

### 5.3 Etudier l'effet de contexte sur les deux classes de mots

Lors de la compréhension d'un texte écrit, les lecteurs ont théoriquement plusieurs sources d'informations distinctes à leur disposition qu'ils peuvent utiliser pour transformer le signal visuel en une interprétation sémantique. Celles-ci incluent les informations pragmatiques, les propriétés sémantiques et syntaxiques des mots, les règles syntaxiques de la langue, les règles orthographiques ou formelles. Les recherches en psycholinguistique tentent de décrire comment ces différentes sources d'informations sont utilisées, et nous avons vu dans le chapitre 1 (section 2) que deux classes de modèles s'opposent quant aux sources d'informations utilisées lors de l'accès au lexique. Les modèles autonomes (Forster, 1976; Forster & Ryder, 1971) suggèrent que l'accès au lexique s'effectue sans le soutien d'informations contextuelles, strictement sur la base des informations issues du décodage de l'entrée sensorielle. A l'opposé, les partisans des modèles interactifs (Morton, 1969; Morton, 1979) postulent que les informations de sources multiples (c'est à dire contextuelle sémantique et syntaxique) peuvent interagir avec celles issues de l'analyse perceptive du signal d'entrée. Alors que dans le chapitre 5, les mots sont présentés dans le contexte de phrases, dans le chapitre 6 nous allons travailler avec des mots isolés. Le principal objectif de cette dernière étude est de définir l'accès au lexique des deux classes de mots lorsque les propriétés lexicales distributionnelle (fréquence d'occurrence) et physique (longueur) sont contrôlées. Ainsi, si les PEs associés aux mots de classe fermée sont différents de ceux obtenus par les mots de classe ouverte, une réponse positive sur la distinction de stockage entre ces deux classes de mots pourra être suggérée. De plus, si des effets de latence en faveur des mots de classe fermée sont mis en évidence, alors nous pourrions conclure que non seulement l'accès au lexique des deux classes de mots est de nature différente, mais aussi que celui-ci est plus précoce pour les mots de classe fermée.



Un troisième objectif est d'étudier l'influence du contexte lors de l'accès au lexique pour les deux classes de mots. L'étude comparée (chapitre 7) des expériences du chapitre 5 et 6 va permettre de mettre en évidence sur quelle composante le contexte linguistique interagit. Si l'hypothèse selon laquelle les mots de classe fermée et ouverte joue un rôle différent au sein d'une phrase est vérifiée, alors les différences entre les PEs de ces deux classes de mots doivent être plus importantes lorsque ceux-ci sont présentés dans un contexte de phrase. De plus, nous pourrions définir si le contexte influence différemment les mots de classe ouverte et fermée. En revanche, si seul l'accès au lexique est l'étape de traitement différenciant les deux classes de mots, alors les différences entre les PEs de ces deux classes de mots doivent être présentes avec et sans contexte linguistique.

#### 5.4 Apport méthodologique: les analyses spatio-temporelles

Un dernier objectif développé dans cette thèse est de nature méthodologique. Les analyses statistiques classiques des PEs s'appuient sur une analyse visuelle des composantes du grand moyennage. C'est à partir de cette analyse que des choix sont effectués, comme par exemple la taille et le rang de latence de la fenêtre d'analyse, ou encore le choix de la localisation des électrodes sur lesquelles vont porter les analyses statistiques ultérieures. Avec l'augmentation considérable du nombre de capteurs enregistrés en simultané, il s'avère de plus en plus difficile de justifier ce type de choix *a posteriori*. Quel est l'intérêt d'enregistrer un nombre conséquent d'électrodes si celles-ci ne sont pas utilisées dans les analyses?

Dans cette thèse nous essayerons de montrer l'utilité d'une autre technique d'analyse statistique, dite spatio-temporelle de segmentation (Lehmann, Ozaki, & Pal, 1987; Lehmann & Skrandeis, 1980). Cette technique d'analyse repose sur l'idée que l'activité électrique du cerveau varie de façon très rapide, et qu'entre ces périodes de fluctuations des périodes de stabilité électrique peuvent être mises en évidence. Puisque deux configurations électriques sur le scalp ne peuvent être obtenues qu'à partir de l'activation de générateurs ou groupes de générateurs différents, alors deux périodes avec des PEs de surface différents représentent des états fonctionnels cognitifs différents. Dans le chapitre 3, nous présentons plus précisément cette méthode dite spatio-temporelle à partir d'un paradigme connu: la présentation de violations sémantiques. Nous avons choisi de travailler sur ce type de violations afin de pouvoir rapporter nos résultats à des effets déjà amplement étudiés au

moyen de la technique des PEs. Par ailleurs, chacune des expériences présentées dans les chapitres suivants est analysée en parallèle au moyen des deux techniques (classique et spatio-temporelle). La présentation des deux analyses permet de mettre en évidence leurs avantages et inconvénients réciproques.

# CHAPITRE 3

## METHODES D'ANALYSE DES PEs: APPLICATION A LA N400

---

### *Résumé*

Le but de cette étude est de comparer deux analyses différentes du décours temporel de la composante N400: les analyses de tracés, et les analyses de segmentation adaptée. Vingt participants adultes, droitiers, de langue maternelle française ont jugé l'acceptabilité sémantique de phrases présentées visuellement. Celles-ci étaient soit correctes (50 %) soit incorrectes (violation de restriction de sélection, 50%). Les PEs sont recueillis à partir de 61 canaux et les cartes de potentiels de surface pour chaque point mesuré (500 Hz soit 1 point toutes les 2 ms) pour chaque sujet sont calculées pour la présentation du mot introduisant les violations ainsi que leurs contrôles. Pour les analyses de tracés, l'amplitude moyenne des PEs pour 9 régions d'intérêt dans 3 rangs de latence (0-300 ms, 300-500 ms, 500-800 ms) est mesurée. Alors que pour l'analyse de segmentation, les PEs sont divisés en segments ou micro-états de configuration électrique stable (i.e. *functional microstates*; Lehmann & Skrandeis, 1980). Les cartes de surface obtenues pour chaque segment sont alors comparées statistiquement avec les PEs des données individuelles des participants pour les deux types de phrases. Si l'analyse de tracés montre un effet N400 classique (compris entre 300 et 500 ms avec un maximum de négativité sur les électrodes localisées en région centro-pariétale droite pour les violations sémantiques), l'analyse de segmentation adaptée indique des différences de distribution sur le scalp pour deux segments ou micro-états: le segment 7 enregistré entre 247 et 338 ms et le segment 8 entre 339 et 434 ms. Les violations sémantiques présentent un effet bipolaire par rapport aux phrases contrôles pour les deux segments: négativité sur les sites postérieurs accompagnée d'une positivité sur les sites antérieurs. Pour le segment 7, la distribution de cet effet est symétrique alors que pour le segment plus tardif (8) l'effet est latéralisé: la négativité postérieure est maximale sur les sites postérieurs droits. En présentation visuelle, et grâce à une analyse de segmentation adaptée, la composante N400 semble bipolaire (pôle négatif et positif) et biphasique (2 micro-états compris entre 247-434 ms). Comparée à une analyse de tracés, la segmentation adaptée nous a permis de préciser les processus sous-jacents au traitement des informations sémantiques véhiculées par un verbe.

## 1 Introduction

Grâce à son excellente résolution temporelle, la technique des PEs est très utile dans l'étude des processus langagiers. La composante N400, découverte par Kutas & Hillyard (1980), correspond à la différence dans les PEs entre le traitement d'une violation sémantique et celui obtenu à partir d'une phrase correcte. Cet effet est généralement caractérisé par différents paramètres: sa polarité (négative), son amplitude (de 2 à 5  $\mu\text{V}$ ), sa latence (maximum de négativité vers 400 ms après le début de la présentation de la violation sémantique en question) et sa distribution sur le scalp (maximum en postérieur droit). Les recherches portant sur cette composante N400 ont permis de mieux comprendre les facteurs qui suscitent son occurrence mais aussi qui modulent ses caractéristiques. Par exemple, Kutas & Hillyard (1984) indiquent que la *Cloze Probability* (CP)<sup>8</sup> d'un mot, influence l'amplitude de la composante N400: moins le mot est prévisible par rapport au contexte présenté, et plus l'amplitude de la N400 sera importante. Comme nous l'avons rapporté dans le chapitre 2, la fréquence d'occurrence d'un mot affecte aussi l'amplitude de la composante N400: plus la fréquence d'un mot est basse et plus l'amplitude de la composante N400 sera importante (Rugg, 1990; van Petten & Kutas, 1990). Paller, Kutas, Shimamura, & Squires (1987) observent également que les mots abstraits disposent d'une composante N400 plus ample par rapport aux mots concrets. Pour finir, Besson, Kutas, & van Petten (1992) indiquent que la répétition des stimuli réduit l'amplitude de la composante N400, même si le mot est incongru sémantiquement. D'après ces différents résultats, il apparaît que cette composante reflète la difficulté qu'un lecteur rencontre pour intégrer les informations sémantiques du mot critique dans son contexte. Bien que de nombreuses études dans ce domaine aient permis de mieux cerner les processus impliqués dans l'enregistrement de la composante N400, d'importants problèmes méthodologiques concernant l'analyse des dimensions temporelles (latence) mais aussi spatiales (topographie) demeurent. Dans cette première partie expérimentale, nous allons soulever ces différents problèmes et essayer d'y remédier en appliquant une technique d'analyse dite "spatio-temporelle". Dans la section qui suit, nous soulevons d'importantes questions relatives à la sélection de la fenêtre temporelle et à la distribution des électrodes.

---

<sup>8</sup> Lorsque l'on veut calculer la *Cloze Probability* (CP) d'un mot dans un contexte, on présente la partie droite du contexte sans le mot critique à un groupe de participants en leur demandant de compléter la phrase par le premier mot qui leur vient à l'esprit. Si un nombre important de participants indique la même réponse (mot de CP élevée), ce résultat suggère que le contexte impose de fortes contraintes.

## 1.1 Limites des analyses de tracés

### 1.1.1 Dimension temporelle: définir la fenêtre d'analyse

Toute analyse portant sur des mesures électrophysiologiques de surface doit effectuer des statistiques dans une fenêtre temporelle précise. La recherche de la fenêtre temporelle appropriée n'est pas simple et a précédemment abouti à de nombreuses disparités et problèmes n'autorisant pas la comparaison directe de résultats émanant de différents laboratoires. A l'origine, les études portant sur la composante N400 utilisent une fenêtre d'analyse temporelle comprise entre 300 et 500 ms (Kutas & van Petten, 1988; Kutas & van Petten, 1994). Afin de répliquer ces résultats princeps, les chercheurs ont tout d'abord conservé cette fenêtre stable, alors que celle-ci ne couvre pas forcément la totalité des effets. Pourquoi utiliser une fenêtre de 300 à 500 ms pour les analyses de la composante N400, alors que cette composante débute en général dès 200 ms et finit après 600 ms? Le premier problème mis en évidence ici est, par conséquent, que la fenêtre d'analyse comprise entre 300 et 500 ms ne couvre pas toute la fenêtre d'intérêt.

Généralement, après visualisation des tracés PEs, la fenêtre d'analyse est choisie arbitrairement afin de représenter au mieux les données. Par exemple, Johnson & Hamm (2000) analysent la composante N400 dans une fenêtre comprise entre 400 et 500 ms, alors que Federmeier & Kutas (1999) utilisent celle comprise entre 350 et 400 ms. De plus, dans la dernière étude, les auteurs Federmeier & Kutas (1999) font varier la fenêtre temporelle en fonction de la variable dépendante étudiée: l'amplitude moyenne est analysée entre 350 et 400 ms, alors que la latence du maximum d'amplitude est étudiée entre 350 et 450 ms. Le problème de cette technique est qu'en réalité les fenêtres sont définies à partir de la grande moyenne. En effet, pour obtenir un PE, il ne suffit pas de réaliser un moyennage sur un nombre d'essais différents (voir chapitre 2, section 2.2). Mais aussi sur différents participants (en général un minimum de 15 sujets). Nous savons qu'il suffit d'un seul sujet montrant des différences de PEs entre deux conditions expérimentales alors que tous les autres sujets ne présentent que peu ou pas de différences (ou encore des différences opposées) pour que l'effet du premier sujet ressorte dans la grande moyenne <sup>9</sup>.

<sup>9</sup> Le fait que les grandes moyennes présentent une différence entre les PEs de deux conditions expérimentales n'indique pas forcément que l'effet est statistiquement significatif.

Comme le choix de la fenêtre d'analyse est déterminé par la visualisation des PEs des grandes moyennes, les résultats de l'analyse statistique peuvent, par conséquent, être entièrement déterminés par les PEs d'un seul sujet.

Un second problème provient de certaines expériences dans lesquelles les manipulations de facteurs suscitent de fines modulations de la composante N400 enregistrées sur des fenêtres beaucoup plus restreintes que celles qui sont traditionnellement fixées entre 300 et 500 ms. Dans ces cas précis, la fenêtre temporelle comprise entre 300 et 500 ms est trop large, et si les auteurs conservent celle-ci, ils risquent de ne pas mettre en évidence leurs effets. Récemment, Steinhauer & Ullman (2002) indiquent des résultats présentant ce problème méthodologique. Ces auteurs utilisent des violations de temps sur des verbes réguliers (*Yesterday I sail Diane's boat to Boston*) ainsi qu'irréguliers (*Yesterday we eat Peter's cake in the kitchen*). Pour les deux types d'erreurs, ces auteurs observent une composante LAN (*Left Anterior Negativity*) comprise entre 300 et 500 ms ainsi qu'une composante P600 (voir chapitre 3 section 2.3.3). Si cette fenêtre d'analyse avait été la seule utilisée, alors ces auteurs n'auraient pu valider leur hypothèse. En revanche, en réduisant la fenêtre d'analyse, ils montrent que la composante LAN présente 2 distributions différentes en fonction de la fenêtre temporelle d'étude, indiquant ainsi des processus différents pour les deux types de verbes (régulier et irrégulier). Entre 300-400 ms, la négativité est similaire à la N400 (postérieure) alors qu'entre 400 et 500 ms celle-ci est frontale. D'après ces auteurs, seule la première distribution (entre 300 et 400 ms) semble spécifique au traitement des informations de régularité des verbes.

La deuxième solution employée est donc de systématiquement diviser la fenêtre 300-500 ms en deux fenêtres plus petites comme l'ont réalisé Steinhauer & Ullman (2002). Avec cette procédure Connolly, Stewart, & Phillips (1990) et Holcomb & Neville (1991) définissent la composante N400 dans deux fenêtres temporelles différentes: une fenêtre précoce (150-350 ms), et une plus tardive (350-600 ms) lors de présentation auditive.

Dans une étude plus récente, Hagoort & Brown (2000) se sont intéressés à la possibilité d'un effet biphasique, (c'est à dire, un effet précoce et tardif) lors du traitement des violations sémantiques. Ils comparent des violations sémantiques présentées soit en milieu, soit en fin de phrases. Ils montrent une dissociation entre deux négativités: une N250 (entre 200 et 300 ms) et une N400 (entre 350 et 500 ms). Comme ces deux composantes ne présentent pas la même topographie sur le scalp, ces auteurs en concluent que ces deux composantes reflètent des processus distincts. A la lumière de ces résultats, on peut suggérer que le processus des violations sémantiques implique une (Friederici, *et al.*, 1993; van Petten, Coulson, Rubin, Plante, & Parks, 1999) ou plusieurs étapes différentes (Connolly, Stewart, & Phillips, 1990; Hagoort & Brown, 2000; Holcomb & Neville, 1991) en fonction du choix de la fenêtre d'analyse. En résumé, il apparaît que la résolution temporelle de cette technique n'a pu être exploitée à sa juste valeur et ce, parce que la fenêtre temporelle pour les analyses est définie arbitrairement. En effet, alors que les mesures peuvent être aussi rapides qu'un point toutes les millisecondes (en général, le rythme d'acquisition des données pour ce genre de paradigme est au maximum de 1 point toutes les 2 ms soit un *sampling* de 500 Hz), le fait d'utiliser des fenêtres de plus de 100 ms en moyenne nous ramène par conséquent, à une acquisition de type un point toutes les 100 ms.

### 1.1.2 Dimension spatiale

Les PEs sont des mesures multidimensionnelles qui varient en latence, en amplitude, mais aussi en topographie (voir chapitre 2, section 2). Cette dernière caractéristique des PEs a été moins étudiée car les études originales enregistraient l'activité électrique sur un nombre restreint d'électrodes. Concernant la topographie de la composante N400, les études décrivent celle-ci comme étant maximale sur les sites centro-pariétaux (Johnson & Hamm, 2000; Kutas & Hillyard, 1983; Kutas & van Petten, 1994). Pour la présentation de mots écrits, la composante N400 apparaît comme latéralisée sur les électrodes relatives à l'hémisphère droit, alors que pour des présentations auditives, celle-ci est latéralisée à gauche (Holcomb & Neville, 1990). Bien que de nombreuses études enregistrent désormais un nombre conséquent de canaux, l'approche générale est toujours de réaliser les analyses sur un nombre limité d'électrodes. Ainsi, un choix est nécessaire concernant le nombre, mais aussi la localisation sur le scalp des électrodes qui vont intervenir dans l'analyse statistique. Par exemple, Münte, *et al.* (2001) enregistrent des PEs à partir de 29 électrodes, mais procèdent aux analyses statistiques sur 19 d'entre-elles. Federmeier, Segal, Lombrozo, & Kutas (2000) analysent 15

des 26 électrodes enregistrées. Enfin, certains chercheurs (Hagoort & Brown, 2000) analysent la latence de la composante N400 sur une seule électrode (CZ)<sup>10</sup>. La question qui nous semble cruciale est: pourquoi enregistrer autant de sites différents alors que seul un nombre restreint d'électrodes va être considéré dans les analyses statistiques?

## 1.2 Quelles solutions adopter?

La méthode d'analyse spatio-temporelle a été proposée par Lehmann & Skrandeis (1980) afin de limiter les choix arbitraires précédemment décrits et, par conséquent, de conserver les avantages des enregistrements PEs. En effet, nous verrons dans la section suivante que d'une part, les fenêtres d'analyses sont directement déterminées par les changements de configurations sur le scalp et d'autre part, que toutes les électrodes sont utilisées dans les analyses statistiques.

### 1.2.1 Analyses spatio-temporelles basées sur la cartographie des PEs

Les cartes de champs de potentiels représentent la distribution des potentiels électriques mesurés simultanément à la surface du scalp par de nombreuses électrodes (Lehmann, 1971; Lehmann, 1972; Ragot & Remond, 1978). Les cartes topographiques sont la représentation de la projection d'une surface 3D (le scalp) sur un plan 2D. Cette méthode de représentation consiste à calculer à partir des valeurs numériques des paramètres mesurés pour chaque localisation d'électrode, les valeurs inter-électrodes par une méthode d'interpolation. Ce calcul s'effectue grâce à des algorithmes et permet de représenter les cartes au moyen de lignes ou zones équipotentiellles (distribution du champ) codées en couleur (Nunez, 1981): le bleu est utilisé pour les valeurs négatives et le rouge, pour les positives. Ce procédé conduit à une illustration graphique de la configuration électrique à chaque instant donné (dépendant du *sampling* d'acquisition). L'intérêt d'une telle représentation est que la répartition des lignes équipotentiellles est indépendante de l'électrode de référence, contrairement aux formes d'ondes temporelles enregistrées au niveau de chaque électrode. Ceci représente un avantage sur les analyses de tracés qui, dépendent du choix de la référence (Nunez, Pilgreen, Westdorp, Law, & Nelson, 1991) contrairement aux analyses spatio-temporelles.

---

<sup>10</sup> Alors que 7 ont été enregistrées.



Nous avons vu (chapitre 2, section 1) que les PEs, et par conséquent, les cartes topographiques de surface, correspondent à un champ électrique global qui est généré par l'ensemble des neurones sous-jacents activés à un instant  $t$ . Une distribution différente des neurones actifs peut produire la même carte topographique de surface. A l'inverse, un changement de configuration du champ électrique ne peut être provoqué que par une modification de la distribution des neurones actifs dans le cerveau. Les différentes topographies de surface observées, peuvent donc être un indice précieux quant à l'activation de neurones spécifiques (Fender, 1987; Lehmann, 1987; Skrandeis, 1986). Si l'on considère qu'une fonction cognitive peut être générée par l'activation de neurones spécifiques localisés dans les aires corticales, alors le résultat de cette activation est un courant de conduction (résultat de la sommation des différents dipôles unitaires, voir chapitre 2, section 1), observable à l'aide d'électrodes disposées sur le scalp. La configuration de ces neurones conduit à l'enregistrement d'une carte de surface qui sera spécifique. Ainsi, l'activation de deux fonctions cognitives différentes engageant deux groupes de neurones distincts doit pouvoir induire des réponses électrophysiologiques spécifiques sur le scalp.

Deux cartes PEs de surface peuvent varier selon deux dimensions: en distribution topographique, mais aussi en puissance (c'est à dire lorsqu'une carte est identique mais que les valeurs du PE varient en parallèle sur tous les sites). Alors que d'un point de vue cognitif, seules les variations de topographie sont intéressantes (des distributions sur le scalp différentes reflètent l'activation de deux groupes de neurones distincts), l'objectif principal des analyses spatio-temporelles est donc d'identifier et de caractériser ces changements dans la distribution topographique des cartes de surface, tout en les discriminant des variations de puissance du champ électrique (conservation de la même configuration). Afin d'aboutir à différencier changement de topographie sur le scalp et changement de puissance des cartes PEs, il nous semble nécessaire d'utiliser des critères objectifs et de pouvoir les décrire de manière numérique.

#### 1.2.1.1 Le *Global Field Power*: GFP

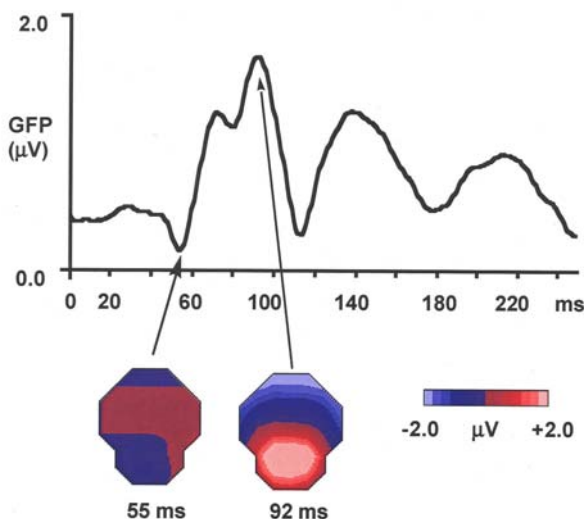
Le champ de potentiels électriques sur le scalp est caractérisé par une certaine puissance. Lehmann & Skrandeis (1980) suggèrent une méthode efficace pour représenter la puissance des topographies des PEs, basée sur une représentation numérique de la puissance du champ électrique. La synchronisation de l'activité neuronale corticale peut être représentée par un

champ électrique fort. A l'inverse, des champs électriques faibles indiquent que l'activité neuronale sous-jacente est plus distribuée et moins synchronisée <sup>11</sup>. Plus précisément, un champ électrique fort est caractérisé sur le scalp par un important gradient (c'est à dire la différence entre la valeur maximale et minimale de la carte de surface est importante) et par conséquent, est représenté avec de nombreuses courbes équipotentielles. Lehmann & Skrandeis (1980) appellent la puissance du champ électrique le *Global Field Power* (GFP). Cet indice correspond à la déviation standard spatiale (ou écart-type) d'une carte lorsque les valeurs de celle-ci sont obtenues à partir d'une référence moyenne (Bertand, Perrin, & Pernier, 1985; Lehmann, 1971; Nunez, 1981). Dans la figure 8, la formule mathématique du GFP est indiquée et peut être calculée pour chaque point au cours du temps. La puissance d'un champ électrique peut être représentée par une seule valeur numérique qui fluctue au cours du temps et est illustrée dans la figure 9 pour une tâche visuelle.

$$GFP = \sqrt{\frac{\sum_{i=1}^n (u_i - \bar{u})^2}{n}} \quad \bar{u} = \frac{\sum_{i=1}^n u_i}{n}$$

avec  $\bar{u}$ : valeur de la référence moyenne;  $n$ : nombre total d'électrodes;  $u_i$ : voltage à l'électrode  $i$

**Figure 8:** Formule du *Global Field Power* (GFP).



**Figure 9:** Illustration du déroulement temporel de la courbe de GFP (*Global Field Power* = puissance de la carte) à partir d'un PE associé à une stimulation du champ visuel central. Deux cartes topographiques de surface, prises à deux instants  $t$  différents dans le temps sont représentées: l'une à 55 msec, avec un GFP faible, l'autre à 92 msec, correspondant à un pic de GFP (champ fort). Le GFP permet d'indiquer chaque point dans le temps où la puissance du champ électrique est élevée. Dans cet exemple, les pics de GFP sont enregistrés à 70, 100, 140 et 200 msec, et correspondent à des composantes bien connues des PEs visuels. L'intérêt de cette mesure numérique du champ électrique (GFP) réside dans le fait que toutes les électrodes sont prises en compte. D'après Michel (1997).

<sup>11</sup> Ce type de carte de surface peut aussi provenir de l'activation de sources neuronales localisées en profondeur dans le cerveau.

Même s'il ne reflète pas des variations de topographie, nous verrons dans nos parties expérimentales que cet indice est utile pour mettre en évidence les différentes composantes que constituent un PE.

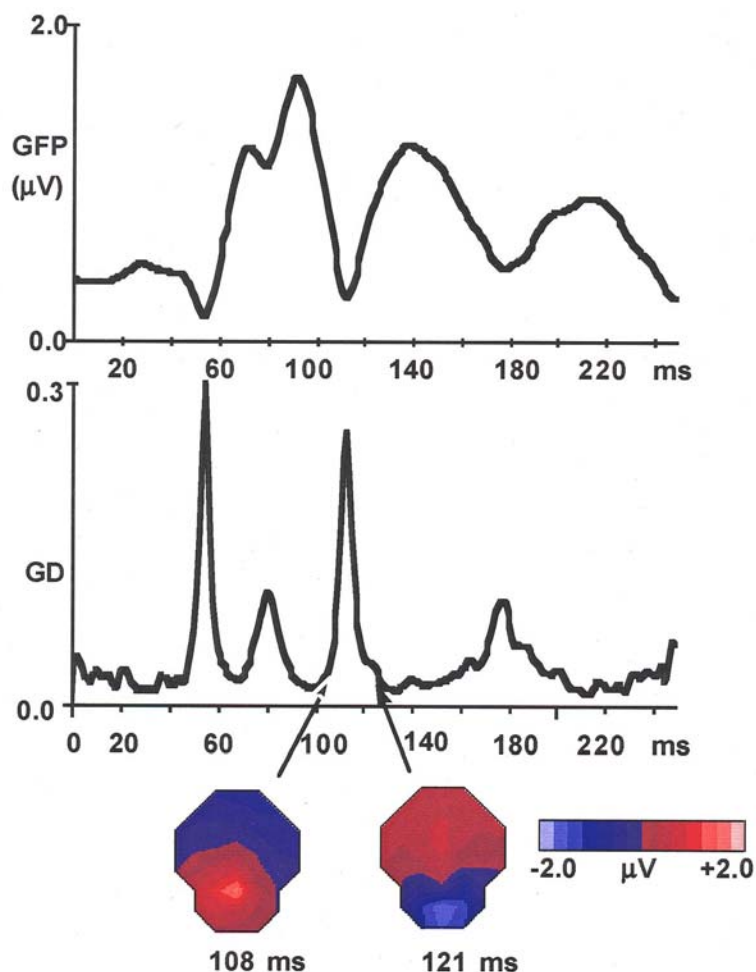
### 1.2.1.2 La *Global Dissimilarity*: GD

La *Global Dissimilarity* (GD), comme son nom l'indique, est un indice qui calcule la dissimilarité entre deux cartes données. Lors de ce calcul, la puissance du champ électrique est réduite à néant car les valeurs des PEs sont normalisées (pour une discussion sur l'utilité de la normalisation des données voir la section 3.1 de ce chapitre). La figure 10 donne la formule mathématique de cet indice.

$$DISS = \sqrt{\frac{\sum_{i=1}^n (u_i - v_i)^2}{n}} \quad u_i = \frac{u_i}{GFP_u} \quad v_i = \frac{v_i}{GFP_v} \quad u_i = U_{org.i} - \bar{u} \quad v_i = V_{org.i} - \bar{v}$$

**Figure 10:** Formule de la *Global Dissimilarity* (GD).

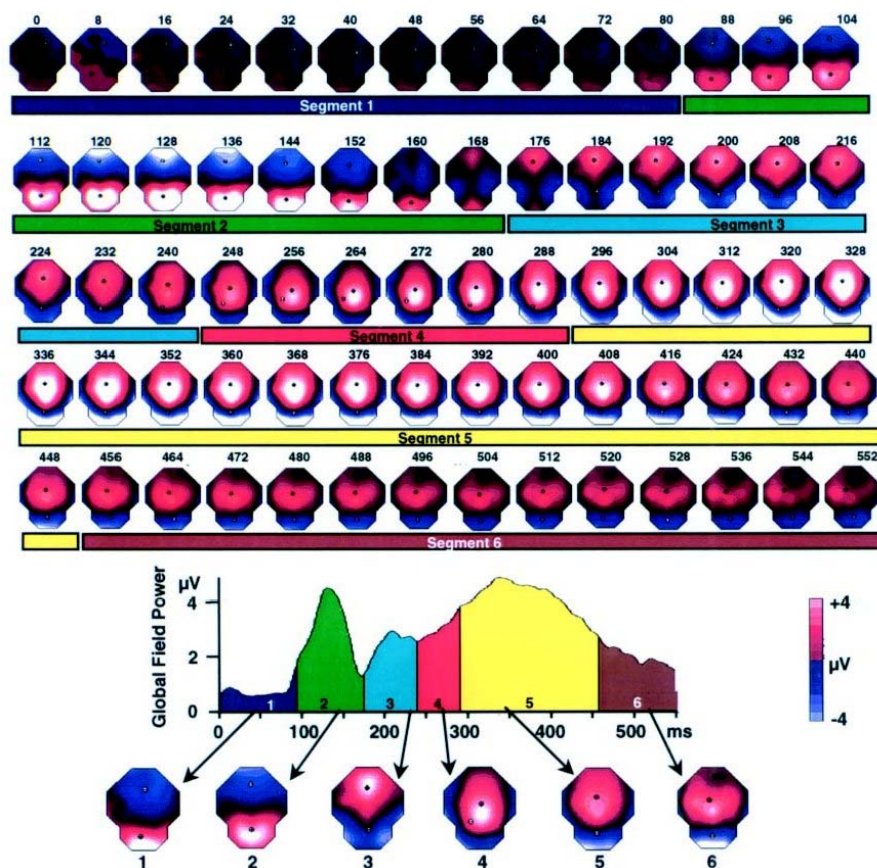
Brandeis, *et al.* (1992) montrent que cet indice est une fonction décroissante du coefficient de corrélation spatiale entre deux cartes. La GD peut être calculée soit entre deux cartes successives d'une même condition soit entre deux cartes obtenues à partir de deux conditions expérimentales différentes. Dans le premier cas, cet indice reflétera les différences topographiques dans une même condition expérimentale et par conséquent nous permettra de caractériser des changements topographiques au cours du temps. En effet, lorsque l'indice de GD est faible, une période de quasi-stabilité dans la configuration électrique des PEs de surface peut être définie. A l'opposé, une forte valeur de GD indique que la carte change de topographie électrique. La figure 11 illustre les relations qui existent entre GD et GFP sur un exemple de PEs visuels. Dans le second cas où la GD est calculée entre deux cartes obtenues à partir de conditions expérimentales différentes, l'indice nous indiquera si les deux cartes proposent des topographies différentes et par conséquent, si celles-ci sont associées à l'activation de groupes de neurones spécifiques.



**Figure 11:** Illustration du GFP et du GD en fonction du temps pour la même condition expérimentale (stimulation visuelle en vision centrale). Les pics de GD (schéma du bas) indiquent l'instant dans le temps où la différence de topographie entre les deux cartes successives ( $t$  et  $t+1$ ) est forte. Deux exemples de cartes sont indiqués pour 108 ms et 121 ms. Comme ces deux cartes de surface entourent un pic de GD, leurs topographies sont très différentes. On peut noter la coïncidence des pics de GD avec les minima du GFP. D'après Michel (1997).

### 1.2.2 Micro-états fonctionnels et segmentation

En inspectant une série de cartes de surface, un phénomène basique peut être observé: les distributions de topographie de surface changent très rapidement (en quelques ms), à la suite desquelles des périodes de plus ou moins longue stabilité électrique peuvent être observées (Lehmann & Skrandeis, 1980). La figure 12 (haut) illustre l'évolution au cours du temps des cartes de surface obtenues à partir d'une stimulation visuelle centrale. Pendant les périodes de stabilité, les cartes augmentent et diminuent en intensité (ou en puissance représentée par le GFP, figure 12, bas), mais la topographie de surface reste quasi constante (figure 12 haut). Ce comportement spatio-temporel de l'activité électrique de surface a été décrit pour l'activité spontanée EEG, mais aussi pour les PEs (pour une revue voir Lehmann, 1987) et est à la base du concept de micro-état fonctionnel (Lehmann & Skrandeis, 1980).



**Figure 12:** Résultat d'une segmentation spatio-temporelle effectuée sur une série de cartes de surface allant de 0 à 552 ms (une carte toutes les 8 ms), associées à une stimulation visuelle centrale. Sur chaque carte, sont représentés les centroïdes indiquant les centres de gravité des voltages positifs et négatifs. En dessous des cartes de surface, sont représentées les périodes de stabilité (une couleur pour chaque) et leurs frontières. Six périodes de stabilité reflètent l'activité lors de cette tâche et représentent des micro-états fonctionnels. Le graphique du bas schématise la courbe de GFP sur laquelle les différents segments ou micro-états sont indiqués. Pour chaque segment, une seule carte de surface est alors calculée. D'après Michel, Seeck, & Landis (1999).

La segmentation est définie comme une procédure permettant de définir les périodes où les cartes de surface présentent les mêmes configurations sur le scalp. Chaque période représente alors un micro-état, autrement appelé "segment". Chacun d'entre eux reflète un état stable de configuration électrique et par conséquent, une activation d'une même population neuronale (Lehmann & Skrandeis, 1980). Le premier postulat suggère qu'un changement de configuration de topographie sur le scalp implique l'activation de groupes de neurones différents dans le cerveau. Le second postulat suggère que les étapes du traitement de l'information sont caractérisées par l'activation de groupes de neurones différents. Par conséquent, un changement de topographie suggère l'activation d'une nouvelle étape dans le traitement des informations (d'où l'appellation de "micro-états fonctionnels"). Il est important de noter que le concept de micro-état fonctionnel ne nécessite pas une conception sérielle du

traitement de l'information. En effet, une carte de surface représentant un micro-état est le résultat de la sommation de tous les neurones actifs dans le cerveau au moment  $t$ . Il est fort probable que plusieurs zones corticales puissent concourir à l'élaboration d'une carte électrique de surface. Par conséquent, un micro-état peut refléter l'activation simultanée de plusieurs zones corticales, et peut donc représenter l'activation en parallèle de différentes étapes.

La question pertinente reste de déterminer comment un micro-état peut être défini dans le profil des cartes de surface de manière objective. Différents critères ont été élaborés afin d'établir les différences topographiques: report des maxima et minima du voltage d'une carte (Koenig & Lehmann, 1996; Lehmann, *et al.*, 1987; Michel & Lehmann, 1993), fluctuation des localisations des centres de gravités (Brandeis & Lehmann, 1994; Brandeis, Lehmann, Michel, & Mingrone, 1995; Brandeis, van Leewen, Rubia, Vitacco, Steger, Pascual-Marqui, & Steinhausen, 1998; Koenig & Lehmann, 1996). Une approche développée par Pascual-Marqui, Michel, & Lehmann (1995) est d'utiliser une analyse statistique en clusters. Pour nos analyses, nous utiliserons une autre solution, la segmentation adaptée, que nous développerons dans la section 3.2 de ce chapitre.

### 1.3 Segmentation et composante N400

Quelques études ont effectué des analyses de segmentation en micro-états fonctionnels avec un paradigme suscitant l'occurrence d'une composante N400. Brandeis, *et al.* (1995) présentent à des sujets sains des phrases de 4 mots qui sont soit terminées par un mot de couleur appropriée (le chat est noir), soit sémantiquement non appropriée (le chat est bleu). Alors que les phrases sont très courtes et qu'une couleur est toujours présentée en fin de phrase, le nom en position finale induit une forte attente en fonction du contexte (*forte Cloze Probability*). Ces auteurs montrent que la différence dans les PEs de phrases correctes et incorrectes s'observe sur deux micro-états ou segments différents: un segment "pré-N400" et un segment "N400". Le segment "pré-N400" enregistré entre 244 et 364 ms (120 ms de durée) est caractérisé par une positivité maximale sur les sites antérieurs gauches et une négativité sur les électrodes postérieures droites. Le second segment, "N400", enregistré entre 364 et 444 ms (80 ms de durée) présente une topographie plus symétrique: positivité antérieure, et négativité postérieure. Avec un paradigme similaire et dans la modalité auditive, Karniski, Vanderploeg, & Lease (1993) décomposent le rang de latence correspondant à la composante

N400 en deux segments spécifiques: "virtual N400" et "slow wave". Le premier segment "virtual N400" enregistré entre 196 et 287 ms (91 ms de durée) présente une positivité antérieure gauche accompagnée sur les sites postérieurs droits par une négativité. Le deuxième segment "slow wave" enregistré entre 418 et 530 ms (112 ms de durée) présente une topographie symétrique comme pour les résultats de Brandeis, *et al.* (1995). Brandeis, *et al.* (1995) et Karniski, *et al.* (1993) utilisent des violations en fin de phrases pour susciter l'occurrence d'une composante N400. Cependant, il n'est pas certain que leurs résultats reflètent uniquement les processus liés au traitement d'une erreur sémantique. En effet, il existe un nombre conséquent d'études montrant que la position finale du mot critique peut engendrer des effets liés à la fin de lecture de la phrase. Ces effets seraient dus à des processus décisionnels ou encore liés avec la réponse motrice (Osterhout, 1997; Osterhout & Mobley, 1995). Par exemple, Osterhout (1997) montre qu'une violation syntaxique placée en milieu d'une phrase engendre l'occurrence d'une composante similaire à la N400 pour le dernier mot de la phrase. Comme le suggèrent Hagoort & Brown (2000), les violations d'ordre sémantique suscitent l'occurrence de N400 variables dans leur amplitude en fonction de la position du mot critique dans la phrase. Puisque la plupart des paradigmes expérimentaux favorisent la présentation des violations sémantiques en fin de phrase (Besson, Faita, Czernasty, & Kutas, 1997; Besson & Kutas, 1993; Besson, *et al.*, 1992; Kutas & Hillyard, 1980; Kutas & Hillyard, 1983; Kutas & van Petten, 1994), il semble intéressant d'investiguer plus amplement la présentation de violations sémantiques en milieu de phrase.

## 1.4 Objectifs

Le but de cette étude est d'explorer plus amplement le composante N400 en comparant deux types d'analyses: la méthode des tracés et la segmentation adaptée. En appliquant l'analyse de segmentation adaptée, nous allons préciser si les violations sémantiques génèrent deux effets de topographies différents sur le scalp (un effet précoce et un plus tardif) comme le suggèrent les résultats de Brandeis, *et al.* (1995) et de Karniski, *et al.* (1993). Alors qu'aucune étude utilisant l'analyse classique des tracés n'observe cet effet biphasique en présentation visuelle des stimuli, cet effet a été enregistré en modalité auditive (Connolly, *et al.*, 1990; Hagoort & Brown, 2000). Connolly, *et al.* (1990) et Hagoort & Brown (2000) émettent l'hypothèse selon laquelle cet effet précoce (PMN/N200) reflète des processus acoustico-phonologiques liés à la modalité de présentation auditive. Pour des expériences utilisant la technique des PEs en modalité visuelle, les phrases doivent être présentées mot par mot au centre de l'écran afin

d'éviter les artefacts liés aux mouvements oculaires. Ainsi, puisque les lecteurs peuvent fixer et lire en une seule fois les mots, les processus acoustico-phonologiques doivent être éliminés. Selon Hagoort & Brown (2000), comme la N200/PMN reflète des processus acoustico-phonologiques et que ce type de traitement est absent lors d'une présentation visuelle, alors la composante N200/PMN est normalement absente des paradigmes en modalité visuelle (Hagoort & Brown, 2000).

Grâce à une analyse de segmentation adaptée, Brandeis, *et al.* (1995) mettent en évidence un effet bi-phasique ("pré-N400" et "N400") lorsque les stimuli sont présentés en modalité visuelle. Que représente le segment pré-N400 enregistré par ces auteurs? Comme ces auteurs utilisent des violations en fin de phrase, il est possible que leurs résultats reflètent en réalité des processus spécifiques liés à la position du mot en fin de phrase. Si une réplification d'un effet bi-phasique est observée lors du processus visuel des violations sémantiques en position non finale dans une phrase, alors nous pourrions, dans ce cas précis, rejeter l'interprétation de Hagoort & Brown (2000). Nous allons appliquer en parallèle des analyses de tracés et de segmentation adaptée sur des PEs enregistrés lors de la présentation visuelle de phrases comportant ou non des violations sémantiques en position non finale. Les analyses de tracés vont nous permettre de comparer nos résultats avec la littérature précédente. En définissant les périodes de stabilité dans les cartes PEs de surface, la segmentation adaptée va nous indiquer les fenêtres temporelles à partir desquelles nous allons appliquer nos analyses statistiques. Si la topographie d'un micro-état est altérée pour les phrases présentant une violation sémantique par rapport à une phrase correcte, alors ce résultat indiquera que différentes structures neuronales ont été activées pendant cette période. Ainsi, on pourrait conclure que les deux types de phrases (violations et correctes) ont suscité l'activation de groupes de neurones séparés (ou partiellement) lors de la période étudiée. Une hypothèse alternative est que les violations sémantiques peuvent induire les mêmes topographies dans les PEs de surface que les phrases correctes, mais qu'un ou plusieurs micro-états soit décalés dans le temps. Pour ce type de données, une analyse de tracés indiquera un effet classique de différence de topographie et ne permettra pas de montrer le décalage dans le temps (Frishkoff & Tucker, 2001). En revanche, une analyse de segmentation adaptée sera à même de détecter des topographies de surface identiques présentant un décalage dans le temps.



## 2 Méthode

### 2.1 Sujets

Vingt étudiants (1 homme et 19 femmes) de la Faculté de Psychologie et des Sciences de l'Education de Genève ont participé bénévolement à cette expérience (moyenne d'âge de 21.6, compris entre 19 et 31 ans). Tous les sujets étaient droitiers, d'après le test de latéralité *Edinburgh Handedness Inventory* (Oldfield, 1971), de langue maternelle française, et présentaient une vision normale ou corrigée.

### 2.2 Stimuli

Cent soixante phrases ont été construites avec une structure identique: adverbe ou préposition + sujet + verbe + complément. La longueur moyenne des phrases était de 7.5 mots (de 5 à 10 mots). La moitié de ces phrases (80) étaient sémantiquement correctes alors que l'autre moitié (80) présentaient des incongruités sémantiques entre le sujet et le verbe (restriction de sélection sur le verbe). La liste complète des stimuli est présentée dans l'annexe 1.

#### **Exemple de phrases :**

Aujourd'hui ma voisine achète la voiture. (phrase de contrôle)

Aujourd'hui ma chatte achète la voiture. (phrase incorrecte)

Deux listes de 80 phrases (40 phrases correctes et 40 violations sémantiques) ont été créées afin que chaque participant ne lise une phrase que dans une seule condition expérimentale (contrôle ou violation sémantique). Dix participants ont été soumis à la première liste, 10 à la seconde. La base de données BRULEX (Content, *et al.*, 1990) a été utilisée pour sélectionner les mots cibles (position du verbe) afin de contrôler la fréquence d'occurrence ( $t_{39} = -.83$ ,  $p = .41$ ) et le nombre de lettres ( $t_{39} = -.88$ ,  $p = .37$ ) des mots entraînant une violation sémantique ( $28786 \pm 8031$  fréquence d'occurrences/million ou encore  $3.83 \pm .9$  log de fréquence,  $8.4 \pm .4$  nombre de lettres) et des mots des phrases de contrôles ( $22000 \pm 4441$  fréquence d'occurrences/million, ou encore  $3.98 \pm .6$  log de fréquence,  $7.9 \pm .3$  nombre de lettres).

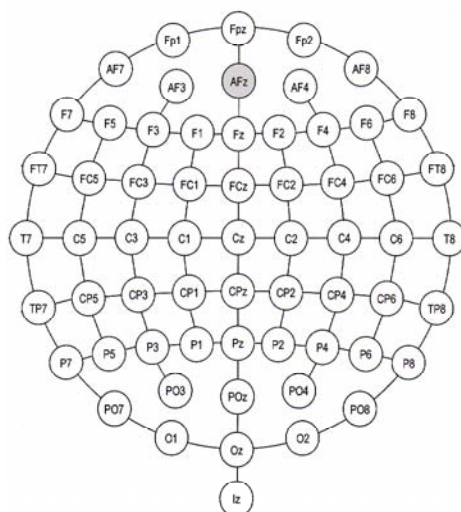
## 2.3 Procédure

Les participants lisaient chaque phrase présentée mot à mot au centre de l'écran d'ordinateur. Chaque mot était visible pendant 300 ms, avec un intervalle entre chaque mot de 500 ms (*Stimulus Onset Asynchrony*, SOA de 800 ms). Chaque phrase commençait avec une croix centrale de fixation (300 ms). Le premier mot de chaque phrase était écrit avec une majuscule, et toutes les autres lettres utilisées étaient des minuscules. Le mot final de chaque phrase était présenté avec un point final. A la fin de chaque phrase, deux points d'interrogation "??" indiquaient au sujet qu'il devait répondre si la phrase était correcte ou non. Les réponses étaient obtenues à l'aide d'un boîtier réponse: la moitié des sujets répondaient avec le pouce de leur main dominante (main droite) lorsque la phrase était correcte, et avec le pouce de l'autre main si la phrase était incorrecte. Pour l'autre moitié, les mains de réponses étaient inversées. Après chaque "??", les participants visualisaient un écran vide pendant 2000 ms, ce qui leur permettait de cligner des yeux et de se reposer. Les phrases étaient mélangées aléatoirement pour chaque liste et divisées en deux blocs de 40 phrases, d'une durée approximative de 5 minutes de stimulation chacun. Avant le début de l'expérience, le participant lisait la consigne écrite et effectuait une phase d'apprentissage d'une dizaine de phrases.

## 2.4 Enregistrement de l'EEG

Les participants étaient installés confortablement dans une pièce isolée acoustiquement et électriquement, à approximativement 1 mètre de l'écran d'ordinateur. L'EEG était enregistré à partir d'un montage de 61 électrodes Ag/AgCl positionnées sur un bonnet (Easy-Cap, FMS, München, Allemagne) et organisées en accord avec le système international 10-20 version révisée (Society, 1994). En plus des électrodes incluses dans le bonnet, nous avons effectué des enregistrements bipolaires pour les EOG (électro-oculogramme) verticaux et horizontaux. Une électrode positionnée sur le mastoïde gauche a été utilisée comme référence pour les enregistrements, l'électrode Afz servait de terre. L'acquisition des données EEG s'est effectuée en continu pendant chaque bloc, à un échantillonnage de 500 Hz (un point toutes les 2 ms) avec des filtres bandes passantes comprises entre 0.1 et 70 Hz (amplificateurs Neuroscan Inc., Herndon, Va., USA). Les artéfacts liés aux mouvements des yeux ont été réduits grâce à l'algorithme fourni par la compagnie Neuroscan (Scan 4.1, Neuroscan Inc.,

Herndon, Va., USA). L'EEG a alors été visualisé afin de rejeter les périodes où l'amplitude des canaux frontaux (3 électrodes: FP1, FPz, FP2) excédait  $\pm 50 \mu\text{V}$ .



**Figure 13:** Illustration de la configuration des 61 électrodes enregistrées sur le scalp. Les nombres impairs sont disposés sur l'hémisphère gauche, et les nombres pairs sur le droit. L'électrode grisée correspond à Afz et a été utilisée comme terre.

Les segments d'EEG libres de tout artefact ont alors été moyennés pour des fenêtres allant de 0 à 800 ms. Le signal a été digitalisé après un filtrage à 30 Hz et la ligne de base utilisée correspond aux 100 ms précédant le stimulus. Cette procédure procure deux jeux (contrôle et violation sémantique) de PES pour chaque participant.

### 3 Analyses des PES

Les PES pour chaque mot introduisant une violation dans la phrase (position du verbe) et leur contrôle ont été soumis à l'analyse statistique. Deux procédures ont été effectuées en parallèle de manière indépendante:

- une analyse de tracés
- une analyse de segmentation adaptée

Pour la première approche, les fenêtres temporelles utilisées pour l'analyse statistique correspondent à celles présentées dans la littérature. Pour la seconde analyse, les fenêtres temporelles ont été définies à partir de l'analyse spatio-temporelle: ce sont les variations des configurations électriques sur le scalp qui définissent celles-ci (voir section 1.2.1).

### 3.1 Analyse des tracés

A partir des PEs, les mesures quantitatives sur deux variables dépendantes (amplitude moyenne et latence de pics) pour chaque sujet et chaque condition expérimentale ont été analysées.

Dans un premier temps, la moyenne de l'amplitude des PEs (en référence mastoïde gauche, et relatif à la ligne de base comprise entre -100 et 0 ms avant la présentation du stimulus critique) a été calculée pour 3 fenêtres temporelles: une première comprise entre 300 et 500 ms, qui correspond au rang de latence de la composante N400; et deux autres fenêtres, une comprise entre 0 et 300 ms et l'autre plus tardive, entre 500 et 800 ms. Trois ANOVAs à mesures répétées séparées ont été réalisées sur la variable dépendante: amplitude moyenne du PEs dans les 3 rangs de latence. Pour chaque analyse, trois facteurs intra-sujets ont été utilisés. Le premier consistait dans le "type de phrase" (contrôle *vs.* violation sémantique). Les deux derniers facteurs avait trait à la localisation topographique des cartes de surface. Les électrodes ont été regroupées en deux facteurs spatiaux ou régions d'intérêt à trois dimensions: "caudalité" (antérieur *vs.* médian *vs.* postérieur) et "latéralité" (gauche *vs.* central *vs.* droit). Les 61 électrodes ont donc été divisées en 9 (3 x 3) régions d'intérêt (tableau 3). Par conséquent, une seule valeur d'amplitude moyenne du PE pour chaque région d'intérêt, pour chaque fenêtre temporelle, pour chaque condition et pour chaque participant a été calculée.

Facteurs Spatiaux		Electrodes
Caudalité	Latéralité	
Antérieur	Gauche	FP1, AF7, AF3, F7, F5, F3
Antérieur	Central	FPz, F1, Fz, F2
Antérieur	Droit	FP2, AF4, AF8, F4, F6, F8
Médian	Gauche	FT7, FC5, FC3, T7, C5, C3, TP7, CP5, CP3
Médian	Central	FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, CP2
Médian	Droit	FC4, FC6, FT8, C4, C6, T8, CP4, CP6, TP8
Postérieur	Gauche	P7, P5, P3, PO7, PO3, O1
Postérieur	Central	P1, Pz, P2, POz, Oz, Iz
Postérieur	Droit	P4, P6, P8, PO4, PO8, O2

**Tableau 3:** Description des facteurs spatiaux utilisés pour les ANOVAs à mesures répétées: "type de phrase" (contrôle *vs.* violation sémantique), "caudalité" (antérieur *vs.* médian *vs.* postérieur) et "latéralité" (gauche *vs.* central *vs.* droit). Aucune électrode n'est rejetée de cette analyse.

Comme préconisé dans la revue de questions sur les PEs de Picton, *et al.* (2000), deux corrections des données ont été appliquées à nos résultats.

Premièrement, une ANOVA à mesures répétées nécessite des données normalement distribuées avec une variance homogène pour chaque groupe. On parle de sphéricité, ou de covariance égale pour chaque paire de la mesure répétée. Cette condition est habituellement violée par les données psychophysologiques (Jennings, 1987). Afin de compenser de telles violations, le degré de liberté peut être réduit en calculant la valeur d'épsilon ( $\epsilon$ , qui correspond à la mesure comprise entre 1 et 0 de l'homogénéité des variances et covariances) selon la méthode de Greenhouse & Geisser (1959; voir Winer, 1971) ou de Huynh & Feldt (1976). Pour chaque ANOVA à mesures répétées, avec un degré de liberté supérieur à 1 au numérateur, nous appliquerons la correction de Greenhouse-Geisser (1959) et nous indiquerons la valeur corrigée de  $p$ .

Deuxièmement, les interactions significatives entre les facteurs "type de phrase" (contrôle *vs.* violation sémantique) et "spatiaux" (régions d'intérêt) peuvent être dues à des différences d'amplitude plutôt que de topographie (McCarthy & Wood, 1985). Si un générateur est localisé en région antérieure gauche, et que dans une condition expérimentale X, il augmente en puissance (par rapport à la condition Y), alors la carte de surface résultante pourra donner des variations de topographie: proche du générateur, les modifications d'amplitude sur le scalp seront importantes alors que les régions éloignées subiront des fluctuations de potentiels moindres. Afin de déterminer quantitativement si les topographies sont réellement différentes, il est nécessaire d'éliminer les différences d'amplitude avant la réalisation de cette analyse. Une méthode couramment utilisée est de normaliser les données avant le calcul de l'ANOVA à mesures répétées. Différentes possibilités existent, et nous avons retenu la procédure de normalisation  $z$ -score, décrite par Rösler, Heil, & Glowalla (1993), équivalente à la procédure de normalisation suggérée par McCarthy & Wood (1985).

Dans un deuxième temps, une analyse de la latence de pics a été réalisée. Ainsi, pour les composantes d'intérêt, la latence des pics pour certaines électrodes précises (choix arbitraire réalisé à partir de la visualisation du grand moyennage ou l'amplitude de la composante est maximum) a été déterminée pour chaque condition et chaque participant. Des ANOVAs à mesures répétées séparées ont été calculées pour chaque composante avec les facteurs

suivants: "type de phrase" (contrôle vs. violation sémantique) et "électrodes" (pour cette analyse un nombre de 3 électrodes a été retenu).

### 3.2 Analyse de segmentation adaptée

Cette analyse fait intervenir 4 étapes différentes que nous allons maintenant expliciter.

Les PEs ont été recalculé avec une référence moyenne. Les PEs de chaque condition sont alors représentés comme une série de cartes de surface qui se succèdent dans le temps, plutôt que comme des courbes multiples dans le temps. La segmentation adaptée est basée sur le principe selon lequel les cartes de surface des PEs restent stables pendant une certaine période de temps, les changements de configuration électrique s'effectuent brusquement (Lehmann, 1987; Lehmann & Skrandeis, 1980). Les périodes stables de configurations de surface vont représenter des micro-états fonctionnels pendant le traitement de l'information (Thut, Hauert, Morand, Seeck, Landis, & Michel, 1999). Les périodes d'instabilité doivent être déterminées afin de définir les frontières des segments ou micro-états: la GD (*Global Dissimilarity*) est calculée entre les paires de cartes de surface successives pour la grande moyenne des PEs des deux conditions expérimentales séparément (contrôle et violation sémantique). Cette mesure permet d'estimer la force des changements de configuration électrique globaux entre deux cartes de PEs (Lehmann & Skrandeis, 1980). Le décours temporel de cette courbe GD indique directement les moments où les cartes de surface changent de topographie et, par conséquent, la limite d'un segment. Un indice bas de GD suggère que les cartes ont des topographies similaires (Lehmann & Skrandeis, 1980). Un indice complémentaire, le GFP (*Global Field Power*), déviation standard spatiale calculée sur toutes les électrodes (Lehmann & Skrandeis, 1980), est utilisé afin d'affiner cette division en micro-états dans le temps. Cet indice est une mesure de la force du champ électrique. Puisque l'instabilité électrique coïncide avec des cartes de surface faibles (c'est à dire de faible GFP), les minimums de GFP peuvent être utilisés comme indices supplémentaires indiquant la frontière d'un segment (Lehmann & Skrandeis, 1980). La segmentation adaptée définit une bordure de segment lorsqu'il y a convergence entre un faible indice de GFP et une GD élevée. Pour chaque période ainsi définie, une carte de surface est alors calculée. Une procédure de "relabelisation" a été employée afin d'éliminer les segments présentant soit des durées courtes (de quelques ms) soit des cartes de surface similaires (Khateb, Annoni, Landis, Pegna, Custodi, Fonteneau, Morand, & Michel, 1999). Cette procédure compare chaque ms d'un PE (grande moyenne)

d'une condition expérimentale avec les cartes moyennes des périodes de stabilité définies et calcule une corrélation spatiale. Nous avons pu utiliser cette procédure introduite par Pegna, Khateb, Spinelli, Seeck, Landis, & Michel (1997) et implémentée dans le logiciel Cartool V2.11 (Functional Brain Mapping Laboratory, Hôpital Cantonal de Genève). A partir de cet indice, les segments présentant des topographies similaires peuvent être regroupés pour ne former qu'un seul micro-état (Khateb, *et al.*, 1999). Après l'obtention d'une solution stable, les frontières de segments sont définies et les cartes moyennes sont calculées afin d'obtenir une représentation unique pour chaque micro-état. La procédure de segmentation adaptée peut alors être caractérisée par une réduction des données dans le temps. Cette première étape de la segmentation adaptée est similaire à celle utilisée par Michel, *et al.* (1992), Brandeis, *et al.* (1995) et Khateb, *et al.* (1999).

La deuxième étape de l'analyse correspond à la détermination des cartes temporelles spécifiques pour la condition expérimentale critique (ici la présentation des violations sémantiques). Pour cela, une procédure de "*fitting*" est utilisée, introduite par Pegna, *et al.* (1997) et implémentée dans le logiciel Cartool V2.11 (Functional Brain Mapping Laboratory, Hôpital Cantonal de Genève). La carte de surface calculée pour la période de stabilité est introduite dans une période temporelle définie des PEs de chaque sujet (données individuelles) et va permettre d'obtenir un coefficient de corrélation spatiale au cours du temps (Khateb, *et al.*, 1999; Khateb, Michel, Pegna, Thut, Landis, & Annoni, 2001; Morand, Thut, de Peralta, Clarke, Khateb, Landis, & Michel, 2000; Pegna, Khateb, Spinelli, Seeck, Landis, & Michel, 1997; Thut, Hauert, Viviani, Morand, Spinelli, Blanke, Landis, & Michel, 2000; Thut, *et al.*, 1999). Le coefficient de corrélation spatiale correspond à un indice de changement de configuration spatiale (Brandeis, *et al.*, 1992; Lehmann & Skrandeis, 1980). Pour chaque carte de la condition de violation sémantique prédéfinie par la segmentation, nous allons "*fitter*" celle-ci dans les données PEs individuelles des conditions expérimentales de contrôle et de violation sémantique. Une visualisation de la variance au cours du temps en résulte pour chacune des cartes des segments stables, et pour chaque participant. La procédure de "*fitting*" permet de récupérer l'indice de "*Best Explained Variance*" (BEV) ou pourcentage maximum de variance expliquée pour chaque sujet et pour les deux conditions expérimentales. Un *t*-test pour des échantillons appariés (contrôle vs. violation sémantique) est calculé séparément pour chaque segment. L'idée sous-jacente est que si une carte est spécifique à la violation sémantique, alors celle-ci doit être moins présente, voire absente dans les données individuelles des sujets de la condition contrôle. Nous suggérons qu'une ou

plusieurs cartes de surface sont spécifiques à la condition de violation sémantique. Ces cartes doivent alors présenter un indice de pourcentage de maximum de variance expliquée (BEV) plus faible lorsque le "fitting" est effectué dans les données individuelles de la condition contrôle par rapport aux données violation sémantique. Ainsi, si l'on obtient un résultat significatif du test de *Student*, alors un des segments présente une topographie de surface différente entre les deux conditions expérimentales.

La troisième étape consiste en la caractérisation des configurations spatiales pour les segments déterminés comme spécifiques à la condition violation sémantique. Une ANOVA à mesures répétées, similaire à celle utilisée pour l'analyse de tracés, est alors calculée (section 3.1 de ce chapitre). L'amplitude moyenne des PEs (variable dépendante) est alors analysée sur la fenêtre temporelle correspondant au segment ou micro-état. Les corrections de Greenhouse-Geisser ainsi que les procédures de normalisation seront aussi réalisées sur ces données (description dans la section 3.1 de ce chapitre).

La dernière étape de l'analyse de segmentation adaptée consiste à déterminer si un segment ou micro-état a pu être décalé dans le temps. Ainsi, la latence du maximum de GFP est calculée pour les segments spécifiques, pour chaque participant et chaque condition expérimentale. Une ANOVA à mesures répétées est alors effectuée en considérant la latence du maximum de GFP comme variable dépendante et le "type de phrase" (contrôle vs. violation sémantique) comme facteur.

## 4 Résultats

### 4.1 Résultats comportementaux

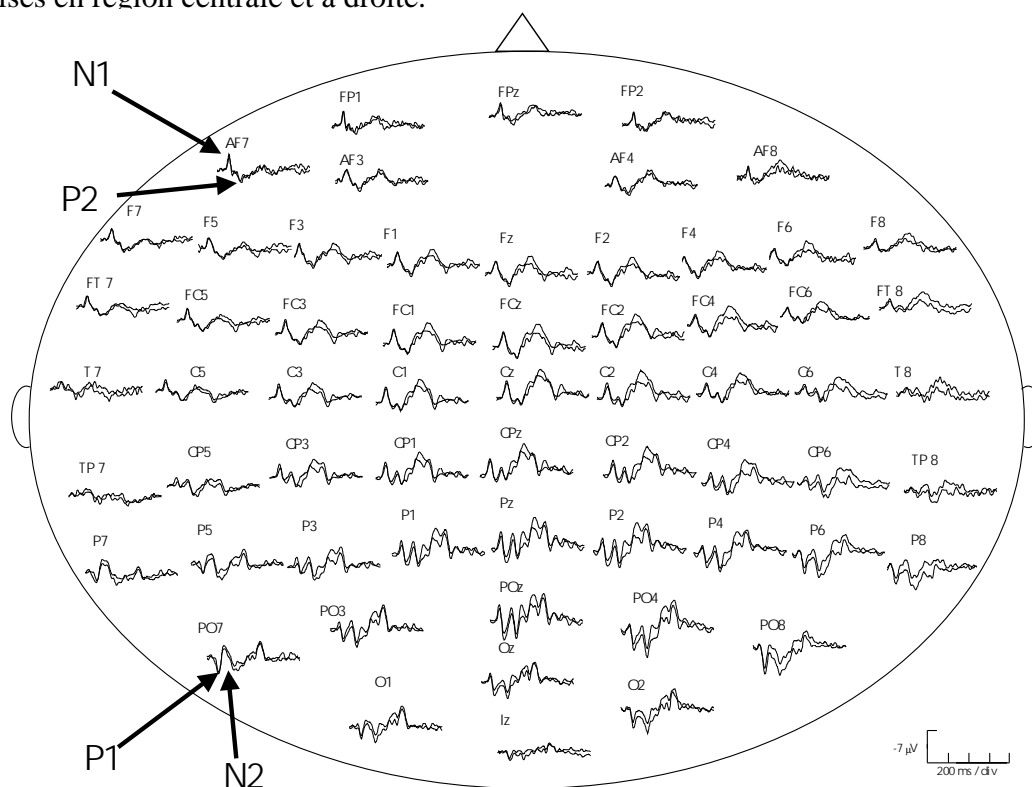
Les moyennes des bonnes réponses ont été soumises à une ANOVA à mesures répétées pour chaque participant, avec comme facteur inter-sujet la "liste" (1 vs. 2) et comme facteur intra-sujet le "type de phrase" (correcte vs. violation sémantique). Les résultats montrent que les phrases présentant une violation sémantique sont mieux détectées (89.1 %) que les phrases contrôles (82.3%,  $F(1,19) = 7.57$ ;  $p < .01$ ). Aucun effet significatif de la "liste" ou d'interaction entre les facteurs "liste" et "type de phrase" n'est observé ( $F < 1$ ): pour l'analyse des PEs les deux listes seront donc regroupées ensembles.



## 4.2 Analyse des tracés

### 4.2.1 Visualisation des tracés

Les grandes moyennes des PEs sont présentées pour les deux conditions expérimentales dans la figure 14. L'analyse visuelle des tracés indique la présence d'une composante précoce P100 avec un maximum en régions postérieures (noté P1 sur la figure 14) et un minimum en régions antérieures (noté N1 sur la figure 14). Par la suite, les PEs s'inversent. En effet, une composante N2 avec un maximum en région antérieure (noté P2 sur la figure 14) et un minimum en région postérieure (noté N2 sur la figure 14) est observée. Ces premières composantes correspondent à une réponse standard lors de stimulation visuelle physique (Picton & Stuss, 1980). Consécutifs à ces composantes précoces, les PEs divergent entre les deux conditions expérimentales: une augmentation de la négativité pour les phrases présentant une violation sémantique est observée (voir figure 14). Cet effet est maximal sur les sites localisés en région centrale et à droite.



**Figure 14:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 61 électrodes enregistrées. Les tracés en gras correspondent aux PEs des violations sémantiques, alors que les tracés en gris clair sont associés à la présentation des phrases contrôles. Les électrodes impaires sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, O pour occipital et I pour inion. Sur cette figure, et sur toutes les

suivantes, la négativité est située vers le haut. Une négativité distribuée sur les sites centraux est présente pour les phrases présentant des violations sémantiques par rapport aux phrases contrôles. Les résultats de l'analyse statistique confirment l'occurrence d'une négativité plus importante sur les sites centro-parietaux droits pour les violations sémantiques que pour les phrases de contrôles.

#### 4.2.2 Analyse topographique par région d'intérêt

Alors que les différences entre les deux conditions ne sont pas significatives entre 0 - 300 ms et 500 - 800 ms ( $F < 1$ ), un résultat significatif est obtenu pour la fenêtre comprise entre 300 - 500 ms. Dans ce rang de latence, les PEs des violations sémantiques sont significativement plus négatifs (- 2.17  $\mu\text{V}$ ) que ceux obtenus pour les phrases contrôles (- 1.14  $\mu\text{V}$ ;  $F(1,19) = 6.05$ ,  $p < .02$ ). L'effet principal interagit avec les deux facteurs spatiaux: "caudalité", données brutes ( $F(2,38) = 2.88$ ,  $p < .08$ ), données normalisées ( $F(2,38) = 3.87$ ,  $p < .03$ ); "latéralité", données brutes ( $F(2,38) = 2.81$ ,  $p < .09$ ), données normalisées ( $F(2,38) = 5.61$ ,  $p < .01$ ). Les comparaisons planifiées (LSD) montrent que les PEs des violations sémantiques sont plus négatifs que ceux enregistrés avec des phrases contrôles dans toutes les régions sur le scalp, exceptée en région antérieure. Les différences sont maximales sur les régions postérieures et droites (voir tableau 4).

	contrôle	violation	différence	valeur de $p$
antérieur	-1.24	-1.74	-0.49	$p < .08$
médian	-1.43	-2.67	-1.24	$p < .01$
postérieur	-0.77	-2.12	-1.35	$p < .01$
gauche	-0.30	-0.94	-0.64	$p < .02$
central	-2.13	-3.10	-0.97	$p < .01$
droite	-1.01	-2.49	-1.48	$p < .01$

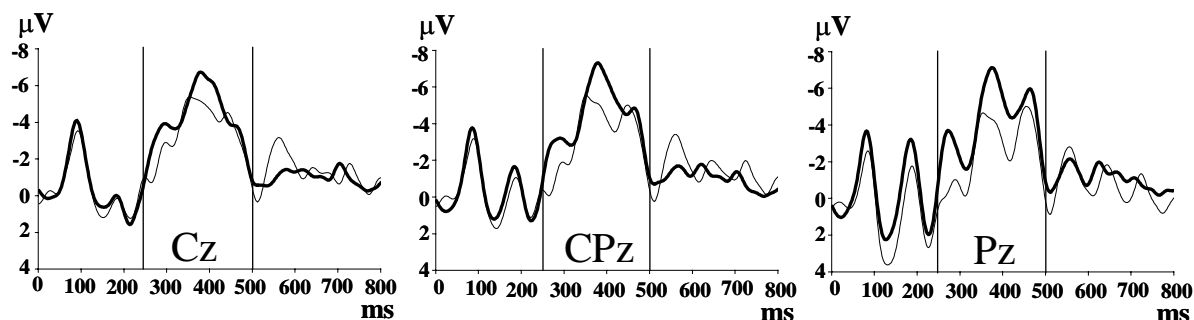
**Tableau 4:** Résultats de la comparaison planifiée (test post-hoc, LSD) pour le rang de latence compris entre 300 et 500 ms pour les différentes régions d'intérêt. Les valeurs indiquées sont en  $\mu\text{V}$ . La colonne "différence" correspond à la valeur du PEs de la condition violation sémantique soustraite à la valeur du PE de la condition contrôle (si la différence est négative, cela indique que les PEs de la condition violation sémantique sont plus négatifs que ceux de la condition contrôle, et vice versa). En gras, sont indiquées les valeurs de  $p$  significatives.

L'analyse des tracés révèle que la composante N400 est restreinte à la fenêtre temporelle comprise entre 300 et 500 ms, et que le maximum de différences entre les violations sémantiques et les phrases contrôles est obtenu dans les régions centro-pariétales. Elle montre aussi que l'effet N400 est latéralisé à droite. La distribution topographique de cet effet N400

correspond précisément aux résultats rapportés avec un nombre conséquent d'électrodes (Curran, Tucker, Kutas, & Posner, 1993; Johnson & Hamm, 2000; Nobre & McCarthy, 1994).

#### 4.2.3 Analyse de latence de pics

Comme la composante N400 est maximale sur les sites centro-pariétaux, nous avons choisi d'analyser la latence des composantes sur 3 électrodes localisées dans ces régions (Cz, CPz, et Pz) et ce, pour les 3 composantes incluses dans le rang de latence compris entre 250 et 500 ms (voir figure 15).



**Figure 15:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour 3 électrodes Cz, CPz, Pz. Les tracés en gras correspondent aux PEs des violations sémantiques, alors que les tracés en gris clair sont associés à la présentation des phrases contrôles. La négativité est située vers le haut. Entre 250 et 500 ms, 3 pics ou composantes sont visibles. Les analyses de latence de pics, ont donc été réalisées sur ces 3 pics (voir tableau 5)

Les résultats des ANOVAs à mesures répétées pour ces 3 composantes et sur les 3 électrodes, ne montrent ni un effet du "type de phrase" (contrôle vs. violation sémantique), ni une interaction entre "type de phrase" x "électrodes" (3) (voir tableau 5).

	contrôle				violation				statistique
	Cz	CPz	Pz	Moyenne	Cz	CPz	Pz	Moyenne	
<b>1ère composante</b>	283.7 (5.8)	276.7 (6.1)	276.7 (5.4)	279.1 (5.7)	282.1 (4.6)	281.2 (4.3)	276.9 (4.1)	280.1 (4.3)	$F(1.19) = 0.04, p > .8$ $F(2.38) = 1.08, p > .3$
<b>2ème composante</b>	369.1 (6.3)	371.9 (5.4)	369.1 (5.1)	370.1 (5.6)	377.7 (6.3)	377.3 (5.6)	373.7 (5.6)	376.3 (5.8)	$F(1.19) = 1.38, p > .2$ $F(2.38) = .59, p > .5$
<b>3ème composante</b>	457.9 (5.9)	453.4 (4.8)	456.2 (3.4)	455.8 (4.7)	460.1 (6.1)	461 (5.2)	461.4 (3.4)	460.8 (4.9)	$F(1.19) = 1.28, p > .2$ $F(2.38) = .52, p > .5$

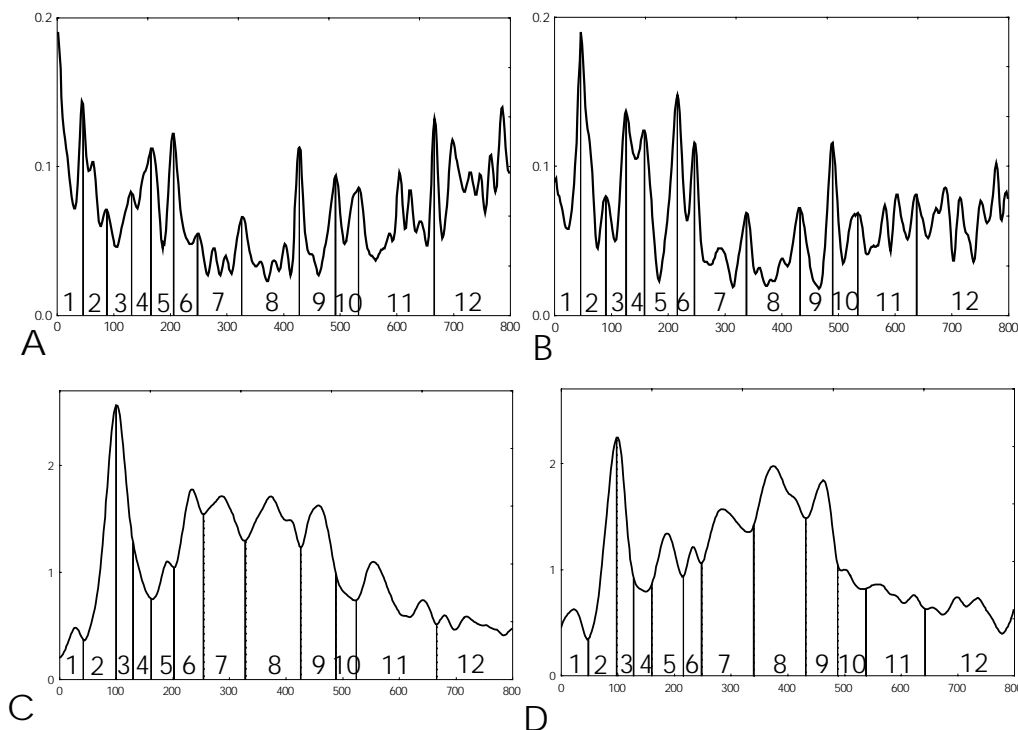
**Tableau 5:** Résultats de l'ANOVA à mesures répétées sur la latence des pics de 3 composantes comprises entre 250 et 500 ms, pour 3 électrodes centrales (Cz, CPZ, PZ). Entre parenthèses sont indiquées l'erreur type. Dans la colonne "statistique" sont rapportées les valeurs de l'ANOVA pour le facteur principal ("type de phrase": contrôle vs. violation sémantique) ainsi que pour l'interaction "type de phrase" x "électrodes".

La première composante présente un maximum de négativité à 279.1 ms et à 280.1 ms; la deuxième composante est enregistrée à 370.1 ms et 376.3 ms et la troisième à 455.8 ms et 460.8 ms pour les phrases contrôles et les violations sémantiques respectivement. Comme le suggèrent les résultats d'autres études (Curran, *et al.*, 1993; Johnson & Hamm, 2000; Nobre & McCarthy, 1994), nos données ne mettent en évidence aucun effet de décalage pour ces 3 composantes (les latences des maximums d'amplitudes pour les phrases contrôles et les violations sémantiques sont identiques).

### 4.3 Analyse spatio-temporelle

#### 4.3.1 Détermination des micro-états

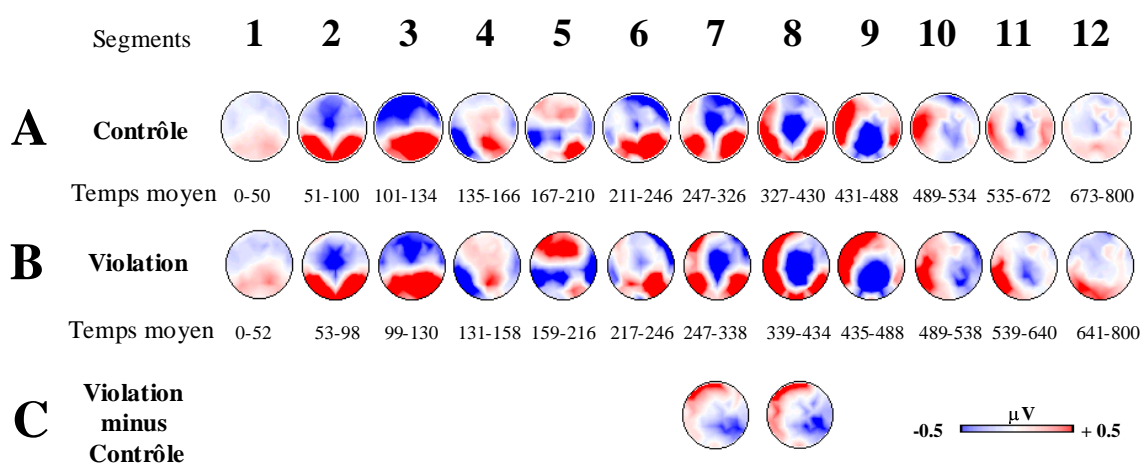
L'analyse de segmentation adaptée est basée sur l'étude des courbes de GD et du GFP des grandes moyennes et est effectuée séparément pour les deux conditions expérimentales. La figure 16 illustre l'évolution des courbes de GD et de GFP au cours du temps pour les deux conditions expérimentales, ainsi que le résultat de l'analyse de segmentation adaptée.



**Figure 16:** (A & B) courbes de *Global Dissimilarity* (GD) pour les grandes moyennes (800 ms d'analyse), pour la condition contrôle (A) et pour la condition où des violations sémantiques sont présentées (B). (C & D), courbes représentant le *Global Field Power* (GFP) pour les grandes moyennes pour la condition contrôle (C) et pour la condition où des violations sémantiques sont présentées (D). Les traits verticaux indiquent la frontière

d'un segment défini par la méthode d'analyse de segmentation adaptée. Les nombres se réfèrent à l'ordre d'apparition des segments ou micro-états dans le temps. La segmentation adaptée révèle que, pour chaque condition expérimentale, 12 configurations différentes et stables sont présentes.

Pour la fenêtre d'analyse comprise entre 0 et 800 ms, nous avons identifié 12 périodes (de moins de 150 ms chacune) représentant les cartes topographiques dominantes dans les grandes moyennes des PEs pour les deux conditions expérimentales (figure 16). Les cartes de surface représentatives de ces 12 micro-états et pour les deux conditions expérimentales sont représentées dans la figure 17.



**Figure 17:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales (A, contrôle, B, violation sémantique) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. La couleur bleue représente des PEs négatifs, alors que les PEs positifs sont représentés en rouge. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, sont représentées les cartes de différence de potentiels (violations sémantiques minus contrôle) pour les segment 7 & 8 où les topographies sont significativement différentes entre les deux conditions expérimentales.

L'analyse visuelle de ces cartes permet de montrer les différentes topographies de composantes connues. Les segments 2 et 3 (entre 51-134 ms pour les phrases correctes et entre 53 et 130 ms pour les violations sémantiques) montrent une positivité (rouge) sur les électrodes localisées en région postérieure et un minimum d'activité (bleue) en région frontale. Ces caractéristiques topographiques et les rangs de latence correspondent à la description de la composante P100 (Brandeis, *et al.*, 1995; Khateb, *et al.*, 1999; Michel, 1997; Michel, Henggeler, & Lehmann, 1992; Michel, *et al.*, 1999; Morand, *et al.*, 2000), ou encore

aux deux pics N1-P1 observés à partir de nos tracés (N1 frontale, et P1 postérieure, figure 14). Le segment 4 présente une négativité sur les électrodes en région postérieure gauche alors que le segment 5 est caractérisé par une négativité bilatérale postérieure. Ces deux derniers segments sont enregistrés entre 135-211 ms pour les phrases contrôles et entre 131-217 ms pour les violations sémantiques. Ces caractéristiques correspondent à la N2-P2 (Brandeis, *et al.*, 1995; Khateb, *et al.*, 1999; Michel, 1997; Michel, *et al.*, 1992; Michel, *et al.*, 1999; Morand, *et al.*, 2000). Dans le rang de latence de la composante N400 (300-500 ms), 4 différents micro-états sont observés. Les segments 6 et 7 montrent une négativité frontale et une positivité sur les régions postérieures. Les segments 8 et 9 présentent des négativités qui semblent plus postérieures que celles enregistrées dans les segments 6 et 7.

#### 4.3.2 Détermination des micro-états spécifiques

Afin de tester si les cartes prédéfinies à partir des grandes moyennes des PEs pour les violations sémantiques sont spécifiques ou non à cette condition expérimentale, nous avons appliqué une procédure de "*fitting*". Ainsi, les cartes de surface correspondant à chaque segment de la condition violation sémantique ont été introduites dans les PEs des données individuelles des deux conditions expérimentales (contrôle et violation sémantique). Cette procédure de "*fitting*" a été réalisée séparément pour chaque segment, et dans une fenêtre temporelle définie sur la base des latences ( $\pm 10$  ms) obtenues sur la grande moyenne (segment 1: 0-60 ms, segment 2: 40-110 ms, segment 3: 90-140 ms, segment 4: 120-170 ms, segment 5: 150-220 ms, segment 6: 200-260 ms, segment 7: 240-350 ms, segment 8: 320-440 ms, segment 9: 420-500 ms, segment 10: 480-550 ms, segment 11: 530-680 ms, segment 12: 630-800 ms). La valeur du BEV (*Best Explained Variance* ou maximum de pourcentage de variance expliquée) pour chaque carte de la condition violation sémantique obtenue dans les données des deux conditions expérimentales de chaque participant est alors comparée statistiquement à l'aide d'un *t*-test à échantillon apparié.

Les résultats montrent que deux cartes présentent des topographies significativement différentes. La carte représentative du segment 7 explique mieux les données des violations sémantiques ( $47.7 \pm 4.9$  %) que les données contrôles ( $38.7 \pm 4.4$  %;  $t = -2.37$ ,  $p < .02$ ). Le même résultat est obtenu pour le segment 8 ( $55.9 \pm 5.3$  % pour les PEs des violations sémantiques,  $43.5 \pm 4.8$  % pour les PEs contrôles;  $t = -2.53$ ;  $p < .02$ ). Aucun des autres

segments ne présente d'effet significatif: les cartes de surface obtenues dans la condition violation sémantique expliquent aussi bien les PEs des données contrôles que ceux obtenus lors de la présentation d'une violation sémantique.

### 4.3.3 Analyse topographique par région d'intérêt

Pour les deux micro-états (segments 7 et 8) définis comme spécifiques au processus d'une violation sémantique, une analyse topographique par région d'intérêt a été réalisée (description dans la section 3.1 de ce chapitre). Deux ANOVAs à mesures répétées séparées ont été effectuées avec l'amplitude moyenne des PEs comme variable dépendante dans les fenêtre temporelle correspondant aux segments (pour le segment 7, contrôle: 247-326 ms, violation sémantique: 247-338 ms; pour le segment 8, contrôle: 327-430 ms, violation sémantique: 339-434 ms). Comme attendu, le fait de considérer des PEs calculés en référence moyenne (la somme de toutes les électrodes pour chaque milliseconde est égale à zéro) ne nous permet pas d'obtenir d'effet significatif du facteur "condition expérimentale" (pour des résultats similaires voir Karniski, *et al.*, 1993). En revanche, des interactions significatives entre le facteur "type de phrase" et les facteurs "spatiaux" sont à noter pour ces deux segments.

Pour le segment 7, l'interaction "type de phrase" x "caudalité" est significative (données brutes:  $F(2,38) = 4.70$ ,  $p < .03$ ; données normalisées:  $F(2,38) = 4.41$ ,  $p < .04$ ) ainsi que "type de phrase" x "latéralité" (données brutes:  $F(2,38) = 2.55$ ,  $p > .11$ ; données normalisées:  $F(2,38) = 4.26$ ,  $p < .02$ ). Les résultats des comparaisons planifiées (LSD, tableau 6) montrent que les PEs des phrases présentant des violations sémantiques sont plus positifs ( $-.17 \mu\text{V}$ ) que ceux des phrases contrôle ( $-.98 \mu\text{V}$ ) sur les régions antérieures ( $p < .01$ ). Sur les régions de l'hémisphère gauche, un effet marginal est obtenu ( $p < .08$ ). Les PEs pour les violations sémantiques situées en région postérieure, montrent une tendance à être plus négatifs que les PEs des phrases contrôles ( $p < .07$ ). Les autres régions d'intérêt ne présentent pas d'effet significatif ( $F < 1$ ).

	contrôle	violation	différence	valeur de $p$
antérieur	-0.98	-0.17	0.81	<b><math>p &lt; .01</math></b>
médian	-0.56	-0.66	-0.10	$p > .1$
postérieur	1.59	0.98	-0.61	$p < .07$
gauche	0.58	1.04	0.45	$p < .08$
central	-1.00	-0.98	0.02	$p > .1$
droite	0.48	0.10	-0.37	$p > .1$

**Tableau 6:** Résultats de la comparaison planifiée (test *post-hoc* LSD) pour le segment 7 (compris entre 247-326 ms pour les phrases de contrôle et entre 247-338 ms pour les violations sémantiques) pour les différentes régions d'intérêt. Les valeurs indiquées sont en  $\mu\text{V}$ . La colonne "différence" correspond à la valeur du PEs de la condition violation sémantique soustraite à la valeur du PEs de la condition contrôle (si la différence est négative, cela indique que les PEs de la condition violation sémantique sont plus négatifs que ceux de la condition contrôle, et vice versa). En gras, sont indiquées les valeurs significatives de  $p$ .

Pour le segment 8, l'interaction "type de phrase" x "caudalité" est significative (données brutes:  $F(2,38) = 4.08$ ,  $p < .04$ ; données normalisées:  $F(2,38) = 5.03$ ,  $p < .01$ ) ainsi que "type de phrase" x "latéralité" (données brutes:  $F(2,38) = 4.82$ ,  $p > .02$ ; données normalisées:  $F(2,38) = 9.08$ ,  $p < .001$ ). Les résultats des comparaisons planifiées (LSD, tableau 7) montrent que, comparés aux PEs des phrases de contrôles, les PEs associés aux violations sémantiques sont plus positifs en région antérieure ( $p < .02$ ) et à gauche ( $p < .02$ ), alors que sur les régions droites, ils sont plus négatifs ( $p < .04$ ) que ceux des phrases contrôles. Aucune différence n'est significative pour les régions médiane, postérieure et centrale ( $F < 1$ ).

	contrôle	violation	différence	valeur de $p$
antérieur	-0.39	0.28	0.67	<b><math>p &lt; .02</math></b>
médian	-0.55	-0.66	-0.11	$p > .1$
postérieur	0.80	0.32	-0.48	$p > .1$
gauche	0.82	1.43	0.61	<b><math>p &lt; .02</math></b>
central	-1.22	-1.20	0.02	$p > .1$
droite	0.26	-0.28	-0.54	<b><math>p &lt; .04</math></b>

**Tableau 7:** Résultats de la comparaison planifiée (test *post-hoc* LSD) pour le segment 8 (compris entre 327-430 ms pour les phrases contrôle et entre 339-434 ms pour les violations sémantiques) pour les différentes régions d'intérêt. Les valeurs indiquées sont en  $\mu\text{V}$ . La colonne "différence" correspond à la valeur du PEs de la condition violation sémantique soustraite à la valeur du PEs de la condition contrôle (si la différence est négative, cela indique que les PEs de la condition violation sémantique sont plus négatifs que ceux de la condition contrôle, et vice versa). En gras, sont indiquées les valeurs significatives de  $p$ .



L'effet N400 semble apparaître dans notre expérience comme un effet biphasique et bipolaire. Biphasique, car deux segments montrent des topographies différentes entre les phrases contrôles et les violations sémantiques. Bipolaire, car pour ces deux segments, un pôle positif et un pôle négatif sont mis en évidence.

#### 4.3.4 Evolution temporelle des micro-états

Nous avons vu que la segmentation adaptée a permis de diviser la fenêtre temporelle comprise entre 300 et 500 ms en 4 différents segments. Bien que seuls deux d'entre eux présentent des topographies spécifiques aux violations sémantiques, le temps d'apparition de ces quatre segments a été l'objet d'une analyse. Le maximum de GFP est calculé pour chaque condition expérimentale et pour chaque participant. Les ANOVAs à mesures répétées séparées pour chaque segment ne montrent aucun effet du facteur "type de phrase" (voir tableau 8).

	Contrôle	Violation	Statistique
Segment 6	227.2 (5.5)	227.9 (4.9)	$F(1,19) = 0.02, p > .9$
Segment 7	289.7 (5.6)	283.4 (5.3)	$F(1,19) = 1.14, p > .3$
Segment 8	382.3 (5.4)	374.6 (2.6)	$F(1,19) = 1.90, p > .1$
Segment 9	464.4 (5.1)	460.2 (4.4)	$F(1,19) = 2.41, p > .1$

**Tableau 8:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments compris entre 200 et 500 ms. Les valeurs indiquées sont en ms avec, entre parenthèses, l'erreur type. Aucune différence de latence entre les deux conditions expérimentales n'est significative.

## 5 Discussion

Les résultats obtenus à partir d'une référence mastoïde gauche avec une analyse de tracés répliquent ceux de la littérature. Une composante N400 est évoquée par la présentation d'une violation sémantique, comparativement à une phrase correcte, dans une fenêtre temporelle comprise entre 300 et 500 ms. La topographie de cet effet (différence entre violation sémantique et phrase contrôle) est maximale sur les électrodes situées sur les régions centro-pariétales et à droite.

## 5.1 La N400 composée d'un ou de deux processus ?

Alors que l'analyse classique sur les tracés met en évidence un effet N400 comme un processus unique, l'analyse spatio-temporelle a révélé une topographie beaucoup plus complexe. Deux segments présentent des topographies différentes entre les violations sémantiques et les phrases contrôles: le segment 7 (247-338 ms pour les violations et 247-326 ms pour les contrôles) et le segment 8 (339-434 ms pour les violations et 327-430 ms pour les contrôles). Pour le segment 7, les violations sémantiques induisent l'occurrence d'une positivité bilatérale antérieure, et d'une négativité postérieure. Au niveau du segment 8, les violations sémantiques sont associées à une positivité significativement plus importante que celle des phrases contrôles, sur les régions antérieures et sur l'hémisphère gauche, et accompagnées d'une négativité sur tout l'hémisphère droit.

Etant donné la présence de deux micro-états significativement différents lors de la présentation de violations sémantiques, l'on doit se poser la question de la fonctionnalité de ces deux segments. Caractérisent-ils la même étape dans le cheminement du traitement de l'information sémantique ou bien peut-on les relier à différentes étapes? Peu d'études ont analysé la composante N400 à l'aide de la segmentation (adaptée ou non) chez des participants sains. A notre connaissance, Brandeis, *et al.* (1995) et Karniski, *et al.* (1993) sont les seuls, et suggèrent deux segments ou micro-états distincts pour l'effet N400. Chaque micro-état de ces trois études (Brandeis, *et al.*, 1995; Karniski, *et al.*, 1993; et notre étude) expose des topographies de surface différentes entre violations sémantiques et phrase contrôles. En accord avec le concept de micro-état fonctionnel (Lehmann, 1987; Lehmann & Skrandeis, 1980), des configurations différentes sur le scalp indiquent que des groupes de neurones différents ont été activés à ce moment précis. Le fait de trouver deux micro-états différents suggère donc que le processus des violations sémantiques implique l'activation de deux étapes différentes. En d'autres termes, la composante N400 reflèterait au moins deux processus.

## 5.2 Topographie des deux segments

Dans la présente étude, nous pouvons noter que les distributions topographiques sur le scalp de l'effet N400 sont différentes de celles obtenues par Brandeis, *et al.* (1995) et Karniski, *et al.* (1993): notre premier micro-état est symétrique, alors que le second est latéralisé. Les études de Brandeis, *et al.* (1995) et de Karniski, *et al.* (1993) indiquent un premier segment latéralisé, puis un deuxième présentant une certaine symétrie. Cette différence peut s'expliquer par les designs expérimentaux utilisés que nous allons expliciter maintenant.

Premièrement, les deux précédentes études utilisent un paradigme similaire: des phrases de 4 mots sont présentées soit visuellement (Brandeis, *et al.*, 1995) soit auditivement (Karniski, *et al.*, 1993), chaque violation sémantique étant positionnée en fin de phrase sur un nom de couleur. Hagoort & Brown (2000) montrent des différences significatives dans la latéralisation de la composante N400 en fonction de la position des violations sémantiques au sein de la phrase. En position finale, la N250 et la N400 présentent des topographies équivalentes avec un maximum de différence sur les sites centro-pariétaux, et légèrement latéralisée sur l'hémisphère gauche. Pour les violations en milieu de phrases, les N250 et N400 sont caractérisées par des négativités postérieures, avec une tendance à une latéralisation sur l'hémisphère droit. Bien qu'Hagoort & Brown (2000) n'aient enregistré les PEs qu'à partir de 7 électrodes, les différences de latéralisation peuvent être pertinentes pour notre réflexion. Ces auteurs n'ont pu dissocier les topographies des composantes N250 et N400, mais leur résultats suggèrent que la position de la violation est un facteur primordial lors de la caractérisation de la latéralisation de ces négativités.

Deuxièmement, dans les études de Brandeis, *et al.* (1995) et de Karniski, *et al.* (1993), toutes les phrases finissent par un mot d'une même catégorie sémantique: un nom de couleur. Les phrases utilisées étaient de 4 mots, donc très courtes, et induisaient par conséquent une forte attente pour la présentation du dernier mot (forte *Cloze Probability*, CP). Il était alors très facile, pour les participants, de détecter la présentation des violations sémantiques. A l'opposé, nous avons enregistré, dans notre étude, des violations sémantiques situées en milieu de phrase portant sur les verbes. Même si la structure syntaxique était constante (adverbe ou préposition, sujet, verbe, complément), notre design expérimental impliquait une recherche dans le lexique mental plus importante de la part du participant. Par ailleurs, nos violations sémantiques portent sur les propriétés sémantiques des verbes, tandis que, dans leur

cas, il s'agit de propriétés sémantiques des noms. Différentes études suggèrent, d'une part, que les PEs des noms et verbes sont différents (Koenig & Lehmann, 1996; Pulvermüller, Lutzenberger, & Preissl, 1999; Pulvermüller, Mohr, & Schleicher, 1999) mais d'autre part, que la composante N400 est différemment latéralisée si elle est enregistrée à partir de violations sémantiques portant sur des noms ou des verbes. En effet, Federmeier, Segal, Lombrozo, & Kutas (2000) et Rösler, Streb, & Haan (2001) indiquent que la composante N400 (entre 200 et 600 ms) associée aux verbes est latéralisée sur l'hémisphère gauche alors que celle obtenue pour les noms apparaît comme bilatérale.

Ces différents points énumérés pourraient expliquer les variabilités de latéralisation de nos deux segments par rapport aux résultats de Brandeis, *et al.* (1995) et de Karniski, *et al.* (1993).

### 5.3 Interprétation fonctionnelle de ces segments

Une question des plus intéressante est de déterminer la fonctionnalité de ces deux segments.

Brandeis, *et al.* (1995) suggèrent que le micro-état présentant une symétrie (entre 364-400 ms) correspond aux caractéristiques de la composante N400 et le nomme donc le segment "N400". En revanche, Karniski, *et al.* (1993) détermine le premier segment (entre 196 et 287 ms) latéralisé, comme représentatif de la N400. Nous pensons, que dans notre étude, le micro-état latéralisé et enregistré entre 339 et 434 ms correspond aux caractéristiques topographiques de la composante N400. Si ce micro-état reproduit les caractéristiques d'une N400, alors ce segment peut refléter le moment où les informations d'un mot sont intégrées dans le contexte précédant afin de construire la représentation de la phrase (Kutas & Hillyard, 1983; Kutas & van Petten, 1994). La fonctionnalité du premier segment reste maintenant à déterminer. Karniski, *et al.* (1993) interprètent le second segment de différence, nommé "slow wave", comme reflétant une analyse plus profonde ou une re-analyse due à la lecture de la violation sémantique. Alors que Brandeis, *et al.* (1995) ne clarifient pas la fonctionnalité de leur segment "pré-N400", notre hypothèse est que notre micro-état 7 (247-338 ms) reflète l'activation des informations spécifiques à chaque mot. Dans le cas présent, les informations activées par un verbe sont multiples et sont regroupées sous l'appellation d'informations thématiques. Ce type d'informations spécifie l'utilisation du verbe et les contraintes qui lui sont imposées en définissant le nombre d'argument obligatoire. Par exemple si le verbe a besoin ou non d'un complément d'objet (verbe transitif/intransitif), ou bien s'il doit être

utilisé avec un sujet animé (restriction de sélection). Lors de la reconnaissance du verbe, ces informations peuvent être disponibles et activées en mémoire. En dépit du fait que le débat reste ouvert quant à la nature des informations stockées en mémoire (acoustique, phonologique, syntaxique, sémantique ou autres), ce type d'informations doit être récupéré.

Connolly & Phillips (1994) et Hagoort & Brown (2000) apportent une réponse possible sur la nature des informations activées. Grâce à des stimulations auditives et à la manipulation de phonèmes initiaux de mots, ils décrivent un effet précoce (entre 200 et 300 ms) sensible aux manipulations acoustico-phonologiques: un effet N200 est enregistré lorsque le mot présenté ne partage pas les phonèmes initiaux du mot cible. Federmeier & Kutas (1999) interprètent différemment l'effet précoce. Ces auteurs comparent les PEs de deux types de violations: intra-catégorie sémantique (le mot présenté appartient à la même catégorie sémantique que le mot approprié) inter-catégorie sémantique (le mot présenté n'appartient pas à la même catégorie sémantique que le mot approprié). Les violations sont présentées en fin de phrases de haute ou de basse *Cloze Probability* (une CP élevée indique que le mot critique est facilement détectable). Pour les phrases de basse CP, les deux types de violations, intra et inter, induisent l'occurrence d'une large et similaire composante N400. En revanche, pour les phrases de haute CP, les violations inter-catégorie sémantique suscitent l'occurrence d'une composante plus large par rapport aux violations intra-catégories. En manipulant les contraintes contextuelles, les auteurs ont mis en évidence la sensibilité de la composante N400 face aux informations activées par le contexte: plus le contexte est contraignant et plus la composante N400 aura une amplitude importante et une période d'enregistrement étendue.

Federmeier & Kutas (1999) suggèrent que cette composante précoce reflète l'activation d'informations apportées par le contexte linguistique. En revanche, Connolly & Phillips (1994) et Hagoort & Brown (2000) indiquent que cet effet précoce peut refléter l'activation des informations acoustico-phonologiques. Comme dans notre étude les phrases étaient présentées visuellement, l'activation des informations acoustico-phonologiques ne semble pas pertinente pour la tâche requise par les participants. Ainsi, l'hypothèse selon laquelle le segment 7, que nous assimilons à la composante précoce N200/PMN (topographie et rang de latence similaire), reflète l'activation des informations acoustico-phonologiques semble invalidée. Alors que notre paradigme expérimental ne nous permet pas de distinguer quels types d'informations sont activés pour susciter l'occurrence de notre premier segment, nous suggérons que ce segment 7 peut refléter l'activation des informations thématiques du verbe.

Ainsi, nous l'avons suggéré, le segment 8, refléterait les processus d'intégration du mot au sein du contexte linguistique présenté. La divergence entre l'activation des informations spécifiques au mot attendu et celles du mot présenté, suscite l'occurrence de ce deuxième micro-état. Mais il est certain que de plus amples investigations dans le domaine sont nécessaires afin de valider nos interprétations.

#### 5.4 Amplitude de l'effet N400

Dans cette étude, l'amplitude de l'effet N400 (différence entre phrase contrôle et violations sémantiques) est maximale pour le segment 7 (approximativement 0.8  $\mu\text{V}$ ) et réduite pour le segment 8 (approximativement 0.5  $\mu\text{V}$ ). Trouver un effet N400 réduit à moins de 1  $\mu\text{V}$  n'est pas surprenant si l'on se réfère à notre paradigme expérimental. Comme les violations sémantiques sont positionnées au niveau du verbe, les participants avaient du mal à déterminer quel verbe pouvait être présenté dans la condition contrôle. Tous les mots de classe ouverte (noms, verbes, adjectifs) suscitent l'occurrence d'une composante N400 (Kutas & van Petten, 1988), mais l'amplitude de celle-ci dépend de la position du mot au sein de la phrase: la composante N400 est maximale en fin par rapport en milieu de phrase (Kutas & Hillyard, 1984; Kutas & van Petten, 1988). Ce résultat suggère que les fluctuations de l'amplitude de la composante N400 reflètent les modifications dans les processus linguistiques lors de la construction de la représentation de la phrase. L'amplitude de la composante N400 diminue au fur et à mesure que le participant accède aux différents mots constituant la phrase: les attentes contextuelles sont d'autant plus importantes en fin qu'au début de chaque phrase. Puisque les contraintes contextuelles sont moindres en milieu de phrase, alors intégrer un nouveau mot qui n'est pas approprié sémantiquement doit être moins difficile par rapport à si l'on se situe en fin de phrase où les contraintes contextuelles sont maximales. Si l'on considère que la composante N400 reflète la difficulté des processus d'intégration sémantique, alors celle-ci doit être moins importante lorsque les participants essaient d'intégrer une violation sémantique en milieu de phrase par rapport à celles présentées en fin de phrase (Hagoort & Brown, 2000; Kutas & Hillyard, 1984; Kutas & van Petten, 1988). De plus, nous avons utilisé des phrases dans lesquelles les participants ne pouvaient pas, ou très peu, anticiper la présentation du mot critique, paradigme qui est aussi une source de réduction de l'amplitude de la composante N400. En effet, la présentation d'une violation sémantique n'est pas une condition nécessaire à l'occurrence d'une

composante N400: si l'on compare l'amplitude des négativités enregistrées lorsqu'un mot est très peu attendu (CP basse) par rapport au contexte précédent (large N400) et lorsqu'une violation sémantique est présente (large N400), la différence sera moindre. Par exemple, Brandeis, *et al.* (1995) utilisent une seule catégorie sémantique en fin de phrase (couleur), permettant aux participants de concentrer leur attention sur ces quelques mots uniquement. La conséquence de ce paradigme est d'activer rapidement l'entrée lexicale appropriée car le champ sémantique des couleurs est déjà actif. L'introduction de violations sémantiques lorsque le participant doit effectuer une recherche lexicale "limitée", résulte en une plus large amplitude de l'effet N400.

Dans notre étude, la position au sein de la phrase de la violation sémantique, ainsi que la présentation de phrases de structure syntaxiques compliquées, peuvent avoir contribué à réduire l'amplitude de l'effet N400.

## 5.5 Contributions méthodologiques

### 5.5.1 Référence moyenne

Les deux analyses utilisées dans cette étude - les tracés et la segmentation adaptée - ont mis en évidence des caractéristiques différentes de la composante N400. Alors que l'analyse de tracés considère la composante N400 comme un processus unique, la segmentation adaptée présente cette composante comme constituée de deux étapes séparées. Cependant, une comparaison des deux méthodes n'est pas si simple. La plupart des effets décrits dans la littérature par des analyses de tracés, utilisent des références mastoïdes gauches, ou encore les oreilles. Puisque ces références ne sont pas équivalentes à zéro au cours du temps, leur activité dépend donc de l'activité cérébrale (Nunez, 1981; Nunez, *et al.*, 1991). La morphologie des composantes et leur distribution spatiale ne représentent donc pas l'activité aux électrodes actives. Pour caractériser l'activité électrique d'une électrode sur le scalp, il semble nécessaire d'utiliser une référence moyenne (Johnson & Hamm, 2000; Lehmann & Skrandeis, 1980; Offner, 1950; Picton, *et al.*, 2000). Dans la présente étude, et avec une référence moyenne, l'effet N400 apparaît non pas comme une déflexion négative (comme souvent décrite dans la littérature; pour un exemple voir Kutas & Hillyard, 1984), mais

comme un effet bipolaire: une déflexion négative sur les sites postérieurs et une déflexion positive sur les sites frontaux (symétrique pour le segment 7, et latéralisée pour le segment 8).

Un nombre limité d'études analysent le pôle positif de l'effet N400 (Brandeis, *et al.*, 1995; Curran, *et al.*, 1993; Johnson & Hamm, 2000). Focaliser les analyses sur le pôle négatif de la composante N400 uniquement a pu incorrectement suggérer que les générateurs de cette composante sont localisés dans les régions postérieures droites du cerveau. Comme le montrent la majorité des études portant sur la localisation des générateurs de la composante N400, les résultats convergent vers une même localisation - grâce à des enregistrements intracrâniens (McCarthy & Wood, 1985), ou bien grâce à la technique du MEG (Schmidt, Arthur, Kutas, & Flynn, 1989) - les sources des générateurs de la composante N400 sont localisées dans les régions médio-temporales gauches du cerveau. Ainsi, l'orientation des deux pôles (positifs et négatifs) de l'effet N400 suggère une telle localisation de générateurs, tout au moins pour le segment présentant une latéralisation, soit le segment 8 (339-434 ms pour les violations sémantiques).

### 5.5.2 La segmentation

La méthode d'analyse de tracés nous a permis d'indiquer qu'une violation sémantique produit des variations de l'activité électrique dans la fenêtre temporelle comprise entre 300 et 500 ms. En revanche, l'utilisation d'une méthode d'analyse basée sur la segmentation des PEs (adaptée), nous précise le déroulement temporel de cet effet, sans devoir recourir à des fenêtres temporelles arbitraires. Toutes les fenêtres utilisées lors de notre analyse sont définies à partir des modifications de topographie des PEs. Les deux segments définis comme spécifiques au processus des violations sémantiques, et présentant des topographies différentes, suggèrent que les générateurs de ces configurations électriques de surface sont séparés ou, en tout cas, ne se chevauchent pas entièrement. Avec une analyse sur la fenêtre temporelle large de 300 à 500 ms, les chercheurs ont conclu que l'effet N400 reflète un même et unique processus. Toutefois, une distinction entre effet précoce et tardif dans la composante N400 a récemment émergé de la littérature (Connolly & Phillips, 1994; Hagoort & Brown, 2000; Kutas & Federmeier, 2000), mais celle-ci reste limitée aux paradigmes basés sur la présentation auditive. Comme Brandeis, *et al.* (1995), nous rapportons avec une analyse de segmentation adaptée une dissociation entre effet précoce et tardif lors du traitement visuel d'une violation sémantique. Alors qu'aucune étude précédente n'a permis de mettre en valeur de tels résultats



avec une analyse de tracés et en modalité visuelle, la segmentation nous semble un outil permettant d'affiner le déroulement temporel et la dynamique des effets connus.

Enfin, la méthode de segmentation adaptée utilise toutes les électrodes enregistrées à chaque étape de l'analyse et, par conséquent, exploite la richesse des données électroencéphalographiques d'un point de vue spatial. De plus, contrairement à l'analyse de tracés, cette technique d'analyse ne repose pas sur des à priori. Elle permet donc de mettre en évidence de nouvelles composantes, ou encore de préciser le déroulement temporel de composantes connues comme, par exemple, la composante N400.

## 6 Conclusions

Les enregistrements PEs à partir d'un nombre conséquent de canaux, permettent le traitement des données comme des distributions de potentiels de surface en 2 dimensions et non plus comme des séries de points univariés (analyses de tracés). Cette technique d'imagerie fonctionnelle 1) possède une excellente résolution temporelle permettant d'étudier les processus cognitifs, 2) permet de caractériser les patterns d'activation de l'activité cérébrale et 3) est sensible aux changements de l'activité cognitive mise en place par le participant lors d'une tâche précise. Du fait de ces avantages, nous allons appliquer cette technique d'imagerie et d'analyse à nos données. Dans un souci de comparaison avec la littérature, les analyses de tracés seront systématiquement réalisées en parallèle sur les données.

# CHAPITRE 4

## Effet de Classe et Propriétés Lexicales non Contrôlées

---

### *Résumé*

Dans ce chapitre sont évalués simultanément les effets de la classe de mot ainsi que des propriétés lexicales (fréquence et longueur). Les PEs sont enregistrés à partir de 61 électrodes alors que les participants lisent et jugent l'acceptabilité des phrases (50 % violations de restriction de sélection, 50 % correctes). Les PEs ont été moyennés sur les phrases correctes uniquement pour les deux classes de mots. Les mots de classe fermée produisent une composante P2 réduite sur les régions antérieures. Une composante N280 maximale sur les régions antérieures est enregistrée pour les deux classes de mots, mais présente des différences de latéralisation: bilatérale pour CF et droite pour CO. Une composante N400 est enregistrée pour les mots de classe ouverte sur les régions postérieures droites. Enfin une composante LAN, maximale sur les sites antérieurs, est mise en évidence pour les mots CF. De plus, nos résultats indiquent que la latence de la composante N280 est plus précoce pour les mots de classe fermée. Des analyses de régression suggèrent que les propriétés lexicales des mots (fréquence et longueur) peuvent prédire la latence de la composante N280. Nos résultats indiquent des différences électrophysiologiques robustes entre les deux classes de mots. Alors que les propriétés lexicales des mots induisent certaines différences électrophysiologiques sur le scalp, tous les effets enregistrés n'ont pu être expliqués par les différences de propriétés lexicales seules. Par conséquent, ces résultats suggèrent que la reconnaissance des mots de classe fermée dans un contexte linguistique, n'active pas les mêmes processus que ceux pour les mots de classe ouverte.

## 1 Introduction

Lorsque l'on compare les PEs associés aux mots de classe ouverte et fermée, la plupart des expériences mènent à la conclusion que les différences s'observent sur plusieurs composantes: la P2, N280, N400 et la LAN. La question non résolue est de déterminer si ces effets rapportés sur le scalp sont caractéristiques d'une différence dans la reconnaissance des deux classes de mots, ou bien s'ils témoignent plutôt de différences au niveau des propriétés lexicales (fréquence, longueur). Alors qu'une seule étude (Neville, *et al.*, 1992) suggère que la composante N280 est réellement spécifique aux mots de classe fermée, toutes les autres expériences ont enregistré des négativités (plus ou moins précoces) sur les sites antérieurs pour les deux classes de mots (voir tableau 2). Parmi les études ayant observé des négativités pour les deux classes de mots, deux courants se sont développés. Le premier rapporte des topographies différentes pour les négativités précoces des deux classes de mots (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; Nobre & McCarthy, 1994; Pulvermüller, *et al.*, 1995) et conclut que cette composante négative est un indicateur précoce spécifique pour les mots de la classe fermée. A l'opposé, le second indique des topographies similaires pour les deux classes de mots (Brown, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002) et suggère un accès au lexique identique pour les deux classes de mots.

Depuis 20 ans, le débat reste ouvert et aucun consensus n'a pu aboutir. Nous proposons donc de poursuivre les recherches sur ces questions afin de caractériser les effets de classe de mots sur le scalp. Notre premier objectif est d'analyser précisément la distinction entre classes de mots à l'aide des PEs. La caractérisation du décours temporel et la répartition spatiale sur le scalp des différentes composantes associées à chacune des classes de mots ont été effectuées grâce à l'enregistrement simultané de 61 canaux<sup>12</sup>. Par ailleurs, nous avons analysé les PEs obtenus au moyen de deux méthodes: celle des tracés et celle de la segmentation adaptée. Cette double analyse est décrite dans le chapitre 3 (section 3). A partir des données récoltées dans le chapitre 3 sur le paradigme de N400, une analyse complémentaire a été réalisée uniquement sur les phrases correctes: les mots ont été regroupés selon qu'ils appartiennent à l'une ou l'autre des classes, et deux PEs différents ont été obtenus.

---

<sup>12</sup> Le nombre moyen de sites d'enregistrements pour les recherches portant sur la distinction entre classe de mots est actuellement de 19.5 électrodes (voir tableau 2, chapitre 2, section 3.2).

L'objectif de cette recherche est de répliquer les résultats obtenus dans les précédentes études n'ayant pas contrôlé les propriétés lexicales des mots, à savoir, leur fréquence et leur longueur (Brown, *et al.*, 1999; Garnsey, 1985; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Kutas & Hillyard, 1983; Kutas, *et al.*, 1988; Münte, *et al.*, 2001; Neville, *et al.*, 1993; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997; van Petten & Kutas, 1991). Comme dans ces études, le matériel de la présente recherche est constitué de mots des deux classes, les mots de la classe fermée étant à la fois plus courts et plus fréquents que les mots de la classe ouverte. Nous faisons l'hypothèse que les mots de classe fermée vont présenter des topographies différentes sur le scalp par rapport aux mots de classe ouverte, et cela, sur une longue période temporelle (incluant une fenêtre précoce et tardive). Une telle prédiction est attendue si des étapes de traitement spécifique sont mises en place pour les deux classes de mots (Garnsey, 1985; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; Kutas & Hillyard, 1983; Kutas, *et al.*, 1988; Nobre & McCarthy, 1994; Pulvermüller, *et al.*, 1995; van Petten & Kutas, 1991). De plus, nous émettons la seconde hypothèse que la latence de la composante N280 (fenêtre précoce) sera plus courte pour les mots de classe fermée que celle enregistrée à partir des mots de classe ouverte (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Nobre & McCarthy, 1994; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997).

Depuis peu, un débat est apparu (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997): la composante N280 pourrait ne pas être uniforme (topographie et latence) à l'intérieur de chaque classe de mots, et refléterait plutôt des différences dans les propriétés lexicales plutôt qu'une différence effective entre ces deux classes de mots. Ainsi, selon cette hypothèse, les caractéristiques de la composante N280 devraient varier en continuum depuis les mots les plus fréquents et les plus courts (tels que les articles et les pronoms) avec en position intermédiaire les mots de moyenne fréquence (par exemple les verbes), puis pour finir avec les mots les moins fréquents et les plus longs (par exemple les noms). Dans cette perspective, certaines recherches (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) ont divisé les deux classes de mots en différentes catégories grammaticales (auxiliaires, noms, verbes, adjectifs, articles, etc.) de différentes fréquences et longueurs. Alors que les analyses sur la topographie ne proposent pas de résultats consistants, deux courants émergent quant aux résultats obtenus sur la latence

de la composante N280. En effet, certaines recherches (King & Kutas, 1998; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) indiquent que la latence de la composante N280 est corrélée avec la fréquence d'occurrence et/ou la longueur: plus la fréquence est élevée (par conséquent, plus le mot est court, par exemple les articles) et plus la latence de la composante N280 est précoce. En revanche, les résultats d'autres groupes (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999) suggèrent que la latence de la composante N280 n'est pas corrélée avec les propriétés lexicales, et reflète donc des différences effectives entre classes de mots.

Face à cette contradiction dans les résultats, notre second objectif était de vérifier si les caractéristiques de la composante N280 (distribution de la négativité sur le scalp et la latence) sont uniformes à l'intérieur de chaque classe de mots. Ainsi, chaque classe a été divisée en différentes catégories grammaticales (classe ouverte: noms, verbes, adverbes; classe fermée: prépositions, quantificateurs, non-quantificateurs). Les analyses vont spécifiquement porter sur le rang de latence de la composante N280. D'une part, les analyses de topographie de la composante N280 pourront nous renseigner sur des différences de traitement entre les catégories grammaticales comme le suggère certains (Cann, 2000; Friederici, 1985; Friederici & Schoenle, 1980; Froud, 2001). D'autre part, des analyses de régression, similaires à celles décrites dans la littérature (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) vont être utilisées afin de déterminer si les propriétés lexicales peuvent prédire la latence de la composante N280. Nous avançons l'hypothèse que la latence de la composante N280 sera corrélée avec les propriétés lexicales des mots des deux classes.

## 2 Méthode

### 2.1 Sujets

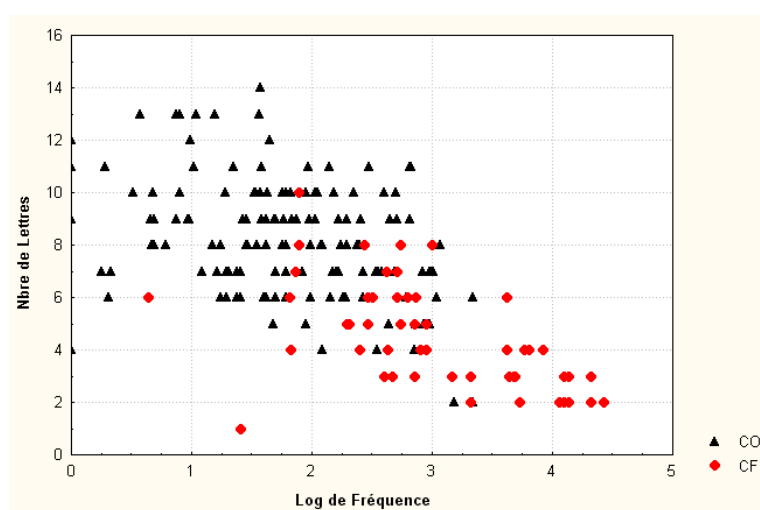
Les participants sont les mêmes que ceux du chapitre 3 (voir section 2.1 chapitre 3 pour un descriptif détaillé).

### 2.2 Stimuli

Les phrases utilisées sont les 160 phrases décrites dans la section 2.2 du chapitre 3, et sont présentées dans l'annexe 1. Les mots de classe ouverte (CO) et fermée (CF) ont été pris en compte uniquement dans les phrases correctes (Devant nous les poules mangeront des graines). Afin de limiter les effets de position dans la phrase, tous les mots utilisés étaient positionnés dans la première moitié des phrases, soit avant, ou équivalent à la présentation du verbe. Ainsi, pour chaque classe de mots, 120 items ont été retenus. Pour les mots de classe ouverte: 40 verbes, 40 noms, 40 adverbes. Pour les mots de classe fermée: 40 prépositions, 40 quantificateurs, 40 non-quantificateurs.

Dans un souci de comparaison avec la littérature précédente (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Kutas & Hillyard, 1983; Kutas, *et al.*, 1988; Neville, *et al.*, 1993; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997; van Petten & Kutas, 1991), les répétitions n'ont pas été contrôlées et, par conséquent, certains mots ont parfois été utilisés plusieurs fois. Les répétitions étaient plus nombreuses pour les mots de classe fermée (85 %) par rapport aux mots de classe ouverte (29.7 %): jusqu'à 97.5 % de répétition pour les quantificateurs (tels que: tous, aucun, beaucoup, peu, quelque), 85 % pour les non-quantificateurs (tels que: le, la, mon, ce), et 72.5 % pour les prépositions (tels que: dans, devant, durant, parmi, pour). Pour la classe ouverte, 70.5 % des adverbes (tels que: parfois, jadis, demain, bientôt), 17.5 % des noms (tel que: enfant, étudiant, voyageur) et 0 % des verbes (tel que: lira, quitter, offrait, passe) étaient répétés. La liste complète des mots utilisés est rapportée dans l'annexe 2.

Grâce à la base de donnée BRULEX (Content, *et al.*, 1990), la fréquence d'utilisation dans la langue française, ainsi que la longueur (nombre de lettres) de chaque mot ont été déterminées. La fréquence d'utilisation observée dans BRULEX pour 100 millions de mots a été ensuite rapportée à un million, puis transformée en log de fréquence, indice plus usité dans la littérature. Une ANOVA séparée avec ces deux variables indique que les mots de classe fermée sont plus courts ( $F(1,278) = 210.21$ ;  $p < .000$  ; avec 4.4 lettres pour CF, et 8.1 pour CO) et plus fréquents ( $F(1,278) = 184.04$ ;  $p < .000$  ; avec 3.15 log de fréquence pour CF et 1.85 pour CO) que les mots de classe ouverte. La répartition des fréquences et longueurs en fonction de la classe de mot est illustrée dans la figure 18.



**Figure 18 :** Répartition des stimuli du chapitre 4 entre mots de classe fermée (CF, cercles rouges) et mots de classe ouverte (CO, triangles noirs). Comme l'analyse statistique le montre, les mots de classe fermée sont plus fréquents et plus courts que ceux de la classe ouverte.

## 2.3 Procédure

La procédure utilisée est celle décrite dans le chapitre 3 (voir section 2.3).

## 2.4 Enregistrement de l'EEG

Les caractéristiques de l'enregistrement EEG sont décrites dans la section 2.4 du chapitre 3.

## 3 Résultats

### 3.1 PEs et classe de mots

Les analyses statistiques réalisées sont identiques à celles présentées dans la section 3 du chapitre 3. Dans un premier temps, nous allons décrire les résultats de l'analyse des tracés puis, ceux obtenus à partir de la segmentation adaptée.

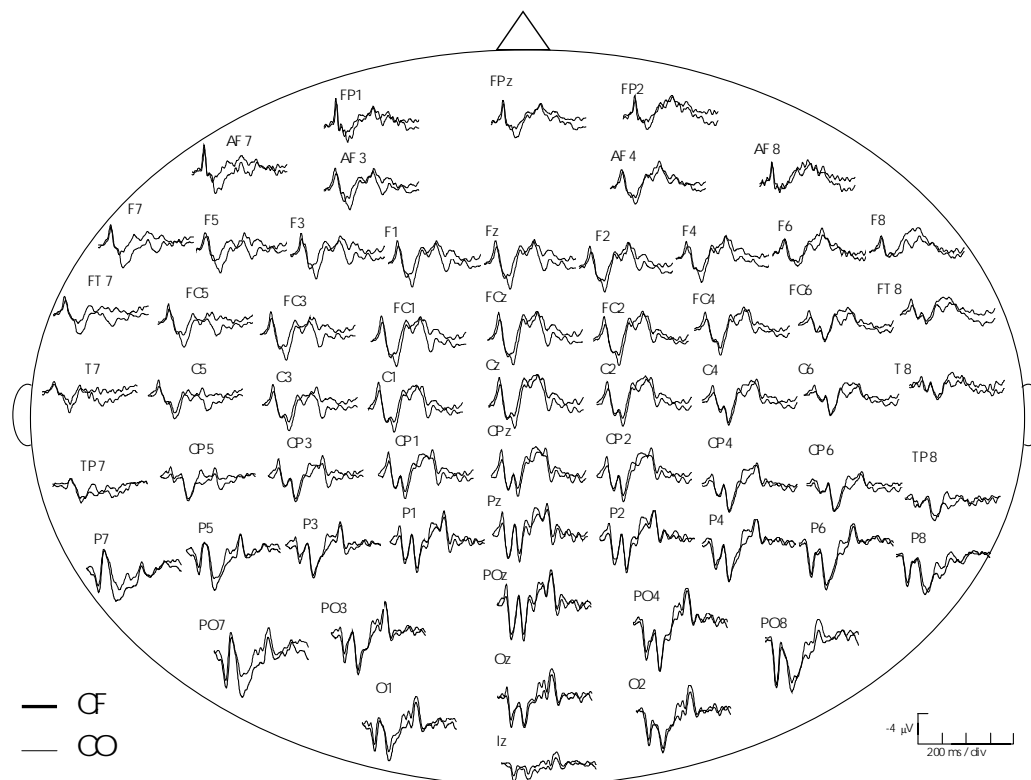
#### 3.1.1 Analyse des tracés

Pour ces analyses, les PEs sont référencés à l'électrode située sur le mastoïde gauche.

##### 3.1.1.1 Visualisation des tracés

La figure 19 présente les tracés obtenus pour les mots de classe ouverte et fermée. Visuellement, les tracés des deux classes de mots révèlent une première composante (P100) avec un minimum sur les sites antérieurs (N1) vers 100 ms accompagnée par une positivité sur les sites postérieurs (P1). A partir de 200 ms, la composante N2 présente une positivité frontale accompagnée par une négativité sur les régions postérieures. Ces premières composantes enregistrées entre 100 et 200 ms reflètent le traitement des caractéristiques physiques lors de la présentation visuelle d'un stimulus (Coles & Rugg, 1995). En ce qui concerne les classes de mots spécifiques, on remarque que les PEs associés aux mots de classe fermée (CF) sont plus négatifs en région frontale, effet qui semble plus prononcé sur les sites antérieurs gauches. De plus, il semble que cet effet soit présent sur deux périodes temporelles différentes: l'une précoce, et l'autre tardive. On observe que les mots de classe ouverte suscitent l'occurrence d'une négativité entre 300 et 500 ms, maximale sur les électrodes disposées en région centro-pariétale droite.





**Figure 19:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 61 électrodes enregistrées. Les tracés en gras correspondent aux PEs des mots de classe fermée (CF), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte (CO). Les électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, O pour occipital et I pour inion. La négativité est située vers le haut.

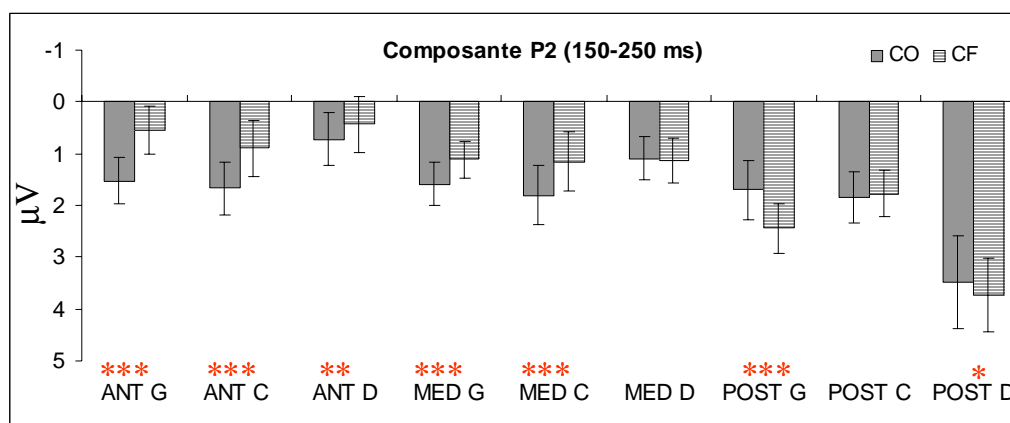
### 3.1.1.2 Analyse topographique par région d'intérêt

Quatre fenêtres temporelles ont été définies afin de réaliser nos analyses statistiques sur la base des analyses de tracés. Etant donné que les études précédentes ont montré que l'effet de classe était enregistré sur 4 composantes différentes, nous avons utilisé des fenêtres reflétant ces composantes: la P2 (entre 150 et 250 ms), la N280 (entre 250 et 350 ms), la N400 (entre 350 et 500 ms), et la LAN (entre 500 et 800 ms). Les résultats des ANOVAs à mesures répétées indiquent que, pour chaque fenêtre temporelle, des différences de topographiques existent entre les deux classes (voir tableau 9).

	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	1.01 - /	0.11 - /	0.60 - /	3.21 - /
Classe x Caudalité	2, 38	<b>4.36 * - 7.48 **</b>	<b>4.67 ** - 3.15 *</b>	<b>6.74 ** - 6.37 **</b>	<b>5.64 ** - 4.68 *</b>
Classe x Latéralité	2, 38	2.90 - 2.18	2.79 - 1.68	<b>10.61 *** - 8.12 **</b>	0.71 - 0.18
Classe x Caudalité x Latéralité	4, 76	<b>7.74 *** - 4.39 **</b>	<b>22.56 *** - 16.50 ***</b>	<b>4.89 ** - 3.01 *</b>	0.50 - 0.91

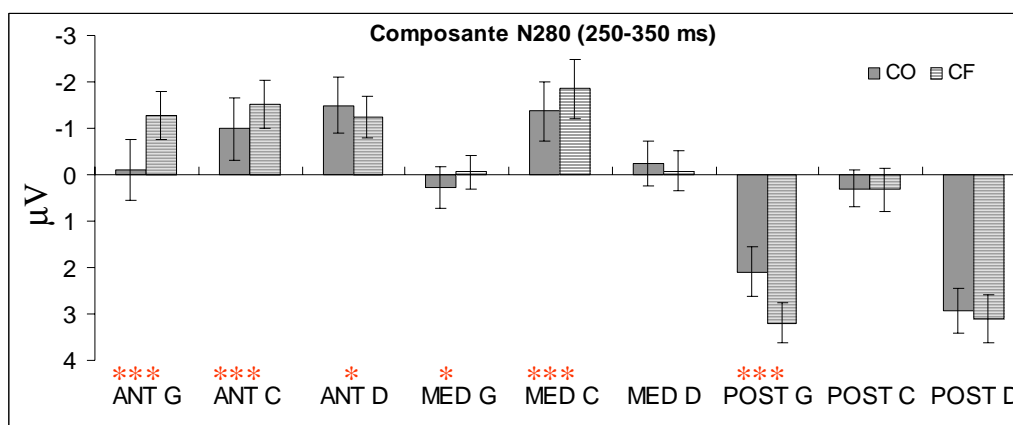
**Tableau 9:** Résultats des ANOVAs à mesures répétées pour les composantes P2 (150-250 ms), N280 (250-350 ms), N400 (350-500 ms), et LAN (500-800 ms) sur l'amplitude du PEs entre les deux classes de mots (CO vs. CF). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

- La P2 (150-250 ms): les résultats de l'ANOVA à mesures répétées entre 150 et 250 ms montrent l'occurrence d'une composante positive (P2) maximale en région postérieure ("caudalité":  $F(2,38) = 4.02$ ;  $p < .02$ ). L'interaction "classe de mots" x "caudalité" ( $F(2,38) = 7.48$ ;  $p < .001$ ) indique une réduction de la composante P2 pour les mots CF sur les régions antérieures ( $0.62 \mu\text{V}$ ) par rapport aux mots CO ( $1.30 \mu\text{V}$ ,  $p < .007$ ). De plus, la triple interaction "classe de mots" x "caudalité" x "latéralité" est significative ( $F(4,76) = 4.39$ ;  $p < .003$ ). La figure 20 présente les résultats des comparaisons planifiées. L'amplitude de la composante P2 des mots CF est réduite par rapport à celle des mots CO sur les régions antérieures (les 3), médianes gauche et centrale. L'effet inverse est observé sur les régions postérieures droite et gauche.



**Figure 20:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante P2 (comprise entre 150 et 250 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

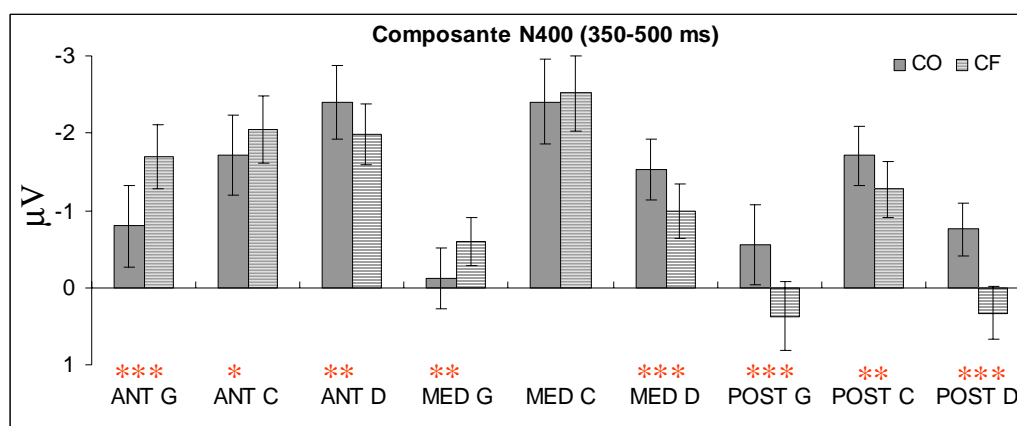
- La N280 (250-350 ms): les résultats de l'ANOVA à mesures répétées indiquent que, dans la fenêtre d'analyse 250-350 ms, la composante négative (N280) est maximale sur les régions antérieuro-médiane ("caudalité":  $F(2,38) = 20.84$ ;  $p < .00001$ ) et centrale ("latéralité":  $F(2,38) = 14.84$ ;  $p < .0001$ ). L'interaction "classe de mots" x "caudalité" ( $F(2,38) = 3.16$ ;  $p < .05$ ) montre que les PEs associés aux mots CF sont plus négatifs ( $-1.34 \mu\text{V}$ ) sur les régions antérieures que ceux des mots CO ( $-0.85 \mu\text{V}$ ,  $p < .03$ ). L'interaction "classe de mots" x "caudalité" x "latéralité" est significative ( $F(4,76) = 16.50$ ;  $p < .0001$ ). La figure 21 présente les résultats des comparaisons planifiées. Une composante négative sur les sites antérieurs est enregistrée pour les deux classes de mots mais avec une latéralisation différente: les PEs des mots CF sont plus négatifs sur les régions antéro-médianes gauche et centrale, alors que les mots CO sont plus négatifs sur les régions antérieures droites. De plus, les PEs associés aux mots CF sont plus positifs que ceux des mots CO sur les régions postérieures gauches.



**Figure 21:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N280 (comprise entre 250 et 350 ms). Les valeurs des PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

- La N400 (350-500 ms): les résultats de l'ANOVA à mesures répétées indiquent, une composante négative, maximale sur les régions antérieuro-médiane dans la fenêtre d'analyse 350-500 ms ("caudalité":  $F(2,38) = 4.09$ ;  $p < .02$ ) et latéralisée à droite ("latéralité":  $F(2,38) = 16.19$ ;  $p < .0001$ ). L'interaction "classe de mots" x "caudalité" ( $F(2,38) = 6.73$ ;  $p < .003$ ) indique que les PEs des mots CO sont plus négatifs ( $-1 \mu\text{V}$ ) que ceux des mots CF sur les régions postérieures ( $-0.18 \mu\text{V}$ ,  $p < .006$ ). L'interaction "classe de mots" x

"latéralité" ( $F(2,38) = 8.12; p < .001$ ) montre que les PEs associés aux mots CO sont plus négatifs à droite ( $-1.55 \mu\text{V}$ ) que ceux des mots CF ( $-0.88 \mu\text{V}, p < .00001$ ). Les comparaisons planifiées pour la triple interaction ( $F(4,76) = 3.00; p < .02$ , voir figure 22) indiquent que les PEs des mots CF sont plus négatifs sur les régions antérieures gauche et centrale, ainsi que médiane gauche, alors ceux associés aux mots CO suscitent l'occurrence d'une négativité largement distribuée sur l'hémisphère droit (3 régions), ainsi que sur les régions médianes et postérieures centrales.



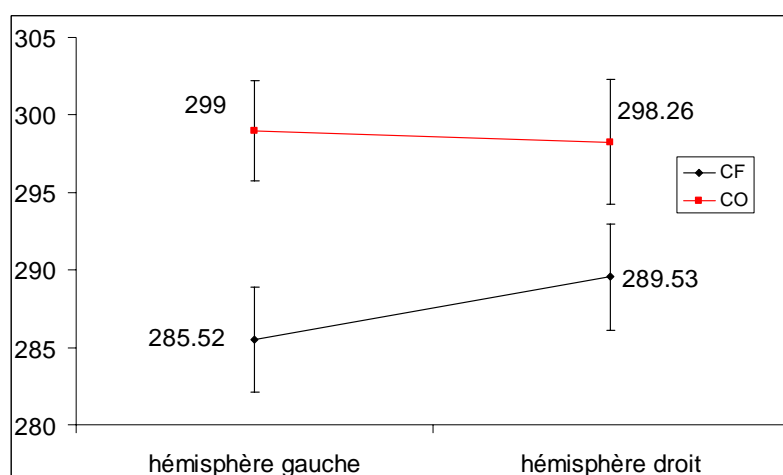
**Figure 22:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N400 (comprise entre 350 et 500 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

- La LAN (500-800 ms): les résultats de l'ANOVA à mesures répétées entre 500 et 800 ms montrent l'occurrence d'une composante négative maximale sur les régions antérieures ("caudalité":  $F(2,38) = 16.16; p < .0001$ ). L'interaction "classe de mots" x "caudalité" est significative ( $F(2,38) = 4.68; p < .0152$ ): les PEs des mots CF sont plus négatifs sur les régions antérieures ( $-0.97 \mu\text{V}$ ) et médianes ( $-0.47 \mu\text{V}$ ) que ceux associés aux mots CO ( $-0.17 \mu\text{V}, p < .0003$  et  $0.27 \mu\text{V}, p < .0007$  respectivement).

Les résultats des analyses de tracés montrent donc l'occurrence de patterns PEs spécifiques multiples pour chaque classe de mots: une composante P2 réduite en région antérieure pour CF, une composante N280 antérieure bilatérale pour CF, et antérieure latéralisée à droite pour CO, une composante N400 distribuée sur les régions centro-postérieures droites pour CO, et enfin, une composante LAN, sur les régions antérieures et médianes pour CF.

### 3.1.1.3 Analyse de latence de la N280

Nos analyses de topographie dans le rang de latence de la N280 (250-350 ms) montrent que la composante négative est localisée sur les sites antérieurs et bilatérale pour les mots CF, alors que pour les mots CO celle-ci est maximale à droite et au centre. Par conséquent, afin d'analyser la latence de cette composante, ainsi que les possibles effets de latéralisation hémisphérique, nous avons déterminé les latences pour les maxima de négativité sur 14 électrodes incluses dans les régions antérieures (FP1, FP2, AF7, AF3, AF4, AF8, F7, F5, F3, F1, F2, F4, F6, F8, en excluant FPz, et Fz <sup>13</sup>). Une ANOVA à mesures répétées avec les facteurs "classe de mots" (CF vs. CO), "hémisphère" (droit vs. gauche) et "électrodes" (7) montre un effet significatif du facteur "classe de mots" ( $F(1,19) = 23.03, p < .0001$ ): la latence de la composante N280 pour les mots de classe fermée ( $287.5 \pm 3.4$  ms) est plus précoce (de 11.1 ms) que celle enregistrée pour les mots de classe ouverte ( $298.6 \pm 3.62$  ms). De plus, le facteur principal "classe de mots" interagit avec l'"hémisphère" ( $F(1,19) = 5.12, p < .03$ ): la latence de la composante N280 des mots CF est plus précoce sur l'hémisphère gauche ( $285.5 \pm 3.3$  ms) que sur l'hémisphère droit ( $289.5 \pm 3.4$  ms,  $p < .01$ ). A l'opposé, la latence de la composante N280 pour les mots CO est identique sur les deux hémisphères (hémisphère gauche:  $299 \pm 3.5$  ms, hémisphère droit:  $298.2 \pm 4$  ms,  $p > .6$ , voir figure 23).



**Figure 23:** Illustration de l'interaction classe de mots x hémisphère sur la variable dépendante latence de la composante N280. les latences sont indiquées en ms.

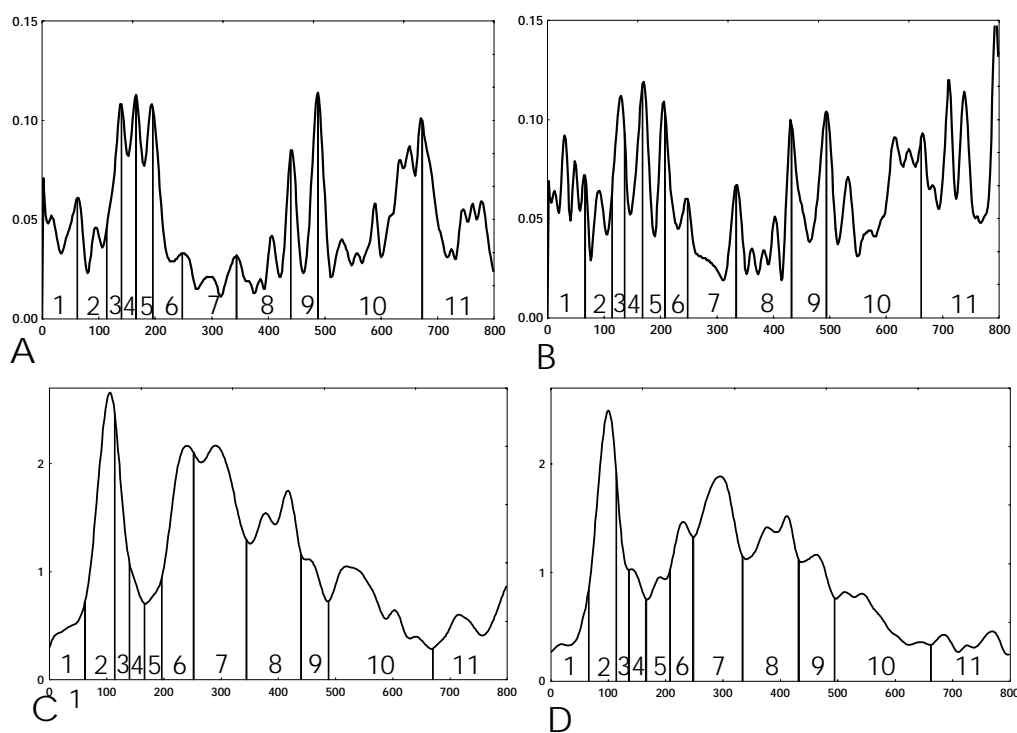
<sup>13</sup> Ces deux électrodes (Fpz et Fz) ont été exclues afin de pouvoir introduire le facteur hémisphère dans l'ANOVA à mesures répétées.

### 3.1.2 Analyse spatio-temporelle

Les analyses spatio-temporelles ont été effectuées sur les PEs calculés à partir de la référence moyenne (la référence moyenne au temps  $t$ , correspond à la moyenne des activités de toutes les électrodes au même instant  $t$ ).

#### 3.1.2.1 Détermination des micro-états

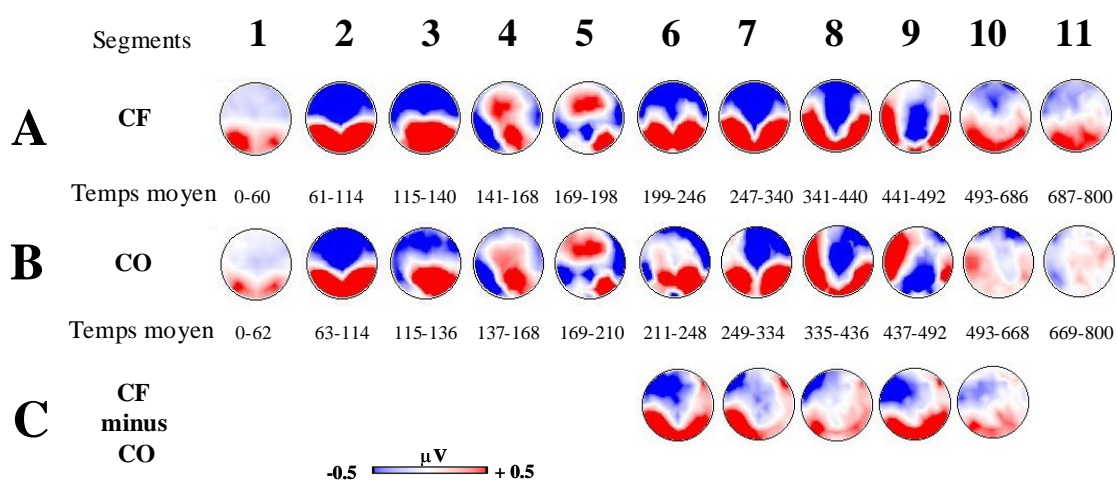
Sur la période de 800 ms correspondant à la fenêtre d'analyse, 11 périodes de stabilité ont été déterminées pour chaque classe de mot (voir figure 24).



**Figure 24:** (A & B) courbes de *Global Dissimilarity* (GD) pour les grandes moyennes (800 ms d'analyse), pour les PEs associés aux mots de classe fermée (A) et de classe ouverte (B). (C & D), courbes représentant le *Global Field Power* (GFP) pour les grandes moyennes pour les mots de classe fermée (C) et de classe ouverte (D). Les traits verticaux indiquent la frontière d'un segment défini par la méthode d'analyse de segmentation adaptée. Les nombres se réfèrent à l'ordre d'apparition des segments ou micro-états dans le temps. La segmentation adaptée révèle, que pour chaque condition expérimentale, 11 configurations différentes et stables sont présentes.

Ainsi, chaque période est équivalente à un segment ou micro-état. Pour chaque segment, une seule carte moyenne est calculée à partir des grandes moyennes et ce, dans les fenêtres

temporelles définies par la segmentation adaptée. De plus, les frontières sont calculées à partir des grandes moyennes pour chaque segment. En observant la figure 25, on s'aperçoit que chaque segment présente une topographie bien caractéristique, correspondant à des composantes que nous avons déjà décrites dans le chapitre 3 (section 4.3.1). Par exemple, les segments 2 et 3 présentent une positivité sur les sites postérieurs accompagnée d'une négativité frontale. Ces segments correspondent clairement à la fenêtre temporelle définie pour la composante P100. Les segments 4 et 5 manifestent une négativité sur les sites postérieurs qui est maximale à gauche. Les rangs de latence de ces segments correspondent au complexe N1/P2 (pour une description plus précise voir le chapitre 3, section 4.3.1).



**Figure 25:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales (A, classe fermée: CF; B, classe ouverte: CO) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. La couleur bleue représente des PEs négatifs, alors que les PEs positifs sont représentés en rouge. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, les cartes de différences de potentiels (classe fermée minus classe ouverte) où les topographies sont significativement différentes (segments 6, 7, 8, 9 et 10) entre les deux conditions expérimentales sont représentées.

### 3.1.2.2 Détermination des micro-états spécifiques

Afin de tester si les cartes des PEs associées aux mots de classe fermée sont spécifiques à cette condition expérimentale, chaque carte de surface de la condition CF a été introduite dans les données individuelles des sujets pour les mots de classe fermée, mais aussi pour ceux de classe ouverte. Cette procédure de "fitting" a été réalisée séparément pour chaque segment, dans une fenêtre temporelle définie sur la base des latences ( $\pm 10$  ms) obtenues sur la grande

moyenne (segment 1: 0-70 ms, segment 2: 50-130 ms, segment 3: 100-150 ms, segment 4: 120-180 ms, segment 5: 150-220 ms, segment 6: 190-270 ms, segment 7: 240-350 ms, segment 8: 320-460 ms, segment 9: 430-500 ms, segment 10: 480-690 ms, segment 11: 660-800 ms). La valeur du BEV (*Best Explained Variance* ou maximum de pourcentage de variance expliquée) pour chaque carte de la condition CF obtenue dans les deux conditions expérimentales de chaque participant, est alors comparée statistiquement à l'aide d'un *t*-test à échantillon apparié. Les résultats montrent que 5 cartes (numérotées de 6 à 10) présentent des topographies significativement différentes entre les mots de classe fermée et ouverte. La carte représentative du segment 6 obtenus à partir des mots CF explique mieux les données CF ( $70.8 \pm .04$  %) par rapport aux données CO ( $53.2 \pm .06$  %;  $t = -3.08$ ,  $p < .001$ ). Le même résultat est obtenu pour les segments 7 ( $71.6 \pm .04$  % pour les PEs des mots CF,  $60.1 \pm .06$  % pour les PEs des mots CO;  $t = -2.58$ ;  $p < .01$ ), 8 ( $65.3 \pm .04$  % pour les PEs des mots CF,  $52.9 \pm .06$  % pour les PEs des mots CO;  $t = -3.37$ ;  $p < .003$ ), 9 ( $47.3 \pm .04$  % pour les PEs des mots CF,  $32.9 \pm .06$  % pour les PEs des mots CO;  $t = -2.75$ ;  $p < .01$ ) et 10 ( $64.5 \pm .04$  % pour les PEs des mots CF,  $49.4 \pm .04$  % pour les PEs des mots CO;  $t = -3.44$ ;  $p < .002$ ).

L'analyse de "*fitting*" a donc permis de mettre en évidence que 5 micro-états (numérotés de 6 à 10) sont spécifiques à la condition de présentation des mots CF. Les latences des frontières pour les segments spécifiques, ainsi que leurs topographies de surface, sont illustrées dans la figure 25.

### 3.1.2.3 Analyse topographique par région d'intérêt

Pour chaque segment qui est spécifique (de 6 à 10) aux mots de classe fermée, nous avons effectué une ANOVA à mesures répétées avec les facteurs suivants: "classe de mots" (fermée vs. ouverte); "caudalité" (antérieur vs. médian vs. postérieur); "latéralité" (gauche vs. centrale vs. droite). Les régions d'intérêt utilisées sont décrites dans le tableau 3 du chapitre 3. Les résultats de ces analyses sont présentés dans le tableau 10.

	ddl	6	7	8	9	10
Classe x Caudalité	2, 38	<b>7.61 *</b> - <b>7.46 **</b>	3.50 * - 2.87	<b>3.45 *</b> - <b>4.26 *</b>	<b>11.88 ***</b> - <b>9.63 **</b>	<b>4.32 *</b> - <b>4.33 *</b>
Classe x Latéralité	2, 38	<b>6.4 **</b> - <b>3.81 *</b>	2.34 - 1.69	<b>6.44 **</b> - <b>4.65 *</b>	<b>13.86 ***</b> - <b>11.82 ***</b>	1.56 - 1.59
Classe x Caudalité x Latéralité	4, 76	<b>15.80 **</b> - <b>13.13 ***</b>	<b>22.49 ***</b> - <b>17.33 ***</b>	<b>6.18 ***</b> - <b>2.94 *</b>	<b>3.09 *</b> - <b>2.51 *</b>	0.96 - 1.08

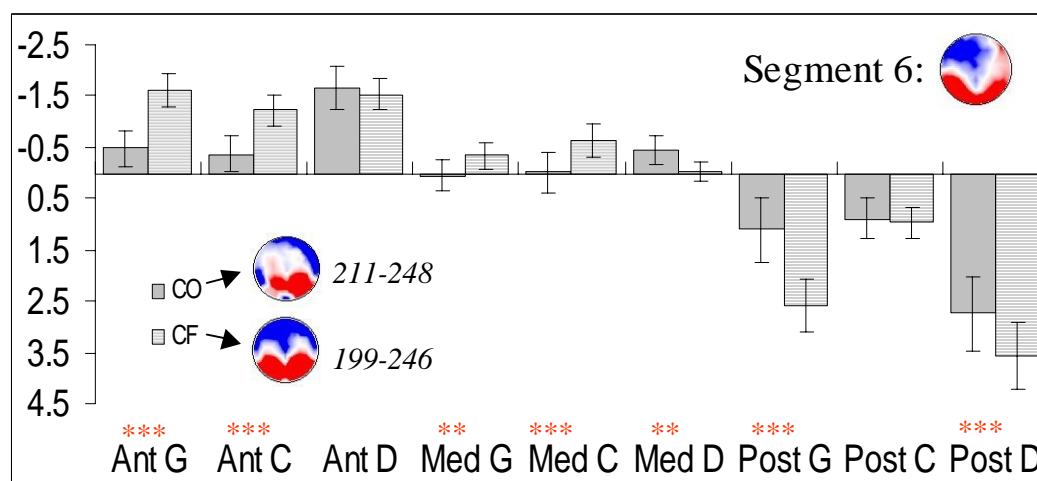
**Tableau 10:** Résultats des ANOVAs à mesures répétées séparées pour chaque segment (6, 7, 8, 9 et 10). Les valeurs de *F* sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique)



après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

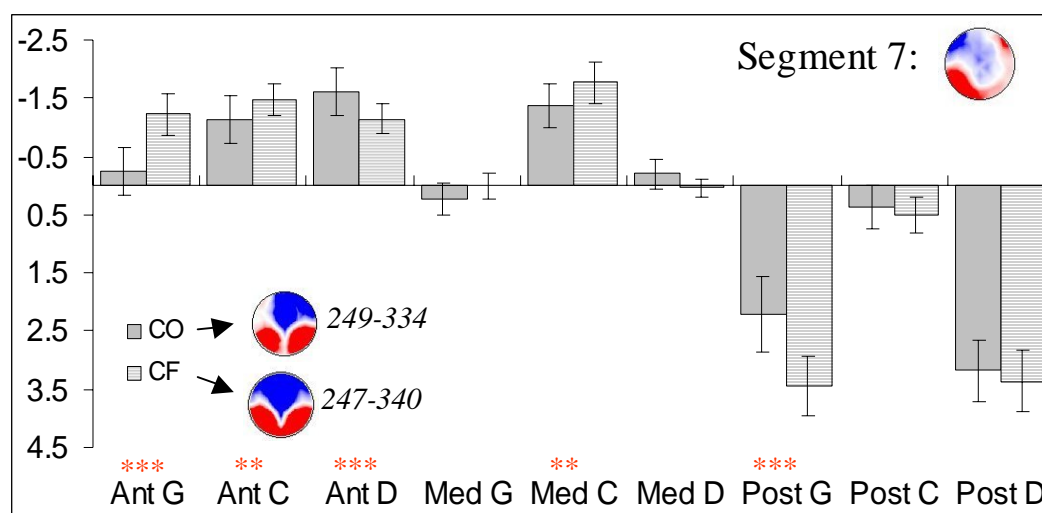
Afin d'éclaircir les différentes interactions significatives observées, nous avons réalisé des comparaisons planifiées (*Post-Hoc*, LSD) que nous allons maintenant décrire pour chaque segment.

- Segment 6 (CF: 199-246 ms, CO: 211-248 ms): dans cette fenêtre d'analyse, les PEs présentent une composante négative maximale en région antéro-médiane ("caudalité":  $F(2,38) = 16.36$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 7.47$ ;  $p < .001$ ) indique que les PEs des mots CF sont plus négatifs ( $-1.45 \mu\text{V}$ ) sur les régions antérieures que ceux des mots CO ( $-0.83 \mu\text{V}$ ,  $p < .02$ ). L'effet inverse est observé sur les régions postérieures (CF:  $2.37 \mu\text{V}$ , CO:  $1.58 \mu\text{V}$ ;  $p < .004$ ). L'interaction "classe de mots" x "latéralité" ( $F(2,38) = 3.82$ ;  $p < .03$ ), ainsi que la triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 13.14$ ;  $p < .0001$ ) montrent que les PEs associés aux mots CF sont plus négatifs que ceux associés aux mots CO sur les régions antéro-médianes gauche et centrale. Sur les régions postérieures gauche et droite et médiane droite, l'effet inverse est observé (voir figure 26).



**Figure 26:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 6 (compris entre 211-248 ms pour CO et 199-246 ms pour CF). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

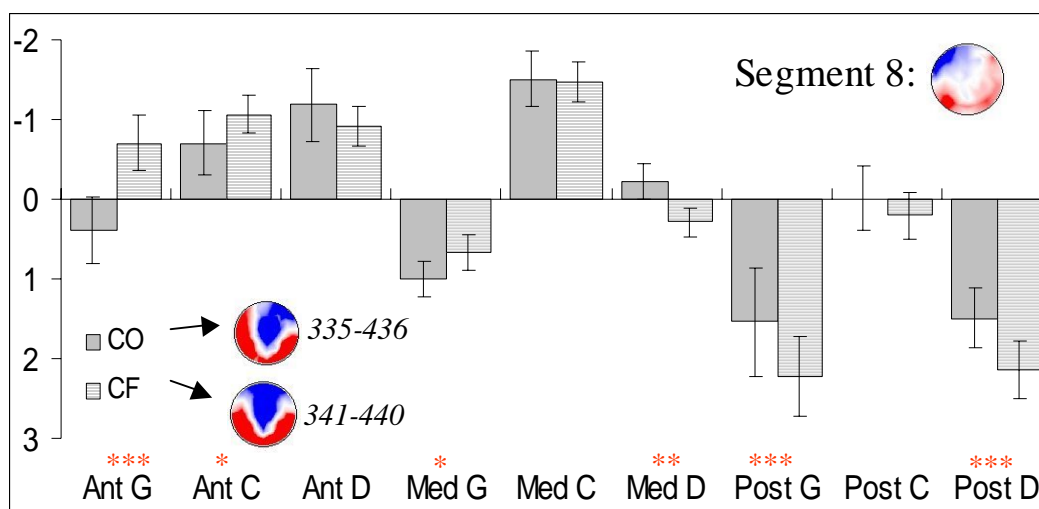
- Segment 7 (CF: 247-340 ms, CO: 249-334 ms): une composante négative est enregistrée dans ce segment, avec un maximum d'activité en région antéro-médiane ("caudalité":  $F(2,38) = 22.76$ ;  $p < .0001$ ) et centrale ("latéralité":  $F(2,38) = 14.61$ ;  $p < .0001$ ). L'analyse de la triple interaction ( $F(4,76) = 17.33$ ;  $p < .0001$ ) montre que les PEs des mots CF sont plus négatifs en régions antérieures gauche et centrale, ainsi que médiane centrale (voir figure 27) par rapport à ceux des mots CO. Il est à noter que sur les sites antérieurs droits, ce sont les PEs des mots CO qui sont plus négatifs que ceux des mots CF. Sur les régions postérieures gauches, les mots CF présentent une positivité plus importante que celle observée à partir des mots CO. Ce segment correspond au rang de latence utilisé pour décrire la composante N280 dans nos analyses de tracés.



**Figure 27:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 7 (compris entre 249-334 ms pour CO et 247-340 ms pour CF). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

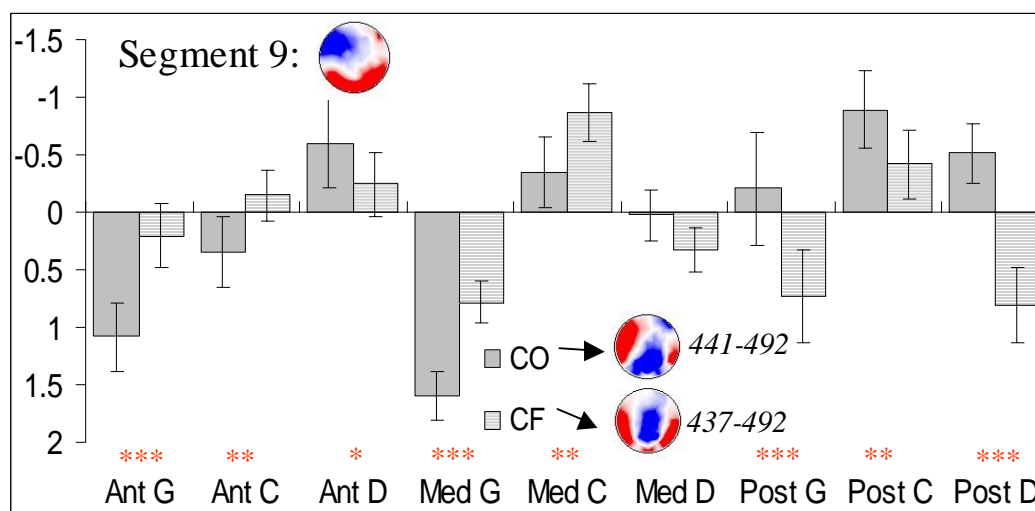
- Segment 8 (CF: 341-440 ms, CO: 335-436 ms): une composante négative en région antéro-médiane ("caudalité":  $F(2,38) = 8.79$ ;  $p < .0007$ ) et maximale en centrale et à droite ("latéralité":  $F(2,38) = 19.76$ ;  $p < .0001$ ) est enregistrée. L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 4.26$ ;  $p < .02$ ) indique que les PEs associés aux mots CO sont plus négatifs ( $1 \mu\text{V}$ ) sur les régions postérieures par rapport à ceux des mots CF ( $1.52 \mu\text{V}$ ,  $p < .04$ ), alors que l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 4.66$ ;  $p < .01$ ) montre que les PEs des mots CO sont plus négatifs ( $0.02 \mu\text{V}$ ) sur les régions à droite par rapport à ceux des mots CF ( $0.50 \mu\text{V}$ ,  $p < .002$ ). La triple interaction "classe de mots" x "caudalité" x

"latéralité" ( $F(4,76) = 2.95; p < .02$ ) indique que les PEs des mots CF sont plus négatifs en régions antérieures gauche et centrale et médiane gauche, alors qu'ils sont plus positifs en régions postérieures gauche et droite ainsi que médiane droite (voir figure 28) par rapport aux PEs associés aux mots CO.



**Figure 28:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 8 (compris entre 335-436 msec pour CO et 341-440 msec pour CF). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

- Segment 9 (CF: 441-492 ms, CO: 437-492 ms): une composante négative maximale sur les régions centrale et droite est enregistrée ("latéralité":  $F(2,38) = 9.29; p < .0005$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 9.68; p < .0004$ ) indique que les PEs des mots CO sont plus négatifs ( $-0.53 \mu\text{V}$ ) sur les régions postérieures par rapport à ceux des mots CF ( $0.37 \mu\text{V}$ ,  $p < .0001$ ). L'interaction "classe de mots" x "latéralité" ( $F(2,38) = 11.32; p < .0001$ ) montre que les PEs associés aux mots CO sont plus négatifs ( $-0.36 \mu\text{V}$ ) sur les régions à droite que ceux des mots CF ( $0.29 \mu\text{V}$ ,  $p < .00001$ ). L'analyse de la triple interaction ( $F(4,76) = 2.51; p < .04$ ) montre que les PEs des mots CF sont plus négatifs en régions antéro-médianes gauche et centrale par rapport à ceux des mots CO. En revanche, les PEs associés aux mots CO présentent une composante négative qui est significativement plus importante sur les régions postérieure (les 3) et antérieure droite par rapport à ceux des mots CF (voir figure 29).



**Figure 29:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 9 (compris entre 441-492 msec pour CO et 437-492 msec pour CF). Les valeurs des PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

La description de ces deux derniers segments (8 et 9) pour la fenêtre d'analyse, ainsi que la topographie des effets rapportés correspondent aux caractéristiques de la composante N400.

- Segment 10 (CF: 493-686 ms, CO: 493-668 ms): une composante négative maximale sur les régions antérieures est enregistrée ("caudalité":  $F(2,38)=13.76$ ;  $p < .0001$ ). Les comparaisons planifiées (*Post-Hoc*, LSD) de l'interaction significative "classe de mots" x "caudalité" ( $F(2,38) = 4.34$ ;  $p < .02$ ) indiquent que les PEs des mots CF sont plus positifs ( $0.89 \mu\text{V}$ ) sur les régions postérieures par rapport à ceux associés aux mots CO ( $0.38 \mu\text{V}$ ,  $p < .009$ ).

L'analyse de segmentation adaptée a permis de mettre en évidence 5 micro-états différents entre les deux classes de mots. Le segment 6 indique une augmentation de la négativité pour les mots CF en région frontale gauche. Pour le segment 7, les deux classes de mots présentent une négativité frontale, la N280: seule celle correspondant aux mots CO est latéralisée à droite. Les segments 8 et 9 indiquent l'occurrence d'une composante négative sur les régions centrale et postérieure droite pour les mots CO, correspondant à la composante N400. Pour le segment 10, la présentation de mots CF suscite l'occurrence d'une positivité plus importante sur les sites postérieurs par rapport aux mots CO.

### 3.1.2.4 Evolution temporelle des micro-états

Les analyses relatives à la latence des pics de GFP ont été réalisées pour les segments 6, 7 et 8. L'analyse sur les segments 9 et 10 a été supprimée car ces pics étaient difficiles à déterminer dans les données individuelles. Une ANOVA à mesures répétées avec comme facteur "classe de mots" (CF vs. CO) a été effectuée pour chaque segment. Les résultats sont décrits dans le tableau 11. La carte caractérisant le segment 7 (comprise entre 247-340 ms pour les mots de classe fermée et entre 249-334 ms pour les mots de classe ouverte) apparaît significativement plus précocement pour les mots de classe fermée que celle des mots de classe ouverte: la différence de latence est de 13 ms. Les autres segments, 6 et 8, ne présentent pas de décalage pour leurs maximums d'activité (voir tableau 11).

segment	classe fermée	classe ouverte	ANOVA
6	207.5 (6.14)	211.1 (5.47)	$F(1,19) = 1.44; p > .25$
7	280.1 (5.43)	293.2 (5.56)	<b><math>F(1,19) = 7.45; p &lt; .01</math></b>
8	410.4 (4.39)	412.5 (4.72)	$F(1,19) = .16; p > .69$

**Tableau 11:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 3 segments (6, 7, 8). Les valeurs indiquées sont en ms avec, entre parenthèses, l'erreur type. Les effets significatifs sont rapportés en gras. Pour le segment 7, la carte associée aux mots CF (280.1 ms) est enregistrée plus précocement (-13 ms) que celle obtenue pour les mots CO (293.2 ms).

## 3.2 PEs et effet de la catégorie grammaticale

Chaque classe de mots a été sous-divisée en 3 catégories grammaticales différentes: les mots de classe fermée (CF) étaient constitués de prépositions (N=40), de quantificateurs (N=40), et de non-quantificateurs (N=40), alors que les mots de classe ouverte (CO) contenaient des noms (N=40), des verbes (N=40), ainsi que des adverbes (N=40).

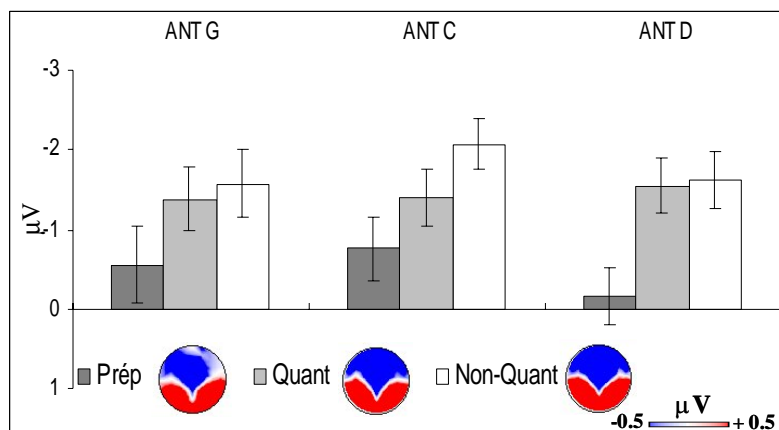
Dans notre analyse de tracés (section 3.1.1 de ce chapitre, figure 21), nous avons montré que la composante N280 était enregistrée pour les deux classes de mots, dans le rang de latence compris entre 250 et 350 ms. Cette négativité précoce, maximale sur les électrodes localisées sur les régions antérieures, présente des différences de latéralisation pour les deux classes de mots (bilatérale pour CF, et centrale et droite pour CO). Afin de déterminer si ces différences de latéralisation sont liées à la classe de mots et donc sont homogènes à l'intérieur de chaque

classe, nous avons choisi d'utiliser les données PEs calculées à partir d'une référence moyenne. De plus, comme notre analyse de segmentation adaptée montre qu'un segment (le 7) coïncide avec la fenêtre utilisée pour l'analyse des tracés, nous avons choisi d'effectuer nos analyses sur les différentes catégories grammaticales uniquement sur ce segment (entre 247 et 340 ms pour les mots de classe fermée, et entre 249 et 334 ms pour les mots de classe ouverte).

### 3.2.1 Topographie du segment 7

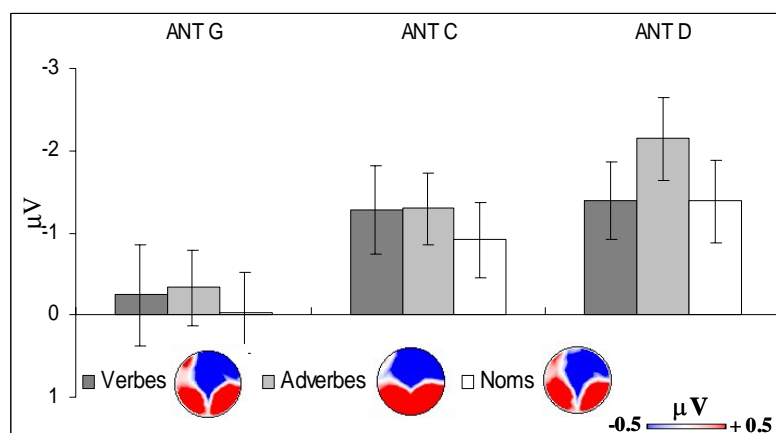
Puisque la composante N280 est distribuée uniquement sur les sites antérieurs, nous avons restreint l'analyse topographique aux électrodes localisées dans cette région. Deux ANOVAs à mesures répétées séparées ont été effectuées pour chaque classe (CO et CF) avec les facteurs suivants: "catégorie grammaticale" (3, CF: prépositions *vs.* quantificateurs *vs.* non-quantificateurs; CO: verbes *vs.* noms *vs.* adverbes) et "latéralité" (gauche *vs.* centrale *vs.* droite). La variable dépendante était l'amplitude moyenne des PEs dans les 3 régions antérieures (voir tableau 3, chapitre 3 section 3.1), entre 247-340 ms pour les mots de classe fermée, et entre 249-334 ms pour les mots de classe ouverte.

Les résultats montrent que, pour les mots de classe fermée, l'effet principal du facteur "catégorie grammaticale" est significatif ( $F(2,38) = 6.10$ ;  $p < .005$ ). Le facteur "latéralité" et l'interaction "catégorie grammaticale" x "latéralité" ne sont pas significatifs ( $F(2,38) = .75$ ;  $p > .4$ ; et  $F(4,76) = 1.4$ ;  $p > .23$ ; respectivement), ce qui indique que les négativités sont réparties équitablement sur le scalp (bilatérales) pour les 3 catégories grammaticales. Les comparaisons planifiées (*post-hoc*, LSD, figure 30) montrent que les PEs associés aux prépositions ( $-0.49 \mu\text{V}$ ) sont (sur les 3 sites antérieurs) moins négatifs que ceux associés aux quantificateurs ( $-1.44 \mu\text{V}$ ,  $p < .01$ ) et aux non-quantificateurs ( $-1.75 \mu\text{V}$ ,  $p < .001$ ). Il n'y a pas de différence entre les PEs associés aux quantificateurs et aux non-quantificateurs ( $p > .4$ ).



**Figure 30:** Illustration de la distribution des négativités des mots de classe fermée, pour le segment 7, compris entre 247 et 340 ms pour les 3 catégories grammaticales. Les valeurs des PEs sont en  $\mu\text{V}$ . Les barres représentent l'erreur type. ANT G pour antérieure gauche, ANT C pour antérieure centrale et ANT D pour antérieure droite. Pour chaque catégorie grammaticale, la carte moyenne de surface est présentée. Les résultats de l'ANOVA à mesures répétées montrent que les négativités associées aux quantificateurs et non quantificateurs sont significativement plus importantes que celle associée aux prépositions sur les 3 régions antérieures.

Pour les mots de classe ouverte, la négativité est identique pour les 3 types de mot: l'effet du facteur "catégorie grammaticale" ( $F(2,38) = .79$ ;  $p > .46$ ) et l'interaction "catégorie grammaticale" x "latéralité" ( $F(4,76) = .69$ ;  $p > .60$ ) ne sont pas significatifs. En revanche, le facteur "latéralité" est significatif ( $F(2,38) = 5.51$ ;  $p < .007$ ): les comparaisons planifiées (*post-hoc* LSD, voir figure 31) indiquent que les PEs associés aux trois catégories grammaticales de classe ouverte sont moins négatifs sur les électrodes situées à gauche ( $-0.19 \mu\text{V}$ ) que sur les électrodes centrales ( $-1.16 \mu\text{V}$ ,  $p < .03$ ) et à droite ( $-1.64 \mu\text{V}$ ,  $p < .002$ ). Aucune différence n'est observée entre les régions centrales et droites ( $p > .28$ ). La négativité pour les mots CO est donc maximale au centre et à droite.



**Figure 31:** Illustration de la distribution des négativités des mots de classe ouverte, pour le segment 7, compris entre 249 et 334 ms pour les 3 catégories grammaticales. Les valeurs des PEs sont en  $\mu\text{V}$ . Les barres représentent l'erreur type. ANT G pour antérieure gauche, ANT C pour antérieure centrale et ANT D pour

antérieure droite. Pour chaque catégorie grammaticale, la carte moyenne de surface est présentée. Les résultats de l'ANOVA à mesures répétées montrent que les négativités associées aux trois catégories grammaticales sont équivalentes, et plus importantes sur les régions antérieures à droite et en centrale.

Les résultats suggèrent que seuls les mots de classe fermée suscitent des PEs différenciés en fonction de leur catégorie grammaticale: les prépositions montrent des PEs moins négatifs sur les sites antérieurs par rapport à ceux des deux autres catégories grammaticales de la classe fermée (quantificateurs et non-quantificateurs).

### 3.2.2 Latence du segment 7

Alors que nous avons vu que la latence de la composante N280 était plus précoce pour les mots CF par rapport à celle des mots CO (section 3.1.1.3 en référence mastoïde gauche et 3.1.2.4 en référence moyenne), nous voulions déterminer si cet effet était homogène entre les différentes catégories grammaticales de chaque classe et, le cas échéant, si d'éventuelles différences de latence peuvent être expliquées par les variations des propriétés lexicales telles que la fréquence et la longueur. Pour chaque participant, le maximum du pic de GFP pour la N280 a été calculé pour les 6 catégories grammaticales, et soumis à une ANOVA à mesures répétées avec, comme facteur principal "catégorie grammaticale" (verbes *vs.* adverbes *vs.* noms *vs.* prépositions *vs.* quantificateurs *vs.* non-quantificateurs). Les résultats montrent que les différences de latence entre les 6 catégories grammaticales sont significatives ( $F(5,95) = 5.52$ ;  $p < .0002$ ). Afin de déterminer les relations entre la latence de cette composante et les propriétés lexicales des mots (fréquence d'occurrence et longueur), nous avons déterminé la fréquence d'occurrence moyenne (log de fréquence) de chaque catégorie grammaticale et sa longueur moyenne (nombre de lettres, voir tableau 12).

Une analyse plus fine a été réalisée afin de déterminer si, à l'intérieur de chaque classe (CO et CF), les propriétés lexicales (fréquence et longueur) étaient homogènes. Les résultats montrent que, pour les mots de classe ouverte, les 3 catégories grammaticales sont de même longueur (verbes: 8.19, adverbes: 7.90, noms: 8.23;  $F(2,157) = 0.27$  ;  $p > .76$ ; pour les valeurs voir le tableau 12) tandis que les fréquences sont différentes ( $F(2,157) = 8.63$  ;  $p < .0003$ ). Les noms ont une fréquence plus basse (3.46 log de fréquence) par rapport aux verbes (3.91 log de fréquence,  $p < .002$ ) et aux adverbes (4.13 log de fréquence,  $p < .00001$ ). Pour les mots de classe fermée, les catégories grammaticales présentent des longueurs différentes ( $F(2,117)$

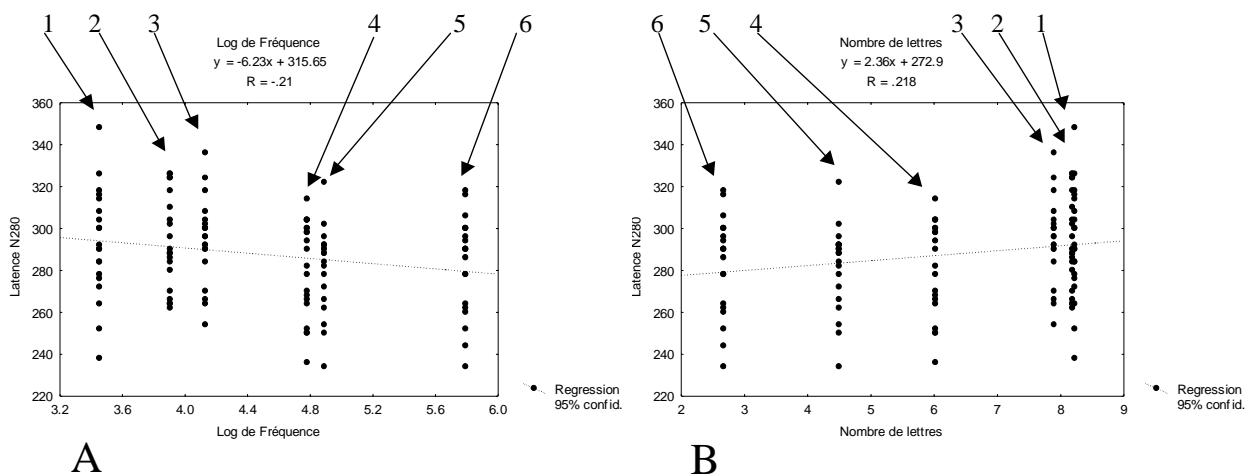


= 53.14 ;  $p < .0001$ ), les non-quantificateurs étant les mots les plus courts (2.68 lettres), puis les prépositions (4.50 lettres) et enfin les quantificateurs (6.03 lettres). Les fréquences d'utilisation des mots de classe fermée sont aussi différentes ( $F(2,117) = 26.62$  ;  $p < .0001$ ). La fréquence d'occurrence des prépositions (4.50 log de fréquence) et quantificateurs (6.03 log de fréquence) sont identiques ( $p > .46$ ) et plus basse que celle des non-quantificateurs (5.79 log de fréquence,  $p < .00001$  et  $p < .00001$  respectivement).

Classe	Type	Longueur	Log Freq
CO	Verbes	8.19 (0.24)	3.91 (0.09)
CO	Adverbes	7.90 (0.41)	4.13 (0.10)
CO	Noms	8.23 (0.31)	3.46 (0.12)
CF	Prépositions	4.50 (0.25)	4.89 (0.12)
CF	Quantificateurs	6.03 (0.29)	4.78 (0.09)
CF	Non-Quantificateurs	2.68 (0.12)	5.79 (0.10)

**Tableau 12:** Chaque catégorie grammaticale est décrite en fonction de leur appartenance à une classe de mot CO: classe ouverte, CF: classe fermée, ainsi qu'en fonction de leur longueur (en nombre de lettres) et leur log de fréquence d'occurrence dans la langue française (Content, *et al.*, 1990). Les chiffres entre parenthèses indiquent l'erreur standard.

Deux analyses de régression ont été réalisées avec, comme prédicteurs, la fréquence d'occurrence (log de fréquence) et la longueur. La figure 32 montre les résultats des équations de régression ainsi que les points indiquant les prédicteurs pour le log de fréquence (gauche) et la longueur (droite). Des corrélations significatives sont observées pour les deux prédicteurs (log de fréquence:  $F(1,118) = 5.37$ ,  $p < .02$ ,  $R = -.21$ ; longueur:  $F(1,118) = 5.83$ ,  $p < .01$ ,  $R = .21$ ). Ces deux régressions reflètent chacune 4 % de la variance expliquée de la latence pour la composante N280. Une analyse de régression additionnelle a été effectuée montrant que la classe de mot est, elle aussi, un prédicteur de la latence ( $F(1,118) = 8.44$ ,  $p < .004$ ,  $R = -.25$ ), un résultat attendu étant donné que le log de fréquence et la longueur sont hautement corrélées avec la classe de mots dans nos stimuli utilisés (86.9 et 89 % respectivement).



**Figure 32:** Résultats des analyses de régression sur la latence de la négativité précoce (ms). A, le prédicteur est le log de fréquence et en B la longueur. Pour les mots CO: 1 = noms, 2 = verbes, 3 = adverbes; pour les mots CF: 4 = quantificateurs, 5 = prépositions, 6 = non quantificateurs. Les deux régressions son significatives.

## 4 Discussion

### 4.1 Tracés et effet de classe de mots

Nos résultats montrent que, dans les 4 fenêtres temporelles testées (150-250 ms, 250-350 ms, 350-500 ms et 500-800 ms), les PEs des deux classes de mots présentent des différences de topographies: des effets maximaux ont été rapportés sur les électrodes localisées en région antérieure gauche pour les fenêtres temporelles 150-250 ms, 250-350 ms et 500-800 ms, alors que, pour la fenêtre 350-500 ms, les différences sont maximales sur les régions centro-postérieures droites. Ces résultats sont concordants avec la littérature précédente portant sur la distinction entre classes de mots (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997).

#### 4.1.1 Effet précoce visuel: la P2

Entre 150 et 250 ms, la composante P2 est réduite sur les sites antérieurs pour les mots CF. Ces résultats répliquent ceux rapportés dans la littérature (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). La composante P2 est généralement associée à l'encodage des caractéristiques visuelles du stimulus lors des processus perceptifs initiaux de catégorisation

des patterns (lettres), plutôt qu'à une analyse lexicale de ces patterns (Hillyard, Teder-Salejarvi, & Münte, 1998). De plus, cette composante s'est avérée sensible aux modulations de l'attention visuelle spatiale et, plus particulièrement, à la position du stimulus dans le champ visuel (Clark & Hillyard, 1996). Ainsi, il semble peu probable que l'effet rapporté dans notre recherche sur la composante P2, soit lié à des processus linguistiques reflétant les distinctions entre mots de classe ouverte et fermée. La fluctuation de l'amplitude de la composante P2 semble plutôt résulter des différences de longueur entre les mots de ces deux classes, ce qui pourrait entraîner de possibles effets visuo-spatiaux attentionnels (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999).

#### 4.1.2 Effet précoce: la N280

Dans le rang de latence compris entre 250 et 350 ms, des différences topographiques émergent également entre les mots de classe fermée et ouverte, bien que tous deux suscitent l'occurrence de négativités dans ce rang de latence. La négativité précoce associée aux mots CF, est équitablement répartie sur les sites antérieurs, alors que celle des mots CO est latéralisée sur les électrodes centrale et droite (figure 21). Ainsi, la différence entre les deux classes de mots est maximale sur les sites médio-antérieurs à gauche et au niveau central. Cette topographie et sa modulation en fonction de la classe de mot est comparable aux données précédemment citées (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Neville, *et al.*, 1992; Nobre & McCarthy, 1994; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). Les différences topographiques dans cette fenêtre temporelle sont généralement reliées aux étapes précoces du traitement d'un mot, lorsque les informations lexico-syntaxiques et lexico-sémantiques sont activées et sont rendues disponibles pour être utilisées plus tardivement lors des processus d'intégration (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). Cette interprétation semble plausible compte tenu du fait que cet effet apparaît après les effets visuels précoces liés à la reconnaissance des patterns physiques des mots (complexe N1/P2), mais avant les étapes d'intégration post-lexicales reflétées, par exemple, par la composante N400 (Kutas & van Petten, 1994).

La plupart des études récentes sur cet effet précoce s'accordent pour suggérer que les négativités précoces sont la manifestation d'un même et unique processus pour les deux classes de mots (Brown, *et al.*, 1999; King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). Nos résultats suggèrent toutefois que, malgré le fait que les deux classes de mots entraînent toutes deux des négativités dans la fenêtre temporelle de la N280, leurs distributions topographiques sont différentes. Ainsi, l'analyse des PEs obtenus sur l'ensemble des 61 électrodes a permis de mettre en évidence l'existence possible d'un traitement spécifique aux mots de classe fermée, lors d'une étape précoce du processus de compréhension. L'interprétation de ces différences topographiques précoces entre les deux classes de mots reste toutefois complexe. Elle est abordée dans la section 4.3 de ce chapitre.

#### 4.1.3 Effet tardif: la N400

Nous avons mis en évidence une composante négative sur les sites centro-postérieurs droits dans la fenêtre temporelle comprise entre 350 et 500 ms, lors de la présentation des mots CO, et réduite voir absente lors de la présentation de mots CF. Ces résultats sont concordants avec ceux de la littérature (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Kutas & Hillyard, 1983; Kutas, *et al.*, 1988; Münte, *et al.*, 2001; Neville, *et al.*, 1993; Neville, *et al.*, 1992; Nobre & McCarthy, 1994; Osterhout, *et al.*, 1997; Pulvermüller, *et al.*, 1995; van Petten & Kutas, 1991). Cette composante a été souvent étudiée et est supposée refléter des processus post-lexicaux d'intégration sémantique dans la phrase: plus un mot est difficile à intégrer, et plus l'amplitude de la composante N400 est importante (Kutas & Hillyard, 1984; Kutas & van Petten, 1994). Pour les mots de classe ouverte, la présence de cette composante N400 pourrait indiquer que ceux-ci disposent d'informations lexico-sémantiques au contraire des mots de classe fermée. Toutefois, il semble intéressant de noter que la fréquence affecte aussi la composante N400: plus la fréquence d'un mot est basse et plus la composante N400 associée sera d'amplitude importante (Rugg, 1990). Comme nos stimuli ne sont pas contrôlés et que les mots de classe ouverte sont de fréquence moins élevée que les mots de classe fermée, nous ne pouvons préciser si l'effet N400 provient ou non de différences de propriétés lexicales.

#### 4.1.4 Effet tardif: la LAN

Dans le rang de latence compris entre 500 et 800 ms, nous avons montré que les mots CF suscitent l'occurrence d'une composante négative largement distribuée sur les sites médio-antérieurs central et gauche (*Left Anterior Negativity*). Ces résultats coïncident avec ceux précédemment obtenus pour des mots inclus dans des phrases ou des textes (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Neville, *et al.*, 1993; Neville, *et al.*, 1992). Comme la LAN associée aux mots CF n'a été enregistrée qu'à partir de phrases ou de textes, cette composante a été classée comme appartenant à la famille de la contingente négative: les lecteurs sont engagés dans des stratégies d'anticipation des mots à venir. Cette interprétation concorde avec le rôle des mots CF au sein d'une phrase. Toutefois, trois récentes études ont enregistré une LAN pour les mots CF, aussi bien dans un contexte de phrases ou de textes que présentés au sein de listes de mots sans lien entre eux (Keurs ter, *et al.*, 2002; Münte, *et al.*, 2001; Osterhout, *et al.*, 1997). Comme la lecture de mots est automatisée, Münte, *et al.* (2001) suggèrent que, même lors de la présentation de listes de mots, des processus d'attente sont mis en place. De plus, le fait que, seuls les mots CF de très haute fréquence suscitent l'occurrence d'une LAN, suggère d'après Münte, *et al.* (2001) que ces derniers induisent des attentes plus importantes que les autres mots (CF de basse fréquence et CO). Alors qu'Osterhout, *et al.* (2002) montrent que la longueur affecte les négativités enregistrées sur toute la période d'analyse (donc incluant aussi la fenêtre d'analyse de la composante LAN), l'interprétation fonctionnelle de cette composante semble encore nécessiter de plus amples recherches.

#### 4.2 Apport de la segmentation

La distinction entre classe de mots montre des différences topographiques sur le scalp commençant dès 199 ms jusqu'à 686 ms. Cette période englobe 5 segments (6, 7, 8, 9, 10) qui semblent spécifiques à la présentation des mots de classe fermée. Maintenant, nous allons comparer les résultats obtenus avec les deux analyses (tracés et segmentation adaptée) afin de mieux caractériser chacun des segments pré-définis.

Le segment 6, enregistré entre 211-248 ms pour les CO et entre 199-246 ms pour les CF, correspond à une portion de la fenêtre d'analyse de la composante P2. Alors qu'avec les

analyses de tracés et une référence mastoïde gauche, tous les PEs étaient positifs (figure 20), la référence moyenne met en évidence la bipolarité (pôle positif et négatif) des composantes (figure 26). De plus, grâce à la segmentation adaptée, nous avons défini une fenêtre temporelle plus restreinte (comparée à l'analyse de tracés 150-250 ms) et, par conséquent, plus précise. Les différences topographiques débutent avec ce segment (à 211 ms pour la CO et 199 ms pour la CF), et non pas à 150 ms comme l'avait défini l'analyse de tracés. Puisque d'une part, une portion de ce segment correspond à la fenêtre temporelle de la composante P2 définie par notre analyse de tracés et que d'autre part, les propriétés lexicales de nos stimuli ne sont pas contrôlées, nous suggérons que ce segment 6 puisse refléter une étape de traitement visuel précoce des patterns du stimulus.

Le segment 7, enregistré entre 249-334 ms pour les CO et entre 247-340 ms pour les CF, correspond à la fenêtre d'analyse de la composante N280 étudiée dans l'analyse des tracés. Dans ce segment, nous avons trouvé, qu'en régions antérieures, les amplitudes moyennes des négativités sont significativement différentes pour les deux classes de mots (figure 27). De plus, l'analyse du maximum de GFP indique que l'activité maximale est plus précoce pour les mots CF comparés aux mots CO. Ces résultats coïncident avec ceux obtenus lors de l'analyse de tracés et nous indiquent que ce segment présente toutes les caractéristiques de l'effet classique N280. Comme nous l'avons indiqué dans la section 3.2.2, la latence du maximum de GFP fluctue dans ce segment avec les propriétés lexicales telles que le log de fréquence et la longueur. Puisque ces propriétés log de fréquence, longueur et classe de mots corrélaient à plus de 80 % dans notre recherche, il semble difficile de préciser si ce délai dans la latence du maximum de GFP du segment 7 est lié à une différence dans l'accès lexical des deux classes de mots, ou bien est plutôt relié à des différences dans les propriétés lexicales des mots. Toutefois, puisque la fréquence est une propriété modifiant la latence de ce segment, et que celle-ci est considérée comme un facteur influençant l'accès au lexique (Balota & Chumbley, 1984; Forster, 1976; Forster, 1979; Glanzer & Ehrenreich, 1979; Morton, 1969), nous suggérons que ce segment 7, correspondant à l'occurrence de la composante N280, reflète les processus d'accès au lexique.

La fenêtre d'analyse de la composante N400 (350-500 ms) utilisée pour l'analyse de tracés coïncide avec les segments (8, 9) de la segmentation adaptée. Alors que pour l'analyse de tracés, une seule étape est considérée dans cet effet, la segmentation adaptée a défini 2 périodes de stabilité sur la base des variations de topographies où les PEs entre les deux

classes de mots différaient. Pour le segment 8 (enregistré entre 335-436 ms pour CO et entre 341-440 ms pour CF) et pour le segment 9 (enregistré entre 436-492 ms pour CO et entre 441-492 pour CF), les PEs associés aux mots CO sont plus négatifs sur les sites postérieurs droits, que ceux associés aux mots CF. Nos résultats indiquent l'occurrence d'une composante N400 pour les mots CO, identiquement distribuée sur les sites postérieurs droits, entre 335-492 ms. Pour le segment 10 (enregistré entre 493-668 ms pour CO et entre 493-686 ms pour CF), les mots CO suscitent l'occurrence d'une négativité maximale en région postérieure et non plus latéralisée mais bilatérale. Ce dernier segment (10) coïncide avec le début de la fenêtre d'analyse de la composante LAN (500-800 ms). L'utilisation de la segmentation adaptée a permis de mettre en évidence que des composantes connues telles que la N400, sont peut-être sous-divisées en différentes sous-composantes (pour une discussion sur la N400 voir chapitre 3, section 5.1). Pour définir la fonction de ces différents segments, la difficulté à l'heure actuelle, réside dans le fait que les interprétations des résultats PEs se font en accord avec les composantes connues révélées par les études de tracés. Nous l'avons vu, ces analyses de tracés (fenêtre temporelle et distribution des électrodes sur le scalp) peuvent ne pas être adaptées aux résultats obtenus (Steinhauer & Ullman, 2002). La segmentation adaptée semble être un bon compromis pour capturer les fines différences topographiques qui peuvent exister. Nous regrouperons ces 3 derniers segments sous l'appellation N400-LAN car sur la base de notre paradigme expérimental, nous ne pouvons pas inférer leur interprétation fonctionnelle. L'occurrence d'une composante N400 pour les mots CO, et l'absence ou une réduction de la N400 pour les mots de classe fermée, peuvent être reliés à l'activation des informations lexico-sémantiques pour les mots CO. Toutefois, la différence dans les propriétés lexicales (fréquence et longueur) entre les deux classes de mots nous empêche d'être catégorique.

Pour finir, l'intérêt d'une telle analyse réside dans le fait que toutes les électrodes sont incluses dans les analyses de topographie et des latences. Nous avons montré que les résultats sur les latences de pics de GFP (décalage de 13 ms pour le segment 7 en faveur des mots CF) coïncident avec ceux obtenus sur les pics de composantes (décalage de 11 ms sur les électrodes antérieures en faveur des mots CF). L'intérêt d'utiliser la variable correspondant au maximum de GFP réside dans le fait qu'aucun choix d'électrode spécifique pour l'analyse statistique n'est effectué. Les résultats indiquent des fluctuations de latence d'apparition de cartes de surface et non pas de décalage de pic au niveau de certaines électrodes locales.

### 4.3 Interprétation de la latence de la N280 et des propriétés lexicales

Les résultats sur l'étude de la latence de la composante N280 et des propriétés lexicales sont primordiaux pour l'interprétation fonctionnelle de cette composante. En effet, la fréquence d'occurrence d'un mot est une propriété qui est implémentée dans les modèles psycholinguistiques lors de l'accès lexical (cette propriété peut affecter différemment la nature ou encore l'organisation des représentations, voir chapitre 1, section 2). Ainsi, déterminer que la fréquence d'occurrence d'un mot affecte la latence de la composante N280 peut suggérer que cette composante reflète des étapes impliquées lors de l'accès lexical. Les études montrant une absence d'effet des propriétés lexicales (fréquence et longueur) sur la latence de la composante négative précoce N280, ont suggéré que cette composante n'est pas directement liée au processus d'accès lexical (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999). A l'opposé, les études montrant une corrélation entre la latence de cette composante négative et les propriétés lexicales des mots, suggèrent que cette fenêtre temporelle précoce reflète des processus lexicaux (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). Toutefois, puisque les différences observées entre les deux classes de mots sont intimement liées aux différences quantitatives des propriétés lexicales, ces derniers auteurs indiquent que les deux classes de mots sont traitées de manière identique lors des processus d'accès au lexique.

Nous avons montré que les 3 facteurs, la classe de mot, le log de fréquence ainsi que la longueur sont de bons prédicteurs de la latence de la composante N280. Comme ces 3 facteurs sont hautement corrélés, nous n'avons pas pu, dans cette expérience, différencier leur impact relatif sur les PEs. Si les propriétés lexicales interviennent lors de l'accès au lexique, comme le suggère Balota & Chumbley (1984), le fait qu'elles influencent la latence de la composante N280 (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997), indique que cette composante, ou plus précisément sa fenêtre temporelle, reflète les processus d'accès au lexique. Nos stimuli n'étant ni contrôlés en fréquence ni en longueur, on ne peut savoir si la différence de latence de la composante N280 entre les deux classes de mots est plutôt liée aux différences quantitatives des propriétés lexicales des deux classes de mots, ou plutôt à des différences qualitatives entre les deux.



Notons toutefois, que nos résultats portant sur l'analyse de régression diffèrent de ceux de King & Kutas (1998), Osterhout, *et al.* (2002) et de Osterhout, *et al.* (1997) dans le pourcentage de variance expliquée pour la latence de la composante N280. Alors que seulement 4 % de notre variance est expliquée par les propriétés lexicales, ces auteurs indiquent des résultats compris entre 46 (Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) et 92 % (King & Kutas, 1998). Ces différences peuvent trouver deux explications. D'une part, les fréquences d'occurrence des mots sont calculées à partir de base de données limitées en nombre. En effet, les trois autres recherches (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) utilisent la base de données BROWN dans laquelle les fréquences sont calculées sur 1 million d'occurrences (Francis & Kucera, 1982). La base de données que nous avons utilisé, BRULEX (Content, *et al.*, 1990), propose des fréquences d'occurrences sur un total de 23.5 millions de mots. Bien que l'estimation des fréquences pour les mots de classe fermée peut ne pas avoir été affectée par de ce type de corpus, c'est un débat actuel que de déterminer si la répartition des fréquences pour les mots de classe ouverte déterminée à partir de ces bases de données restreintes, permettent de capturer l'étendue des différentes fréquences d'occurrence. D'autre part, Osterhout, *et al.* (2002) et Osterhout, *et al.* (1997) réalisent les analyses de régression sur 10 électrodes incluant des électrodes localisées en région postérieure, alors que King & Kutas (1998) réalisent leur régression sur une seule électrode (F7). Ainsi, les différences dans les bases de données ainsi que le nombre et/ou l'emplacement des électrodes sur lesquelles portent les régressions, peuvent avoir entraîné les différences dans les pourcentages de variance expliquée entre notre recherche et ceux obtenus dans la littérature précédente (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997).

#### 4.4 Les prépositions, des mots de classe fermée ?

Lors de notre étude sur les différentes catégories grammaticales, nous avons mis en évidence que les négativités précoces sont latéralisées sur les sites antérieurs droits et centraux pour les trois catégories grammaticales de mots CO (verbes, adverbes, et noms), tandis que celles des mots de classe fermée semblent moins uniformes en ce qui concerne les différentes catégories (prépositions, quantificateurs, non-quantificateurs). En effet, nous avons montré que les prépositions suscitaient l'occurrence d'une composante négative de même topographie (sites antérieurs) que les quantificateurs et les non-quantificateurs, mais d'amplitude restreinte par rapport à ceux-ci. Les catégories grammaticales des mots de classe ouverte présentent des

négativités précoces ayant la même topographie malgré des différences de fréquence (entre les noms d'une part et les verbes et adverbess d'autre part, voir tableau 12). La fréquence d'occurrence ne semble donc pas affecter la topographie des négativités dans cette fenêtre temporelle (250-350 ms) pour les catégories grammaticales des mots CO. Les prépositions, ayant une fréquence intermédiaire entre les quantificateurs et les non-quantificateurs, présentent une amplitude différente de celles obtenues à partir des deux autres catégories grammaticales. Ces résultats suggèrent donc que la fréquence d'occurrence ne semble pas non plus affecter la topographie des négativités précoces des mots de la classe fermée.

Nous avons vu, dans la section 3 du chapitre 1 que la catégorisation des prépositions en tant que mots de classe fermée ne semble pas si évidente. En effet, Garrett & Kean (1980) montrent que les erreurs de production spontanée portant sur les prépositions, présentent des caractéristiques typiques à la fois des erreurs sur les mots de classe ouverte et des erreurs sur les mots de classe fermée. Ce double comportement témoigne de l'ambiguïté du rôle des prépositions dans la structure syntaxique de la phrase (dans l'étude de la construction syntaxique des phrases, les prépositions sont considérées comme lexicales, ou appartenant à la classe ouverte) mais aussi de leur statut phonologique (les prépositions présentent une structure phonologique similaire aux mots de classe fermée). Bien que les recherches en psycholinguistique et en linguistique indiquent une possible différence dans le traitement des prépositions (Cann, 2000; Friederici, 1985; Froud, 2001), nos résultats portant sur la fenêtre d'analyse de la N280, suggèrent plutôt des similitudes entre les 3 catégories grammaticales de la classe fermée. Puisque nous avons trouvé la même topographie pour les 3 catégories grammaticales des mots de classe fermée, nous pouvons conclure que, dans ce rang de latence, la même étape de traitement a été activée. Si l'activation d'une configuration spécifique de neurones induit une carte de surface, alors nous pouvons suggérer que les mêmes générateurs ou groupes de générateurs ont été activés durant cette étape, mais que leur puissance n'a pas été engagée au même titre (il est possible aussi qu'une combinaison de générateurs ait été activée mais différemment). Ce résultat ne remet pas en cause la possibilité d'une différence dans le traitement des prépositions, comparées aux autres catégories grammaticales de la classe fermée. En effet, il semble possible de suggérer que ces différences interviennent plus tardivement dans les processus de reconnaissance du mot.

## 5 Conclusions

Pour finir, nous indiquerons que nos résultats répliquent ceux obtenus dans la littérature sur l'effet de classe mots (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). En revanche, comme nous l'avons déjà souligné plusieurs fois, il semble impossible de déterminer si les effets (topographies et latences) obtenus sur le scalp sont uniquement liés à des différences dans la reconnaissance des deux classes de mots, ou s'ils sont associés à des différences entre les propriétés lexicales, telles que la fréquence et la longueur qui sont observées entre nos deux classes de stimuli. La recherche présentée dans le chapitre suivant va nous permettre de dissocier les effets liés aux propriétés lexicales d'un mot de ceux liés à la classe du mot.

# CHAPITRE 5

## Effet de Classe et Propriétés Lexicales Contrôlées en Contexte

---

### *Résumé*

Le chapitre présent investigate la reconnaissance des mots de classe fermée et ouverte, d'une part lorsque ceux-ci sont intégrés dans un contexte de phrase, et d'autre part lorsque les propriétés lexicales sont contrôlées (fréquence d'occurrence et longueur). Les potentiels évoqués sont enregistrés à partir de 61 électrodes alors que les participants lisent les phrases. L'analyse des tracés indique des différences de topographie et de latence entre les PEs des deux classes de mots dans la fenêtre comprise entre 250 et 350 ms (correspondant à la composante N280). Une composante négative est enregistrée sur les régions antéro-médianes pour les deux classes de mots: alors que cette négativité est maximale sur les régions à droite pour les mots CF, elle est maximale à gauche pour les mots CO. De plus, la composante N280 est plus précoce pour les mots CF que celle associée aux mots CO. L'analyse de segmentation corrobore ces résultats (différence de topographie et de latence) mais suggère que deux micro-états précoces présentent des différences de topographie: le segment 6 (compris entre 199 et 256 ms pour CF et entre 201 et 250 ms pour CO, négativité antérieure) et le segment 7 (compris entre 257 et 338 ms pour CF et entre 251 et 334 ms pour CO, négativité antéro-médiane). L'analyse séparée de l'effet de classe pour les mots de haute et de basse fréquences indique que les effets précoces de topographie (segment 6 et 7) et de latence (segment 7 uniquement) sont liés à la reconnaissance des mots de classe fermée de haute fréquence seulement. Ces résultats suggèrent que dans un contexte de phrase l'accès au lexique des mots de classe fermée de haute fréquence présente d'une part, des distributions de champs de potentiels sur le scalp différents (effet de topographie) mais d'autre part, des décalages dans le temps (effet de latence) par rapport aux mots de classe ouverte de fréquence équivalente (> 400 occurrences/million).

## 1 Introduction

Une des distinctions les plus évidentes entre les deux classes de mots réside dans leur différence de fréquence d'occurrence et de longueur: les mots de classe ouverte tendent à être plus longs et de fréquence moins élevée par rapport aux mots de classe fermée. La fréquence d'un mot semble affecter les processus d'accès lexicaux (Chambers & Forster, 1975; Forster & Chambers, 1973; Howes & Solomon, 1951; Scarborough, *et al.*, 1979) par conséquent, les modèles psycholinguistiques implémentent cette propriété lors de l'accès au lexique (Forster, 1976; Glanzer & Ehrenreich, 1979; Morton, 1969). Puisque les propriétés lexicales (fréquence et longueur) sont intimement liées à la classe de mots, les différences de traitement entre les deux classes de mots, observées à partir de nombreux paradigmes, peuvent en réalité être les conséquences de différences dans les propriétés lexicales. Ainsi, déterminer si l'information codée par la classe de mots permet d'activer des processus différents lors de l'accès au lexique est capital pour l'élaboration des modèles de compréhension du langage.

Les recherches portant sur la distinction entre classe de mots indiquent avec la technique des PEs des différences électrophysiologiques débutant dès 200 ms jusqu'à 800 ms (Brown, *et al.*, 1999; Garnsey, 1985; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Neville, *et al.*, 1992; Nobre & McCarthy, 1994; Osterhout, 1997; Osterhout, *et al.*, 2002; Pulvermüller, *et al.*, 1995). Alors qu'une composante négative précoce, la N280, a été avancée comme reflétant le traitement spécifique des mots de classe fermée (Neville, *et al.*, 1992), les résultats de certains auteurs (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997) montrent que la latence d'apparition de cette composante N280 varie en fonction de la fréquence d'occurrence des mots. Dans notre chapitre 4, nous avons montré des résultats qui soutiennent cette interprétation: la latence de la composante N280 semble effectivement être influencée non seulement par la fréquence d'occurrence mais aussi par la longueur des mots <sup>14</sup>.

---

<sup>14</sup> Nous avons aussi montré une corrélation entre la classe de mots et la latence de la composante N280.

Ces différents résultats suggèrent donc que la N280 ne reflète pas, ou en tout cas pas uniquement, l'activation de processus spécifiques à la classe de mots fermée (King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). Au contraire, les résultats de certains (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999) suggèrent une indépendance de la latence de la composante N280 face aux propriétés lexicales, et donc que la classe de mots est l'information pertinente pour caractériser la précocité de cette composante.

Alors que la fréquence d'occurrence affecte les réponses comportementales mais aussi les PEs (Besson & Kutas, 1993; Kutas & van Petten, 1994; Rugg, 1990 ; Rugg & Coles, 1995), Neville, *et al.* (1992) et Osterhout, *et al.* (2002) montrent que la longueur est aussi une propriété des mots modulant les PEs. L'ambiguïté des résultats issus des analyses de régressions provient du fait que la classe de mots, la fréquence d'occurrence et la longueur sont intimement liées. Ainsi, dans le rang de fréquence élevée et de longueur courte, seuls des mots de classe fermée sont présents (voir figure 32, chapitre 4). Afin de déterminer si les effets observés sur le scalp sont indépendants des propriétés lexicales de ces mots, à savoir leur fréquence d'occurrence et leur longueur, il semble nécessaire de pouvoir former deux groupes de mots présentant les mêmes propriétés lexicales (fréquences et longueurs) mais de classes différentes. Deux études ont tenté de relever ce challenge (Garnsey, 1985; Pulvermüller, *et al.*, 1995)<sup>15</sup> et présentent des données confortant l'idée que malgré un contrôle des propriétés lexicales, les PEs reflètent la distinction entre classe de mots. Alors que Garnsey (1985) montre des différences tardives (entre 350 et 600 ms), Pulvermüller, *et al.* (1995) présentent l'évidence d'une distinction dans le traitement des deux classes dès 160 ms après la présentation des mots de classe fermée. Alors que seulement deux expériences (Garnsey, 1985; Pulvermüller, *et al.*, 1995) ont contrôlé les propriétés lexicales des deux classes de mots, aucune n'a utilisé les mots dans un contexte de phrase ou de texte. Ce point peut en réalité apporter des réserves quant à l'interprétation des résultats. Les linguistes et psycholinguistes s'accordent en général pour reconnaître que les deux classes de mots jouent des rôles différents au sein de la phrase. Par exemple, Garrett (1975) indique que lors de la production, les mots de classe fermée sont utilisés pour la construction syntaxique de la phrase alors que les mots de classe ouverte sont intégrés plus tardivement dans cette structure (voir chapitre 1 section 4.1.1.2).

---

<sup>15</sup> Nous n'incluons pas l'étude de (Nobre & McCarthy, 1994) dans ce groupe, car seule la longueur était contrôlée.

Si ces différences apparaissent au niveau de la phrase, il semble primordial d'examiner les processus de traitement de ces mots lorsqu'ils sont dans le contexte d'une phrase, et non pas seulement lorsqu'ils sont isolés. A l'heure actuelle aucune recherche portant sur l'effet de classe de mots n'a présenté des items contrôlés en fréquence et longueur dans un contexte de phrase. Le premier objectif de ce chapitre est de combler cette carence dans la littérature. Nous allons contrôler la fréquence d'occurrence et la longueur des deux classes de mots et présenter ces mots dans le contexte de phrases à des positions fixes.

Les travaux de Bradley (1978) et de Bradley, *et al.* (1980) suggèrent une absence de sensibilité pour les mots de classe fermée à la fréquence d'occurrence, alors que les mots de classe ouverte y seraient sensibles. Comme nous l'avons indiqué (chapitre 1, section 4.2), ces résultats ont suscité bon nombre de critiques dans la littérature et ont été à l'origine d'une série d'études apportant des résultats contradictoires (Gordon, 1983; Gordon & Caramazza, 1982; Gordon & Caramazza, 1985; Segui, *et al.*, 1987; Segui, *et al.*, 1982). De plus, Gordon (1983) et Gordon & Caramazza (1985) montrent qu'il existe un effet de seuil ou de saturation dans les relations linéaires entre le log de fréquence et le temps de réaction pour une tâche de décision lexicale, et ce, pour les deux classes de mots. Ainsi, en dessous de 2,6 log de fréquence ( $< 400$  occurrences/million), la linéarité semble conservée pour les deux classes de mots alors qu'au-dessus de cette limite les patterns observés se différencient pour les deux classes de mots. En effet, au-dessus de 2,6 log de fréquence, les mots de classe fermée apparaissent relativement insensibles à la fréquence tandis que les mots de classe ouverte continuent de montrer une certaine variabilité en fonction de la fréquence (Gordon, 1983; Gordon & Caramazza, 1985). Comme très peu de mots de classe fermée sont de basse fréquence ( $< 2,6$  log de fréquence), il est fort probable que les recherches n'ayant pas contrôlé la fréquence utilisent des mots de classe fermée de très haute fréquence, ce qui expliquerait pourquoi un effet de fréquence est présent pour les mots de classe ouverte mais absent pour ceux de classe fermée.

Face à ces conclusions, nous avons mis en place notre expérience afin de manipuler la fréquence d'occurrence lors de l'étude de la distinction entre classes de mots. Ainsi, nous allons comparer l'effet de classe pour des mots de basse fréquence ( $< 2,6$  log de fréquence) et de haute fréquence ( $> 2,6$  log de fréquence). Les mots de classe ouverte et fermée ont été sélectionnés et répartis en deux sous-catégories: des mots de haute et de basse fréquence.

Comme les propriétés lexicales seront contrôlées dans cette expérience, en comparant avec les résultats du chapitre 4, nous pourrions identifier précisément les patterns PEs spécifiques à la distinction entre classe de mots. Ainsi, nous émettons la première hypothèse que les deux classes de mots vont susciter l'occurrence de composantes précoces, négatives, distinctes (topographie) sur le scalp. Ce résultat indiquera que les mots de classe fermée et de classe ouverte sont reconnus de manière différente dans un contexte de phrases. En ce qui concerne la latence d'occurrence de la composante précoce, nous avons montré que non seulement elle corrèle avec la fréquence d'occurrence et la longueur d'un mot, mais aussi avec la classe de celui-ci. Dans notre chapitre 4, ainsi que dans les études (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997), il était impossible de différencier les trois propriétés. Puisque dans cette recherche les propriétés lexicales des mots seront contrôlées, nous allons pouvoir indiquer précisément si la latence de la composante N280 varie indépendamment des propriétés lexicales. Nous suggérons que la latence de la N280 est affectée par les 3 propriétés (classe, fréquence et longueur). Par conséquent lorsque seule la classe de mot est étudiée, la latence sera plus précoce pour les mots de classe fermée que celle des mots de classe ouverte.

Notre deuxième hypothèse est que les processus associés aux mots de classe fermée sont différents en fonction du rang de fréquence étudiée, alors que celui des mots de classe ouverte ne diffère pas (Gordon, 1983; Gordon & Caramazza, 1985). Selon Gordon (1983) et Gordon & Caramazza (1985) seuls les mots de classe fermée se comportent différemment en fonction du rang de fréquence étudié<sup>16</sup>. Autrement dit, le traitement des mots de classe fermée diffère en fonction du rang de fréquence étudiée, alors que celui des mots de classe ouverte reste similaire quelque soit la fréquence étudiée. Si tel est le cas, alors nous devrions pouvoir mettre en évidence des effets de classes (distinction entre mots de classe ouverte et fermée) différents pour ces deux rangs de fréquence (haute et basse fréquence). Nous allons alors pouvoir enregistrer des différences de topographie et/ou de latence de composantes spécifiques pour chaque rang de fréquence d'occurrence utilisée (> et < à 2,6 log de fréquence).

<sup>16</sup> L'observation d'un effet de seuil pour les mots de haute fréquence de classe fermée uniquement est liée au fait que très peu de mots de classe ouverte sont distribués dans les fréquences supérieures à 400 occurrences/million. Par conséquent ces mots n'atteignent pas le niveau de saturation.



## 2 Méthode

### 2.1 Sujets

Vingt étudiants de la Faculté de Psychologie et des Sciences de l'Éducation de Genève, dont 19 femmes, ont participé bénévolement à cette expérience (moyenne d'âge de 21,9 compris entre 19 et 31). Tous les sujets étaient droitiers, selon le test de latéralité *Edinburgh Handedness Inventory* (Oldfield, 1971), de langue maternelle française, et présentaient une vision normale ou corrigée.

### 2.2 Stimuli

Un corpus de 160 mots a été constitué (Palix, 2000). Le choix de ces mots a été réalisé à partir de la base de donnée BRULEX (Content, *et al.*, 1990) répertoriant 35 000 mots. Ainsi, nous avons sélectionné 80 mots de classe fermée (CF) et 80 autres de classe ouverte (CO). La fréquence d'utilisation ainsi que la longueur de chaque mot ont été contrôlées. La fréquence d'utilisation fournie par BRULEX correspond au nombre d'occurrences d'une chaîne de caractères dans un échantillonnage de textes (littéraires, romans, essais, etc.) constitué à partir de 23,5 millions de mots. Ce nombre est rapporté à un total de 100 millions pour obtenir l'indice de fréquence. Chaque fréquence a été par la suite rapportée à 1 million, puis transformée en log de fréquence, échelle plus usitée dans la littérature (King & Kutas, 1998; Osterhout, *et al.*, 1997). Une échelle logarithmique permet d'accentuer l'importance des basses fréquences ce qu'une échelle linéaire ne peut faire. Par exemple, la différence entre deux mots, un de fréquence 2 et l'autre de fréquence 1 est plus importante que celle entre deux mots de fréquence 2002 et 2001 (pour la première paire la différence est de 0.3003 alors que pour la seconde cette différence est de 0.0002). Puisque dans une langue il existe un nombre plus conséquent de mots de basse fréquence, il semble judicieux d'utiliser cette échelle afin de les dissocier correctement. La longueur des mots a aussi été contrôlée en terme de nombre de lettres. La liste complète des mots est rapportée en annexe 3. Les caractéristiques générales des stimuli sont présentées dans le tableau 13. L'analyse de variance confirme qu'il n'y a pas de différences de fréquence ( $F < 1$ ) ni de longueur ( $F < 1$ ) entre nos deux classes de mots. De plus, au sein de chaque classe de mots, nous avons sous-divisé les 80 items en mots de haute fréquence (compris entre 2,7 et 3,6 log de fréquence, > 400 occurrences/million) et de basse fréquence (compris entre 1,5 et 2,6 log de fréquence, < 400 occurrences/million).

Classe de mot	Nombre d'items	Descriptif	Log de fréq	Longueur
CO	80	61.25 % Noms	2.58 (.05)	4.83 (.12)
		22.5 % Verbes		
		16.25 % Adjectifs		
CF	80	51.25 % Prépositions	2.59 (.06)	4.76 (.16)
		18.75 % Pronoms		
		13.75 % Conjonctions		
		8.75 % Déterminants		
		7.5 % Locutions-Interjections		

**Tableau 13:** Caractéristiques des mots choisis pour la classe ouverte (CO) et fermée (CF). Chaque classe présentait un nombre équivalent d'items (80), de fréquence d'utilisation et de nombre de lettres équivalents. Entre parenthèses est inscrite l'erreur type. La composition en catégorie grammaticale pour chaque classe de mot est décrite.

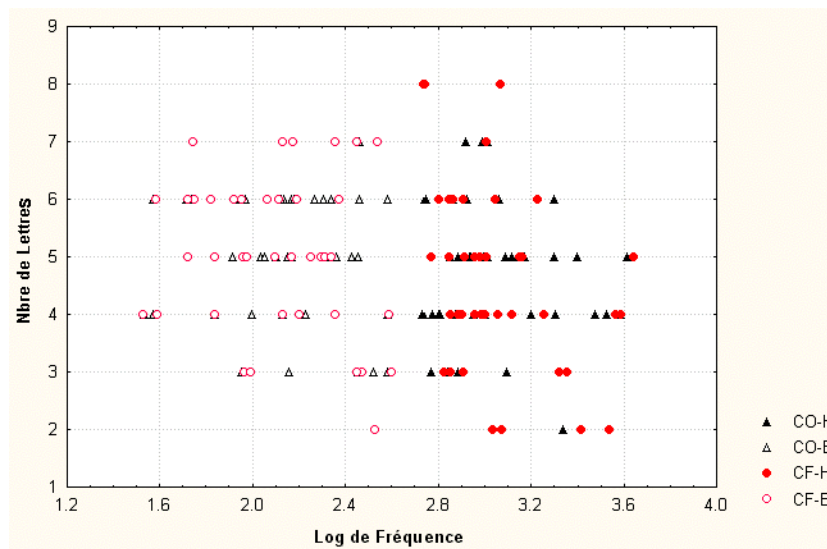
Ainsi, 4 catégories différentes ont été créées: mots de classe ouverte de basse fréquence (CO-B, 40 items), mots de classe ouverte de haute fréquence (CO-H, 40 items), mots de classe fermée de basse fréquence (CF-B, 40 items) et mots de classe fermée de haute fréquence (CF-H, 40 items). Les caractéristiques générales des stimuli sont rapportées dans le tableau 14.

	Nombre de mots	Classe de mots	Noms	Log de fréq	Longueur
Haute Fréquence	40	CO	CO-H	3.05 (.03)	4.67 (.18)
	40	CF	CF-H	3.07 (.04)	4.47 (.23)
Basse Fréquence	40	CO	CO-B	2.12 (.04)	5.00 (.17)
	40	CF	CF-B	2.10 (.03)	5.05 (.21)

**Tableau 14:** Descriptif des différentes fréquences et longueurs pour chaque sous-groupe de classe de mots. Entre parenthèses est inscrite l'erreur type.

Deux analyses de variance séparées ont été réalisées. La première avait comme facteurs "classe de mots" (ouverte vs. fermée) et "fréquence" (haute vs. basse) et révèle que l'effet du facteur "classe de mots" n'est pas significatif ( $F < 1$ ) alors que l'effet de la "fréquence" l'est ( $F(1,156) = 467.7, p < .001$ ) sans interaction entre "classe de mots" x "fréquence" ( $F < 1$ ). Les mots de haute fréquence sont plus fréquents que les mots de basse fréquence, effet significatif à la fois pour les mots de classe fermée et de classe ouverte.

La deuxième analyse de variance a été réalisée sur la variable dépendante longueur (nombre de lettres) et nous a permis de vérifier si les mots de haute fréquence présentaient des longueurs plus courtes comparés aux mots de basse fréquence. Une analyse avec comme facteurs "classe de mots" (ouverte vs. fermée) et "fréquence" (haute fréquence vs. basse fréquence) a été réalisée. L'effet du facteur "classe de mots" n'est pas significatif ( $F < 1$ ) alors que l'effet du facteur "fréquence" l'est ( $F(1,156) = 4.96, p < .02$ ): les mots de haute fréquence sont non seulement plus fréquents que ceux de basse fréquence, mais ils sont aussi plus courts. L'interaction "classe de mots" x "fréquence" n'est pas significative ( $F < 1$ ): les mots de classe ouverte de haute fréquence (CO-H) présentent la même longueur que les mots de classe fermée haute fréquence (CF-H). Il n'y a pas de différence de longueur entre les mots de classe ouverte et fermée de basse fréquence. La répartition de ces stimuli en fonction de leur longueur, de leur log de fréquence ainsi que de leur classe est illustrée dans la figure 33.



**Figure 33:** Répartition des stimuli du chapitre 5 entre mots de classe fermée (CF, cercles vides et pleins) et mots de classe ouverte (CO, triangles vides et pleins). Comme l'analyse statistique le montre, les mots de classe fermée de haute fréquence (CF-H, cercles pleins) sont plus fréquents et plus courts que les mots de classe fermée de basse fréquence (CF-B, cercles vides). Cet effet est aussi observé pour les mots de classe ouverte (CO-H, triangles pleins, CO-B, triangles vides). Aucune différence de log de fréquence ou de longueur n'est significative entre les mots CF-H (cercles pleins) et CO-H (triangles pleins) ainsi qu'entre les mots CF-B (cercles vides) et CO-B (triangles vides).

## 2.3 Procédure

Les mots expérimentaux (classe ouverte et classe fermée) ont été insérés dans des phrases correctes à une position fixe: le 4<sup>ème</sup> mot. La liste complète des phrases utilisées est présentée dans l'annexe 4. Les participants lisaient chaque phrase présentée mot par mot au centre de l'écran d'ordinateur. Chaque mot était visible pendant 300 ms, avec un intervalle de 500 ms (*Stimuli Onset Asynchrony*, SOA de 800 ms). Chaque phrase commençait avec une croix centrale de fixation (300 ms). Le premier mot de chaque phrase était écrit avec une majuscule, et toutes les autres lettres utilisées étaient des minuscules. Le mot final de chaque phrase était suivi d'un point final. La tâche du participant était de lire les phrases et d'en comprendre le sens. Pour 20 % des phrases, nous avons présenté après le mot final une question de compréhension à choix forcé (réponse sur le boîtier de réponse). Par exemple, pour la présentation de la phrase "Ta maman a de jolis cheveux blonds", la question de compréhension était "Ta maman est... 1) blonde 2) brune. Ce paradigme nous a permis de contrôler si les participants réalisaient correctement la tâche de lecture et s'ils analysaient les phrases en profondeur. A la fin de chacune des phrases, les participants visualisaient un écran vide pendant 2000 ms qui leur permettait de cligner des yeux et de se reposer. Les phrases étaient randomisées et divisées en deux blocs de 80 phrases d'une durée approximative de 8 minutes chacune. Avant le début de l'expérience, le participant lisait la consigne écrite et effectuait une phase d'apprentissage d'une dizaine de phrases.

## 2.4 Enregistrement de l'EEG

Les caractéristiques de l'enregistrement sont identiques à celles utilisées dans les chapitres 3 et 4 (voir section 2.4 chapitre 3). La procédure de moyennage a produit 6 jeux de PEs pour chaque participant: CO, CF, CF-H, CO-H, CF-B, CO-B. Dans un premier temps, nous allons étudier l'effet de classe général lorsque les fréquences et longueurs sont contrôlées (CO vs. CF). Nous analyserons ensuite l'effet de classe pour les mots de haute fréquence (CO-H vs. CF-H) et enfin l'effet de classe pour les mots de basse fréquence (CO-B vs. CF-B).

### 3 Résultats

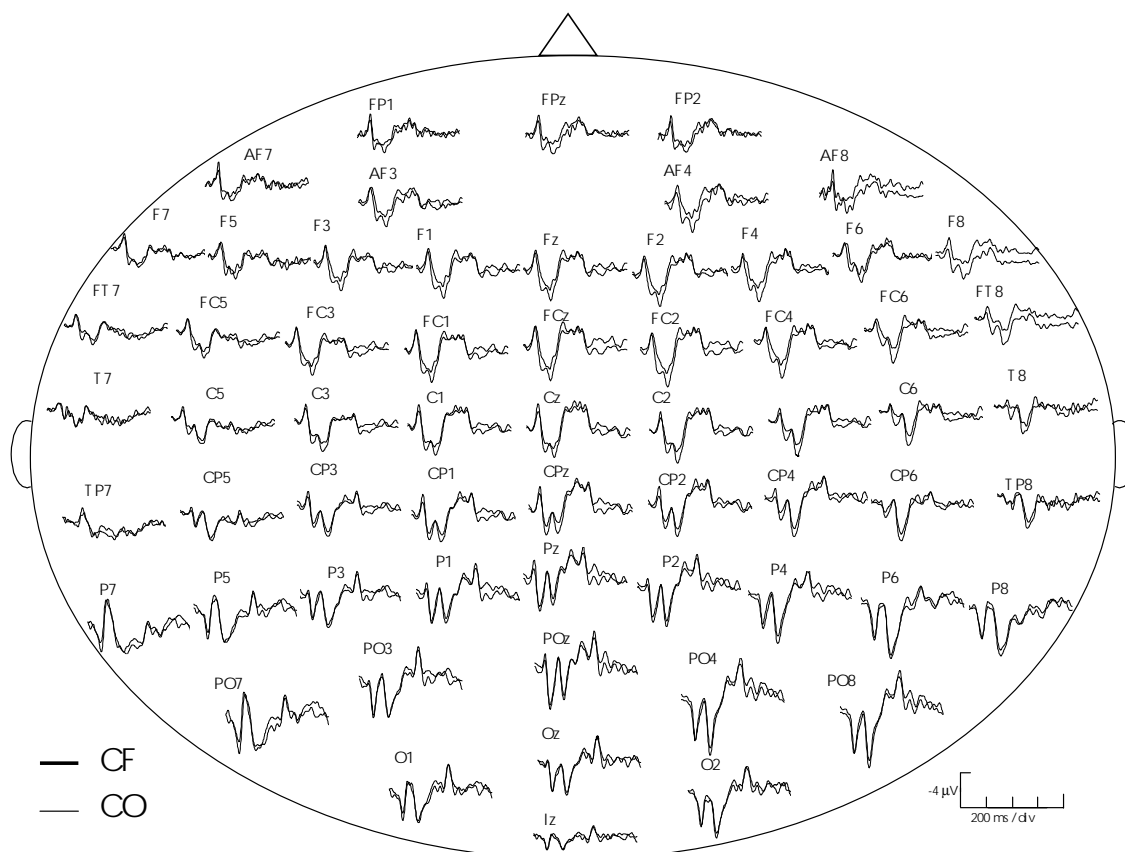
#### 3.1 Comparaison CF vs. CO

##### 3.1.1 Analyse des tracés

Pour ces analyses, les PEs sont référencés à l'électrode située sur le mastoïde gauche.

##### 3.1.1.1 Visualisation des tracés

La figure 34 illustre les PEs obtenus pour les mots de classe ouverte (CO) et fermée (CF).



**Figure 34:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 61 électrodes enregistrées. Les tracés en gras correspondent aux PEs des mots de classe fermée (CF), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte (CO). Les électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, O pour occipital et I pour inion. La négativité est située vers le haut.

Visuellement, les tracés révèlent les composantes que nous avons décrites précédemment à savoir: P100, N2 et P2 (section 4.2 du chapitre 3, et section 3.1.1 chapitre 4). Comparés aux tracés obtenus pour les deux classes de mots lorsque les propriétés lexicales ne sont pas contrôlées (figure 19), les différences sont plus restreintes à la fois en amplitude, mais aussi en distribution sur le scalp. Dès l'occurrence de la P2, on remarque que les PEs des mots de classe fermée (CF) sont plus négatifs (autrement dit, la composante P2 est moins importante) en région frontale que ceux des mots de classe ouverte (CO). Dans une fenêtre plus tardive, nous pouvons observer que les PEs des mots CF sont plus négatifs que ceux des mots CO, effet qui semble plus important sur les sites antérieurs droits.

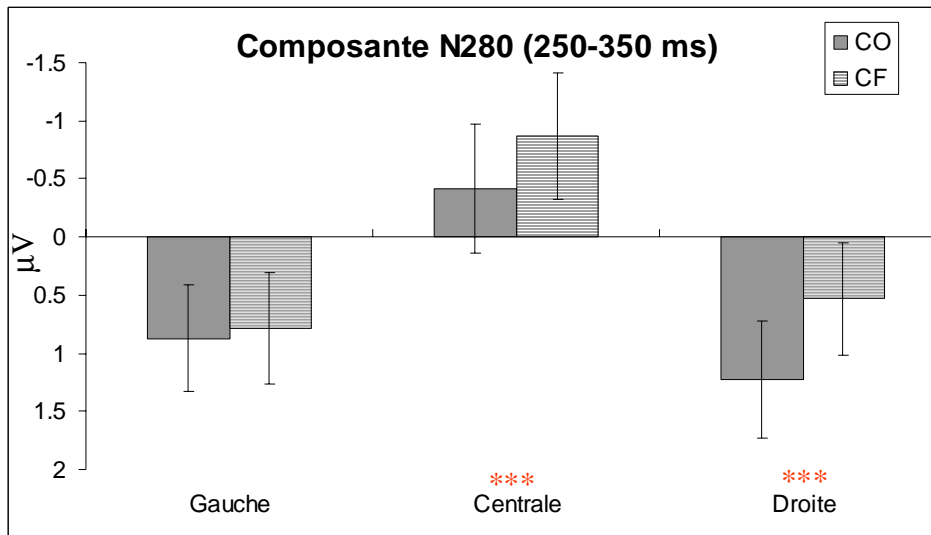
### 3.1.1.2 Analyse de topographie par région d'intérêt

Après une analyse visuelle des composantes, les 4 fenêtres temporelles définies pour les chapitres 3 et 4 ont été conservées pour réaliser nos analyses statistiques: la P2 (150 et 250 ms), la N280 (250 et 350 ms), la N400 (350 et 500), et la LAN (500 et 800 ms). Les résultats des ANOVAs à mesures répétées séparées pour chaque fenêtre temporelle sont présentés dans le tableau 15.

	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	3.45 - /	1.51 - /	0.01 - /	1.96 - /
Classe x Caudalité	2, 38	2.24 - 1.77	4.32 * - 2.16	0.45 - 0.01	0.44 - 0.80
Classe x Latéralité	2, 38	5.61 ** - 1.95	<b>7.10 ** - 4.93 *</b>	0.71 - 2.07	2.40 - 1.75
Classe x Caudalité x Latéralité	4, 76	0.73 - 0.63	0.15 - 0.36	0.80 - 1.50	0.68 - 0.94

**Tableau 15:** Résultats des ANOVAs à mesures répétées pour les composantes P2 (150-250 ms), N280 (250-350 ms), N400 (350-500 ms), et LAN (500-800 ms) sur l'amplitude du PEs entre les deux classes de mots (CO vs. CF). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Seule une interaction entre les facteurs "classe de mots" x "latéralité" est significative et ce dans la fenêtre temporelle de la composante N280 (250-350 ms;  $F(2,38) = 4.93$ ,  $p < .01$ ). Dans cette fenêtre, une composante négative présente un maximum d'activité en régions antéro-médianes ("caudalité":  $F(2,38) = 12.75$ ;  $p < .0001$ ) et centrales ("latéralité":  $F(2,38) = 16.66$ ;  $p < .0001$ ). Les comparaisons planifiées de l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 4.93$ ;  $p < .01$ ) indiquent que les PEs associés aux mots de classe fermée sont plus négatifs que ceux des mots CO sur les régions centrales (CF:  $-0.86 \mu\text{V}$ , CO:  $-0.41 \mu\text{V}$ ,  $p < .0003$ ) et droites (CF:  $0.52 \mu\text{V}$ , CO:  $1.22 \mu\text{V}$ ,  $p < .0003$ ; voir figure 35).



**Figure 35:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour l'interaction "classe de mots" x "latéralité" pour la composante N280 (comprise entre 250 et 350 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 3 régions d'intérêt: gauche, centrale et droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

L'analyse de tracés indique que les différences de topographie sont enregistrées dans une seule fenêtre temporelle restreinte à celle de la composante N280 (250-350 ms). Alors que les deux classes de mots suscitent l'occurrence d'une composante négative dans la fenêtre d'analyse comprise entre 250 et 350 ms, maximale en région antéro-médiane, les mots CF suscitent l'occurrence d'une négativité plus importante en région centrale et à droite par rapport aux mots CO.

### 3.1.1.3 Analyse de latence de la N280

Nous avons reconduit la même analyse que celle décrite dans la section 3.1.1.3 du chapitre 4. La latence de la composante N280 a été analysée sur les maxima de négativité de 14 électrodes incluses dans les régions antérieures (FP1, FP2, AF7, AF3, AF4, AF8, F7, F5, F3, F1, F2, F4, F6, F8, en excluant FPz, et Fz). Une ANOVA à mesures répétées avec les facteurs "classe de mots" (CF vs. CO), "hémisphère" (droit vs. gauche) et "électrodes" (7) montre un effet significatif du facteur "classe de mots" ( $F(1,19) = 7.79, p < .01$ ): la latence de la composante N280 pour les mots de classe fermée (CF:  $298.5 \pm 3.4$  ms) est plus précoce (de 7.2 ms) que celle enregistrée pour les mots de classe ouverte (CO:  $305.7 \pm 4.1$  ms). De plus, le facteur "hémisphère" est significatif ( $F(1,19) = 1.85, p < .03$ ) sans interaction avec le facteur "classe de mots" ( $F < 1$ ): la latence de la composante négative pour les deux classes de mots est plus précoce pour les électrodes localisées sur l'hémisphère gauche (CF:  $296.7 \pm 3.5$  ms,

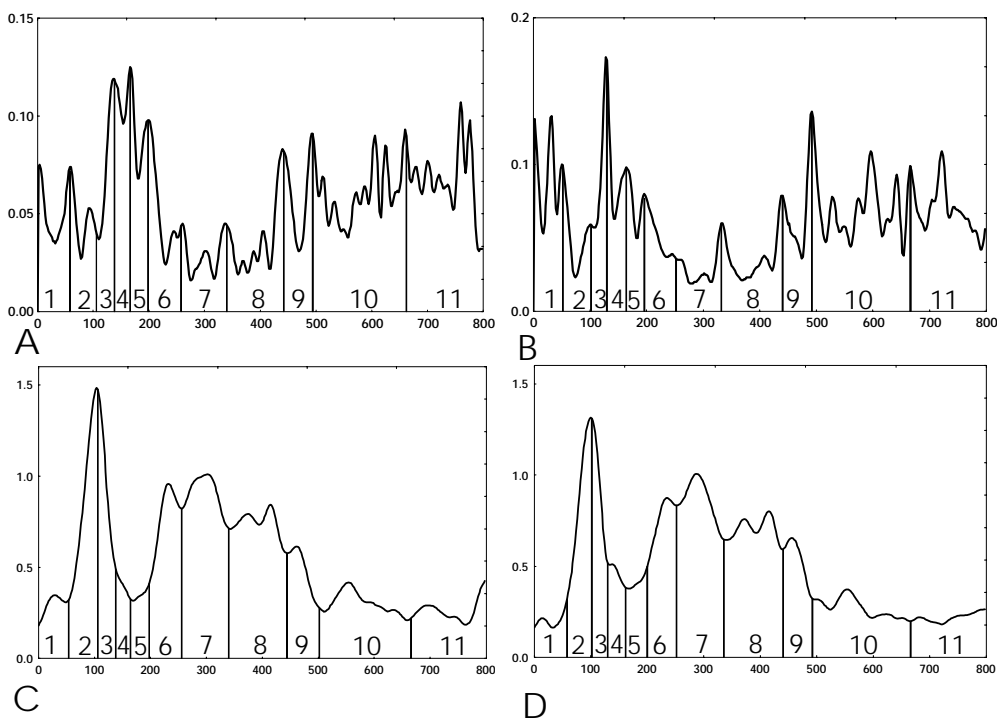
CO:  $303.7 \pm 4.1$  ms) que celles distribuées sur l'hémisphère droit (CF:  $300.3 \pm 3.3$  ms, CO:  $307.7 \pm 3.9$  ms).

### 3.1.2 Analyse spatio-temporelle

Les analyses spatio-temporelles sont effectuées sur les PEs calculés à partir de la référence moyenne.

#### 3.1.2.1 Détermination des micro-états

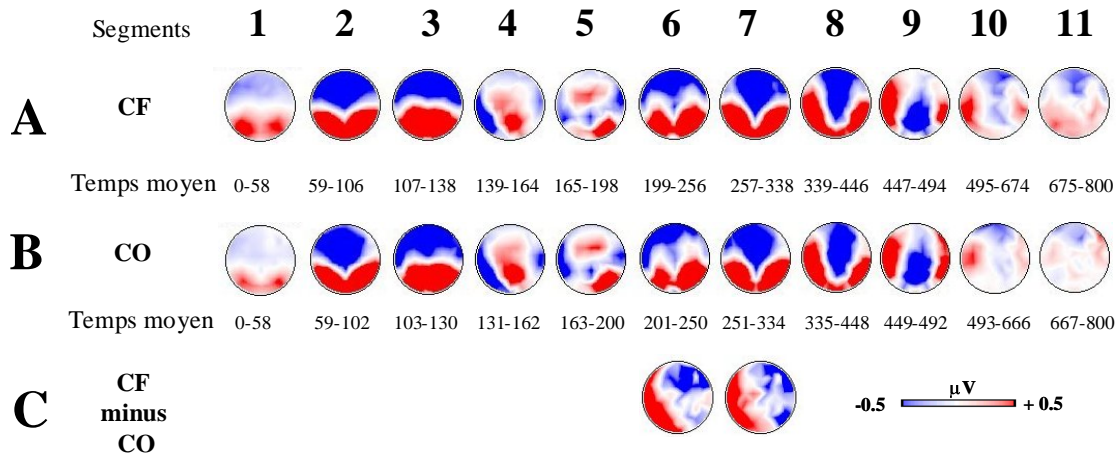
Sur la période de 800 ms correspondant à la fenêtre d'analyse, 11 périodes de stabilité ont été déterminées séparément pour chaque classe de mots (voir figure 36).



**Figure 36:** (A & B) courbes de *Global Dissimilarity* (GD) pour les grandes moyennes (800 ms d'analyse), pour les PEs associés aux mots de classe fermée (A) et de classe ouverte (B). (C & D), courbes représentant le *Global Field Power* (GFP) pour les grandes moyennes pour les mots de classe fermée (C) et de classe ouverte (D). Les traits verticaux indiquent les frontières d'un segment défini par la méthode d'analyse de segmentation adaptée. Les nombres se réfèrent à l'ordre d'apparition des segments ou micro-états dans le temps. La segmentation adaptée révèle que pour chaque condition expérimentale, 11 configurations différentes et stables sont présentes.



Pour chaque segment, une seule carte moyenne est calculée à partir des PEs des grandes moyennes, et ce dans les fenêtres temporelles définies par la segmentation adaptée. De plus, à partir de la grande moyenne, les frontières de chaque segment sont définies (voir figure 37).



**Figure 37:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales (A, classe fermée: CF; B, classe ouverte: CO) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. La couleur bleue représente des PEs négatifs, alors que les PEs positifs sont représentés en rouge. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, sont représentées les cartes de différences de potentiels (classe fermée minus classe ouverte) pour les segments 6, 7 où les topographies sont significativement différentes entre les deux conditions expérimentales.

### 3.1.2.2 Détermination des micro-états spécifiques

Afin de tester si les cartes des PEs associées aux mots de classe fermée sont spécifiques à cette condition expérimentale, chaque carte de surface de la condition CF a été introduite dans les données individuelles des sujets pour les mots de classe fermée, mais aussi pour ceux de classe ouverte. La procédure de "fitting" a été réalisée séparément pour chaque segment, et dans une fenêtre temporelle définie sur la base des valeurs ( $\pm 10$  ms) obtenues sur la grande moyenne (segment 1: 0-70 ms, segment 2: 50-120 ms, segment 3: 100-150 ms, segment 4: 120-180 ms, segment 5: 150-220 ms, segment 6: 190-270 ms, segment 7: 240-350 ms, segment 8: 320-460 ms, segment 9: 440-510 ms, segment 10: 480-690 ms, segment 11: 590-800 ms). La valeur du BEV (*Best Explained Variance* ou maximum de pourcentage de variance expliquée) pour chaque carte de la condition CF obtenue à partir de la procédure de "fitting" dans les deux conditions expérimentales de chaque participant est alors comparée statistiquement à l'aide d'un *t*-test à échantillon apparié. Les résultats montrent que 2 cartes

(segments 6 et 7) présentent des topographies significativement différentes pour les mots de classe fermée et ouverte. La carte représentative du segment 6 rend mieux compte des données CF ( $62.1 \pm .04$  %) par rapport aux données CO ( $53,7 \pm .05$  %;  $t = -2.15$ ,  $p < .04$ ). Le même résultat est obtenu pour le segment 7 ( $67.7 \pm .04$  % pour les PEs des mots CF,  $59.8 \pm .04$  % pour les PEs des mots CO;  $t = -2.58$ ;  $p < .01$ ). Aucun autre segment ne montre d'effet significatif. Les latences des frontières pour les 2 segments spécifiques, ainsi que leurs topographies de surface sont illustrées dans la figure 37.

### 3.1.2.3 Analyse topographique par région d'intérêt

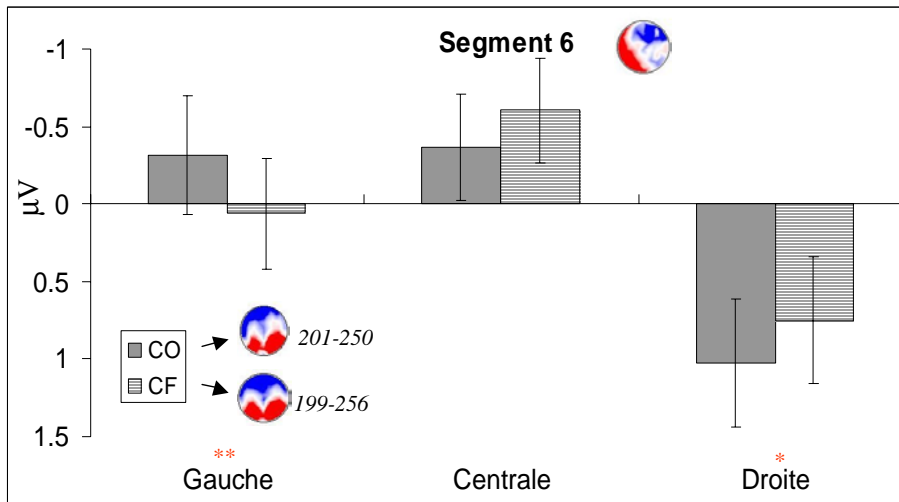
Pour chaque segment qui est spécifique (6 et 7) aux mots de classe fermée, nous avons effectué une ANOVA à mesures répétées avec les facteurs suivants: "classe de mots" (fermée vs. ouverte); "caudalité" (antérieur vs. médian vs. postérieur); "latéralité" (gauche vs. centrale vs. droite). Les régions d'intérêt utilisées sont décrites dans le tableau 3 du chapitre 3. Les résultats de ces analyses sont présentés dans le tableau 16.

	ddl	6	7
Classe x Caudalité	2, 38	2.86 - 2.30	0.96 - 0.44
Classe x Latéralité	2, 38	<b>7.51 ** - 4.31 *</b>	<b>8.69 *** - 6.70 **</b>
Classe x Caudalité x Latéralité	4, 76	0.42 - 0.81	0.46 - 0.32

**Tableau 16:** Résultats des ANOVAs à mesures répétées séparées pour chaque segment (6 et 7). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

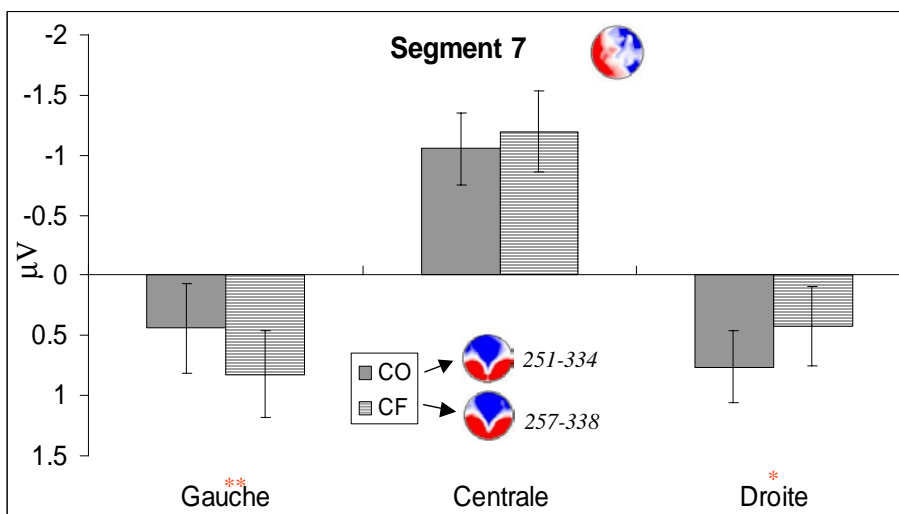
Afin d'éclaircir les différents effets observés, nous allons étudier séparément chaque segment.

- Segment 6 (CF: 199-256 ms, CO: 201-250 ms): une composante négative est enregistrée sur les régions antérieures ("caudalité": ( $F(2,38) = 21.02$ ;  $p < .0001$ ) et maximale à gauche et en central ("latéralité": ( $F(2,38) = 8.09$ ;  $p < .001$ ). L'analyse de l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 4.32$ ;  $p < .02$ ) montre deux effets différents sur le scalp. Alors que les PEs des mots CF sont plus négatifs à droite ( $0.74 \mu V$ ) par rapport aux mots CO ( $1.02 \mu V$ ,  $p < .04$ ), ceux-ci sont moins négatifs (ou plus positifs) à gauche (CF:  $0.06 \mu V$ , CO:  $-0.31$ ;  $p < .007$ ). Aucune différence n'est observée en région centrale (voir figure 38).



**Figure 38:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 6 (compris entre 201-250 ms pour CO et 199-256 ms pour CF). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 3 régions d'intérêt: gauche, centrale et droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF et CO (en bas) ainsi que la carte de différences (CF minus CO, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

- Segment 7 (CF: 257-338 ms, CO: 251-334 ms): la composante négative enregistrée dans ce segment présente un maximum d'activité en région antéro-médiane ("caudalité":  $F(2,38) = 31.97$ ;  $p < .0001$ ) et centrale ("latéralité":  $F(2,38) = 21.52$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 6.71$ ;  $p < .003$ ) montre deux effets différents sur le scalp (voir figure 39). Alors que les PEs des mots CF sont plus négatifs à droite ( $0.42 \mu\text{V}$ ) par rapport aux mots CO ( $0.76 \mu\text{V}$ ,  $p < .01$ ), ceux-ci sont plus positifs à gauche (CF:  $0.82 \mu\text{V}$ ; CO:  $0.43$ ,  $p < .004$ ). Aucune différence n'est observée en région centrale ( $p > .05$ ).



**Figure 39:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 7 (compris entre 251-334 ms pour CO et 257-338 ms pour CF). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 3 régions d'intérêt: gauche, centrale et droite. Les barres verticales

indiquent l'erreur type. Les cartes de surface sont illustrées pour CF et CO (en bas) ainsi que la carte de différences (CF minus CO, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.1.2.4 Evolution temporelle des micro-états

Les analyses sur la latence des pics de GFP ont été réalisées pour les segments 6, 7, 8 et 9. Une ANOVA à mesures répétées pour chaque segment avec comme facteur "classe de mots" (fermée vs. ouverte) a été réalisée. Les résultats sont décrits dans le tableau 17. Une seule carte apparaît plus précocement pour les mots de classe fermée par rapport aux mots de classe ouverte. Ainsi, la carte caractérisant le segment 7 (comprise entre 257-338 ms pour les mots de classe fermée et entre 251-334 ms pour les mots de classe ouverte) est significativement plus précoce pour les mots de classe fermée que pour les mots de classe ouverte: la différence de latence est de 10.5 ms.

Segment	Classe Fermée	Classe Ouverte	ANOVA
6	233.3 (4.37)	238.1 (3.71)	$F(1,19) = 1.01; p > .32$
7	285.7 (4.27)	296.2 (4.80)	<b><math>F(1,19) = 7.45; p &lt; .01</math></b>
8	405.7 (3.80)	406 (4.24)	$F(1,19) = .01; p > .91$
9	468.5 (3.94)	463.1 (3.78)	$F(1,19) = 1.79; p > .19$

**Tableau 17:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments (6, 7, 8 et 9). Les valeurs indiquées sont en ms, avec entre parenthèses l'erreur type. Les effets significatifs sont rapportés en gras. Pour le segment 7, la carte associée aux mots CF est enregistrée plus précocement (-10.5 ms) que celle obtenue pour les mots CO.

### 3.1.3 Synthèse CF vs. CO

Grâce à une analyse de tracés nous avons montré que l'effet de classe de mots (lorsque les propriétés lexicales sont contrôlées) est restreint dans la fenêtre d'analyse de la composante N280 (250-350 ms). Bien que les deux classes de mots suscitent l'occurrence d'une composante négative maximale en région antéro-médiane, des différences de latéralisation sont mises en évidence: la négativité associée aux mots CF est plus importante sur les régions centrale et droite par rapport à celle des mots CO qui est maximale sur les régions à gauche. De plus, l'analyse de latence indique que cette composante N280 est plus précoce de 7.2 ms pour les mots CF. Sur la base des variations de topographie, l'analyse de segmentation adaptée a tout d'abord permis de diviser la fenêtre d'analyse de la composante N280 en deux micro-états. En effet, alors que pour un premier segment 6 (199-256 ms pour CF, 201-250 ms pour CO) les deux classes de mots présentent des négativités maximales en régions

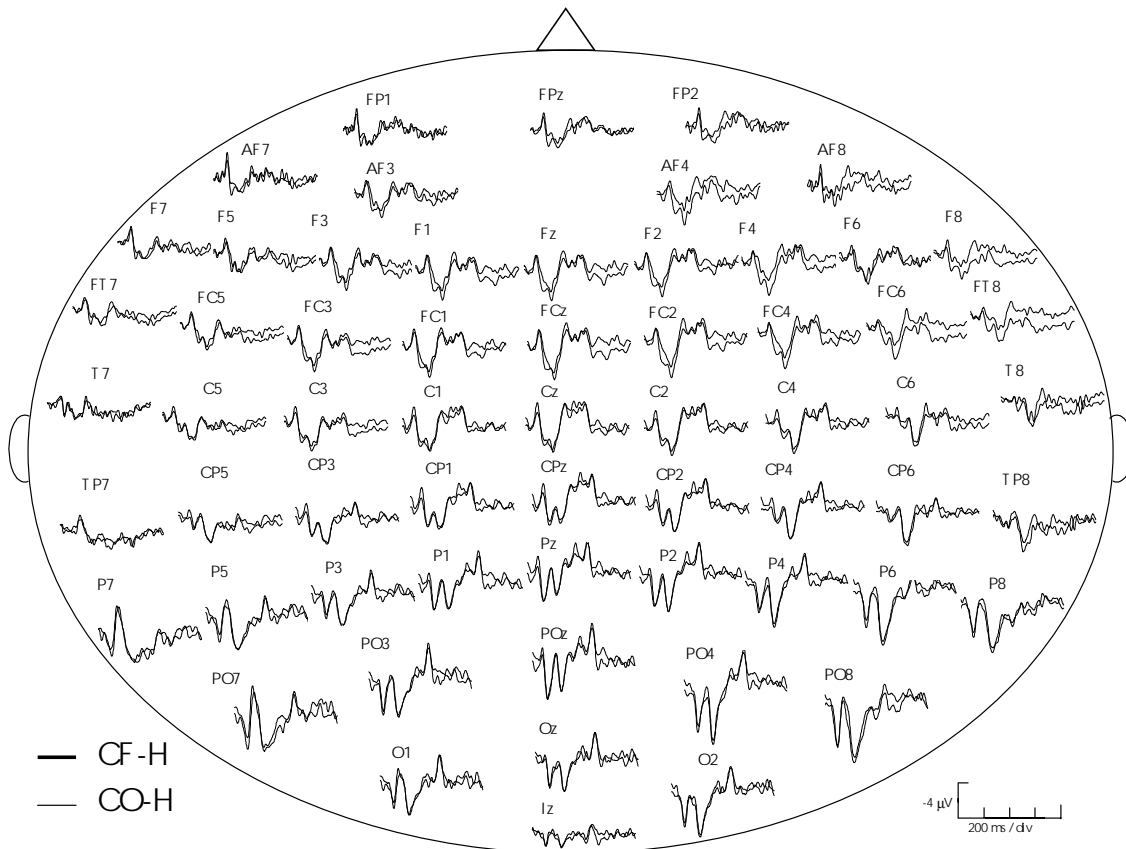
antérieures, ces négativités pour le segment 7 (257-338 ms pour CF, et 251-334 ms pour CO) s'étendent aux régions antérieure et médiane. Pour ces deux segments, des différences de topographies sont présentes entre les deux classes de mots. Pour le segment 6, la négativité des mots CF est plus importante à droite alors que celle des mots CO est maximale en régions gauche et centrale. Pour le segment 7, les PEs des mots CF sont plus négatifs à droite alors que pour les mots CO, la négativité est maximale en région centrale. De plus l'analyse du maximum de GFP indique qu'un seul segment (7) est enregistré plus précocement (10.5 ms) pour les mots CF.

## 3.2 Comparaison CF vs. CO haute fréquence

### 3.2.1 Analyses des tracés

Pour ces analyses, les PEs sont référencés à l'électrode située sur la mastoïde gauche.

#### 3.2.1.1 Visualisation des tracés



**Figure 40:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 61 électrodes enregistrées. Les tracés en gras correspondent aux PEs des mots de classe fermée de haute fréquence (CF-H), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte de haute fréquence (CO-H). Les

électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sont distribuées sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, O pour occipital et I pour inion. La négativité est située vers le haut.

La figure 40 présente les PEs pour les mots de haute fréquence (fréquence d'occurrence > 2.6 log de fréquence) de classe fermée (CF-H) et de classe ouverte (CO-H). Par rapport à la figure 34 qui illustre les PEs des mots CO et CF lorsque des hautes et basses fréquences sont moyennées ensemble, les effets semblent similaires dans le temps, mais aussi dans la répartition sur le scalp, malgré une certaine diminution en amplitude.

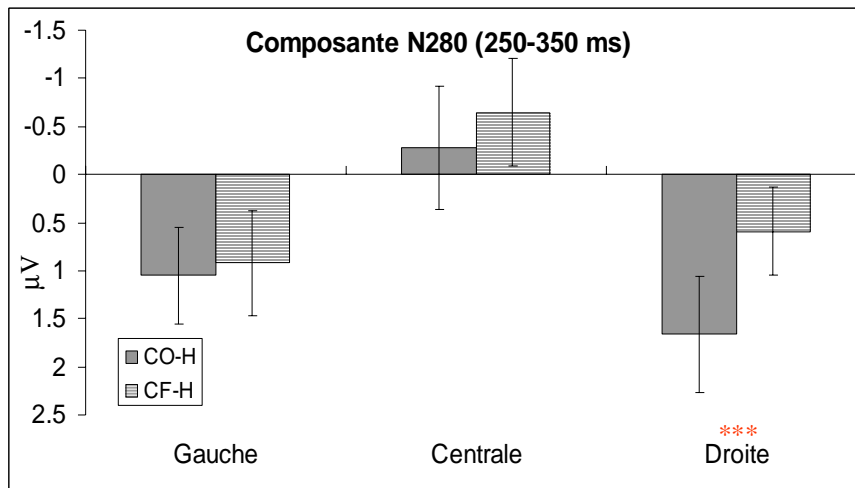
### 3.2.1.2 Analyse de topographie par région d'intérêt

Les résultats des ANOVAs à mesures répétées, pour les 4 fenêtres temporelles: P2 (150 et 250 ms), N280 (250 et 350 ms), N400 (350 et 500 ms), et LAN (500 et 800 ms), sont présentés dans le tableau 18.

	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	2.11 - /	2.53 - /	0.42 - /	2.50 - /
Classe x Caudalité	2, 38	0.72 - <i>1.03</i>	1.75 - <i>2.47</i>	0.30 - <i>0.01</i>	1.01 - <i>1.15</i>
Classe x Latéralité	2, 38	2.01 - <i>0.89</i>	<b>5.28 **</b> - <b>5.37 **</b>	1.39 - <i>1.20</i>	0.82 - <i>0.44</i>
Classe x Caudalité x Latéralité	4, 76	0.50 - <i>0.29</i>	0.68 - <i>0.37</i>	0.56 - <i>0.47</i>	0.56 - <i>0.58</i>

**Tableau 18:** Résultats des ANOVAs à mesures répétées pour les composantes P2 (150-250 ms), N280 (250-350 ms), N400 (350-500 ms), et LAN (500-800 ms) sur l'amplitude du PE entre les deux classes de mots de haute fréquence (CO-H vs. CF-H). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Seule une interaction "classe de mots" x "latéralité" est significative et ce, dans le rang de latence compris entre 250 et 350 ms correspondant à la composante N280 ( $F(2,38) = 5.38$ ;  $p < .008$ ). Dans ce rang de latence une composante est enregistrée avec un maximum de négativité sur les régions antéro-médianes ("caudalité":  $F(2,38) = 13.72$ ;  $p < .0001$ ) et centrales ("latéralité":  $F(2,38) = 16.19$ ;  $p < .00001$ ). Les comparaisons planifiées de l'interaction "classe de mots" x "latéralité" (voir figure 41) indiquent que les PEs associés aux mots de classe fermée sont moins positifs que ceux des mots CO sur l'hémisphère droit (CF-H:  $0.58 \mu\text{V}$ , CO-H:  $1.66 \mu\text{V}$ ;  $p < .00001$ ). Aucune différence n'est significative pour les régions centrale et gauche ( $p > .05$ ).



**Figure 41:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour l'interaction "classe de mots" x "latéralité" dans la fenêtre comprise entre 250 et 350 ms. Les valeurs du PE (en µV) pour les mots de classe fermée (CF-H) et ouverte (CO-H) de haute fréquence sont indiquées pour les 3 régions d'intérêt: gauche, centrale, droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

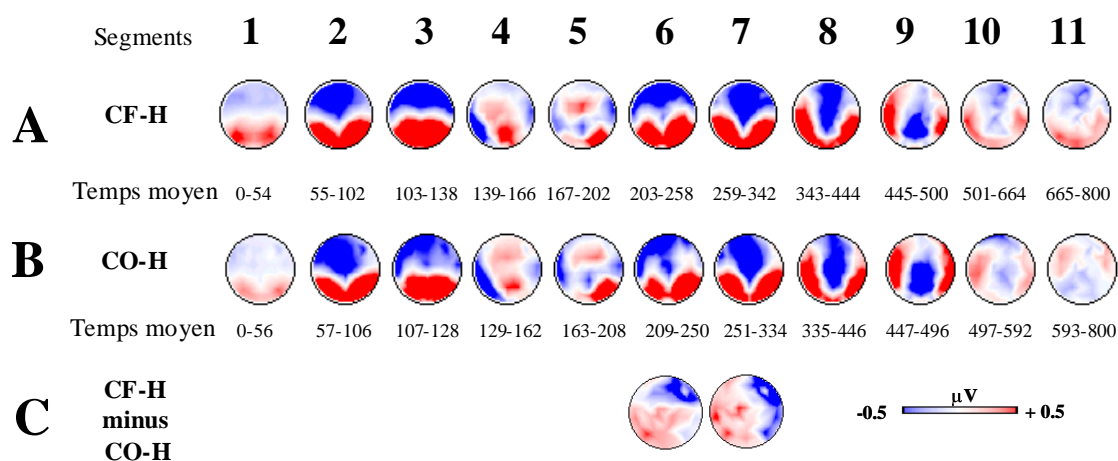
### 3.2.1.3 Analyse de latence de la N280

Les résultats de l'ANOVA à mesures répétées avec les facteurs "classe de mots" (CF-H vs. CO-H), "hémisphère" (droit vs. gauche) et "électrodes" (7) montrent un effet significatif du facteur "classe de mots" ( $F(1,19) = 5.73, p < .02$ ): la latence de la composante N280 pour les mots de classe fermée (CF-H:  $293.1 \pm 3.6$  ms) est plus précoce (de 8.6 ms) que celle enregistrée pour les mots de classe ouverte (CO-H:  $301.7 \pm 3.3$  ms). De plus, le facteur "hémisphère" est significatif ( $F(1,19) = 8.51, p < .009$ ) sans interaction avec le facteur "classe de mots" ( $F(1,19) = 1.3, p > .3$ ): la latence de la composante négative pour les deux classes de mots est plus précoce pour les électrodes localisées sur l'hémisphère gauche (CF-H:  $291.5 \pm 3.4$  ms, CO-H:  $301.1 \pm 3.3$  ms) que celles placées sur l'hémisphère droit (CF-H:  $294.7 \pm 3.7$  ms, CO-H:  $302.4 \pm 3.3$  ms).

## 3.2.2 Analyse spatio-temporelle

### 3.2.2.1 Détermination des micro-états

La procédure de segmentation adaptée a déterminé 11 segments différents pour chaque condition expérimentale. Les segments ainsi que leur carte moyenne et leurs frontières (en ms) sont présentés dans la figure 42.



**Figure 42:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales de haute fréquence (A, classe fermée: CF-H; B, classe ouverte: CO-H) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. La couleur bleue représente des PEs négatifs, alors que les PEs positifs sont représentés en rouge. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, sont représentées les cartes de différence de potentiels (CF-H minus CO-H) pour les segments 6, 7 où les topographies sont significativement différentes entre les deux conditions expérimentales.

### 3.2.2.2 Détermination des micro-états spécifiques

La procédure de "fitting" a été réalisée comme décrite dans le paragraphe 3.1.2.2 de ce chapitre. Les résultats montrent que 2 cartes (segment 6 et 7) présentent des topographies significativement différentes entre les mots de classe fermée et ouverte de haute fréquence. La carte représentative du segment 6 explique mieux les données CF-H ( $64.4 \pm .03$  %) par rapport aux données CO-H ( $57,4 \pm .04$  %;  $t = -2.43$ ,  $p < .02$ ). Le même résultat est obtenu pour le segment 7 ( $68.1 \pm .04$  % pour les PEs des mots CF-H,  $61.4 \pm .04$  % pour les PEs des mots CO-H;  $t = -2.32$ ;  $p < .03$ ). Aucun autre segment ne montre d'effet significatif. Les latences des frontières pour les 2 segments spécifiques, ainsi que leurs topographies de surface, sont illustrées dans la figure 42.

### 3.2.2.3 Analyse topographique par région d'intérêt

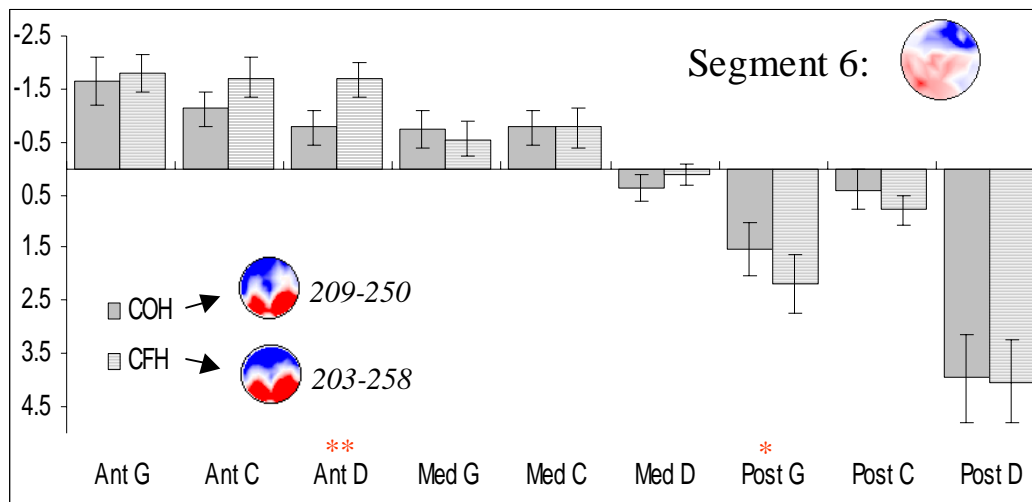
Les résultats des ANOVAs à mesures répétées pour les deux segments spécifiques (6 et 7) sont présentés dans le tableau 19.



	ddl	6	7
Classe x Caudalité	2, 38	2.08 - 1.88	0.73 - 1.07
Classe x Latéralité	2, 38	1.18 - 1.26	<b>6.26 ** - 6.91 **</b>
Classe x Caudalité x Latéralité	4, 76	0.23 - 0.11	0.48 - 0.27

**Tableau 19:** Résultats des ANOVAs à mesures répétées séparées pour chaque segment (6 et 7). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

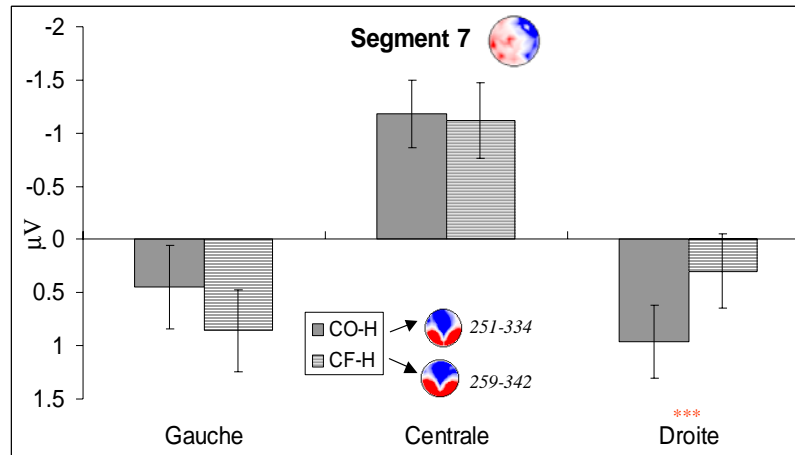
Pour le segment 6 (CF-H: 203-258 ms, CO-H: 209-250 ms), une composante négative maximale en région antérieure ("caudalité":  $F(2,38) = 26.12$ ;  $p < .0001$ ) à gauche et en central ("latéralité":  $F(2,38) = 9.53$ ;  $p < .0004$ ) est enregistrée. Alors que la procédure de "fitting" indiquait des différences de topographie pour le micro-état 6 (compris entre 209-250 ms pour CO-H et 203-258 ms pour CF-H), les résultats de l'ANOVA à mesures répétées n'indiquent aucune interaction significative entre les facteurs "classe de mots" et "caudalité" et/ou "latéralité". Afin d'élucider le problème, nous avons conduit des ANOVAs à mesures répétées séparées pour chaque région d'intérêt avec comme facteur "classe de mots" (CF-H vs. CO-H). Les résultats de cette analyse sont présentés dans la figure 43: les PEs des mots CF-H sont plus négatifs ( $-1.68 \mu V$ ) en région antérieure droite par rapport aux mots CO-H ( $-0.77 \mu V$ ,  $p < .007$ ) et plus positifs en postérieure gauche (CF-H:  $2.18 \mu V$ , CO-H:  $1.53 \mu V$ ;  $p < .04$ ).



**Figure 43:** Résultats des ANOVAs à mesures répétées séparées pour chaque région d'intérêt (ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite) pour le segment 6 (compris entre 209-250 ms pour CO-H et 203-258 ms pour CF-H). Les valeurs des PEs sont en  $\mu V$ , les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-H et CO-H (en bas) ainsi que la carte de différences (CF-H minus CO-H, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Pour le segment 7 (CF-H: 259-342 ms, CO-H: 251-334 ms), la négativité enregistrée est maximale en régions antéro-médianes ("caudalité":  $F(2,38) = 34.95$ ;  $p < .00001$ ) centrales

("latéralité":  $F(2,38) = 20.13$ ;  $p < .0001$ ). L'analyse (*Post-Hoc* LSD, voir figure 44) de l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 6.91$ ;  $p < .002$ , voir tableau 19) montre que les PEs des mots CF sont plus négatifs à droite ( $0.42 \mu\text{V}$ ) par rapport aux mots CO ( $0.76 \mu\text{V}$ ,  $p < .01$ ). Aucune différence n'est observée en région centrale ou encore à droite ( $p > .05$ ).



**Figure 44:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 7 (compris entre 251-334 ms pour CO-H et 259-342 ms pour CF-H). Les valeurs des PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-H) et ouverte (CO-H) sont indiquées pour les 3 régions d'intérêt: gauche, centrale et droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-H et CO-H (en bas) ainsi que la carte de différences (CF-H minus CO-H, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.2.2.4 Evolution temporelle des micro-états

Les analyses sur la latence des pics de GFP ont été réalisées pour les segments 6, 7, 8 et 9. Une ANOVA à mesures répétées pour chaque segment avec comme facteur "classe de mots" (fermée *vs.* ouverte) a été effectuée pour les mots de haute fréquence. Les résultats sont décrits dans le tableau 20. Une seule carte apparaît plus précocement pour les mots de classe fermée (CF-H) par rapport aux mots de classe ouverte (CO-H). Il s'agit de la carte caractérisant le segment 7 (compris entre 251-334 ms pour CO-H et 259-342 ms pour CF-H) qui est significativement plus précoce pour les mots de classe fermée par rapport aux mots de classe ouverte: la différence de latence est de 7.2 ms.

Segment	CF-H	CO-H	ANOVA
6	230.1 (4.48)	232.2 (4.35)	$F(1,19) = .32; p > .57$
7	<b>285.2 (4.19)</b>	<b>292.4 (4.60)</b>	<b><math>F(1,19) = 8.01; p &lt; .01</math></b>
8	406.8 (4.81)	406.5 (5.72)	$F(1,19) = .01; p > .95$
9	465.1 (3.06)	467.1 (3.77)	$F(1,19) = .58; p > .45$

**Tableau 20:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments (6, 7, 8 et 9). Les valeurs indiquées sont en ms, avec entre parenthèse l'erreur type. Les effets significatifs sont rapportés en gras. Pour le segment 7, la carte associée aux mots CF-H est enregistrée plus précocement (-7.2 ms) que celle obtenue pour les mots CO-H.

### 3.2.3 Synthèse CF-H vs. CO-H

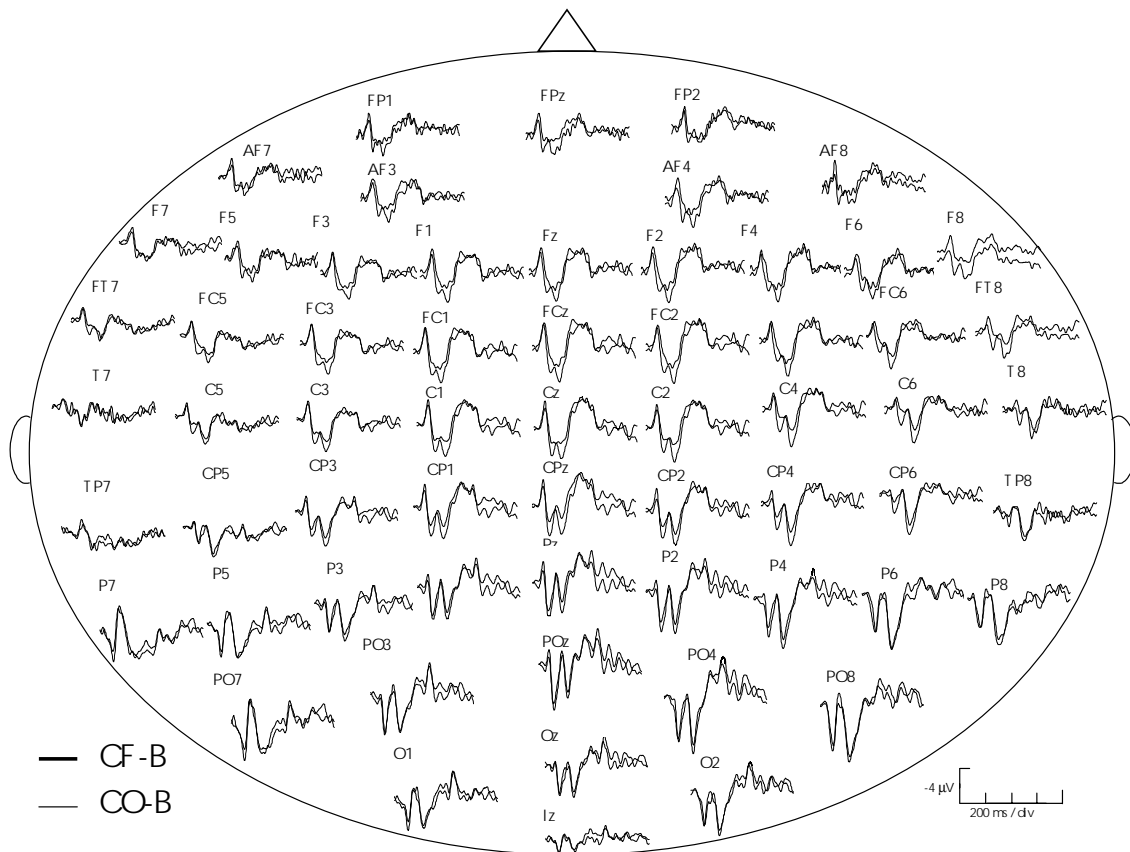
L'analyse portant sur les mots de haute fréquence (> 2.6 log de fréquence) nous indique des résultats similaires à ceux reportés lorsque des mots de haute et de basse fréquences sont moyennés ensemble (CO vs. CF, voir section 3.1 de ce chapitre). L'analyse de tracés indique que seul le rang de latence correspondant à la composante N280 est caractéristique de l'effet de classe: alors que les deux classes de mots suscitent l'occurrence de négativités antéro-médianes, les PEs des mots CF-H sont plus négatifs à droite que ceux des mots CO-H. De plus, la composante N280 est enregistrée 8,6 ms plus précocement pour les mots CF-H que pour les mots CO-H. L'analyse de segmentation adaptée divise cette fenêtre temporelle en deux micro-états (segment 6 et 7) indiquant que pour le premier segment (segment 6), les négativités sont plus focales (région antérieure) puis se distribuent plus largement sur le scalp (régions antérieure et médiane) pour le deuxième segment (7). Des différences de topographies sont enregistrées dans ces deux segments. Pour les segment 6, les PEs des mots CF-H sont plus négatifs en antérieur droit et plus positifs en postérieur gauche que ceux des mots CO-H. Dans la fenêtre correspondant au segment 7, les PEs des mots CF-H sont plus négatifs à droite que ceux des mots CO-H. Pour finir, l'analyse du maximum de GFP indique qu'un seul segment est enregistré plus précocement (7.2 ms) pour les mots CF-H par rapport aux mots CO-H.

### 3.3 Comparaison CF vs. CO basse fréquence

#### 3.3.1 Analyse des tracés

##### 3.3.1.1 Visualisation des tracés

La figure 45 présente les PEs pour les mots de basse fréquence de classe fermée (CF-B) et de classe ouverte (CO-B).



**Figure 45:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 61 électrodes enregistrées. Les tracés en gras correspondent aux PEs des mots de classe fermée de basse fréquence (CF-B), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte de basse fréquence (CO-B). Les électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, O pour occipital et I pour inion. La négativité est située vers le haut.

### 3.3.1.2 Analyse de topographie par région d'intérêt

Les résultats des ANOVAs à mesures répétées pour les 4 fenêtres temporelles P2 (150 et 250 ms), N280 (250 et 350 ms), N400 (350 et 500 ms), et LAN (500 et 800 ms) sont présentés dans le tableau 21.

	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	2.51 - /	0.73 - /	0.12 - /	2.50 - /
Classe x Caudalité	2, 38	1.82 - <i>1.24</i>	3.03 ** - <i>1.01</i>	0.01 - <i>0.22</i>	1.01 - <i>1.15</i>
Classe x Latéralité	2, 38	3.71 * - <i>1.44</i>	3.75 ** - <i>0.55</i>	0.09 - <i>0.27</i>	0.82 - <i>0.44</i>
Classe x Caudalité x Latéralité	4, 76	0.58 - <i>0.63</i>	0.51 - <i>0.48</i>	1.26 - <i>0.89</i>	0.56 - <i>0.58</i>

**Tableau 21:** Résultats des ANOVAs à mesures répétées pour les composantes P2 (150-250 ms), N280 (250-350 ms), N400 (350-500 ms), et LAN (500-800 ms) sur l'amplitude du PEs entre les deux classes de mots de basse fréquence (CO-B vs. CF-B). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Quelle que soit la fenêtre étudiée, les résultats ne montrent pas d'effet de la classe de mots et ce facteur n'interagit avec aucun des autres facteurs étudiés (caudalité et latéralité).

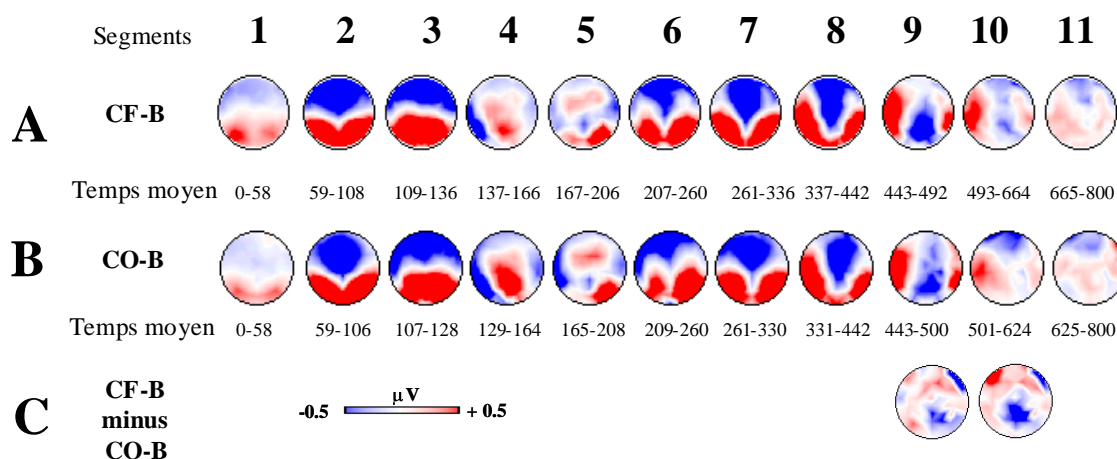
### 3.3.1.3 Analyse de latence de la N280

Les résultats de l'ANOVA à mesures répétées avec les facteurs "classe de mots" (CF-B vs. CO-B), "hémisphère" (droit vs. gauche) et "électrodes" (7) montrent que l'effet du facteur principal "classe de mots" n'est pas significatif ( $F(1,19) = 2.2, p > .16$ ): la latence de la composante négative pour les mots de classe fermée (CF-B:  $291.7 \pm 3.6$  ms) est équivalente à celle enregistrée par les mots de classe ouverte (CO-B:  $296.5 \pm 2.8$  ms). Par ailleurs, le facteur "hémisphère" est significatif ( $F(1,19) = 13.56, p < .0001$ ): la latence de la composante négative pour les deux classes de mots est plus précoce pour les électrodes localisées sur l'hémisphère gauche (CF-B:  $290.2 \pm 3.6$  ms, CO-B:  $294.8 \pm 2.9$  ms) que sur le droit (CF-B:  $293.2 \pm 3.6$  ms, CO-B:  $298.2 \pm 2.7$  ms). Toutefois, cet effet est le même pour les deux classes de mots, comme l'indique l'absence d'interaction entre le facteur "hémisphère" et le facteur "classe de mots" ( $F < 1$ ).

### 3.3.2 Analyse spatio-temporelle

#### 3.3.2.1 Détermination des micro-états

La procédure de segmentation adaptée a déterminé 11 segments différents pour chaque condition expérimentale. Les segments, ainsi que leur carte moyenne et leurs frontières (ms), sont présentés dans la figure 46.



**Figure 46:** Cartes PE de surface des grandes moyennes (N=20) pour les deux conditions expérimentales de basse fréquence (A, Classe fermée: CF-B ; B, Classe ouverte: CO-B) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, sont représentées les cartes de différence de potentiels (CF-B minus CO-B) pour les segments 9 et 10 où les topographies sont significativement différentes entre les deux conditions expérimentales.

#### 3.3.2.2 Détermination des micro-états spécifiques

La procédure de "fitting" a été réalisée comme décrite dans le paragraphe 3.1.2.2 de ce chapitre. Les résultats montrent que 2 cartes (segments 9 et 10) présentent des topographies significativement différentes pour les mots de classe fermée et pour ceux de classe ouverte de basse fréquence. La carte représentative du segment 9 explique mieux les données CF-B ( $43.6 \pm .04$  %) que celles des données CO-B ( $34.1 \pm .04$  %;  $t = -2.28$ ,  $p < .03$ ). Le même résultat est obtenu pour le segment 10 ( $38.1 \pm .03$  % pour les PEs des mots CF-B,  $25.8 \pm .03$  % pour les PEs des mots CO-B;  $t = -3.43$ ;  $p < .002$ ). Aucun autre segment ne montre d'effet significatif. Les latences des 2 segments spécifiques, ainsi que leurs topographies de surface sont illustrées dans la figure 46.

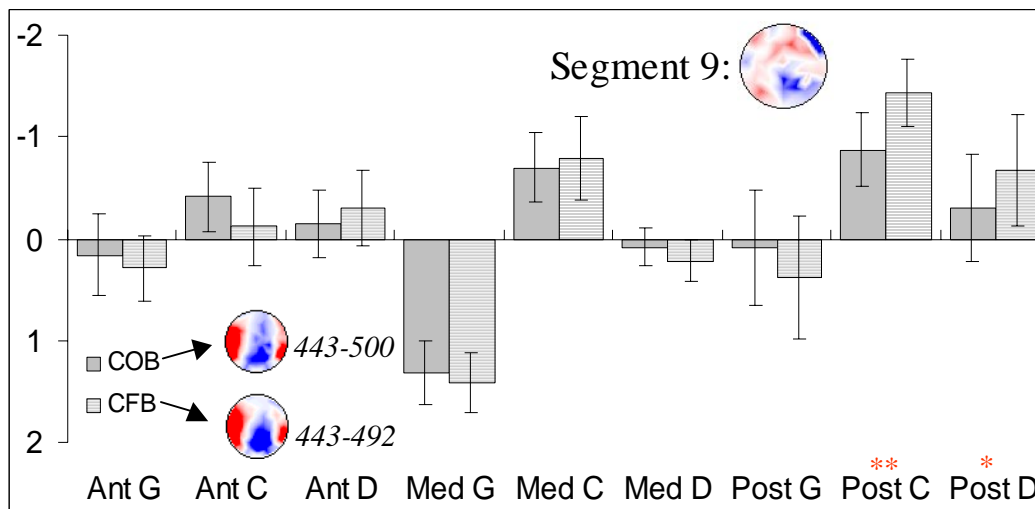
### 3.3.2.3 Analyse topographique par région d'intérêt

Les résultats des ANOVAs à mesures répétées pour les deux segments spécifiques (9 et 10) sont présentés dans le tableau 22.

	ddl	9	10
Classe x Caudalité	2, 38	0.33 - 0.25	2.18 - 1.41
Classe x Latéralité	2, 38	0.59 - 0.68	1.48 - 1.31
Classe x Caudalité x Latéralité	4, 76	<b>2.76 * - 2.45 *</b>	1.96 - 1.37

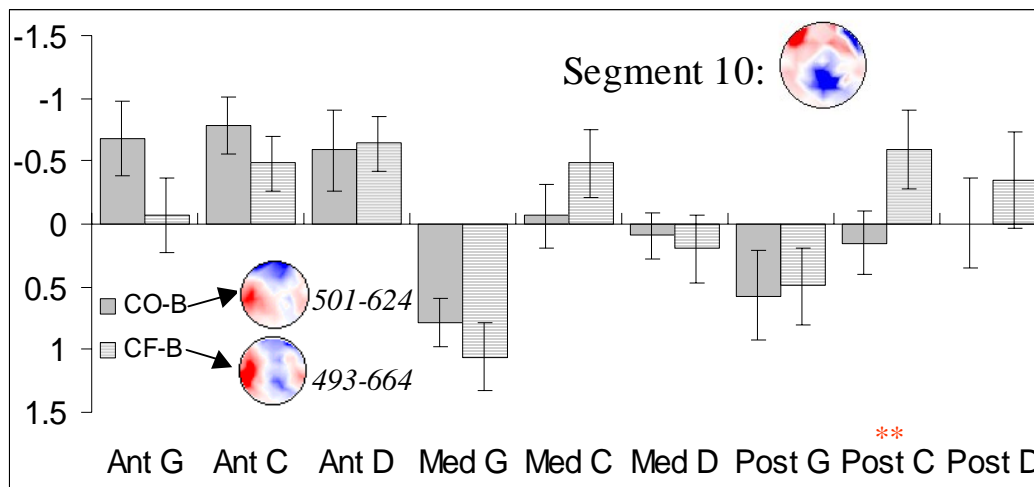
**Tableau 22:** Résultats des ANOVAs à mesures répétées séparées pour chaque segment (9 et 10). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Pour le segment 9 (CO-B: 443-500 ms, CF-B: 443-492 ms), une composante négative est enregistrée sur les régions centrale et droite ("latéralité":  $F(2,38) = 9.52$ ;  $p < .0004$ ). L'analyse de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 2.45$ ;  $p < .05$ ) indique que les PEs des mots CF-B sont plus négatifs que ceux des mots CO-B en régions postérieures centrale (CF-B:  $-1.42 \mu\text{V}$ , CO-B:  $-0.87 \mu\text{V}$ ;  $p < .002$ ) et droite (CF-B:  $-0.66 \mu\text{V}$ , CO-B:  $-0.29 \mu\text{V}$ ;  $p < .04$ ; voir figure 47).



**Figure 47:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 9 (compris entre 443-500 ms pour CO-B et 443-492 ms pour CF-B). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-B et CO-B (en bas) ainsi que la carte de différences (CF-B minus CO-B, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Pour le segment 10 (CO-B: 501-624 ms, CF-B: 493-664 ms), une composante négative est maximale sur les régions antérieures ("caudalité" :  $F(2,38) = 3.39$ ;  $p < .04$ ) centrales et droites ("latéralité":  $F(2,38) = 6.42$ ;  $p < .004$ ). Le tableau 22 indique que pour ce segment, aucune interaction n'est significative avec le facteur "classe de mots". Afin de vérifier les résultats de la procédure de "fitting", nous avons réalisé des ANOVAs à mesures répétées séparées pour chaque région d'intérêt avec comme facteur "classe de mots" (CF-B vs. CO-B). Les résultats de cette analyse sont présentés dans la figure 48. Le seul effet notable est que les PEs des mots CF-B sont plus négatifs en région postérieure centrale ( $-0.59 \mu\text{V}$ ) par rapport à ceux des mots CO-B ( $0.15 \mu\text{V}$ ,  $p < .002$ ).



**Figure 48:** Résultats des ANOVAs à mesures répétées séparées pour chaque région d'intérêt (ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite) pour le segment 10 (compris entre 501-624 ms pour CO-B et 493-664 ms pour CF-B). Les valeurs du PEs sont en  $\mu\text{V}$ , les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-B et CO-B (en bas) ainsi que la carte de différences (CF-B minus CO-B, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.3.2.4 Evolution temporelle des micro-états

Les analyses sur la latence des pics de GFP ont été réalisées pour les segments 6, 7, 8 et 9. Une ANOVA à mesures répétées pour chaque segment avec comme facteur "classe de mots" (CF-B vs. CO-B) a été réalisée. Les résultats sont décrits dans le tableau 23. Aucune carte n'est significativement plus précoce pour les mots de classe fermée (CF-B) par rapport aux mots de classe ouverte (CO-B).



Segment	CF-B	CO-B	ANOVA
6	232.9 (4.84)	232.5 (5.29)	$F(1,19) = .01; p > .90$
7	290.3 (4.32)	290.4 (4.48)	$F(1,19) = .01; p > .96$
8	403.4 (3.98)	406 (4.15)	$F(1,19) = .55; p > .46$
9	467.4 (4.45)	465 (4.09)	$F(1,19) = .31; p > .58$

**Tableau 23:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments (6, 7, 8 et 9). Les valeurs indiquées sont en ms, avec entre parenthèses l'erreur type. Les effets significatifs sont rapportés en gras. Aucun segment ne montre de décalage dans le maximum de GFP.

### 3.3.3 Synthèse CF-B vs. CO-B

L'analyse portant sur les mots de basse fréquence ( $< 2.6$  log de fréquence) nous montre des résultats totalement différents de ceux rapportés lorsque seuls des mots de haute fréquence sont moyennés (voir section 3.2.3). Alors que l'analyse de tracés indique aucun effet de topographie ou encore de décalage de latence entre les deux classes de mots, la segmentation adaptée a capturé des différences de topographie dans deux segments tardifs. Pour le premier segment, numéro 9 (enregistré entre 443-500 ms pour CO-B et entre 443-492 ms pour CF-B), les PEs montrent des négativités localisées en régions centrale et droite, alors que pour le segment 10 (enregistré entre 501-624 ms pour CO-B et entre 493-664 ms pour CF-B), les négativités sont maximales en régions antérieures centrale et droite. Pour ces deux segments, les différences de topographie entre les deux classes de mots sont focales et concentrées en région postérieure (centrale et droite pour le segment 9 et uniquement centrale pour le segment 10): sur ces régions, les mots CF-B sont plus négatifs par rapport aux mots CO-B. Aucun décalage n'est mis en évidence entre les deux classes de mots avec l'analyse sur le maximum de GFP.

## 4 Discussion générale

### 4.1 L'effet de classe ne repose pas uniquement sur les propriétés lexicales

#### 4.1.1 Les analyses des tracés

Les analyses classiques (de tracés) ont montré que lorsque les propriétés lexicales sont contrôlées et les mots présentés dans un contexte de phrases, les différences de topographie entre les deux classes de mots sont restreintes à la fenêtre temporelle correspondant à la composante N280 (250-350 ms): la négativité antéro-médiane associée aux mots CF est plus

importante en régions centrale et à droite que celle des mots CO (voir figure 35). De plus, la latence de cette négativité N280 est plus précoce de 7.2 ms pour les mots de classe fermée par rapport aux mots de classe ouverte.

Dans notre expérience précédente (chapitre 4) où les propriétés lexicales ne sont pas contrôlées, nous avons montré des différences de topographie entre les deux classes de mots sur une fenêtre beaucoup plus large, allant de 150 à 800 ms. La comparaison des résultats de ces deux expériences suggère donc que les différences de topographie observées dans les fenêtres temporelles incluant 150-250 ms (P2), 350-500 ms (N400) et 500-800 ms (LAN) ne soient pas directement liées à des différences qualitatives entre les deux classes de mots, c'est à dire, au niveau de leur fonction dans la phrase. Comme dans l'expérience du chapitre 4 de nombreux facteurs n'étaient pas contrôlés (fréquence d'occurrence, longueur, répétition, position dans la phrase, tâche), ces différentes fenêtres peuvent refléter des variations de différentes natures: 1) liées aux propriétés lexicales distributionnelles des mots (leur fréquence) et/ou à leurs propriétés physiques (leur longueur), 2) liées à la répétition des mots CF qui était plus importante, 3) liées à la présentation de violation sémantique. Puisque dans ce chapitre de tels facteurs ne sont plus confondus, nos résultats suggèrent que c'est au niveau de la fenêtre temporelle allant de 250 à 350 ms (incluant la N280) que ressort la spécificité fonctionnelle des mots de classe fermée par rapport aux mots de classe ouverte.

Les analyses de régression effectuées dans le chapitre 4 ont montré que la latence de la composante précoce négative N280 (enregistrée entre 250 et 350 ms) est effectivement corrélée avec la classe de mots, mais qu'elle est aussi corrélée avec la fréquence et la longueur de ces mots. Cette nouvelle expérience montre que le contrôle de la fréquence et de la longueur des mots ne supprime pas entièrement la différence de latence entre les deux classes de mots. Toutefois, si l'on examine les effets rapportés dans ces deux chapitres, la différence de latence entre les deux classes de mots apparaît plus importante (11.1 ms) lorsque les propriétés lexicales ne sont pas contrôlées comparée à celle observée lorsque les propriétés lexicales sont contrôlées (7.2 ms). En conclusion, il semble que la fréquence et la longueur affectent bel et bien la latence de la composante N280 enregistrée entre 250 et 350 ms, mais ces deux facteurs à eux seuls ne rendent pas entièrement compte de l'activité électrique observée dans cette fenêtre. La différence entre les mots de classe fermée et ceux de classe ouverte au niveau de la fonction qu'ils jouent dans la langue semble, elle aussi, être un facteur déterminant dans la définition d'une part de la topographie, mais aussi de la précocité de cette composante.

### 4.1.2 L'analyse de segmentation adaptée

Dans un premier temps, nous aimerions noter un résultat non dépourvu d'intérêt. La méthode de segmentation adaptée réalisée sur 2 sets de données totalement indépendants (chapitres 4 et 5) indique une distribution spatiale sur le scalp, ainsi qu'un décours temporel (frontières des segments) des cartes PEs qui sont similaires pour ces deux expériences (figure 25 et figure 37). D'une manière objective, nous pouvons donc être confiants quant à la détermination des frontières pour nos fenêtres temporelles utilisées lors de nos analyses statistiques.

Lorsque les propriétés lexicales sont contrôlées, les résultats de la segmentation adaptée indiquent que deux segments (6 et 7) présentent des différences de topographie dans la fenêtre temporelle allant de 199 à 338 ms pour les mots CF. Pour le premier segment (199-256 ms, segment 6), la négativité antérieure est équitablement répartie pour les mots CF alors que pour les mots CO celle-ci est maximale à gauche et en central. Le deuxième segment (257-338 ms, segment 7) montre que la négativité antéro-médiane des mots CF est aussi bilatérale alors que pour les mots CO, elle est concentrée en régions centrales. De plus, le segment 7 présente un pic de maximum de GFP qui est plus précoce de 10.5 ms pour les mots CF. Grâce à cette méthode, nous pouvons affiner les résultats obtenus à partir de l'analyse de tracés. En effet, les analyses de tracés nous ont permis de conclure que les différences de topographie s'étendaient entre 250 et 350 ms. Mais si l'on se réfère aux résultats de la segmentation adaptée, définissant les fenêtres à partir des fluctuations des PEs de surface, ces différences sont enregistrées dès 200 ms, soit dès le début du segment 6. La fenêtre arbitraire utilisée lors de l'analyse de tracés pour la composante P2 (150-250 ms) s'est donc révélée soit trop large, soit mal ciblée, pour pouvoir rendre compte des différences de topographie aussi précoces que 200 ms entre les deux classes de mots (telles que celles mises en évidence au niveau du segment 6 lors de l'analyse de segmentation). Ce résultat souligne les limites de l'analyse des tracés qui n'a pas permis de mettre en évidence l'existence de différences précoces, c'est à dire antérieures à la N280, entre les deux classes de mots. De plus, la segmentation adaptée a permis de diviser la fenêtre caractéristique de la composante N280 en deux micro-états. Alors que dans le segment 6, les négativités associées sont localisées sur les régions antérieures seulement, pour le segment 7, celles-ci sont observées sur les régions antérieures et médianes. Ce dernier résultat suggère qu'en réalité cette composante N280, ou plutôt, la fenêtre temporelle correspondant à cette composante, peut refléter plusieurs processus différents.

Si l'on compare les effets obtenus dans le chapitre 4 et ceux de ce chapitre, leurs distributions sur le scalp ne sont pas similaires. Alors que dans le chapitre 4, la différence de négativité entre les deux classes de mots était maximale sur les régions restreintes en antérieure et médiane gauche et centrale (figure 26 et 27), lorsque les propriétés lexicales sont contrôlées, celles-ci sont enregistrées sur les régions englobant tout l'hémisphère droit (figure 38 et 39). Si l'on regarde attentivement les PEs pour les régions antérieures séparément pour chaque classe de mots, on s'aperçoit qu'en réalité les mots CF montrent pour les deux expériences une négativité également répartie (bilatérale) pour les deux segments et dans les deux expériences (figures 25 et 37). En revanche, les mots CO semblent présenter des différences de latéralisation flagrantes entre les deux expériences. En effet, dans le chapitre 4 (figure 25), les négativités associées aux mots de classe ouverte sont d'abord maximales en région antérieure droite (segment 6) puis en région antérieure droite et centrale (segment 7). Dans le chapitre 5 (figure 37), la négativité des mots CO est enregistrée sur les régions antérieures gauche et centrale puis uniquement centrale. Autrement dit, les différences topographiques rapportées entre les deux expériences semblent reposer entièrement sur les mots de la classe ouverte. Afin d'expliquer ces différences de latéralisation pour les mots CO, nous avons émis plusieurs hypothèses.

Une première hypothèse porte sur les différentes catégories grammaticales composant la classe de mots CO. Alors que dans le chapitre 4 les mots CO étaient constitués de 33.3 % de noms, 33.3 % de verbes et 33.3 % d'adverbes (voir section 2.2 chapitre 4), dans le chapitre 5 nous avons utilisé 61.5 % de noms, 22.5 % de verbes et 16.25 % d'adjectifs (tableau 13, section 2.2 de ce chapitre). Cette différence dans l'élaboration de nos deux groupes de mots de classe ouverte a pu affecter la latéralisation de la composante négative précoce. Toutefois, l'analyse des catégories grammaticales opérée dans le chapitre 4 (section 3.2) indique que les négativités des 3 catégories grammaticales pour les mots CF sont différentes (en amplitude) sur les régions antérieures pour le segment 7; aucune différence n'a été rapportée pour les différentes catégories grammaticales des mots CO. Ces résultats nous permettent de rejeter l'hypothèse selon laquelle les proportions inégales dans les catégories grammaticales utilisées dans les deux expériences auraient influencé la latéralisation de la composante négative précoce pour les mots de classe ouverte <sup>17</sup>.

---

<sup>17</sup> ce qui n'est pas le cas pour les différentes catégories grammaticales des mots de classe fermée, voir section 3.2 chapitre 4.

Une seconde hypothèse est que les paradigmes expérimentaux utilisés dans les deux expériences ont suscité ces différences. En effet, le matériel utilisé mais aussi la tâche demandée ont pu créer ces différences. Alors que dans l'expérience du chapitre 4 la moitié des phrases présentait des violations sémantiques et l'autre moitié correspondait à des phrases correctes, dans le chapitre 5 toutes les phrases étaient correctes. De plus, pour le chapitre 4, les participants devaient effectuer un jugement sémantique sur chacune des phrases alors que dans la seconde expérience (chapitre 5), les participants devaient répondre à des questions de compréhension pour 20 % des phrases. Si l'on considère que les mots de classe ouverte véhiculent les informations référentielles dans la phrase, le fait d'avoir présenté des violations sémantiques dans l'expérience du chapitre 4 a pu influencer la stratégie des participants. Osterhout, *et al.* (2002) apportent des résultats qui corroborent cette hypothèse. Dans une histoire courte, la moitié des phrases présentait des violations soit d'ordre syntaxique soit d'ordre sémantique. Les auteurs ont manipulé la tâche demandée: les participants devaient soit réaliser un jugement d'acceptabilité sur les phrases, soit simplement les lire. Les résultats indiquent que le type de tâche influence les PEs des mots CO et CF entre 200-500 ms: les PEs lors de la lecture simple sont plus négatifs par rapport à ceux obtenus lors d'un jugement d'acceptabilité. Mais surtout, le type de tâche influence différemment les PEs des deux classes de mots: les différences entre les deux tâches pour les mots CO sont plus importantes sur les sites antérieurs alors qu'elles sont plus distribuées sur le scalp pour les mots CF. Comme les participants devaient réaliser un jugement à la fois d'acceptabilité syntaxique et sémantique, il ne semble pas surprenant que ce type de tâche puisse affecter différemment les PEs des deux classes de mots. En revanche, puisque dans notre expérience, seules des violations sémantiques étaient utilisées, il est possible que les participants aient engagé des stratégies d'analyse sémantique poussée alors que l'analyse syntaxique de la phrase était reléguée au second plan. Si l'on considère que les mots CO apportent les informations sémantiques de la phrase alors que les mots CF sont utilisés lors de l'ébauche de la construction syntaxique (Friederici & Schoenle, 1980; Garrett & Kean, 1980), alors la différence de type de tâche ainsi que la nature des phrases employées dans nos deux expériences (chapitre 4 et 5) a pu produire des variabilités observées dans les PEs des mots CO uniquement.

## 4.2 Hypothèses quant à l'interprétation fonctionnelle des effets

La segmentation adaptée définit deux segments de topographie différente entre les mots de classe ouverte et fermée (segments 6 et 7). Comme ces deux micro-états présentent des topographies différentes (négativité antérieure pour le segment 6 puis antéro-médiane pour le 7), nous avançons l'hypothèse que ces deux fenêtres d'analyse (segment 6 et 7) reflètent l'activation de deux étapes différentes lors de la reconnaissance des mots. De plus, comme nous avons montré des différences significatives de topographie entre les deux classes de mots, nous suggérons que ces deux micro-états sont différents pour les mots de classe fermée et ouverte.

Dans le chapitre 4, nous avons suggéré, que le segment 6 refléterait plutôt une étape de traitement visuel précoce. Dans ce chapitre, nous avons enregistré ce segment entre 199 et 256 ms pour les mots CF, fenêtre d'analyse qui correspond avec celle du segment 6 du chapitre 4 (199-246 ms). Alors que dans ce chapitre les mots utilisés ont des propriétés lexicales contrôlées, l'interprétation selon laquelle ce segment puisse refléter des différences de traitement visuel n'est plus valable. De plus, notons que la fenêtre temporelle de ce segment ne correspond pas totalement à celle de la composante P2 (150-250 ms) et est un peu plus tardive (200-250 ms). Dans cette période et à partir de violations de catégorie syntaxique, Friederici & Mecklinger (1996) et Friederici, Pfeifer, & Hahne (1993) observent l'occurrence d'une composante précoce avec une négativité maximale sur les sites antérieure gauche, effet qu'ils nomment ELAN pour *Early Left Anterior Negativity*. Ces auteurs suggèrent que cette composante reflète l'activation des informations de catégorie syntaxique lors de construction de la représentation structurale de la phrase. La topographie ainsi que la fenêtre temporelle correspond à l'effet que nous avons enregistré dans le segment 6. Alors que nous ne présentons pas de violations de catégorie syntaxique, nous émettons l'hypothèse que l'effet observé dans ce segment correspond à l'activation spécifique des informations de la classe du mot. Si les mots de classe fermée participent à la construction de la représentation syntaxique de la phrase (Friederici & Schoenle, 1980; Garrett, 1975), alors il est nécessaire que précocement une étape dans les processus d'identification des deux classes de mots soit différente: l'effet enregistré à partir du segment 6 peut en être la manifestation sur le scalp.

Le deuxième segment (7) présentant des différences de topographie est enregistré entre 257 et 338 ms pour les mots CF et entre 251 et 334 ms pour les mots CO. Comme nous l'avons déjà indiqué dans le chapitre précédent, cette fenêtre temporelle, ainsi que la distribution sur le

scalp des PEs correspond aux caractéristiques de la composante N280 (Brown, *et al.*, 1999; King & Kutas, 1998; Neville, *et al.*, 1993; Neville, *et al.*, 1992; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997; Pulvermüller, *et al.*, 1995). Nous avons émis l'hypothèse dans le chapitre précédent que cette composante reflète des processus d'accès lexicaux. Si tel est le cas, alors le fait d'avoir enregistré une latence plus précoce pour les mots CF par rapport aux mots CO indique que l'accès aux mots CF est plus rapide que celui des mots CO. L'interprétation de cet effet de latence reposerait sur une différence dans le nombre de candidats à vérifier lors de l'accès au lexique. Si les mots de classe fermée et ouverte sont stockés dans deux lexiques séparés et puisque les mots de classe fermée sont moins nombreux que les mots de classe ouverte, alors la recherche dans le lexique mental où les mots de classe fermée sont présents doit être plus rapide que celle effectuée lors de la vérification de la liste incluant les mots de classe ouverte. Nos résultats indiquant des différences de topographie et de latence dans le segment 7 suggèrent que les mots de classe fermée et ouverte sont stockés séparément dans deux lexiques différents et que la reconnaissance des mots CF est plus rapide que celle des mots CO. Toutefois, une deuxième interprétation pour le décalage du segment 7 est possible si l'on considère que le contexte peut avoir affecté les PEs dans cette fenêtre. Selon Friederici & Schoenle (1980) et Shillcock & Bard (1993), le contexte syntaxique peut influencer la reconnaissance des mots de classe fermée alors qu'aucun effet de contexte sur la reconnaissance des mots de classe ouverte n'est attendu. Pour ces auteurs, les mots de classe fermée sont associés à la construction de la structure syntaxique qui est automatique et indépendante de tout autre type d'information que celle de nature syntaxique (Forster, 1976; Forster, 1979). Ainsi les différences de latence en faveur des mots CF obtenues dans notre expérience peuvent aussi s'expliquer par le fait que le contexte linguistique (selon Friederici & Schoenle, 1980 et Shillcock & Bard, 1993; uniquement les informations de nature syntaxique) a permis une activation plus rapide des mots CF par rapport aux mots CO.

A la lumière de nos résultats, nous ne pouvons nous prononcer en faveur de l'une ou de l'autre de ces deux interprétations, mais le chapitre 6 tentera d'éclairer ce débat.

### 4.3 L'effet de classe varie avec la fréquence

Bien que nous ayons mis en évidence un effet de classe "général" en comparant des mots de CF ayant une même fréquence moyenne que des mots de CO, nos résultats indiquent que cet effet de classe n'est pas identique pour des mots de haute fréquence et de basse fréquence.

En effet, nous avons montré que pour des mots de haute fréquence ( $> 2.6$  log de fréquence, ou  $> 400$  occurrences/million) l'effet de classe de mots est enregistré dans la fenêtre temporelle correspondant à la composante N280 (250-350). L'analyse de segmentation adaptée montre d'une part des différences de topographie pour les segments 6 et 7 (qui correspondent aux fenêtres 203-258 ms et 259-342 ms pour CF-H; 209-250 ms et 251-334 ms pour CO-H respectivement), et d'autre part que les mots de classe fermée de haute fréquence (CF-H) présentent une composante négative pour le segment 7 qui est plus précoce (de 7.2 ms) par rapport à celle enregistrée pour les mots de classe ouverte de haute fréquence. Notons que l'analyse de tracés a elle aussi montré une différence significative (de 8.6 ms) dans la fenêtre 250-350 ms. A l'opposé, l'effet de classe pour des mots de basse fréquence ( $< 2,6$  log de fréquence, ou encore  $< 400$  occurrences/million) présente des différences de topographie plus tardives que celles observées pour l'effet de classe de haute fréquence. L'analyse de tracés n'indique aucune différence de topographie dans les différentes fenêtres temporelles étudiées (P2, N280, N400 et LAN). La segmentation adaptée a néanmoins capturé des différences de topographie toutefois très sporadiques. Nous émettons des réserves quant à ces différences topographiques entre les deux classes pour les mots de basse fréquence. Dans les fenêtres correspondantes aux deux segments en question, les segments 9 et 10, le GFP (*Global Field Power*) indique des valeurs très basses: la puissance des cartes est, par conséquent, très peu élevée. Le bruit dans les PEs a pu être à la base des résultats significatifs. De plus, la procédure de "fitting" indique des taux assez bas de pourcentage de variance expliquée (BEV) pour ces deux cartes, même lorsque la carte des mots CF-B est testée dans les propres données individuelles des mots CF-B (43.6 % pour le segment 9 et 38.1 % pour le segment 10). Ainsi, il nous semble nécessaire de répliquer ces différences topographiques tardives sur des mots de basse fréquence ( $< 2.6$  log de fréquence). Enfin, nous n'avons pas mis en évidence de différence de latence d'occurrence que ce soit au niveau des analyses spécifiques des électrodes situées en région antérieure (analyse de tracés), ou bien grâce au pic de maximum de GFP (analyse de segmentation adaptée).



## 5 Conclusions

Les différences topographiques et de latences observées entre les deux classes de mots lorsque les deux rangs de fréquence sont moyennés apparaissent donc reposer principalement sur les mots de fréquence d'occurrence supérieure à 400 par million ( $> 2,6 \log$  de fréquence). L'effet de classe pour des fréquences supérieures à  $2,6 \log$  de fréquence se résume donc à deux étapes ou micro-états de topographie différente sur le scalp dont l'une d'entre elles est plus précoce pour les mots de classe fermée par rapport aux mots de classe ouverte. Pour les mots de basse fréquence, nous pouvons conclure que des différences tardives de topographie sont enregistrées, mais que celles-ci nécessitent d'être répliquées. Nos résultats suggèrent donc que les deux classes de mots ne sont pas traitées de manière identique selon que les mots ont une basse ou une haute fréquence d'occurrence. Ces résultats confortent l'hypothèse de Gordon & Caramazza (1983; 1985) selon laquelle les mots de classe fermée et ouverte réagissent différemment en fonction du rang de fréquence étudié. Les mots de classe fermée de haute fréquence présentent deux micro-états ou étapes qui diffèrent des mots de classe ouverte de même longueur et fréquence d'occurrence. Nous avons émis l'hypothèse que le premier segment pourrait refléter des processus liés à la construction de la structure syntaxique de la phrase, et que le second représente l'accès au lexical mental. Afin de pouvoir soutenir nos hypothèses sur les interprétations fonctionnelles de ces deux segments, il nous faut vérifier si ceux-ci appartiennent bien aux étapes incluses lors de l'accès lexical et non pas à des processus plus tardifs comme l'intégration. Le chapitre qui suit est consacré à cette vérification puisque les mots seront cette fois présentés de manière isolée, sans présence de contexte linguistique.

# CHAPITRE 6

## Effet de Classe et Propriétés Lexicales Contrôlées en Isolation

---

### *Résumé*

Des mots de classe ouverte et fermée, contrôlés en fréquence d'occurrence ainsi qu'en longueur, sont présentés dans une liste de mots. Les potentiels évoqués sont recueillis à partir de 62 électrodes, alors que les participants réalisent une tâche de lecture des mots présentés visuellement sur un écran. L'analyse de tracés indique que des différences topographiques sont significatives dans le rang de latence 250-350 ms et 350-500 ms. De plus, la composante négative, enregistrée dans la fenêtre temporelle 250-350 ms, présente une latence plus précoce (7 ms) pour les mots CF. La segmentation adaptée indique que 3 segments, ou micro-états, montrent des différences topographiques entre les deux classes de mots. Le segment 6 (297-366 ms pour CO, 295-386 ms pour CF) indique la présence d'une composante négative antérieure gauche et centrale pour les deux classes de mots, mais plus importante sur les 3 sites antérieurs pour les mots CF. Pour le segment 7 (367-452 ms pour CO, 387-452 ms pour CF), alors que les mots CO évoquent une composante négative antérieure, celle-ci est latéralisée sur les régions gauches pour les mots CF. Pour le segment 8 (453-595 ms pour CO, 453-628 ms pour CF), deux négativités sont enregistrées sur le scalp. D'une part, les PEs associés aux mots CO sont plus négatifs que ceux des mots CF sur les régions postérieures droites, alors que l'effet inverse est observé sur les régions gauches postérieure et médiane. Deux segments (5 et 6) présentent des latences plus précoces pour les mots CF par rapport aux mots CO (3.9 et 8.8 ms respectivement). L'analyse séparée de l'effet de classe indique que les effets topographiques précoces sont essentiellement liés à une différence dans la reconnaissance des mots de basse fréquence. Les différences de latence sont liées au traitement des mots de classe fermée de haute fréquence. Nos résultats suggèrent que l'accès au lexique des mots de classe fermée de basse fréquence est différente de celle des mots de classe ouverte de fréquence équivalente lorsque ceux-ci sont présentés dans une liste de mots. Une différence de rapidité dans le traitement des mots de classe fermée de haute fréquence est obtenue par rapport aux mots de classe ouverte de même fréquence.

## 1 Introduction

Alors qu'un rôle sémantique est attribué aux mots de classe ouverte, les mots de classe fermée semblent être liés à la construction de la structure syntaxique (Friederici & Schoenle, 1980; Garrett, 1975; Garrett & Kean, 1980). Toutefois, selon ces théories, il n'est pas certain qu'il existe des différences au niveau de l'accès lexical de ces deux classes de mots. Des différences pourraient apparaître après l'identification du mot à savoir, lors de l'intégration du mot dans son contexte. Si tel est le cas, lorsque ces derniers sont présentés en isolation, la probabilité de trouver des différences de traitement entre ces deux classes doit être réduite. De nombreux arguments plaident, en effet, en faveur de l'hypothèse selon laquelle la distinction entre les mots de classe fermée et ouverte ne serait effective que dans un contexte de phrase (Brown, *et al.*, 1999; King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997).

Grâce à la technique des PEs, quelques expériences (6 au total, voir tableau 2 chapitre 2) ont essayé de répondre à la question de savoir si ces deux classes de mots sont traitées différemment dans le cerveau lorsque ceux-ci sont présentés en isolation. Deux d'entre elles (Münte, *et al.*, 2001; Osterhout, *et al.*, 1997) concluent que l'effet précoce enregistré sur la composante N280 est similaire pour les deux classes de mots, alors que les 4 autres montrent des composantes différentes pour les deux classes de mots (Garnsey, 1985; Keurs ter, *et al.*, 2002; Nobre & McCarthy, 1994; Pulvermüller, *et al.*, 1995). Comme nous l'avons souligné dans le chapitre 2, la variabilité des résultats peut provenir de différents facteurs tels que les propriétés lexicales, les répétitions, la tâche. Nous avons vu que la fréquence et la longueur interfèrent avec l'effet de classe (chapitres 4 et 5). Afin de limiter les effets des propriétés lexicales, 2 des 6 expériences citées, contrôlent la fréquence et la longueur de leurs stimuli (Garnsey, 1985; Pulvermüller, *et al.*, 1995). Alors que Garnsey (1985) constate que les PEs associés aux mots CF sont plus négatifs à partir de 350 ms sur les 8 électrodes enregistrées par rapport à ceux des mots CO, Pulvermüller, *et al.* (1995) trouvent deux composantes négatives précoces, la N160 et la N260, enregistrées pour les deux classes de mots, mais différemment latéralisées: maximale à gauche pour les mots de classe fermée et bilatérale pour les mots de classe ouverte. En conclusion, comme deux composantes différentes sont enregistrées sur le scalp après la présentation des mots de classe ouverte et fermée et ce, dès 160 ms, ces auteurs

concluent que des processus différents sont activés lors de l'accès au lexique des deux classes de mots. A l'opposé, les différences tardives mises en évidence par Garnsey (1985) suggèrent un accès lexical similaire pour les deux classes de mots, mais des processus d'intégration différents. Même si ces 2 expériences (Garnsey, 1985; Pulvermüller, *et al.*, 1995) suggèrent qu'un processus différent (lié ou non à l'accès lexical) est mis en place lors du traitement des mots de classe ouverte et fermée, il ne faut pas oublier que Münte, *et al.* (2001) et Osterhout, *et al.* (1997) proposent des interprétations opposées. Nous tenons à noter un dernier point qui peut rendre les résultats de Garnsey (1985) et de Pulvermüller, *et al.* (1995) moins irréfutables: dans ces 2 expériences, les auteurs utilisent des tâches de décision lexicale (les items critiques sont mélangés avec des non-mots). Nous avons suggéré que ce type de tâche peut influencer la manière dont les participants analysent les différents mots, et par conséquent, ne pas refléter uniquement la reconnaissance des mots. Il est fort probable que le fait d'avoir demandé aux participants de réaliser une telle tâche<sup>18</sup> et de mélanger les items avec des non-mots, ai pu accentuer les effets obtenus. Deux expériences (Keurs ter, *et al.*, 2002; Osterhout, 1997) présentent les mots en isolation sans tâche additionnelle (Keurs ter, *et al.*, 2002), ou seulement avec des questions de compréhension (Osterhout, 1997). Osterhout (1997) indiquent des similitudes dans les composantes négatives précoces, alors que Keurs ter, *et al.* (2002) aboutissent aux conclusions opposées. Quoiqu'il en soit, puisque ces recherches n'ont pas contrôlé la fréquence et la longueur des mots utilisés, il est impossible de déterminer si les effets observés sur le scalp sont liés à la classe de mots, ou plutôt aux propriétés lexicales.

L'expérience que nous allons maintenant décrire participe à ce débat et a pour premier objectif d'étudier la distinction entre classe de mots lorsque ceux-ci sont présentés dans une liste de mots. A l'opposé de Garnsey (1985) et de Pulvermüller, *et al.* (1995), aucune tâche additionnelle ne sera demandée aux participants, qui devront simplement lire les items.

---

<sup>18</sup> En comparant nos résultats du chapitre 4 et 5, nous avons montré des différences dans la latéralisation des négativités précoces des mots CO par rapport aux mots CF. Nous avons émis l'hypothèse que ces différences proviennent des différences de tâches employées dans nos deux expériences. De tels résultats coïncident avec ceux de Osterhout, *et al.* (2002).

Le deuxième objectif de cette expérience est de vérifier si l'effet de classe (c'est à dire la distinction entre mots de classe ouverte et fermée) est différent lorsqu'on étudie des mots de fréquence élevée ( $> 2,6 \log$  de fréquence) ou basse ( $< 2,6 \log$  de fréquence). En présence de contexte, les différences topographiques précoces et les différences de latences sont exclusivement enregistrées pour les mots de haute fréquence (chapitre 5). Si ces effets sont liés à des différences lors de l'identification des deux classes de mots, alors nous devrions obtenir de tels résultats lorsque les mots sont présentés dans des listes de mots.

## 2 Méthode

### 2.1 Sujets

Vingt étudiants de la Faculté de Psychologie et des Sciences de l'Éducation de Genève, dont 15 femmes, ont participé bénévolement à cette expérience. La moyenne d'âge était de 24,1 ans (compris entre 18 et 39 ans). Tous les sujets étaient droitiers, selon le test de latéralité *Edinburgh Handedness Inventory* (Oldfield, 1971), de langue maternelle française, et présentaient une vision normale ou corrigée.

### 2.2 Matériel

Le corpus utilisé dans cette expérience est décrit dans le chapitre 5, section 2. La liste complète des mots est présentée dans l'annexe 3. La succession des mots a été réalisée de sorte que les participants ne puissent former des groupes de mots, que se soit d'un point de vue syntaxique (par exemple: déterminant + nom ou encore nom + verbe) ou encore sémantique (par exemple: adjectif + nom; voir annexe 5).

### 2.3 Procédure

Chacun des items test était présenté pendant 300 ms, avec un SOA (*Stimulus Onset Asynchrony*) de 1,1 secondes. Entre chaque mot, une croix de fixation était présentée pendant 700 ms (SOA de 1,5 secondes). Les participants étaient informés que les mots n'avaient pas de lien entre eux et que leur tâche était simplement de les lire. Avant le début de l'expérience

(7 minutes de stimulation), le participant lisait une consigne écrite et effectuait une phase d'apprentissage sur une dizaine de mots.

## 2.4 Enregistrement de l'EEG

Après la pose du bonnet à électrodes (Quik-Cap, Neuroscan Inc., Herndon, Va., USA), le participant était installé dans une salle insonorisée et isolée électriquement (cage de Faraday<sup>19</sup>), à environ 100 cm en face d'un écran d'ordinateur. L'électroencéphalogramme (EEG) a été enregistré à partir d'un montage constitué de 62 électrodes Ag/AgCl dérivé du système 10-20 international (Society, 1994). La configuration des électrodes est un peu différente de celle utilisée dans les autres expériences (3, 4, 5 voir figure 13). En effet, une électrode a été supprimée (IZ en occipital), et deux nouvelles positions en postérieur gauche (PO5, située entre PO3 et PO7) et en postérieur droit (PO6, située entre PO4 et PO8) ont été ajoutées. L'électrode Afz a été utilisée comme terre. L'électrode située sur le mastoïde gauche a été utilisée pendant l'enregistrement comme référence. Afin de contrôler les clignements et/ou mouvements des yeux durant l'acquisition de l'EEG, 2 électro-oculogrammes (EOG) horizontaux et verticaux ont été enregistrés à l'aide d'un montage bipolaire d'électrodes externes Ag/AgCl. L'acquisition, ainsi que l'obtention des moyennages, est identique aux expériences précédentes (voir chapitre 3, section 2.4) et résultent en 6 jeux de PEs pour chaque participant: CO, CF, CO-H, CF-H, CO-B, CF-B.

Dans un premier temps, nous allons étudier l'effet de classe lorsque les hautes et basses fréquences sont mélangées mais contrôlées (CF vs. CO). Puis, nous analyserons séparément l'effet de classe pour les mots de haute fréquence (CF-H vs. CO-H) et de basse fréquence (CF-B vs. CO-B).

---

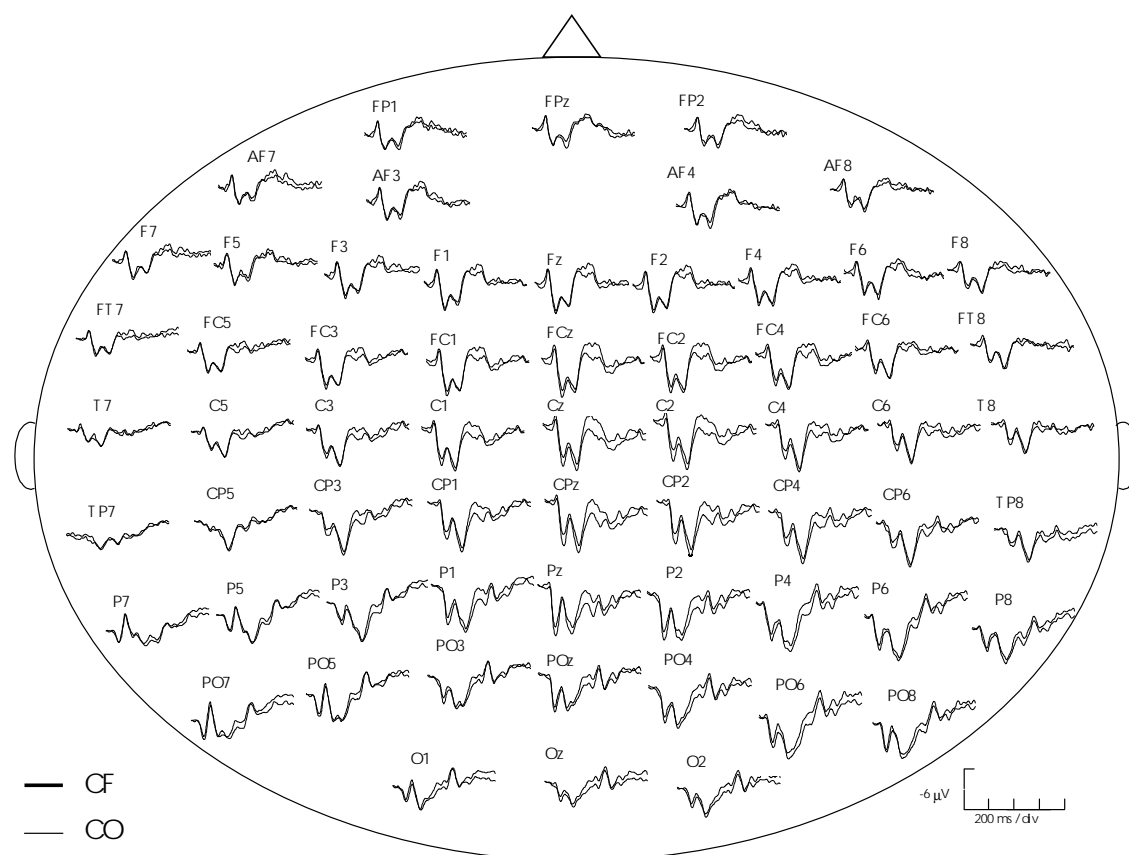
<sup>19</sup> Cette dernière expérience s'est déroulée dans les locaux de la FAPSE à Uni-mail, alors que toutes les autres expériences ont eu lieu dans les anciens locaux sur le site de Batelle à Carouge.

### 3 Résultats

#### 3.1 Comparaison CF vs. CO

##### 3.1.1 Analyses des tracés

##### 3.1.1.1 Visualisation des tracés



**Figure 49:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 62 électrodes enregistrées. Les tracés en gras correspondent aux PEs des mots de classe fermée (CF), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte (CO). Les électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sont sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, et O pour occipital. La négativité est située vers le haut.

La figure 49 présente les tracés obtenus pour les mots de classe ouverte (CO) et fermée (CF) en isolation. Visuellement, les tracés révèlent les composantes que nous avons décrites

précédemment P100, N2 et P2 (chapitres 3, 4 et 5). La comparaison des PEs pour les deux classes de mots indique qu'entre 300-500 ms les PEs des mots de classe ouverte (CO) sont plus négatifs en régions centrales par rapport à ceux des mots de classe fermée (CF). Il semble intéressant de noter, qu'avant 300 ms, les PEs des deux classes de mots sont très positifs et qu'aucune composante négative n'est visible sur les sites antérieurs. De plus, aucune différence ne semble évidente avant la fenêtre temporelle de 300-500 ms.

### 3.1.1.2 Analyse topographique par région d'intérêt

Après une analyse visuelle des composantes, les 4 fenêtres temporelles définies pour le chapitre 4 ont été conservées pour réaliser nos analyses statistiques: la P2 (150 et 250 ms), la N280 (250 et 350 ms), la N400 (350 et 500 ms), et la LAN (500 et 800 ms). Puisque trois nouvelles positions d'électrodes ont été utilisées dans cette expérience, nous avons réorganisé les facteurs spatiaux pour les régions postérieures (voir tableau 24).

Facteurs Spatiaux		Electrodes
Caudalité	Latéralité	
Antérieur	Gauche	FP1, AF7, AF3, F7, F5, F3
Antérieur	Central	FPz, F1, Fz, F2
Antérieur	Droit	FP2, AF4, AF8, F4, F6, F8
Médian	Gauche	FT7, FC5, FC3, T7, C5, C3, TP7, CP5, CP3
Médian	Central	FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, CP2
Médian	Droit	FC4, FC6, FT8, C4, C6, T8, CP4, CP6, TP8
Postérieur	Gauche	P7, P5, P3, PO7, PO5, PO3
Postérieur	Central	P1, Pz, P2, POz, O1, Oz, O2
Postérieur	Droit	P4, P6, P8, PO4, PO6, PO8

**Tableau 24:** Description des facteurs spatiaux utilisés pour l'ANOVAs à mesures répétées: "classe de mots" (CO vs. CF), "caudalité" (antérieur vs. médian vs. postérieur) et "latéralité" (gauche vs. central vs. droit). Aucune électrode n'est rejetée de cette analyse.

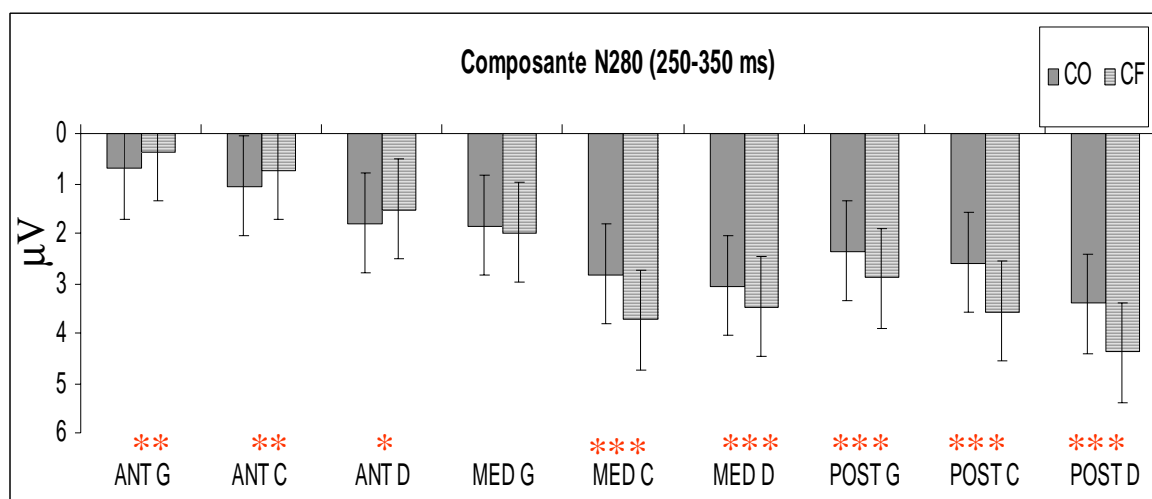
Les résultats des ANOVAs séparées à mesures répétées, pour chaque fenêtre temporelle, avec les facteurs "classe de mots" (CO vs. CF), "caudalité" (antérieur vs. médian vs. postérieur) et "latéralité" (gauche vs. central vs. droit) sont présentés dans le tableau 25.



	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	1.48 - /	1.56 - /	<b>8.90</b> ** - /	3.10 - /
Classe x Caudalité	2, 38	0.82 - 1.61	<b>8.32</b> ** - <b>4.99</b> *	0.67 - 0.39	0.83 - 0.90
Classe x Latéralité	2, 38	3.34 * - 0.92	3.10 - 1.17	<b>4.26</b> * - <b>3.53</b> *	0.04 - 0.34
Classe x Caudalité x Latéralité	4, 76	2.97 * - 1.75	<b>3.43</b> * - <b>2.90</b> *	<b>3.49</b> * - <b>2.93</b> *	2.21 - 2.35

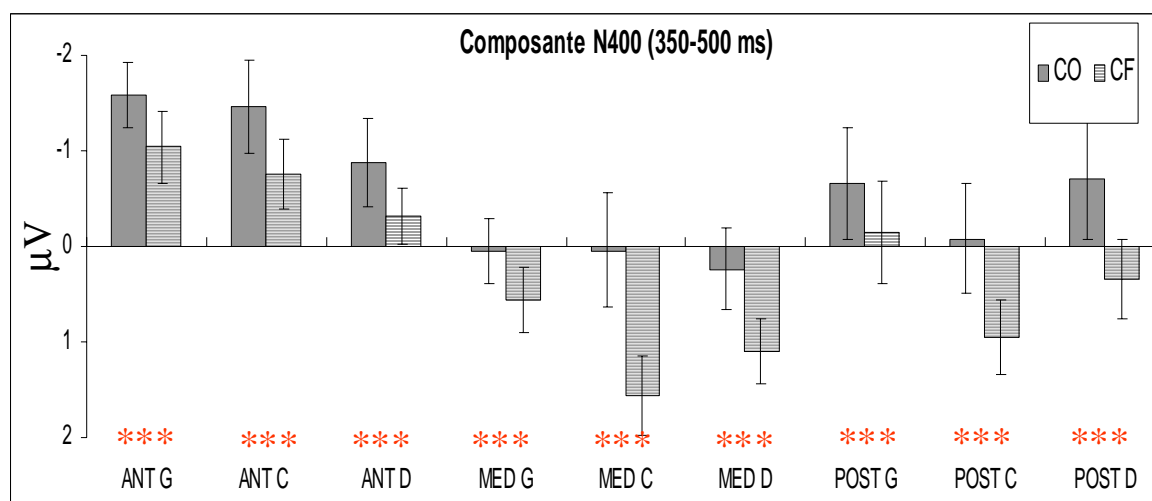
**Tableau 25:** Résultats des ANOVAs à mesures répétées pour les différentes fenêtres d'analyses (150-250 ms, 250-350 ms, 350-500 ms et 500-800 ms) sur l'amplitude du PEs entre les deux classes de mots (CO vs. CF). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italiques) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Dans le rang de latence de la composante N280 (250-350 ms), les PEs sont plus positifs en régions médiane et postérieure ("caudalité";  $F(2,38) = 13.64$ ;  $p < .0001$ ) ainsi qu'à droite ("latéralité";  $F(2,38) = 13.75$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 5.00$ ;  $p < .01$ ) indique que les PEs des mots CF sont plus positifs en région médiane ( $p < .02$ ) et postérieure ( $p < .0002$ ) par rapport à ceux des mots CO. Les comparaisons planifiées (*Post-Hoc*, LSD, figure 50) de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 2.91$ ;  $p < .02$ ) indiquent, qu'en régions antérieures (les 3) les PEs des mots CF sont moins positifs que ceux des mots CO, alors qu'en régions postérieures (les 3) et médianes centrale et droite, l'effet inverse est observé.



**Figure 50:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N280 (comprise entre 250 et 350 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Entre 350 et 500 ms, une négativité est enregistrée sur les régions antérieures ("caudalité";  $F(2,38) = 7.61$ ;  $p < .001$ ) et à gauche ("latéralité";  $F(2,38) = 3.62$ ;  $p < .03$ ). Dans cette fenêtre temporelle, l'effet du facteur principal "classe de mots" est significatif: les PEs des mots CO sont plus négatifs ( $-0.59 \mu V \pm 0.49$ ) sur tout le scalp, par rapport à ceux des mots CF ( $0.24 \mu V \pm 0.38$ ). La figure 51 présente les résultats des comparaisons planifiées (*Post-Hoc*, LSD) pour l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 2.93$ ;  $p < .02$ ). En régions frontales, les PEs des deux types de mots suscitent l'occurrence de négativités, mais l'amplitude est plus importante pour les mots CO que pour les mots CF. En régions postérieures, les PEs des mots CO suscitent l'occurrence d'une négativité, alors que ceux des mots CF sont positifs. En régions médianes, les PEs des mots CF sont plus positifs par rapport à ceux des mots CO.

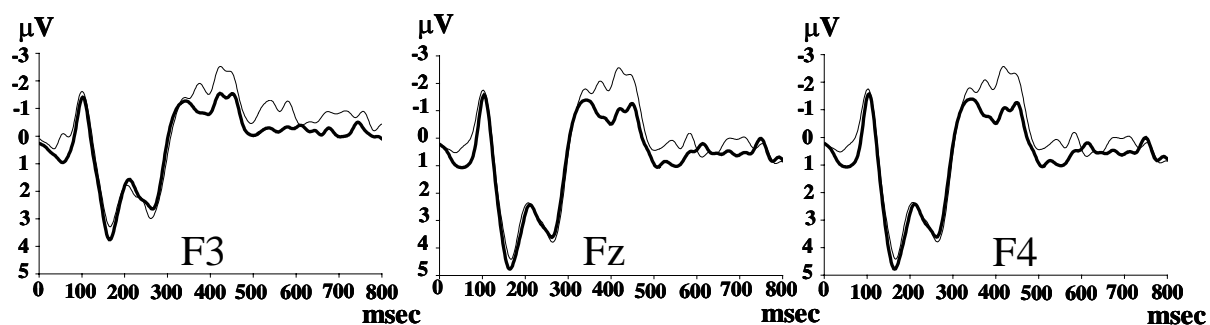


**Figure 51:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N400 (comprise entre 350 et 500 ms). Les valeurs du PEs (en  $\mu V$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.1.1.3 Analyse de latence

Alors que, dans la fenêtre temporelle comprise entre 250 et 350 ms, les PEs des deux classes de mots sont positifs (figure 50), entre 350 et 500 ms, des composantes négatives sur les régions antérieures sont enregistrées pour les deux classes de mots (figure 51). Si l'on observe plus précisément le déroulement temporel des PEs sur quelques électrodes frontales (F3, Fz, F4,

voir figure 52), on note qu'une composante négative est enregistrée pour les deux classes de mots et ce, uniquement après 300 ms.



**Figure 52:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour 3 électrodes localisées en région frontale: F3 (gauche), Fz (centrale), et F4 (droite). Les tracés en gras correspondent aux PEs des mots de classe fermée (CF), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte (CO). La négativité est située vers le haut.

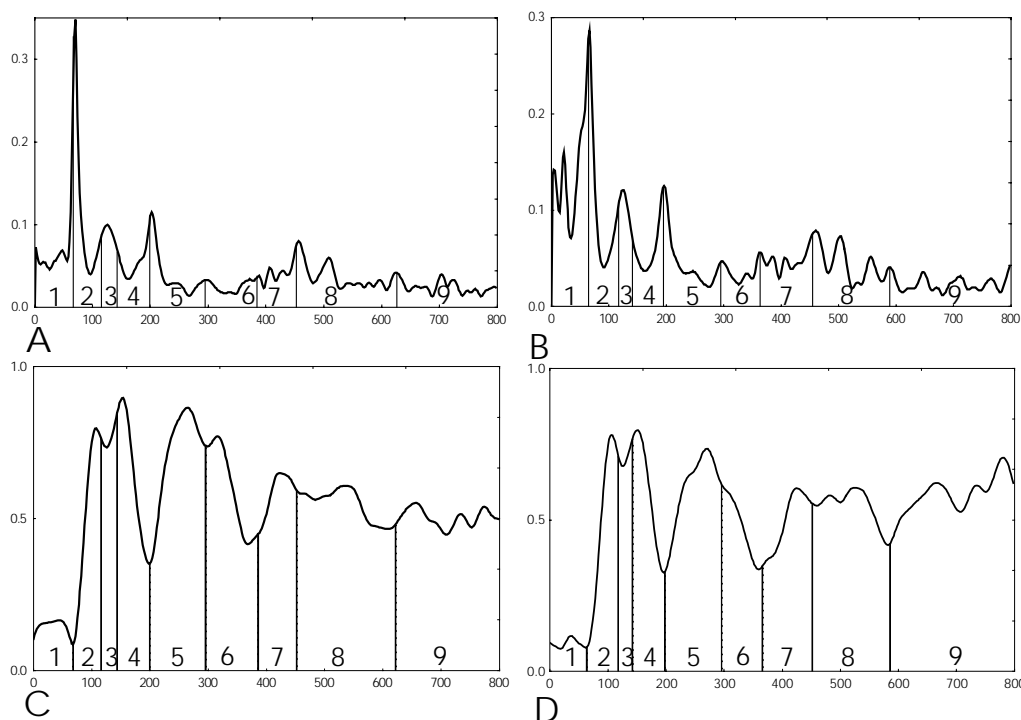
Suite à ces observations, l'analyse de latence de pic de la composante négative a donc été réalisée sur le premier pic négatif compris entre 300 et 400 ms. Nous avons reconduit la même analyse que celle décrite dans les sections 3.1.1.3 des chapitres 4 et 5. La latence des composantes négatives a été analysée sur les maximums de négativité de 14 électrodes incluses dans les régions antérieures (FP1, FP2, AF7, AF3, AF4, AF8, F7, F5, F3, F1, F2, F4, F6, F8, en excluant FPz, et Fz). Une ANOVA à mesures répétées, avec les facteurs "classe de mots" (CF vs. CO), "hémisphère" (droit vs. gauche) et "électrodes" (7), montre un effet significatif du facteur "classe de mots" ( $F(1,19) = 5.34, p < .03$ ): la latence de la composante négative pour les mots de classe fermée (CF:  $316 \pm 4.6$  ms) est plus précoce (de 7 ms) que celle enregistrée par les mots de classe ouverte (CO:  $323 \pm 4.8$  ms). Le facteur "hémisphère" ainsi que l'interaction "classe de mots" x "hémisphère" ne sont pas significatifs ( $F < 1$ ): la latence de la composante négative pour les deux classes de mots est identique pour les électrodes distribuées sur l'hémisphère gauche (CF:  $315.7 \pm 4.3$  ms, CO:  $322.3 \pm 4.8$  ms) et droit (CF:  $316.3 \pm 4.9$  ms, CO:  $323.6 \pm 4.9$  ms).

Les résultats des analyses classiques de tracés indiquent que, dans la fenêtre temporelle de la composante N280 (250-350 ms), les PEs associées aux mots CF sont plus négatifs sur les régions antérieures et, plus positifs sur les régions postérieures que ceux des mots CO. Entre 350-500 ms, les PEs des mots CO sont plus négatifs sur toutes les régions du scalp par rapport à ceux des mots de classe fermée. De plus, la composante négative des mots CF apparaît plus précocement (7 ms) que celle des mots CO.

### 3.1.2 Analyse de segmentation adaptée

#### 3.1.2.1 Détermination des micro-états

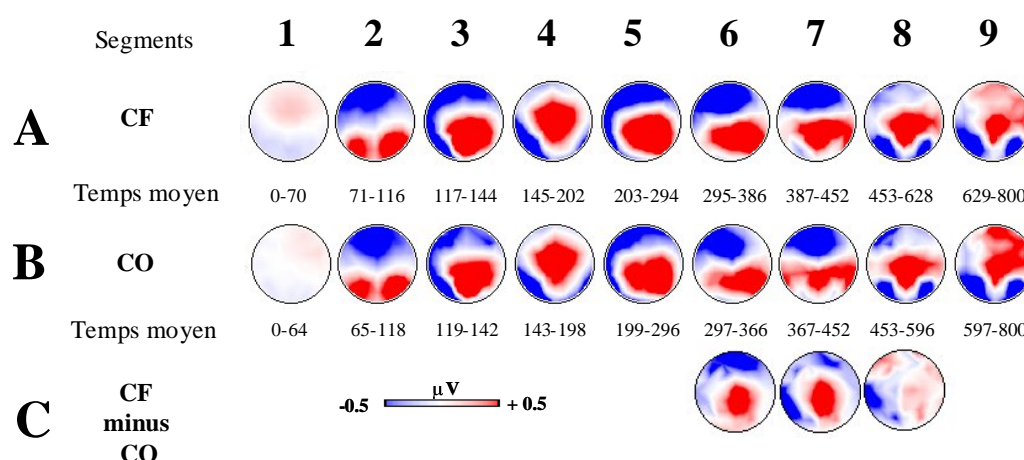
Sur la période de 800 ms correspondant à la fenêtre d'analyse, 9 périodes de stabilité ont été déterminées séparément, pour chaque classe de mots. Ainsi, chaque période équivaut à un segment ou micro-état (figure 53).



**Figure 53:** (A & B) courbes de *Global Dissimilarity* (GD) pour les grandes moyennes (800 ms d'analyse), pour les PEs associés aux mots de classe fermée (A) et de classe ouverte (B). (C & D), courbes représentant le *Global Field Power* (GFP) pour les grandes moyennes pour les mots de classe fermée (C) et de classe ouverte (D). Les traits verticaux indiquent les frontières d'un segment défini par la méthode d'analyse de segmentation adaptée. Les nombres se réfèrent à l'ordre d'apparition des segments, ou micro-états dans le temps. La segmentation adaptée révèle que, pour chaque condition expérimentale, 9 configurations différentes et stables sont présentes.

Pour chaque segment, une seule carte moyenne est calculée à partir des PEs des grandes moyennes et ce, dans les fenêtres temporelles définies par la segmentation adaptée. De plus, pour chaque segment, les frontières sont aussi calculées à partir des grandes moyennes. La figure 54 présente les différentes topographies, pour chaque segment défini. Le segment 2 correspond à la composante P100, qui est généralement plus large sur les sites occipitaux latéraux (PO7, PO8), et présente un pic vers 100 ms. La composante P100 est ici latéralisée sur les sites postérieurs droits, et est enregistrée entre 71-116 ms pour les mots CF et entre 65-118 ms pour les mots CO. Le segment 3 correspond à la composante P130 dont le maximum

de positivité est situé sur les sites centro-postérieurs (Pz, CPz) légèrement latéralisée sur les régions droites. Le segment 3 est enregistré entre 117-144 ms pour les mots CF, et entre 119-142 ms pour les mots CO. Le segment 4 correspond à l'occurrence de la composante N200 maximale sur les sites postérieurs latéraux, et latéralisée à gauche. Sa latence est généralement de 160 ms. Le segment 4 est enregistré entre 145-202 ms pour les mots CF, et entre 143-198 ms pour les mots CO. La distribution sur le scalp, ainsi que les rangs d'apparition de ces composantes visuelles précoces (définies par la segmentation adaptée), correspondent à la description faite par Curran, *et al.* (1993) et Nobre & McCarthy (1994).



**Figure 54:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales (A, classe fermée: CF; B, classe ouverte: CO) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. La couleur bleue représente des PEs négatifs, alors que les PEs positifs sont représentés en rouge. Les frontières (en ms) des segments sont indiquées à partir des données des grandes moyennes. En C, sont représentées les cartes de différences de potentiels (classe fermée minus classe ouverte) pour les segments 6, 7, et 8 où les topographies sont significativement différentes entre les deux conditions expérimentales.

### 3.1.2.2 Détermination des micro-états spécifiques

Afin de tester si les cartes des PEs associés aux mots de classe fermée sont spécifiques à cette condition expérimentale, chaque carte de surface de la condition CF a été introduite dans les données individuelles des sujets pour les mots de classe fermée, mais aussi pour ceux de la classe ouverte. La procédure de "fitting" a été réalisée séparément pour chaque segment, et dans une fenêtre temporelle définie sur la base des valeurs ( $\pm 10$  ms) obtenues sur la grande

moyenne (segment 1: 0-80 ms, segment 2: 56-130 ms, segment 3: 108-154 ms, segment 4: 134-212 ms, segment 5: 190-308 ms, segment 6: 280-400 ms, segment 7: 358-468 ms, segment 8: 440-640 ms, segment 9: 588-800 ms). La valeur du BEV (*Best Explained Variance* ou maximum de pourcentage de variance expliquée) pour chaque carte de la condition CF obtenue dans les deux conditions expérimentales de chaque participant est alors comparée statistiquement à l'aide d'un *t*-test à échantillon apparié. Les résultats montrent que seulement trois des neuf segments prédéterminés semblent spécifiques à la classe fermée (CF). Ainsi, la carte 6, obtenue à partir des mots de classe fermée (CF), explique mieux les données PEs des mots CF ( $64.1 \pm .04$  %) que celles des mots CO ( $54.9 \pm .04$  %;  $t = -2.43$ ,  $p < .02$ ). Des résultats identiques sont obtenus pour les segments 7 (CF:  $54.6 \pm .04$  %, CO:  $46.3 \pm .05$  %;  $t = -2.46$ ,  $p < .02$ ) et 8 (CF:  $55.7 \pm .04$  %, CO:  $49.9 \pm .04$  %;  $t = -2.24$ ,  $p < .03$ ). Les frontières (début et fin du micro-état) des 3 cartes spécifiques sont définies sur la base des grandes moyennes, pour les deux classes de mots: 6 (295-386 ms; 297-366 ms), 7 (387-452 ms; 367-452 ms), et 8 (453-628 ms; 453-596 ms).

### 3.1.2.3 Analyse topographique par région d'intérêt

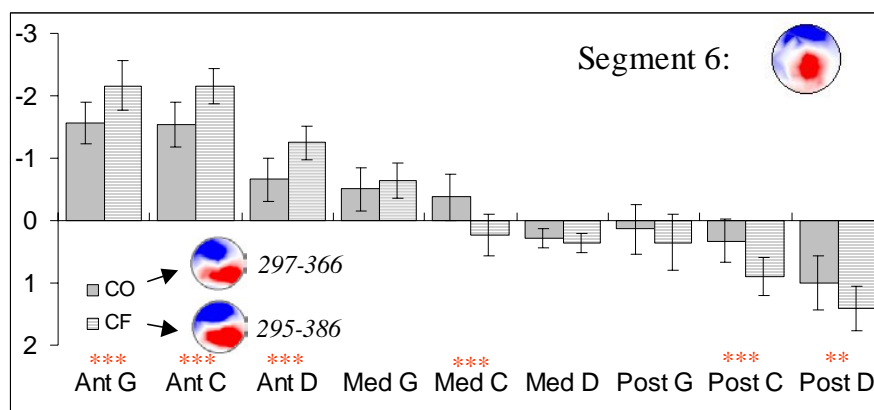
Pour chaque segment spécifique aux mots de classe fermée (6, 7 et 8), nous avons effectué une ANOVA à mesures répétées, avec les facteurs suivants: "classe de mots" (fermée vs. ouverte); "caudalité" (antérieur vs. médian vs. postérieur); "latéralité" (gauche vs. centrale vs. droite). Les régions d'intérêt utilisées sont les mêmes que celles utilisées lors de l'analyse de tracés, et sont décrites dans le tableau 24. Les résultats de ces analyses sont présentés dans le tableau 26.

	ddl	6	7	8
Classe x Caudalité	2, 38	<b>5.01 *</b> - <b>4.57 *</b>	1.37 - 0.82	0.01 - 0.50
Classe x Latéralité	2, 38	2.06 - 0.68	<b>5.98 **</b> - <b>4.47 *</b>	1.91 - 1.17
Classe x Caudalité x Latéralité	4, 76	<b>2.49 *</b> - <b>2.94 *</b>	<b>5.30 **</b> - <b>4.61 **</b>	<b>4.33 **</b> - <b>5.01 **</b>

**Tableau 26:** Résultats des ANOVAs à mesures répétées séparées pour chaque segment (6, 7 et 8). Les valeurs de *F* sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

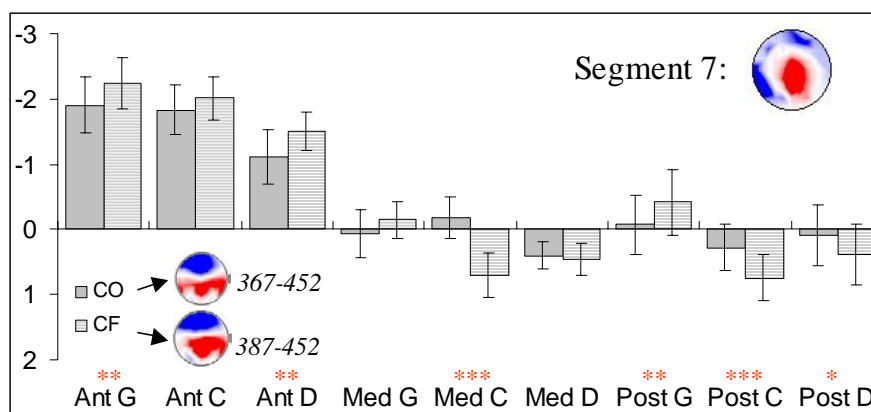
Afin d'éclaircir les différentes interactions significatives observées, nous avons réalisé des comparaisons planifiées que nous allons maintenant décrire (*Post-Hoc*, LSD), pour chaque segment.

- Segment 6 (CF: 295-386 ms, CO: 297-366 ms): dans cette fenêtre temporelle, les PEs montrent une négativité maximale sur les régions antérieures ("caudalité";  $F(2,38) = 13.45$ ;  $p < .0001$ ) à gauche et en central ("latéralité";  $F(2,38) = 7.87$ ;  $p < .001$ ). L'interaction "classe de mots" x "caudalité" ( $F(2,38) = 4.58$ ;  $p < .01$ ) indique que les PEs des mots CF sont plus négatifs ( $-1.85 \mu\text{V}$ ) sur les régions antérieures par rapport à ceux des mots CO ( $-1.25 \mu\text{V}$ ;  $p < .01$ ). Les comparaisons planifiées (*Post-Hoc*, LSD, voir figure 55) de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 2.94$ ;  $p < .02$ ) montrent que les PEs des mots CF sont plus négatifs sur les régions antérieures (les 3) que ceux des mots CO, alors que sur les régions postéro-médianes centrale et postérieure droite, l'effet inverse est observé.



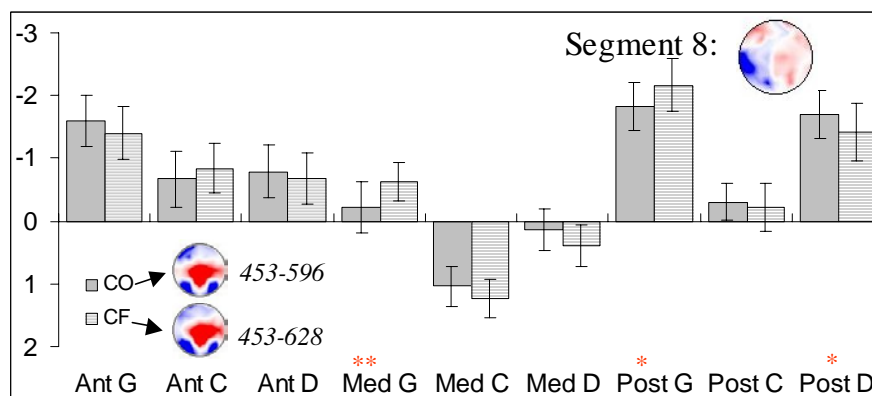
**Figure 55:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 6 (compris entre 297-366 ms pour CO et 295-386 ms pour CF). Les valeurs de PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF et CO (en bas) ainsi que la carte de différences (CF minus CO, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

- Segment 7 (CF: 387-452 ms, CO: 367-452 ms): les PEs présentent une négativité maximale sur les sites antérieurs ("caudalité";  $F(2,38) = 10.20$ ;  $p < .0003$ ). L'analyse de l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 4.48$ ;  $p < .01$ ) indique que les PEs associés aux mots CF sont plus négatifs ( $-0.93 \mu\text{V}$ ) à gauche que ceux des mots CO ( $-0.63 \mu\text{V}$ ;  $p < .04$ ) alors, qu'en région centrale, l'effet inverse est observé (CF:  $-0.18 \mu\text{V}$ , CO:  $-0.58 \mu\text{V}$ ;  $p < .002$ ). L'analyse (*Post-Hoc*, LSD, voir figure 56) de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 4.62$ ;  $p < .002$ ) montre que les PEs des mots CF sont plus négatifs en régions antérieures gauche et droite ainsi qu'en postérieur gauche et plus positifs en régions postéro-médiane centrale et postérieure droite que ceux des mots CO.



**Figure 56:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 7 (compris entre 367-452 ms pour CO et 387-452 ms pour CF). Les valeurs du PE (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF et CO (en bas), ainsi que la carte de différences (CF minus CO, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

- Segment 8 (CF: 453-628 ms, CO: 453-595 ms): dans cette fenêtre temporelle, les PEs montrent des négativités à la fois sur les régions antérieures et postérieures à gauche et à droite ("caudalité";  $F(2,38) = 7.50$ ;  $p < .001$ ; "latéralité";  $F(2,38) = 23.53$ ;  $p < .0001$ ). Les comparaisons planifiées (*Post-Hoc*, LSD, voir figure 57) de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 5.01$ ;  $p < .001$ ) montrent que les PEs des mots CF sont plus négatifs en régions postéro-médianes gauches que ceux des mots CO. En région postérieure droite, les PEs des mots CO sont plus négatifs que ceux enregistrés par les mots CF.



**Figure 57:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 8 (compris entre 453-596 ms pour CO et 453-628 ms pour CF). Les valeurs du PE (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale;



ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF et CO (en bas) ainsi que la carte de différences (CF minus CO, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.1.2.4 Evolution temporelle des micro-états

Les analyses sur la latence des pics de GFP ont été réalisées pour les segments 5, 6, 7 et 8. Pour chaque segment, une ANOVA à mesures répétées avec comme facteur "classe de mots" (fermée vs. ouverte) a été réalisée. Les résultats sont décrits dans le tableau 27. Deux cartes sont enregistrées plus précocement pour les mots CF par rapport aux mots CO. Ainsi, la carte 5 (comprise entre 203-294 ms pour les mots CF et entre 199-296 ms pour les mots CO) est significativement (-3.9 ms) plus précoce pour les mots CF que pour les mots CO. La carte 6 (comprise entre 295-386 ms pour les mots CF et entre 297-366 ms pour les mots CO) est enregistrée 8.8 ms plus précocement pour les mots CF.

Segment	Classe Fermée	Classe ouverte	ANOVA
5	<b>254.4 (4.21)</b>	<b>258.3 (5.23)</b>	<b><math>F(1,19) = 4.65; p &lt; .04</math></b>
6	<b>324.3 (4.17)</b>	<b>333.1 (5.41)</b>	<b><math>F(1,19) = 4.88; p &lt; .03</math></b>
7	425.4 (4.06)	426.1 (4.63)	$F(1,19) = .02; p > .88$
8	540.4 (7.06)	542.1 (6.12)	$F(1,19) = .07; p > .79$

**Tableau 27:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments (5, 6, 7 et 8). Les valeurs indiquées sont en ms, avec, entre parenthèses, l'erreur type. Les effets significatifs sont rapportés en gras.

### 3.1.3 Synthèse CF vs. CO

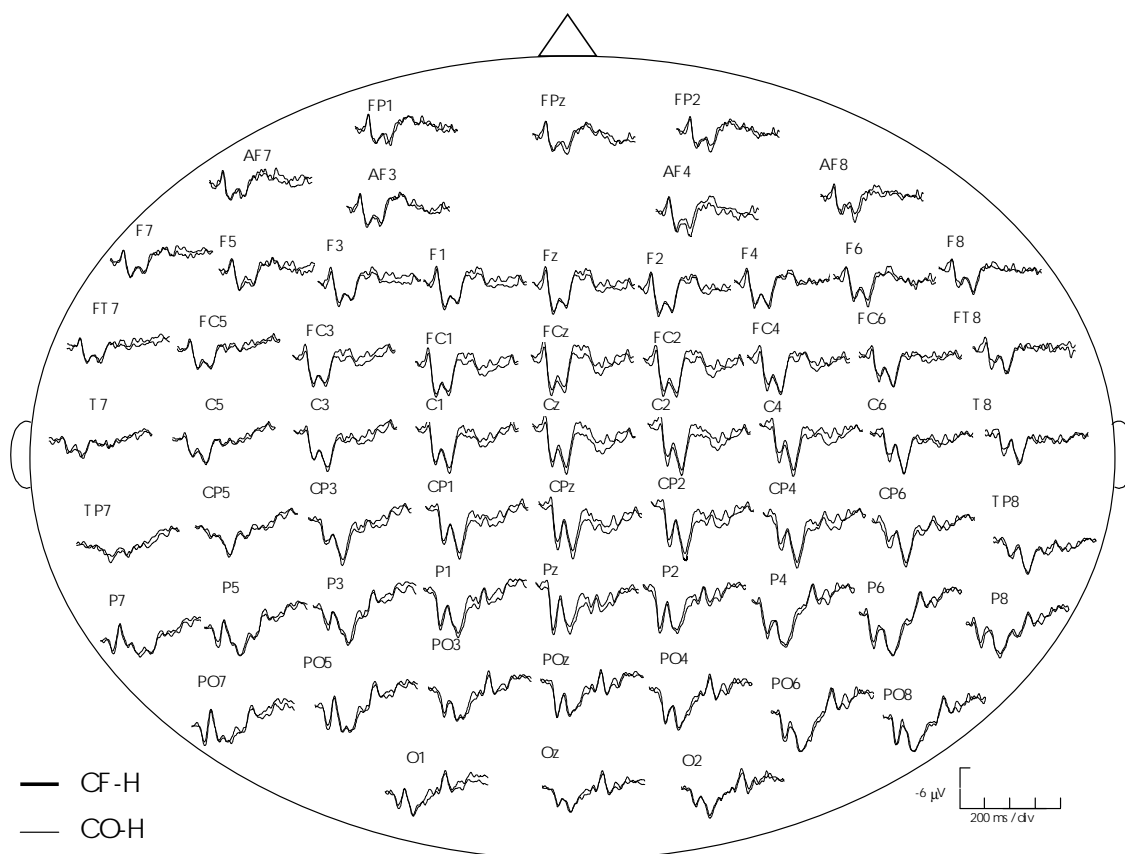
La segmentation adaptée montre que 3 segments (6, 7, 8) présentent des différences topographiques significatives entre les deux classes de mots. Alors que, pour le segment 6, des négativités sur les régions antérieures gauche et centrale sont enregistrées pour les deux classes de mots, celle accompagnant les mots CF est plus importante sur les sites antérieurs. Dans la fenêtre temporelle du segment 7, des négativités sont enregistrées en régions antérieures, et différemment latéralisées pour les deux classes de mots: les PEs des mots CF sont plus négatifs à gauche que ceux des mots CO. Pour le dernier segment 8, deux négativités sont observées: une première sur les régions antérieures avec un maximum à

gauche, et une deuxième en régions postérieures, avec un maximum à droite. La négativité observée en région antérieure est plus importante pour les mots CF, que pour les mots CO sur les sites postéro-médians gauches. En région postérieure droite, se sont les PEs des mots CO qui sont plus négatifs que ceux des mots CF. L'analyse du maximum de GFP indique que deux cartes correspondant aux segments 5 et 6, sont enregistrées plus précocement pour les mots de classe fermée (3.9 et 8.8 ms respectivement) que pour les mots CO.

## 3.2 Comparaison CF vs. CO haute fréquence

### 3.2.1 Analyse des tracés

#### 3.2.1.1 Visualisation des tracés



**Figure 58:** Grande moyenne (N=20) des PEs en référence mastoïde gauche pour les 62 électrodes enregistrées. Les tracés en gras correspondent aux PEs des mots de classe fermée de haute fréquence (CF-H), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte de haute fréquence (CO-H). Les électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sont sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, et O pour occipital. La négativité est située vers le haut.

La figure 58 présente les PEs pour les mots de haute fréquence de classe fermée (CF-H) et de classe ouverte (CO-H). Par rapport à la figure 49 où les mots de haute et de basse fréquences sont moyennés ensemble, les effets semblent plus réduits dans le temps, mais aussi sur le scalp. De légères différences semblent être présentes en région antérieure droite: les mots de classe fermée semblent plus négatifs que les mots de classe ouverte. Sur les quelques électrodes des sites centraux, nous retrouvons une négativité qui semble plus importante pour les mots de classe ouverte, entre 200 et 400 ms, par rapport aux mots de classe fermée.

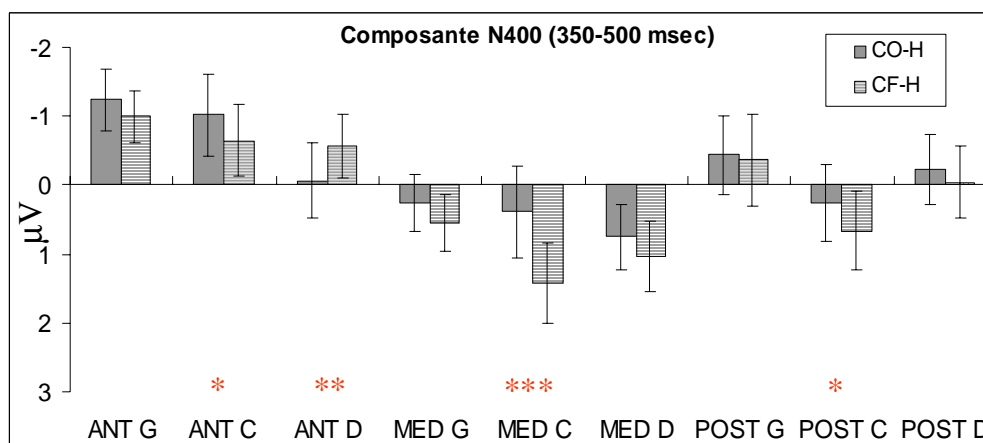
### 3.2.1.2 Analyse topographique par région d'intérêt

Les résultats des ANOVAs à mesures répétées, avec les facteurs "classe de mots" (CO-H vs. CF-H), "caudalité" (antérieur vs. médian vs. postérieur) et "latéralité" (gauche vs. central vs. droit) sont présentés dans le tableau 28, pour les 4 fenêtres temporelles: P2 (150 et 250 ms), N280 (250 et 350 ms), N400 (350 et 500 ms), et LAN (500 et 800 ms).

	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	0.01 - /	0.24 - /	0.55 - /	0.93 - /
Classe x Caudalité	2, 38	0.47 - <i>1.25</i>	2.48 - <i>2.57</i>	0.74 - <i>0.53</i>	0.31 - <i>0.15</i>
Classe x Latéralité	2, 38	0.72 - <i>0.7</i>	1.69 - <i>1.12</i>	2.49 - <i>1.68</i>	1.72 - <i>1.43</i>
Classe x Caudalité x Latéralité	4, 76	1.21 - <i>1.19</i>	1.19 - <i>1.28</i>	<b>2.25 - 2.72 *</b>	1.22 - <i>1.27</i>

**Tableau 28:** Résultats des ANOVAs à mesures répétées pour les différentes fenêtres d'analyses (150-250 ms, 250-350 ms, 350-500 ms et 500-800 ms) sur l'amplitude du PEs, entre les deux classes de mots de haute fréquence (CO-H vs. CF-H). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Dans la fenêtre d'analyse correspondant à la composante N400 (350-500 ms), une négativité est enregistrée avec un maximum en région antérieure ("caudalité";  $F(2,38) = 5.59$ ;  $p < .007$ ) et latéralisée à gauche ("latéralité";  $F(2,38) = 3.90$ ;  $p < .02$ ). Les comparaisons planifiées (*Post Hoc* LSD, figure 59) de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 2.72$ ;  $p < .03$ ) indiquent que les PEs des mots CF-H sont plus négatifs, en régions antérieures droites et, plus positifs en régions postéro-médianes centrales, que ceux associés aux mots CO-H. Les PEs des mots CO-H sont plus négatifs en antérieur central par rapport aux PEs des mots CH-H.



**Figure 59:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N400 (comprise entre 350 et 500 ms). Les valeurs des PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-H) et ouverte (CO-H) de haute fréquence sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.2.1.3 Analyse de latence

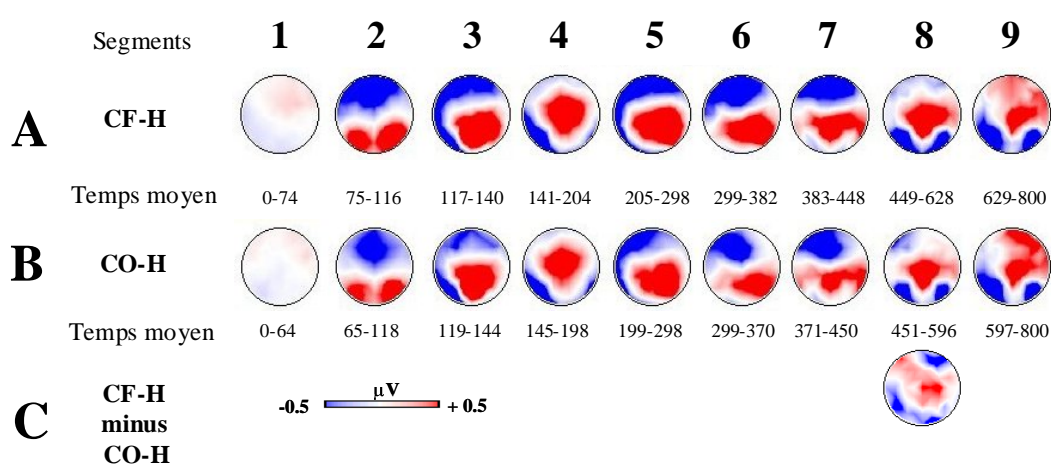
L'ANOVA à mesures répétées avec les facteurs "classe de mots" (CF-H vs. CO-H), "hémisphère" (droit vs. gauche) et "électrodes" (7) montre un effet significatif du facteur "classe de mots" ( $F(1,19) = 22.47, p < .0001$ ): la latence de la composante négative est plus précoce (de 11.1 ms) pour les mots de classe fermée (CF-H:  $314.8 \pm 4.6$  ms) que celle enregistrée pour les mots de classe ouverte (CO-H:  $325.9 \pm 4.7$  ms). Le facteur "hémisphère" ( $F(1,19) = 0.43, p > .52$ ) ainsi que l'interaction "classe de mots" x "hémisphère" ( $F(1,19) = 0.92, p > .35$ ) ne sont pas significatifs: la latence de la composante négative est identique sur l'hémisphère gauche (CF-H:  $313.7 \pm 4.3$  ms, CO-H:  $326.1 \pm 4.4$  ms) et droit (CF-H:  $315.9 \pm 4.9$  ms, CO-H:  $325.8 \pm 4.9$  ms).

L'analyse de tracés indique que, pour les mots de fréquence supérieure à 400 occurrences par million ( $> 2.6$  log de fréquence), les différences topographiques entre les deux classes sont enregistrées dans un rang de latence tardif (350-500 ms). Un décalage de latence (11.1 ms) indiquant des effets plus précoces pour les mots CF que pour les mots CO, est aussi observé.

### 3.2.2 Analyse de segmentation adaptée

#### 3.2.2.1 Détermination des micro-états

La procédure de segmentation adaptée a déterminé 9 segments différents pour chaque condition expérimentale. Les segments, ainsi que leur carte moyenne et leurs frontières, sont présentés dans la figure 60.



**Figure 60:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales (A, classe fermée haute fréquence: CF-H; B, classe ouverte haute fréquence: CO-H) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, est représentée la carte de différence de potentiels (classe fermée minus classe ouverte) pour le segment 8, où les topographies sont significativement différentes entre les deux conditions expérimentales.

#### 3.2.2.2 Détermination des micro-états spécifiques

La procédure de "fitting" a été réalisée comme décrite dans la section 3.1.2.2 de ce chapitre. Les résultats montrent qu'une seule carte (segment 8: 449-628 ms pour CF-H et 451-596 ms pour CO-H) présente des topographies significativement différentes entre les mots de classe fermée et ouverte de haute fréquence. La carte représentative du segment 8, obtenue à partir des données CF-H, explique mieux les données CF-H ( $57.5 \pm .03$  %) que celle des mots CO-H ( $43.9 \pm .04$  %;  $t = -2.97$ ,  $p < .007$ ). Aucun autre segment ne montre d'effet significatif.

Pour ce segment spécifique, les latences des frontières, et la topographie de surface sont illustrées dans la figure 60.

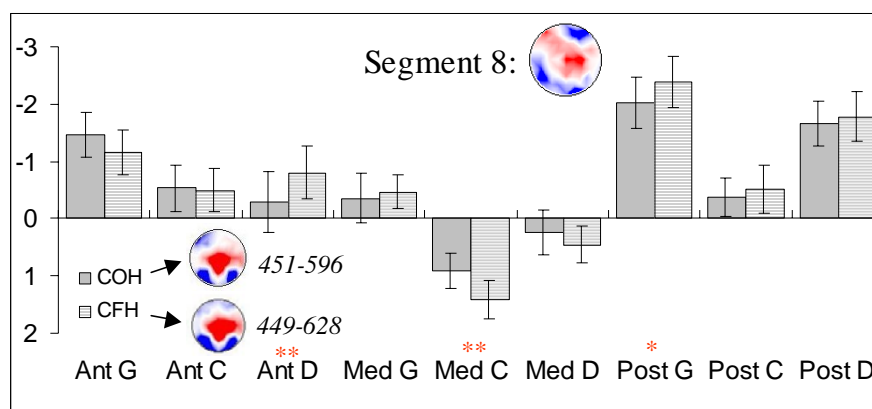
### 3.2.2.3 Analyse topographique par région d'intérêt

Pour le segment spécifique aux mots de classe fermée de haute fréquence (segment 8), nous avons effectué une ANOVA à mesures répétées avec les facteurs suivants: "classe de mots" (fermée vs. ouverte); "caudalité" (antérieur vs. médian vs. postérieur); "latéralité" (gauche vs. centrale vs. droite). Les résultats des différentes interactions sont présentés dans le tableau 29.

	ddl	8
Classe x Caudalité	2, 38	0.46 - <i>1.01</i>
Classe x Latéralité	2, 38	0.45 - <i>0.20</i>
Classe x Caudalité x Latéralité	4, 76	<b>3.30 *</b> - <b>4.13 **</b>

**Tableau 29:** Résultats des ANOVAs à mesures répétées pour le segment (8) spécifique aux mots de classe fermée de haute fréquence (CF-H). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Dans la fenêtre temporelle correspondant au segment 8, les PEs présentent des négativités à la fois sur les régions postérieure et antérieure ("caudalité";  $F(2,38) = 8.84$ ;  $p < .0007$ ) mais latéralisées à gauche ("latéralité";  $F(2,38) = 19.92$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 4.13$ ;  $p < .004$ ; voir figure 61) nous indique que les PEs des mots CF-H sont plus négatifs, sur les régions antérieure droite et postérieure gauche et, plus positifs en région médiane centrale que les mots CO-H.



**Figure 61:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 8 (compris entre 451-596 ms pour CO-H et 449-628 ms pour CF-H). Les valeurs des PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-H) et ouverte (CO-H) de haute fréquence sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche;

ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-H et CO-H (en bas) ainsi que la carte de différence (CF-H minus CO-H, en haut). \*\*\*,  $p < .001$ , \*\*,  $p < .01$ , \*,  $p < .05$ .

### 3.2.2.4 Evolution temporelle des micro-états

Les analyses portant sur la latence des pics de GFP (voir tableau 30) montrent que deux des segments sont enregistrés plus précocement pour les mots CF-H que pour les mots CO-H. Le segment 5 présente un décalage de 5.5 ms, alors que le décalage est plus important pour le segment 6: 9.4 ms.

Segment	CF-H	CO-H	ANOVA
5	253.0 (4.46)	258.5 (5.49)	<b><math>F(1,19) = 6.48; p &lt; .01</math></b>
6	322.8 (4.71)	332.2 (5.79)	<b><math>F(1,19) = 5.21; p &lt; .03</math></b>
7	421.7 (5.44)	422.8 (4.91)	$F(1,19) = .05; p > .82$
8	540.4 (7.62)	540.3 (6.75)	$F(1,19) = .00; p > .99$

**Tableau 30:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments (5, 6, 7 et 8). Les valeurs sont indiquées en ms, avec, entre parenthèses, l'erreur type. Les effets significatifs sont rapportés en gras.

### 3.2.3 Synthèse CF vs. CO haute fréquence

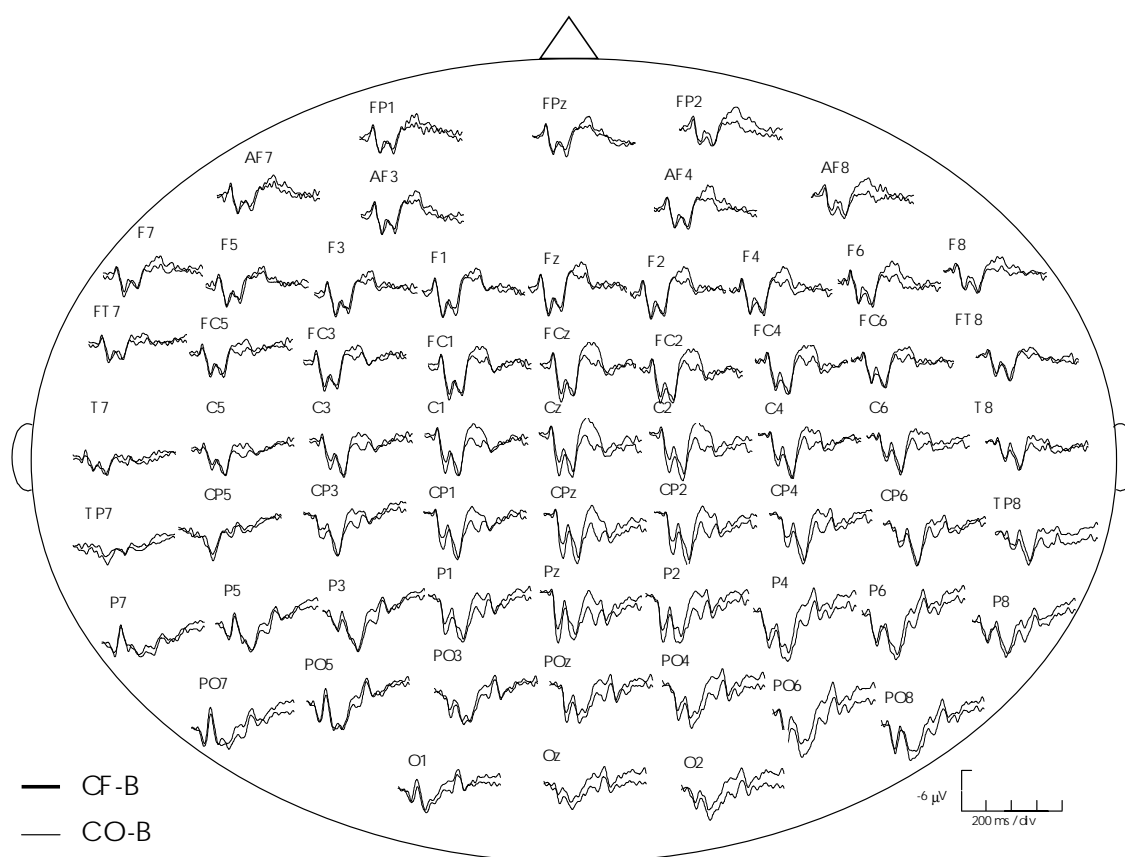
Les résultats de la segmentation adaptée suggèrent que, pour un segment tardif (numéro 8), des différences topographiques sont présentes en région antérieure droite et postérieure gauche entre les deux classes de mots. De plus, deux segments sont enregistrés plus précocement pour les mots CF-H, par rapport aux mots CO-H. Les différences de latence sont de 5.5 (segment 5) et 9.4 ms (segment 6).

## 3.3 Comparaison CF vs. CO basse fréquence

### 3.3.1 Analyse des tracés

### 3.3.1.1 Visualisation des tracés

La figure 62 présente les PE pour les mots de classe fermée (CF-B) et de classe ouverte (CO-B) de basse fréquence. Par rapport à la figure 58, où seuls les mots de haute fréquence sont pris en compte, les effets (c'est à dire la distinction entre mots CF-B et CO-B) semblent plus importants sur le scalp. Ainsi, l'on observe l'occurrence d'une négativité comprise entre approximativement 300 et 500 ms sur la majorité des électrodes pour les mots CO-B. Cet effet semble d'amplitude maximale sur les électrodes localisées en régions centrale et droite.



**Figure 62:** Grande moyenne (N=20) des PE en référence mastoïde gauche pour les 62 électrodes enregistrées. Les tracés en gras correspondent aux PE des mots de classe fermée de basse fréquence (CF-B), alors que les tracés en gris clair sont associés à la présentation des mots de classe ouverte de basse fréquence (CO-B). Les électrodes avec un nombre impair sont distribuées sur l'hémisphère gauche, alors que les électrodes portant les nombres pairs sur l'hémisphère droit. FP, AF, F pour frontal, FC, C et CP pour central, P et PO pour pariétal, et O pour occipital. La négativité est située vers le haut.



### 3.3.1.2 Analyse topographique par région d'intérêt

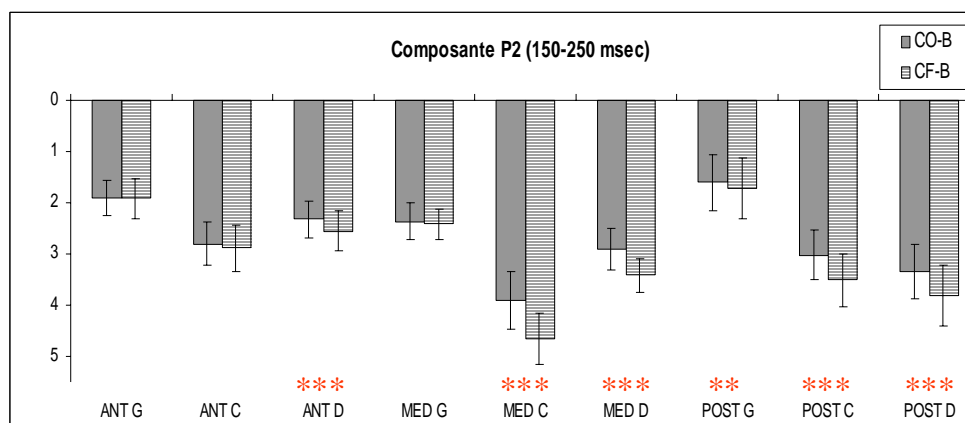
Les résultats des ANOVAs à mesures répétées, pour les 4 fenêtres temporelles: P2 (150 et 250 ms), N280 (250 et 350 ms), N400 (350 et 500 ms), et LAN (500 et 800 ms), sont présentés dans le tableau 31.

	ddl	P2 (150-250)	N280 (250-350)	N400 (350-500)	LAN (500-800)
Classe	1, 19	2.28 - /	1.18 - /	<b>7.42 * - /</b>	1.88 - /
Classe x Caudalité	2, 38	<b>0.88 - 3.50 *</b>	<b>7.14 *** - 5.63 **</b>	0.31 - 0.38	0.90 - 0.83
Classe x Latéralité	2, 38	3.36 * - 1.23	3.49 - 1.06	<b>3.79 * - 3.45 *</b>	1.38 - 0.10
Classe x Caudalité x Latéralité	4, 76	<b>5.52 *** - 3.99 **</b>	<b>4.73 ** - 3.82 **</b>	<b>3.41 * - 3.33 *</b>	<b>3.27 * - 3.25 *</b>

**Tableau 31:** Résultats des ANOVAs à mesures répétées pour les différentes fenêtres d'analyse (150-250 ms, 250-350 ms, 350-500 ms et 500-800 ms) sur l'amplitude du PEs entre les deux classes de mots de basse fréquence (CO-B vs. CF-B). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

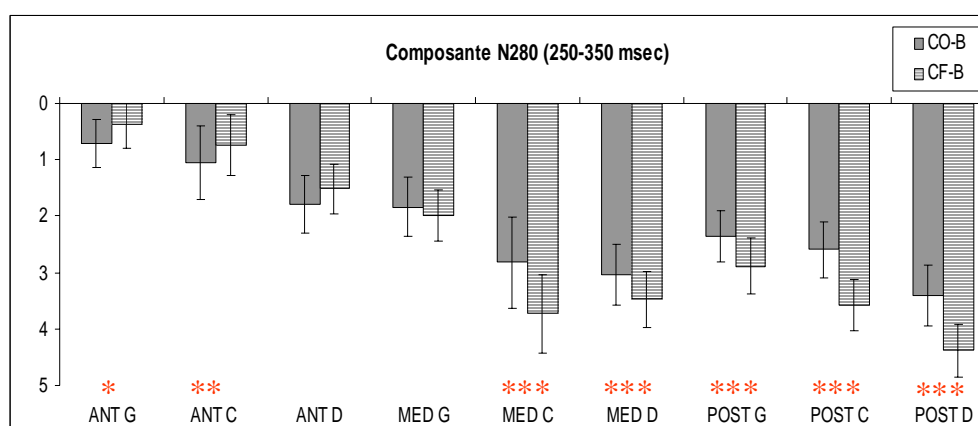
Pour chaque fenêtre d'analyse une interaction est au moins significative. Nous allons donc procéder à des comparaisons planifiées, afin de préciser les différences de topographie sur le scalp.

Entre 150 et 250 ms, une positivité est enregistrée avec un maximum sur les régions situées à gauche ("latéralité";  $F(2,38) = 18.92$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 3.50$ ;  $p < .04$ ) montre que les PEs des mots CF-B sont plus positifs sur les régions médiane (CO-B:  $3.02 \mu\text{V}$ , CF-B:  $3.67 \mu\text{V}$ ;  $p < .01$ ) et postérieure (CO-B:  $2.50 \mu\text{V}$ , CF-B:  $3.32 \mu\text{V}$ ;  $p < .001$ ), que ceux des mots CO-B. L'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 4.00$ ;  $p < .005$ ) indique que les PEs associés aux mots CF-B sont plus positifs sur les 3 régions situées à droites et en postéro-médiane centrale par rapport à ceux des mots CO-B (voir figure 63).



**Figure 63:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante P2 (comprise entre 150 et 250 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

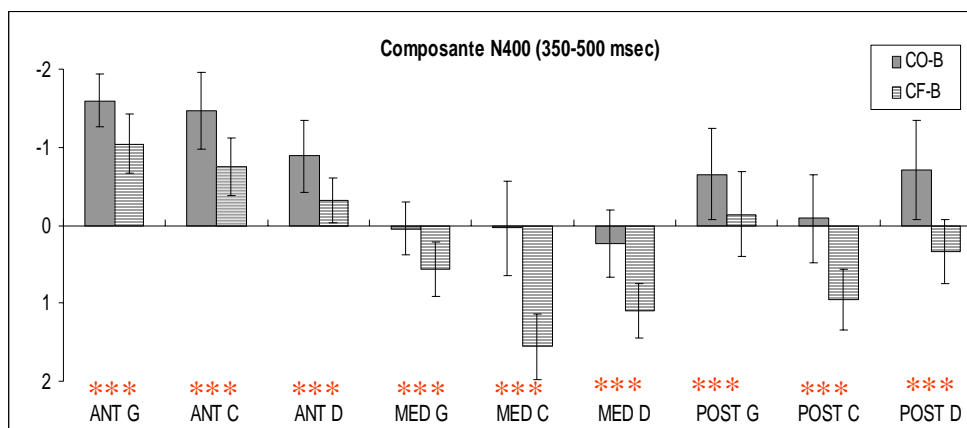
Entre 250 et 350 ms, une positivité est enregistrée avec un maximum sur les régions en postéro-médianes ("caudalité";  $F(2,38) = 14.91$ ;  $p < .0001$ ), au centre et à droite ("latéralité";  $F(2,38) = 9.45$ ;  $p < .0005$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 5.63$ ;  $p < .007$ ) montre que les PEs des mots CF-B sont plus positifs ( $3.83 \mu\text{V}$ ) en région postérieure que ceux des mots CO-B ( $2.69 \mu\text{V}$ ;  $p < .0001$ ). L'analyse de la triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 3.82$ ;  $p < .006$ , voir figure 64) indique que les PEs des mots CF-B sont plus négatifs en région antérieure gauche et centrale par rapport aux PEs des mots CO-B. En régions postérieures (3) ainsi qu'en médiane centrale et droite l'effet inverse est observé.



**Figure 64:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N280 (comprise entre 250 et 350 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) sont

indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

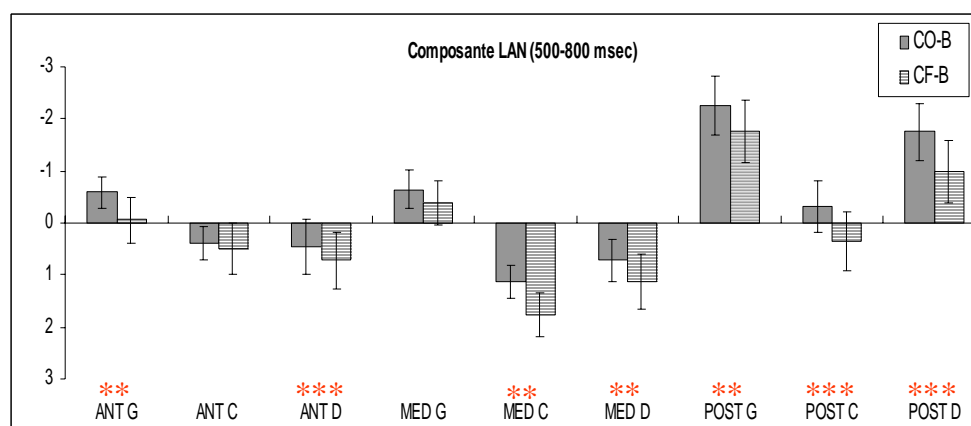
Entre 350 et 500 ms, une composante négative maximale est enregistrée sur les régions antérieures ("caudalité";  $F(2,38) = 8.53$ ;  $p < .0009$ ). L'effet du facteur principal "classe de mots" est significatif ( $F(1,19) = 7.43$ ;  $p < .01$ ), indiquant que les PEs des mots CO-B (- 0.97  $\mu V$ ) sont plus négatifs que ceux associés aux mots CF-B (0.42  $\mu V$ ). De plus, l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 3.45$ ;  $p < .04$ ) indique que cet effet est maximal sur les régions au centre et à droite. La triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 3.32$ ;  $p < .01$ ; figure 65) indique que les PEs des mots CO-B sont plus négatifs en régions antérieures et postérieures que ceux des mots CF-B. En région médiane, ce sont les PEs des mots CF-B qui sont plus positifs.



**Figure 65:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante N400 (comprise entre 350 et 500 ms). Les valeurs du PEs (en  $\mu V$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Entre 500 et 800 ms, une composante négative est enregistrée avec un maximum d'activité sur les régions postérieures ("caudalité";  $F(2,38) = 5.36$ ;  $p < .008$ ) et latéralisée à gauche ("latéralité";  $F(2,38) = 21.83$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 3.25$ ;  $p < .01$ , figure 66) indique que les PEs des mots CO-B sont plus négatifs en régions postérieures (3) et antérieure gauche par rapport à ceux

associés aux mots CF-B. En régions médianes centrale et droite ainsi qu'en région antérieure droite, les PEs des mots CF-B sont plus positifs que ceux des mots CO-B.



**Figure 66:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour la composante LAN (comprise entre 500 et 800 ms). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.3.1.3 Analyse de latence

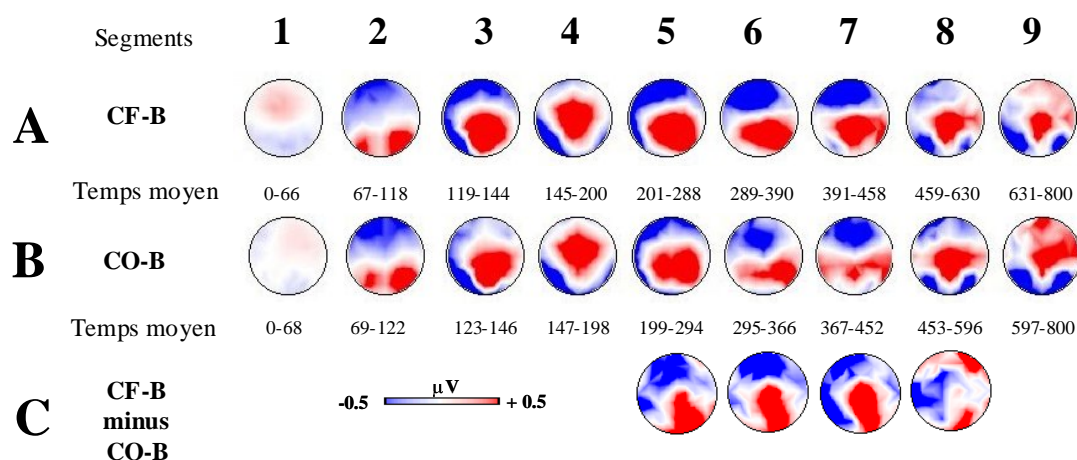
Les résultats de l'ANOVA à mesures répétées avec les facteurs "classe de mots" (CF-B vs. CO-B), "hémisphère" (droit vs. gauche) et "électrodes" (7) montrent que l'effet du facteur principal "classe de mots" n'est pas significatif ( $F(1,19) = 2.57, p > .13$ ): la latence de la composante négative est similaire pour les deux classes de mots (CF-B:  $314.2 \pm 4.4$  ms; CO-B:  $318.9 \pm 4.6$  ms). De plus, l'effet du facteur "hémisphère" et l'interaction "classe de mots" x "hémisphère" ne sont pas significatifs ( $F(1,19) = 2.13, p > .16$ ;  $F(1,19) = 0.64, p > .43$ , respectivement): la latence de la composante négative est équivalente pour les électrodes localisées sur les deux hémisphères quelque soit la classe considérée (hémisphère gauche: CF-B,  $312.4 \pm 4.2$  ms et CO-B,  $317.9 \pm 4.9$  ms; hémisphère droit: CF-B,  $315.9 \pm 4.6$  ms et CO-B:  $319.9 \pm 4.2$  ms).

Pour les mots ayant une fréquence inférieure à 400 occurrences par million ( $< 2.6 \log$  de fréquence), l'effet de classe se traduit par des différences topographiques dans les rangs de latence correspondant aux composantes P200, N280, N400 et LAN. Aucun effet de latence n'est observé.

### 3.3.2 Analyse de segmentation adaptée

#### 3.3.2.1 Détermination des micro-états

La procédure de segmentation adaptée a permis de mettre en évidence 9 segments différents pour chaque condition expérimentale. Les segments, leurs cartes moyennes et leurs frontières (valeur de début et de fin) sont présentés dans la figure 67.



**Figure 67:** Cartes PEs de surface des grandes moyennes (N=20) pour les deux conditions expérimentales (A, classe fermée basse fréquence: CF-B; B, classe ouverte basse fréquence: CO-B) définies par la procédure de segmentation adaptée. Les cartes sont représentées dans l'ordre d'apparition dans le temps (de gauche à droite). La partie gauche des cartes représente les électrodes situées à gauche. Les frontières (en ms) des segments sont indiquées sur la base des données des grandes moyennes. En C, sont représentées les cartes de différences de potentiels (CF-B minus CO-B) pour les 4 segments où les topographies sont significativement différentes entre les deux conditions expérimentales.

#### 3.3.2.2 Détermination des micro-états spécifiques

Les résultats de la procédure de "fitting" montrent les deux classes de mots induisent 4 cartes (segments 5, 6, 7 et 8) avec des topographies significativement différentes. Pour les mots CF-B, la carte représentative du segment 5 explique mieux les PEs des données CF-B ( $69.6 \pm .04$  %) par rapport à ceux des données CO-B ( $61.1 \pm .04$  %;  $t = -2.31$ ,  $p < .03$ ). Le même résultat est obtenu pour le segment 6 ( $64.9 \pm .04$  % pour les PEs des mots CF-B,  $56.1 \pm .04$  % pour les PEs des mots CO-B;  $t = -2.19$ ;  $p < .04$ ), 7 ( $56.1 \pm .05$  % pour les PEs des mots CF-B,  $45.9 \pm$

.04 % pour les PEs des mots CO-B;  $t = -2.40$ ;  $p < .02$ ) et 8 ( $53.1 \pm .03$  % pour les PEs des mots CF-B,  $46.1 \pm .04$  % pour les PEs des mots CO-B;  $t = -2.58$ ;  $p < .01$ ). Aucun autre segment ne montre d'effet significatif. Pour les 4 segments spécifiques, les latences des frontières, et leurs topographies de surface sont illustrées dans la figure 67.

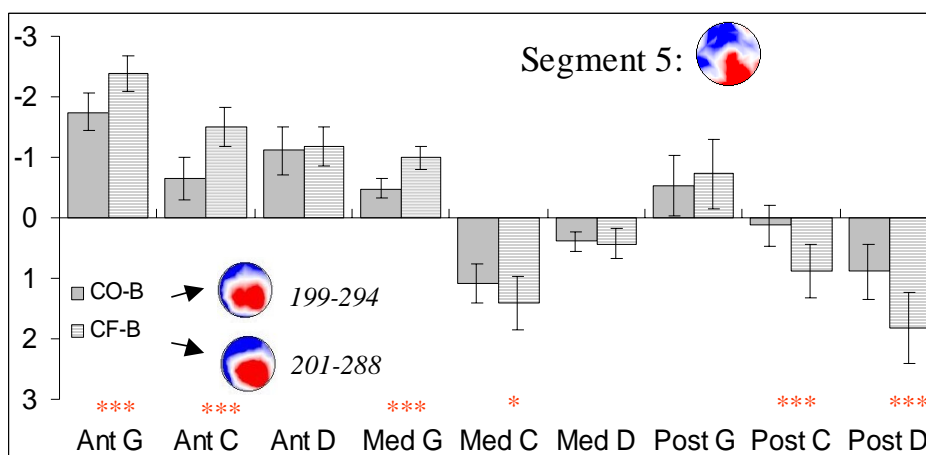
### 3.3.2.3 Analyse topographique par région d'intérêt

Pour les segments spécifiques aux mots de classe fermée de basse fréquence (segments 5, 6, 7 et 8), nous avons effectué une ANOVA à mesures répétées avec les facteurs suivants: "classe de mots" (fermée vs. ouverte); "caudalité" (antérieur vs. médian vs. postérieur); "latéralité" (gauche vs. centrale vs. droite). Les résultats des différentes interactions sont présentés dans le tableau 32.

	ddl	5	6	7	8
Classe x Caudalité	2, 38	<b>3.85 * - 3.81 *</b>	<b>4.81 * - 6.15 **</b>	1.13 - 0.85	0.40 - 0.44
Classe x Latéralité	2, 38	4.25 * - 1.37	1.57 - 0.60	<b>6.37 ** - 6.41 **</b>	3.49 - 2.83
Classe x Caudalité x Latéralité	4, 76	<b>6.39 *** - 4.78 **</b>	<b>3.47 * - 3.80 *</b>	<b>5.95 *** - 5.62 **</b>	<b>3.27 * - 3.40 *</b>

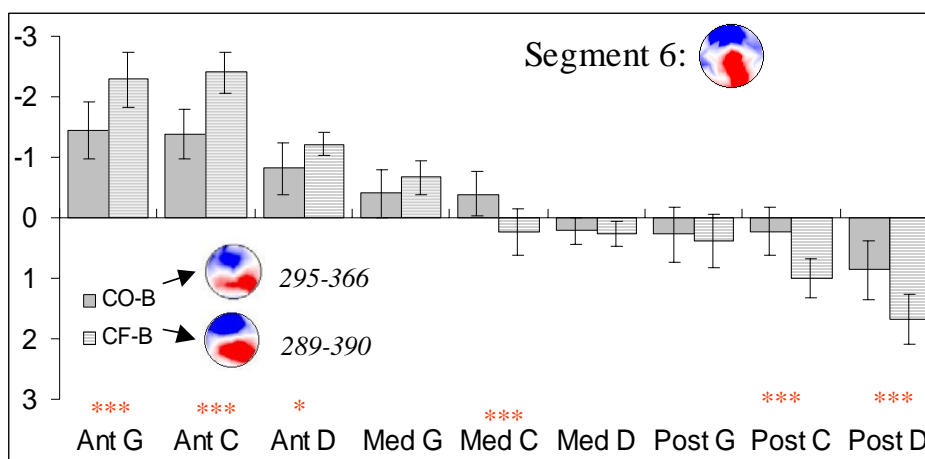
**Tableau 32:** Résultats des ANOVAs à mesures répétées pour les 4 segments (5, 6, 7 et 8) spécifiques aux mots de classe fermée de basse fréquence (CF-B). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Pour le segment 5, une négativité est enregistrée avec un maximum sur les régions antérieures ("caudalité";  $F(2,38) = 7.76$ ;  $p < .001$ ) et latéralisée à gauche ("latéralité";  $F(2,38) = 23.39$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 3.82$ ;  $p < .03$ ) montre que les PEs des mots CF-B sont plus négatifs ( $-1.67 \mu\text{V}$ ) en région antérieure que ceux des mots CO-B ( $-1.16 \mu\text{V}$ ;  $p < .05$ ). L'analyse de la triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 4.79$ ;  $p < .001$ , voir figure 68) indique que les PEs des mots CF-B sont plus négatifs en région antéro-médiane gauche et antérieure centrale par rapport à ceux des mots CO-B. Sur les régions postéro-médiane centrale et postérieure droite l'effet inverse est observé.



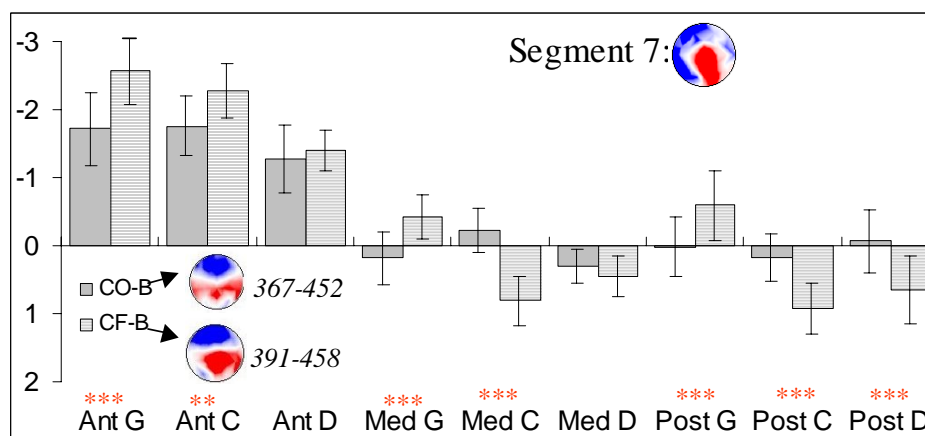
**Figure 68:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 5 (compris entre 199-294 ms pour CO-B et 201-288 ms pour CF-B). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) de basse fréquence sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-B et CO-B (en bas) ainsi que la carte de différence (CF-B minus CO-B, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Le segment 6 présente une négativité dont le maximum est enregistré sur les régions antérieures ("caudalité";  $F(2,38) = 12.88$ ;  $p < .0001$ ), à gauche et au centre ("latéralité";  $F(2,38) = 5.39$ ;  $p < .008$ ). L'analyse de l'interaction "classe de mots" x "caudalité" ( $F(2,38) = 6.16$ ;  $p < .004$ ) montre que les PEs associés aux mots CF-B sont plus négatifs ( $-1.96 \mu\text{V}$ ) en région antérieure que ceux des mots CO-B ( $-1.21 \mu\text{V}$ ;  $p < .01$ ). L'analyse de la triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 3.81$ ;  $p < .007$ , voir figure 69) indique que les PEs des mots CF-B sont plus négatifs sur les 3 régions antérieures, et plus positifs en régions postéro-médiane centrale et postérieure droite par rapport à ceux des mots CO-B.



**Figure 69:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 6 (compris entre 295-366 ms pour CO-B et 289-390 ms pour CF-B). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) de basse fréquence sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-B et CO-B (en bas) ainsi que la carte de différence (CF-B minus CO-B, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Le segment 7 présente une négativité dont le maximum est enregistré sur les régions antérieures ("caudalité";  $F(2,38) = 11.31$ ;  $p < .0001$ ). L'analyse de l'interaction "classe de mots" x "latéralité" ( $F(2,38) = 6.41$ ;  $p < .004$ ) montre que les PEs des mots CF-B sont plus négatifs ( $-1.19 \mu\text{V}$ ) à gauche par rapport à ceux des mots CO-B ( $-0.50 \mu\text{V}$ ;  $p < .006$ ). L'analyse de la triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 5.62$ ;  $p < .0005$ , voir figure 70) indique que les PEs des mots CF-B sont plus négatifs sur les 3 régions situées à gauche ainsi qu'en antérieure centrale, et plus positifs en régions postéro-médiane centrale et postérieure droite par rapport à ceux des mots CO-B.

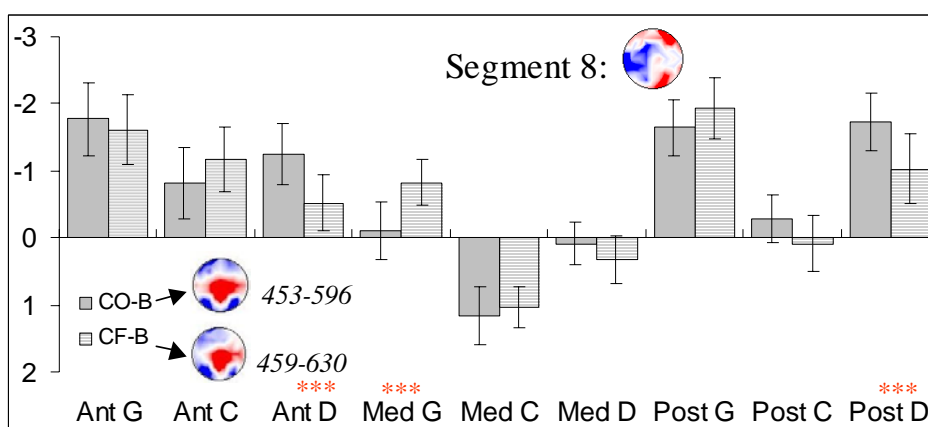


**Figure 70:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 7 (compris entre 367-452 ms pour CO-B et 391-458 ms pour CF-B). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) de basse fréquence sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche;



ANT C, antérieure centrale; ANT D, antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-B et CO-B (en bas) ainsi que la carte de différence (CF-B minus CO-B, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Le segment 8 présente une négativité dont le maximum est enregistré sur les régions antérieure et postérieure ("caudalité";  $F(2,38) = 5.95$ ;  $p < .005$ ) et latéralisée à gauche ("caudalité";  $F(2,38) = 16.96$ ;  $p < .0001$ ). L'analyse de la triple interaction "classe de mots" x "caudalité" x "latéralité" ( $F(4,76) = 3.40$ ;  $p < .01$ , voir figure 71) indique que les PEs des mots CO-B sont plus négatifs sur les régions antérieure et postérieure droites par rapport à ceux associés aux mots CF-B. En région médiane gauche l'effet inverse est observé (les PEs des mots CF-B sont plus négatifs que ceux associés aux mots CO-B).



**Figure 71:** Résultats du test *post-hoc* LSD des comparaisons planifiées pour le segment 8 (compris entre 453-596 ms pour CO-B et 459-630 ms pour CF-B). Les valeurs du PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF-B) et ouverte (CO-B) de basse fréquence sont indiquées pour les 9 régions d'intérêt: ANT G, antérieure gauche; ANT C, antérieure centrale; ANT D antérieure droite; MED G, médiane gauche; MED C, médiane centrale; MED D, médiane droite; POST G, postérieure gauche; POST C, postérieure centrale; POST D, postérieure droite. Les barres verticales indiquent l'erreur type. Les cartes de surface sont illustrées pour CF-B et CO-B (en bas) ainsi que la carte de différence (CF-B minus CO-B, en haut). \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.3.2.4 Evolution temporelle des micro-états

Les analyses portant sur la latence des pics de GFP ne montrent pas de différences significatives quant à l'apparition des segments 5, 6, 7, et 8 (voir tableau 33).

Segment	CF-B	CO-B	ANOVA
5	256.5 (4.22)	257.9 (5.32)	$F(1,19) = .35; p > .55$
6	326.1 (4.29)	330.6 (4.53)	$F(1,19) = 1.23; p > .28$
7	425.0 (4.64)	431.4 (5.35)	$F(1,19) = 1.71; p > .20$
8	540.9 (7.29)	542.1 (8.71)	$F(1,19) = .02; p > .88$

**Tableau 33:** Résultats des ANOVAs à mesures répétées séparées sur la latence du maximum de GFP pour les 4 segments (5, 6, 7 et 8). Les valeurs indiquées sont en ms avec, entre parenthèses, l'erreur type. Les effets significatifs sont rapportés en gras.

### 3.3.3 Synthèse CF vs. CO basse fréquence

Les résultats de la segmentation adaptée suggèrent que 4 segments (5, 6, 7, et 8) présentent des différences topographiques significatives sur le scalp, entre les deux classes de mots. Alors que pour le segment 5 (CO-B: 199-294 ms, CF-B: 201-288 ms), la négativité associée aux mots CF-B est plus importante sur les régions antéro-médiane gauche et antérieure centrale par rapport aux mots CO-B; pour le segment 6 (CO-B: 295-366 ms, CF-B: 289-390 ms), les différences sont observées sur toutes les régions antérieures. Dans la fenêtre temporelle correspondant au segment 7 (CO-B: 367-452 ms, CF-B: 391-458 ms), la négativité antérieure associée aux mots CF-B est plus importante que celle des mots CO-B, sur les régions situées sur l'hémisphère gauche. Pour le dernier segment 8 (CO-B: 453-596 ms, CF-B: 459-630 ms), des négativités sont enregistrées à la fois sur les régions antérieures et postérieures. Pour ce micro-état, les PEs des mots CO-B sont plus négatifs sur les régions droites antérieure et postérieure par rapport à ceux des mots CF-B. En région médiane gauche, se sont les PEs des mots CF-B qui sont plus négatifs par rapport à ceux associés aux mots CO-B. Aucun délai dans l'apparition des cartes n'est mis en évidence.

## 4 Discussion

### 4.1 Effet de classe en isolation: apport de la segmentation adaptée

Les analyses de tracés montrent que, lorsque les mots sont présentés en isolation, deux fenêtres d'analyses indiquent des différences topographiques entre les deux classes de mots. Entre 250 et 350 ms, les PEs associés aux mots CF sont plus négatifs que ceux des mots CO

sur les régions antérieures, alors que l'effet inverse est observé sur les régions postérieures du scalp (les PEs des mots CF sont plus positifs par rapport à ceux des mots CO).

Entre 350 et 500 ms, comparé aux PEs des mots de classe fermée (CF), ceux des mots de classe ouverte (CO) sont plus négatifs sur les régions antérieures et postérieures, alors que sur les sites en médians, les PEs des mots CF sont plus négatifs.

Ces résultats sont confirmés et affinés par l'analyse de segmentation. En effet, celle-ci indique que 3 segments présentent des différences topographiques entre les deux classes de mots. Pour le segment 6 (enregistré entre 295-386 ms pour CF et entre 297-366 ms pour CO), les PEs des mots CF sont plus négatifs sur les régions antérieures, alors qu'en régions postéro-médiane centrale et postérieure droite, ils sont plus positifs que ceux associés aux mots CO-B. Pour le segment 7 (387-452 ms pour CF et entre 367-452 ms pour CO), la négativité des mots CF est plus importante sur les sites distribués sur l'hémisphère gauche par rapport aux mots CO. Pour le dernier segment 8 (453-628 ms pour CF et entre 453-596 ms pour CO), les PEs des mots CF sont plus négatifs sur les régions localisées en postéro-médianes gauches par rapport à ceux des mots CO. Les PEs des mots CO présentent une négativité plus importante sur les régions postérieures droites que celle suscitée lors de la présentation des mots CF.

Grâce à la segmentation adaptée nous avons défini le commencement des différences topographiques plus précisément (295-297 ms), alors qu'avec les analyses de tracés classiques, les fenêtres fixes suggéraient des effets dès 250 ms. De plus, l'analyse de segmentation adaptée montre que 3 segments présentent des topographies différentes entre les deux classes de mots, alors que l'analyse de tracés ne définit que deux effets différents (entre 250-350 ms et entre 350-500 ms). Grâce à une réduction adéquate et non arbitraire des fenêtres temporelles, l'analyse de segmentation adaptée a permis la précision du début des différences topographiques entre les deux classes de mots. Ainsi, comme l'indique Steinhauer & Ullman (2002), il est parfois utile de réduire la fenêtre d'analyse afin de montrer des effets plus fins dans les PEs. Puisque les propriétés lexicales sont contrôlées, nous ne pouvons assigner ces effets à des différences de fréquence et/ou de longueur entre les deux classes. Nos résultats suggèrent donc, qu'en isolation, le traitement des mots de classe fermée présente des différences par rapport à celui des mots de classe ouverte.

L'analyse de la latence de pic à partir des électrodes frontales a montré qu'une composante négative (distribuée sur les sites antérieurs) est plus précoce (de 7 ms) pour les mots de classe fermée (316 ms) que pour les mots de classe ouverte (323 ms). L'analyse du pic de GFP suggère que deux des segments enregistrés pour les mots de classe fermée sont plus précoces. En effet, le segment 5 indique une différence de 3.9 ms (254.4 ms pour CF, 258.3 ms pour CO) et le segment 6, une différence de 8.8 ms, celui-ci étant plus précoce pour les mots CF (324.3 ms) que pour CO (333.1 ms). Pour les électrodes frontales et dans le rang de latence du segment 5 (compris entre 203-294 ms pour CF et entre 199-296 ms pour CO), nous ne pouvions mettre en évidence de pic négatif avec l'analyse de tracés. En effet, sur ces électrodes et avec une référence mastoïde gauche, tout le PE est positif (figure 50 et 52). Par conséquent, analyser la latence de ce pic ne pouvait être réalisable (tout au moins sur les électrodes situées en régions antérieures). A l'opposé, avec une référence moyenne, la carte représentative du segment 5 paraît bipolaire (figure 54), et un pic de GFP a pu être déterminé (figure 53). A notre connaissance, aucune autre étude ne présente de résultats indiquant un décalage de deux composantes ou segments pour les mots de classe fermée par rapport aux mots de classe ouverte, que ce soit en contexte de phrases, ou en isolation. Ainsi, ces résultats restent à être confirmés, cependant l'usage de la segmentation adaptée (ainsi que l'utilisation d'une référence moyenne) nous semble primordial, afin de révéler la finesse des décours temporels respectifs des deux classes de mots. Nos résultats indiquent la précocité de deux segments, ou micro-états, lors de la lecture en isolation des mots de classe fermée, par rapport aux mots de classe ouverte.

## 4.2 Interprétation fonctionnelle des effets

Nos résultats montrent que, pour le segment 5 (203-294 ms pour CF, et 199-296 ms pour CO), les PEs des deux classes de mots sont équivalents en topographie sur le scalp, mais qu'un décalage de 3.9 ms est enregistré. Ce segment 5 correspond à la topographie du segment 6 du chapitre 5 (199-256 ms pour CF et 201-250 ms pour CO). De plus la fenêtre temporelle de ce segment correspond à une partie de celle de la composante P200 (Curran, *et al.*, 1993; Frishkoff & Tucker, 2001; Nobre & McCarthy, 1994). Généralement, les modulations d'amplitude de la composante P200 ne sont pas directement liées à des étapes purement linguistiques: elles sont considérées comme reflétant l'encodage des caractéristiques du stimulus (traitement perceptif), ou encore la catégorisation des patterns du stimulus en lettres plutôt qu'à un traitement à un niveau lexical (Keurs ter, *et al.*, 2002; Kutas & Hillyard,

1983). Dans notre chapitre précédent, nous avons indiqué que l'effet dans la fenêtre 200-250 ms correspond aussi avec l'enregistrement de la composante ELAN, reflétant l'activation des informations de catégorie syntaxique (Friederici & Mecklinger, 1996; Friederici, *et al.*, 1993). Ainsi, nous avons suggéré, que ce segment reflète l'activation de l'information de la classe du mot lors de la construction de la structure syntaxique de la phrase. En isolation, nos résultats suggèrent que même si aucune représentation n'est construite ce type d'information est malgré tout activé.

Le segment 6 (295-386 ms pour CF et 297-366 ms pour CO) présente un effet robuste de la classe de mots, tant d'un point de vue topographique que de la latence d'apparition de la carte spécifique. Bien que les deux classes de mots suscitent l'occurrence de composantes négatives en régions antérieures (figure 55), les PEs associés aux mots CF sont plus négatifs sur les 3 régions antérieures par rapport à ceux des mots CO. Cette fenêtre d'analyse et les caractéristiques topographiques de ces négativités, sont similaires à la description de la composante N280 (Keurs ter, *et al.*, 2002; Münte, *et al.*, 2001; Nobre & McCarthy, 1994). Des différences topographiques présentes dans cette fenêtre temporelle ont été associées aux processus d'identification des mots de classe ouverte, et de ceux de la classe fermée (Brown, *et al.*, 1999; Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Münte, *et al.*, 2001; Neville, *et al.*, 1992; Osterhout, 1997). Hors contexte, nos résultats suggèrent que cette étape de traitement présente donc des dissemblances entre les deux classes de mots. De plus l'identification des mots de classe fermée est plus rapide de 8.8 ms par rapport à celle des mots de classe ouverte. En présence de contexte (chapitre 5), nous avons montré des différences de latéralisation pour le segment 7 (257-338 ms pour CF et 251-334 ms pour CO) mais aussi des différences de latence d'apparition. Puisque le rang d'enregistrement du segment 7, chapitre 5, semble correspondre au segment 6 de ce chapitre, des différences topographiques pour ces deux segments indiquent la présence d'un processus commun aux deux expériences. Comme aucun contexte linguistique n'est utilisé dans le chapitre présent, nous pouvons donc relier ce segment 6 à un processus d'accès lexical.

Le segment 7 (387-452 ms pour CF et 367-452 ms pour CO) montre que les mots de classe fermée suscitent l'occurrence d'une composante négative plus importante sur les sites antérieurs droits et gauches ainsi que postérieurs gauches par rapport à celle associée aux mots CO. Deux études récentes seulement ont enregistré cet effet tardif hors contexte de phrase (Keurs ter, *et al.*, 2002; Münte, *et al.*, 2001). La distribution sur le scalp, ainsi que le

délai d'apparition correspondent aux caractéristiques de l'effet de classe tardif obtenues dans un contexte de phrase, la LAN (Brown, *et al.*, 1999; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Kutas & Hillyard, 1983; Münte, *et al.*, 2001; Neville, *et al.*, 1993; Neville, *et al.*, 1992; Nobre & McCarthy, 1994; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997; Pulvermüller, *et al.*, 1995). Alors que cette différence de topographie est généralement attribuée à des effets post-lexicaux, il semble concevable que, malgré l'absence de contexte linguistique dans le cas présent, les mots présentés conservent toute leur spécificité et que, par conséquent, ils soient traités de manière automatique comme dans un contexte de phrase (Münte, *et al.*, 2001).

Le segment 8 (453-628 ms pour CF et 453-596 ms pour CO) indique l'occurrence d'une composante négative sur les sites postérieurs droits pour les mots de classe ouverte. Cette composante négative correspond aux caractéristiques de la composante N400 (Kutas & van Petten, 1994; Tucker, 1993). La comparaison des PEs pour les mots lus dans un contexte normal et dans des listes de mots sans lien apparent nous révèle que les mots de classe ouverte suscitent l'occurrence d'une composante N400, d'amplitude plus large dans la condition où les mots n'ont pas de lien apparent. Ce résultat réplique ceux de Osterhout, *et al.* (1997) et de van Petten & Kutas (1991) et suggère que dans un texte cohérent, les mots de classe ouverte suscitent une composante N400 d'amplitude réduite, comparée à lorsque ceux-ci sont lus en isolation. Ces résultats corroborent l'interprétation fonctionnelle de la composante N400: plus un mot est facilement intégré d'un point de vue sémantique, et plus la composante N400 est réduite (Kutas & Hillyard, 1980).

### 4.3 Effet de classe en fonction du rang de fréquence

Nos résultats indiquent, qu'en fonction du rang de fréquence étudié, l'effet de classe de mots (distinction entre mots de classe ouverte et fermée) n'est pas similaire sur le scalp.

En effet, nous avons montré que pour des mots de haute fréquence, l'effet de classe est enregistré dans la fenêtre temporelle correspondant à la composante N400 (350-500). L'analyse de segmentation adaptée montre des différences topographiques pour le segment 8 (qui correspond aux fenêtres 449-628 ms pour CF-H et 451-596 ms pour CO-H). De plus, pour les mots de haute fréquence, deux segments sont plus précoces pour les mots de classe fermée (CF-H) que pour les mots de classe ouverte (segment 5, différence de 5.5 ms; segment 6 différence de 9.4 ms). Pour des mots de basses fréquences, l'effet de classe présente des

topographies différentes de celles observées pour l'effet de classe de haute fréquence. Alors que l'analyse de tracés indique des différences topographiques dans les 4 fenêtres temporelles testées (P2, N280, N400 et LAN), la segmentation adaptée montre que 4 segments (5, 6, 7 et 8) présentent des topographies différentes entre les deux classes de mots. Enfin, aucun décalage n'a été mis en évidence, que ce soit au niveau des analyses spécifiques des électrodes situées en région antérieure (analyse de tracés), ou bien grâce au pic de maximum de GFP (analyse de segmentation adaptée).

Lorsque les deux rangs de fréquence sont moyennés, les différences topographiques observées sont donc principalement dues à des différences effectives entre les deux classes de mots, pour le rang de fréquence situé dans la partie inférieure, c'est à dire pour des fréquences inférieures à 400 occurrences par million ( $< 2,6 \log$  de fréquence). Pour des fréquences inférieures à  $2,6 \log$  de fréquence, l'effet de classe se résume donc à 4 micro-états différents sur le scalp. En revanche, les effets de latence sont principalement liés à des différences de traitement obtenus sur le rang de fréquence supérieur à 400 occurrences par million.

## 5 Conclusions

Notre recherche indique qu'en isolation, les mots de classe fermée de basse fréquence ( $< 400$  occurrences/million) sont analysés plus précocement (dès 200 ms) par rapport aux mots de classe ouverte de fréquence comparable. Pour les mots de haute fréquence, les différences topographiques enregistrées entre les deux classes de mots sont plus tardives, indiquant que des processus identiques sont mis en place lors de l'identification de ces deux classes. En revanche, deux micro-états sont plus précoces pour les mots de classe fermée. Ce résultat suggère des processus plus rapides pour les mots de classe fermée par rapport aux mots de classe ouverte de haute fréquence.

# CHAPITRE 7

## LE CONTEXTE AFFECTE-T-IL IDENTIQUEMENT LES DEUX CLASSES DE MOTS?

---

### *Résumé*

Ce chapitre explore l'influence du contexte linguistique sur la reconnaissance des deux classes de mots. Pour ce faire, nous comparons directement les résultats obtenus en contexte de phrases (chapitre 5) et en isolation (chapitre 6). Un test de complétion ou de *Cloze Procedure* permet de rendre compte des contraintes apportées par le contexte avant la présentation des mots cibles. Dans un deuxième temps, une comparaison statistique des PEs obtenus à partir de la présentation des mots en contexte de phrase (chapitre 5) et en isolation (chapitre 6) est réalisée. Les résultats indiquent que, d'une part, le contexte linguistique impose plus de contraintes sur l'identité et sur la classe de mots lorsque ceux-ci sont de classe ouverte. Les analyses électroencéphalographiques suggèrent que la présence d'un contexte linguistique module spécifiquement les PEs des deux classes de mots dans une fenêtre précoce (0-150 ms) et une plus tardive (250-350 ms). Le premier effet peut être lié soit à des différences de processus attentionnels, soit à des différences dans des processus purement psycholinguistiques. Le deuxième effet a été attribué aux attentes spécifiques créées par le contexte linguistique sur les informations de la classe de mot.



## 1 Introduction

Ce chapitre vise à étudier l'impact différencié du contexte sur les deux classes de mots: la présence d'un contexte linguistique influence-t-elle différemment les mots de classe ouverte des mots de classe fermée? Si tel est le cas, il s'agira alors de définir à quelle(s) étape(s) de la reconnaissance des mots, cet effet différencié du contexte intervient. Pour répondre à ces questions, nous avons réalisé deux analyses.

D'une part, les contraintes développées par le contexte linguistique (chapitre 5) avant la présentation du mot cible sont estimées. Le terme contrainte est utilisé par Schwanenflugel & Shoben (1985) afin d'expliquer la facilitation qu'un contexte exerce sur la reconnaissance des mots. Selon ces auteurs, certaines phrases induisent davantage de contraintes que d'autres. En conséquence, les attentes pour certains mots sont plus ou moins importantes. Les contraintes qu'un contexte impose à l'apparition d'un mot sont généralement mesurées grâce au test de complétion ou de *Cloze Procedure*. Ce test indique la probabilité que les participants complètent un fragment de phrase par un mot particulier (Fischler & Bloom, 1979; Schwanenflugel & Shoben, 1985). De leurs résultats obtenus à partir de nombreuses tâches (Schwanenflugel & Shoben, 1985) concluent que, plus une phrase impose de contraintes, et plus les effets de contexte seront importants. Etant donné que les mots de classe fermée constituent le cadre syntaxique de la phrase et n'assurent aucune fonction référentielle, et que à l'opposé, les mots de classe ouverte sont liés à des référents extralinguistiques, nous avançons l'hypothèse que le contexte linguistique impose des contraintes différentes lors de la reconnaissance de ces deux classes de mots. Si tel est le cas, alors nous devrions trouver des mesures de complétion différentes pour les deux classes de mots.

D'autre part, nous allons étudier les modulations électroencéphalographiques consécutives à la présentation du contexte linguistique. Pour ce faire, nous allons comparer les résultats obtenus lors d'une présentation en isolation (chapitre 6) aux résultats obtenus lors d'une présentation en contexte de phrase (chapitre 5). Seules deux autres études (Osterhout, *et al.*, 1997; van Petten & Kutas, 1991) ont réalisé directement des analyses statistiques entre mots présentés en isolation et mots présentés dans un contexte de phrase. Van Petten & Kutas (1991) comparent l'effet de différents types de contexte, sur les deux classes de mots. Les informations du contexte sont manipulées en présentant des phrases normales (informations sémantiques et syntaxiques conservées, contexte sémantique), des phrases dans lesquelles les

mots de classe ouverte sont remplacés par des mots de la même classe (informations syntaxiques conservées, contexte syntaxique), et des phrases dans lesquelles les mots étaient mélangés (aucune information conservée, liste de mots). Ces auteurs indiquent que le contexte affecte les PEs en diminuant la négativité de la composante N400 (entre 300 et 500 ms) comparée à celle observée en liste de mots. De plus, la présence d'un contexte sémantique affecte les PEs des deux classes de mots alors que, pour le contexte syntaxique, seuls les PEs des mots CF présentent des variations d'amplitudes. Pour finir, ces auteurs montrent des différences topographiques dans la répartition des effets du contexte: alors que les PEs des mots CO sont affectés sur les régions situées sur l'hémisphère droit, pour les PEs des mots CF, les effets divergent sur les régions de l'hémisphère gauche. Entre 500 et 700 ms, le contexte affecte les PEs des mots CF en augmentant la négativité sur les régions frontales bilatérales, par rapport à une présentation en isolation. Osterhout, *et al.* (1997) indiquent des résultats comparables: le contexte augmente la négativité des mots CF entre 500 et 700 ms, alors que pour les mots CO le contexte diminue la négativité sur les régions postérieures entre 350 et 450 ms.

Les résultats de ces deux études semblent indiquer des résultats en faveur de l'hypothèse selon laquelle le contexte affecte différemment le traitement des deux classes de mots. Malheureusement, aucun contrôle de la fréquence et/ou de la longueur des stimuli n'est effectué. Comme notre matériel est contrôlé en fréquence et en longueur, nous allons pouvoir apporter des précisions par rapport aux études de Osterhout, *et al.* (1997) et de van Petten & Kutas (1991). Ainsi, l'enregistrement de différences sur des composantes précoces (telles que la composante P200), supposées refléter des étapes de catégorisation de stimuli (Hillyard, *et al.*, 1998), plaiderait fortement en faveur des théories interactives qui suggèrent que des informations de différentes natures (sémantiques, syntaxiques, contextuelles) sont utilisées dès qu'elles sont disponibles afin de construire une représentation de la phrase (Morton, 1969; Morton, 1979). A l'opposé, l'observation de différences tardives, par exemple au niveau de la composante N400 ou plus tardivement, suggérerait que l'impact différencié du contexte intervient après le processus de reconnaissance lexicale lui-même, c'est à dire, au niveau des processus post-lexicaux impliqués dans le traitement de ces deux classes de mots.

## 2 Test de complétion (*Cloze Procedure*, CP)

### 2.1 Méthode

Un test de complétion (CP) a été réalisé sur les phrases utilisées dans le chapitre 5 (mots critiques en 4<sup>ème</sup> position). Vingt participants<sup>20</sup> ont lu les 160 fragments de phrases constitués des trois premiers mots de chacune des phrases testées. Il leur était demandé de noter par écrit le premier mot qui leur venait à l'esprit pour compléter la phrase. La consigne spécifiait clairement que plusieurs mots pouvaient être ajoutés et, par conséquent, que les phrases pouvaient être constituées de plus de 4 mots. Dans ce dernier cas, les juges devaient rapporter l'ensemble des mots qui leur venaient à l'esprit. Les réponses ont été analysées en fonction de 2 facteurs: (1) l'identité (le mot utilisé par les participants est identique ou non au mot présenté dans notre expérience), (2) la classe de mot ouverte ou fermée (le mot utilisé par les participants appartient-il ou non à la même classe que le mot présenté dans notre expérience). Des ANOVAs ont été réalisées séparément pour chaque critère (identité, classe de mots) avec comme facteurs "classe de mots" (CO vs. CF) et "fréquence" (Haute vs. Basse).

### 2.2 Résultats

Les résultats du test de complétion sont rapportés dans le tableau 34. Globalement, le matériel expérimental utilisé dans le chapitre 5 était très peu contraignant pour les mots critiques en 4<sup>ème</sup> position. En effet, une faible proportion (comprise entre 2.8 % et 7 %) de phrases présentait le mot proposé par les participants. En revanche, la classe des mots est plus justement proposée par nos juges (comprise entre 77.6 % et 89.6 %). Alors que notre matériel ne semble pas apporter d'information (ou très peu) sur l'identité du mot, la classe de mots a pu être donnée avant la présentation du mot critique et ce, grâce à la lecture du contexte linguistique.

---

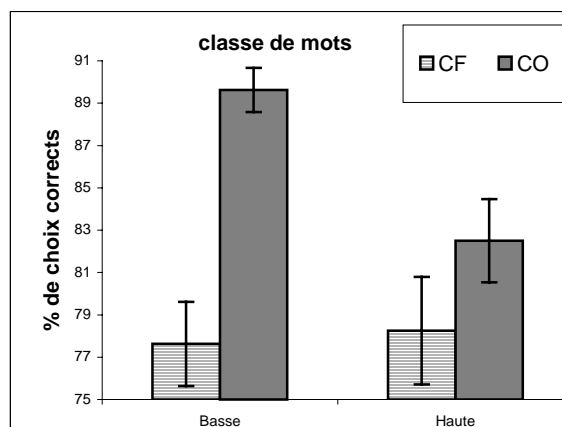
<sup>20</sup> Les participants n'ont pas contribué à l'expérimentation PEs du chapitre 5.

	CF-B	CF-H	CO-B	CO-H
identité	2.88	5.63	5.88	7.00
classe	77.60	78.30	89.60	82.50

**Tableau 34:** Résultats du test de complétion (N=20). Les chiffres rapportés sont des pourcentages de réponses correctes pour la prédictibilité. L'identité correspond aux essais où le participant a donné le mot exact, la classe correspond aux essais où le participant a donné un mot de la même classe que le mot présenté dans l'étude du chapitre 5. CF-B: mots de classe fermée de basse fréquence, CF-H: mots de classe fermée de haute fréquence, CO-B: mots de classe ouverte de basse fréquence, CO-H: mots de classe ouverte de haute fréquence.

Pour la variable dépendante identité, les résultats indiquent que les effets du facteur "classe de mots" ( $F(1,19) = 13.26$ ,  $p < .002$ ) et du facteur "fréquence" sont significatifs ( $F(1,19) = 6.86$ ,  $p < .01$ ) sans interaction "classe de mots" x "fréquence" ( $F < 1$ ). Les participants réussissent mieux à prédire les mots de classe ouverte (6.44 %) que ceux de la classe fermée (4.25 %). Par ailleurs, les mots de haute fréquence donnent lieu à un plus grand nombre de choix identiques à la cible (8.76%), que les mots de basse fréquence (6.33%, voir tableau 34).

Pour la variable dépendante classe de mots, les résultats montrent que les effets du facteur "classe de mots" ( $F(1,19) = 7.48$ ,  $p < .01$ ) et du facteur "fréquence" sont significatifs ( $F(1,19) = 11.13$ ,  $p < .0003$ ). L'interaction "classe de mots" x "fréquence" est significative ( $F(1,19) = 6.15$ ,  $p < .02$ ). Les participants ont plus souvent tendance à choisir un mot de classe ouverte, alors que le mot cible était effectivement un mot de classe ouverte (86.06 %) qu'ils n'ont tendance à choisir un mot de classe fermée, alors que la cible était un mot de classe fermée (77.94 %). L'effet de la fréquence n'est observé que sur les mots de classe ouverte (voir figure 72): tandis que la fréquence influence le choix des mots lorsque ceux-ci sont de classe ouverte (CO-B: 89.60%, CO-H: 82.50%;  $p < .004$ ), la fréquence n'influence pas le choix des mots lorsque ceux-ci sont de classe fermée (CF-H: 78.30 %, CF-B: 77.60%;  $p > .7$ ; voir tableau 34). Cette interaction témoigne du fait que le contexte contraint plus fortement l'apparition du mot cible, lorsque ce dernier est de la classe ouverte et, en particulier, lorsqu'il est de basse fréquence: le mot cible est plus prédictible lorsqu'il fait partie de la classe ouverte et lorsqu'il est de basse fréquence.



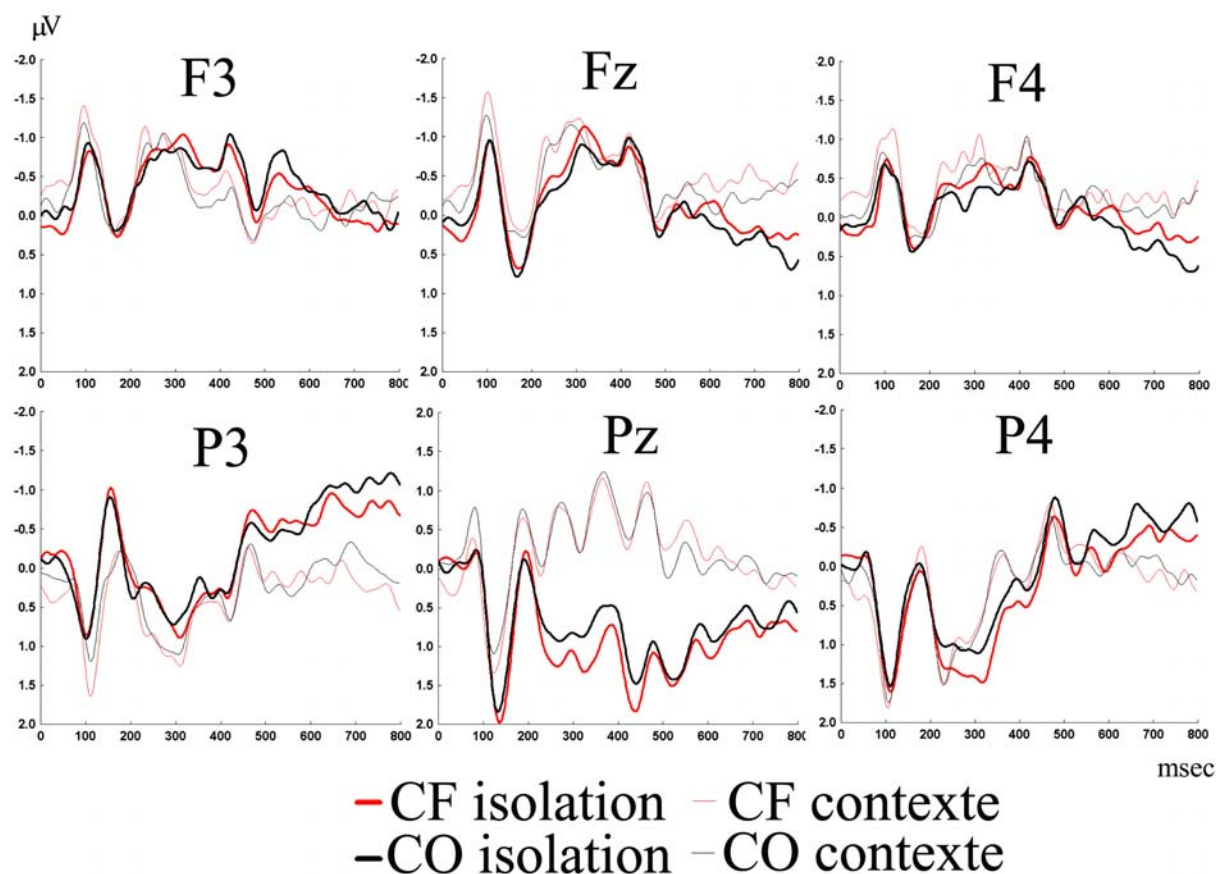
**Figure 72:** Illustration de l'interaction "classe de mots" x "fréquence" sur la variable dépendante classe de mots. En ordonnée est reporté le pourcentage de choix corrects pour la classe de mots.

### 3 Données électroencéphalographiques

#### 3.1 Analyses

Afin d'éviter les problèmes liés à une référence non stable, telle que celle située en mastoïde gauche (voir chapitre 2), les analyses statistiques ont été réalisées sur les données en référence moyenne. De plus, comme la segmentation adaptée a montré des variations dans les frontières des segments obtenus pour ces deux expériences (voir chapitre 5, figure 37 et chapitre 6, figure 54), il nous semblait difficile de choisir les fenêtres temporelles adéquates pour les analyses statistiques. Par conséquent, les fenêtres temporelles utilisées sont celles employées dans les analyses de tracés classiques.

Dans la figure 73, les PEs des mots de classe fermée et ouverte sont rapportés pour les deux types de présentations (isolé et en contexte), et pour 6 électrodes (3 frontales, F3, Fz et F4; 3 postérieures, P3, Pz et P4). Sur ces tracés, on note que, dès la présentation des stimuli, les PEs des mots présentés en isolation paraissent moins négatifs que ceux enregistrés dans un contexte de phrase: cet effet semble absent sur l'électrode P3 (postérieur gauche), et plus tardif pour l'électrode F3 (antérieure gauche). A l'opposé, les PEs des mots présentés dans un contexte de phrase montrent une négativité accrue sur l'électrode Pz (postérieure centrale) entre 200 et 800 ms.



**Figure 73:** Grande moyenne (N=20) des PEs en référence moyenne pour 6 électrodes: F3, Fz, et F4 situées en région antérieure et P3, Pz, P4 situées en région postérieure. Les tracés en gras correspondent aux PEs des mots enregistrés en isolation (chapitre 6), alors que les tracés fins sont associés à la présentation des mots dans un contexte de phrases (chapitre 5). Les PEs rouges sont associés à la présentation des mots de classe fermée, et ceux en noir aux mots de classe ouverte. La négativité est représentée vers le haut.

Des ANOVAs à mesures répétées ont été réalisées sur la variable dépendante amplitude des PEs avec, comme facteur inter-sujets le "type de présentation" (contexte vs. isolé) et comme facteurs intra-sujets, la "classe de mots" (CO vs. CF), la "caudalité" (antérieur vs. médian vs. postérieur) et la "latéralité" (gauche vs. central vs. droite) pour 5 fenêtres temporelles différentes (0-150, 150-250, 250-350, 350-500, 500-800 ms). Afin de ne pas surcharger nos résultats de détails, seuls les résultats relatant les effets globaux du contexte, impliquant le facteur principal "type de présentation", et ceux spécifiques à chaque classe de mots, impliquant des interactions entre "type de présentation" x "classe de mots" avec les autres facteurs intra-sujets ("caudalité" et/ou "latéralité") sont discutés. Les interactions sont analysées par de simples ANOVAs. De plus, afin de définir si les effets de contexte observés proviennent des mots de haute et/ou de basse fréquence, nous avons calculé les ANOVAs à

mesures répétées pour les deux fréquences séparément, pour les rangs de latence où un effet significatif a été trouvé.

## 3.2 Résultats

### 3.2.1 CF vs. CO

Les résultats des ANOVAs à mesures répétées avec le facteur principal inter-sujet "type de présentation" (résultats du chapitre 5 vs. résultats du chapitre 6) ainsi que les interactions avec les facteurs intra-sujets ("classe de mots", "caudalité", "latéralité") sont rapportés dans le tableau 35. L'effet du facteur principal "type de présentation" (contexte vs. isolé) est significatif dans les 5 fenêtres d'analyses choisies (de 0 à 800 ms). Toutefois, ce facteur interagit avec les facteurs "caudalité" et "latéralité": les interactions "type de présentation" x "caudalité" et "type de présentation" x "latéralité" sont significatives de 0 à 800 ms. Ces interactions indiquent que les PEs des mots présentés en isolation sont plus négatifs que les PEs des mots présentés en contexte sur les régions postérieures, ainsi qu'à gauche. De 150 à 800 ms, l'interaction "type de présentation" x "caudalité" x "latéralité" est aussi significative: les PEs des mots présentés en isolation sont plus négatifs sur les régions postérieures gauche et droite, alors que ceux des mots présentés dans un contexte, sont plus négatifs en région centrale médiane.

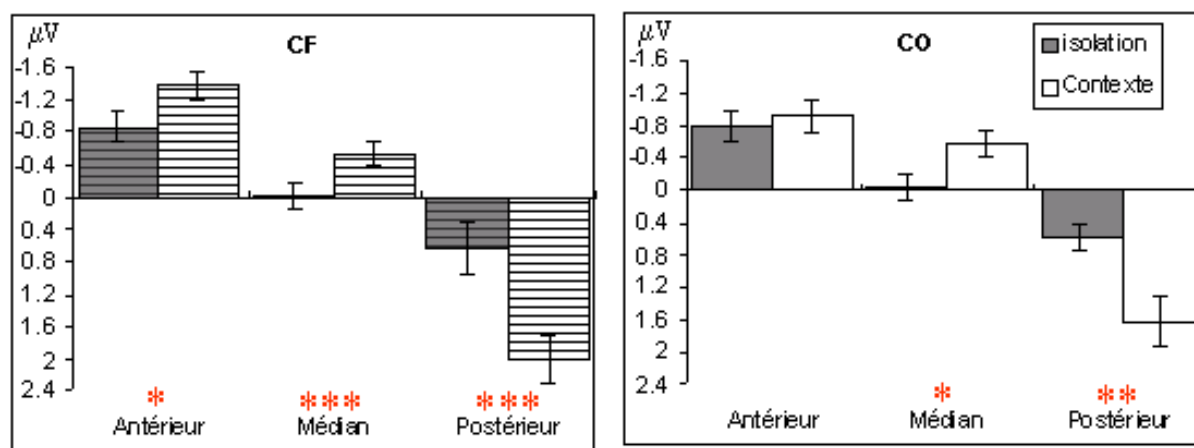
	ddl	0 - 150	150 - 250	250 - 350	350 - 500	500 - 800
présentation	1, 38	<b>35.77***</b> - /	<b>28.09***</b> - /	<b>29.89***</b> - /	<b>12.44***</b> - /	<b>7.22**</b> - /
présentation x classe	1, 38	0.97 - 2.36	1.71 - 0.14	0.36 - 0.78	2.71 - 0.67	0.27 - 0.47
présentation x classe x caudalité	2, 76	<b>2.79*</b> - <b>4.68*</b>	0.06 - 0.02	1.35 - 0.94	0.03 - 0.09	0.02 - 0.23
présentation x classe x latéralité	2, 76	<b>5.49***</b> - <b>3.74*</b>	7.42** - 2.70	7.19** - 2.61	3.23* - 0.72	1.23 - 0.35
présentation x classe x caudalité x latéralité	4, 152	0.68 - 0.34	2.10 - 1.17	<b>2.79*</b> - <b>2.89*</b>	2.52 - 2.33	2.12 - 2.15

**Tableau 35:** Résultats des ANOVAs à mesures répétées séparées pour chaque fenêtre temporelle (0-150, 150-250, 250-350, 350-500 et 500-800 ms) avec le facteur inter-sujet "type de présentation" (contexte vs. isolé) et pour les facteurs intra-sujets "classe de mots" (CO vs. CF), "caudalité" (antérieur vs. médian vs. postérieur) et "latéralité" (gauche vs. central vs. droite). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et pour les données normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Les interactions significatives entre les facteurs "type de présentation" x "classe de mots" et "caudalité" et/ou "latéralité" suggèrent que le contexte affecte différemment les PEs des deux

classes de mots, sur deux fenêtres d'analyse: entre 0-150 et entre 250 –350 ms (voir tableau 35).

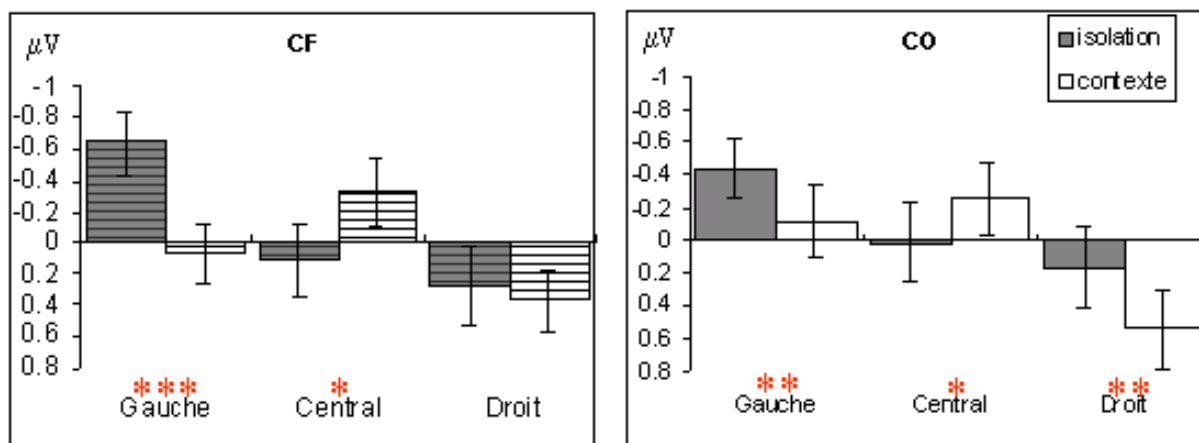
Entre 0 et 150 ms, l'analyse de l'interaction "type de présentation" x "classe de mots" x "caudalité" ( $F(2,76) = 4.68; p < .01$ ) indique, qu'au niveau des régions antérieures, le contexte affecte l'amplitude (augmentation) de la négativité pour les mots CF, tandis qu'aucun effet significatif du contexte n'est observé pour les mots CO (voir figure 74). En revanche, en régions médiane et postérieure, le contexte modifie les PEs de façon similaire pour les deux classes de mots: augmentation de la négativité en médian et augmentation de la positivité en postérieur.



**Figure 74:** Résultats des ANOVAs simples pour chaque région d'intérêt illustrant l'interaction "type de présentation" x "classe de mots" x "caudalité" pour la fenêtre 0-150 ms. Les valeurs des PEs (en µV) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées; les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

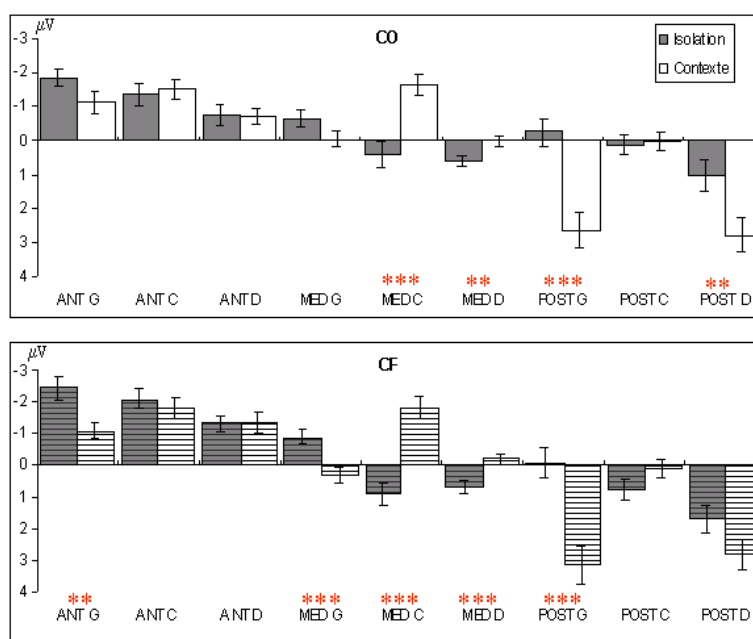
Comme le montre l'interaction "type de présentation" x "classe de mots" x "latéralité" ( $F(2,76) = 3.74; p < .02$ ), entre 0 et 150 ms, l'effet du type de présentation en fonction de la classe de mots est également modulé par la latéralité des électrodes. Le contexte augmente l'amplitude de la positivité pour les mots CO sur les régions situées à droite, alors qu'aucun effet n'est observé pour les mots CF (voir figure 75). Sur les régions à gauche et en centrale, le contexte affecte similairement les PEs des deux classes de mots par une diminution de la négativité sur les régions situées à gauche et une augmentation de la négativité en région centrale.





**Figure 75:** Résultats des ANOVAs simples pour chaque région d'intérêt illustrant l'interaction "type de présentation" x "classe de mots" x "latéralité" pour la fenêtre 0-150 ms. Les valeurs des PE (en µV) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées; les barres verticales indiquent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

Entre 250 et 350 ms, l'interaction "type de présentation" x "classe de mots" x "caudalité" x "latéralité" est significative ( $F(4,152) = 2.89$ ;  $p < .02$ ): lors de la présentation des mots CF dans un contexte de phrase, la négativité enregistrée en régions antéro-médianes gauches est réduite, comparée à la condition de présentation isolée de ces mêmes mots (voir figure 76). Cet effet est absent pour les mots CO. En région postérieure droite, le contexte affecte spécifiquement les PE des mots CO: une augmentation de la positivité est observée, comparée à la présentation en isolation.



**Figure 76:** Résultats des ANOVAs simples pour chaque région d'intérêt illustrant l'interaction "type de présentation" x "classe de mots" x "caudalité" x "latéralité" pour la fenêtre d'analyse comprise entre 250-350 ms.

Les valeurs des PEs (en  $\mu\text{V}$ ) pour les mots de classe fermée (CF) et ouverte (CO) sont indiquées pour les 9 régions d'intérêt. Les barres verticales représentent l'erreur type. \*\*\*,  $p < .001$ ; \*\*,  $p < .01$ ; \*,  $p < .05$ .

### 3.2.2 CF vs. CO haute fréquence

Les résultats des ANOVAs à mesures répétées réalisées pour les mots de haute fréquence, avec le facteur principal inter-sujet "type de présentation" (comparaison des résultats de l'expérience du chapitre 5 vs. chapitre 6) sont rapportés dans le tableau 36, pour les fenêtres temporelles 0-150 ms et 250-350 ms.

	ddl	0 - 150	250 - 350
présentation	1, 38	<b>37.95***</b> - /	<b>28.89***</b> - /
présentation x classe	1, 38	0.91 - <i>0.04</i>	2.40 - <i>1.24</i>
présentation x classe x caudalité	2, 76	2.02 - <i>0.42</i>	0.20 - <i>0.20</i>
présentation x classe x latéralité	2, 76	<b>3.91*</b> - <b>5.21***</b>	1.82 - <i>2.01</i>
présentation x classe x caudalité x latéralité	4, 152	0.49 - <i>0.98</i>	0.20 - <i>0.51</i>

**Tableau 36:** Résultats des ANOVAs à mesures répétées séparées pour les fenêtres temporelles (0-150, et 250-350 ms) avec le facteur inter-sujet "type de présentation" (contexte vs. isolé) et pour les facteurs intra-sujets "classe de mots de haute fréquence" (CO-H vs. CF-H), "caudalité" (antérieur vs. médian vs. postérieur) et "latéralité" (gauche vs. central vs. droite). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et pour les données normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Seule l'interaction entre les facteurs "type de présentation" x "classe de mots" x "latéralité" est significative pour la fenêtre 0-150 ms ( $F(2,76) = 5.21$ ;  $p < .007$ ). Alors qu'aucun effet n'est observé pour les mots CF-H, le contexte augmente l'amplitude de la positivité sur les électrodes situées sur l'hémisphère droit pour les mots CO-H.

### 3.2.3 CF vs. CO basse fréquence

Pour les fenêtres temporelles 0-150 et 250-350 ms, les résultats des ANOVAs à mesures répétées avec le facteur principal inter-sujet "type de présentation" sont rapportés dans le tableau 37 pour les mots de basse fréquence.

	ddl	0 - 150	250 - 350
présentation	1, 38	<b>22.51***</b> - /	<b>26.77***</b> - /
présentation x classe	1, 38	0.01 - <i>1.28</i>	0.13 - <i>0.94</i>
présentation x classe x caudalité	2, 76	<b>0.90 - 5.57*</b>	2.71 - <i>2.45</i>
présentation x classe x latéralité	2, 76	1.67 - <i>0.41</i>	5.28** - <i>1.11</i>
présentation x classe x caudalité x latéralité	4, 152	1.68 - <i>1.13</i>	<b>5.51** - 3.64**</b>

**Tableau 37:** Résultats des ANOVAs à mesures répétées séparées pour les fenêtres temporelles (0-150, et 250-350 ms) avec le facteur inter-sujet "type de présentation" (contexte *vs.* isolé) et pour les facteurs intra-sujets "classe de mots de basse fréquence" (CO-B *vs.* CF-B), "caudalité" (antérieur *vs.* médian *vs.* postérieur) et "latéralité" (gauche *vs.* central *vs.* droite). Les valeurs de  $F$  sont indiquées pour les données brutes (nombres normaux) et pour les données normalisées (nombres en italique) après correction de Greenhouse-Geisser. Les effets significatifs sont rapportés en gras, avec \*,  $p < .05$ ; \*\*,  $p < .01$ ; \*\*\*,  $p < .001$ .

Une première interaction "type de présentation" x "classe de mots" x "caudalité" est significative dans la fenêtre précoce comprise entre 0-150 ms ( $F(2,76) = 5.57$ ;  $p < .005$ ): le contexte affecte les PEs de toutes les régions à l'exception des régions antérieures pour les mots CO-B.

Dans la fenêtre comprise entre 250 et 350 ms, l'interaction "type de présentation" x "classe de mots" x "caudalité" x "latéralité" est significative ( $F(4,152) = 3.64$ ;  $p < .007$ ): comme pour les résultats CO *vs.* CF (voir figure 76), le contexte influence spécifiquement les PEs des mots CF-B sur les régions antéro-médianes gauches (réduction de négativité), alors qu'en régions postérieures droites, seuls les mots CO-B subissent cette influence (augmentation de positivité).

## 4 Discussion

### 4.1 Test de complétion

Nos résultats de complétion indiquent que l'identité des mots de classe ouverte est plus facilement prédite que celle des mots de classe fermée. Les participants choisissent plus souvent le mot exact pour les hautes fréquences que pour les basses fréquences. Les résultats relatifs à la classe de mot sont différents de ceux obtenus sur l'identité. Premièrement, le contexte linguistique est plus contraignant sur la classe lorsque des mots de classe ouverte sont présentés par rapport à ceux de classe fermée. Deuxièmement, aucun effet de fréquence

n'est observé pour la classe fermée alors que, pour les mots de classe ouverte, la classe des mots de basse fréquence est mieux prévisible par rapport aux mots de haute fréquence.

De manière générale, les participants choisissent plus souvent un mot identique, et/ou de même classe que le mot cible utilisé dans notre matériel lorsque ce dernier appartient à la classe ouverte comparée à la classe fermée. Plusieurs études rapportent des résultats contradictoires avec notre test de complétion, indiquant des taux de prédictions plus élevés pour les mots de classe fermée que pour ceux de classe ouverte (Aborn, Rubenstein, & Sterling, 1959; Friederici, 1985; Gough, 1983). Seule l'étude de Swinney, Zurif, & Cutler (1980) s'accorde avec nos résultats (les taux du test de complétion sont rapportés dans l'étude de Friederici, 1985).

Plusieurs explications peuvent être émises quant aux différences entre nos résultats et ceux de la littérature précédente. Si l'on observe les différents items utilisés dans nos stimuli, on peut noter que plus de 61 % des mots de classe ouverte sont des noms (49 items sur 80, chapitre 5 tableau 13). De plus, si l'on détaille un peu plus les phrases utilisées et rapportées dans l'annexe 4, on s'aperçoit que 46 des 49 items précédemment cités sont précédés, soit par un article, soit par un déterminant, soit par un adjectif possessif. Les contraintes liées à la lecture de tels mots imposent forcément qu'un nom ou encore un adjectif va suivre. Ainsi plus de 57 % des items de la classe ouverte entraînent dans ce critère pouvant expliquer pourquoi l'identité mais aussi la classe pour les mots de classe ouverte, sont mieux prédites que celle des mots de classe fermée dans notre test de complétion. La position à laquelle le mot critique est présenté peut aussi affecter la mesure du test de complétion. Alors que, par exemple, Friederici (1985) présentent les mots critiques entre la 8<sup>ème</sup> et 14<sup>ème</sup> position dans la phrase, Swinney, *et al.* (1980) utilisent des positions (entre la 5<sup>ème</sup> et la 6<sup>ème</sup> position) très proches de celles utilisées dans notre étude (4<sup>ème</sup> position). Comme l'indique Aborn, *et al.* (1959), la longueur de l'amorce affecte le taux de prédictibilité: plus celle-ci est longue (avec un niveau de saturation indiqué à 12 mots par ces auteurs), et plus la prédictibilité des mots est importante.

Par conséquent, la structure syntaxique locale et la position dans la phrase ont pu engendrer les différences de résultats avec ceux de la littérature (Aborn, *et al.*, 1959; Friederici, 1985; Gough, 1983).

Nos résultats du test de complétion suggèrent que le contexte utilisé dans notre étude a joué un rôle important en imposant plus de contraintes, tant sur l'identité que sur la classe, pour les mots de classe ouverte par rapport aux mots de la classe fermée. Toutefois, cet effet reste restreint aux mots de basse fréquence. Des résultats similaires sont obtenus par Becker (1979), indiquant que les mots de basse fréquence profitent plus de l'apport du contexte que les mots de haute fréquence. Grâce à nos résultats, nous pouvons préciser que le traitement des mots de classe ouverte de basses fréquences est celui qui bénéficie le plus de l'apport du contexte. Globalement, nos résultats du test de complétion suggèrent qu'un phénomène d'anticipation est à l'œuvre lors du processus de compréhension de phrase, dans notre matériel utilisé et que ce phénomène est plus important lors du traitement des mots de classe ouverte de basse fréquence que pour les mots de classe fermée de fréquences similaires. Si tel est le cas, nous devrions trouver des résultats sur les PEs qui confirment ce test off-line.

#### 4.2 Observation des effets de classe pour les deux expériences

En présence d'un contexte, les résultats du chapitre 5 indiquent des différences topographiques entre les deux classes de mots dès 203 ms et ce, jusqu'à 334 ms (englobant deux segments différents). De plus, un segment est enregistré plus précocement (de 10.5 ms) pour les mots de classe fermée. A l'opposé, les mêmes stimuli présentés en isolation induisent des différences topographiques plus tardives et sur une durée plus longue: à partir de 295 ms et jusqu'à 596 ms, incluant 3 segments dont deux sont enregistrés plus précocement pour les mots de classe fermée. Dès 200 ms, l'effet du contexte est donc enregistrable d'une part, sur la topographie des PEs et, d'autre part, sur la latence du segment 5 (chapitre 6). Comme aucune tâche additionnelle, hormis la compréhension n'était demandée dans nos deux expériences, nous ne pouvons attribuer ces effets à des différences de tâches. En revanche, il semble possible d'avancer l'hypothèse que le contexte linguistique a pu affecter les PEs des deux classes de mots et ce, dès 200 ms. Par conséquent, nos résultats suggèrent que le contexte agit dès 200 ms post-stimulus, et qu'il influence différemment le traitement des deux classes de mots. De plus, l'avantage du segment 5 observé (décalage de 3.9 ms) pour les mots de classe fermée présentés en isolation (chapitre 6), est réduit à néant lorsque ceux-ci sont lus dans un contexte de phrase (pas de décalage observé sur le segment 6 pour le chapitre 5). Grâce au contexte linguistique (chapitre 5), le processus mis en place lors du segment 6, est différent (topographie différente), mais tout aussi rapide pour l'une ou l'autre des classes de mots.

### 4.3 Effet général du contexte sur les PEs

Lorsque l'on compare les PEs obtenus lors de la présentation de mots au sein d'un contexte et en isolation, deux résultats ressortent. Sur la fenêtre d'analyse comprise entre 0 et 800 ms, les PEs des mots en isolation sont généralement plus négatifs sur les régions postérieures et gauches. D'autre part, entre 150 et 800 ms, les PEs des mots présentés dans un contexte de phrase sont plus négatifs sur les régions centro-médianes par rapport aux PEs des mots présentés en isolation. Ce dernier effet présente les caractéristiques (topographie et rang de latence) de la CNV, pour *Contingent Negative Variation* (Coles & Rugg, 1995; Kutas & van Petten, 1994). Pour obtenir une telle composante, le paradigme prototypique est de présenter, à intervalles fixes, un stimulus "alerte" à partir duquel le sujet doit donner sa réponse. Pendant cette période de préparation, une CNV peut être enregistrée. La motivation, ou encore la complexité de la tâche, modifie l'amplitude de cette composante (Irwin, Knott, McAdam, & Rebert, 1966; McAdam, Irwin, Rebert, & Knott, 1966; McCallum & Papakostopoulos, 1973). Les premières interprétations fonctionnelles de la composante CNV indiquent que celle-ci reflète les processus cognitifs d'attente, et/ou de préparation du stimulus à venir (Hillyard, 1973). Une telle interprétation semble applicable à nos résultats. En effet, les conditions expérimentales des expériences en contexte et en isolation n'étaient pas identiques. Premièrement, aucune tâche n'était demandée aux participants lors de la lecture de mots isolés, alors que pour l'expérience présentant les phrases, les participants devaient répondre, dans 20 % des cas, à une question de compréhension. Cette différence a pu induire une variation dans le niveau d'implication des participants. Deuxièmement, comparée à la lecture de mots en isolation, la lecture d'une phrase implique des niveaux de traitement plus élaborés, quelle que soit la tâche demandée (van Petten & Kutas, 1990; van Petten & Kutas, 1991). Troisièmement, pour les deux expériences, les mots sont présentés à un rythme fixe <sup>21</sup>. Le rythme est identique pour les deux classes de mots alors que, lors de la lecture normale de phrases, les lecteurs omettent plus souvent de fixer les mots de classe fermée (Carpenter & Just, 1983).

---

<sup>21</sup> Pour l'expérience en contexte de phrase - chapitre 5 - le rythme est plus rapide (SOA de 800 ms) que celui utilisé pour l'expérience en isolation - chapitre 6 - (SOA de 1100 ms).

Ainsi, ce paradigme aurait du entraîner une période d'attente plus importante pour les mots de classe fermée que pour les mots de classe ouverte. Cependant, comme nous le montre le test de complétion, l'identité et la classe des mots de classe ouverte ont pu être mieux anticipées par la présentation du contexte. Par conséquent, enregistrer une composante CNV pour nos deux classes de mots, lorsque ceux-ci sont présentés dans un contexte de phrase ne semble pas inattendu.

#### 4.4 Effet spécifique sur la composante P100: 0-150 ms

Nous avons mis en évidence un effet du contexte linguistique, spécifique pour chaque classe de mots, dès la présentation du stimulus critique et ce, jusqu'à 150 ms. Cette fenêtre d'analyse couvre une composante précoce qui présente une négativité, sur les sites situés en région frontale (N100) et une positivité, sur les sites postérieurs (P100). Nos résultats indiquent que, sur les régions situées sur l'hémisphère droit, le contexte augmente la positivité des PEs pour les mots de classe ouverte de haute fréquence (CO-H), alors qu'aucun effet n'est observé pour les PEs des mots de classe fermée de même fréquence (CF-H). En région antérieure, la présentation d'un contexte augmente la négativité des PEs des mots de classe fermée de basse fréquence (CF-B), par rapport aux PEs des mots de classe ouverte de fréquence équivalente (CO-B). Alors qu'à partir d'une analyse visuelle des PEs, van Petten & Kutas (1991) ne relatent pas les effets précoces, Osterhout, *et al.* (1997) indiquent une composante N100 réduite pour les deux classes de mots lorsque les mots sont présentés en isolation (les résultats statistiques ne sont pas significatifs).

Les modulations de la composante P100 sont souvent liées à des variations attentionnelle. Il semble maintenant bien établi que, diriger spécifiquement l'attention vers un hémichamp visuel résulte dans des détections et/ou discriminations des stimuli plus exactes dans cette location (LaBerge, 1995). De récentes investigations électrophysiologiques ont essayé de préciser les mécanismes neuronaux à l'origine de tels effets. Dans les études où des stimuli sont présentés rapidement et de manière imprévisible, diriger l'attention dans l'un des hémichamps conduit à accroître l'amplitude de la P100 précoce (75-130 ms) et/ou de la négativité N2 (150-190 ms; Hillyard, *et al.*, 1998; Mangun, 1995). Ainsi, l'augmentation de l'amplitude de la positivité en région postérieure, et/ou de la négativité rapportée en région antérieure, semblent indiquer que, pour des étapes très précoces, l'attention visuelle sélective spatiale agit comme un mécanisme amplificateur qui module (en augmentant ou en

diminuant) l'importance de l'activité neuronale évoquée (Luck & Ford, 1998; Vogel & Luck, 2000).

Dans notre étude, l'hypothèse explicative de la modulation entre 0 et 150 ms de la négativité pour CF, et de la positivité pour CO, pourrait être liée à des modulations de l'attention sélective attribuée à chaque mot. Le test de complétion indique que des attentes spécifiques ont été élaborées pour les deux classes de mots. Pour chaque classe de mots, les attentes spécifiques induites par le contexte linguistique ont pu se répercuter différemment sur les PEs. Alors que pour les mots de classe fermée, ce gain est enregistré sur les sites antérieurs, uniquement pour les mots de basse fréquence (voir figure 74), les variations d'amplitude sont enregistrées sur l'hémisphère droit pour les mots de classe ouverte de haute fréquence (voir figure 75).

Une autre interprétation semble possible si l'on considère l'étude de Friederici & Mecklinger (1996) dans laquelle ces auteurs montrent l'occurrence d'une composante très précoce (dès 100 ms), la ELAN (*Early Left Anterior Negativity*), lors la présentation d'une violation de catégorie syntaxique grammaticale. Cette composante présente à la fois, la latence mais aussi la distribution antérieure de notre effet obtenu pour les mots de classe fermée. Ces auteurs associent cette composante à des traitements syntaxiques automatiques, pendant lesquels la structure syntaxique de la phrase est construite à partir des informations de catégories syntaxiques des mots (Friederici, 2002). Si cette interprétation est retenue, alors nos résultats suggèrent que, très précocement, les deux classes de mots ne sont pas traités de manière identique en contexte de phrase. Cette dernière interprétation semble applicable à nos résultats, mais il est certain que de futures études sont nécessaires, afin de pouvoir départager l'interprétation attentionnelle (c'est à dire non linguistique) de l'interprétation ELAN (c'est à dire hautement psycholinguistique).

#### 4.5 Effet spécifique sur la composante N280: 250-350 ms

Pour la fenêtre d'analyse 250 à 350 ms, nos résultats indiquent que le contexte affecte spécifiquement les PEs des mots CF en régions antéro-médianes gauches (diminution de la négativité), alors que celui-ci augmente la positivité des mots CO en région postérieure droite. Van Petten & Kutas (1991) rapportent des résultats similaires aux nôtres mais dans une fenêtre temporelle plus tardive et plus large (300-500 ms) que celle correspondant à



l'occurrence de la composante N280. Malheureusement, ces auteurs ne nous renseignent pas sur des effets plus précoces. Il semble possible que le chevauchement de leur fenêtre et celles utilisées dans notre analyse (250-350 et 350-500 ms) soit à l'origine du décalage des effets observés entre les deux études. Osterhout, *et al.* (1997) relatent que les PEs des mots CF sont plus négatifs entre 300 et 400 ms, lorsque ceux-ci sont présentés visuellement en isolation par rapport à un contexte (effet non significatif). Ainsi, nos résultats concordent avec les tendances observées dans la littérature. Toutefois, nous apportons une précision supplémentaire par rapport aux études de Osterhout, *et al.* (1997) et de van Petten & Kutas (1991) qui n'ont pas tenu compte de la fréquence des mots manipulés. Nous avons montré que le contexte module spécifiquement la négativité en régions antéro-médianes gauches des mots CF-B, alors que celui-ci affecte la positivité en région postérieure droite des mots CO-B. Dans cette fenêtre temporelle, le contexte n'affecte pas différemment les PEs pour les mots des deux classes lorsque la fréquence est supérieure à 2.6 log de fréquence. Nos résultats suggèrent que, dans cette fenêtre (250-350 ms), les effets de contexte sont essentiellement liés au traitement différencié des deux classes de mots, pour des fréquences inférieures à 2.6 log de fréquence (soit < 400 occurrences/million).

Les résultats du test de complétion indiquent que les participants arrivent mieux à prédire la classe et l'identité des mots de classe ouverte, par rapport aux mots de classe fermée et ce, pour les basses fréquences uniquement. Par conséquent, dans la fenêtre temporelle 250-350 ms, l'anticipation de la classe, ou encore de l'identité des mots à venir (prédictibilité), peut être un facteur influençant la distribution des effets sur le scalp. Ces résultats ne remettent pas en cause le fait que la composante N280 puisse refléter des processus d'accès lexicaux. En effet, la composante N400 s'avère, par exemple, sensible à des facteurs multiples, tant au niveau lexical (fréquence d'occurrence, longueur, etc.) qu'au niveau post-lexical (probabilité d'occurrence, répétition, etc.; Chwilla, *et al.*, 1998; Coulson & Federmeier, soumis).

Dans ce chapitre, nous avons vu que le contexte affecte les PEs des deux classes de mots, sur des régions différentes du scalp et ce, dans le rang de latence compris entre 250 et 350 ms. Ainsi, si l'hypothèse selon laquelle cette fenêtre temporelle correspond à une étape de traitement lors de l'accès au lexique s'avère vérifiée, alors nous pouvons proposer que le contexte linguistique affecte de manière spécifique cette étape. En effet, nos résultats attestent de modulations différentes de PEs, pour les deux classes de mots. Les informations contextuelles (prédictibilité) participent lors de la reconnaissance des mots de façon précoce.

Alors que les théories autonomistes (Forster, 1979) postulent une totale indépendance des processus d'accès au lexique face aux informations contextuelles, à l'opposé, les théories interactives (Morton, 1969; Morton, 1979) suggèrent que toutes les sources d'informations (sémantiques, syntaxiques, contextuelles) sont utilisées dès lors qu'elles sont disponibles, afin de construire une représentation de la phrase. Comme le postulent les modèles interactifs de la reconnaissance lexicale (Morton, 1969; Morton, 1979), nos résultats suggèrent une utilisation précoce (dès 250 ms) des informations apportées par le contexte linguistique.

## 5 Conclusions

Le contexte linguistique semble affecter différemment les PEs des deux classes de mots dans deux fenêtres temporelles différentes, entre 0-150 ms et entre 250-350 ms. Pour la première fenêtre (0-150 ms), les effets sont obtenus pour les deux rangs de fréquence (basse et haute), et deux interprétations possibles, attentionnelle ou purement psycholinguistique ont été émises. La deuxième fenêtre (250-350 ms) ne présente que des effets sur les basses fréquences et, comme le test de complétion l'indique, pourraient refléter des effets contextuels liés au processus de prédictibilité de la classe ou de l'identité du mot.

# CHAPITRE 8

## DISCUSSION GENERALE ET CONCLUSIONS

---

### *Résumé*

Les résultats combinés de nos différentes expériences motivent les conclusions de ce chapitre et suggèrent que la distinction entre mots de classe ouverte et fermée joue un rôle important dans l'organisation du lexique mental et/ou dans les processus d'accès. Cette dissociation entre les deux classes de mots ne peut pas être le résultat de différences dans les propriétés lexicales (fréquence et longueur). C'est à la lumière du modèle de Friederici (2002) que nos résultats seront repris et permettront l'adaptation d'un modèle neurocognitif définissant des représentations et/ou des processus d'accès au lexique séparés pour les deux classes de mots. Ce chapitre sera clôturé par des perspectives élargissant notre recherche au domaine auditif ainsi qu'à la détermination plus précise du type d'information contextuelle impliqué dans les résultats de notre chapitre 7.

## 1 Récapitulatif

Les mots de classe fermée et ouverte manifestent des propriétés linguistiques distinctes dans de multiples domaines tels que la phonologie, la morphologie, la syntaxe et la sémantique (chapitre 1, section 3). Cette constatation est désormais également reconnue dans le domaine de la psycholinguistique. Par exemple, à partir de nombreux paradigmes expérimentaux en psychologie cognitive (voir chapitre 1, section 4) des différences dans la reconnaissance des deux classes de mots ont pu émerger, et les données issues des neurosciences cognitives, plus particulièrement à partir des techniques électrophysiologiques, émettent l'hypothèse d'une composante permettant une dissociation entre les deux classes de mots, la N280 (voir chapitre 2, section 3). La latence ainsi que la distribution sur le scalp de cette composante ont été considérées comme spécifiques aux mots de classe fermée. Cette dernière hypothèse semble très séduisante, car la fenêtre d'apparition de cette composante, comprise entre 250 et 350 ms en moyenne, est estimée comme précoce et, faisant part entière avec les processus d'accès lexicaux (Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998).

Alors que la fréquence est une propriété essentielle dans la détermination de la rapidité et l'exactitude de la reconnaissance d'un mot, il a souvent été confondu les notions de classe de mots et fréquence, les mots de classe fermée étant en général plus fréquents et plus courts que les mots de classe ouverte. Aux balbutiements de ce millénaire, la revue de littérature ne permet aucune affirmation catégorique quant à savoir si la classe de mots permet une séparation des représentations et/ou des processus d'accès au lexique ou non. Comment alors affirmer que les propriétés intrinsèques aux deux classes de mots, et non pas leurs propriétés lexicales telles que la fréquence et/ou la longueur, sont responsables des résultats précédemment cités?

Cette thèse participe au débat portant sur la clarification du type d'informations permettant de structurer les représentations stockées dans le lexique lors de la compréhension du langage écrit. L'un de nos **premiers objectifs** était de définir si la classe de mot permet de structurer le lexique mental. Une première expérience (chapitre 4) nous a permis d'enregistrer les caractéristiques électrophysiologiques pour des mots représentatifs de chaque classe. En se basant sur les variations naturelles de la fréquence des stimuli au sein des deux classes, nous

avons déterminé que les caractéristiques de la composante N280 varient en fonction de la fréquence et/ou de la longueur des stimuli mais aussi en fonction de la classe de mots. La deuxième expérience, utilisant des items avec des propriétés lexicales contrôlées (chapitre 5), nous indique que la classe de mots permet de modifier les caractéristiques de la N280 (distribution spatiale sur le scalp mais aussi temporelle). Notre **deuxième objectif** était de manipuler la fréquence (haute et basse) pour chaque classe de mots afin de vérifier si l'effet de classe était homogène. Nos résultats suggèrent qu'il n'en est rien et que la fréquence d'occurrence modifie fortement l'effet de classe de mots. Le **troisième objectif** était de définir si le contexte affecte de façon similaire les PEs des deux classes de mots. Pour cela, nous avons comparé (chapitre 7) les mêmes mots présentés soit dans un contexte de phrase (chapitre 5) soit en isolation (chapitre 6). Nous avons pu définir que le contexte affecte très précocement et différemment les PEs lors de l'accès au lexique des deux classes de mots. Le **dernier objectif** développé dans notre thèse porte sur une contribution méthodologique face aux mesures électrophysiologiques. En comparant systématiquement les analyses de tracés et de segmentation adaptée, nous avons défini plus précisément les fenêtres temporelles dans lesquelles des différences de topographie apparaissaient.

Les résultats de notre recherche vont être analysés à partir du modèle neurocognitif développé par Friederici (2002; sous presse).

## 2 Modèle neurocognitif de Friederici 2002

Friederici (2002) propose un modèle neurocognitif autonome (Forster, 1976; Forster, 1979) de la perception d'une phrase dans lequel deux types de contraintes des données empiriques sont utilisés. D'une part, les restrictions temporelles sont issues de résultats à partir de données électrophysiologiques multiples telles que les PEs et la MEG (voir figure 77), alors que les restrictions spatiales sont issues des résultats à partir de données issues de la TEP et de l'IRMf. Dans cette section nous ne traiterons que des données temporelles qui sont le cœur de notre recherche. Ce modèle comporte quatre phases que nous allons brièvement discuter; le but étant d'ouvrir une perspective explicative face à nos résultats.

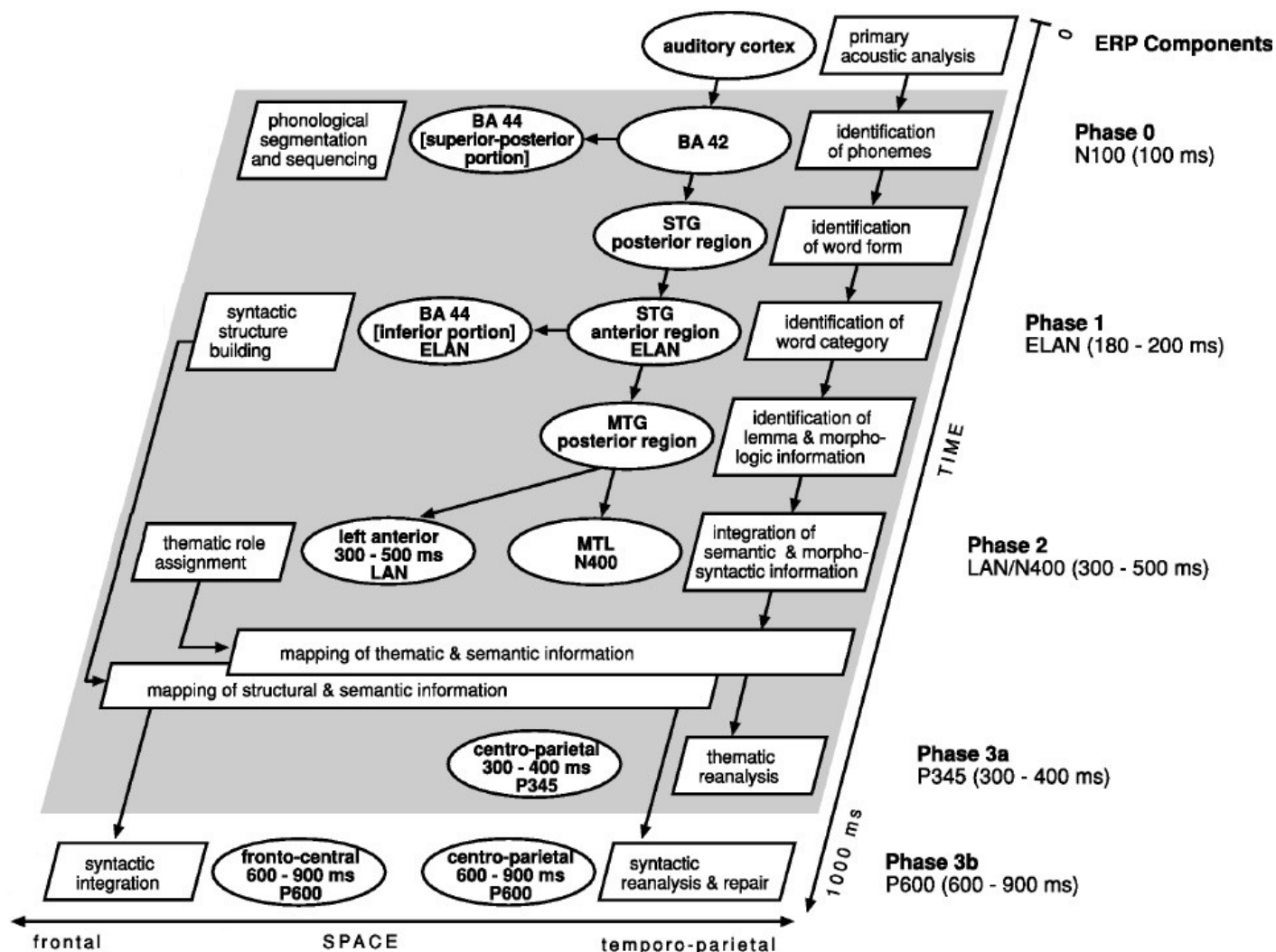


Figure 77: Modèle neurocognitif des processus de compréhension de phrase. D'après Friederici (2002).

Les différentes phases du modèle de Friederici sont déterminées, à partir de résultats d'expériences présentant des violations spécifiques. La première phase, **phase 0**, comprise entre 0 et 100 ms correspond à l'occurrence d'une composante N100. Celle-ci reflète l'analyse primaire acoustique ou visuelle du stimulus. Cette première phase perceptive n'est pas plus analysée par ces auteurs car non spécifique du traitement de la phrase. L'étape suivante ou **phase 1** correspond à la fenêtre temporelle comprise entre 150-300 ms. Lors de la présentation de violations de la catégorie syntaxique du mot, une composante précoce (180-200 ms) négative et maximale sur les régions antérieures gauches est enregistrée, la ELAN (*Early Left Anterior Negativity*). Selon Friederici, une première étape d'analyse utilise les informations de catégorie syntaxique du mot afin de construire la structure syntaxique de la phrase. Chaque nouveau mot est alors intégré dans cette structure, et si la catégorie syntaxique du mot ne correspond pas à la structure élaborée, alors une composante ELAN est enregistrée.

La **phase 2** correspond à l'étape de traitement où les informations lexico-sémantiques et morpho-syntaxiques sont disponibles. Après l'identification du mot dans le lexique mental, les autres informations qui lui sont associées sont disponibles: sémantique, et syntaxique tels que le genre des noms ou encore les informations sur la structure de l'argument du verbe. Selon Friederici, les informations sémantiques et morpho-syntaxiques sont utilisées en même temps, mais intégrées en parallèle dans deux représentations différentes sans qu'il n'y ait possibilité d'échanges d'informations. La composante N400, enregistrée lors de violation sémantique entre 300 et 500 ms, reflète l'activation des propriétés lexico-sémantiques du mot, et son amplitude suggère la difficulté d'intégration sémantique au sein du contexte linguistique. Dans cette même fenêtre temporelle, des violations d'accord entre le sujet et le verbe suscitent l'occurrence d'une composante LAN reflétant l'activation des informations morpho-syntaxiques. La dernière phase ou **phase 3** est définie par l'occurrence d'une composante tardive, la P600, et reflète la difficulté d'intégration des représentations structurale syntaxique, sémantique et morpho-syntaxique. Lorsque ces différentes représentations ne peuvent être correctement intégrées, alors une ré-analyse de la structure initiale se met en place afin de réparer les erreurs.

### 3 Analyse de nos résultats à la lumière du modèle de Friederici

#### 3.1 Mise en correspondance des segments et des phases

A partir des résultats du chapitre 4 nous allons mettre en relation nos micro-états avec les différentes phases du modèle de Friederici (2002). L'expérience du chapitre 4, bien que ne contrôlant ni la fréquence ni la longueur des mots, nous apporte tout de même des renseignements précieux en indiquant des différences de topographies dans 5 segments ou micro-états compris entre 199 et 686 ms (voir tableau 38). Puisque chaque période temporelle présente des distributions différentes sur le scalp, nous avons suggéré que ces 5 micro-états reflètent l'activation de générateurs différents ou tout au moins non superposés. Si l'on accepte l'hypothèse selon laquelle un processus cognitif va activer spécifiquement un groupe de neurones, alors nous pouvons conclure que nos différents micro-états reflètent l'activation d'étapes différentes au sein du traitement d'un mot dans un contexte de phrase (Fender, 1987; Lehmann, *et al.*, 1987).

Chapitre 4			Chapitre 6 - Isolé						Chapitre 5 - contexte					
Seg	Lat	Topo	Seg	Lat	HF		BF		Seg	Lat	HF		BF	
					Topo	Lat	Topo	Lat			Topo	Lat	Topo	Lat
6		CF: Ant Med G-C 199 246	5	-5.5			CF: Ant G-C et Med G 201 288		6		CF: Ant D 203 258			
7	-13.1	CF: Ant G-C et Med C 247 340	6	-9.4			CF: Ant (3) 289 390		7	-7.2	CF: Ant Med C-D 259 342			
8		CF: Ant G-C et Med G 341 440	7				CF: Gauche + Ant C 391 458		8					
9		CF: Ant Med G-C CO: Ant D Post (3) 441 492 493	8		CF: Ant D et Post G 449		CF: Med G CO: Ant D Post D 459		9				CF: Post C-D 443 492	
10		CO: Post (3) 493 686							10				CF: Post C 493 664	

**Tableau 38:** Tableau récapitulatif de nos résultats portant sur la distinction entre classe de mots. Pour chaque expérience sont rapportés les résultats significatifs (effets de classe de mot) pour les analyses portant sur la topographie (Topo) et le décalage de latence (Lat). Les frontières des segments sont rapportées à droite dans la colonne Topo. Seules les différences de négativités sont indiquées entre les deux classes de mots. CF: indique que les PEs des mots de classe fermée sont plus négatifs que ceux de classe ouverte. CO: indique que les PEs des mots de classe ouverte sont plus négatifs que ceux de classe fermée. Les distributions de ces effets sont indiquées avec les abréviations suivantes. Ant : antérieur, Post : postérieur, Med: médian, G: gauche, D: droite, C: central. HF: mots de haute fréquence, BF: mots de basse fréquence. Seg: segment, Lat: latence, Topo: topographie. Les différences de latence négative indiquent que la carte représentative du segment est plus précoce pour les mots de classe fermée que celle des mots de classe ouverte.

Le premier segment, numéro 6 (compris entre 199 et 246 ms) indique la présence d'une composante négative de distribution antérieuro-médiane gauche pour les mots de classe fermée (voir tableau 38). La topographie ainsi que la fenêtre temporelle de notre segment 6 correspondent aux critères définis pour la composante ELAN enregistrée dans la phase 1 du modèle de Friederici (2002). Dans cette expérience nous ne présentons pas de violations de catégorie syntaxique, mais il semble possible de suggérer que dans un contexte de phrase, les informations spécifiques de la classe du mot soient activées et participent à la construction d'une première structure syntaxique.

Le deuxième segment, numéro 7, présente des résultats similaires à Keurs ter, *et al.* (1999; voir tableau 2, chapitre 2, section 3.2). En effet, nous indiquons l'enregistrement de deux composantes N280 distinctes en topographie pour les deux classes de mots. En revanche, contrairement à Brown, *et al.* (1999) et à Keurs ter, *et al.* (1999), la latence de cette composante précoce est corrélée et varie linéairement avec la fréquence d'occurrence des mots ainsi que leur longueur (King & Kutas, 1998; Münte, *et al.*, 2001; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997; chapitre 4, figure 34).



La fréquence d'occurrence étant un facteur affectant les processus d'accès lexicaux, cette composante négative frontale précoce, enregistrée entre 247 et 340 ms, pourrait refléter des processus d'accès au lexique comme le suggèrent certaines études (Keurs ter, *et al.*, 2002; Keurs ter, *et al.*, 1999; King & Kutas, 1998; Osterhout, *et al.*, 2002; Osterhout, *et al.*, 1997). La fenêtre temporelle d'apparition de ce segment 7 correspond à celle de la composante N280 et au descriptif de l'étape 1 du modèle de Friederici (2002) dans laquelle l'identification du mot prendrait place. Alors qu'avec un paradigme de présentation de violation aucune composante spécifique n'a pu être mise en évidence, nos résultats suggèrent que la composante N280 diffère pour les deux classes de mots. Cette composante met en évidence l'accès à des représentations et/ou à des processus d'accès lexical séparés et qui sont plus rapides (de 13.1 ms) pour les mots de classe fermée. N'oublions pas que la latence de cette composante est corrélée avec la fréquence et la longueur des mots, résultat incitant à conclure que cette composante ne reflète tout au moins pas uniquement l'accès à des représentations et/ou à des processus d'accès lexical séparés mais aussi des indices de fréquence et de longueur.

Les fenêtres temporelles correspondant aux segments 8 (341 - 440 ms), 9 (441 - 492 ms) et 10 (493 - 686 ms) sont incluses dans la phase 2 du modèle de (Friederici, 2002). Pour les mots de classe fermée, une composante LAN est enregistrée lors des segments 8 et 9, alors que pour les mots de classe ouverte, une composante N400 est observée lors des segments 9 et 10. Les composantes enregistrées dans ces segments suggèrent donc une intégration des informations morpho-syntaxiques pour les mots de classe fermée (LAN) et lexico-sémantiques pour les mots de classe ouverte (N400). Puisque deux composantes, N400 et LAN, sont enregistrées respectivement pour chaque classe de mots dans un contexte de phrase, nous suggérons que les mots des deux classes ne portent pas les mêmes informations et que ceux-ci participent à l'élaboration de représentations séparées.

### 3.2 Accès lexical sans apport de contexte

Nous allons tout d'abord commenter les résultats du chapitre 6 lorsque les mots sont présentés en isolation, ceci afin d'observer l'effet de classe sans que le contexte linguistique ne puisse affecter ces processus.

Les mots de haute fréquence présentent des différences de latence dans l'apparition des segments précoces (5 et 6), et des différences de topographies tardives pour le segment 8 (449-628 ms, voir tableau 38, résultats pour le chapitre 6). Pour le segment 5 (201-288 ms), nos résultats indiquent que la négativité antéro-médiane gauche représentative des PEs associés aux mots de classe fermée est enregistrée 5.5 ms plus précocement que celle des mots de classe ouverte. Comme les mots de classe fermée de haute fréquence sont très représentatifs de cette classe, et si l'on considère que ceux-ci participent à l'élaboration de la structure syntaxique de la phrase, alors ce résultat peut s'accorder avec l'hypothèse selon laquelle l'information de la classe fermée doit être activée rapidement afin de permettre l'élaboration de la structure syntaxique de la phrase. Nos résultats indiquent que même en isolation, cette voie spécifique est activée. Le segment 6 présente lui aussi un décalage (9.4 ms) dans l'apparition du potentiel associé aux mots de classe fermée par rapport à celui de la classe ouverte. Ce résultat suggère donc que l'activation de la représentation et/ou du processus d'accès lexical est plus rapide pour les mots de classe fermée que celui des mots de classe ouverte.

Pour le segment 6 et pour les basses fréquences uniquement, nous indiquons une latéralisation différente de la composante N280 pour les mots de classe fermée par rapport à celle des mots de classe ouverte. Nos résultats corroborent donc ceux de Keurs ter, *et al.* (2002), Münte, *et al.* (2001), Nobre & McCarthy (1994) et de Pulvermüller, *et al.* (1995; voir tableau 2). Les résultats rapportés dans cette thèse suggèrent donc fortement que des représentations séparées ou bien des processus d'accès différents existent pour les deux classes de mots pour des fréquences inférieures à 400 occurrences par million, et apportent des arguments en faveur des hypothèses de Bradley (1978), Friederici & Schoenle (1980) et de Garrett (1975).

Le lecteur aura pu remarquer que nous concluons à des "distinctions entre représentation lexicales et/ou des processus d'accès lexicaux". L'une des limites des recherches en psycholinguistique présentées dans le chapitre 1 est qu'elles ne permettent pas de séparer représentations et processus d'accès malgré la très bonne résolution temporelle de la technique utilisée (voir chapitre 1, tableau 1). Il est capital pour les théories de pouvoir dissocier ces deux effets (Rapp & Caramazza, 1993), mais cela est malheureusement actuellement impossible expérimentalement. Nos résultats montrant des différences de topographie entre les deux classes de mots peuvent s'adapter aussi bien aux deux interprétations suivantes: 1) les deux classes de mots sont stockées dans deux lexiques

mentaux différents et les mêmes processus d'accès sont utilisés; ou bien 2) les deux classes de mots sont stockées dans un même et unique lexique mental, mais deux processus d'accès différents sont utilisés. De futures expérimentations, et même peut être l'élaboration de nouvelles techniques d'imageries cérébrales différemment sensibles, devront être réalisées afin de pouvoir départager ces deux interprétations.

### 3.3 Accès lexical avec apport de contexte

Alors qu'en isolation les mots de basses fréquences présentaient des différences de topographies très précoces (201 ms), avec l'apport du contexte celles-ci sont enregistrées à partir de 443 ms. Il semble intéressant de noter que notre test de complétion indique que l'amorce de contexte utilisé avant le mot critique induit des contraintes plus importantes quant à l'apparition de la classe et de l'identité des mots de classe ouverte de basse fréquence par rapport aux mots de classe fermée de fréquence équivalente. Nos résultats en contexte de phrase n'indiquant pas de différence topographique précoce pour les mots de bases fréquences (voir tableau 38) peuvent s'accorder avec l'interprétation selon laquelle l'amorce de contexte a facilité l'accès aux informations de la classe et ce préférentiellement pour les mots de classe ouverte de basse fréquence. Si les mots de classe fermée sont reconnus via une voie spécifique lors de l'accès lexical (Friederici, 2002), alors nos résultats indiquent que grâce à une amorce de contexte les mots de classe ouverte ont pu aussi bénéficier de celle-ci.

Nous avons aussi montré qu'en isolation, deux micro-états sont plus précoces pour les mots de classe fermée par rapport aux mots de classe ouverte de haute fréquence, alors que dans un contexte de phrase un seul segment présente des différences de latence entre les deux classes de mots (tableau 38). Nous pouvons donc suggérer que les informations véhiculées par le contexte ont aussi affecté les PEs des mots de haute fréquence, ceci en permettant d'annuler le décalage de latence pour le premier micro-état observé pour les mots de classe fermée.

Une analyse statistique comparative de ces résultats (chapitre 7) indique que les informations apportées par le contexte peuvent influencer différemment la topographie des PEs pour les deux classes de mots dans la fenêtre temporelle comprise entre 250 et 350 ms. Ces résultats suggèrent que les informations contextuelles agissent de façon précoce dans la reconnaissance des mots des deux classes. Par conséquent, nos résultats penchent vers l'hypothèse d'une perméabilité de l'accès lexical face aux informations contextuelles (Friederici & Schoenle,

1980; Herron & Bates, 1997; Shillcock & Bard, 1993) et valident par conséquent les modèles plus généraux sur la reconnaissance des mots de type interactif (Morton, 1969; Morton, 1979). Le problème émis ici est que nous ne pouvons aisément spécifier quels types d'informations incluses dans le contexte sont responsables de cette influence. Par exemple, Friederici & Schoenle (1980) concluent que les informations contextuelles syntaxiques affectent la reconnaissance des mots de classe fermée, alors que les informations contextuelles sémantiques n'affectent que la reconnaissance des mots de classe ouverte. Nos résultats ne nous permettent pas de trancher entre les deux types d'informations incluses dans le contexte phrastique. Nous verrons dans nos perspectives quelle nouvelle expérience pourrait résoudre ce problème.

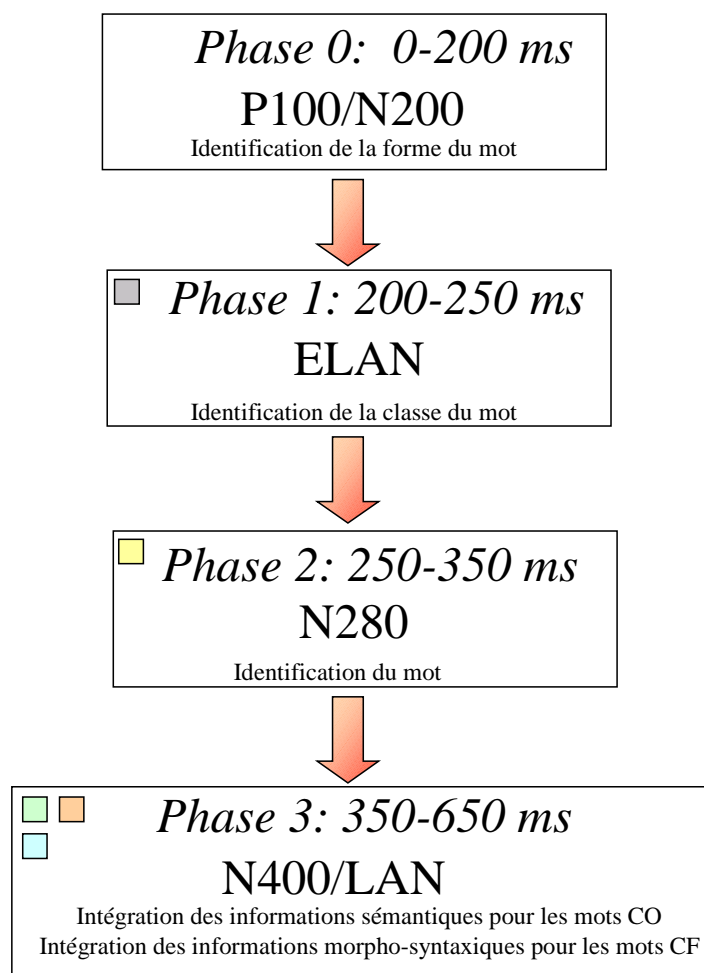
## 4 Contribution de nos recherches

### 4.1 Adaptation du modèle neurocognitif de Friederici

Dans un premier temps nous tenons à préciser que notre adaptation du modèle de Friederici (2002) n'est pas fondée sur des résultats portant sur la présentation de violations, contrairement au modèle originel. Au contraire, celui-ci est développé à partir de situation se rapprochant de la lecture normale et s'appuie sur les résultats de l'effet de classe de mots pour ajouter des contraintes. Malgré la présence de violations sémantiques dans notre chapitre 4, les analyses de l'effet de classe de mots ne sont réalisées qu'à partir des phrases contrôles. Quatre phases ont été retenues et sont présentées dans la figure 78.

- **Phase 0:** P100/N200 enregistrée entre 0 et 200 ms.

Cette phase correspond à l'enregistrement des composantes précoces perceptives et englobe la composante P100, avec un maximum sur les sites postérieur et la composante N2 avec un minimum sur les sites postérieurs gauches. A travers nos différentes expériences aucune différence de topographie entre les deux classes de mots n'est apparue dans cette fenêtre temporelle, suggérant que pour des processus visuels, cette première étape n'est pas spécifique à l'une ou l'autre de classe de mots. Cette première phase ne sera donc pas plus explicitée ici. Notons tout de même que nous avons enregistré un effet de la présentation du contexte (chapitre 7) sur cette première phase qui est spécifique à chaque classe de mot.



**Figure 78:** Adaptation du modèle neurocognitif de Friederici au processus de compréhension de phrase écrite. Les carrés de couleur correspondent aux différences de topographies enregistrées entre les mots de classe fermée et ouverte à partir de nos différentes expériences résumées dans le tableau 38.

- **Phase 1:** ELAN comprise entre 200 et 250 ms.

Alors que pour cette fenêtre temporelle, Friederici (2002) indique l'activation des informations de la catégorie syntaxique du mot, nous proposons que ce micro-état (numéro 6 pour le chapitre 4 et 6, et numéro 5 pour le chapitre 6) soit défini par l'activation des informations de la classe du mot. Nous suggérons que cette étape reflète l'activation des informations de classe de mots qui vont être utilisées lors de l'analyse syntaxique primaire de la phrase. Alors que pour les mots de basse fréquence l'activation de deux représentations ou de deux processus différents est observée, pour les mots de haute fréquence cette étape semble identique pour les deux classes de mots, quoique plus rapide pour les mots de classe fermée que pour ceux de classe ouverte. Nous avons émis l'hypothèse que pour les mots de haute fréquence la rapidité d'accès aux mots de classe fermée reflète la mise en place d'un processus automatique. Nous avons appelé cette étape ELAN, pour *Early Left Anterior Negativity*. Avec les données que nous avons en main, nous ne pouvons définir précisément si la négativité enregistrée dans nos expériences est similaire à celle de Friederici. En revanche, nous pouvons suggérer qu'elles reflètent des aspects syntaxiques des processus de traitement de la phrase bien que très certainement à des niveaux différents.

- **Phase 2:** N280 comprise entre 250 et 350 ms

Cette fenêtre temporelle correspond à la deuxième partie de la phase 1 du modèle de Friederici (2002). Dans cette période les auteurs proposent l'activation du lemma. Pour des basses fréquences, la composante N280 présente des distributions sur le scalp différentes pour les deux classes de mots suggérant des représentations et/ou des processus d'accès séparés. Pour des mots de fréquence supérieure à 400 occurrences par million, cette étape est activée plus rapidement pour les mots de classe fermée que pour ceux de classe ouverte. Nous avons montré que le contexte affecte différemment les PEs des classes de mots de basse fréquence, suggérant ainsi que l'identification de la représentation au sein du lexique puisse être modulée par le contexte. En revanche, nous ne pouvons préciser la nature des informations contextuelles permettant cette influence.

- **Phase 3:** N400/LAN comprise entre 250 et 650 ms

Après l'identification de l'entrée lexicale dans le lexique mental, les informations spécifiques qui lui sont associées sont disponibles. A partir de la présentation de violations d'accord entre le sujet et verbe (Kutas & Hillyard, 1983; Osterhout & Mobley, 1995) ou encore d'accord de genre entre le déterminant et le nom (Gunter, Friederici, & Schriefers, 2000) une composante LAN est enregistrée. Cette composante reflète l'activation des informations morpho-syntaxiques du mot et sa difficulté d'intégration dans la représentation syntaxique de la phrase (Kutas & Hillyard, 1983; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995). A l'opposé, la difficulté d'intégration des informations lexico-sémantiques montrent l'occurrence d'une composante N400 (Kutas & Hillyard, 1980). Ces deux dernières composantes sont le support de la phase 2 du modèle de Friederici (2002). Sur la base de nos résultats sur les mots de basses fréquences présentés en isolation, nous indiquons l'occurrence d'une composante LAN pour les mots de classe fermée et une composante N400 pour les mots de classe ouverte. Bien qu'aucun processus d'intégration ne soit requis ici puisque les mots sont présentés dans des listes, il semble possible que ces deux classes de mots aient activé des informations qui leur sont spécifiques. Notons toutefois que cette ségrégation du type d'informations n'est évidente que pour les mots de basse fréquence. Si l'hypothèse selon laquelle ces deux composantes LAN et N400 reflètent l'activation et la difficulté d'intégration des informations morpho-syntaxiques et lexico-sémantiques, alors lorsque les mots sont lus en contexte et facilement intégrés, ces deux composantes devraient être réduites. Les résultats en présence de contexte corroborent cette hypothèse en indiquant une totale absence d'effet N400 et LAN pour les mots de haute et de basse fréquence.

## 4.2 La segmentation adaptée

Le chapitre 3 de cette thèse apporte une contribution méthodologique aux analyses portant sur les PEs. Nos analyses en parallèle, des tracés et de segmentation adaptée, sur nos différentes expériences ont permis de montrer l'utilité de la segmentation adaptée, mais aussi l'utilité parfois de combiner les deux méthodes d'analyses. En outre, l'analyse de segmentation a pu apporter des précisions temporelles intéressantes comparée à l'analyse de tracés. Les principaux points positifs que nous retiendrons de la segmentation adaptée sont les suivants. D'une part, cette analyse permet de définir des fenêtres temporelles motivées par les modulations topographiques électriques dans chaque condition expérimentale et ne reposent pas sur des choix arbitraires. D'autre part, elle inclut pour chaque étape de l'analyse toutes les électrodes enregistrées. Cette analyse permet donc d'éviter des choix arbitraires nécessaires lors de l'analyse de tracés et par conséquent peut révéler des effets non encore mis en évidence. Plus concrètement, la segmentation adaptée a permis de diviser une fenêtre temporelle large telle que celle de la composante N400 (comprise en générale entre 300 et 500 ms) en deux micro-états différents. Ces résultats élargissent ceux de Brandeis *et al.* (1995) à des violations non plus en fin mais au milieu de phrases. Alors que, par exemple cette division de la composante N400 en deux sous-processus pour des violations en milieu de phrases n'était évidente qu'à partir de présentation auditive (Hagoort & Brown, 2000), la découverte de ce résultat dans la modalité visuelle permet d'ouvrir le débat sur des interprétations fonctionnelles non dépendantes de la modalité de présentation des stimuli. Nous avons également pu définir des micro-états qui à l'heure actuelle non pas encore pu être précisément caractérisés d'un point de vue fonctionnel. Par exemple, nous avons émis des hypothèses quant à l'interprétation de ce premier segment présentant des différences de topographie, le segment 7 (247 - 338 ms), mais il est évident que de futures recherches doivent être mises en place afin de vérifier nos hypothèses.

## 5 Perspectives

Dans ce travail, nous nous sommes restreints à l'étude de la modalité visuelle. Etant donné que dans notre introduction nous indiquions que la compréhension du langage oral était la modalité la plus utilisée pour communiquer, il nous semble pertinent de questionner la possibilité d'une dissociation des représentations et/ou processus d'accès lexical entre les

deux classes de mots dans la modalité auditive. Une première réponse peut être apportée par les observations de patients cérébrolésés indiquant des déficits spécifiques à la fois pour une classe de mots (par exemple des verbes) et pour une modalité par exemple la production (Caramazza & Hillis, 1991). Dans cette lignée, Rapp & Caramazza (1997) observent un patient avec une production orale caractérisée par des difficultés pour pratiquement tous les mots exceptés ceux de la classe fermée. A l'opposé, la production écrite de ce patient présente un pattern complémentaire à travers lequel les mots de classe fermée sont absents ainsi que les noms mal utilisés. Sur la base de ces observations, ces auteurs indiquent que pour réaliser de telles performances ce patient dispose de deux copies du lexique mental, comme le suggère Morton (1969; 1979). A l'opposé, Herron (1998) et Herron & Bates (1997) postulent que les propriétés acoustico-phonologiques des mots permettent de mettre en place très rapidement des processus spécifiques pour chaque classe de mots. Selon ces auteurs, il n'est en aucun cas nécessaire d'activer deux lexiques mentaux séparés lors de la reconnaissance des deux classes de mots. Si tel est le cas, les différences de topographies entre les deux classes de mots doivent être très précoces et incluses dans la phase 0 (entre 0 et 200 ms) et non plus apparaître dans les étapes suivantes comme nous l'avons montré dans des processus visuels. Une expérience intéressante serait donc de présenter nos mots contrôlés en fréquence et en longueur dans la modalité auditive afin de vérifier si nous obtenons des composantes distinctives similaires par rapport à la modalité visuelle. Le cas échéant, cette dernière expérience pourra alors nous permettre de valider l'hypothèse selon laquelle les propriétés acoustico-phonologiques des mots ne sont pas suffisantes pour activer des processus spécifiques aux deux classes de mots.

D'autre part, la présentation auditive des stimuli peut aussi nous permettre de tester des populations différentes. Alors que notre recherche a porté uniquement sur des adultes sains, la présentation de nos stimuli contrôlés à différentes populations peut apporter des indices précieux sur la reconnaissance des deux classes de mots. Les résultats de Neville, *et al.* (1993) et de Neville & Mills (1997) semblent converger vers l'évidence d'une distinction précoce (dès 36 mois) entre les deux classes de mots. Il serait donc intéressant de déterminer si, chez ces enfants, cette distinction existe indépendamment des propriétés lexicales confondues. De plus, les populations présentant des déficits du langage (tels que des patients aphasiques), ou encore des troubles de développement du langage (Troubles Spécifiques du Langage, TSL) sont plus faciles à tester à l'oral qu'à l'écrit. Ces études pourront apporter de nouvelles



sources et contraintes pour les modèles de reconnaissance portant sur les deux classes de mots.

Une autre perspective possible de notre travail est de mieux cerner quel type d'informations contextuelles affecte la topographie de la composante N280. Ainsi nous proposons de tester l'hypothèse selon laquelle les informations sémantique et syntaxique ne vont pas affecter de manière similaire les PEs des deux classes de mots. En utilisant toujours nos deux listes de mots contrôlés, nous proposons de modifier les informations sémantiques présentes dans le contexte. Pour ce faire, nous utiliserons la méthode décrite par van Petten & Kutas (1991) dans laquelle ces auteurs remplacent tous les mots de classe ouverte par un mot de même catégorie grammaticale. Ainsi les informations syntaxiques contextuelles sont-elles conservées, mais celles sémantiques ne sont plus disponibles. Dans ce cas précis, nous avançons l'hypothèse que seuls les PEs des mots CF seront affectés dans la fenêtre temporelle de la composante N280 par la présence d'informations syntaxiques alors ceux de classe ouverte ne le seront pas. Ainsi, ce type d'hypothèse penche vers une explication des effets de contexte de type interactif, mais dans lequel les informations syntaxiques ne peuvent affecter que la reconnaissance des mots de classe fermée (Friederici, 1985; Friederici, 1986; Friederici & Schoenle, 1980; Shillcock & Bard, 1993).

## 6 Conclusions

Notre recherche portant sur la distinction entre classe de mots a permis de mettre en évidence l'imbrication étroite des effets de classe, de contexte et de fréquence. Toutefois, cette thèse a permis de lever l'ambiguïté entre les effets des propriétés lexicales (fréquence et longueur) de ceux de classe de mots. Bien que les propriétés lexicales (fréquence et longueur) semblent jouer un rôle important pour caractériser les PEs sur le scalp, d'après nos recherches, en contexte de phrase et en isolation, nous pouvons avancer l'idée qu'elles ne sont pas suffisantes pour rendre compte des différences entre les deux classes ouverte et fermée. En effet, la classe de mots affecte la topographie ainsi que la latence de cette composante négative précoce. Ce résultat non des moindres va permettre d'ouvrir le champ vers de nouvelles recherches portant sur l'étude des propriétés linguistiques spécifiques (phonologie, morphologie, syntaxe, sémantique) permettant cette différenciation dans les représentations et/ou des processus d'accès au lexique.

BIBLIOGRAPHIE

---

- Aborn, M., Rubenstein, H., & Sterling, T.D. (1959). Sources of contextual constraint upon words in sentences. *Journal of Experimental Psychology*, 57(3), 171-180.
- Allison, T., Wood, C.C., & McCarthy, G.M. (1986). The central nervous system. In S.W. Porges (Ed.), *Psychophysiology: Systems, processes, and applications* (pp. 5-25). New York: Guilford.
- Andreewsky, E., & Seron, X. (1975). Implicit processing of grammatical rules in a classical case of agrammatism. *Cortex*, 9, 379-390.
- Andrews, S. (1989). Frequency and neighborhood effects on lexical access: activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 802-814.
- Balota, D. (1994). Visual word recognition: the journey from features to meaning. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego: Academic Press.
- Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of frequency in the neglected decision stage. *Journal of Experimental Psychology: Human, Perception and Performance*, 10, 340-357.
- Bastiaansen, M.C.M., & Hagoort, P. (2000). Event-related power changes during sentence processing. *Annual Report* (pp. 195). Nijmegen: Max Planck Institut fur Psycholinguistik.
- Bastiaansen, M.C.M., & Hagoort, P. (2001). Induced EEG responses to open-class and closed-class words. *Annual Report* (pp. 188). Nijmegen: Max Planck Institut fur Psycholinguistik.
- Bastiaansen, M.C.M., van Berkum, J.J.A., & Hagoort, P. (2002). Event-related theta responses in the human EEG differentiate between gender and number agreement violations during online sentence processing. *Neuroscience Letters*, 323(1), 13-16.
- Bates, E., & Wulfeck, B. (1989). Crosslinguistic studies of aphasia. In E. Bates (Ed.), *The crosslinguistic studies of sentence processing*. (pp. 328-374). New York: Cambridge University Press.
- Becker, C.A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human, Perception and Performance*, 5(2), 252-259.
- Benson, D.F. (1979). Neurologic correlates of anomia. In H. Whitaker (Ed.), *Studies in neurolinguistics* (Vol. 4): Academic Press.
- Berger, H. (1929). Uber das Elektrenkephalogramm das menchen. *Archiv für Psychiatrie*, 87, 527-570.
- Bertand, O., Perrin, F., & Pernier, J. (1985). A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalography and Clinical Neurophysiology*, 62, 462-464.
- Besner, D. (1988). Visual word recognition: Special-purpose mechanisms for the identification of open and closed class items? *Bulletin of the Psychonomic Society*, 26(2), 91-93.
- Besson, M., Faita, F., Czternasty, C., & Kutas, M. (1997). What's in a pause: event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology*, 46(1), 3-23.

- Besson, M., & Kutas, M. (1993). The many facets of repetition: a cued-recall and event-related potential analysis of repeating words in same versus different sentence contexts. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*(5), 1115-1133.
- Besson, M., Kutas, M., & van Petten, C. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience*, *4*(2), 132-149.
- Biassou, N., Obler, L.K., Nespoulous, J.L., Dordain, M., & Harris, K.S. (1997). Dual processing of open- and closed-class words. *Brain and Language*, *57*(3), 360-373.
- Bloom, L. (1970). *Language development: form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bolinger, D. (1975). *Aspect of language*. New-York: Hartcourt Brace Janovich.
- Bowers, J.S. (2000). In defense of abstractionist theories of repetition priming and word identification. *Psychonomic Bulletin*, *7*(1), 83-99.
- Bradley, D.C. (1978). *Computational distinctions of vocabulary type*. Unpublished doctoral dissertation, MIT.
- Bradley, D.C., & Garrett, M.F. (1983). Hemisphere differences in the recognition of closed and open class words. *Neuropsychologia*, *21*(2), 155-159.
- Bradley, D.C., Garrett, M.F., & Zurif, E.B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes*. Cambridge, Mass: MIT Press.
- Brandeis, D., & Lehmann, D. (1994). ERP mapping: A tool for assessing language disorders? (Chap 10). In G.R. Mangun (Ed.), *Cognitive Electrophysiology* (pp. 239-247). Boston: Birkhäuser.
- Brandeis, D., Lehmann, D., Michel, C.M., & Mingrone, W. (1995). Mapping event-related brain potential microstates to sentence endings. *Brain Topography*, *8*(2), 145-159.
- Brandeis, D., Naylor, H., Halliday, R., Callaway, E., & Yano, L. (1992). Scopolamine effects on visual information processing, attention, and event-related potential map latencies. *Psychophysiology*, *29*(3), 315-336.
- Brandeis, D., van Leewen, T.H., Rubia, K., Vitacco, D., Steger, J., Pascual-Marqui, R.D., & Steinhausen, H.-C. (1998). Neuroelectric mapping reveals precursor of stop failures in children with attention deficits. *Behavioural Brain Research*, *94*(1), 111-125.
- Brown, C.M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience*, *11*(3), 261-281.
- Cann, R. (2000). Functional versus lexical: a cognitive dichotomy. In R.D. Borsley (Ed.), *The nature and function of syntactic categories*. London: Academic Press.
- Caramazza, A. (1996). Neuropsychology. The brain's dictionary. *Nature*, *380*(6574), 485-486.
- Caramazza, A., & Berndt, R.S. (1978). Semantic and syntactic processes in aphasia: A review of the literature. *Psychological Bulletin*, *85*(4), 898-918.
- Caramazza, A., & Hillis, A. (1991). Lexical organisation of nouns and verbs in the brain. *Nature*, *349*, 788-790.
- Caramazza, A., & Shelton, J.R. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*(1), 1-34.
- Carpenter, P., & Just, M. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading* (pp. 275-307). New York: Academic.
- Carreiras, M., Perea, M., & Grainger, J. (1997). Orthographic neighborhood effects on visual word recognition in Spanish: cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 857-871.
- Chambers, S.M., & Forster, K.I. (1975). Evidence for lexical access in a simultaneous matching task. *Memory and Cognition*, *3*, 549-559.

- Chiarello, C., & Nuding, S. (1987). Visual field effects for processing content and function words. *Neuropsychologia*, 25(3), 539-548.
- Chomsky, N. (1995). Categories and transformations. *Minimalist Program*. Cambridge, MA: MIT Press.
- Chwilla, D.J., Brown, C.M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology*, 32(3), 274-285.
- Chwilla, D.J., Hagoort, P., & Brown, C.M. (1998). The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *The Quarterly Journal of Experimental Psychology*, 51A(3), 531-560.
- Clark, V.P., & Hillyard, S.A. (1996). Spatial selective attention affect early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, 8, 387-402.
- Coles, M., Gratton, G., & Fabiani, M. (1990). Event-related potentials. In L.G. Tassinari (Ed.), *Principles of psychophysiology: physical, social, and inferential elements*. New York: Cambridge University Press.
- Coles, M.G.H., & Rugg, M.D. (1995). Event-related brain potentials: an introduction. In M.G.H. Coles (Ed.), *Electrophysiology of mind: event-related brain potentials and cognition*. (pp. 1-26). Oxford: University Press.
- Coltheart, M. (1980). Deep dyslexia: A review of the syndrome. In J.C. Marshall (Ed.), *Deep dyslexia* (pp. 22-47). London: Routledge & Kegan Paul.
- Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535-555). Hillsdale, NJ: Erlbaum.
- Connolly, J.F., & Phillips, N.A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6(3), 256-266.
- Connolly, J.F., Stewart, S.H., & Phillips, N.A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and Language*, 39, 302-318.
- Content, A., Mousty, P., & Radeau, M. (1990). Brulex: une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, 90, 551-566.
- Curran, T., Tucker, D.M., Kutas, M., & Posner, M.I. (1993). Topography of the N400: brain electrical activity reflecting semantic expectancy. *Electroencephalography and Clinical Neurophysiology*, 88(3), 188-209.
- Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research*, 22(2), 109-131.
- Cutler, A., & Norris, D.G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human, Perception and Performance*, 14, 113-121.
- Dell, G.S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5(4), 313-349.
- Donchin, E., & Heffley, E. (1978). Multivariate analysis of event-related potential data: a tutorial review. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related brain potential research*. (pp. 555-572). Washington DC: US Government Printing Office.
- Drewnowski, A., & Healy, A.F. (1977). Detection errors on the and and: Evidence for reading units larger than the word. *Memory and Cognition*, 5(6), 636-647.
- Duchek, J.M., & Neely, J.H. (1989). A dissociative word-frequency x levels-of-processing interaction in episodic recognition and lexical decision task. *Memory and Cognition*, 17, 148-162.

- Federmeier, K.D., & Kutas, M. (1999). A rose by any other name: long term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469-495.
- Federmeier, K.D., Segal, J.B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class ambiguous words in context. *Brain*, *123*, 2552-2566.
- Fender, D.H. (1987). Source localization of brain activity. In A. Remond (Ed.), *Handbook of Electroencephalography and Clinical Neurophysiology (Vol 1: Methods of analysis of brain electrical and magnetic signals)* (pp. 355-399). Amsterdam: Elsevier.
- Ferrand, L. (2001). *Cognition et lecture. Processus de base de la reconnaissance des mots écrits chez l'adulte*. Bruxelles: DeBoeck Université.
- Fischler, I., & Bloom, P.A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behaviour*, *18*, 1-20.
- Flores d'Arcais, G.B. (1981). The acquisition of meaning of the connectives. In W. Deutsch (Ed.), *The child's construction of language*. London: Academic Press.
- Fodor, J.A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: Bradford.
- Forster, K.I. (1976). Accessing the mental lexicon. In E.W. Walker (Ed.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North-Holland.
- Forster, K.I. (1979). Levels of processing and the structure of the language processor. In E.C.T. Walker (Ed.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Forster, K.I. (1981a). Frequency blocking and lexical access. *Journal of Verbal Learning and Verbal Behavior*, *20*, 190-203.
- Forster, K.I. (1981b). Priming and the effects of sentence and lexical contexts on naming time: evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology*, *33a*, 465-495.
- Forster, K.I. (1989). Basic issues in lexical processing. In N. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 75-107). Cambridge, MA: MIT University Press.
- Forster, K.I., & Chambers, S.M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*, 627-635.
- Forster, K.I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *10*, 680-698.
- Forster, K.I., & Olbrei, I. (1973). Semantic heuristic analysis. *Cognition*, *2*, 319-347.
- Forster, K.I., & Ryder, L.A. (1971). Perceiving the structure and meaning of sentences. *Journal of Verbal Learning and Verbal Behavior*, *10*, 285-296.
- Francis, W.N., & Kucera, H. (1982). *Computational analysis of english usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Friederici, A.D. (1983). Children's sensitivity to function words during sentence comprehension. *Linguistics*, *21*, 717-739.
- Friederici, A.D. (1985). Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition*, *19*, 133-166.
- Friederici, A.D. (1986). Autonomy and automaticity: Accessing function words during sentence comprehension. In E. Andreewsky (Ed.), *Perspectives in Cognitive Neuropsychology*. Hillsdale, N.J.: Lawrence Erlbaum.
- Friederici, A.D. (2002). Toward a neural basis of auditory sentence processing. *Trends in Cognitive Science*, *6*(2), 78-84.
- Friederici, A.D. (sous presse). The neural basis of sentence processing: A neurocognitive model. In D. Poeppel (Ed.), *Towards a New Functional Anatomy of Language. Special Issue, Cognition*.

- Friederici, A.D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: first-pass and second-pass parsing processes. *Journal of Psycholinguistic Research*, 25(1), 157-176.
- Friederici, A.D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1(3), 183-192.
- Friederici, A.D., & Schoenle, P.W. (1980). Computational dissociation of two vocabulary types: evidence from aphasia. *Neuropsychologia*, 18, 11-20.
- Frishkoff, G.A., & Tucker, D.M. (2001). *Anatomy of the N400: Brain Electrical Activity in Propositional Semantics* (No 01-1). The Institut of Cognitive and Decision Sciences.
- Fromkin, V.A. (1987). The lexicon: evidence from acquired dyslexia. *Language*, 63, 1-22.
- Froud, K. (2001). Prepositions and the lexical/functional divide: aphasic evidence. *Lingua*, 111, 1-28.
- Gardner, H., & Zurif, E.B. (1975). Bee but not be: Oral reading of single words in aphasia and alexia. *Neuropsychologia*, 13, 181-190.
- Garnsey, S. (1985). *Function words and content words: Reaction time and evoked potential measures of word recognition* (Cognitive Science Technical Report URCS-29). Unpublished doctoral dissertation, University of Rochester.
- Garrett, M.F. (1975). The analysis of sentence production. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 19). New York: Academic Press.
- Garrett, M.F. (1982). Syntactic processes in sentence production. In M. Garrett (Ed.), *Perspectives on mental representation*. Cambridge, MA: MIT Press.
- Garrett, M.F., & Kean, M.-L. (1980). Levels of representation and the analysis of speech errors. In M. Aronoff (Ed.), *Juncture*. San Francisco: AMNI Libri.
- Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, 113(2), 256-281.
- Geschwind, N. (1972). Language and the brain. *Scientific American*, 22(4), 76-83.
- Glanzer, M., & Ehrenreich, S.L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 18, 381-398.
- Gleitman, L., & Wanner, E. (1982). Language acquisition: the state of the state of the art. In L. Gleitman (Ed.), *Language acquisition: the state of the art* (pp. 1-50). Cambridge: University Press.
- Goodman, G.O., McClelland, J.L., & Gibbs, J.R.W. (1981). The role of syntactic context in word recognition. *Memory and Cognition*, 9(6), 580-586.
- Gordon, B. (1983). Lexical access and lexical decision: mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior*, 22, 24-44.
- Gordon, B., & Caramazza, A. (1982). Lexical decision for open- and closed-class words: failure to replicate differential frequency sensitivity. *Brain and Language*, 15(1), 143-160.
- Gordon, B., & Caramazza, A. (1983). Closed- and open-class lexical access in agrammatic and fluent aphasics. *Brain and Language*, 19, 335-345.
- Gordon, B., & Caramazza, A. (1985). Lexical access and frequency sensitivity: frequency saturation and open/closed class equivalence. *Cognition*, 21(2), 95-115.
- Gough, P.B. (1972). One second of reading. In I.G. Mattingly (Ed.), *Language by eye and by ear*. Cambridge, MA: MIT Press.
- Gough, P.B. (1983). Context, form, and interaction. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 203-211). New York: Academic Press.

- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, 29, 228-244.
- Greenhouse, W.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95-112.
- Grodzinski, Y. (1984). The syntactic characterization of agrammatism. *Cognition*, 16, 99-120.
- Gunter, T.C., Friederici, A.D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, 12(4), 556-568.
- Hagoort, P., & Brown, C.M. (2000). ERP effects of listening to speech: semantic ERP effects. *Neuropsychologia*, 38(11), 1518-1530.
- Halgren, E., Clarke, J.M., & Hervé, N. (1992). Rôle des enregistrements profonds dans la localisation, l'agencement séquentiel et la caractérisation des étapes de traitement de l'information dans le cerveau. *Psychologie Française*, 37(2), 149-166.
- Haveman, A. (1997). *The open-/closed-class distinction in spoken-word recognition*. Unpublished doctoral dissertation.
- Healy, A.F., & Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human, Perception and Performance*, 9, 413-426.
- Herron, D.T. (1998). *Context, word class, and prosody in the recognition of open- and closed-class words*. Unpublished doctoral dissertation, California.
- Herron, D.T., & Bates, E.A. (1997). Sentential and acoustic factors in the recognition of Open- and Closed-class words. *Journal of Memory and Language*, 37, 217-239.
- Hillyard, S.A. (1973). The CNV and human behavior: A review. *Electroencephalography and Clinical Neurophysiology Supplement*, 33, 161-171.
- Hillyard, S.A., Teder-Salejarvi, W.A., & Münte, T.F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, 8(2), 202-210.
- Höhle, B., & Weissenborn, J. (sous presse). German-learning ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*.
- Holcomb, P.J. (1993). Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology*, 30, 47-61.
- Holcomb, P.J., & Neville, H. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, 5, 281-312.
- Holcomb, P.J., & Neville, H.J. (1991). Natural speech processing: an analysis using event-related brain potentials. *Psychobiology*, 19, 286-300.
- Howes, D.H., & Solomon, R.L. (1951). Word frequency, personal values and visual duration thresholds. *Psychological Review*, 58, 256-270.
- Huynh, H., & Feldt, L.S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and splitplot designs. *Journal of Educational Statistics*, 1, 69-82.
- Irwin, D.A., Knott, J.R., McAdam, D.W., & Rebert, C.S. (1966). Motivational determinants of the "contingent negative variation". *Electroencephalography and Clinical Neurophysiology*, 21, 538-543.
- Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110, 306-340.
- Jennings, J.R. (1987). Editorial policy on analyses of variance with repeated measures. *Psychophysiology*, 24, 474-475.
- Johnson, B.W., & Hamm, J.P. (2000). High-density mapping in an N400 paradigm: evidence for bilateral temporal lobe generators. *Clinical and Neurophysiology*, 111(3), 532-545.

- Karniski, W., Vanderploeg, R., & Lease, L. (1993). "Virtual N400" and slow wave topography to auditory sentence incongruence. *Brain and Language*, *44*(1), 58-79.
- Kean, M.-L. (1979). Agrammatism: A phonological deficit? *Cognition*, *5*, 9-46.
- Keurs ter, M., Brown, C.M., & Hagoort, P. (2002). Lexical processing of vocabulary class in patients with Broca's aphasia: an event-related brain potential study on agrammatic comprehension. *Neuropsychologia*, *40*(9), 1547-1561.
- Keurs ter, M., Brown, C.M., Hagoort, P., & Stegeman, D.F. (1999). Electrophysiological manifestations of open- and closed-class words in patients with Broca's aphasia with agrammatic comprehension: An event-related brain potential study. *Brain*, *122*, 839-854.
- Khateb, A., Annoni, J.-M., Landis, T., Pegna, A., Custodi, M.-C., Fonteneau, E., Morand, S., & Michel, C.M. (1999). Spatio-temporal analysis of electric brain activity during semantic and phonological word processing. *International Journal of Psychophysiology*, *32*(3), 215-31.
- Khateb, A., Michel, C.M., Pegna, A.J., Thut, G., Landis, T., & Annoni, J.M. (2001). The time course of semantic category processing in the cerebral hemispheres: an electrophysiological study. *Cognitive Brain Research*, *10*(3), 251-264.
- King, J.W., & Kutas, M. (1998). Neural plasticity in the dynamic of human visual word recognition. *Neuroscience Letters*, *244*, 61-64.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance : a review and analysis. *Brain Research Review*, *29*, 169-195.
- Klosek, J. (1979). Two unargued linguistic assumptions in Kean's phonological interpretation of agrammatism. *Cognition*, *7*, 61-68.
- Koenig, T., & Lehmann, D. (1996). Microstates in language-related brain potential maps show noun-verb differences. *Brain and Language*, *53*(2), 169-182.
- Kolk, H.H., & Blomert, L. (1985). On the Bradley hypothesis concerning agrammatism: The nonword-interference effect. *Brain and Language*, *26*(1), 94-105.
- Koriat, A., Greenberg, S.N., & Goldshmid, Y. (1991). The missing-letter effect in hebrew: word frequency or word function? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17*(1), 66-80.
- Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *4*(12), 463.
- Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.
- Kutas, M., & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, *11*(5), 539-550.
- Kutas, M., & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161-163.
- Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as an index of stimulus evaluation time. *Science*, *197*, 792-795.
- Kutas, M., & van Petten, C. (1988). Event-related brain potential studies of language. In M.G.H. Coles (Ed.), *Advances in Psychophysiology* (Vol. 3, pp. 139-187). Greenwich: CT: JAI Press.
- Kutas, M., van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology*, *69*(3), 218-233.
- Kutas, M., & van Petten, C.K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. (Vol. 4, pp. 83-143). San Diego: Academic Press.



- LaBerge, D. (1995). *Attentional processing: the brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
- Lehmann, D. (1971). Multichannel topography of human alpha EEG fields. *Electroencephalography and Clinical Neurophysiology*, *31*, 439-449.
- Lehmann, D. (1972). Human scalp EEG fields: Evoked, alpha, sleep and spike-wave patterns. In M.A.B. Brazier (Ed.), *Synchronization of EEG activity in epilepsies* (pp. 301-325). Berlin: Springer.
- Lehmann, D. (1987). Principles of spatial analysis. In A. Remond (Ed.), *Handbook of Electroencephalography and Clinical Neurophysiology (Vol 1: Methods of analysis of brain electrical and magnetic signals)* (pp. 309-354). Amsterdam: Elsevier.
- Lehmann, D., Ozaki, H., & Pal, I. (1987). EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and Clinical Neurophysiology*, *67*(3), 271-288.
- Lehmann, D., & Skrandeis, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609-621.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M., Praamstra, P., Meyer, A.S., Helenius, P., & Salmelin, R. (1998). An MEG Study of Picture Naming. *Journal of Cognitive Neuroscience*, *10*(5).
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-75.
- Lorente de No, R. (1947). Action potential of the motoneurons of the hypoglossus nucleus. *Journal of Cellular and Comparative Physiology*, *29*, 207-287.
- Luck, S.J., & Ford, M.A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences*, *95*(3), 825-830.
- Marshall, J.C., & Newcombe, F. (1966). Syntactic and semantic errors in paralexia. *Neuropsychologia*, *4*, 169-176.
- Matthei, E.H., & Kean, M.L. (1989). Postaccess processes in the open vs. closed class distinction. *Brain and Language*, *36*(2), 163-180.
- McAdam, D.W., Irwin, D.A., Rebert, C.S., & Knott, J.R. (1966). Conative control of the contingent negative variation. *Electroencephalography and Clinical Neurophysiology*, *21*, 194-195.
- McCallum, W.C., & Papakostopoulos, D. (1973). The CNV and reaction time in situations of increasing complexity. *Electroencephalography and Clinical Neurophysiology Supplement*, *33*, 179-185.
- McCarthy, G., & Wood, C.C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203-208.
- McClelland, J.L., & Rumelhart, D.E. (1982). An interactive activation model of context effects in letter perception, Part 1: An account of basic findings. *Psychological Review*, *88*, 375-405.
- McMahon, A. (1994). *Understanding language change*. Cambridge: University Press.
- Meyer, D.E., Osman, A.M., Irwin, D.E., & Yantis, S. (1988). Modern mental chronometry. *Biological Psychology*, *26*, 3-67.
- Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227-334.
- Michel, C.M. (1997). *Dynamic functional brain mapping: Spatio-temporal analysis of the electric activity of the human brain*. Thèse d'habilitation, Faculté de Médecine.

- Michel, C.M., Henggeler, B., & Lehmann, D. (1992). 42-channel potential map series to visual contrast and stereo stimuli: perceptual and cognitive event-related segments. *Internal Journal of Psychophysiology*, *12*(2), 133-145.
- Michel, C.M., & Lehmann, D. (1993). Single doses of piracetam affect the 42-channel event-related potential microstates maps in a cognitive paradigm. *Neuropsychobiology*, *28*(4), 212-221.
- Michel, C.M., Seeck, M., & Landis, T. (1999). Spatiotemporal dynamics of human cognition. *News in Physiological Sciences*, *14*(5), 206-214.
- Mohr, B., Pulvermuller, F., & Zaidel, E. (1994). Lexical decision after left, right and bilateral presentation of function words, content words and non-words: evidence for interhemispheric interaction. *Neuropsychologia*, *32*(1), 105-124.
- Monsell (1985). Repetition and the lexicon. In W.A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2, pp. 147-195). London: Erlbaum.
- Morand, S., Thut, G., de Peralta, R.G., Clarke, S., Khateb, A., Landis, T., & Michel, C.M. (2000). Electrophysiological evidence for fast visual processing through the human koniocellular pathway when stimuli move. *Cerebral Cortex*, *10*(8), 817-825.
- Morgan, J.L., Shi, R., & Alopenna, P. (1996). Perceptual basis of rudimentary grammatical categories: toward a broader conceptualization of bootstrapping. In K. Demuth (Ed.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 263-283). Mahwah, NJ: Erlbaum.
- Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, *76*, 165-178.
- Morton, J. (1979). Some experiments of facilitation in word and picture recognition and their relevance for the evolution of a theoretical position. In H. Bouma (Ed.), *The processing of visual language*. New York: Plenum Press.
- Morton, J., & Patterson, K. (1980). Little words - No! In J.C. Marshall (Ed.), *Deep dyslexia*. London: Routledge & Kegan Paul.
- Munte, T.F., Wieringa, B.M., Weyerts, H., Szentkuti, A., Matzke, M., & Johannes, S. (2001). Differences in brain potentials to open and closed class words: class and frequency effects. *Neuropsychologia*, *39*(1), 91-102.
- Neeleman, A., & Weerman, F. (1999). *Flexible syntax: A theory of case and arguments*. Dordrecht: Kluwer.
- Nespoulous, J.-L., Dordain, M., Perron, C., Ska, B., Dub, D., Caplan, D., Mehler, J., & Lecours, A.R. (1988). Agrammatism in sentence production without comprehension deficits: Reduced availability of syntactic structures and/or grammatical morphemes? A case study. *Brain and Language*, *33*, 273-295.
- Neville, H.J., Coffey, S.A., Holcomb, P.J., & Tallal, P. (1993). The neurobiology of sensory and language processing in language-impaired children. *Journal of Cognitive Neuroscience*, *5*(2), 235-253.
- Neville, H.J., & Mills, D.L. (1997). Epigenesis of language. *Mental Retardation and Developmental Disabilities Research Reviews*, *3*, 282-292.
- Neville, H.J., Mills, D.L., & Lawson, D.S. (1992). Fractionating language: different neural subsystems with different sensitive periods. *Cerebral Cortex*, *2*(3), 244-258.
- Nobre, A.C., & McCarthy, G. (1994). Language-Related ERPs: Scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience*, *6*(3), 233-255.
- Norris, D.G. (1984). The effects of frequency, repetition and stimulus quality in visual word recognition. *Quarterly Journal of Experimental Psychology*, *36A*, 50-518.
- Nunez, P.L. (1981). *Electrical fields of the brain*. New York: Oxford University Press.

- Nunez, P.L., Pilgreen, K.L., Westdorp, A.F., Law, S.K., & Nelson, A.V. (1991). A visual study of surface potentials and laplacians due to distributed neocortical sources: Computer simulations and evoked potentials. *Brain Topography*, 4(2), 151-168.
- Offner, F.F. (1950). The EEG as potential mapping: The value of the average monopolar reference. *Electroencephalography and Clinical Neurophysiology*, 2, 215-216.
- Oldfield, R.C. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology*, 18, 340-353.
- Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Osterhout, L. (1997). On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain and Language*, 59(3), 494-522.
- Osterhout, L., Allen, M., & McLaughlin, J. (2002). Words in the brain: lexical determinants of word-induced brain activity. *Journal of Neurolinguistics*, 15, 171-187.
- Osterhout, L., & Bersick, M. (1999). Words - sentences = ? *Behavioral and Brain Sciences*, 22, 298-299.
- Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: word length and frequency predict the latency of an early negativity. *Biological Psychology*, 46(2), 143-168.
- Osterhout, L., & Holcomb, P.J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785-806.
- Osterhout, L., & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739-773.
- Palix, J. (2000). *Apport des potentiels évoqués dans la distinction entre deux classes de mots: les mots de classe ouverte et le mots de classe fermée*. Mémoire de DES de psychologie cognitive et expérimentale.
- Paller, K.A., Kutas, M., Shimamura, A.P., & Squires, L.R. (1987). Brain responses to concrete and abstract words reflect processes that correlates with later performance on test of recall and stemhyphen completion priming. *Electroencephalography and Clinical Neurophysiology, Suppl. 40*, 360-365.
- Pascual-Marqui, R.D., Michel, C.M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Transaction on Biomedical Engineering*, 42(7), 658-665.
- Pechmann, T., & Zerbst, D. (2002). The activation of word class information during speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(1), 233-243.
- Pegna, A., Khateb, A., Spinelli, L., Seeck, M., Landis, T., & Michel, C.M. (1997). Unravelling the cerebral dynamics of mental imagery. *Human Brain Mapping*, 5(1), 1-12.
- Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., & Taylor, M.J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, 37(2), 127-152.
- Picton, T.W., & Stuss, D.T. (1980). The component structure of the human event-related potentials. In L. Deecke (Ed.), *Motivation, Motor and Sensory Processes of the Brain: Electrical Potentials, Behavior and Clinical Use*. (Vol. 54, pp. 17-49). Amsterdam, New York, Oxford: Elsevier.
- Pinker, S. (1999). *L'instinct du langage*. Paris: Odile Jacob.
- Pollack, I., & Pickett, J. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior*, 3, 79-84.

- Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Electrocortical distinction of vocabulary types. *Electroencephalography and Clinical Neurophysiology*, 94(5), 357-370.
- Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex*, 9(5), 497-506.
- Pulvermüller, F., Mohr, B., & Schleicher, H. (1999). Semantic or lexico-syntactic factors: what determines word-class specific activity in the human brain? *Neuroscience Letters*, 275(2), 81-84.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J.A. (1972). *A Grammar on Contemporary English*. London: Harlow: Longman.
- Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell, B.
- Ragot, R.A., & Remond, A. (1978). EEG field mapping. *Electroencephalography and Clinical Neurophysiology*, 45, 417-421.
- Rao, K.V., & Proctor, R.W. (1984). Study-phase processing and the word frequency effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 386-394.
- Rapp, B., & Caramazza, A. (1993). On the distinction between deficits of access and deficits of storage: a question of theory. *Cognitive Neuropsychology*, 10, 113-141.
- Rapp, B., & Caramazza, A. (1997). The modality-specific organization of grammatical categories: Evidence from impaired spoken and written sentence production. *Brain and Language*, 56(2), 248-286.
- Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory and Cognition*, 9, 443-448.
- Rosenberg, B., Zurif, E.B., Brownell, H., Garrett, M., & Bradley, D.C. (1985). Grammatical class effects in relation to normal and aphasic sentence processing. *Brain and Language*, 26, 179-211.
- Rösler, F., Heil, M., & Glowalla, U. (1993). Monitoring retrieval from long-term memory by slow event-related brain potentials. *Psychophysiology*, 30, 170-182.
- Rösler, F., Streb, J., & Haan, H. (2001). Event-related brain potentials evoked by verbs and nouns in a primed lexical decision task. *Psychophysiology*, 38(4), 694-703.
- Rubenstein, H., Garfield, L., & Millikan, J.A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 9, 487-494.
- Rubin, D.C. (1980). 51 properties of 125 words: A unit analysis of verbal behavior. *Journal of Verbal Learning and Verbal Behavior*, 19, 736-755.
- Rugg, M.D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory and Cognition*, 18(4), 367-379.
- Rugg, M.D., & Coles, M.G.H. (1995). The ERP and cognitive psychology: Conceptual issues. In M.G.H. Coles (Ed.), *Electrophysiology of mind: Event-related brain potentials and cognition*. (pp. 27-39). Oxford: University Press.
- Saint-Aubin, J., & Poirier, M. (1997). The influence of word function in the missing-letter effect: Further evidence from French. *Memory and Cognition*, 25(5), 665-676.
- Scarborough, D.L., Cortese, C., & Scarborough, H.J. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human, Perception and Performance*, 3, 1-17.
- Scarborough, D.L., Gerard, L., & Cortese, C. (1979). Accessing lexical memory: the transfer of word repetition effects across task and modality. *Memory and Cognition*, 7, 3-12.
- Schmauder, A. (1996). Ability to stand alone and processing of open-class and closed-class words: Isolation versus context. *Journal of Psycholinguistic Research*, 25(4), 443-481.

- Schmauder, A.R., Morris, R.K., & Poynor, D.V. (2000). Lexical processing and text integration of function and content words: Evidence from priming and eye fixations. *Memory and Cognition*, 28(7), 1098-1108.
- Schmidt, A.L., Arthur, D.L., Kutas, M., & Flynn, E. (1989). Neuromagnetic responses during reading meaningful and nonmeaningful sentences. *Psychophysiology*, 26, S6.
- Schwanenflugel, P.J., & Shoben, E.J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming word. *Journal of Memory and Language*, 24, 232-252.
- Segalowitz, S.J., & Lane, K.C. (2000). Lexical access of function versus content words. *Brain and Language*, 75(3), 376-389.
- Segui, J. (1989). L'accès au lexique: Théories et données. In Calliope (Ed.), *La parole et son traitement automatique*. coll. Technique et Scientifique, CNET/ENST, Paris: Masson.
- Segui, J., Frauenfelder, U.H., Lainé, C., & Mehler, J. (1987). The word frequency effect for open- and closed-class items. *Cognitive Neuropsychology*, 4(1), 33-44.
- Segui, J., Mehler, J., Frauenfelder, U.H., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia*, 20(6), 615-627.
- Seidenberg, M.S., Waters, G.S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory and Cognition*, 12(4), 315-328.
- Shapiro, K.A., & Jensen, L.R. (1986). Processing open and closed class-headed nonwords: left hemisphere support for separate vocabularies. *Brain and Language*, 28, 318-327.
- Shi, R., Morgan, J.L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, 25(1), 169-201.
- Shi, R., & Werker, J.F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, 12(1), 70-75.
- Shi, R., Werker, J.F., & Morgan, J.L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2), B11-B21.
- Shillcock, R.C., & Bard, E.G. (1993). Modularity and the processing of closed-class words. In R.C. Shillcock (Ed.), *Cognitive models of speech processing: The second spurlong meeting* (Vol. 9, pp. 163-185). Hove: UK: Erlbaum.
- Shillcock, R.C., Bard, E.G., & Spensley, F. (1988). *Some prosodic effects on human word recognition in continuous speech*. Paper presented at the Speech '88, 7th FASE Symposium, Edinburg, UK.
- Skrandeis, W. (1986). Visual evoked potential topography: Methods and results. In F.H. Duffy (Ed.), *Topographic mapping of brain electrical activity* (pp. 1-30). Boston: Butterworths.
- Society, A.E. (1994). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 11, 111-113.
- Sorensen, J.M., Cooper, W.E., & Paccia, J.M. (1978). Speech timing of grammatical categories. *Cognition*, 6, 135-153.
- Steinhauer, K., & Ullman, M. (2002). Consecutive ERP effect of morpho-phonology and morpho-syntax. *Brain and Language*, 83(1), 62-65.
- Stemberger, J.P. (1984). Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology*, 1(4), 281-313.
- Strömqvist, S., Ragnarsdottir, H., & Richthoff, U. (1984). Input and production of function words. In B. Höhle (Ed.), *Approaches to bootstrapping: phonological, lexical, syntactic and neurophysiological aspects of early language acquisition*. (Vol. 2, pp. 155-177). Amsterdam: Benjamins.
- Sutton, S., Braren, M., Zubin, J., & John, E.R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, 150, 1187-1188.

- Swinney, D.A. (1982). Lexical processing during sentence comprehension: Effects of higher order constraints and implications for representation. In J. Anderson (Ed.), *The cognitive representation of speech*. Amsterdam: North-Holland.
- Swinney, D.A., Zurif, E.B., & Cutler, A. (1980). Effects of sentential stress and word class upon comprehension in Broca's aphasics. *Brain and Language*, *10*(1), 132-144.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, *7*, 263-272.
- Taft, M., & Forster, K.I. (1976). Lexical storage and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, *15*, 607-620.
- Tanenhaus, M.K., Garnsey, S.M., & Boland, J.E. (1990). Combinatory lexical information and language comprehension. In G.M.T. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 383-408). Cambridge, MA: MIT Press.
- Tanenhaus, M.K., & Lucas, M.M. (1987). Context effects in lexical processing. Special Issue: Spoken word recognition. *Cognition*, *25*, 213-234.
- Thibadeau, R., Just, M.A., & Carpenter, P.A. (1983). A model of the time course and content of reading. *Cognitive Science*, *6*, 157-203.
- Thierry, G. (2000). *Exploration des processus phonologiques, lexicaux et sémantiques: De compréhension du langage oral en potentiels évoqués et en imagerie par résonance magnétique fonctionnelle*. Unpublished doctoral dissertation, Université de Toulouse II.
- Thut, G., Hauert, C., Viviani, P., Morand, S., Spinelli, L., Blanke, O., Landis, T., & Michel, C. (2000). Internally driven vs. externally cued movement selection: A study on the timing of brain activity. *Cognitive Brain Research*, *9*(3), 261-269.
- Thut, G., Hauert, C.-A., Morand, S., Seeck, M., Landis, T., & Michel, C. (1999). Evidence for interhemispheric motor-level transfer in a simple reaction time task: An EEG-study. *Experimental Brain Research*, *128*(1-2), 256-261.
- Tucker, D.M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology*, *87*(3), 154-163.
- van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Timecourse of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394-417.
- van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition*, *18*, 380-393.
- van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition*, *19*(1), 95-112.
- Vogel, E.K., & Luck, S.J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*(2), 190-203.
- Winer, B.J. (1971). *Statistical principles in experimental design*. New-York: McGraw-Hill.
- Winnick, W.A., & Daniel, S.A. (1970). Two kinds of response priming in tachistoscopic word recognition. *Journal of Experimental Psychology*, *84*, 74-81.
- Zurif, E.B., & Caramazza, A. (1976). Psycholinguistic structures in aphasia: Studies in syntax and semantics. In H.A. Whitaker (Ed.), *Studies in Neurolinguistics*. New York: Academic Press.

**ANNEXES**

---

**ANNEXE 1:** Phrases du chapitre 3

**ANNEXE 2:** Stimuli du chapitre 4

**ANNEXE 3:** Stimuli des chapitres 5 et 6

**ANNEXE 4:** Phrases du chapitre 5

**ANNEXE 5:** Ordre des stimuli du chapitre 6

## - ANNEXE 1 -

## Phrases du chapitre 3

Chaque phrase est numérotée (de 1 à 80) avec la lettre "b" indiquant les phrases contrôles et "m" les phrases avec la violation sémantique. Chaque participant ne voyait une phrase que dans une condition seule condition expérimentale. Deux listes ont été créées: la première est constituée des phrases contrôle numérotées 1 à 40 (b) ainsi que des phrases présentant les violations sémantiques numérotées de 41 à 80 (m). La deuxième liste est constituée du reste des phrases (41-80 b et 1-40 m).

- 1 b A New York aucun pianiste n'écouterait le concert.
- 1 m A New York aucun fromage n'écouterait le concert.
- 2 b Après la pluie aucun cinéaste ne tournait un nouveau film.
- 2 m Après la pluie aucun maillot ne tournait un nouveau film.
- 3 b Après le travail la secrétaire payera la facture.
- 3 m Après le travail la plage payera la facture.
- 4 b Au dix-huitième siècle tous les employés signaient la pétition.
- 4 m Au dix-huitième siècle toutes les fourchettes signaient la pétition.
- 5 b Au travail les techniciens montaient les escaliers.
- 5 m Au travail les lunettes montaient les escaliers.
- 6 b Aujourd'hui cet aveugle voyage en première classe.
- 6 m Aujourd'hui cet oreille voyage en première classe.
- 7 b Aujourd'hui ma voisine achète la voiture.
- 7 m Aujourd'hui ma chatte achète la voiture.
- 8 b Auparavant ce comptable prenait la voiture.
- 8 m Auparavant ce matériel prenait la voiture.
- 9 b Auparavant son copain invitait ses amis.
- 9 m Auparavant son œil invitait ses amies.
- 10 b Autour de nous peu d'hommes construisaient une maison.
- 10 m Autour de nous peu d'ombres construisaient une maison.
- 11 b Autour du comptoir tous les responsables payaient le pourboire.
- 11 m Autour du comptoir tous les verres payaient le pourboire.
- 12 b Autrefois cette réceptionniste envoyait les lettres.
- 12 m Autrefois cette bougie envoyait les lettres.
- 13 b Bientôt de nombreux ouvriers repeindront ces murs.
- 13 m Bientôt de nombreux champignons repeindront ces murs.
- 14 b Bientôt la plupart des entrepreneurs fumeront la pipe.
- 14 m Bientôt la plupart des échecs fumeront la pipe.
- 15 b Chez elle la maîtresse corrige les exercices.
- 15 m Chez elle la banquette corrige les exercices.
- 16 b Chez lui son gendre boit du café.
- 16 m Chez lui son art boit du café.
- 17 b Dans la classe les élèves parlent français.
- 17 m Dans la classe les tables parlent français.
- 18 b Dans la forêt quelques bûcherons coupent du bois.
- 18 m Dans la forêt quelques feuilles coupent du bois.
- 19 b Dans quelques heures beaucoup d'invités mangeront un gâteau.



- 19 m Dans quelques heures beaucoup de pantalons mangeront un gâteau.
- 20 b Dans trois jours cet acteur quittera la scène.
- 20 m Dans trois jours ce coude quittera la scène.
- 21 b Déjà cette femme va à la plage.
- 21 m Déjà cette inclinaison va à la plage.
- 22 b Demain midi quelques grévistes reprendront leur travail.
- 22 m Demain midi quelques porches reprendront leur travail.
- 23 b Demain tous les gardiens regarderont la télévision.
- 23 m Demain tous les camions regarderont la télévision.
- 24 b Depuis hier son fils réparait les ordinateurs.
- 24 m Depuis hier son agenda réparait les ordinateurs.
- 25 b Depuis un mois trois amis revenaient à Genève.
- 25 m Depuis un mois trois rames revenaient à Genève.
- 26 b Derrière la chaise le chat mangera la souris.
- 26 m Derrière la chaise le sel mangera la souris.
- 27 b Dès demain soir cet étudiant lira le livre.
- 27 m Dès demain soir cet accordéon lira le livre.
- 28 b Dès vendredi prochain cet homme prendra le bus.
- 28 m Dès vendredi prochain cet œuf prendra le bus.
- 29 b Désormais beaucoup d'étudiants découperont du papier.
- 29 m Désormais beaucoup de fronts découperont du papier.
- 30 b Désormais près de la fenêtre le client déjeune tout seul.
- 30 m Désormais près de la fenêtre le planché déjeune tout seul.
- 31 b Devant les spectateurs le magicien donne ses secrets.
- 31 m Devant les spectateurs le plafond donne ses secrets.
- 32 b Devant nous les poules mangent des graines.
- 32 m Devant nous les conclusions mangent des graines.
- 33 b Devant ses amis la mariée portait une robe longue.
- 33 m Devant ses amis la corvée portait une robe longue.
- 34 b Dorénavant chaque fillette pensera à son devoir.
- 34 m Dorénavant chaque élection pensera à son devoir.
- 35 b Durant la soirée les trois femmes dansaient le rock.
- 35 m Durant la soirée les trois montres dansaient le rock.
- 36 b Durant les vacances quelques voyageurs attendront le bus.
- 36 m Durant les vacances quelques disques attendront le bus.
- 37 b En France les enfants partent en vacances.
- 37 m En France les cartons partent en vacances.
- 38 b Environ sous le pont ce conducteur percutera une autre voiture.
- 38 m Environ sous le pont cet arbre percutera une autre voiture.
- 39 b Face à la presse tous les libraires répondent aux questions.
- 39 m Face à la presse tous les postes répondent aux questions.
- 40 b Hier après-midi quelques photographes recevaient une récompense.
- 40 m Hier après-midi quelques classeurs recevaient une récompense.
- 41 b Hier soir de nombreux diplomates buvaient du champagne.
- 41 m Hier soir de nombreux crayons buvaient du champagne.
- 42 b Hormis Hélène toutes les fillettes pleuraient beaucoup.
- 42 m Hormis Hélène toutes les fuites pleuraient beaucoup.
- 43 b Jadis de nombreux artisans tapissaient le salon.
- 43 m Jadis de nombreuses diligences tapissaient le salon.
- 44 b Jadis le Père-Noël offrait des cadeaux.

- 44 m Jadis le poil offrait des cadeaux.  
45 b Jadis tes oiseaux chantaient dans le ciel.  
45 m Jadis tes cordons chantaient dans le ciel.  
46 b Jamais le client ne donnait des idées.  
46 m Jamais le cartable ne donnait des idées.  
47 b Là beaucoup d'enfants jouent au ballon.  
47 m Là beaucoup de brancards jouent au ballon.  
48 b Maintenant la plupart des opposants organiseront une grève.  
48 m Maintenant la plupart des ballons organiseront une grève.  
49 b Maintenant trois locataires louent un appartement.  
49 m Maintenant trois cigares louent un appartement.  
50 b Néanmoins chaque architecte présentait ses plans.  
50 m Néanmoins chaque épingle présentait ses plans.  
51 b Néanmoins mon psychiatre conservera sa théorie.  
51 m Néanmoins mon addition conservera sa théorie.  
52 b Par la fenêtre le poète regardera les étoiles.  
52 m Par la fenêtre le malheur regardera les étoiles.  
53 b Par le passé peu de travailleurs déterminaient les conditions.  
53 m Par le passé peu de rouleaux déterminaient les conditions.  
54 b Parfois chaque enfant achetait du pain.  
54 m Parfois chaque alarme achetait du pain.  
55 b Parfois trois propriétaires raconteront leurs histoires.  
55 m Parfois trois plats raconteront leurs histoires.  
56 b Parfois un professeur finissait son cours.  
56 m Parfois un bol finissait son cours.  
57 b Parmi les jeunes une demoiselle dépense toutes ses économies.  
57 m Parmi les jeunes une urgence dépense toutes ses économies.  
58 b Partout les pharmaciens vendaient des médicaments.  
58 m Partout les saluts vendaient des médicaments.  
59 b Pendant le week-end prochain de nombreux standardistes raccrocheront le téléphone.  
59 m Pendant le week-end prochain de nombreux ciseaux raccrocheront le téléphone.  
60 b Pendant ses cours cet étudiant passe des examens.  
60 m Pendant ses cours cet ordinateur passe des examens.  
61 b Plus tard peu de reporters commenteront un discours.  
61 m Plus tard peu d'armoires commenteront un discours.  
62 b Pour travailler ces salariés portent un pantalon.  
62 m Pour travailler ces mouchoirs portent un pantalon.  
63 b Près des maisons tous les chiens aboieront dans la cour.  
63 m Près des maisons tous les prêts aboieront dans la cour.  
64 b Presque toujours la plupart des étrangers oublient leurs papiers.  
64 m Presque toujours la plupart des valises oublient leurs papiers.  
65 b Presque tous les pilotes décollent avec leur avion.  
65 m Presque tous les filets décollent avec leur avion.  
66 b Prochainement le voyageur visitera Paris.  
66 m Prochainement le téléphone visitera Paris.  
67 b Rapidement tous les vendeurs annonçaient les promotions.  
67 m Rapidement toutes les portes annonçaient les promotions.  
68 b Régulièrement cet individu tachait sa cravate.  
68 m Régulièrement cet incon vénient tachait sa cravate.  
69 b Sauf exception aucune boucherie vendra du beurre.

- 69 m Sauf exception aucune anémone ne vendra du beurre.  
70 b Selon eux chaque spectateur louait sa place.  
70 m Selon eux chaque éclipse louait sa place.  
71 b Selon lui beaucoup d'alpinistes achètent du matériel.  
71 m Selon lui beaucoup de fraises achètent du matériel.  
72 b Simultanément la secrétaire tapait une lettre.  
72 m Simultanément la synthèse tapait une lettre.  
73 b Sous son bureau le directeur posera un roman.  
73 m Sous son bureau le placard posera un roman.  
74 b Souvent aucune fille ne partait en vacances.  
74 m Souvent aucun arbre ne partait en vacances.  
75 b Souvent le clown allait sur scène.  
75 m Souvent le sable allait sur scène.  
76 b Sur le tableau les enfants dessinent avec les craies.  
76 m Sur le tableau les oursins dessinent avec les craies.  
77 b Sur son lit ce garçon collait des timbres.  
77 m Sur son lit cet espoir collait des timbres.  
78 b Toujours cet adolescent répare sa bicyclette.  
78 m Toujours cet acide répare sa bicyclette.  
79 b Toutefois la plupart des livreurs ramenaient les meubles.  
79 m Toutefois la plupart des chaussettes ramenaient les meubles.  
80 b Toutefois peu de concierges éteindront la lumière.  
80 m Toutefois peu de citrons éteindront la lumière.

- ANNEXE 2 -  
Stimuli du chapitre 4

## Classe ouverte

Verbes	Présentation	Fréquence/1 mns	Log de Fréquence	Nombre de lettres
aboieront	1	7.48	0.87	9
achetait	1	61.73	1.79	8
achète	1	61.73	1.79	6
achètent	1	61.73	1.79	8
allait	1	2163.41	3.34	6
annonçaient	1	94.44	1.98	11
attendront	1	494.60	2.69	10
boit	1	122.69	2.09	4
buvaient	1	122.69	2.09	8
chantaient	1	112.18	2.05	10
collait	1	20.42	1.31	7
commenteront	1	9.74	0.99	12
conservera	1	57.22	1.76	10
construisaient	1	37.52	1.57	14
corrige	1	19.56	1.29	7
coupent	1	60.66	1.78	7
dansaient	1	59.00	1.77	9
décollent	1	4.55	0.66	9
découperont	1	10.42	1.02	11
déjeune	1	83.76	1.92	7
dépense	1	16.50	1.22	7
dessinent	1	38.20	1.58	9
déterminaient	1	36.88	1.57	13
donnait	1	965.90	2.98	7
donne	1	965.90	2.98	5
écouterà	1	251.94	2.40	8
envoyait	1	118.48	2.07	8
éteindront	1	33.48	1.52	10
finissait	1	194.59	2.29	9
fumeront	1	41.60	1.62	8
invitait	1	57.09	1.76	8
jouent	1	268.83	2.43	6
lira	1	0.59	-0.23	4
louait	1	25.56	1.41	6
louent	1	25.56	1.41	6
mangent	1	166.68	2.22	7
mangera	1	166.68	2.22	7
mangeront	1	166.68	2.22	9
montaient	1	253.21	2.40	9
offrait	1	168.72	2.23	7
organiseront	1	44.88	1.65	12
oublie	1	174.59	2.24	8
parlent	1	987.68	2.99	7
partait	1	344.51	2.54	7
partent	1	344.51	2.54	7
passé	1	728.38	2.86	5

payaient	1	97.72	1.99	8
payera	1	97.72	1.99	6
pensera	1	833.97	2.92	7
percutera	1	0.34	-0.47	9
pleuraient	1	107.33	2.03	10
portait	1	380.67	2.58	7
portent	1	380.67	2.58	7
posera	1	192.89	2.29	6
prenait	1	1029.63	3.01	7
prendra	1	1029.63	3.01	7
présentait	1	153.75	2.19	10
quittera	1	196.97	2.29	8
raccrocheront	1	7.53	0.88	13
raconteront	1	139.62	2.14	11
ramenaient	1	89.04	1.95	10
recevaient	1	221.48	2.35	10
regardera	1	654.86	2.82	9
regarderont	1	654.86	2.82	11
réparait	1	17.61	1.25	8
répare	1	17.61	1.25	6
repeindront	1	0.85	-0.07	11
répondent	1	445.89	2.65	9
reprendront	1	298.82	2.48	11
revenaient	1	402.92	2.61	10
signaient	1	28.37	1.45	9
tachait	1	2.12	0.33	7
tapait	1	23.73	1.38	6
tapissaient	1	1.91	0.28	11
tournait	1	240.88	2.38	8
va	1	2163.41	3.34	2
vendaient	1	49.98	1.70	9
vendra	1	49.98	1.70	6
visitera	1	29.05	1.46	8
voyage	1	19.65	1.29	6
<b>Total/Moyenne</b>	<b>80</b>	<b>253.93</b>	<b>1.91</b>	<b>8.19</b>

<b>Adverbes</b>	<b>Présentation</b>	<b>Fréquence/1 mns</b>	<b>Log de Fréquence</b>	<b>Nombre de lettres</b>
aujourd'hui	2	666.44	2.82	11
auparavant	2	33.90	1.53	10
autrefois	1	106.10	2.03	9
bientôt	2	160.13	2.20	7
déjà	1	708.17	2.85	4
demain	2	144.22	2.16	6
désormais	2	95.25	1.98	9
dorénavant	1	4.76	0.68	10
environ	1	25.56	1.41	7
hier	2	348.16	2.54	4
jadis	3	90.44	1.96	5
jamais	1	1105.10	3.04	6
là	1	1526.96	3.18	2
maintenant	2	502.64	2.70	10
néanmoins	2	41.94	1.62	9

parfois	3	269.46	2.43	7
partout	1	149.32	2.17	7
presque	2	484.05	2.68	7
prochainement	1	3.78	0.58	13
rapidement	1	43.18	1.64	10
régulièrement	1	15.69	1.20	13
simultanément	1	8.04	0.91	13
souvent	2	359.06	2.56	7
toujours	1	1158.19	3.06	8
toutefois	2	75.47	1.88	9
<b>Total/Moyenne</b>	<b>40</b>	<b>325.04</b>	<b>2.07</b>	<b>8.12</b>

Noms	Présentation	Fréquence/1 mns	Log de Fréquence	Nombre de lettres
acteur	1	40.28	1.61	6
adolescent	1	19.01	1.28	10
bûcheron	1	4.89	0.69	8
client	1	41.60	1.62	6
comptable	1	4.93	0.69	9
concierge	1	26.88	1.43	9
diplomate	1	9.65	0.98	9
directeur	1	49.05	1.69	9
élève	1	47.94	1.68	5
employé	1	50.24	1.70	7
enfant	3	612.92	2.79	6
entreprise	1	61.04	1.79	10
étudiant	2	28.54	1.46	8
femme	1	859.92	2.93	5
filles	1	439.21	2.64	5
fillette	1	15.01	1.18	8
garçon	1	185.06	2.27	6
gardien	1	24.12	1.38	7
invité	1	2.04	0.31	6
locataire	1	9.44	0.97	9
maîtresse	1	69.09	1.84	9
opposant	1	6.16	0.79	8
père-noël	1	505.96	2.70	9
pianiste	1	4.72	0.67	8
professeur	1	66.53	1.82	10
psychiatre	1	3.27	0.51	10
réceptionniste	1	11.01	1.04	13
responsable	1	38.33	1.58	11
salarié	1	1.78	0.25	7
secrétaire	1	37.35	1.57	10
spectateur	1	34.33	1.54	10
standardiste	1	0.21	-0.68	12
technicien	1	7.95	0.90	10
travailleur	1	22.54	1.35	11
voisine	1	12.30	1.09	7
voyageur	2	35.22	1.55	8
<b>Total/Moyenne</b>	<b>40</b>	<b>121.43</b>	<b>1.45</b>	<b>8.24</b>

**Classe fermée**

<b>Préposition</b>	<b>Présentation</b>	<b>Fréquence/1 mns</b>	<b>Log de Fréquence</b>	<b>Nombre de lettres</b>
a	1	25.31	1.40	1
après	2	900.85	2.95	5
au	2	5320.59	3.73	2
autour	2	322.69	2.51	6
chez	2	796.58	2.90	4
dans	4	8389.96	3.92	4
depuis	2	627.89	2.80	6
derrière	1	273.00	2.44	8
dès	2	398.97	2.60	3
devant	3	724.72	2.86	6
durant	2	65.68	1.82	6
en	1	11332.81	4.05	2
face	1	248.66	2.40	4
hormis	1	4.33	0.64	6
par	2	4853.00	3.69	3
parmi	1	204.25	2.31	5
pendant	2	416.36	2.62	7
plus	1	5875.48	3.77	4
pour	1	6417.15	3.81	4
près	1	429.51	2.63	4
sauf	1	67.09	1.83	4
selon	2	196.76	2.29	5
sous	1	901.83	2.96	4
sur	2	4751.70	3.68	3
<b>Total/Moyenne</b>	<b>40</b>	<b>2335.93</b>	<b>2.81</b>	<b>4.36</b>

<b>Quantificateurs</b>	<b>Présentation</b>	<b>Fréquence/1 mns</b>	<b>Log de Fréquence</b>	<b>Nombre de lettres</b>
aucun	2	290.65	2.46	5
aucune	2	290.65	2.46	6
beaucoup	4	549.57	2.74	8
chaque	4	507.41	2.71	6
nombreuses	1	77.98	1.89	10
nombreux	3	77.98	1.89	8
peu	4	1457.95	3.16	3
plupart	4	73.89	1.87	7
quelques	4	1003.46	3.00	8
tous	6	4177.32	3.62	4
toutes	2	4177.32	3.62	6
trois	4	543.10	2.73	5
<b>Total/Moyenne</b>	<b>40</b>	<b>1102.27</b>	<b>2.68</b>	<b>6.33</b>

<b>Non-Quantificateur</b>	<b>Présentation</b>	<b>Fréquence/1 mns</b>	<b>Log de Fréquence</b>	<b>Nombre de lettres</b>
ce	2	12465.23	4.10	2
ces	1	12465.23	4.10	3
cet	8	708.94	2.85	3

cette	2	708.94	2.85	5
la	4	26626.08	4.43	2
le	9	20927.31	4.32	2
les	6	20927.31	4.32	3
ma	1	2084.41	3.32	2
mon	1	2084.41	3.32	3
son	3	4377.23	3.64	3
tes	1	467.42	2.67	3
un	1	13540.43	4.13	2
une	1	13540.43	4.13	3
<b>Total/Moyenne</b>	<b>40</b>	<b>10071.03</b>	<b>3.71</b>	<b>2.77</b>



- ANNEXE 3 -  
Stimuli des chapitres 5 et 6

## Mots de Classe Fermée

<b>mots</b>	<b>fréquence</b>	<b>fréquence/1mn</b>	<b>log de fréquence</b>	<b>Nombre de lettres</b>
afin	b	227	2.36	4
auprès	b	129	2.11	6
auquel	b	83	1.92	6
ayant	b	205	2.31	5
ceci	b	160	2.20	4
certes	b	115	2.06	6
chacun	b	235	2.37	6
ci	b	336	2.53	2
combien	b	135	2.13	7
dedans	b	56	1.75	6
dehors	b	155	2.19	6
dès	b	399	2.60	3
dessus	b	53	1.72	6
duquel	b	38	1.58	6
durant	b	66	1.82	6
grâce	b	218	2.34	5
guère	b	148	2.17	5
hélas	b	69	1.84	5
hors	b	134	2.13	4
jadis	b	90	1.96	5
lors	b	68	1.83	4
lorsque	b	347	2.54	7
mien	b	39	1.59	4
nôtre	b	53	1.72	5
nul	b	93	1.97	3
parmi	b	204	2.31	5
partout	b	149	2.17	7
puisque	b	282	2.45	7
quant	b	126	2.10	5
quel	b	389	2.59	4
quoique	b	55	1.74	7
selon	b	197	2.29	5
sien	b	34	1.53	4
sinon	b	95	1.98	5
soi	b	297	2.47	3
soudain	b	226	2.35	7

tantôt	b	89	1.95	6
tel	b	282	2.45	3
tôt	b	98	1.99	3
voici	b	179	2.25	5
ainsi	h	817	2.91	5
alors	h	941	2.97	5
après	h	901	2.95	5
aussi	h	1'440	3.16	5
avant	h	583	2.77	5
avec	h	3'849	3.59	4
beaucoup	h	550	2.74	8
ça	h	1'069	3.03	2
car	h	806	2.91	3
cela	h	1'138	3.06	4
celui	h	696	2.84	5
cet	h	709	2.85	3
chez	h	797	2.90	4
comme	h	4'375	3.64	5
contre	h	697	2.84	6
dans	h	8'390	3.92	4
déjà	h	708	2.85	4
depuis	h	628	2.80	6
devant	h	725	2.86	6
donc	h	703	2.85	4
dont	h	1'306	3.12	4
encore	h	1'684	3.23	6
ici	h	663	2.82	3
jamais	h	1'105	3.04	6
jusque	h	803	2.90	6
leur	h	1'790	3.25	4
même	h	3'663	3.56	4
moi	h	2'244	3.35	3
mon	h	2'084	3.32	3
notre	h	1'010	3.00	5
ou	h	2'600	3.42	2
puis	h	761	2.88	4
quand	h	1'397	3.15	5
quelque	h	1'003	3.00	7
si	h	3'425	3.53	2
sous	h	902	2.96	4
te	h	1'181	3.07	2
toujours	h	1'158	3.06	8
très	h	998	3.00	4
vers	h	956	2.98	4

## Mot de classe ouverte

mots	fréquence	fréquence/1mn	log de fréquence	Nombre de lettres
ail	b	69	1.84	4
ancien	b	218	2.34	6
assis	b	142	2.15	5
cas	b	286	2.46	3
chair	b	109	2.04	5
cheveu	b	136	2.13	6
confus	b	52	1.71	6
couple	b	37	1.57	6
cri	b	144	2.16	3
détail	b	90	1.96	6
élan	b	69	1.84	4
fièvre	b	53	1.72	6
fini	b	170	2.23	4
fond	b	381	2.58	4
garçon	b	185	2.27	6
garde	b	113	2.05	5
genre	b	126	2.10	5
juste	b	208	2.32	5
lettre	b	384	2.58	6
moyen	b	268	2.43	5
nation	b	93	1.97	6
neuf	b	134	2.13	4
nom	b	382	2.58	3
nord	b	100	2.00	4
notion	b	88	1.95	6
parti	b	231	2.36	5
propos	b	132	2.12	6
public	b	148	2.17	6
retour	b	151	2.18	6
rond	b	37	1.57	4
rue	b	333	2.52	3
saison	b	54	1.74	6
sœur	b	148	2.17	5
soie	b	34	1.53	4
sourire	b	288	2.46	7
suivre	b	291	2.46	6
total	b	82	1.91	5

ville	b	285	2.45	5
vin	b	90	1.95	3
vivant	b	204	2.31	6
aimer	h	766	2.88	5
air	h	584	2.77	3
autre	h	2'484	3.40	5
beau	h	741	2.87	4
bien	h	2'982	3.47	4
chose	h	1'390	3.14	5
coup	h	627	2.80	4
croire	h	1'147	3.06	6
dire	h	3'849	3.59	4
dit	h	761	2.88	3
faire	h	4'101	3.61	5
fait	h	1'584	3.20	4
femme	h	860	2.93	5
fois	h	1'001	3.00	4
heure	h	869	2.94	5
homme	h	1'991	3.30	5
main	h	892	2.95	4
monde	h	1'023	3.01	5
mort	h	772	2.89	4
mot	h	685	2.84	3
nouveau	h	818	2.91	7
parler	h	988	2.99	5
passer	h	728	2.86	6
pensée	h	554	2.74	6
penser	h	834	2.92	6
petit	h	1'476	3.17	5
peut	h	3'363	3.53	4
point	h	999	3.00	5
premier	h	966	2.98	7
prendre	h	1'023	3.01	7
savoir	h	1'987	3.30	6
sens	h	590	2.77	4
temps	h	1'215	3.08	5
tenir	h	710	2.85	5
tête	h	638	2.80	4
va	h	2'163	3.34	2
venir	h	1'297	3.11	5
vie	h	1'241	3.09	3
voir	h	1'998	3.30	4
voix	h	538	2.73	4

- ANNEXE 4 -  
Phrases du chapitre 5

**Classe Fermée**

1. Je te punirais **si** tu mens encore.
2. Tu rangeras tout **avant** les prochaines vacances.
3. Je fais tout **ou** ne fais rien.
4. Le vendredi treize **te** porte souvent bonheur.
5. Il verra ce **dont** je suis capable.
6. Tu apprends tout **ça** pour demain matin.
7. Je vais bien **car** tu es parti.
8. Elle est venue **vers** les onze heures.
9. Tous ces actes **leur** ont été interdits.
10. Je me réfugie **sous** un grand arbre.
11. Je vais souvent **chez** ma grande sœur.
12. Le pelage de **cet** âne est gris.
13. Il a abandonné **mon** vieux chien malade.
14. Il a rit **puis** rentra chez lui.
15. Je suis faible **donc** souvent fatigué.
16. Tu es toujours **très** près de moi.
17. Je suis heureux **quand** tu vas bien.
18. Tu rêves quand **moi** je me concentre.
19. Tu es distrait **comme** tous les autres.
20. Je prendrais le **même** train que toi.
21. Je cherche encore **celui** qui me guérira.
22. Ces filles sont **aussi** arrivées trop tôt.
23. Le passé est **toujours** présent en moi.
24. Elle ne boit **jamais** de vin blanc.
25. Il cherche encore **quelque** chose de mieux.
26. Nous souffrons car **notre** ventre est plein.
27. Je suis reposée **après** avoir bien dormi.
28. Elle pense toujours **ainsi** et pas autrement.
29. Tu voudrais bien **encore** gagner de l'argent.
30. Le chat attend **devant** son assiette vide.
31. Tu crois que **cela** ne vaut rien.
32. Ne joue pas **avec** cette méchante fille.
33. Je travaillerai dur **jusque** tard ce soir.
34. Elle se donne **beaucoup** plus de peine.
35. Il parle beaucoup **alors** je me tais.
36. Je mange encore **dans** mon jardin.
37. Elle était résolument **contre** ce parti politique.
38. Tu y penses **déjà** depuis bien longtemps.
39. Je suis ridée **depuis** bientôt dix ans.
40. On est mieux **ici** que par-là.
41. Tu le feras **dès** que tu pourras.
42. Il aimerait savoir **quel** est ton âge.
43. Tu te réveilles **lorsque** je rentre tard.
44. Prends cette chose **ci** et non l'autre.

45. On garde pour **soi** bien des secrets.
46. Tu es premier **grâce** à tes qualités.
47. Tu dois douter **puisque** tu en parles.
48. Il se battra **tel** le vieux lion.
49. On te suivra **partout** où tu iras.
50. Tu vois l'homme **duquel** il a hérité.
51. Après la fête **chacun** rentre chez soi.
52. Je me prive **afin** de rester mince.
53. On marchait et **soudain** il est tombé.
54. Elle mourût en **ayant** connu la richesse.
55. Les fleurs que **voici** sont pour vous.
56. Il faut choisir **selon** vos propres goûts.
57. Je trouve que **ceci** vous convient bien.
58. Il doit rester **dehors** par beau temps.
59. Tu en as **guère** plus que moi.
60. Elle sait bien **combien** ça me coûte.
61. Ma tante habite **hors** de la ville.
62. Elle est restée **auprès** de sa mère.
63. Je me culpabilise **quant** à sa situation.
64. Il sait que **certes** je m'en irai.
65. Ton frère est **parmi** les plus doués.
66. Je me lève **tôt** toute la semaine.
67. Tu manges tout **sinon** je me fâche.
68. Je pars mais **nul** ne sait où.
69. Les guerres de **jadis** étaient très longues.
70. Ton fils viendra **tantôt** te rendre visite.
71. Un grand progrès **auquel** je crois beaucoup.
72. On est venu **hélas** tu étais parti.
73. On a chanté **lors** de cette fête.
74. Il a plu **durant** de longs jours.
75. Je suis resté **dedans** car il neige.
76. Elle se marie **quoique** très peu motivée.
77. Il choisit le **nôtre** pour vous énerver.
78. Tu dois monter **dessus** pour mieux sauter.
79. Tu prends le **mien** sans me prévenir.
80. Après tout le **sien** était mieux avant.

## Classe Ouverte

81. Elle prend un **élan** important pour sauter.
82. Je mange une **aile** de poulet grillée.
83. Il entend un **cri** dans la nuit.
84. Elle passe la **saison** hivernale en station.
85. Une très grande **fièvre** peut être dangereuse.
86. Le voisin reste **confus** après ce vacarme.
87. Elle tourne en **rond** depuis ton départ.
88. Je préfère la **soie** rouge en dentelle.
89. Elle porte son **nouveau** bonnet de laine.
90. Il assume son **premier** échec en amour.
91. Tu as une **vie** banale et monotone.
92. Je déplore le **fait** qu'il soit sévère.
93. En courant elle **peut** battre le record.
94. Elle traverse ce **petit** bois sans crainte.
95. Elle entend une **voix** dans le lointain.
96. Il a toujours **dit** bonjour au monsieur.
97. Tu as un **point** d'avance sur moi.
98. Tu prends trois **fois** l'avion cette année.
99. Mon chat est **mort** de vieillesse récemment.
100. Il est une **heure** moins le quart.
101. Elle donne la **main** à sa maman.
102. Je peux te **faire** gagner un prix.
103. Il écrira un **mot** pour te remercier.
104. Nous allons donc **voir** les artistes jouer.
105. Ce vase est **beau** et pas cher.
106. Tu ignores le **sens** de cette expérience.
107. Elle prend le **temps** de se reposer.
108. Tu dois lui **dire** de faire attention.
109. Ma sœur est **bien** une artiste étonnante.
110. Elle veut un **homme** à sa taille.
111. Il faut vraiment **savoir** qui il est.
112. Il donna un **coup** de coude malencontreux.
113. Je connais cet **air** depuis mon enfance.
114. Il ne peut **aimer** que cette femme.
115. Elle a une **pensée** pour sa mère.
116. Tu dois te **tenir** à la barre.
117. Elle achète une **autre** robe que nous.
118. Il lave cette **chose** horrible et sale.
119. On va bientôt **venir** pour vous parler.
120. Elle va te **croire** sans aucun problème.
121. Le temps va **passer** plus vite désormais.
122. Avoir la grosse **tête** après une nomination.
123. Je ne peux **penser** à rien maintenant.
124. Il veut te **parler** de quelque chose.
125. Nous avons un **monde** qui nous sépare.
126. Nous allons ensuite **prendre** un café liégeois.
127. Depuis peu il **va** seul à l'école.

128. Elle est la **femme** de ma vie.
129. Il chante au **nom** de la patrie.
130. Elle observe du **fond** de son trou.
131. Il reçut ta **lettre** hier à midi.
132. Il habite la **rue** devant la gare.
133. Voilà un grave **cas** de grippe aiguë.
134. Son avis est **juste** mais très sévère.
135. Le garçon devra **suivre** la piste fléchée.
136. Elle va en **ville** acheter des bas.
137. Je trouve ce **public** calme et froid.
138. Il a un **sourire** qui me glace.
139. Sa vie de **couple** est un échec.
140. Tu connais le **moyen** d'ouvrir ce coffre.
141. Il soutient le **parti** d'un leader communiste.
142. Elle expose un **ancien** vase en terre.
143. Il est bien **vivant** grâce à toi.
144. Je vais avoir **fini** mes devoirs demain.
145. Voilà un beau **garçon** qui me plaît.
146. Il est resté **assis** tout le temps.
147. Je serai de **retour** dès ce soir.
148. Elle a une **sœur** qui étudie beaucoup.
149. Je vois un **cheveu** dans ma soupe.
150. Ce pull est **neuf** et déjà effiloché.
151. Ce sont des **propos** plutôt très controversés.
152. Tu es le **genre** d'homme à éviter.
153. Il monte la **garde** devant le château.
154. Le péché de **chair** est sévèrement puni.
155. Je parcours le **nord** de ce pays.
156. Tu admires cette **nation** qui mérite confiance.
157. Tu préfères le **vin** des cépages bordelais.
158. Ce tout petit **détail** m'a soudain échappé.
159. Tu perds la **notion** du mot travailler.
160. Nous avons au **total** une somme rondelette.



- ANNEXE 5 -  
Ordre des stimuli du chapitre 6

<b>ordre</b>	<b>stimuli</b>	<b>classe</b>	<b>ordre</b>	<b>stimuli</b>	<b>classe</b>	<b>ordre</b>	<b>stimuli</b>	<b>classe</b>	<b>ordre</b>	<b>stimuli</b>	<b>classe</b>
1	quant	CF	44	confus	CO	87	va	CO	130	fond	CO
2	cet	CF	45	même	CF	88	garçon	CO	131	public	CO
3	total	CO	46	ci	CF	89	propos	CO	132	point	CO
4	sous	CF	47	chair	CO	90	nom	CO	133	tête	CO
5	moi	CF	48	soudain	CF	91	mon	CF	134	peut	CO
6	quel	CF	49	air	CO	92	dire	CO	135	croire	CO
7	ainsi	CF	50	fièvre	CO	93	vin	CO	136	quelque	CF
8	dit	CO	51	avec	CF	94	petit	CO	137	selon	CF
9	si	CF	52	soi	CF	95	ceci	CF	138	couple	CO
10	des	CF	53	chose	CO	96	quoique	CF	139	soie	CO
11	contre	CF	54	savoir	CO	97	durant	CF	140	cas	CO
12	ayant	CF	55	mot	CO	98	depuis	CF	141	leur	CF
13	ou	CF	56	neuf	CO	99	très	CF	142	donc	CF
14	rond	CO	57	tenir	CO	100	pensée	CO	143	fini	CO
15	coup	CO	58	juste	CO	101	lettre	CO	144	homme	CO
16	autre	CO	59	car	CF	102	auprès	CF	145	voix	CO
17	jamais	CF	60	guère	CF	103	aile	CO	146	sœur	CO
18	chez	CF	61	ville	CO	104	faire	CO	147	nul	CF
19	premier	CO	62	dessus	CF	105	heure	CO	148	alors	CF
20	fois	CO	63	comme	CF	106	cela	CF	149	certes	CF
21	vivant	CO	64	passer	CO	107	tôt	CF	150	aussi	CF
22	voir	CO	65	nôtre	CF	108	jadis	CF	151	celui	CF
23	prendre	CO	66	avant	CF	109	hors	CF	152	vers	CF
24	tantôt	CF	67	après	CF	110	mien	CF	153	devant	CF
25	vie	CO	68	nation	CO	111	grâce	CF	154	auquel	CF
26	penser	CO	69	cri	CO	112	combien	CF	155	main	CO
27	toujours	CF	70	dont	CF	113	sourire	CO	156	élan	CO
28	beau	CO	71	garde	CO	114	tel	CF	157	parler	CO
29	beaucoup	CF	72	voici	CF	115	femme	CO	158	parti	CO
30	genre	CO	73	assis	CO	116	duquel	CF	159	saison	CO
31	nouveau	CO	74	parmi	CF	117	bien	CO	160	ça	CF
32	jusque	CF	75	détail	CO	118	lors	CF			
33	sens	CO	76	monde	CO	119	dedans	CF			
34	dans	CF	77	partout	CF	120	puisque	CF			
35	quand	CF	78	rue	CO	121	cheveu	CO			
36	lorsque	CF	79	sien	CF	122	afin	CF			
37	moyen	CO	80	retour	CO	123	nord	CO			
38	dehors	CF	81	notre	CF	124	aimer	CO			
39	notion	CO	82	hélas	CF	125	te	CF			
40	temps	CO	83	ici	CF	126	ancien	CO			
41	encore	CF	84	fait	CO	127	puis	CF			
42	sinon	CF	85	venir	CO	128	chacun	CF			