



# Exploration des mécanismes non conscients de la perception de la parole : approches comportementales et électroencéphalographiques

Carine Signoret

► **To cite this version:**

Carine Signoret. Exploration des mécanismes non conscients de la perception de la parole : approches comportementales et électroencéphalographiques. Psychologie. Université Lumière - Lyon II, 2010. Français. <tel-00562541>

**HAL Id: tel-00562541**

**<https://tel.archives-ouvertes.fr/tel-00562541>**

Submitted on 15 Nov 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Exploration des mécanismes non conscients de la perception de la parole : approches comportementales et électroencéphalographiques

Thèse de doctorat de Psychologie

*Mention Psychologie Cognitive*

Carine SIGNORET

Sous la direction de Dr. Fabien PERRIN et Dr. Barbara TILLMANN

*CNRS – UMR 5020 Laboratoire de Neurosciences Sensorielles, Comportement et Cognition*

École Doctorale 476 : Neurosciences et Cognition

Présentée et soutenue publiquement le 10 juin 2010

Traduction du 30 Septembre 2011

Sous l'évaluation de : Dr. Catherine TALLON-BAUDRY, Rapporteur

Pr. Jonathan GRAINGER, Rapporteur

Pr. Axel CLEEREMANS, Examineur

Dr. Barbara TILLMANN, Directrice de thèse

Dr. Fabien PERRIN, Directeur de thèse



*À ceux qui comptent :*

*à Adrian,*

*à Luc,*

*Jag Älskar Dig.*





# REMERCIEMENTS

De tous les mercis, mon premier va à Fabien,  
Qui m'a insufflé un courage indispensable,  
Afin que ce travail devienne réalisable.

Mon second merci se réserve pour Barbara,  
Source d'idées ingénieuses et intarissables,  
Qui transmet une rigueur vraiment incomparable.

Merci aussi aux membres d'une équipe CAP-tivante,  
Actuels et anciens, renommés et prochains,  
Dont voici une liste condensée de quelques-uns :

Nico, Ben, Katrin, Géraldine, Fred, Lisianne, Aymeric, Jérôme, Lauranne, Philippe et  
Johann.

Par ailleurs, je tiens à remercier chaleureusement Marie-Claude et Pierre, ma tendre Mamie, ma très chère Julie ainsi que Fred et Caro, qui ont chacun à leur manière, su m'apporter le soutien plus que nécessaire à la réalisation de cette thèse. Je remercie également Franck, Cécile, Kelly, Will, Dorothée, Betty, Loïc, Laëtitia, Lio, Grogn', Émilie sans oublier mon Lapin et mes ches voisins du quatrième qui ont TOUS su créer des îlots de joie et d'évasion au milieu d'un torrent de travail.

Mes ultimes remerciements sonneront plutôt comme des excuses envers ma famille la plus proche : pour ne pas avoir été souvent présente auprès de vous aussi bien physiquement que psychologiquement, pour avoir passé des nuits et des week-ends au laboratoire, pour avoir donné la priorité à la recherche plutôt qu'à ma vie de mère et de femme. Je pense tout particulièrement à mon fils, *Adrian*, avec qui j'aurai souhaité partager bien plus de temps que je ne me suis donné la permission de le faire et à mon mari, *Luc*, qui a su gérer d'une façon exemplaire, rassurante et intelligente, aussi bien matériellement que sentimentalement, la situation angoissante induite par le travail de thèse et surtout par la période de rédaction !

J'adresse enfin une pensée singulièrement reconnaissante aux membres du jury de thèse, aux 328 volontaires inscrits aux expériences, à tous les membres du laboratoire NSS devenu entre-temps NSCC, au volcan Eyjafjallajokull de s'être réveillé juste à temps, à Manue et Hannah pour leur patience, et enfin à Étienne : des milliers de mercis ne pourront jamais refléter la hauteur de son incommensurable soutien, aussi bien amical que scientifique.

À l'âge de 14 ans, devant la mairie de Roche-la-Molière, en écoutant parler les plus grands de leur devenir professionnel, j'ai décidé que mon avenir contiendrait la réalisation d'une Thèse en Psychologie, sans trop savoir sur quel chemin cette envie allait m'engager réellement. J'étais heureuse d'avoir pris cette décision, qui me semblait être comme un rêve lointain... Me voilà aujourd'hui encore plus heureuse, et surtout soulagée, d'avoir réalisé ma thèse, mon rêve ; 15 ans plus tard.

*Saint-Étienne, le 10 Juillet 2010*

# TABLE DES MATIÈRES

REMERCIEMENTS .....	5
TABLE DES MATIÈRES .....	7
RÉSUMÉ .....	13
LISTE DES ABBRÉVIATIONS .....	15
INTRODUCTION .....	17
CHAPITRE 1 SIGNAL SONORE ET SENSATION AUDITIVE .....	21
1.1 Signal sonore .....	21
1.1.1 <i>Les différents signaux</i> .....	21
a. Son pur .....	22
b. Son complexe.....	23
c. Son de parole.....	25
1.1.2 <i>La sensation auditive</i> .....	27
a. Hauteur .....	27
b. Sonie.....	28
c. Audiométrie.....	31
1.2 Anatomie fonctionnelle .....	32
1.2.1 <i>Au niveau périphérique</i> .....	32
1.2.2 <i>Les voies auditives ascendantes</i> .....	34
1.2.3 <i>Au niveau cortical</i> .....	36
1.2.4 <i>Projections corticales descendantes</i> .....	39
CHAPITRE 2 PERCEPTION DES SONS DE PAROLE.....	42
2.1 Méthodes de mesure .....	42
2.1.1 <i>Les mesures comportementales</i> .....	43
a. Aperçu des différentes tâches expérimentales.....	43
b. Approches subjective et objective.....	46
2.1.2 <i>Les mesures d'électroencéphalographie</i> .....	49

a. L'origine de l'EEG.....	50
b. Analyse du signal.....	52
2.1.3 <i>Les autres méthodes d'investigation fonctionnelle</i> .....	61
2.2 Analyse cognitive des sons de parole .....	62
2.2.1 <i>Le concept de lexique mental</i> .....	63
a. Données de psycholinguistique .....	63
b. Confirmation des données par la neurophysiologie .....	64
2.2.2 <i>Les étapes du traitement</i> .....	65
a. Modèles de traitement autonome séquentiel .....	66
b. Modèles de traitement autonome parallèle.....	69
2.2.3 <i>Les apports de la neuropsychologie et de la neurophysiologie</i> .....	72
2.3 Influence non consciente des connaissances impliquées lors du traitement des sons de parole .....	76
2.3.1 <i>Données expérimentales</i> .....	77
a. La restauration phonémique .....	77
b. La catégorisation de phonèmes ambigus.....	78
c. Les contraintes lexicales.....	79
2.3.2 <i>Le modèle « TRACE »</i> .....	80
2.3.3 <i>Vers un nouveau modèle</i> .....	82
2.4 Traitements liés à une perception non consciente des sons de parole .....	83
2.4.1 <i>Influence du contexte ignoré</i> .....	84
2.4.2 <i>Influence du contexte non perçu consciemment</i> .....	85
a. Perceptions des sons de parole lors de l'atténuation ou de la diminution de la conscience .....	86
b. Perceptions subliminales à l'éveil.....	87
c. Corrélats neuronaux de la perception consciente et non consciente.....	90
d. Données en modalité auditive .....	94
CHAPITRE 3 PROBLÉMATIQUE.....	97

CHAPITRE 4 EXPLORATIONS DE L'INFLUENCE DES PROCESSUS <i>TOP-DOWN</i> SUR LA	
DÉTECTION DES SONS DE PAROLE .....	101
4.1 Étude I : Performances de détection des mots, pseudo-mots et sons	
complexes.....	102
4.1.1 <i>Méthodologie</i> .....	103
Participants.....	103
Procédure.....	103
4.1.2 <i>Résultats</i> .....	104
4.1.3 <i>Discussion</i> .....	105
4.2 Étude II : « Facilitated auditory detection for speech ».....	106
<i>Abstract</i> .....	108
4.2.1 <i>Introduction</i> .....	109
4.2.2 <i>Experiment 1</i> .....	112
a. Method .....	114
b. Results.....	118
c. Discussion .....	123
4.2.3 <i>Experiment 2</i> .....	125
a. Method .....	126
b. Hypotheses .....	127
c. Results .....	128
d. Discussion .....	131
4.2.4 <i>Loudness model</i> .....	132
4.2.5 <i>General Discussion</i> .....	134
a. Speech Detection Effect.....	135
b. Recognition without awareness?.....	136
c. Conclusion.....	137
<i>Acknowledgements</i> .....	138
<i>References</i> .....	138
4.3 Bilan.....	145

## CHAPITRE 5 EXPLORATIONS DES PROCESSUS ACTIVÉS À LA SUITE D'UNE PERCEPTION

AUDITIVE NON CONSCIENTE .....	147
5.1 Étude III : Les différentes mesures du niveau de conscience .....	148
5.1.1 <i>Méthodologie</i> .....	149
Participants.....	149
Procédure.....	150
5.1.2 <i>Résultats</i> .....	150
5.1.3 <i>Discussion</i> .....	154
5.2 Étude IV : « <i>Subliminal semantic priming in speech</i> » .....	155
<i>Abstract</i> .....	157
5.2.1 <i>Introduction</i> .....	157
5.2.2 <i>Methods</i> .....	159
a. Participants .....	159
b. Stimuli.....	160
c. Procedure.....	161
5.2.3 <i>Results</i> .....	163
a. Pretest.....	163
b. Experiment .....	164
5.2.4 <i>Discussion</i> .....	166
a. Subliminal semantic priming in speech.....	166
b. Subliminal repetition priming in speech .....	167
c. Rapidity effect.....	169
d. Measures of Awareness.....	170
<i>Acknowledgments</i> .....	170
<i>References</i> .....	170
5.3 Étude V : « <i>Electrical brain dissociation for consciously and unconsciously categorized auditory stimuli</i> » .....	175
<i>Acknowledgements</i> .....	177
<i>Abstract</i> .....	177
5.3.1 <i>Introduction</i> .....	178
5.3.2 <i>Materials and Methods</i> .....	179
a. Participants .....	179
b. Stimuli.....	179
c. Experimental Procedure .....	180

d. EEG recordings .....	181
e. Data analysis .....	181
5.3.3 <i>Results</i> .....	183
a. Behavioral data.....	183
b. Event-related potentials (ERPs) .....	184
c. Event-related Spectral Perturbations (ERSPs) .....	186
5.3.4 <i>Discussion</i> .....	188
<i>References</i> .....	192
CHAPITRE 6 CONCLUSIONS ET PERSPECTIVES .....	196
6.1 Influences des caractéristiques énergétiques des signaux.....	197
6.2 Traitements cognitifs des sons de parole .....	199
6.3 Influences des connaissances sur la perception auditive .....	201
6.3.1 <i>Influences des connaissances sur les traitements de hauts niveaux</i> .....	201
6.3.2 <i>Influences des connaissances sur les traitements de bas niveaux</i> .....	202
6.4 Influence de la perception auditive inconsciente sur la perception ultérieure .....	205
6.4.1 <i>Conditions expérimentales utilisées</i> .....	205
6.4.2 <i>Perception non consciente des stimulations lexicales</i> .....	206
a. Analyse de bas niveau .....	206
b. Traitement sémantique .....	208
6.4.3 <i>Niveau de conscience</i> .....	209
6.5 Conclusions.....	211
<b>ANNEXES .....</b>	<b>ERROR! BOOKMARK NOT DEFINED.</b>
ANNEXE 1 .....	214
Procédure pour audiométrie tonale .....	214
ANNEXE 2 .....	216
Résultats supplémentaires de l'Étude V .....	216
LISTES DES FIGURES .....	219
BIBLIOGRAPHIE.....	224





# RÉSUMÉ

Tandis que de nombreuses informations sont disponibles dans notre environnement à chaque instant, toutes ne donnent pas lieu à une perception consciente. Il est alors légitime de se demander quels mécanismes entrent en jeu dans le phénomène de perception. Sur la base de quels processus une stimulation sensorielle sera-t-elle perçue de façon consciente ? Que deviennent les stimulations qui ne sont pas perçues consciemment ?

Ce présent travail de thèse vise à apporter des éléments de réponse à ces deux questions dans la modalité auditive. À travers plusieurs études utilisant des approches comportementales mais aussi électroencéphalographiques, nous suggérons que les connaissances pourraient exercer une influence *top-down* facilitant les hauts comme les bas niveaux de traitement (comme la détection) des stimulations auditives complexes. Les stimulations pour lesquelles nous avons des connaissances (phonologiques et sémantiques) sont mieux détectées que les stimulations ne contenant ni caractéristique phonologique ni caractéristique sémantique. Nous montrons également que l'activation des connaissances influence la perception des stimulations ultérieures, et ce, même lorsque le contexte n'est pas perçu consciemment. En effet nous avons pu mettre en évidence un effet d'amorçage sémantique subliminal et nous avons observé des modifications neuronales oscillatoires dans la bande de fréquence bêta concomitante au traitement lexical de stimulations non catégorisées consciemment.

L'ensemble des perceptions auditives ne serait alors que le produit d'une interaction permanente entre le contexte environnemental et les connaissances, ce qui nous conduirait à percevoir préférentiellement ce que nous connaissons déjà.



# LISTE DES ABBRÉVIATIONS

ANOVA : *ANalysis Of VAriance*

CCE : Cellules Ciliées Externes

CCI : Cellules Ciliées Internes

CGM : Corps Genouillé Médian

EEG : ÉlectroEncéphaloGraphie

GTS : Gyrus Temporal Supérieur

Hz : Hertz

IRMf : Imagerie fonctionnelle par Résonance Magnétique

LSD : *Least Significant Difference*

ms : millisecondes

PWSE : *Pseudo-Word Superiority Effect*

SDE : *Speech Detection Effect*

STS : Sillon Temporal Supérieur

TEP : Topographie par Émissions de Positrons

WDE : *Word Detection Effect*

WSE : *Word Superiority Effect*

SD : *Standard Deviation*

TDS : Théorie de la Détection du Signal

dB : déciBel

SPL : *Sound Pressure Level*

SL : Sensation Level

HL : Hearing Level



# INTRODUCTION

Deux cent vingt milliards de minutes de conversation sont échangées chaque année dans le monde via le téléphone. Force est donc de constater que la parole constitue un outil de communication essentiel à toute relation humaine, impliqué de façon importante dans notre vie quotidienne. Le langage est également un système très complexe dont les mécanismes, notamment ceux à la base de la perception de la parole, restent en partie énigmatiques. Outre la complexité des paramètres physiques qui déterminent la structure des sons de parole, la richesse des informations qu'ils véhiculent nécessite des mécanismes de traitement élaborés qui nous permettent d'accéder très rapidement à toutes les informations (prosodiques, sémantiques, sociales, culturelles...) contenues dans le langage. Ceci suggère que des processus automatiques et facilitateurs participent à la perception. De la même façon, la multitude d'événements sensoriels contenus dans notre environnement implique que, seule une partie de ces événements aboutit à une représentation consciente. Toutefois, nous avons tous fait le constat que les événements "ignorés" pouvaient tout de même influencer notre perception consciente, suggérant par conséquent que ces événements ont été analysés de façon non consciente. En effet, l'idée selon laquelle notre comportement pourrait être influencé par des stimulations perçues inconsciemment est ancienne, et a été explicitement exprimée en 1704 par le philosophe Gottfried Leibniz dans ses *Nouveaux Essais sur l'entendement humain* :

*« Des centaines d'indications nous permettent de conclure qu'à chaque moment existe en nous une infinité de perceptions, non accompagnées par la conscience »  
(Leibniz, 1886).*

À partir de ces constats, on peut donc se demander :

- Quels sont les mécanismes non conscients qui participent à, ou influencent, la perception des sons de parole ? Est-ce que nos connaissances langagières facilitent cette perception ?
- À quel point les sons de parole dont aucune représentation consciente n'est établie sont traités inconsciemment ?

Un des enjeux majeurs de la psychologie et des neurosciences cognitives de ces 30 dernières années a été d'apporter des données expérimentales de l'existence de processus non conscients pouvant influencer notre comportement, qu'il soit cognitif ou neuronal. Ce travail de thèse se concentre sur deux mécanismes non conscients qui participent à la perception des sons de parole : l'influence des connaissances linguistiques et l'influence du contexte (c'est-à-dire des stimulations précédentes) non perçu consciemment.

Si la psychophysique permet de contrôler finement les paramètres des stimulations utilisées dans les études visuelles, il est encore difficile de bien contrôler les caractéristiques acoustiques du signal de parole (Chapitre 1). Connaître les caractéristiques physiques du signal sonore (Section 1.1), mais aussi les phénomènes de transduction neuronale et de projections ascendantes sur les structures auditives primaires (Section 1.2), permet de comprendre les impressions psychologiques liées à la perception auditive. Nous verrons qu'une originalité de ce travail de thèse repose sur le fait que les propriétés physiques des stimulations auditives utilisées ont été finement contrôlées afin de les rendre les plus comparables possibles.

Notre système cognitif met en jeu des processus de traitement qui permettent d'extraire le sens de la stimulation auditive perçue (Chapitre 2). En psychologie cognitive, il existe différentes approches comportementales permettant d'appréhender le fonctionnement du système cognitif. L'étude du fonctionnement cognitif peut également être explorée par la neurophysiologie, par exemple grâce aux techniques électroencéphalographiques qui permettent d'étudier, avec une précision de l'ordre de la milliseconde, le déroulement temporel du traitement perceptif d'une stimulation (Section 2.1). Ces diverses méthodes d'investigations ont permis de mettre en évidence l'influence des connaissances mais aussi l'influence du contexte sur la perception auditive consciente. Nous présenterons quelques pertinents modèles psycholinguistiques de la perception et de la compréhension des sons de parole qui se distinguent selon leur aspect sériel ou parallèle dans le traitement de l'information de parole (Section 2.2). Un seul modèle propose une influence *top-down* des processus de traitement de haut niveau vers les processus de traitement de niveau plus précoce. Dans ce travail de thèse, nous avons recherché si nos connaissances pouvaient aussi influencer les bas niveaux de traitement comme la détection auditive (Section 2.3). Nous présenterons enfin les données qui montrent que le contexte perçu sans attention

dirigée ou sans conscience, peut influencer la perception ultérieure. Nous préciserons qu'une multitude d'expériences ont étudié la perception non consciente en modalité visuelle, mais qu'il existe peu de littérature présentant des études ayant étudié la perception non consciente en modalité auditive, et qu'aucun signe de traitement sémantique en perception auditive subliminale n'a pour l'instant été démontré (Section 2.4).

Les travaux présentés dans cette thèse se situent à l'interface de plusieurs disciplines : la psychologie cognitive, la psychoacoustique – issue de la rencontre entre psychologie et acoustique –, la psycholinguistique – issue de la rencontre entre psychologie et linguistique – et les neurosciences. L'approche pluridisciplinaire du phénomène de perception des stimulations auditives complexes comme les sons de parole s'est révélée indispensable pour comprendre le fonctionnement des processus non conscients mis en jeu dans la compréhension des stimulations auditives, et particulièrement des sons de parole.





# Chapitre 1

## Signal sonore et sensation auditive

Le son de parole est un son complexe qui se définit selon plusieurs paramètres physiques. Afin de mettre en évidence les difficultés matérielles liées à l'investigation de la perception en modalité auditive, détaillons tout d'abord les propriétés physiques propres aux stimulations sonores (section 1.1.1). Nous pourrons alors décrire comment ces propriétés physiques sont transcrites par le système récepteur du signal sonore et par quelles voies elles sont transférées jusqu'au système cérébral (section 1.1.2).

### 1.1 Signal sonore

Le son est un phénomène physique résultant de la vibration acoustique des molécules de l'air. Cette excitation mécanique peut être produite par différents types de sources sonores comme un instrument de musique ou une personne qui parle. Lorsque l'onde acoustique se propage dans l'air, les particules de celui-ci subissent une impulsion qui les met en mouvement dans une direction donnée. Sur leur parcours, elles rencontrent d'autres molécules qu'elles poussent, formant ainsi une zone de compression. L'air traversé par l'onde acoustique est alors le siège de compressions et dépressions successives.

#### 1.1.1 Les différents signaux

Les signaux acoustiques peuvent être décomposés selon leurs paramètres physiques. Il est nécessaire de décrire la notion de son pur pour comprendre les caractéristiques des sons complexes et des sons de parole.

## a. Son pur

Le son pur correspond à une onde sinusoïdale, comparable aux ondes créées à la surface d'une étendue d'eau par un objet entrant en contact avec cette surface. Outre sa durée, l'onde est définie par trois paramètres : la fréquence, l'amplitude et la phase initiale (Figure 1).

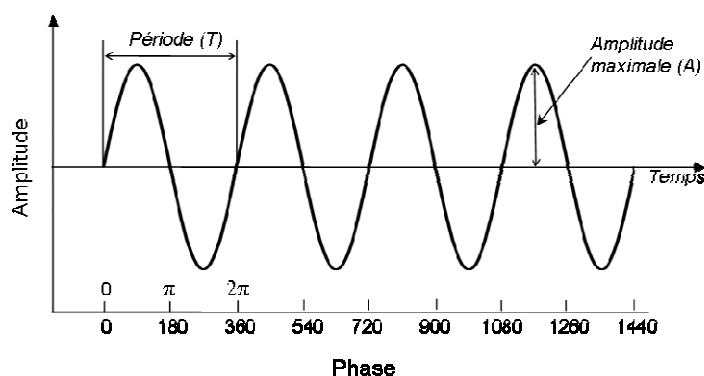


Figure 1 : Représentation d'une onde de vibration de forme sinusoïdale. Quatre cycles sont représentés, bien que cette onde peut se répéter à l'infini. L'amplitude instantanée est donnée par l'expression  $A \sin(2\pi f t + \varphi)$ , où  $t$  désigne le temps,  $f$  la fréquence,  $A$  l'amplitude maximale et  $\varphi$  la phase initiale. La phase se mesure en degrés (notés  $^\circ$ ) ou radians. Un cycle complet correspond à  $360^\circ$  ou  $2\pi$  radians. (Adapté de Moore, 2003).

La fréquence correspond à la vitesse de vibration dans l'air. La fréquence est calculée selon le nombre de cycles (ou périodes) effectués en une seconde. Elle équivaut ainsi au nombre d'oscillations par seconde (équivalent à  $1/T$ ,  $T$  étant la période) et s'exprime en Hertz (Hz). Plus la fréquence est élevée, plus le son est aigu. L'oreille humaine est sensible aux fréquences comprises entre 16 et 20 000 Hz (Figure 3). En dessous de 16 Hz, les infrasons sont ressentis par la paroi abdominale. Au-dessus de 20 000 Hz, les ultrasons peuvent être perçus par le système auditif de certains animaux (comme les chiens, les chauves-souris ou les dauphins).

Le niveau sonore se mesure en dB-SPL (*Sound Pressure Level*) qui est une unité logarithmique de pression acoustique. Lorsque l'amplitude de pression de la source sonore est multipliée par 2, le niveau sonore augmente de 6 dB. La notion de dB-SPL est toutefois insuffisante pour rendre compte de la sensation sonore effectivement perçue par l'oreille humaine (voir section 1.1.2).

La phase initiale permet de définir le moment de l'attaque d'un son pur. Un décalage entre deux sons purs au moment de leur attaque correspond à une différence de phase initiale.

Bien que les différentes mesures effectuées sur les sons (fréquence et amplitude) fassent référence au son pur, ces derniers sont rares dans notre environnement. En effet, la plupart des sources sonores produisent des sons complexes qui sont constitués d'un ensemble, éventuellement infini, de sons purs ayant chacun leur fréquence, amplitude et phase propres.

## **b. Son complexe**

Un signal sonore est rarement constitué d'une seule fréquence mais plutôt d'un ensemble de composantes fréquentielles. Souvent ces dernières sont des multiples entiers d'une même fréquence, appelée la fréquence fondamentale ( $F_0$ ). On parle alors de son complexe harmonique. Les sons complexes sont inharmoniques lorsque leurs composantes fréquentielles ne sont pas multiples de la fréquence fondamentale, ce qui est commun dans notre environnement (bruits de soufflerie, sons percussifs comme celui de la cymbale...). La fréquence fondamentale d'un son complexe est perçue comme la hauteur fondamentale du son (section 1.1.2a). De façon générale, on peut considérer tout son complexe comme une superposition de sons purs ayant des fréquences et des amplitudes variables ce qui peut résulter en des formes d'ondes très irrégulières (Figure 2.a). Au cours du temps, l'amplitude du son complexe varie selon l'amplitude des harmoniques qui le composent, ce qui influence également son enveloppe temporelle (Figure 2.b).

Les sons complexes harmoniques peuvent donc être caractérisés par des paramètres similaires à ceux des sons purs : la fréquence (fondamentale), l'amplitude et la phase de chaque fréquence harmonique. De plus, la répartition des harmoniques (ou spectre) détermine une caractéristique inhérente aux sons complexes : le timbre. Si les harmoniques les plus intenses sont en basses fréquences, le son paraît sombre mais si les harmoniques les plus intenses sont en hautes fréquences, le son paraît clair. C'est grâce au timbre qu'une note jouée par un piano pourra être distinguée de la même note jouée par un violon, ou qu'un locuteur pourra être différencié d'un autre.

L'analyse de Fourier permet d'effectuer mathématiquement une décomposition fréquentielle des sons complexes, indépendamment du temps. Au contraire, le sonagramme ou spectrogramme (Figure 2.c) permet une représentation spectrale du son complexe variant selon ses composantes fréquentielles au cours du temps. Il est notamment possible d'obtenir l'enveloppe spectrale du son complexe qui indique l'amplitude des différentes fréquences pour une fenêtre temporelle donnée. L'enveloppe spectrale ne sera donc pas la même suivant que la fenêtre temporelle est longue (par exemple pour la totalité du signal, Figure 2.d) ou courte (par exemple, pour une fenêtre de 15 millisecondes –ms– du même signal, Figure 2.e). Notons que la fréquence fondamentale se retrouve sur ce type de représentation par la distance fréquentielle séparant deux harmoniques. Ce type de représentation est réaliste dans la mesure où le système auditif décompose effectivement le signal sonore selon ses fréquences grâce à des filtres auditifs.

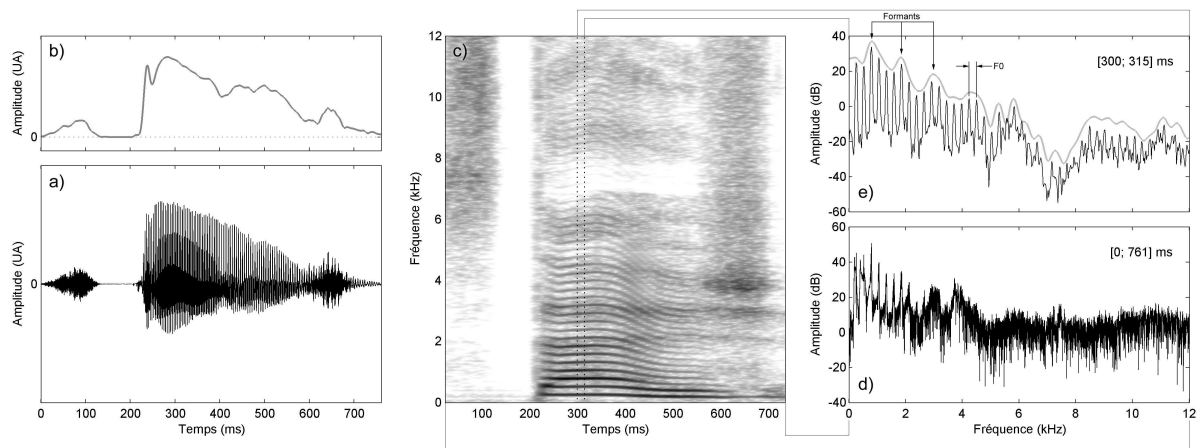


Figure 2 : Représentations acoustiques du mot « stage ». a) Amplitude globale de toutes les fréquences variant au cours du temps. b) Enveloppe temporelle représentant l'amplitude générale du son. c) Sonagramme obtenu par la transformée de Fourier. Les niveaux de gris indiquent l'amplitude de chaque fréquence au cours du temps. Les fréquences harmoniques correspondent aux traits noirs horizontaux les plus foncés. La distance observée entre deux fréquences harmoniques correspond à la valeur de la fréquence fondamentale (ici, 217 Hz). d) Amplitude donnée pour chaque fréquence d'un son complexe. e) Amplitude moyenne pour chaque fréquence sur une partie d'un son complexe. La courbe définie par les amplitudes maximales de chaque fréquence correspond à l'enveloppe spectrale. La distance observée entre deux pics fréquentiels correspond à la valeur de la fréquence fondamentale (ici, 217 Hz).

Les filtres auditifs correspondent à des mécanismes qui amplifient une certaine fréquence (la valeur centrale du filtre) tout en atténuant progressivement les autres fréquences (voir page 34). Ils sont caractérisés par leur largeur qui définit la résolution

fréquentielle du système auditif. Cette largeur, appelée ERB (*Equivalent Rectangular Bandwidth*), dépend de la fréquence : par exemple, pour une fréquence de 100 Hz la largeur est de 35.5 Hz et pour une fréquence de 2000 Hz, elle est de 240 Hz. Ainsi, chaque filtre auditif pourra être excité par plusieurs harmoniques d'un même son complexe. Si chacun des filtres auditifs excités par un son complexe reçoit un grand nombre d'harmoniques, le son est qualifié de non-résolu. Au contraire, si les harmoniques excitent chacune individuellement un filtre auditif différent, le son est considéré comme résolu. Le percept de hauteur fondamentale évoqué par un son complexe résolu est plus clair que celui évoqué par un son complexe non-résolu (Moore, 2003).

Une autre caractéristique importante des sons complexes est le fait qu'ils véhiculent du sens. C'est le cas par exemple des sons environnementaux, des sons musicaux ou des sons de parole.

### **c. Son de parole**

La voix produit des sons complexes harmoniques. Les valeurs de F0 sont en moyenne différentes chez l'homme (150 Hz) et la femme (250 Hz), ainsi que chez l'enfant (350 Hz). Les sons de parole comportent des indices acoustiques qui permettent d'identifier le locuteur selon son âge, son sexe ou son origine (francophone ou non-francophone par exemple). La liste des indices acoustiques du français a été établie par Delattre en 1958 et est remise à jour par les nombreux travaux actuels (Snoeren, Hallé, & Segui, 2006). Elle est divisée en deux catégories, les indices vocaliques (par exemple, les formants qui correspondent aux harmoniques ayant les amplitudes les plus élevées, par exemple pour le son /a/ contenu dans le mot « stage », Figure 2.e) et les indices consonantiques (par exemple, le *Voice Onset Time* qui correspond au temps d'attaque vocale, caractérisé par une bande claire dans le sonagramme, par exemple avant le /t/ du mot « stage », Figure 2.c). Les signaux de parole sont construits à partir d'une combinaison complexe de différents indices acoustiques qui permettent de classifier les différents phonèmes utilisés dans notre langue. Le phonème (Table 1) est défini comme la plus petite unité distinctive qui sert à créer des différences de sens entre les signaux de parole. Par exemple, le phonème /b/ existe car il est mis en jeu dans une opposition distinctive : /bo/-/po/. Le

changement d'un seul phonème entraîne un changement de sens dans un même contexte. Cependant, les propriétés des signaux acoustiques correspondant à des sons de parole sont bien plus complexes et bien plus variés que cette catégorisation en phonèmes. Aucune correspondance simple ne peut être faite entre les propriétés acoustiques et les propriétés phonétiques car la production d'un phonème est modifiée par les sons qui le précèdent : c'est le phénomène de coarticulation.

Table 1 : Phonèmes de la langue française. Les distinctions vocaliques entre /e/ et /ɛ/, /ø/ et /œ/, et /o/, /a/ et /ɑ/ ainsi qu'entre /œ̃/ et /ɛ̃/ ne sont pas faites par tous les locuteurs du français, suivant les accentuations régionales.

Voyelles	/i/ : il	/ɔ/ : mordre	/ə/ : le	/œ/ : peur
	/e/ : blé	/o/ : mot	/ɛ̃/ : plein	/ɑ/ : pâte
	/ɛ/ : lait	/u/ : genou	/ɑ̃/ : sans	/ø/ : peu
	/a/ : plat	/y/ : rue	/ɔ̃/ : bon	/œ̃/ : brun
Semi-consonnes	/j/ : yeux	/w/ : oui	/ɥ/ : lui	
Consonnes	/p/ : père	/R/ : rond	/f/ : feu	/l/ : lent
	/t/ : terre	/m/ : main	/s/ : sale	/g/ : gare
	/k/ : cou	/n/ : nous	/ʃ/ : chat	/ŋ/ : gai
	/b/ : bon	/ʒ/ : agneau	/v/ : vous	/z/ : je
	/d/ : dans	/h/ : hop	/z/ : zéro	

La syllabe est construite à partir de plusieurs phonèmes associés selon les règles de la langue. La structure de la syllabe comporte généralement une attaque, qui est définie par une consonne ou un assemblage de consonnes, et une rime. La rime est elle-même formée d'un noyau, qui est constitué d'une voyelle, et généralement d'une coda qui est composé d'une ou plusieurs consonnes (Table 2). Le noyau est le point le plus proéminent de la syllabe. Il correspond généralement à une voyelle. La coda comprend la (ou les) consonne(s) suivant le noyau. Par exemple, le mot « arc » est composé d'une syllabe qui n'a pas d'attaque mais seulement une rime. Codé phonétiquement /aRk/, la rime est composée d'un noyau formé par la voyelle /a/ et d'une coda formée par deux consonnes : /R/ et /k/.

Table 2 : Représentation structurale de la syllabe. L'attaque, qui n'est pas indispensable, peut se composer d'une consonne (C) ou d'une double-consonne (CC). Le noyau est en général associé à une voyelle (V). Il est suivi d'une coda, qui n'est pas indispensable, mais peut comprendre jusqu'à 3 consonnes.

Syllabe française		
Attaque	Rime	
	Noyau	Coda
C ou ø ou (CC)	V	C ou ø ou (CCC)

L'identification d'un son de parole repose sur des analyses complexes dont les bases s'acquièrent dès le plus jeune âge. Plusieurs modèles vont proposer une explication à la reconnaissance des sons de parole (voir Chapitre 2.2).

## 1.1.2 La sensation auditive

Les attributs psychologiques relatifs à l'expérience sensorielle sont différents des attributs physiques des sons. Si les aspects physiques des sons peuvent être déterminés avec précision, il est difficile de cerner les impressions psychologiques qui en découlent. Grossièrement, la fréquence renvoie à la hauteur du son (grave, médium ou aiguë), l'amplitude renvoie à son intensité (faible, forte). La phase initiale d'un son pur isolé ne donne pas lieu à un percept spécifique mais est utile pour la localisation de la source sonore. La sensation auditive fait référence aux paramètres physiques des signaux acoustiques mais dépend également des caractéristiques individuelles. Le seuil de perception d'un son (intensité minimale pour laquelle un son est perçu) dépendra alors aussi bien de son intensité que de sa fréquence.

### a. Hauteur

La périodicité d'un son induit une perception de hauteur, à la condition que la périodicité soit supérieure à 30 Hz (Pressnitzer, Patterson, & Krumbholz, 2001). Le système auditif humain est très sensible aux différences de hauteur entre deux sons. Par exemple un son pur de 1 000 Hz est discriminable d'un son de 1 002 Hz (Moore, 2003). La hauteur est la dimension relative à la perception de la mélodie en musique. Un chroma, qui définit la note dans la gamme musicale (par exemple Ré ou Sol#), et une tonie, qui définit l'octave (par exemple Ré3 ou Ré4), peuvent alors être associés à la hauteur. Notons que certaines personnes sont dotées d'une capacité à nommer la



hauteur perçue d'un son suivant les termes musicaux de chroma et tonie : on parle alors d'oreille absolue.

La saillance de la hauteur définit la clarté de la perception de cette hauteur. Dans un son complexe harmonique, la saillance de la hauteur peut être manipulée en modifiant les relations de phase entre les harmoniques. Par ailleurs, des sons ayant le même contenu fréquentiel mais des relations de phases différentes peuvent induire une sensation de hauteur différente (Moore, 2003).

## **b. Sonie**

La sonie représente l'intensité subjective d'un son perçu par l'auditeur. Elle est dépendante de l'amplitude et de la fréquence du son mais aussi des caractéristiques individuelles de l'auditeur. La sensibilité de l'oreille varie selon les fréquences dans l'ensemble du domaine audible : pour une même amplitude, les sons de la zone conversationnelle sont mieux perçus que les sons graves ou aigus. Par exemple, le seuil de perception d'un son pur de 20 Hz est d'environ 80 dB, celui d'un son pur de 20 000 Hz est d'environ 30 dB, alors que celui d'un son pur de 2 000 Hz est de 0 dB. Le seuil de perception d'un son dépend également de plusieurs autres paramètres, à savoir : la fatigue de l'auditeur, le moment de la journée, l'environnement sonore... Lorsque le son atteint un niveau d'intensité élevé, il devient douloureux : c'est la limite supérieure de perception. Au-delà de ce seuil de douleur, la perception n'est plus auditive mais devient nociceptive.

La sonie d'un son d'une fréquence donnée peut être comparée à celle qu'il aurait à une autre fréquence : des courbes d'isonie sont ainsi obtenues (Figure 3). Sur chaque courbe représentant un niveau d'intensité sonore, la même sonie est perçue à chaque fréquence. Les fréquences moyennes sont favorisées par l'oreille humaine car il faut moins de pression acoustique pour les percevoir. Ainsi, un son de 100 Hz présenté à 35 dB-SPL ne sera pas perçu. Par ailleurs, plus l'intensité du son est élevée, plus la sonie sera stable entre les fréquences. Le phone est une unité de mesure logarithmique proposée pour désigner la sonie d'un son indépendamment de sa fréquence. Ainsi, deux sons de fréquences différentes ont le même nombre de phones si leur sonie est perçue de manière égale. Pour rendre compte de la sonie, une première approche simpliste consiste à corriger le niveau sonore (dB-SPL) grâce à

une pondération qui dépend de la fréquence mais aussi de l'intensité du signal sonore :

- la pondération A donne des résultats mesurés en dB-A représentatifs des sensations perçues dans les niveaux sonores faibles de 25 à 55 dB. Cette pondération est couramment utilisée pour rendre compte du filtrage effectué par l'oreille externe et moyenne (voir section 1.2.1);
- la pondération B donne des résultats mesurés en dB-B représentatifs des sensations perçues dans les niveaux sonores moyens de 55 à 85 dB;
- la pondération C donne des résultats mesurés en dB-C représentatifs des sensations perçues dans les niveaux sonores forts de 85 à 100 dB.

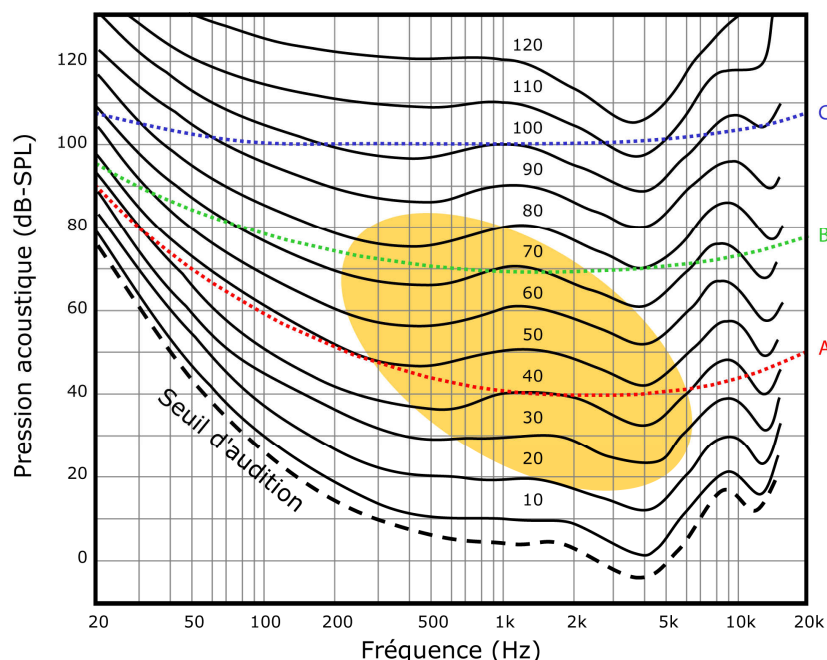


Figure 3 : Fréquences sonores (en Hertz) perçues par l'oreille humaine en fonction de l'intensité (en dB-SPL). La même sensation de sonie est ressentie de manière égale tout au long de chaque courbe d'isophonie représentée. La zone orange correspond à la zone conversationnelle dans laquelle les sons sont les mieux perçus. Les sons situés en dessous du seuil d'audition ne sont pas perçus par l'oreille humaine. La ligne rouge représente la pondération A, la ligne verte représente la pondération B et la ligne bleue représente la pondération C. (Adapté de Fletcher & Munson, 1933)

À la différence des pondérations qui se basent sur la sonie des sons purs, plusieurs modèles ont proposé des méthodes de calcul de la sonie réellement perçue pour les sons complexes comme les voyelles (Kewley-Port, 1991) ou les mots (Glasberg & Moore, 2002). Ces modèles se basent sur une modélisation du système

auditif périphérique, et visent donc à représenter l'énergie induite par le signal sonore au niveau du nerf auditif. Le modèle de Glasberg et Moore (2002) prend également en compte le phénomène d'intégration temporelle qui consiste en l'augmentation de la sonie d'un son avec la durée de ce son. Il est ainsi possible de calculer plusieurs sonies : la sonie instantanée, la sonie à court-terme, la sonie à long-terme (Figure 4) pour lesquelles on peut également obtenir la sonie moyenne et la sonie maximale du signal sonore.

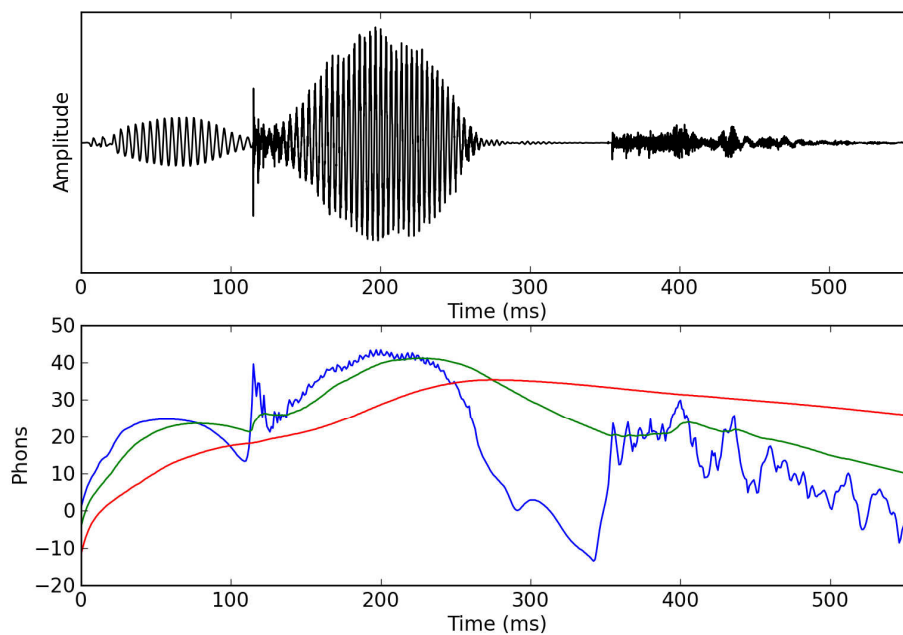


Figure 4 : Le pseudo-mot « gucre » a été utilisé dans ces représentations. *En haut* : Forme de l'onde temporelle d'un son complexe. *En bas* : Sonie mesurée au cours du temps pour un son complexe. La courbe bleue représente la sonie instantanée, la courbe verte la sonie à court-terme et la courbe rouge la sonie à long-terme.

En psychologie expérimentale, aucune étude ne s'est pour l'instant intéressée à la détectabilité des sons de parole. Comme nous l'avons vu précédemment, le son complexe est composé de plusieurs fréquences ayant chacune une amplitude et une phase particulière. Tous ces paramètres étant interdépendants, il est difficile de maîtriser parfaitement les propriétés acoustiques de ce type de son et donc de présenter des stimulations auditives complexes qui sont perçues de façon égale.

### c. Audiométrie

Pour s'assurer que le signal sonore est perçu par le système auditif d'un individu, une audiométrie peut être réalisée. Cette méthode subjective nécessite la coopération active du participant pour déterminer son seuil de perception. Le seuil de perception est représenté par une valeur, exprimée en décibel, à partir de laquelle le participant détecte le son présenté. On peut distinguer deux catégories de seuils :

- absolu (seuil de détection) : il réfère à l'intensité à partir de laquelle le participant perçoit un son donné. La tâche de l'auditeur est de définir à partir de quelle intensité il perçoit le son en répondant systématiquement à chaque fois qu'il perçoit un son. L'intensité la plus faible pour laquelle il répond est considérée comme étant la valeur de seuil de détection.
- différentiel (seuil de discrimination) : il réfère à l'intensité à partir de laquelle le participant perçoit une différence entre deux sons. La tâche de l'auditeur est de définir entre deux stimulations sonores laquelle est la plus longue, ou laquelle est la plus aiguë, ou laquelle est la plus forte par exemple.

La perception du signal sonore est très différente d'un individu normo-entendant à l'autre (Veuillet, Collet, & Duclaux, 1991). Il est donc primordial de s'assurer que les participants inclus dans une expérimentation ont une audiométrie correcte et homogène. Hormis le fait que l'audiométrie tonale nous permet de vérifier le bon fonctionnement du système auditif des individus, elle permet d'établir des seuils absolus de détection pour chaque fréquence testée. Un test d'audiométrie tonale a une durée d'environ 15 minutes. Cette procédure d'audiométrie fait référence à des normes spécifiques (ANSI, 2004) décrites en Annexe (voir page 214).

Un autre type d'audiométrie existe : l'audiométrie vocale. Elle consiste à étudier la perception de la parole en ayant pour but de tester l'intelligibilité de la parole et ainsi la compréhension de l'auditeur. La tâche de l'auditeur est de répéter des stimulations auditives émises à différentes intensités sonores. Pour ce test, le matériel phonétique peut être soit des phrases, soit des mots monosyllabiques ou dissyllabiques, soit des pseudo-mots. Par exemple, les 40 listes créées par Fournier (1951) sont constituées de 10 mots chacune. Si une erreur est commise sur le mot, celui-ci est considéré comme faux. Le résultat se compte en pourcentage par liste : le nombre de mots correctement répétés dans une liste est multiplié par 10 pour obtenir

le pourcentage d'intelligibilité. En audition monaurale, un participant normo-entendant obtient 100 % d'intelligibilité à 20 dB-SL (*Sensation Level*, c'est-à-dire au-dessus de son seuil de perception) alors qu'à 5 dB-SL, ses performances d'intelligibilité sont proches de 0 %. Le seuil d'intelligibilité représente l'intensité à partir de laquelle 50 % des mots sont compris.

Bien que les différentes listes de Fournier soient équilibrées en terme de phonèmes, d'importantes variations de performances inter-listes ont été constaté entre les listes. De plus, ces listes ne sont pas équilibrées en termes de sonie. D'autres listes existent (par exemple Combescure, 1981 ; Lafon, 1967) mais aucune ne propose un équilibrage en terme énergétique. Enfin, peu de mots sont contenus dans ces listes.

## 1.2 Anatomie fonctionnelle

Grâce à la neurophysiologie, il a été possible d'obtenir une fine description des structures anatomiques et des mécanismes impliqués dans le traitement du signal sonore. Lorsque l'onde acoustique est captée par le système sensoriel, elle se propage tout d'abord dans trois parties de l'oreille (Figure 5) : l'oreille externe, l'oreille moyenne et l'oreille interne contenant la cochlée. Les informations encodées par la cochlée sont dirigées vers différents noyaux successifs via le nerf auditif jusqu'au niveau cortical. Le traitement de l'information s'effectue alors par le cortex auditif, situé dans le lobe temporal. Nous allons présenter le rôle de ce système périphérique lors de la perception des sons de parole avant de décrire le rôle du système central dont les voies auditives ascendantes et descendantes permettant de moduler l'information perçue à des bas niveaux de traitement.

### 1.2.1 Au niveau périphérique

Partie visible du système auditif, l'oreille externe est formée du pavillon, du conduit auditif et de la membrane du tympan. La forme de l'oreille externe permet d'amplifier les fréquences sonores comprises entre 1 500 Hz et 7 000 Hz par des phénomènes de résonance; l'amplification étant maximale pour 3 000 Hz, fréquence moyenne des sons de parole. Sous l'effet de la pression acoustique, la membrane du

tympan se met à vibrer. L'oreille moyenne est une cavité osseuse remplie d'air dans laquelle s'articulent trois osselets. Les différents osselets ont pour fonction de transmettre mécaniquement les oscillations du tympan entre le milieu aérien de l'oreille externe et le milieu liquidien de l'oreille interne. Par ce système, la pression exercée au niveau de la membrane tympanique est amplifiée à l'entrée de l'oreille interne à un niveau 26 fois supérieur. Les osselets sont rattachés à la boîte crânienne par de petits muscles qui peuvent, lors de sons trop intenses, bloquer en partie cette transmission afin de protéger l'oreille. La contraction de ces muscles, appelée réflexe stapédien, est involontaire. L'oreille interne comporte deux parties anatomiquement liées mais fonctionnellement distinctes : l'appareil vestibulaire et la cochlée, organe récepteur de l'audition. La cochlée transforme les ondes de pression en influx nerveux.

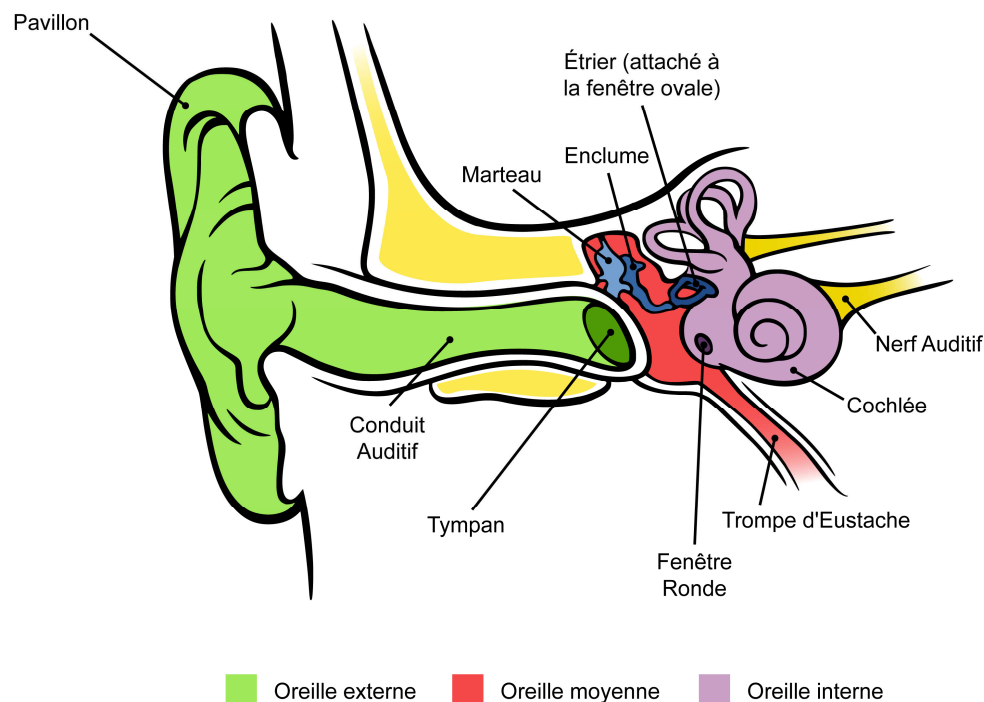


Figure 5 : Schéma de l'oreille externe, moyenne et interne, chez l'humain. (Adapté de Chittka & Brockmann, 2005)

La cochlée est un conduit enroulé en spirale dont le canal cochléaire soutient la membrane basilaire. Longue de 35 mm, elle supporte l'organe de Corti qui contient les cellules réceptrices : les cellules ciliées externes (CCE) et les cellules ciliées internes (CCI). La vibration sonore provoque une onde de déformation spécifique de la membrane basilaire qui se propage de la base à l'apex. Une correspondance précise existe entre le lieu de déformation de la membrane basilaire et la fréquence du son :

c'est une organisation tonotopique. Les CCI oscillent en fonction de la déformation de la membrane basilaire au niveau du maximum de son onde. Elles sont à l'origine du codage de l'information auditive. Si la pression acoustique sonore est insuffisante, elle n'engendrera pas d'action des CCI. Si la pression acoustique est forte, les CCI auront une activité plus étendue, qui sera limitée par l'action des CCE. Des étapes de traitement et d'encodage des signaux sonores sont alors engagées. Apte à décoder les différentes fréquences d'un son en temps réel (par l'intermédiaire des CCI) et dotée d'une sélectivité fréquentielle considérable (en raison de l'action des CCE), la cochlée constitue un banc de filtres qui fournit au système auditif une analyse fréquentielle et temporelle des signaux acoustiques comparable à ce que l'on peut observer sur un spectrogramme. Les cellules ciliées sont directement connectées par leur base au nerf auditif. Chacune des fibres afférentes du nerf auditif présente des propriétés de sélection à la fréquence du son, appelée la fréquence caractéristique. La fréquence des potentiels d'action de ces fibres augmente en fonction de l'intensité du son.

## **1.2.2 Les voies auditives ascendantes**

Les noyaux cochléaires sont le premier relais auditif de la voie auditive primaire (Figure 6). À ce niveau s'effectue un travail de décodage du message en fonction des paramètres de durée, d'intensité et de fréquence. Les noyaux cochléaires suivent une organisation tonotopique : les neurones sont sélectifs d'une fréquence et présentent des réponses particulières en fonction de l'intensité. Les réponses de certains neurones semblent extraire le bruit de fond d'une stimulation complexe. Les noyaux cochléaires sont à l'origine de plusieurs voies de projection binaurale sur le thalamus : une qui fait relais dans le complexe olivaire supérieur, une autre dans le lemnicus latéral et une dernière dans le colliculus inférieur. Ainsi, dès le tronc cérébral, le système auditif est organisé de façon parallèle. La fonction principale du complexe olivaire supérieure réside dans la localisation spatiale des sons. Le lemnicus latéral reçoit également des afférences binaurales. Les neurones du colliculus inférieur montrent une sélectivité tonale remarquable, et entraînent des réponses différentes pour des fréquences proches, par exemple entre 10 125 Hz d'une fréquence de 10 000 Hz. Certains neurones ne répondent d'ailleurs qu'à des sons modulés en fréquence, d'autres (appelées les cellules à délai) seulement à des sons de durée spécifique. Les voies auditives ascendantes font ensuite relais dans le corps genouillé médian (CGM) du

thalamus dont les neurones présentent deux types de sélectivité : une sélectivité spécifique à des sons complexes présentant des combinaisons particulières de fréquences, et une sélectivité pour des intervalles de temps spécifiques entre deux fréquences d'un signal sonore.

À partir du nerf auditif, une autre voie de projections ascendantes et de connexions secondaires existerait. Cette voie, appelée la voie réticulaire, n'est pas dédiée exclusivement à la modalité auditive : elle est polymodale. Elle reçoit diverses informations sensorielles et permet une sélection du type d'information à traiter en priorité. Par contre, la voie auditive primaire est une voie spécifique à la fonction auditive qui relie la cochlée au cortex auditif primaire.

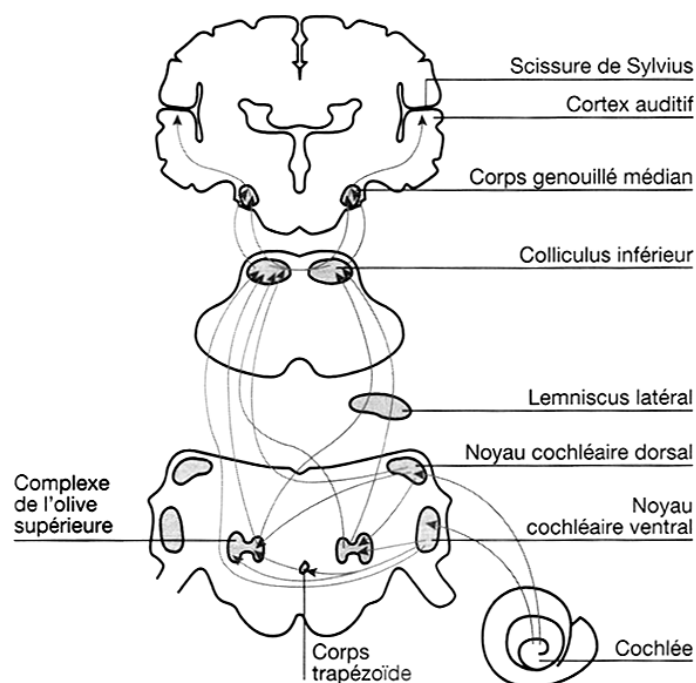


Figure 6 : Présentation schématique des voies auditives ascendantes, de la cochlée jusqu'au cortex auditif primaire (adaptée de Bonnet & Chancier, 1999). Indication de plusieurs relais de ces voies. Les fibres auditives de chaque oreille se projettent dans le cortex gauche et dans le cortex droit.

L'électroencéphalographie (EEG) a permis de mettre en évidence le fonctionnement des différents relais sous-corticaux nécessaires au traitement du signal sonore (Figure 6). Les potentiels évoqués auditifs (PEA, voir page 54) de latence précoce reflètent les étapes du traitement sous-cortical du signal sonore. Les variations électriques enregistrées résultent de la propagation et du traitement du message nerveux auditif dans le système auditif jusqu'au niveau cortical. Le signal recueilli représente la réponse des noyaux successifs des voies auditives qui se traduit



par une série de déflexions successives (Picton, Hillyard, Krausz, & Galambos, 1974). Le signal issu des PEA de latence précoce dépend à la fois de la structure acoustique de la stimulation, ainsi que de l'état fonctionnel du système auditif du participant. Ces PEA de latence précoce ont lieu dans les 10-12 ms qui suivent la stimulation sonore (Figure 7) dans le nerf auditif (I et II), puis le noyau cochléaire et le complexe olivaire supérieur (III), le lemniscus latéral et le colliculus inférieur (IV et V).

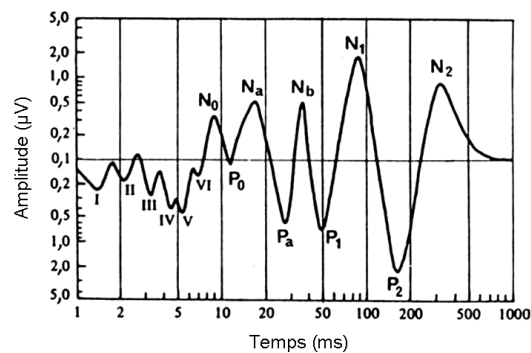


Figure 7 : Potentiels évoqués de latence précoce (I à VI) de moyenne latence (N0 à Nb) et de latence tardive (P1 à N2). Ces composantes sont considérées comme sensorielles ou obligatoires car elles apparaissent à la suite de toutes stimulations auditives. (Adapté de Picton et al., 1974).

### 1.2.3 Au niveau cortical

Depuis le CGM au niveau thalamique, les fibres se projettent sur le cortex auditif. Le cortex auditif fait partie du lobe temporal du cerveau. Il est situé dans le gyrus temporal supérieur, au niveau de la scissure de Sylvius. Il est constitué du cortex auditif primaire (également appelé A1) situé au niveau du gyrus de Heschl et du cortex auditif secondaire (également appelé A2) situé dans la partie postérieure du lobe temporal supérieur (Figure 8).

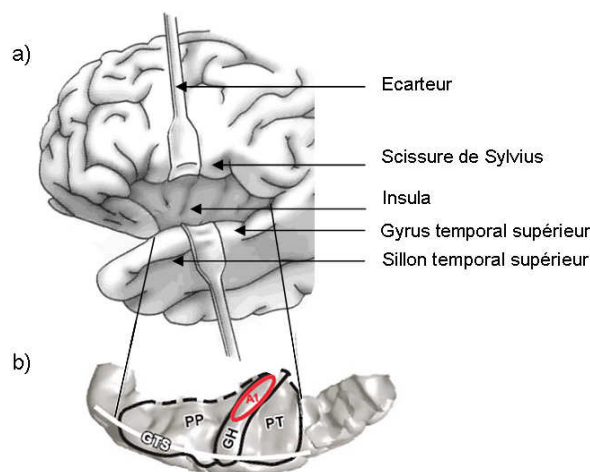


Figure 8 : a) Vue latérale du cortex cérébral gauche situant le sillon temporal supérieur (STS), le gyrus temporal supérieur (GTS), la scissure de Sylvius et laissant apparaître l'insula à l'aide des écarteurs (schéma adapté de Purves et al., 2004) b) Vue de dessus du plan temporal supérieur situant l'aire auditive primaire (A1) au niveau du gyrus de Heschl (GH). Les aires auditives associatives se situent dans la partie latérale du gyrus de Heschl, dans les planum polare (PP) et temporale (PT) et dans le gyrus temporal supérieur (GTS) (schéma adapté du cerveau humain standard du Montréal Neurological Institut et Bidet-Caulet, 2007).

Les PEA de latence moyenne (N0, P0, Na, Pa et Nb) qui succèdent aux PEA de latence précoce reflèteraient les différentes étapes de traitement de l'information dans A1 et se produisent dans les 10 à 50 ms qui suivent la stimulation (Figure 7). La source exacte des réponses de moyenne latence est mal connue, mais elle correspondrait vraisemblablement à l'activité de générateurs de A1, les sources des Na/Pa se trouvant plutôt à la pointe du gyrus de Heschl (Figure 9). L'arrivée des premières informations au cortex auditif pourrait alors avoir lieu dès l'onde Na, soit 16 ms après le début de la stimulation sonore. Les PEA de latence moyenne sont en général considérés comme des composantes sensorielles (donc obligatoirement produits à la suite d'une stimulation sonore) car ils sont associés aux étapes précoces du filtrage de l'information acoustique, en relation avec les caractéristiques physiques de la stimulation (Näätänen, 1990). L'absence, ou la détérioration, d'un des pics des PEA de latence précoce ou moyenne peut donc être reliée à un dysfonctionnement du traitement du signal sonore au niveau correspondant des voies auditives (Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994). Le PEA tardif N1 reflèterait également l'activation d'A1 : ses propriétés sont développées dans le chapitre suivant (Section 2.1.2) car cette composante serait influencée par des processus attentionnels.

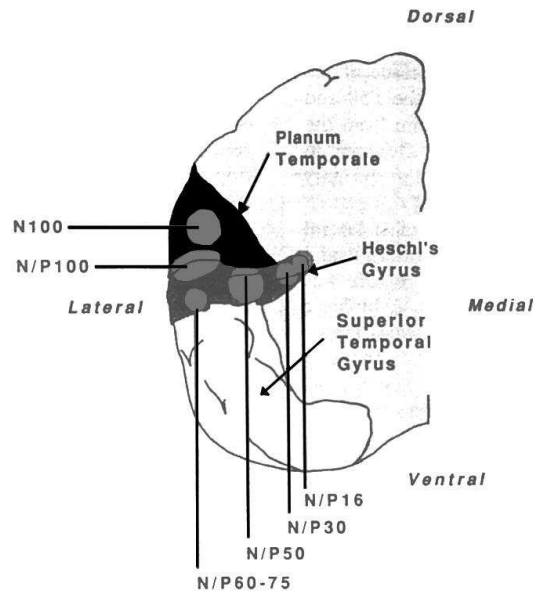


Figure 9 : Représentation schématique de la localisation anatomique des générateurs des potentiels évoqués de latence moyenne (N/P16 pouvant correspondre au Na et N/P30 pouvant correspondre aux Pa et Nb observés en EEG) et tardifs (N/P50 pouvant correspondre au P1a et N/P60-75 pouvant correspondre au P1b et N/P100 pouvant correspondre au N1 observés en EEG) dans le gyrus de Heschl et le planum temporale (Liégeois-Chauvel et al., 1994).

A1 est organisé tonotopiquement, reflétant ainsi l'organisation tonotopique cochléaire. En effet, A1 reçoit des projections point par point en provenance du CGM ce qui induit une organisation tonotopique précise. Cette organisation tonotopique (Figure 10) a été confirmée par des études de neuroimagerie (Humphries, Liebenthal, & Binder, 2010). Le gyrus de Heschl est activé bilatéralement par toutes stimulations sonores, aussi bien par des sons de parole que par d'autres sons, comme des sons environnementaux ou des sons purs (par exemple Binder et al., 2000). Le cortex auditif primaire constitue ainsi la voie d'entrée principale de toutes les stimulations sonores de l'environnement.

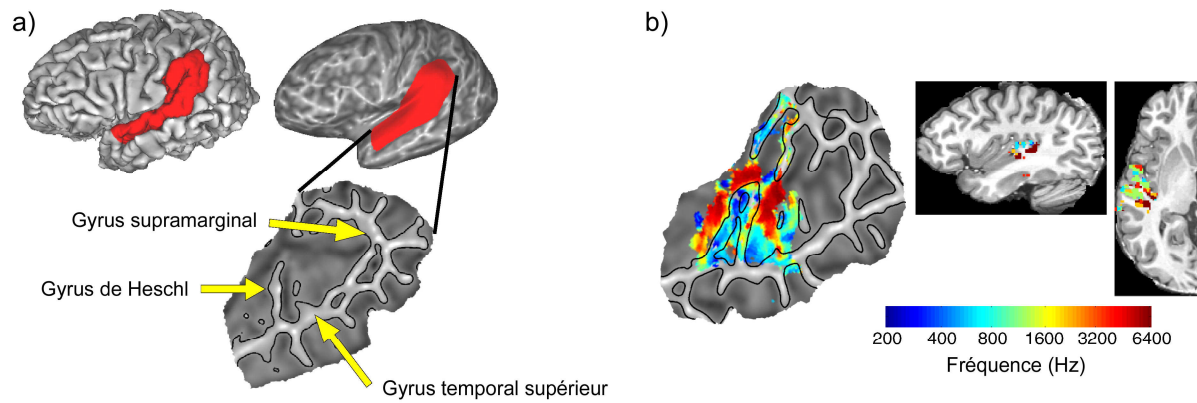


Figure 10 : a) Localisation du gyrus de Heschl au niveau du sillon temporal supérieur. b) organisation tonotopique du gyrus de Heschl en surface, sur une coupe latérale et sur une coupe horizontale (d'après Humphries et al., 2010)

## 1.2.4 Projections corticales descendantes

Comme chacun des cinq sens, l'audition est dotée d'un système ascendant que nous avons décrit précédemment, remontant l'information depuis les récepteurs sensoriels périphériques jusqu'aux aires corticales. En parallèle de ces processus ascendants, il existe également des processus descendants empruntant les voies efférentes qui permettent de moduler le message nerveux afférent afin de réaliser une analyse complexe du signal sonore. Les voies auditives efférentes ont été largement décrites chez l'animal grâce à des techniques de marquage neuronal (Huffman & Henson, 1990 ; Warr, Guinan, & White, 1986).

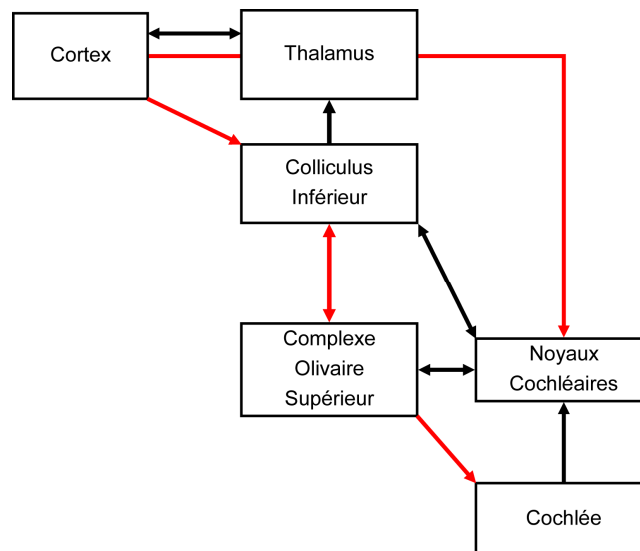


Figure 11 : Voies ascendantes (en noir) et descendantes (en rouge) du système auditif. Une première voie efférente descend du cortex au colliculus inférieur, puis au complexe olivaire supérieur pour ensuite atteindre la cochlée. Une seconde voie efférente se projette directement du cortex aux noyaux cochléaires qui vont ensuite atteindre le complexe olivaire supérieur avant de redescendre au niveau de la cochlée. Seule la voie descendante olivo-cochléaire est étudiée de manière non-invasive chez l'humain (d'après Collet, 2005).

Les fibres auditives efférentes se projettent du cortex auditif au colliculus inférieur, du colliculus inférieur au complexe olivaire supérieur et du complexe olivaire supérieur à la cochlée. Les fibres efférentes sont directement connectées aux CCE et permettent la modulation de leur activité. Globalement, ces voies descendantes peuvent exercer une activité excitatrice ou inhibitrice sur les structures sur lesquelles elles font synapse. Chez l'humain, il est impossible d'explorer ces voies de manière non invasive sauf pour le système efférent olivo-cochléaire (Khalifa, 1999 ; Micheyl, Carbonnel, & Collet, 1995 ; Moulin, Collet, & Duclaux, 1993). Le système efférent médian chemine du noyau olivaire supérieur médian par le nerf vestibulaire jusqu'à l'organe de Corti (Warr et al., 1986) où il se connecte aux CCE de la cochlée. C'est par le biais de la contraction des CCE que le système nerveux central pourrait exercer une influence descendante sur les mécanismes d'amplification cochléaire. Plusieurs travaux ont été menés au sein de notre laboratoire (Khalifa et al., 2001 ; Perrot et al., 2006) pour mettre en évidence l'implication de ce système efférent, dont le rôle n'est pas encore totalement élucidé. Ces efférences, permettant un contrôle en retour d'un relais sur le relais précédent, pourraient intervenir dans le traitement sélectif de sons, l'adaptation, la détection du signal dans le bruit, l'amélioration de la discrimination, la protection contre la sur-stimulation ou le traitement attentionnel

(Avan & Bonfils, 1992 ; Giard, Collet, Bouchet, & Pernier, 1994 ; Meric & Collet, 1994). Quelques résultats mettent en évidence un rétrocontrôle cortical au niveau du tronc cérébral dans la perception des sons de parole (Akhoun, 2008 ; Krishnan, Xu, Gandour, & Cariani, 2005) ou des sons de musique (Marmel, 2009 ; Wong, Skoe, Russo, Dees, & Kraus, 2007). Ainsi, le système efférent reflète potentiellement une voie d'accès *top-down* du niveau cortical au niveau périphérique (Davis & Johnsruide, 2007).

# Chapitre 2

## Perception des sons de parole

Un des principes fondamentaux de la psychologie cognitive postule que le traitement de l'information sensorielle se fait par une série d'opérations distinctes, plus ou moins hiérarchisées, qui se déroulent par étapes successives et/ou parallèles. Plusieurs paradigmes expérimentaux ont été mis au point afin de pouvoir observer les capacités cognitives des participants (section 2.1.1). Une première série de modèles mettent en évidence des processus ascendants permettant de comparer le signal sonore aux connaissances stockées en mémoire (section 2.2). Cependant, des indices expérimentaux questionnent le fonctionnement purement ascendant des mécanismes impliqués dans la reconnaissance des sons de parole. Un modèle de reconnaissance des sons de parole propose alors que des mécanismes *top-down* pourraient influencer les traitements sur la reconnaissance des sons de parole (section 2.3). Enfin, nous verrons que la reconnaissance des sons de parole peut également être influencée par une perception auditive non consciente du contexte (section 2.4). L'étude du fonctionnement cognitif par l'intermédiaire de l'EEG permet de mettre en évidence le déroulement temporel des activités cognitives observées de façon comportementale (section 2.1.2). Nous verrons que l'analyse des potentiels évoqués peut être complétée par l'analyse en temps-fréquence du signal EEG. Ces différentes mesures sont particulièrement utiles dans la mise en évidence des mécanismes non conscients impliqués dans la perception des sons de parole.

### 2.1 Méthodes de mesure

Les études de psychologie cognitive utilisent classiquement des paradigmes expérimentaux fournissant des données comportementales permettant d'explorer les différentes étapes du traitement cognitif et d'établir des hypothèses sur leur nature et leur déroulement temporel. Les traitements liés à la perception concernent la détection de la présence d'une stimulation, son identification et sa catégorisation sémantique.

Pour étudier l'aspect cognitif de la perception auditive, les performances des participants peuvent être recueillies par l'utilisation de tâches directes et de tâches indirectes. Les performances des participants peuvent alors être analysées soit à l'aide d'une approche subjective soit à l'aide d'une approche objective. Des mesures de l'activité cérébrale sont aussi utilisées en psychologie cognitive. Alors que l'EEG est la technique d'investigation permettant de rendre compte du déroulement temporel précis du traitement cérébral des stimulations sensorielles, d'autres techniques (imagerie cérébrale, études lésionnelles) permettent de rendre compte de la localisation des traitements cognitifs.

### **2.1.1 Les mesures comportementales**

Le rôle des mesures comportementales est d'évaluer les performances des participants lors de la réalisation d'une tâche. Les tâches comportementales peuvent se séparer en deux classes : les tâches directes et les tâches indirectes. De plus, deux façons de mesurer les performances des participants se distinguent : l'approche subjective et l'approche objective.

#### **a. Aperçu des différentes tâches expérimentales**

La nature de la tâche permet d'explorer directement ou indirectement les différents processus réalisés par le participant.

##### ***Les tâches directes***

Les tâches directes permettent d'explorer les performances du participant à cette tâche suite à la présentation d'une stimulation. Ces performances peuvent être évaluées par des tâches de *détection*, de *discrimination*, d'*échelonnage*, de *jugement* ou de *choix*. La tâche de *détection* détermine la capacité du participant à déceler l'absence ou la présence d'une stimulation, et n'implique pas l'identification de la stimulation. En fonction des paramètres de l'expérience, la tâche peut révéler des performances distribuées selon une courbe appelée fonction psychométrique. La tâche de *discrimination* détermine la capacité du participant à distinguer deux stimulations. La tâche d'*échelonnage* utilise une échelle de mesure afin de quantifier la variation des sensations correspondant aux variations des caractéristiques physiques des



stimulations qui en sont la source. La tâche de *jugement* fait appel à une évaluation portée sur la stimulation. La tâche de *choix* propose une opération de sélection, entre plusieurs éventualités de réponses déterminées par l'expérimentateur. Par exemple, la tâche de catégorisation sémantique et la tâche de décision lexicale sont des tâches de choix forcé. Pour la première, les participants doivent choisir si la stimulation présentée appartient à une catégorie sémantique (animal ou objet par exemple) et pour la seconde, les participants doivent choisir si la stimulation présentée est un mot ou un pseudo-mot.

Le problème des tâches directes est que, pour une même impression subjective, la réponse du participant peut dépendre de la confiance qu'il accorde à sa perception. Par exemple lors d'une tâche de détection par procédure oui/non, deux types de stratégie se distinguent : soit les participant adoptent une stratégie conservatrice (c'est-à-dire qu'il donne une réponse négative lorsqu'il n'est pas sûr de sa perception), soit les participants adoptent une stratégie libérale (c'est-à-dire qu'il donne une réponse positive même s'il n'est pas sûr de sa perception). Ainsi, un biais de réponse peut être observé à travers les réponses des participants. Pour pallier à cette critique, une échelle de réponse permettant d'analyser l'utilisation d'une stratégie peut être mise en place. Elle peut par exemple comporter quatre points de mesure : aucune perception, peut-être perçue, perçue mais non identifiée, identifiée (par exemple, Summerfield, Jack, & Burgess, 2002). Dans le même but, d'autres auteurs (par exemple Adams & Adams, 1960) utilisent une échelle de confiance permettant au participant d'indiquer le niveau de certitude qu'il accorde à sa réponse.

Bien que de nombreuses techniques et paradigmes, dont certains ont été expliqués dans cette partie, aient été développés dans le but d'évaluer la part des stratégies décisionnelles impliquées dans l'orientation du choix de réponse à une tâche directe, l'utilisation de tâches indirectes permet de minimiser la mise en place de stratégies décisionnelles.

### ***Les tâches indirectes***

Les tâches indirectes permettent d'observer l'effet ou l'influence du traitement d'une information sur le traitement d'une stimulation. En pratique, les participants sont impliqués dans une tâche pour laquelle les stimulations sont manipulées de telle

sorte qu'elles exercent ou non une influence sur les réponses des participants. Ces effets sont observables dans les tâches dites de Stroop et dans les paradigmes d'amorçage. Dans les tâches de Stroop (Stroop, 1935), le principe est de présenter une stimulation contenant en parallèle une information contradictoire ou non. Des mots désignant des couleurs sont écrits dans une police de couleur soit congruente avec le mot (par exemple, le mot "bleu" écrit en bleu) soit incongrue (par exemple, le mot "bleu" écrit en rouge). La consigne donnée aux participants est d'identifier la couleur du mot sans lire le mot lui-même. Le pourcentage d'erreurs est plus élevé lorsque la couleur de la police et le mot sont incongrus (répondre "bleu" lorsque le mot présenté est bleu). Les temps de réaction, indicateur du temps nécessaire à l'identification de la couleur avec laquelle le mot est écrit, sont également plus longs lorsque le mot et la couleur de la police sont incongrus que lorsqu'ils sont congruents ou partagent une relation de neutralité (le mot « lion » écrit en rouge). Il existe donc un effet d'interférence sémantique, ou effet Stroop, provoqué par la lecture automatique du mot.

Dans les paradigmes d'amorçage (Neely, 1977), le principe est de présenter deux stimulations, une première appelée *amorçe* et une seconde appelée *cible*. Les participants sont engagés dans une tâche concernant la seconde stimulation. Aucune tâche n'est demandée concernant l'amorçe dans ce type de paradigme, bien que des tâches ultérieures puissent être réalisées pour contrôler la qualité ou le niveau de perception de l'amorçe par exemples. Les performances des participants sont mesurées sur le traitement de la seconde stimulation, qui est influencée ou non par le traitement de l'amorçe. Ces deux stimulations peuvent être soit reliée (par la répétition, la phonologie, le sens) soit non reliées. Prenons un exemple concret : l'amorçe présentée est le mot « jour ». Dans les cas où l'amorçe et la cible sont en relation sémantique (« nuit »), ou phonologique (« four ») ou de répétition (« jour »), les taux de réponses correctes seront plus élevés et les temps de réaction aux réponses correctes seront plus courts que dans les cas où l'amorçe et la cible ne sont pas reliée (« vase »). En parallèle du paradigme d'amorçage est née une série de paradigmes proposant des tâches diverses (complétion de mots, de phrases) qui ont permis d'étudier le fonctionnement du traitement cognitif des sons de parole, mais aussi des sons de musique (Tillmann, Hoch, & Marmel, sous presse, sous presse) et des sons environnementaux (Ballas, 1993).

Les tâches indirectes ont l'avantage de mesurer avec objectivité les performances des participants. Néanmoins, l'utilisation de tâches directes s'avère indispensable pour évaluer la perception subjective des participants. L'analyse des performances des participants peut alors se réaliser soit par une approche subjective soit par une approche objective permettant d'extraire la stratégie décisionnelle des réponses des participants.

## **b. Approches subjective et objective**

L'approche subjective propose que l'interprétation subjective du participant soit la plus représentative de ses capacités, et l'approche objective conçoit que les performances des participants peuvent être biaisées par des processus décisionnels stratégiques.

### ***Mesure subjective***

Les performances des participants peuvent s'analyser en calculant les taux de réponses correctes et incorrectes. Les temps de réaction ou les temps de réponses peuvent également être analysés. La différence entre le temps de réaction et le temps de réponses réside dans l'énoncé de la consigne. Si le participant a pour consigne de répondre le plus rapidement possible, on parle du recueil des temps de réaction. Si le participant n'a pas de consigne particulière, on parle du recueil des temps de réponses.

La mesure de seuil permet de définir les performances de perception en fonction de la qualité d'une stimulation. Dans les méthodes de seuil telles que la procédure oui/non, la performance observée peut être le résultat de deux facteurs : la sensibilité du système sensoriel et la stratégie adoptée par le participant. En conséquence, le seuil subjectif de détection d'une stimulation par exemple sera plus élevé pour un participant adoptant une stratégie conservatrice que pour un participant adoptant une stratégie libérale

### ***Une mesure objective, la théorie de la détection du signal***

La théorie de la détection du signal (TDS, Green & Swets, 1974 ; Macmillan & Creelman, 2005) est une mesure objective qui permet de dissocier la sensibilité à une stimulation de l'utilisation d'une stratégie dans un modèle d'analyse des performances. La méthode proposée par ce modèle consiste à calculer la capacité du

participant à discriminer une stimulation par rapport à une autre (Figure 12). La TDS prend en compte non seulement les performances des participants mais également les stratégies décisionnelles impliquées dans l'élaboration de la réponse du participant. Pour ce faire, une unité de mesure appelée  $d'$  permet de calculer la différence de perception entre deux stimulations. Prenons un exemple précis. Les participants doivent effectuer une tâche de détection d'une stimulation. Certains essais comportent une stimulation et d'autres non. En conformité avec la TDS, les performances de discrimination entre la présence et l'absence de la stimulation sont analysées selon quatre critères (Table 3) :

- la participant a détecté la stimulation présentée, c'est une réponse correcte;
- le participant n'a pas détecté la stimulation présentée, c'est une omission;
- le participant a détecté une stimulation alors que celle-ci était absente, c'est une fausse-alarme;
- le participant n'a pas détecté la stimulation et celle-ci était absente, c'est un rejet correct.

L'application de la TDS a aussi été extrapolée à des tâches visant à discriminer deux stimulations l'une de l'autre. Dans ces conditions, une tâche de décision lexicale peut permettre par exemple d'obtenir une mesure de discrimination entre des mots et des non-mots, les mots étant considérés comme un signal présent et les non-mots étant considérés comme une absence de signal.

Table 3 : La prise en compte des quatre possibilités entre présence ou non du signal (première colonne) et réponse de l'observateur (bruit ou signal, première ligne) permet de calculer le  $d'$  (mesure de discriminabilité) et le  $c$  (mesure du critère).

	<i>Réponse Oui</i>	<i>Réponse Non</i>
<i>Signal présent</i>	Réponse Correcte	Omission
<i>Signal absent</i>	Fausse-alarme	Rejet Correct

La théorie de la détection du signal propose qu'à tout évènement sensoriel corresponde une réponse du participant. Pour une même stimulation présentée plusieurs fois au cours de différents essais, les réponses du participant se distribueront selon une loi dite normale. Ainsi, deux stimulations différentes engendreront deux distributions de même variance et dont la distance entre leur moyenne est une mesure de sensibilité. La TDS postule que les participants orientent leurs réponses,

correspondant à l'une ou l'autre des deux distributions suivant la stimulation présentée, en fonction d'un seuil interne ( $t$ ). Ce seuil interne représente la stratégie adoptée par le participant qui a plus ou moins tendance à répondre pour une distribution plutôt que l'autre.

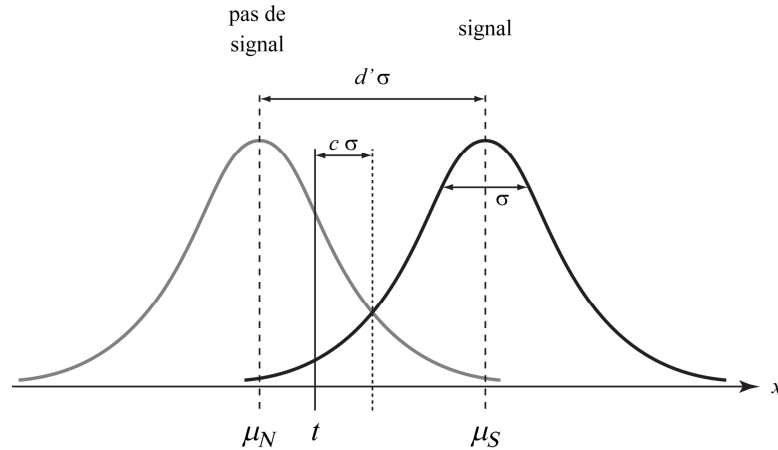


Figure 12 : La distribution des réponses des participants se répartit sur l'axe  $x$  suivant une distribution normale de moyenne  $\mu$  et de même variance  $\sigma$ . La moyenne de la distribution normale lorsqu'il y a une absence de signal est notée  $\mu_N$  et  $\mu_S$  lorsque le signal est présent. La distance entre les moyennes des deux distributions représente la valeur  $d' \sigma$ . Le biais est noté  $c$  et est lié au seuil interne  $t$  du participant. En deçà du seuil interne, les participants répondent « pas de signal » et au-delà du seuil interne, les participants répondent « signal ».

Les taux de réponses correctes et de fausses-alarmes sont ainsi normalisés en valeurs  $z$ , fonction inverse de la loi normale réduite ( $\Phi$ ). Le calcul du  $d'$  se fait ensuite par la soustraction du score  $z$  de fausses-alarmes (égales au pourcentage de « oui » lorsqu'aucun signal n'était présenté) au score  $z$  de réponses correctes (égales au pourcentage de « oui » lorsqu'un signal était présenté). Plus la valeur de  $d'$  est élevée, plus le participant est considéré comme étant capable de discriminer les deux stimulations présentées (pour l'exemple donné, de discriminer un signal par rapport à un silence).

$$\begin{aligned}
 d' &= z(\Pr(\text{Oui} \mid \text{signal})) - z(\Pr(\text{Oui} \mid \overline{\text{signal}})) \\
 &= z\left(1 - \Phi\left(\frac{t - \mu_S}{\sigma}\right)\right) - z\left(1 - \Phi\left(\frac{t - \mu_N}{\sigma}\right)\right) \\
 &= -z\left(\Phi\left(\frac{t - \mu_S}{\sigma}\right)\right) + z\left(\Phi\left(\frac{t - \mu_N}{\sigma}\right)\right) \\
 &= -\frac{t - \mu_S}{\sigma} + \frac{t - \mu_N}{\sigma} \\
 &= \frac{\mu_S - \mu_N}{\sigma}
 \end{aligned}$$

Le  $d'$  est donc indépendant du seuil interne du participant, et s'exprime comme la distance entre les deux distributions, normalisée par l'écart-type ( $\sigma$ ). Le  $d'$  représente donc bien la discriminabilité du signal par rapport au bruit. Quant au biais de stratégie, il représente la valeur du seuil interne du participant par rapport au critère idéal qui s'exprime :  $(\mu_N + \mu_S)/2$ .

Dans le cas particulier d'une tâche de détection à plusieurs niveaux de présentation, à chaque niveau de présentation correspond une valeur de  $\mu_S$ . Les proportions de réponses « oui » en fonction du niveau de présentation suivent une courbe gaussienne cumulée. Une telle fonction peut alors être ajustée aux mesures réelles pour estimer la sensibilité et la détectabilité en fonction de la stimulation, indépendamment du niveau de présentation. Une mesure de sensibilité correspond à la valeur de la pente de la courbe, alors que la mesure de détectabilité correspond à la valeur de la position horizontale de la courbe. La pente de la courbe nous renseigne sur la relation entre les variations de performances et les variations du niveau de présentation, c'est-à-dire sur l'influence de la stimulation sur la perception. La position horizontale nous renseigne sur les performances associées à un niveau de présentation, c'est-à-dire sur le seuil interne des participants. Un avantage de cette méthode réside dans le fait de réduire l'importance des problèmes d'échantillonnage et de saturation si l'ajustement est fait selon la méthode du maximum de vraisemblance.

Alors que les mesures comportementales permettent d'évaluer les différents niveaux de traitement cognitif des stimulations sensorielles, la technique de l'EEG nous permet d'observer le déroulement temporel des processus cérébraux (notamment perceptifs) et éventuellement d'estimer la localisation des activités cérébrales liées à ces traitements.

## 2.1.2 Les mesures d'électroencéphalographie

Outre les données issues de l'investigation comportementale, les processus impliqués dans le traitement des sons de parole peuvent être observés par l'activité électrique cérébrale qu'ils engendrent. Les activités neuronales oscillatoires recueillies à la surface du scalp (ou en profondeur dans le cas de la stéréo-

électroencéphalographie) reflètent ainsi les processus successifs du traitement de la stimulation, d'un niveau strictement sensoriel à des niveaux de plus haute intégration. Ces activités peuvent être analysées de deux façons : par une analyse de l'activité évoquée et par une analyse de l'activité induite.

### a. L'origine de l'EEG

L'origine de l'EEG est la découverte, chez l'animal, de variations électriques provenant du cerveau (voir notamment Caton, 1875). Hans Berger enregistra également une telle activité à la surface du scalp chez l'homme. Dans son célèbre papier de 1930, il écrivit :

*« The electroencephalogram represents a continuous curve with continuous oscillations in which...one can distinguish larger first order waves with an average duration of 90 milliseconds and smaller second order waves of an average duration of 35 milliseconds. The larger deflections measure at most 150 to 200 microvolts. »*

Berger appela ce rythme de large amplitude ayant une rythmicité d'environ 10 cycles par seconde, et qui était induit par la fermeture des yeux chez le participant éveillé mais au repos, le rythme alpha ( $\alpha$ ) car il constituait le premier rythme qu'il avait observé (Figure). Il dénomma par la suite les ondes de plus petite amplitude, présentes quand les yeux étaient ouverts, rythme bêta ( $\beta$ ).

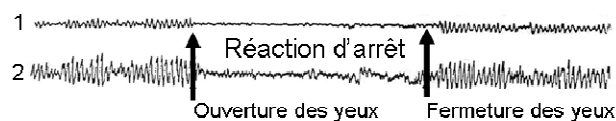


Figure 13 : Enregistrement de l'activité oscillatoire spontanée à la surface du scalp dans une région frontale (1) et dans une région occipitale (2). Apparition du rythme alpha observé lorsque l'individu éveillé ferme les yeux. Apparition du rythme bêta (et disparition du rythme alpha) observé lorsque l'individu éveillé ouvre les yeux. Le fait que l'activité alpha disparaisse lorsque les yeux sont ouverts a été nommé réaction d'arrêt (d'après Berger, 1930).

Dans sa première communication sur l'EEG humaine, il introduisit le sujet de son chapitre par une question :

*« Will it be possible to demonstrate intellectual processes by means of the EEG? »*

Il donna une réponse positive dans la même publication en montrant le blocage du rythme alpha durant différents processus mentaux. Il devint rapidement évident que le signal électrique enregistré à la surface du scalp variait en fonction de l'état des participants (Berger, 1930). Actuellement, il est bien déterminé que des courants ioniques dans les neurones (essentiellement dans les cellules pyramidales) sont à la base de l'activité électrique corticale enregistrée par l'EEG. Il est généralement admis que l'activité électrique mesurée sur le scalp ne provient pas principalement des courants ioniques intracellulaires associés aux potentiels d'action de par leur brièveté (1-2 ms) et leur irrégularité. L'activité enregistrée par l'EEG provient essentiellement des courants extracellulaires résultant des potentiels post-synaptiques excitateurs et inhibiteurs (Figure 14). Le potentiel extracellulaire engendré par ces courants est de très faible amplitude, et diminue en fonction du carré de la distance à laquelle il est enregistré. Toutefois, il semble que l'on puisse enregistrer à la surface du scalp l'activité des cellules pyramidales. En effet, les cellules pyramidales ont de longues dendrites apicales ayant toutes une orientation perpendiculaire à la surface du scalp. La sommation de leurs potentiels extracellulaires crée une activité suffisamment importante pour être recueillie sur le scalp. Ces macrocolonnes corticales, d'environ trois millimètres de diamètre et trois millimètres d'épaisseur, constituées de milliers de neurones, peuvent être assimilées à un dipôle équivalent de courant (Nunez, 1989). Ce dipôle crée un champ électrique qui se propage passivement au travers du cortex, du liquide céphalo-rachidien, de l'os et qui peut être recueilli par des électrodes à la surface du scalp. Les variations de potentiels recueillies sur le scalp, chez l'homme, résultent donc de la sommation d'activité de nombreuses populations de neurones.



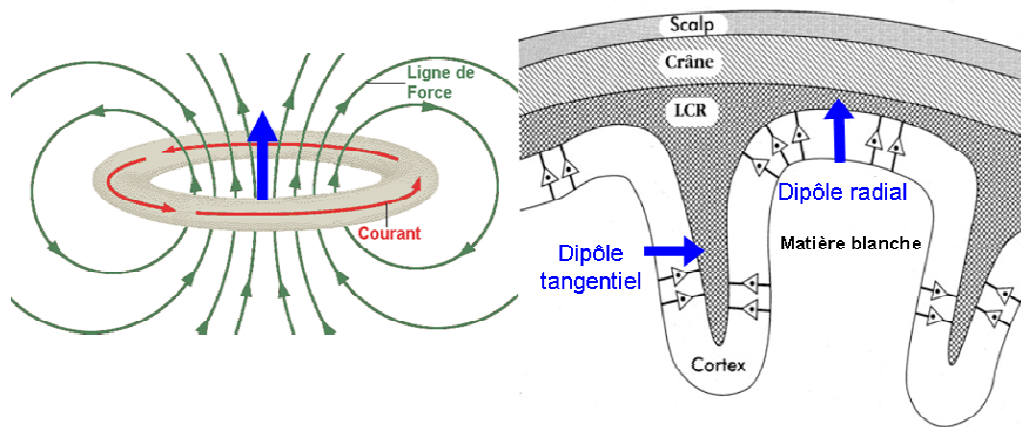


Figure 14 *À gauche* : La notion de dipôle permet de comprendre que les neurones créent une activité électrique post-synaptique (en rouge) ainsi qu'une activité magnétique (en vert). *À droite* : L'activité électrique générée par les neurones Les neurones orientés parallèlement à la surface du scalp (comme les cellules pyramidales du cortex), créent des dipôles de courant radiaux lorsqu'ils sont situés sur un gyrus ou des dipôles tangentiels lorsqu'ils sont situés sur une scissure.

Notons au passage que l'activité électrique peut également être enregistrée grâce à la magnétoencéphalographie (MEG). Elle repose sur le fait qu'un courant électrique génère un champ magnétique dont la valeur est directement proportionnelle à son intensité. Le capteur neuro-magnétomètre composé de fils supraconducteurs refroidis à l'hélium liquide permet d'enregistrer avec une grande précision les champs magnétiques provoqués par les modifications des potentiels électriques des membranes des neurones au niveau du cortex cérébral. Avec cette technique, les dipôles enregistrés sont les dipôles radiaux, parallèles à la surface du scalp alors que la technique de l'EEG permet d'enregistrer les dipôles tangentiels, perpendiculaires à la surface du scalp (Figure 14).

## b. Analyse du signal

L'analyse du signal EEG permet d'explorer les processus successifs mis en jeu lors de la perception sensorielle des stimulations d'une part mais aussi lors de l'exécution cognitive de la tâche d'autre part. Depuis les découvertes séminales de Hans Berger, des activités oscillatoires ont été documentées dans le cerveau de nombreuses espèces mammifères et dans de larges gammes fréquentielles recouvrant des oscillations très lentes ayant des périodes de plusieurs minutes jusqu'à des oscillations très rapides atteignant 600 Hz (Buzsáki & Draguhn, 2004). La première

classification fonctionnelle de ces rythmes cérébraux a été introduite par les experts de la Fédération Internationale des Sociétés pour l'Électroencéphalographie et la Neurophysiologie Clinique en 1974. Suivant la tradition de Berger, les rythmes cérébraux découverts (Figure 15b) depuis lors ont été dénommés avec des lettres grecques, et les limites entre les différentes bandes de fréquence ont été dessinées arbitrairement (delta,  $\delta$  : 0.5-4 Hz; thêta,  $\theta$  : 4-8 Hz; alpha,  $\alpha$  : 8-12 Hz; bêta,  $\beta$  : 12-30 Hz; gamma,  $\gamma$  > 30 Hz). Bien que la classification internationale conserve une importance pratique, elle regroupe potentiellement des activités ayant des significations fonctionnelles différentes suivant les structures cérébrales et les espèces concernées. Pour des raisons techniques, la classification des rythmes EEG n'a inclus que tardivement les rythmes de la bande gamma, alors que ceux-ci avaient été décrits antérieurement : en 1942, Adrian avait rapporté que l'application de substances odorantes sur la muqueuse olfactive donnait naissance à un train d'oscillations sinusoïdales pendant toute la durée de la stimulation. Cette activité induite était généralement comprise entre 30 et 60 Hz et a été interprétée comme étant le reflet de la représentation consciente des stimulations olfactives.

Un signal EEG est constitué de la superposition d'activités neuronales oscillant selon différents rythmes dont l'amplitude varie au cours du temps selon une multitude de paramètres physiologiques (état de vigilance, environnement sensoriel...) et/ou physiopathologiques (lésions cérébrales, foyers épileptiques...). Les différents rythmes constituant le signal brut de l'EEG (Figure 15a) peuvent être retrouvés par l'intermédiaire d'une décomposition fréquentielle du signal. Grâce à l'application d'une transformée de Fourier (Figure 15), les composantes fréquentielles du signal brut de l'EEG peuvent être retrouvées. Une décomposition du signal brut de l'EEG selon ses paramètres temporels et fréquentiels peut également être obtenue grâce à l'application d'une transformée en ondelettes dont la résolution temporelle varie en fonction de la résolution spectrale (voir page 58).

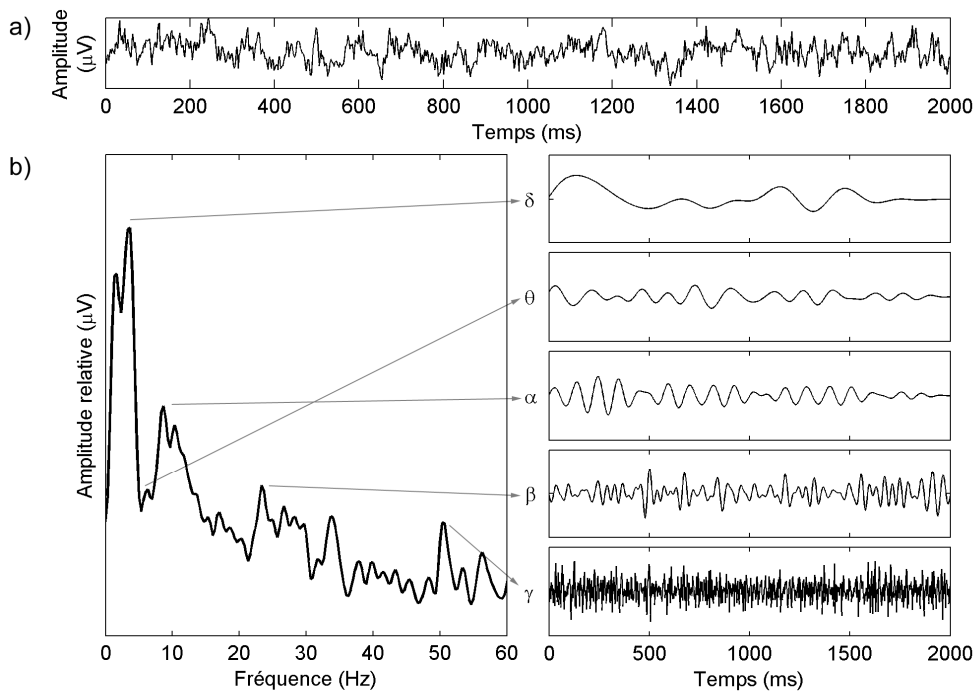


Figure 15 a) Deux secondes d'enregistrement brut de signal EEG. b) Ce signal est composé de plusieurs activités fréquentielles que l'on peut identifier grâce à une transformée de Fourier, et isoler par filtrage (à droite).

### *Analyse du signal évoqué*

Toute stimulation sensorielle provoque des modifications de l'activité cérébrale qui peut être recueillie à la surface du scalp. La technique des potentiels évoqués permet non seulement l'exploration fonctionnelle objective des systèmes sensoriels (voir Chapitre 1-2.a) mais aussi l'exploration du système cognitif, c'est-à-dire des processus corticaux. Pour que l'activité corrélée avec les stimulations se distingue de l'activité spontanée, la méthode classique consiste à répéter, dans des conditions identiques, un nombre élevé de stimulations et à effectuer la moyenne des segments du signal EEG qui suit chaque stimulation (Picton et al., 2000). Le signal aléatoire, apparu indépendamment des stimulations, s'élimine lors du moyennage, et seul le signal qui apparaît en corrélation de phase avec le début de la stimulation émerge. L'intérêt majeur d'enregistrer des potentiels évoqués réside dans le fait que les paramètres des potentiels évoqués (latence, amplitude et topographie) permettent d'aborder la dynamique des opérations mentales (Renault, Ragot, Lesevre, & Remond, 1982). La présence de composantes (pics ou potentiels évoqués) témoigne de l'activité électrique d'une ou de plusieurs zones du cerveau. Les composantes permettent de rendre compte des activités cérébrales successives sous-jacentes aux

processus de traitement de la stimulation ou de la réalisation de la tâche. Les potentiels évoqués sont caractérisés par leur polarité (notée N pour une négativité et P pour une positivité), leur latence (par rapport au début de la présentation de la stimulation) et leur amplitude (exprimée en microvolts,  $\mu\text{V}$ ). Actuellement, cette technique permet d'enregistrer, d'une manière non invasive avec les électrodes de scalp, les activités cérébrales avec une précision de l'ordre de la milliseconde dès le tronc cérébral.

Les PEA tardifs apparaissent après 70 ms et peuvent s'étendre jusqu'à une seconde après le début de la stimulation. Ces composantes sont davantage sujettes à l'état d'attention et à la vigilance du participant et sont souvent impliquées dans les processus cognitifs (mémorisation, traitement linguistique, etc.) du cerveau. Ainsi, les sources et les propriétés des composantes tardives sont d'autant plus complexes qu'elles sont associées à un traitement cérébral élaboré : en général, plus une composante de scalp est tardive, plus grand sera le nombre de ses générateurs sous-corticaux et/ou corticaux et plus complexes seront ses propriétés fonctionnelles. Compte tenu du nombre croissant d'investigations en potentiels évoqués, il est difficile de référencer toutes les composantes auditives tardives. Toutefois, nous pouvons suggérer une classification des composantes les plus souvent rencontrées, selon leur implication dans les étapes du traitement cognitif, qui vont de l'évaluation de la stimulation (Hillyard, Hink, Schwent, & Picton, 1973) à la clôture de la période cognitive (Desmedt, 1980 ; Verleger, 1988).

### ***Principaux potentiels évoqués auditifs tardifs observés en réponse aux sons de parole***

La première composante corticale, désignée par le complexe N1-P2, est considérée comme obligatoire en raison de son apparition systématique après une stimulation auditive (Figure 16a et b). Elle correspond à un double processus : respectivement une négativité à 100 ms et une positivité à 200 ms. La composante N1 auditive regrouperait l'activité des aires primaires et secondaires. Elle présente une amplitude maximum au vertex provenant de générateurs situés dans chaque hémisphère au niveau du gyrus de Heschl et du planum temporale (Howard et al., 2000 ; Liégeois-Chauvel, Laguitton, Badier, Schwartz, & Chauvel, 1995 ; Liégeois-Chauvel et al., 1994). Par l'étude de cette composante, de nombreuses études en EEG

ont confirmé la présence d'une organisation tonotopique des fréquences dans le cortex auditif (Liégeois-Chauvel, Musolino, & Chauvel, 1991 ; Vaughan & Ritter, 1970 ; Verkindt, Bertrand, Perrin, Echallier, & Pernier, 1995). Des générateurs frontaux ont également été mis en évidence (Giard et al., 1994). L'amplitude de la composante N1 semble non seulement être en rapport direct avec l'intensité du signal (Picton et al., 1974) mais également influencée par des processus attentionnels. Cette différence de localisation reflèterait une autre composante, appelée  $N_d$ , qui viendrait se superposer à la N1 (Hansen & Hillyard, 1980 ; Näätänen, Gaillard, & Mäntysalo, 1978). La composante P2 quant à elle, a pu être mise en relation avec des processus de mémorisation à court terme : lorsque la stimulation est répétée, l'amplitude de la composante P2 diminue (Picton & Hillyard, 1974).

La composante N2 (Figure 16a) est observée sur les régions centrales du scalp. Sa latence d'apparition est une mesure du temps nécessaire à la détection de la stimulation (Ritter, Simson, Vaughan, & Friedman, 1979) et son amplitude est d'autant plus grande que l'attention portée sur la stimulation est importante (Näätänen, 1990). La composante P3 (ou P300) culmine environ 300 ms après la présentation d'une stimulation et présente une topographie centro-pariétale (Figure 16a). Ces deux dernières ondes sont considérées comme endogènes du fait que leur topographie et leur comportement ne varient pas en fonction des caractéristiques physiques des stimulations elles-mêmes, mais plutôt en fonction de la tâche à effectuer. Sutton et ses collègues (1965) ont mis en évidence la composante P3 à la suite de la présentation, parmi un train de tonalités fréquentes, d'une tonalité rare sur laquelle le participant devait porter son attention. La P3 peut aussi apparaître après des stimulations de parole, souvent avec une latence plus tardive du fait que les stimulations de parole sont plus longues que les sons purs. Elle pourrait être en relation directe avec l'accès aux connaissances sémantiques, signe d'une catégorisation de la stimulation auditive (Kotchoubey & Lang, 2001). Les réponses P300 ont été montrées plus amples pour les stimulations lexicales que pour les stimulations non lexicales, notamment dans l'hémisphère gauche (Ziegler, Besson, Jacobs, Nazir, & Carr, 1997). Deux autres composantes cérébrales ont été mise en relation avec le traitement spécifique des sons de parole : il s'agit des composantes N400 et P600. La N400 (Figure 16b) est peut-être l'effet le plus étudié dans les études de potentiel évoqué portant sur l'intégration sémantique. Plusieurs études (Kutas &

Hillyard, 1980 ; McNamara, 2005 pour une revue) rapportent par exemple qu'une phrase sémantiquement incorrecte (« le chien dort dans la *tasse* »), provoque une plus grande amplitude de la composante N400 après le mot incongru comparativement à une phrase sémantiquement correcte (« le chien dort dans la *niche* »). Par ailleurs, une phrase grammaticalement incorrecte (« le chien apprend *de* sauter le mur ») provoque une plus grande positivité de la composante P600 après le mot « de » comparativement à une phrase grammaticalement correcte (« le chien essaie *de* sauter le mur »). La N400 semble donc liée à la congruence sémantique (Kutas & Federmeier, 2000), alors que la P600 à l'agrammaticalité (Osterhout & Holcomb, 1992).

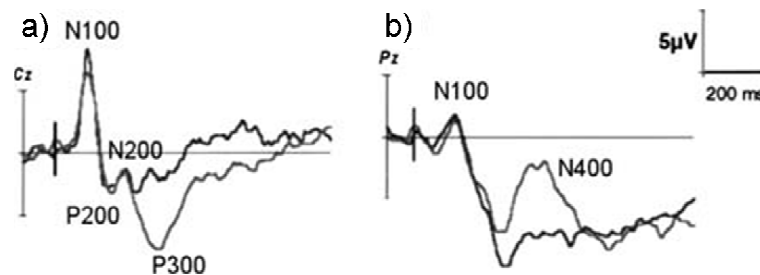


Figure 16 : a) Observation sur Cz des potentiels évoqués N100, P200, P300 et N400 suite à la présentation d'une stimulation auditive rare (en gris) comparativement à la présentation d'une stimulation auditive fréquente (en noir) pour laquelle seuls les potentiels évoqués N100, P200 et N200 sont observés. b) Suite à la présentation d'un mot sémantiquement incongru (en noir), observation sur Pz du potentiel évoqué N400 qui n'apparaît pas pour le mot sémantiquement congruent (d'après Vanhauzenhuysse, Laureys, & Perrin, 2008).

Les potentiels évoqués se révèlent particulièrement intéressants pour mettre en évidence des différences d'activités liées à l'analyse d'une stimulation auditive. Une décomposition temporelle et fréquentielle peut être réalisée à partir du potentiel évoqué (Figure 17c.). Cette représentation indique quelle activité oscillatoire est évoquée par la stimulation, en phase avec l'apparition de la stimulation. La méthode utilisée pour mettre en évidence les modifications d'activité oscillatoire consécutive à une stimulation est basée sur une décomposition temporelle et fréquentielle utilisant des ondelettes. Pour réaliser cette analyse en temps-fréquence de façon optimale, Tallon-Baudry et Bertrand (1999) ont proposé de convoluer le signal avec une ondelette gaussienne complexe de Morlet.

### *Analyse du signal induit*

Au lieu d'observer l'activité temporelle et fréquentielle moyenne évoquée par une stimulation comme le fait l'analyse du signal évoqué, l'analyse du signal induit nous renseigne sur l'évolution temporelle et fréquentielle de l'énergie du signal qui n'est pas en phase avec l'apparition de la stimulation. En effet, l'analyse du signal induit repose sur la décomposition en temps-fréquence réalisée à partir chaque essai. Elle permet ensuite de moyennner, par participant et par condition expérimentale, l'activité induite par la stimulation, c'est-à-dire l'activité qui peut être éliminée lors du moyennage des activations évoquées par la stimulation car elle n'est pas en phase avec l'apparition de la stimulation (Figure 17). À partir des représentations temps-fréquence moyennes par participant et par condition expérimentale, des analyses statistiques de groupe dans des bandes de fréquences d'intérêt sont réalisées.

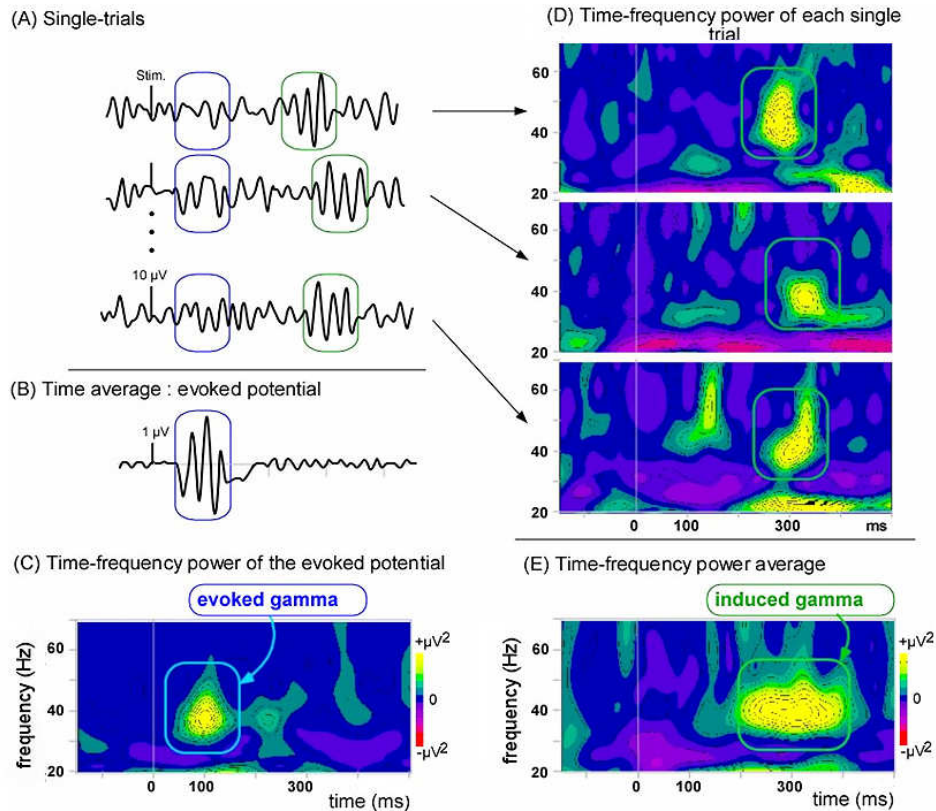


Figure 17 : Représentation schématique des réponses évoquées et induites par une stimulation après moyennage classique des signaux pour l'étude des réponses gamma (Tallon-Baudry & Bertrand, 1999). A) Les essais EEG présentent de faibles réponses gamma dont la latence d'apparition est calée après l'apparition de la stimulation (en bleu) ainsi que des réponses gamma plus amples dont la latence d'apparition est variable (en vert). B) Les méthodes classiques de moyennage mènent à la mise évidence d'un potentiel évoqué classique. C) Représentation de la réponse évoquée avec le temps en abscisse et la fréquence en ordonnée. L'échelle de couleur code la variation de puissance corrigée en ligne de base (pré-stimulation). D) Analyse temps-fréquence réalisée sur chaque essai individuellement. E) le moyennage entre les essais permet de révéler la réponse gamma induite.

Les activités induites sont décomposées selon la même classification fréquentielle correspondant à l'activité oscillatoire spontanée. Différentes fonctions ont été attribuées aux différentes bandes de fréquence. Par exemple, les activités de la bande de fréquence thêta ont été mises en relation avec les processus d'encodage sensoriels et mnésiques (pour une revue, voir Buzsáki, 2006) alors que les processus attentionnels ont été mis en relation avec les activités de la bande de fréquence alpha (pour une revue, voir Klimesch et al., 2006). Les activités observées dans la bande de fréquence bêta ont été proposées comme étant impliquées dans des synchronisations longue-distance (Bibbig, Traub, & Whittington, 2002). Les activités observées dans la bande de fréquence gamma ont été proposées comme étant impliquées dans le liage



informationnel (Engel, Fries, König, Brecht, & Singer, 1999) qui sous-tend l'hypothèse selon laquelle les synchronisations temporelles seraient à la base d'un processus intégratif :

*« The hypothesis is that synchronization of neuronal discharges can serve for the integration of distributed neurons into cell assemblies and that this process may underlie the selection of perceptually and behaviorally relevant information. »*

Le fait que les neurones oscillent de manière synchronisée dans et à travers différentes régions cérébrales, c'est-à-dire que l'activité oscillatoire neuronale corrèle en phase et/ou en amplitude dans une bande de fréquence et/ou entre bandes de fréquence, (pour une revue, voir Le Van Quyen & Bragin, 2007) pourrait de surcroît constituer une mise en relation des traitements effectués par chaque assemblée de neurones séparément. Les différentes parties cérébrales pourraient ainsi mettre en commun différentes informations sur la stimulation perçue (Engel & Singer, 2001). Les études de Crick et Koch (1990) ont même proposé que l'émergence de la perception consciente soit le résultat d'une activité synchronisée entre différentes régions cérébrales dans la bande de fréquence gamma (pour une revue plus complète, voir section 2.4.2c).

Suite à la présentation des sons de parole, les activités observées dans la bande de fréquence alpha au niveau des cortex sensoriels ont été proposées comme étant impliquées dans des processus de reconnaissance des sons de parole (Engel et al., 1999) alors que les activités observées dans les bandes d'activité bêta seraient en lien avec des opérations d'intégration linguistique (Bastiaansen & Hagoort, 2006 ; Pulvermüller, Lutzenberger, & Preissl, 1999). La perception des mots serait à ce propos caractérisée par une activité désynchronisée dans les bandes d'activité alpha et bêta, accompagnée par une activité synchronisée dans les bandes d'activité thêta et gamma au niveau temporal et frontal (Edwards et al., 2009 ; Luo, Husain, Horwitz, & Poeppel, 2005 ; Shahin, Picton, & Miller, 2009).

Alors que l'EEG nous renseigne de façon précise sur le décours temporel des activités cérébrales électriques, et ce pour différentes fréquences d'oscillations neuronales, il fournit une représentation anatomique moins détaillée. La résolution du problème inverse permet d'estimer par l'application de différents algorithmes la localisation des sources de dipôles observés à la surface du scalp (par exemple

LORETA, Pascual-Marqui et al., 1999, ou LAURA, *Local autoregressive average*, Grave de Peralta Menendez, Gonzalez Andino, Lantz, Michel, & Landis, 2001).

### 2.1.3 Les autres méthodes d'investigation fonctionnelle

D'autres techniques permettent de fournir des renseignements anatomiques et fonctionnels précis, comme la tomographie par émission de positons (TEP) ou l'imagerie fonctionnelle par résonance magnétique (IRMf) ou encore les études lésionnelles. La TEP offre des renseignements quant au fonctionnement des tissus et des organes du corps humain, comme le cerveau. Cette technique consiste à injecter un produit radioactif émetteur de positons. La désintégration radioactive de cet isotope aboutit à l'émission de rayons qui sont détectables et quantifiables à l'aide d'une caméra. Toutefois, si cette technique permet de nous renseigner sur les régions cérébrales activées au cours d'une tâche particulière avec une résolution spatiale de l'ordre de cinq millimètres, elle présente une très faible résolution temporelle. En effet, l'image obtenue en TEP correspond à une activité moyenne du cerveau, acquise pendant une période allant d'une minute à une heure. Devant le coût et l'aspect invasif de cette technique, l'IRMf s'est particulièrement développée ces dernières années. Cette technique permet une résolution temporelle inférieure à la seconde et ne nécessite pas d'injection de produit de contraste. Son principe repose sur l'enregistrement des variations d'oxygénation du sang (ou variations du signal BOLD, pour *Blood Oxygen Level Dependent*). De plus, la plus grande résolution spatiale et l'acquisition simultanée d'une IRM morphologique rendent cette technique appropriée pour l'étude des données individuelles. Cette technique permet ainsi d'imager les variations de signal d'origine vasculaire qui sont causées par les changements de l'activité mentale.

Une autre méthode consiste à étudier les altérations du comportement résultant de la lésion sélective d'une région cérébrale. Ces études dites lésionnelles apportent également de nombreuses données quant à la localisation des processus cognitifs (voir section 2.2.3) car si une structure cérébrale intervient dans une tâche, une dégradation de cette tâche doit être observée suite à la détérioration de son fonctionnement. La pathologie humaine a fourni des informations essentielles sur les relations entre le cerveau et le comportement. Broca et Wernicke ont été les premiers chercheurs à mettre en évidence les relations entre certaines parties de l'hémisphère gauche et le

langage. Les lésions cérébrales peuvent avoir pour origine des problèmes vasculaires, provoqués quand l'irrigation sanguine est interrompue, par exemple à cause d'une thrombose (réduction locale du débit sanguin à cause de dépôts graisseux sur la paroi), d'une embolie (débit sanguin diminué quand un corps étranger ou un caillot pénètre dans une artère), d'un accident hémorragique (rupture d'un vaisseau sanguin cérébral, par suite d'hypertension, d'anévrisme congénital ou de malformation artérioveineuse congénitale), des tumeurs (c'est-à-dire des masses de tissu qui se développent de façon anormale), des troubles dégénératifs ou infectieux ou des traumatismes provoqués par des accidents. La localisation des lésions peut être mise en rapport avec les troubles du comportement observés. Ces derniers sont évalués par les paradigmes de la psychologie cognitive, qui mesurent les grandes fonctions (comme le langage) ou les opérations élémentaires qui sont à la base des grandes fonctions (traitements sémantique, syntaxique par exemples).

Bien que ces techniques soient très précises quant à la localisation spatiale des activités neuronales, elles demeurent néanmoins peu précises quant au déroulement temporel des activités neuronales. Or l'étude du déroulement de processus mentaux, comme la perception ou le traitement cognitif de stimulations, nécessite des techniques simples et non invasives mettant en évidence des événements cérébraux très précis, de l'ordre de la milliseconde. La technique de l'EEG sera par conséquent employée dans ce travail de thèse afin de découvrir quels corrélats neuronaux peuvent être rattachés aux différents traitements auditifs des sons de parole.

## **2.2 Analyse cognitive des sons de parole**

L'analyse des sons de parole peut être décomposée en plusieurs étapes de traitement. La première étape de transduction sensorielle qui se déroule au niveau de l'oreille et des voies ascendantes jusqu'au cortex auditif primaire (section 1.2) aboutirait à une analyse globale des sons entendus sur la base de leurs propriétés acoustiques. L'analyse de ces propriétés permettrait ensuite de distinguer différentes classes d'objets auditifs : les sons de parole, les sons de musique ou les sons environnementaux. Concernant les sons de parole, notre système central met en œuvre plusieurs processus de traitement que ce chapitre va présenter. Les modèles

psycholinguistiques proposent d'expliquer l'accès à la compréhension des sons de parole par la comparaison au lexique mental (section 2.2.1). Deux types de modèles de reconnaissance des sons de parole seront présentés (section 2.2.2) : les modèles autonomes mettant en jeu des traitements sériels, et les modèles autonomes mettant en jeu des traitements parallèles. La compréhension des sons de parole nécessiterait une analyse intégrative, via les traitements phonologiques, sémantiques, grammaticaux et syntaxiques, qui serait assurée par de vastes réseaux associatifs distribués au sein des deux hémisphères cérébraux, et plus particulièrement à gauche (section 2.2.3).

## **2.2.1 Le concept de lexique mental**

### **a. Données de psycholinguistique**

Le concept de lexique mental, utilisé notamment en psycholinguistique, a été énoncé pour la première fois par Anne Treisman en 1960. Ce concept propose que les sons de parole soient compris grâce aux informations perçues par notre système auditif, mais aussi grâce aux connaissances linguistiques stockées en mémoire. Le lexique mental est décrit comme un système dans lequel se déroulerait l'appariement entre le signal physique de parole et le sens associé stocké en mémoire à long terme. Ce concept est assimilé à une sorte de dictionnaire mental qui contiendrait les connaissances, qu'elles soient sémantiques, orthographiques, phonologiques, syntaxiques ou pragmatiques, qu'un individu possède sur les mots d'une langue. Connaître un mot de sa langue reviendrait à connaître la relation entre une forme (visuelle ou auditive) et une signification. Le lexique mental serait organisé de façon hiérarchique ce qui permettrait de récupérer, de façon extrêmement rapide et efficace, les relations entre ces connaissances, chaque fois qu'un mot entre dans le système. Lors de la perception d'un son de parole, le lexique entre en jeu pour permettre l'accès au sens, il permet de sélectionner le bon candidat parmi des dizaines de milliers de mots en seulement quelques millisecondes (Grosjean, 1980). Les informations contenues dans le lexique mental doivent donc être organisées pour répondre à une telle vitesse de recherche. Depuis l'acceptation de l'idée de l'existence du lexique mental, les psycholinguistes n'ont cessé de s'interroger sur son mode de fonctionnement et sur son architecture interne.

## **b. Confirmation des données par la neurophysiologie**

Trois groupes de données issues des études de neurophysiologie étayent la notion de lexique mental. Premièrement, les patients ayant une aphasie de Wernicke font des erreurs de production de langage, appelées paraphasies sémantiques. Ces patients utilisent un mot d'une même catégorie sémantique à la place du mot voulu (par exemple « tulipe » à la place de « rose »). Les patients atteints de démence sémantique progressive ont des troubles sévères de dénomination et de compréhension des mots et une perte des concepts verbaux et non verbaux. Par ailleurs, ces patients ont un langage fluide, grammaticalement correct, une préservation de la répétition, lecture et écriture des mots irréguliers. Ces patients ont des difficultés à affecter des objets à une catégorie sémantique. En outre, ils donnent souvent le nom d'une catégorie alors qu'on leur demande de dire ce qu'il y a sur une image (pour un cheval ils vont dire « animal »). Le second groupe de données est issu des études de patients cérébro-lésés qui montrent des troubles sémantiques spécifiques à certaines catégories. Ces patients ont une grande difficulté à nommer des animaux, des visages célèbres ou des outils. Ces déficits corrélerent avec la lésion de régions spécifiques à gauche. Une étude réalisée en TEP (Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996) a montré que des régions différentes du lobe temporal gauche sont activées spécifiquement : le pôle temporal pour la catégorie des personnes célèbres, la partie postéro-latérale du lobe inféro-temporal pour la catégorie des animaux, la jonction temporo-pariéto-occipitale pour la catégorie des outils (Figure 18).

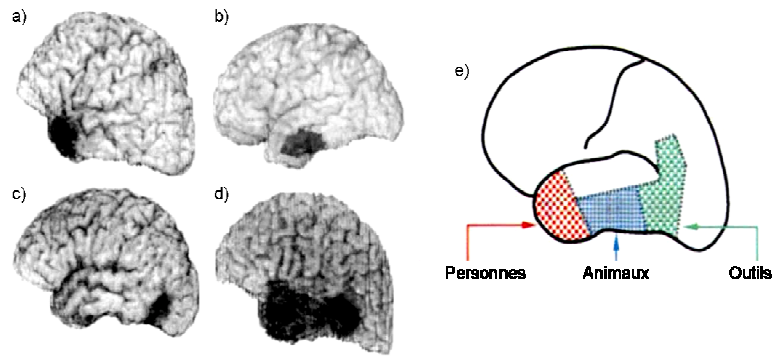


Figure 18 : Reconstruction 3D des cerveaux des patients ayant un déficit de catégorisation (d'après Damasio et al., 1996). a) Régions concernant les déficits de catégorisation des personnes célèbres, b) Régions concernant les déficits de catégorisation des animaux, c) Régions concernant les déficits de catégorisation des outils, d) Régions concernant les déficits de catégorisation des personnes célèbres, des animaux et des outils. e) Schéma récapitulatif des régions concernées par les déficits de catégorisation des personnes (en rouge), des animaux (en bleu) et des outils (en vert).

Enfin, un dernier groupe d'études liées au paradigme d'amorçage (Neely, 1977) sont en faveur du concept de lexique mental. Ces expériences montrent que le participant décide plus rapidement de la catégorie lexicale d'un mot si une relation sémantique lie les deux mots utilisés dans le paradigme d'amorçage (par exemple « banane-jaune ») que si aucune relation n'existe entre les deux mots (par exemple « banane-classeur »). Cet amorçage pourrait être dû à une propagation automatique de l'activation des mots appartenant à la même catégorie au sein du réseau sémantique (McNamara, 1992). L'électrophysiologie cognitive a mis en évidence la composante N400 qui est d'autant plus grande en amplitude que la paire de mots est incongrue (Kutas & Hillyard, 1980). Elle signerait le moment (bien avant la fin du mot) où le participant détecte la discordance sémantique.

Avant même que les données issues de la neurophysiologie apportent des corrélats anatomo-fonctionnels du traitement des sons de parole, les données issues des études de psychologie cognitive ont servi à créer des modèles explicatifs du traitement des sons de parole

## 2.2.2 Les étapes du traitement

Plusieurs étapes de traitement sont nécessaires avant d'arriver à une représentation linguistique d'un signal de parole. Le modèle classique de

compréhension des sons de parole met en jeu plusieurs niveaux : le niveau acoustique où les indices acoustiques composés de traits articulatoires sont analysés, le niveau phonologique qui correspond à la représentation abstraite des phonèmes, le niveau lexical où l'association des phonèmes forme des mots, puis les niveaux syntaxique et morphologique qui participent à la compréhension de la parole. Pour chacun de ces niveaux de traitement, l'information serait comparée au lexique mental qui contiendrait les connaissances sur les stimulations du langage.

### **a. Modèles de traitement autonome séquentiel**

#### ***La théorie de l'affinage phonétique***

La théorie de l'affinage phonétique est un modèle de perception des sons de parole (Pisoni, 1978). Trois étapes indépendantes sont nécessaires pour le traitement des sons de parole : le traitement auditif, le traitement phonétique et le traitement phonologique (Figure 19). À l'étape d'encodage du signal acoustique, le signal acoustique est converti en signal neurologique appelé *code brut* qui conserve les caractéristiques physiques du son (fréquence, intensité, structure harmonique). Au cours de l'étape du traitement phonétique, le système cognitif de l'auditeur active ses connaissances concernant la prononciation des sons. C'est au cours de cette étape que s'effectue la segmentation du signal de parole nécessaire pour délimiter les frontières phonémiques. Enfin, au cours de l'étape de traitement phonologique, l'auditeur active ses connaissances sur la manière dont la langue utilise le phonème perçu. C'est au cours de cette étape que les phénomènes de coarticulation sont résolus.

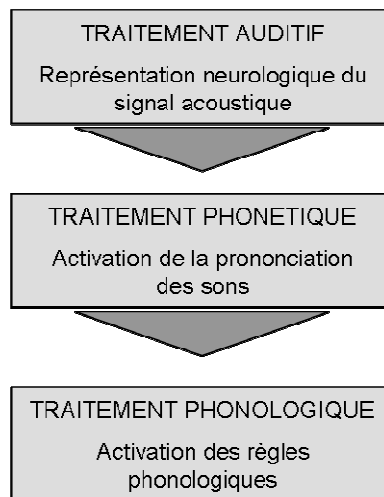


Figure 19 : Représentation du modèle de la parole proposé par Pisoni et ses collègues (Pisoni, 1978 ; Pisoni & Sawusch, 1975).

Notons qu’il n’est toutefois pas certain que la reconnaissance auditive implique nécessairement l’identification strictement séquentielle des unités qui composent le mot perçu. Pour certains auteurs, l’information sensorielle sur laquelle les procédures d’identification lexicale se basent ne correspond pas obligatoirement à la première partie du mot entendu. Ce serait plutôt l’information saillante sur le plan physique, telle que la syllabe accentuée (voir par exemple Cutler & Norris, 1988 ; Grosjean & Gee, 1987) ou la racine pour les mots morphologiquement complexes (Meunier, 1997) qui seraient utilisées pour accéder aux représentations lexicales.

### ***Le modèle COHORT***

Le modèle COHORT fait partie des modèles mettant seulement en jeu des processus cognitifs ascendants pour expliquer la reconnaissance des sons de parole. Ces modèles sont séquentiels, c’est-à-dire que le traitement des sons de parole se fait depuis les niveaux inférieurs de traitement (traitement acoustique) vers les niveaux supérieurs (traitement sémantique). Également appelé *Active Direct Access Model*, le modèle COHORT (Marslen-Wilson & Welsh, 1978) a été l’un des premiers à introduire l’idée que l’accès au lexique mental se ferait grâce à un processus de sélection d’un candidat parmi un ensemble de candidats potentiels, appelé la cohorte initiale. Selon cette théorie, le signal est analysé de façon séquentielle, c’est-à-dire au fur et à mesure de son entrée dans le système auditif. Dès les premiers phonèmes qu’il traite, l’auditeur active une cohorte de mots qui correspondent à un ensemble de candidats potentiels, c’est-à-dire commençant par les premiers phonèmes entendus.



Ensuite, au fur et à mesure que le système accumule des informations (c'est-à-dire que l'auditeur perçoit les phonèmes suivants du son de parole), les candidats qui ne correspondent plus à la stimulation acoustique sont éliminés. La reconnaissance du mot se fait lorsqu'il ne reste plus qu'un seul candidat : c'est le point de reconnaissance ou point d'unicité. Dans ce modèle, certains mots peuvent être identifiés avant que tous leurs phonèmes aient été entendus, du moment que tous les autres candidats ont été éliminés. L'existence du point d'unicité a été mise en évidence par les études menées par Grosjean (1980). L'auteur a montré qu'un mot pouvait être reconnu avant d'être présenté en entier. Par exemple, en français, le mot « crocodile », peut être reconnu dès que le deuxième /o/ est produit, aucun autre mot de la langue ne commençant par la séquence /kRoko/. Chaque mot est donc reconnu dès que l'information acoustique disponible est suffisante. Une relation entre les temps de réponse obtenus dans une tâche de détection de phonème et le point d'unicité des mots a également été mise en évidence (Marslen-Wilson, 1984).

Le modèle de la cohorte est exclusivement basé sur les propriétés acoustico-phonétiques du son de parole et met fortement l'accent sur le premier phonème du mot. Ainsi, un mot dont le phonème initial serait dégradé ou masqué ne pourrait pas être reconnu (Norris, 1994). Pour pallier à cette limite, la version II du modèle COHORT (Marslen-Wilson, 1987) met en place l'activation des candidats potentiels selon un principe de niveaux d'activation (dont le degré minimal n'est pas spécifié). Le candidat ayant la plus grande correspondance avec le signal acoustique est choisi. Par conséquent, le modèle n'exige plus que l'appariement entre le signal sonore et la représentation lexicale soit parfait, mais il prend en compte la qualité d'ajustement entre les deux. Le nombre de compétiteurs activés peut donc avoir une influence sur le temps nécessaire à la reconnaissance d'un mot particulier. Ainsi, lorsque la cohorte initiale contient un ou plusieurs compétiteurs plus fréquents que le mot-cible, le temps nécessaire à sa reconnaissance sera plus important que celui nécessaire pour reconnaître un mot de même fréquence mais dont les compétiteurs sont moins fréquents (Marslen-Wilson, 1990). Le fait que la qualité de l'appariement entre le signal acoustique et l'entrée lexicale soit exprimée par un degré d'activation permet par ailleurs de rendre compte de l'effet de la fréquence de surface en reconnaissance auditive (Taft & Hambly, 1986). Plus un mot est fréquent, plus son niveau d'activation de base est élevé. Sa reconnaissance en est ainsi d'autant plus rapide.

Bien que les applications de ce modèle puissent expliquer un grand nombre d'effets (effet de fréquence, de compétition lexicale), les données issues de la neuropsychologie cognitive permettent de mettre en évidence le fait que certains aspects des processus impliqués dans le fonctionnement de la perception auditive consciente des sons de parole ne s'exécutent pas de façon sérielle (pour des revues voir Grabowski & Damasio, 2000 ; Zatorre & Binder, 2000).

## **b. Modèles de traitement autonome parallèle**

### *Le modèle d'Ellis et Young*

Concernant la distinction entre les étapes de traitement acoustique et les étapes de traitement lexical, plusieurs cas de neuropsychologie reportés par Ellis et Young (1988) montrent que ces deux étapes sont fonctionnellement séparables. Dans le cas des agnosies auditives des sons de parole, les patients sont incapables d'effectuer une discrimination phonémique. Par conséquent, il ne leur sera pas possible d'effectuer une tâche de reconnaissance. Mais dans certains cas, l'existence d'une déficience perceptive apparaît sans déficience de la reconnaissance. Par exemple, un patient pourra facilement identifier la musique d'une chanson (comme « Au clair de la lune »), mais sera incapable de la différencier d'une autre musique (comme « Frère Jacques ») dans une tâche de discrimination. D'autres patients montrent que seul le niveau de la reconnaissance peut être affecté. Concrètement, ces patients pourront décrire très précisément le son qu'ils perçoivent mais seront incapables de l'identifier, ni même décider s'ils l'ont déjà entendu. Ainsi, un patient sera capable d'imiter le miaulement d'un chat mais incapable de le reconnaître.

L'observation de patients cérébro-lésés constitue une source de données permettant de tester les modèles existants ou de formuler de nouvelles hypothèses. Sur la base des données du champ de la neuropsychologie, Ellis et Young (1988) ont proposé un modèle mettant en jeu cinq composantes utiles à la reconnaissance des sons de parole :

- le système d'analyse auditive serait chargé d'extraire les phonèmes des sons de parole ;
- le système des entrées lexicales auditives reconnaîtrait directement les mots familiers (c'est-à-dire ayant une fréquence d'occurrence élevée) ;

- le système sémantique contiendrait les connaissances sur le sens des mots ;
- le système des sorties lexicales stockerait les connaissances sur la prononciation des mots ;
- le système des sorties phonémique fournirait les connaissances sur les sons des mots.

L'originalité de ce modèle réside dans le fait que trois voies de reconnaissance serait impliquées dans la reconnaissance des sons de parole. La première voie utilise les entrées lexicales auditives, le système sémantique et le système des sorties phonémique et représente la voie usuelle de la reconnaissance des mots. La seconde voie utilise le système des entrées lexicales auditives et le système des sorties lexicales. Les patients utilisant cette voie sont alors incapables de comprendre ce qu'ils entendent (mais peuvent comprendre ce qu'ils lisent) même s'ils peuvent le répéter (Ellis, 1984). La troisième voie utilise directement le système d'analyse auditive et le système des sorties phonémique. Les patients utilisant cette voie sont incapables de répéter un mot entendu et donnent alors un synonyme.

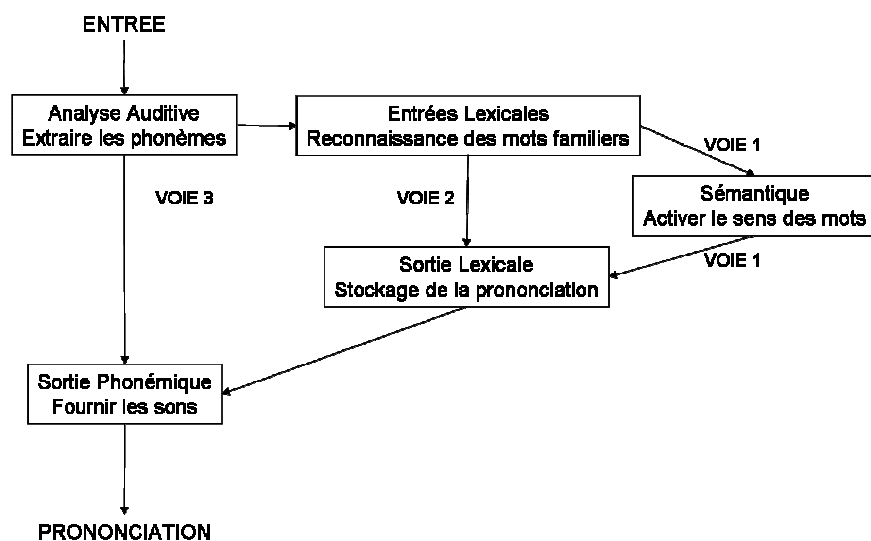


Figure 20 : Représentation du modèle d'Ellis et Young pour la reconnaissance des mots. Ce modèle postule que les mots soient reconnus grâce à la mise en œuvre de plusieurs systèmes.

Ces données sont donc en faveur d'un traitement parallèle et non sériel des indices acoustiques contenu dans les sons de parole, chaque traitement étant réalisé par une voie particulière. La critique principale de ce modèle est qu'il n'est pas explicatif, dans le sens où il décrit schématiquement les troubles neuropsychologiques

(Shallice, 1988). Par ailleurs, les données issues de la pathologie ne peuvent pas expliquer à elles seules le fonctionnement cérébral et perceptif normal. D'une part car il n'existe pas deux patients ayant le même type de lésions et le même patron de dysfonctionnement, et d'autre part car les effectifs des groupes sont faibles. Les effets ascendants ou/et *top-down* se doivent d'être abordés chez des participants ne présentant aucun trouble lésionnel ni auditif du fait de la plasticité cérébrale qui permet une réorganisation du système central en cas de lésions. Pour finir, la principale critique réside dans le fait que chaque nouvelle pathologie engendre de nouveaux composants et de nouvelles voies. Des modèles connexionnistes lui sont alors préférés.

### ***Les modèles connexionnistes autonomes***

Le modèle *Short List* (Norris, 1994) est un modèle connexionniste autonome proposant deux étapes de traitement. La première étape de traitement consiste à activer une liste réduite de candidats potentiels obtenue grâce à un système de scores d'ajustement entre le signal acoustique reçu et les entrées lexicales. La seconde étape de traitement est une étape de compétition lexicale entre les mots appartenant à la liste réduite. L'appartenance d'un candidat à la liste réduite dépend de sa congruité avec le signal. Les candidats les moins appariés au signal d'entrée sortent automatiquement de la liste réduite du fait de leur faible score d'ajustement. La liste réduite est donc constituée d'un ensemble de candidats potentiels qui, dans un second temps, s'inhibent entre eux grâce aux boucles de rétroaction qui restreignent le nombre de candidats activés. Les mots présentant les mêmes phonèmes (recouvrement phonémique) vont s'inhiber de façon proportionnelle au nombre de phonèmes qu'ils partagent. L'avantage de ce modèle est que la taille réduite du nombre de candidats potentiels additionnée à un nombre restreint de connexions relayant l'activation, permet au modèle de fonctionner avec un lexique de moindre taille. Le problème de la segmentation de la parole est résolu grâce aux processus de compétitions lexicales.

Une extension de ce modèle a été proposée par Norris, McQueen & Cutler (2000) avec le modèle *Merge* qui modélise les décisions phonémiques sans retour *top-down* du lexique sur les niveaux de traitement inférieurs. Ce modèle propose que le traitement phonologique et le traitement lexical ne se feraient pas de manière sérielle mais en parallèle. Les deux types d'information viendraient alors se rassembler à un

niveau décisionnel pour définir quel est le meilleur candidat. Le niveau décisionnel reçoit alors des informations en provenance de tous les niveaux de traitement et peut alors fusionner les informations afin de décider de ce qui a été perçu.

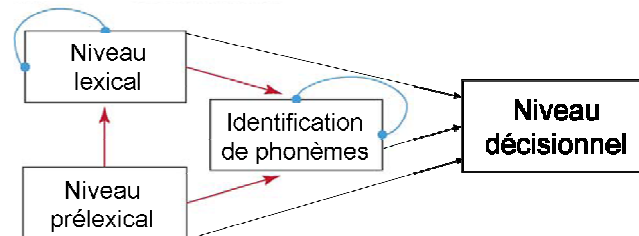


Figure 21 : Représentation schématique du modèle autonome proposé par Norris, McQueen et Cutler (2000). Les flèches rouges indiquent le sens d'activation entre les différents niveaux alors que les flèches bleues indiquent une inhibition à un seul niveau (adapté de McClelland, Mirman, & Holt, 2006).

L'approche de ce modèle permet de rendre compte d'un traitement rapide des sons de parole. Puisque les informations sont combinées de façon continue, il n'est pas nécessaire d'attendre que les analyses d'une voie soient entièrement terminées pour que les jugements puissent se réaliser. Cependant, afin d'éviter qu'une décision soit prise sans que les informations provenant directement du signal soient prises en compte, aucune activation du niveau décisionnel ne peut être réalisée en absence de support prélexical. Le problème principal de ce modèle repose sur le fait que le niveau décisionnel est activé lorsqu'une tâche directe est demandée au participant.

### 2.2.3 Les apports de la neuropsychologie et de la neurophysiologie

Historiquement, deux aires cérébrales situées dans l'hémisphère gauche ont été décrites comme étant spécialisées dans le traitement des sons de parole : l'aire de Broca et l'aire de Wernicke (Figure 22). L'aire de Broca est située au niveau du gyrus frontal inférieur alors que l'aire de Wernicke, dont le composant essentiel est le planum temporale, est située sur le gyrus temporal supérieur entre le cortex auditif primaire et le lobe pariétal inférieur. Les données de neuropsychologie (notamment des patients aphasiques) suggèrent que l'aire de Broca jouerait un rôle important dans le traitement phonologique et syntaxique et serait surtout impliquée dans la production de la parole. L'aire de Wernicke serait engagée dans l'analyse sémantique

des sons de parole ainsi que dans la représentation de séquences phonétiques. En effet, Démonet et ses collaborateurs (1992) ont montré lors d'une étude réalisée en TEP, que le cortex temporal supérieur est impliqué dans l'écoute des sons de parole, que ceux-ci aient un sens ou non. Les aires de Broca et Wernicke ne seraient pas les seules régions impliquées dans le traitement du langage. Le lobe pariétal inférieur, connecté à l'aire de Broca et à l'aire de Wernicke par d'importants faisceaux de fibres nerveuses (Geschwind, 1965 ; Price, 2000), serait également indispensable dans le traitement des sons de parole. Il se compose du gyrus angulaire et du gyrus supralimbral. Le gyrus angulaire serait davantage impliqué dans le traitement sémantique (de concert avec le gyrus cingulaire postérieur) alors que le gyrus supramarginal serait plus précisément impliqué dans le traitement phonologique et articulatoire des mots. La situation anatomique du lobe pariétal inférieur lui permet d'être en relation étroite avec les aires occipitales visuelles ainsi que le cortex somatosensoriel. De plus, le lobe pariétal inférieur contient des neurones multimodaux capables de traiter les informations visuelles, auditives et tactiles.

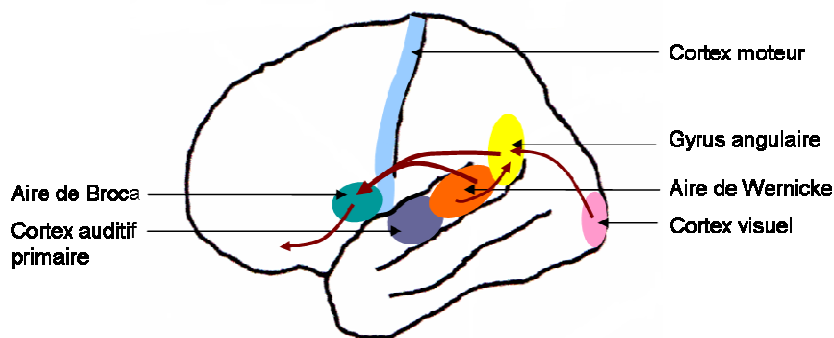


Figure 22 : Localisation de l'aire de Broca au niveau du gyrus frontal extérieur et de l'aire de Wernicke (comprenant le planum temporale) au niveau du gyrus temporal supérieur.

Les travaux de Wernicke et Broca concernant les patients aphasiques ont suggéré la prépondérance de l'hémisphère gauche dans le traitement du langage et de l'hémisphère droit dans le traitement de la prosodie. Une étude sur la dominance du langage chez des participants normaux droitiers a montré que 94 % des participants normaux avaient une dominance de l'hémisphère gauche pour le langage, aucune dominance inverse n'étant observée (Springer et al., 1999). Bien que l'activation au niveau du gyrus du planum temporale s'observe pour tous les sons, des différences latérales ont été mises en évidence : le traitement des dimensions temporelles

s'effectuant plutôt à gauche et le traitement des variations fréquentielles s'effectuant plutôt à droite (Johnsrude, Zatorre, Milner, & Evans, 1997 ; Zatorre & Belin, 2001). Il a été montré que des lésions de l'hémisphère gauche provoquent des aphasies alors que des lésions de l'hémisphère droit correspondant aux aires de Broca et de Wernicke pourraient rendre les patients aprosodiques ou amusiques. En accord avec ce dernier résultat, l'équipe canadienne de Zatorre a montré en 1992, grâce à la TEP, que l'analyse de la hauteur d'un mot activait le cortex préfrontal droit (image obtenue en soustrayant une tâche de détection de hauteur de mot à une tâche passive d'écoute de mots), alors que l'analyse de la structure phonétique finale des mots activait des structures frontales gauches (image obtenue en soustrayant une tâche de détection phonétique de mot à une tâche passive d'écoute de mots). Des résultats plus récents obtenus par IRMf ont confirmé l'implication de l'hémisphère droit, plus précisément du lobe frontal inférieur, dans le traitement prosodique des sons de parole (Buchanan et al., 2000).

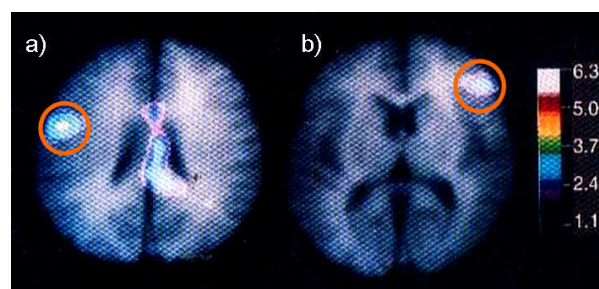


Figure 23 : Activités cérébrales issues d'une TEP (Zatorre et al., 1992). a) Activation de régions cérébrales situées à gauche lors de l'analyse phonétique. b) Activation de régions cérébrales situées à droite lors de l'analyse de la hauteur.

Les études de Binder et ses collaborateurs (2000) ont mis en évidence un modèle hiérarchique du traitement auditif des mots au niveau du lobe temporal : (1) un premier traitement, purement acoustique, serait effectué au niveau de la partie dorsale du gyrus temporal supérieur (GTS) qui est activé de façon bilatérale pour toutes les stimulations linguistiques ou non ; (2) un deuxième traitement distinguant les stimulations verbales et non verbales se ferait au niveau de la partie ventrale du GTS. Une activation bilatérale est observée au niveau de cette région mais de façon un peu plus étendue au niveau de l'hémisphère gauche. À ce niveau, un certain degré de spécialisation hémisphérique peut être observé ; (3) le troisième niveau de traitement correspondrait à la reconnaissance des stimulations lexicales auditives impliquant l'activation des informations stockées en mémoire, tels que le statut

lexical, la fonction syntaxique et le sens, et ce, indépendamment des caractéristiques physiques du mot. Les régions anatomiques responsables de ce traitement comprendraient les aires temporales adjacentes au cortex auditif, postérieurement le gyrus angulaire, ventralement les gyri temporaux moyens et inférieurs et antérieurement le pôle temporal. Une convergence de données provenant d'études différentes montre, en outre, l'importance de la portion moyenne du STS qui répond préférentiellement aux stimulations auditives verbales. Une étude en IRMf (Giraud et al., 2004) utilisant des stimulations auditives mieux contrôlées que celles utilisées lors des études précédentes, a montré que le traitement des traits acoustiques se ferait dans les régions dorsales du lobe temporal. La compréhension des sons de parole se ferait de façon bilatérale, dans les parties médianes et inférieures du lobe temporal. L'aire de Wernicke serait impliquée aussi bien dans la compréhension que dans l'analyse acoustique.

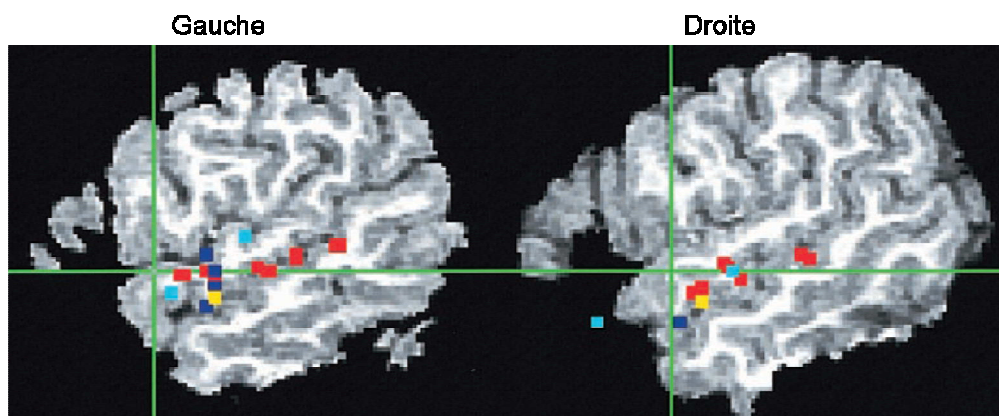


Figure 24 Résumé des activations observées pour les stimulations de parole versus les stimulations de non parole. Les résultats des études de Zatorre et collaborateurs (1992) sont reportées en bleu clair, les résultats des études de Démonet et collaborateurs (1992) sont reportées en bleu marine, les résultats des études de Binder et collaborateurs (1997) sont reportées en jaune et les résultats des études de Binder et collaborateurs (2000) sont reportées en rouge (d'après Binder et al., 2000).

Par analogie au système visuel, des études (pour une revue voir Arnott, Binns, Grady, & Alain, 2004) ont suggéré qu'il existerait deux voies différentes qui émergeraient du cortex auditif primaire, l'une ventrale et l'autre dorsale : la voie du « quoi ? » et la voie du « où ? » (ou du « comment ? »). La voie du « quoi ? » se projetterait antérieurement à A1 (le long de la scissure antérieure temporale supérieure gauche) puis dans le gyrus frontal inférieur. La voie du « où ? » se projetterait postérieurement (le long de la scissure postérieure temporale supérieure gauche) puis



dans le gyrus pariétal supérieur et enfin le gyrus frontal supérieur. Par analogie au fonctionnement du système visuel, il a été proposé que la voie du « quoi ? » soit impliquée dans la reconnaissance alors que la voie du « où ? » serait impliquée dans la localisation des sources sonores (Bidet-Caulet & Bertrand, 2005 ; Kaas & Hackett, 1999).

## **2.3 Influence non consciente des connaissances impliquées lors du traitement des sons de parole**

Analyser les sons de parole consiste à apparier le signal sonore aux représentations linguistiques stockées en mémoire. Différents modèles ont été mis au point pour tenter de décrire l'appariement entre les informations sensorielles qui parviennent au cortex auditif et les représentations lexicales. L'information acoustique parvient à l'oreille de façon séquentielle et peut commencer à être traitée avant que l'ensemble du signal sonore ait été entendu. Les modèles proposant que les processus impliqués dans le traitement des sons de parole agissent de façon ascendante défendent deux approches : l'une sérielle (par exemple *COHORT*, Marslen-Wilson, 1987) et l'autre parallèle (par exemple *Merge*, Norris et al., 2000). Cependant, des indices expérimentaux questionnent le fonctionnement ascendant. En effet, certaines données expérimentales (2.3.1) sont à même d'appuyer une théorie selon laquelle une influence *top-down* des niveaux supérieurs agirait sur de plus bas niveau de traitement. Trois exemples d'effets ont donné lieu à une discussion sur le fonctionnement de processus rétroagissant des représentations lexicales vers les représentations phonémiques : (1) la catégorisation de phonèmes ambigus, (2) la restauration phonémique, et (3) les effets des contraintes lexicales. Ces études témoignent des effets *top-down* des niveaux lexicaux vers les niveaux prélexicaux lors de l'appariement du signal acoustique aux représentations correspondantes. Une rétroaction des représentations lexicales sur les représentations phonémiques peut alors être modélisée (section 2.3.2). Cette approche interactive a donné naissance au modèle TRACE (McClelland & Elman, 1986).

### 2.3.1 Données expérimentales

Plusieurs effets expérimentaux montrent que la reconnaissance des sons de parole s'établit à l'aide de l'activation de nos connaissances (voir Remez, Rubin, Berns, Pardo, & Lang, 1994, pour une revue).

#### a. La restauration phonémique

L'un des résultats les plus souvent cités pour mettre en évidence les influences *top-down* de nos connaissances est l'effet de restauration phonémique, mis en évidence par Warren (1970) et ses collaborateurs (Obusek & Warren, 1973 ; Warren & Obusek, 1971 ; Warren & Warren, 1970). Leurs diverses expériences utilisaient des stimulations auditives dont un phonème était remplacé par un bruit. Les participants affirmaient entendre le mot intact, avec un bruit superposé. De plus, les auteurs ont mis en évidence que le contexte sémantique peut influencer l'identification du phonème perçu : si \* désigne le bruit, dans le contexte « *it was found that the \*eel was on the axle* », les participants entendent « *wheel* » (roue), alors que si « *axle* » (essieu) est remplacé par « *table* » (table), les participants perçoivent « *meal* » (repas). Warren et ses collaborateurs ont également montré qu'une partie de la stimulation, plus grande que le phonème, pouvait être remplacée par un bruit sans que les participants ne puissent le repérer. Toutefois, si un silence se substituait à la partie manquante du mot, alors les participants détectaient que le mot avait été coupé. Samuel (1996) a testé si les participants étaient capables de distinguer un bruit seul d'un phonème présenté avec un bruit superposé. Les résultats ont montré que les participants éprouvaient plus de difficultés à distinguer un bruit seul d'un phonème avec un bruit superposé dans le cas où le phonème était contenu dans un mot plutôt que dans un pseudo-mot. Cette diminution de discriminabilité perceptive entre un phonème auquel un bruit est superposé et un phonème qui a été remplacé par du bruit constituerait un indice de rétroaction du lexique mental contenant les connaissances lexicales vers les niveaux de représentations phonémiques. Ce type d'effet de restauration phonémique suggère à la fois des processus ascendants dans l'analyse du signal auditif et des processus descendants d'interprétation de ce signal.

## b. La catégorisation de phonèmes ambigus

Une démonstration de l'influence du traitement lexical sur le traitement phonémique est illustrée par l'expérience de Ganong (1980) qui a montré qu'un phonème ambigu est perçu en fonction de son contexte. Si \* désigne un phonème ambigu entre le /t/ et le /d/, il est perçu /t/ dans le contexte /\*ask/ (formant le mot *task* (tâche) par préférence au la *dask*), mais à l'opposé, il est perçu comme /d/ dans le contexte /\*ash/, (formant le mot *dash* (tiret) plutôt que le pseudo-mot *tash*). Ganong met ainsi en évidence une tendance chez les participants à donner une réponse en accord avec une interprétation lexicale de la séquence de phonèmes. Cet effet serait expliqué par une influence *top-down* des connaissances sur les niveaux de traitement prélexicaux. Cependant, l'interprétation de ce type de résultats a aussi donné lieu à une interprétation autonome (pour une revue, voir Pitt & Samuel, 1993). Plus récemment, des études réalisées en neurophysiologie ont confirmé que les informations lexicales pourraient influencer des niveaux de traitement précoces (Bonte, Valente, & Formisano, 2009 ; Gow, Segawa, Ahlfors, & Lin, 2008 ; van Linden, Stekelenburg, Tuomainen, & Vroomen, 2007 ; Myers & Blumstein, 2008). L'activation du gyrus temporal supérieur a par exemple été associée à l'influence lexicale sur un jugement sémantique (Myers & Blumstein, 2008). De plus, Gow et ses collaborateurs (2008) ont montré que l'activité oscillatoire dans la bande de fréquence gamma (= 40 Hz) du gyrus supramarginal (associée à la représentation des mots) influencerait l'activité oscillatoire du gyrus temporal supérieur postérieur (associée au traitement phonétique).



Figure 25 : Interactions prédites entre les traitements phonétiques (STG, gyrus temporal supérieur), lexicaux (SMG, gyrus supramarginal) et décisionnels (IFG, gyrus frontal inférieur) pour une explication top-down de l'effet Ganong, représentée schématiquement sur une reconstruction de surface corticale (d'après Gow et al., 2008).

### c. Les contraintes lexicales

La tâche de détection de phonèmes a permis de mettre en évidence un accès aux connaissances lexicales dans les processus d'identification phonémique. Au cours de cette tâche, un phonème particulier contenu dans une stimulation doit être détecté. Dans l'étude séminale de Cattell (1886), des mots et des non-mots ont été présentés sur une très courte durée (5-10 millisecondes). La tâche des participants était de rapporter autant de lettres qu'ils le pouvaient. L'auteur a constaté un effet de supériorité du mot (appelé *Word Superiority Effect*, ou WSE) : les performances de reconnaissance, identification ou rappel des lettres sont meilleures quand les lettres font partie d'un mot que quand elles font partie d'un non-mot - ensemble de lettres imprononçables - ou d'un pseudo-mot - ensemble de lettres prononçables - (Grainger & Jacobs, 1994 ; Grainger, Bouttevin, Truc, Bastien, & Ziegler, 2003 ; McClelland, 1976 ; Reicher, 1969 ; Wheeler, 1970) que le phonème apparaisse au début, au milieu ou à la fin des stimulations. Cette découverte suggère que des limitations de mémoire à court terme puissent être plus facilement surmontées pour des mots que pour des pseudo-mots ou des non-mots car nos connaissances lexicales aident à reconstruire le mot, et ainsi les lettres manquantes dans le mot. Cet effet peut être mis en relation avec l'effet de fréquence des mots qui apparaît dans pratiquement toutes les tâches. Plus un mot a une fréquence d'occurrence élevée, plus sa reconnaissance sera rapide (Forster & Chambers, 1973 ; Whaley, 1978). Ainsi, un mot comme « foin » sera reconnu plus lentement qu'un mot plus fréquent comme « foie ». L'effet du voisinage, consiste à observer un temps de reconnaissance plus long pour les mots comportant des voisins orthographiques que pour les mots n'en comportant pas (Grainger, O'Regan, Jacobs, & Segui, 1989). Un voisin est un mot (*foin*) pour lequel en changeant une seule lettre (*f*), on obtient un autre mot (*loin*). Les participants répondant à une tâche de décision lexicale ont des performances meilleures lorsque le mot n'a pas de voisins orthographiques que lorsqu'il a des voisins orthographiques ayant une fréquence d'occurrence plus élevée que la sienne.

De façon similaire, un effet de supériorité des pseudo-mots (*Pseudo-Word Superiority Effect*, ou PWSE) a été mis en évidence (Baron & Thurston, 1973 ; Grainger & Jacobs, 1994 ; McClelland, 1976 ; McClelland & Johnston, 1977). Cet effet se réfère aux facilitations de perception liées aux caractéristiques phonologiques de la stimulation : les performances d'identification des lettres sont meilleures quand

les lettres font partie d'un pseudo-mot que quand elles font partie d'un non-mot (Maris, 2002 ; McClelland & Rumelhart, 1981). Certains auteurs ont mis en évidence l'existence d'effets lexicaux pour des expériences dans lesquelles les stimulations présentées étaient toutes des pseudo-mots. Connine, Titone, Deelman et Blasko (1997), utilisant une tâche de détection de phonème dans des pseudo-mots, observent des temps de réaction plus rapides dans les pseudo-mots qui sont des voisins orthographiques des mots, que dans des pseudo-mots qui ne ressemblent à aucun mot. Par exemple, la détection de phonèmes est plus facile dans un pseudo-mot proche d'un mot de la langue (« foiture », proche de « voiture » et de « toiture ») que dans un pseudo-mot qui ne ressemble à aucun mot réel de la langue (« gdale »). Ceci suggère alors l'intervention de processus d'accès au lexique mental au cours du traitement des stimulations de parole qui n'ont pas de représentation lexicale. La facilitation de perception liée aux caractéristiques phonologiques des stimulations a même été observée à un niveau plus local (c'est-à-dire lorsque la stimulation n'est constituée que d'un seul caractère) : les lettres sont mieux reconnues que des pseudo-lettres ou des symboles (Burgund, Guo, & Aurbach, 2009 ; Mason, 1982 ; Reingold & Jolicoeur, 1993 ; Tydgate & Grainger, 2009). Les connaissances linguistiques amélioreraient donc les traitements de haut niveau, comme l'identification de lettre ou la discrimination lexicale des stimulations sensorielles.

Une récente étude réalisée en EEG (Coch & Mitra, 2010) a mis en évidence des effets de supériorité liés au traitement des mots et des pseudo-mots sur l'amplitude des potentiels évoqués. Des différences liées au traitement orthographique (entre des stimulations linguistiques et des non-mots) sont observées dès 150 ms, alors que des différences liées au traitement lexical (entre les mots et les pseudo-mots) apparaissent dès 200 ms et se retrouvent sur la composante N400. Les auteurs de cette étude concluent alors que l'observation de ces différences précoces serait en lien avec des activations *top-down*.

### **2.3.2 Le modèle « TRACE »**

Afin de pouvoir identifier correctement les phonèmes dans le signal de parole, l'une des solutions qui a été envisagée pour faciliter l'appariement d'un signal acoustique avec des représentations phonémiques ou phonologiques consiste à faire intervenir des procédures de rétroaction des niveaux lexicaux vers les niveaux de

représentation prélexicaux. Ainsi, des connaissances de haut niveau faciliteraient la tâche d'identification phonémique en guidant les choix perceptifs. Le modèle TRACE (Figure 26), basé sur le modèle à activation interactive de McClelland et Rumelhart (1981) établi pour la reconnaissance du langage écrit, supporte cette approche connexionniste interactive (McClelland & Elman, 1986).

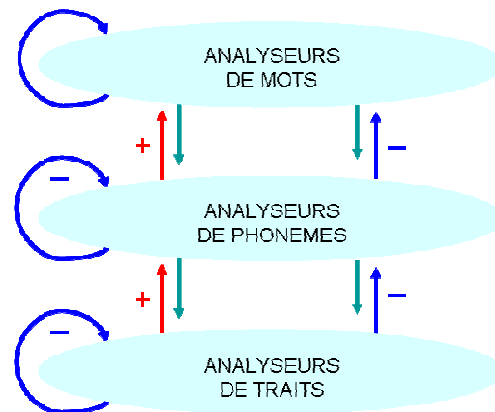


Figure 26 : Représentation schématique du modèle TRACE proposé par McClelland et Elman en 1986 pour rendre compte de la reconnaissance des sons de parole. Les liens activateurs entre les niveaux sont en rouge, les liens inhibiteurs sont en bleu et les liens représentant une influence rétrograde sont en vert.

TRACE est un modèle entièrement interactif dans lequel les informations circulent à la fois des niveaux les plus bas vers les niveaux les plus hauts et inversement. Ce modèle présente trois niveaux de représentation interconnectés : le niveau des traits acoustiques, le niveau des unités phonémiques et le niveau des mots. Les connexions entre niveaux sont excitatrices alors que les connexions entre unités d'un même niveau sont inhibitrices. Chaque unité doit atteindre un certain niveau d'activation avant de pouvoir activer d'autres unités à son tour. La décharge d'activation dépend du niveau d'activation de l'unité activatrice et de la force du lien existant entre les deux unités. L'état d'activation de chaque unité de chaque niveau à chaque instant est déterminé par les activations et inhibitions qu'elle reçoit des autres unités auxquelles elle est connectée. L'activation d'une unité au niveau des phonèmes provient des activations au niveau des traits (influences ascendantes guidées par le percept sensoriel) mais aussi des activations provenant des unités activées au niveau des mots (influences lexicales descendantes). TRACE propose donc un traitement interactif de l'information linguistique en prenant en compte l'information périphérique sensorielle mais aussi le retour des informations lexicales. En parallèle de ces activations, des inhibitions ont lieu entre les unités d'un même niveau. Un

phonème activé va inhiber les autres phonèmes et cela d'autant plus efficacement qu'il est fortement activé, et que les connexions sont fortes. Au final, les unités les plus activées entretiennent leur activation et inhibent les unités les plus faiblement activées, aboutissant à l'élimination des hypothèses lexicales les moins probables. Un mot sera reconnu lorsque son niveau d'activation domine le niveau d'activation des autres mots du lexique. Le seuil d'activation est déterminé par la règle de Luce (1959) : la probabilité de réponse d'une unité particulière est calculée en divisant la force de réponse de cette unité par la somme des forces de réponse de toutes les unités pour un temps donné.

L'explication des effets expérimentaux tels que la restauration phonémique, la catégorisation de phonèmes ambigus ainsi que la détection de phonèmes est donnée grâce à un modèle qui inclut des facilitations *top-down* des connaissances à tous les hauts niveaux de traitement des sons de parole (TRACE). Cependant, une explication autonome (Merge, Norris et al., 2000) qui implique un niveau décisionnel permettrait également d'expliquer ces effets sans intervention d'influence *top-down*. Ces deux types d'explications sont actuellement au centre d'un des principaux débats en psycholinguistique, à savoir celui de l'utilité des activations *top-down* dans les modèles de reconnaissance de la parole.

### 2.3.3 Vers un nouveau modèle

D'après les données anatomiques, neurophysiologiques, expérimentales et de modélisation rapportées dans cette partie, les connaissances influenceraient le traitement des sons de parole dans des tâches demandant un haut niveau de traitement, comme des tâches de reconnaissances ou d'identification. À notre connaissance, seulement deux études en modalité visuelle (Doyle & Leach, 1988 ; Merikle & Reingold, 1990) et aucune en modalité auditive, se sont intéressées à l'influence des connaissances lexicales ou phonologiques sur le premier niveau de traitement des sons de parole que représente la détection. Dans la modalité visuelle, Doyle et Leach (1988) d'une part, et Merikle et Reingold (Merikle & Reingold, 1990) d'autre part, ont rapporté un avantage des mots sur les non-mots pour la détection, appelé l'effet de détection des mots (*Word Detection Effect*, ou WDE). Dans l'étude de Doyle et Leach (1988), les participants devaient détecter une stimulation (un mot ou un non-mot) qui était brièvement affiché à l'écran. Les résultats suggèrent que les mots étaient plus

faciles à détecter que les non-mots. Toutefois, les deux groupes de stimulations n'ont pas été appariés en nombre de lettres, ce qui suggère que la différence de détection pourrait être due à une différence entre des caractéristiques physiques des deux ensembles de stimulations. En revanche, Merikle et Reingold (1990) ont utilisé des mots et des non-mots qui ont été appariés en nombre de lettres. Dans cette étude, un WDE est également observé : quand le délai entre l'apparition de la cible visuelle et le masque suivant était si court que les participants avaient des difficultés à détecter la cible visuelle, les mots étaient plus facilement détectés que les non-mots. À notre connaissance, aucune étude n'a observé ce genre de facilitation en modalité auditive, à savoir si l'influence des connaissances lexicales et/ou phonologiques peut s'observer sur un bas niveau de traitement que représente la détection. Les données collectées dans la modalité visuelle ne nous permettent pas de prédire les résultats pour la modalité auditive puisque les voies nerveuses visuelles et auditives sont différentes, au moins à partir des organes sensoriels jusqu'aux cortex sensoriels primaires. Ainsi, il reste à démontrer si un WDE ou, plus généralement, un effet de détection de la parole (*Speech Detection Effect*, ou SDE) peuvent également être observés dans la modalité auditive.

L'investigation expérimentale relative à cette question est présentée au Chapitre 4.

## **2.4 Traitements liés à une perception non consciente des sons de parole**

Par le fait que nous soyons immergés dans un environnement riche en stimulations sensorielles, il est rare qu'une stimulation auditive soit présentée seule, hors contexte. Au contraire, nous sommes en permanence entourés de stimulations sensorielles qui entretiennent entre elles des relations séquentielles (notamment pour les sons de parole pour lesquels la composante temporelle est déterminante) et/ou des relations parallèles (plusieurs stimulations peuvent survenir au même moment, ou être intrinsèquement liées comme la vision des lèvres et l'écoute des sons de parole). Le traitement séquentiel des informations contenues dans un contexte implique une activation permanente de nos connaissances qui vont elles-mêmes influencer par anticipation les informations suivantes. Cet effet, qui peut être facilitateur ou



inhibiteur, est observé par exemple dans les expériences utilisant des paradigmes d'amorçage (pour une revue, voir Neely, 1991). Généralement, cette influence du contexte se réalise sans rapport explicite de la part des participants. Par exemple, cet effet s'observe dans des paradigmes manipulant l'attention : de nombreuses études montrent qu'un contexte ignoré, sur lequel aucune attention n'est portée, peut activer nos connaissances et parfois même réorienter le focus attentionnel (Cherry, 1953). Les données, principalement issues des études en modalité visuelle mais aussi des études des états de conscience modifiée, suggèrent que l'influence du contexte pourrait aussi s'observer lorsque le contexte apparaît sous le seuil de perception consciente.

### 2.4.1 Influence du contexte ignoré

Le paradigme d'*amorçage* est un paradigme permettant de mettre en évidence l'influence d'un traitement d'une première stimulation sur le traitement d'une stimulation suivante. En modalité auditive, la perception d'une première stimulation amorce permet de traiter plus rapidement les caractéristiques d'une stimulation cible suivante si une relation (de type répétition, phonologique, associatif, sémantique) existe entre les deux stimulations consécutives (Grosjean & Frauenfelder, 1997). Cet effet persiste lorsque les participants focalisent leur attention sur la cible (pour laquelle une tâche est demandée) et qu'ils ignorent l'amorce. Les connaissances en relation avec le traitement de la première stimulation s'activeraient donc sans même que les participants en aient explicitement conscience. La perception d'informations volontairement ignorées a été initialement démontrée dans les situations d'écoute dichotique (Cherry, 1953) dans lesquelles un message différent est transmis à chaque oreille. La tâche demandée oblige les participants à se concentrer sur le message transmis sur une oreille et d'ignorer le message transmis à l'autre oreille. Il a été montré que des stimulations auditives présentées dans l'oreille ignorée peuvent réorienter le focus attentionnel de façon automatique à la condition que la stimulation soit pertinente pour l'auditeur. Cet effet de réorientation est appelé effet *cocktail party*. L'exemple le plus fréquemment donné est celui de la perception du propre prénom prononcé en-dehors du focus attentionnel de l'auditeur. On observe, mais seulement pour un tiers des participants (Moray, 1959), que même si les participants sont concentrés sur une autre conversation, son attention sera automatiquement redirigée vers le locuteur prononçant le prénom du participant. Plus généralement, des

stimulations auditives comportant une information pertinente (comme une composante émotionnelle par exemple) pourraient influencer la perception auditive jusqu'à réorienter l'attention sur la stimulation ignorée (Grandjean et al., 2005).

Treisman (1964) a utilisé les paradigmes d'écoute dichotique pour élaborer un modèle d'attention sélective précoce : le modèle d'atténuation. Ce modèle propose que les propriétés physiques des stimulations sont analysés en parallèle de façon pré-attentive et que seules les stimulations pertinentes, comme la stimulation cible, sont analysées de façon attentive en assignant des significations à la stimulation sélectionnée. Un filtre attentionnel viendrait bloquer les informations non pertinentes. D'autres théories ont proposé un filtrage attentionnel plus tardif, permettant une analyse plus élaborée des stimulations ignorées (Norman, 1968) voire une modulation des ressources attentionnelles, l'attention représentant une quantité fixe de ressources qui sont distribuées selon les exigences de l'activité cognitive de l'individu (Kahneman, 1973).

Les études neurophysiologiques apportent des données sur les traitements réalisées en dehors du champ attentionnel. Grâce à l'utilisation d'un paradigme *oddball* (qui consiste en une série de présentation de la même stimulation au milieu de laquelle survient de façon rare une stimulation déviante), Näätänen et ses collaborateurs (1990) ont mis en évidence un potentiel évoqué appelé *Mismatch Negativity* (MMN) qui correspond au traitement automatique d'une stimulation rares et déviante vers 150 ms, alors que les participants sont engagés dans une tâche cognitive principale. Aucune diminution des performances des participants n'a été rapporté lors du traitement cérébral des stimulations déviantes, laissant penser que le traitement de ces stimulations n'a pas interagit avec les ressources attentionnelles nécessaires à la tâche principale.

## **2.4.2 Influence du contexte non perçu consciemment**

La question de l'influence des stimulations non perçues consciemment reste encore débattue aujourd'hui. Cette question a été majoritairement étudiée sous l'angle de la modalité visuelle (pour une revue, voir Kouider & Dehaene, 2007), très peu sous l'angle de la modalité auditive. Pourtant, des études réalisées chez des patients dans

des états de conscience diminuée ou altérée laissent à penser que la perception auditive non consciente est possible en audition.

#### **a. Perceptions des sons de parole lors de l'atténuation ou de la diminution de la conscience**

De nombreuses études ont suggéré que la perception auditive, et plus particulièrement des sons de parole, perdurait pendant le sommeil. La technique des potentiels évoqués a permis de mettre en évidence que les participants endormis pouvaient détecter la présence de leur propre prénom au sein d'une série d'autres prénoms (pour une revue, voir Bastuji, Perrin, & García-Larrea, 2002). L'onde N400, qui signe la détection d'une incongruence sémantique a également été observée dans les différents stades du sommeil, suggérant donc que les mécanismes d'analyse linguistique perdurent pendant le sommeil (Perrin, Bastuji, & García-Larrea, 2002).

L'ensemble des pathologies associées à une altération de la conscience (chez le patient comateux, en état végétatif, ou en état de conscience minimale) ou à une condition apparentée (comme le *locked-in* syndrome) apporte des données quant aux capacités d'intégration cognitive (comme l'analyse du sens des stimulations) en modalité auditive. Différentes études en EEG et en neuro-imagerie ont montré qu'il était possible d'observer chez certains patients un traitement des sons de parole (pour une revue en neuroimagerie, voir Owen & Coleman, 2008 ; pour une revue en EEG, voir Vanhaudenhuyse et al., 2008). Ces études suggèrent que les mécanismes à la base de la perception des sons de parole perdurent lorsque la conscience est altérée, comme observé par exemple par Perrin et collaborateurs (2006) chez deux tiers des patients en état végétatif et chez tous les patients en état de conscience minimale (Figure 27).

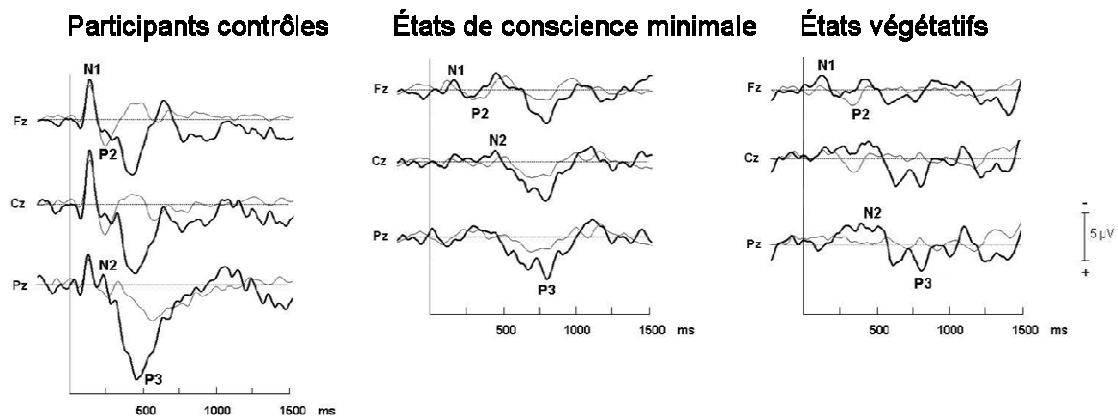


Figure 27 : Réponses N1-P2-N2 et P3 observées chez des participants contrôlés ainsi que chez des patients dont l'état de conscience est altéré : dans des états de conscience minimale et dans des états végétatifs (d'après Perrin et al., 2006).

Toutefois, montrer que des perceptions des sons de parole surviennent dans des états de conscience diminuée (comme pour le sommeil) ou altérée (dans le cas des patients) n'implique pas forcément qu'à l'état d'éveil, des perceptions auditives non conscientes existent. En effet, ces différents états physiologiques et/ou physiopathologiques sont régis par des mécanismes cellulaires et cérébraux différents ayant des conséquences différentes sur les mécanismes perceptifs.

## b. Perceptions subliminales à l'éveil

Les recherches pionnières dans le domaine de la perception subliminale utilisaient par exemple un tableau placé suffisamment loin pour que les participants ne puissent pas lire ce qui était écrit dessus. En les interrogeant, les participants répondaient qu'ils ne voyaient rien. Mais lorsque les participants étaient forcés de choisir entre deux propositions ce qui était écrit sur le tableau (par exemple une lettre ou un chiffre), les participants donnaient la bonne réponse de façon supérieure au hasard (Sidis, 1898). Au cours du siècle dernier, « Mangez du pop-corn » a sans doute été le plus célèbre message défini comme subliminal (Vicary, 1957). Projeté en 1957 dans une salle de cinéma des États-Unis, le message était intercalé entre les images d'un film et apparaissaient si brièvement que les spectateurs ne pouvaient pas le traiter consciemment. Les propriétaires du cinéma ont annoncé que leurs ventes de pop-corn avaient décuplé. Cependant, ceux qui ont voulu répliquer la manipulation ont échoué et les propriétaires eux-mêmes ont avoué la supercherie dix ans plus tard. Mais le mythe était né : il reposait sur l'idée qu'on ne peut pas contrôler les perceptions

subliminales parce qu'elles s'adressent à l'inconscient cognitif. La façon de mettre en évidence une perception subliminale est alors devenue la principale problématique.

### ***Mesurer la perception subliminale***

La façon la plus courante de montrer qu'une stimulation est perçue en absence de conscience est de mettre en évidence une dissociation entre deux mesures de perception (Erdelyi, 1985). Une première tâche permet de mesurer le niveau de conscience que les participants ont de la stimulation. Pour réduire la perception consciente de la stimulation, un masquage est classiquement utilisé en modalité visuelle (Forster & Davis, 1984). La seconde tâche permet de mesurer le niveau de traitement de la stimulation présentée lors de la première tâche. On parle de dissociation lorsque les performances du participant à la seconde tâche sont supérieures au hasard (soit  $d' > 0$ ) alors que les performances du participant à la première tâche sont égales au hasard (soit  $d' = 0$ ).

Différentes tâches ont été utilisées pour mesurer la perception consciente. Elles diffèrent à la fois par le niveau de traitement demandé et par le choix des réponses. Certaines études ont voulu démontrer l'absence de détection consciente de la stimulation (par exemple, Marcel, 1983), l'absence de catégorisation lexicale (par exemple, Kouider & Dupoux, 2005), voire l'absence d'identification (par exemple, McCauley, Parmelee, Sperber, & Carr, 1980). Certaines proposent deux choix de réponse par procédure oui/non (par exemple, Marcel, 1983), d'autres des échelles à quatre points (aucune expérience visuelle/brève expérience visuelle/image presque claire/image très claire) (par exemple, Summerfield et al., 2002). Enfin d'autres études ont suggéré qu'il était préférable de mesurer le niveau de confiance du participant dans la perception (par exemple, Cheesman & Merikle, 1984) ou de lui demander de parier sur sa réponse (par exemple, Persaud, McLeod, & Cowey, 2007). Toutefois des résultats très récents suggèrent que les mesures de perception seraient des mesures de conscience plus exhaustives que ces deux derniers index (Sandberg, Timmermans, Overgaard, & Cleeremans, 2010).

Plusieurs approches ont été proposées pour mesurer l'existence éventuelle de processus inconscients. Toutes essaient de montrer que la perception subliminale active des connaissances, et de ce fait influence le comportement ou la perception ultérieure (Frith, Perry, & Lumer, 1999). Une première approche consiste à effectuer

une mesure directe des performances des participants à la tâche de perception, par exemple à une tâche d'identification. Cette approche est souvent utilisée lors de la mise en évidence de perceptions visuelles subliminales des stimulations non langagières. Une deuxième approche consiste à mesurer indirectement les performances des participants à la tâche de perception, par exemple par l'intermédiaire d'une présentation des stimulations dans un paradigme d'amorçage. Comme pour le paradigme d'amorçage classique (Neely, 1977), le principe du paradigme d'amorçage subliminal repose sur la mesure de l'influence de la stimulation amorce sur une stimulation cible. Dans une étude princeps publiée en 1983, Marcel a rapporté la possibilité de mesurer des perceptions non conscientes grâce à l'utilisation d'un paradigme d'amorçage subliminal. Dans un tel paradigme, l'amorce est visuellement présentée pendant quelques dizaines de millisecondes, immédiatement précédée et suivie par la présentation d'un masque visuel (Figure 28). La cible est ensuite présentée de façon à ce que les participants la détectent consciemment afin de réaliser une tâche de catégorisation par exemple. Marcel a pu observer que lorsque les participants devaient déterminer si la cible était un mot ou un pseudo-mot, ils répondaient plus vite lorsque la cible était précédée par une amorce sémantiquement reliée que lorsque la cible et l'amorce n'étaient pas sémantiquement reliées. Ces premiers résultats ont cependant été remis en cause car les stimulations utilisées comme amorces ont été suspectées d'être perçues consciemment (Holender, 1986a).



Figure 28 : Paradigme de présentation subliminale masquée. À chaque essai, un premier masque constitué de chaînes de caractères précède la stimulation amorce qui est flashée durant quelques dizaines de millisecondes (entre 29 et 41 ms selon les études). En présentant un second masque immédiatement après l'amorce, on supprime la perception consciente de l'amorce (d'après Ferrand, 2001).

Ces premières controverses ont permis d'introduire des méthodes rigoureuses de l'évaluation de la perception des amorces, notamment grâce à l'application de la

théorie de la détection du signal (voir page 46). Dans ce cadre, la perception du participant est considérée comme non consciente lorsque ce dernier n'est pas capable de distinguer une stimulation d'une autre, c'est-à-dire que la valeur de son score  $d'$  est égale à zéro. Cette valeur étant difficile à obtenir, plusieurs études utilisent une méthode de régression linéaire (Greenwald, Klinger, & Schuh, 1995) qui permet de projeter les performances des participants à une valeur de  $d'$  égale à zéro. Plusieurs études ont réussi à mettre en évidence des processus sémantiques pour des mots ou des nombres pour lesquels l'absence de perception consciente était correctement contrôlée (Dehaene et al., 1998 ; Greenwald, Draine, & Abrams, 1996 ; Reynvoet, Brysbaert, & Fias, 2002). Toutefois, ces résultats peuvent être interprétés comme des effets d'amorçage non sémantique. Dans la plupart de ces expériences, des associations directes de type « stimulus-réponse » sont à l'œuvre et court-circuitent l'analyse sémantique (Damian, 2001). Ainsi, Abrams et Greenwald (2000) ont montré que leurs effets d'amorçage masqué étaient entièrement expliqués par le fait que les mots amorces, qui étaient également utilisés comme mots cibles dans d'autres essais, étaient associés à un code de réponse. Dès lors qu'ils utilisaient comme stimulations amorces des mots jamais utilisés comme cibles, les effets d'amorçage disparaissaient. Cette démonstration de la nature non sémantique des représentations des stimulations amorces masquées fut complétée en montrant qu'après avoir catégorisé les mots « *smut* » (saleté) et « *bile* » (bile), comme négatifs sur le plan émotionnel, le mot « *smile* » (sourire), constitué de fragments des mots « *smut* » et « *bile* », amorçait la réponse négative, et non, comme le voudrait son sens, la réponse positive. Toutefois, quelques récentes études ont montré un effet d'amorçage sémantique inconscient visuel satisfaisant toutes les critiques, notamment celles de l'absence de répétition et de l'utilisation d'une valeur  $d'$  égale à zéro pour la mesure du niveau de conscience de l'amorce (par exemple, Dell'Acqua & Grainger, 1999 ; Kiefer & Brendel, 2006).

### **c. Corrélats neuronaux de la perception consciente et non consciente**

Si les bases cérébrales des processus conscients et inconscients ne sont pas encore expliquées, de nombreux travaux ouvrent des perspectives d'explications très intéressantes. Sur la base d'études de neuropsychologie, une première hypothèse d'explication du fonctionnement des traitements non conscients a été de proposer que les informations traitées inconsciemment n'activeraient pas les mêmes réseaux

neuronaux que les informations traitées consciemment. C'est ce qui a notamment mis en évidence grâce aux études de neuropsychologies qui reportent des pathologies telles que la vision aveugle (Weiskrantz, Warrington, Sanders, & Marshall, 1974), l'héminégligence (Vuilleumier & Schwartz, 2001) ou la prosopagnosie (Damasio, Damasio, & Van Hoesen, 1982). Ces différentes études nous apprennent que lorsque le patient est forcé à réaliser une tâche alors qu'il n'est pas conscient de la cible à traiter, ses performances sont toujours supérieures au hasard. La différence entre une perception consciente et inconsciente résiderait alors dans l'activation de réseaux cérébraux différents : la voie ventrale (occipito-temporale) permettrait l'élaboration d'une représentation consciente alors que la voie dorsale (occipito-pariétale) permettrait l'élaboration d'une représentation non consciente (Goodale, Milner, Jakobson, & Carey, 1991 ; Milner & Goodale, 2008). Le paradigme de rivalité binoculaire a permis de montrer cette spécificité chez des participants sains (Fang & He, 2005 ; Tong, Nakayama, Vaughan, & Kanwisher, 1998). Cependant, cette dissociation est encore discutée à la vue de résultats contradictoires (Dehaene & Naccache, 2001). D'autres hypothèses peuvent être avancées. Selon Dehaene et ses collaborateurs (2006), la différence entre une perception consciente et une perception inconsciente reposerait davantage sur la modulation de l'activation de structures cérébrales communes plutôt que sur une dichotomie anatomo-fonctionnelle (Cohen et al., 2000 ; Dehaene & Naccache, 2001). De plus, il serait nécessaire que ces réseaux cérébraux, éloignés anatomiquement mais fortement interconnectés, présentent une activité cohérente pour que la stimulation émerge à la conscience (Dehaene & Changeux, 2004). Cette dernière hypothèse suggère qu'une synchronisation cérébrale serait indispensable pour réaliser la conjonction des différentes analyses élémentaires d'une stimulation, qui sont anatomiquement distribuées (Singer & Gray, 1995). La signature neuronale de la perception non consciente correspondrait à une propagation locale de l'activité neuronale le long des voies sensorielles, tandis que la perception consciente nécessiterait une coordination globale à longue distance de l'activité neuronale (Dehaene et al., 2006). Par exemple, Melloni et ses collaborateurs (2007) ont présentés, dans une étude d'EEG, des résultats qui vont dans le sens de cette hypothèse. Les mots perçus et non perçus consciemment induisaient une augmentation locale similaire de l'activité gamma, mais seuls les mots perçus consciemment induisaient une synchronisation à longue distance de l'activité gamma entre des régions cérébrales largement séparées (Figure 29). D'autres études sont en



accord avec ce résultat (Meador, Ray, Echauz, Loring, & Vachtsevanos, 2002 ; Rodriguez et al., 1999 ; Wyart & Tallon-Baudry, 2008).

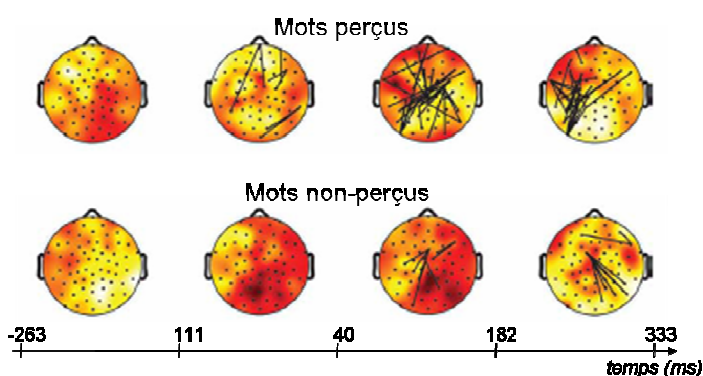


Figure 29 : Synchronisation de l'activité gamma entre les différentes électrodes de scalp pour les mots perçus consciemment (en haut) et les mots non perçus consciemment (en bas).

Si de nombreuses études suggèrent que la représentation consciente d'une stimulation nécessiterait la synchronisation des structures participant à la perception dans la bande de fréquence gamma, d'autres études ont montré que les activités dans les bandes de fréquence alpha et bêta pourraient également jouer un rôle (Gaillard et al., 2009 ; Dan Glaser & Scherer, 2008 ; Luo et al., 2005 ; Palva, Linkenkaer-Hansen, Näätänen, & Palva, 2005). C'est le cas par exemple des études portant sur le phénomène de capture attentionnelle. L'incapacité à détecter la seconde cible a été associée avec une synchronisation de phase réduite dans la bande de fréquence bêta dans un réseau à large échelle incluant les régions fronto-pariétales (Gross et al., 2004 ; Nakatani, Ito, Nikolaev, Gong, & van Leeuwen, 2005).

Le concept d'espace de travail global a été initialement développé dans le but d'expliquer comment une perception émerge à la conscience. Ce modèle s'appuie sur la présence de plusieurs systèmes spécialisés qui accomplissent chacun leur tâche indépendamment. L'accès à la conscience serait rendu possible lorsque ces différents systèmes mettent en commun certaines informations dans un même « espace de travail global ». La mise en commun des informations issues des différents systèmes dépendrait de la synchronisation transitoire d'assemblées neuronales largement distribuées (Engel et al., 1999 ; Thompson & Varela, 2001). Dehaene et ses collaborateurs (2006) proposent que l'espace de travail global s'articule autour d'un ensemble de neurones pyramidaux. En effet, ces neurones possédant de longs axones permettraient une connectivité étendue en reliant entre elles différentes aires

corticales. Dans ce modèle, l'attention joue un rôle déterminant d'amplification descendante sur l'activité neuronale des différentes aires corticales. Ce mécanisme permettrait de mobiliser les neurones et/ou de maintenir une activité suffisamment élevée. L'existence d'un seuil d'activité à franchir est aussi un point central dans ce modèle qui permet alors à l'activité neuronale de générer une représentation consciente. Comme nous l'avons vu précédemment, ce modèle connexionniste permet de décrire un continuum d'états de conscience sur lequel se distinguent trois états en particulier (Figure 30) :

- l'activation n'est pas suffisante pour déclencher un état d'activation générale dans le réseau : les stimulations sont traitées localement, de façon subliminale (au sens étymologique, sous le seuil de conscience);
- l'activation est suffisante pour déclencher un état d'activation générale dans le réseau mais l'attention ne remplit pas son rôle d'amplification descendante; c'est un niveau de préconscience;
- l'activation est suffisante pour déclencher un état d'activation générale dans le réseau et l'attention remplit son rôle d'amplification descendante; les stimulations sont traitées de façon consciente. L'espace de travail global est activé car les informations préconscientes ont reçu suffisamment d'amplification attentionnelle pour franchir le seuil de la conscience.

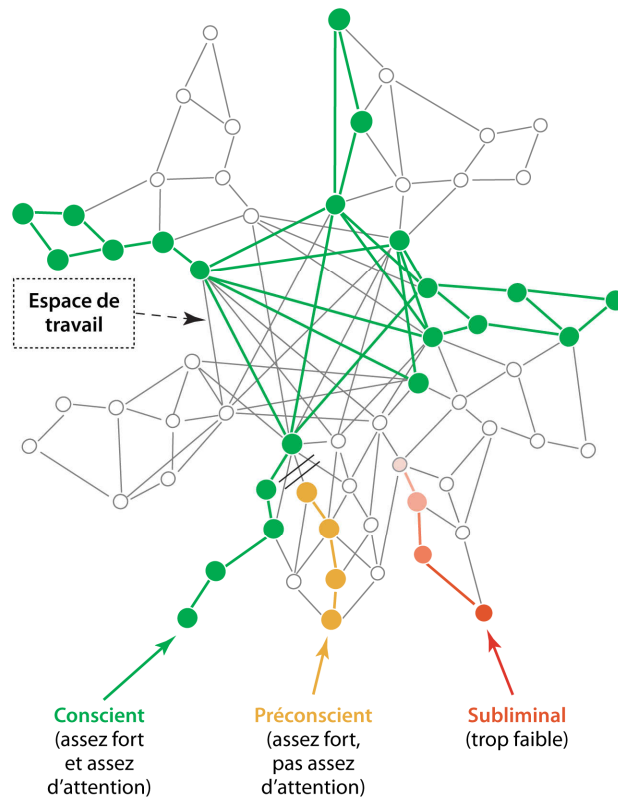


Figure 30 : Représentation schématique du fonctionnement de l'accès à la conscience dans un modèle neuronal d'espace de travail global (d'après Deheane et al., 2006).

#### d. Données en modalité auditive

En modalité auditive, Urban (1992) a rapporté les études qui ont tenté de mettre en évidence une influence des perceptions auditives non conscientes sur la perception auditive consciente ou sur des processus d'apprentissage et de mémoire. Plusieurs techniques de présentation sonore ont été testées : la diminution de l'intensité de la stimulation, la présentation de stimulations à des fréquences inaudibles par l'oreille humaine, la présentation inversée des stimulations sonores (*reversed-speech*). Aucune des études rapportées n'a mis en évidence des signes de traitement auditif non conscient. Deux autres études (Lindner, 1968 ; Shipley, 1965) ont présenté des sons purs et déterminé si une dissociation était présente entre leur détection et leur reconnaissance. Là aussi, des résultats contradictoires ont été observés : alors que Shipley (1965) ne reporte aucune reconnaissance sans détection, contrairement à Lindner (1968). Plus récemment, Kouider et Dupoux (2005) ont mis au point un paradigme d'amorçage subliminal auditif pour des sons de parole. Par analogie au paradigme d'amorçage masqué utilisé en modalité visuelle, un masquage auditif

entourait la présentation d'une stimulation amorce compressée à 65 %. Immédiatement après la fin de la stimulation amorce était présentée une stimulation cible clairement discriminable (plus forte de 15 dB par rapport à l'intensité de la stimulation amorce et des masques auditifs). Les auteurs ont contrôlé le niveau de conscience de la perception de l'amorce en proposant deux post-tests sur l'intelligibilité de l'amorce. Ces tests comprenaient une tâche de décision lexicale et une tâche de décision de parole (entre parole et bruit) sur les amorces utilisées. Ils ont montré, par l'utilisation de la méthode de régression linéaire, que lorsque l'amorce compressée et la cible avaient une relation de répétition, les temps de réponse des participants étaient plus rapides. Aucun effet d'amorçage subliminal sémantique n'a en revanche été mis en évidence. Cette étude est la seule à ce jour et à notre connaissance, montrant l'influence du traitement d'une stimulation subliminale auditive sur la perception consciente auditive. Toutefois, les voies de traitement auditif et visuel étant différents, au moins du système périphérique jusqu'au niveau cortical, les techniques utilisées en modalité visuelle ne peuvent pas être transposées à la modalité auditive. Aussi, la compression des stimulations auditives ne représente pas une situation naturelle d'écoute, ce qui pourrait expliquer que l'accès à la représentation sémantique soit difficile. En manipulant différemment les caractéristiques physiques des stimulations auditives dans le but de proposer une présentation écologique, nous pouvons penser que des mécanismes de perception inconsciente, équivalents à ceux observés en modalité visuelle, pourraient être obtenus en modalité auditive.

Les études proposées au Chapitre 5 tentent d'apporter des éléments de réponse quant à l'observation d'un traitement auditif subliminal.



# Chapitre 3

## Problématique

L'étude des processus impliqués dans le traitement des stimulations auditives est encore peu abordée en psychologie cognitive. Pourtant, la voie auditive représente une modalité majeure d'accès aux informations utiles à la communication humaine. La perception des sons de parole nécessiterait de multiples processus d'analyses du signal qui se dérouleraient de façon sérielle et/ou parallèle ainsi que des processus de comparaison aux connaissances (phonologiques ou sémantiques par exemple) organisées dans le lexique mental. Deux autres processus pourraient intervenir dans le phénomène de perception. D'une part, les connaissances pourraient influencer le traitement d'une stimulation isolée (c'est-à-dire sans qu'elle soit amorcée par d'autres stimulations ou précédée d'un contexte particulier), si des connaissances sont détenues sur cette stimulation isolée. D'autre part, l'activation des connaissances pourrait influencer l'analyse des stimulations suivantes via des processus *top-down*, si les stimulations partagent des caractéristiques (phonologiques ou sémantiques par exemple) communes. L'influence des connaissances liées au contexte, sur la perception, pourrait se dérouler lorsque l'attention n'est pas dirigée sur les stimulations du contexte, voire même lorsque la perception du contexte n'est pas consciente. Afin de mieux comprendre ces deux processus non conscients qui influencent la perception auditive, des études comportementales et électroencéphalographiques ont été mises en œuvre au cours de ce travail doctoral.

Le premier axe de recherche a consisté à définir l'influence *top-down* des connaissances sur le traitement de stimulations auditives isolées. Peu d'études ont montré l'influence *top-down* des connaissances en modalité visuelle sur le premier niveau de traitement que représente la détection de stimulations présentées de façon isolée. De plus, aucune étude en modalité auditive ne s'est intéressée à l'influence de nos connaissances lexicales ni à l'influence des connaissances phonologiques sur le premier niveau de traitement que représente la détection. Notre premier axe de recherche (Chapitre 4) propose de déterminer si les connaissances de l'auditeur,

qu'elles soient lexicales ou phonologiques, peuvent avoir un effet précoce sur les premiers niveaux de traitement auditif, notamment sur la détection des événements sonores complexes. Si nos connaissances lexicales et phonologiques ont une influence sur la détection, alors (i) les performances de détection des mots seront supérieures à celles des pseudo-mots, et (ii) les performances de détection des mots et des pseudo-mots seront supérieures de celles des sons complexes. Une première étude a consisté à présenter, à différents niveaux d'intensité sonore des mots, des pseudo-mots et des sons complexes sans composante phonologique (Étude I) et à demander aux participants de réaliser une tâche de détection de stimulations. Deux autres expériences dans lesquelles certains paramètres expérimentaux (tels que les caractéristiques énergétiques des stimulations, l'ordre de présentation, les conditions d'écoute) ont été manipulés et ont permis d'affiner les résultats (Étude II). Un point fondamental permettant de réaliser une comparaison entre les différentes natures de stimulations a été de s'assurer que les stimulations auditives ne différaient que par leur aspect lexical et/ou phonologique. De nombreux paramètres physiques et phonologiques ont été contrôlés afin d'utiliser des stimulations auditives comparables dans les différentes conditions expérimentales. Cet important travail de psychoacoustique a été réalisé lors de la création du matériel utilisé mais aussi *a posteriori*.

Le second axe de recherche consiste à rechercher si la présentation subliminale d'une stimulation auditive peut activer les connaissances et donc influencer la perception ultérieure (Chapitre 5). Alors que de nombreuses études ont montré qu'une stimulation visuelle pouvait être analysée inconsciemment et influencer le traitement des stimulations ultérieures, très peu d'études se sont intéressées à la modalité auditive. De plus, la question des mesures utilisées pour définir le niveau de conscience, la question des tâches et des paradigmes les plus appropriés, la question du niveau de traitement atteint suite à une perception subliminale et enfin la question des corrélats neuronaux d'un traitement inconscient sont encore débattues. Sur la base de ce qui a été observé en perception auditive dans les états de conscience diminuée ou altérée, des signes de traitement sémantique suite à une perception subliminale auditive peuvent être attendus, lorsque les participants ne sont pas capables de catégoriser consciemment la stimulation subliminale. Nous nous proposons d'examiner l'existence d'une perception auditive subliminale en prenant en compte

les spécificités du système auditif (notamment la composante temporelle des sons de parole). De plus, cette problématique a été abordée à l'aide de différentes mesures (directes/indirectes), différents indices du niveau de conscience (détection/catégorisation), différentes approches (objectives/subjectives) permettant d'évaluer les performances, différentes natures de stimulations (mots/pseudo-mots/sons complexes) et différents contextes d'écoute. L'Étude II précédente a permis de rechercher d'éventuelles dissociations entre la détection des mots, des pseudo-mots et des sons complexes, et leurs reconnaissances (mesure directe). Après avoir évalué précisément les seuils de détection et de catégorisation (Étude III), la perception auditive inconsciente a été évaluée par une mesure indirecte de la perception, à savoir lors de la réalisation d'une tâche de décision lexicale dans un paradigme d'amorçage dans lequel la stimulation amorce n'est pas catégorisée consciemment (Étude IV). L'Étude V se propose d'examiner, sur la base des résultats comportementaux précédents, les corrélats électroencéphalographiques associés à une perception auditive consciente ou non consciente. Dans cette dernière étude, des silences ainsi que des mots, pseudo-mots et sons complexes sont présentés à deux niveaux d'intensité sonore différents. Un premier niveau d'intensité sonore permettait une catégorisation consciente de la stimulation auditive présentée (stimulations catégorisables), et un second niveau d'intensité sonore ne le permettait pas (stimulations non catégorisables).

L'aspect pluridisciplinaire de ce travail de thèse, mis en avant dans la partie théorique, sera également présent dans la partie expérimentale. Nous verrons que l'approche expérimentale en modalité auditive soulève des questions scientifiques aussi bien acoustiques que psychologiques et physiologiques. Les résultats présentés dans ce travail de thèse sont issus de la combinaison des méthodologies expérimentales de ces différentes approches qui se sont révélées complémentaires. L'ensemble des résultats sera discuté à la lumière d'études antérieures provenant des trois disciplines. Les conclusions obtenues proviendront de la confrontation entre ces diverses disciplines et n'auraient pas pu être envisagées sans cette vision interdisciplinaire.





## Chapitre 4

# Explorations de l'influence des processus *top-down* sur la détection des sons de parole

L'influence des connaissances sur la détection de stimulations présentées de façon isolée, c'est-à-dire indépendamment du contexte, a été démontrée en modalité visuelle pour la première fois par Catell (1886). En revanche, seulement deux études ont étudié l'influence possible des connaissances lexicales ou phonologiques (voir section 2.3) sur les stades précoces du traitement sensoriel (Doyle & Leach, 1988 ; Merikle & Reingold, 1990). À notre connaissance, aucune étude n'a étudié ce genre de facilitation dans la modalité auditive. En particulier, aucune étude n'a étudié l'influence des connaissances phonologiques sur la détection, à savoir si les pseudo-mots sont mieux détectés que des stimulations de non-parole (stimulations complexes ne comportant pas d'information phonologique). Les données collectées dans la modalité visuelle ne nous permettent pas de prédire les résultats pour la modalité auditive puisque les voies nerveuses visuelles et auditives sont différentes, au moins à partir des organes sensoriels jusqu'aux cortex sensoriels primaires. Ainsi, il reste à démontrer si un effet de détection des mots ou, plus généralement, un effet de détection de la parole peuvent être observés dans la modalité auditive.

Les deux études de ce premier axe de recherche analysent l'influence des connaissances phonologiques et lexicales sur la détection auditive, en utilisant trois types de stimulations : des mots, des pseudo-mots et des sons complexes sans composante phonologique. Afin de minimiser les différences de traitement dues à des caractéristiques sensorielles, les propriétés énergétiques des trois types de stimulations ont été égalisées le plus finement possible (en termes d'intensité, durée, enveloppe temporelle, et spectre moyen). En outre, les mots et les pseudo-mots étaient composés des mêmes phonèmes afin de contrôler les différences phonologiques. Ces différents

contrôles permettent ainsi de s'assurer que d'éventuelles différences de détection entre les stimulations soient seulement dues aux caractéristiques phonologiques et sémantiques. Dans ces conditions, nous avons émis l'hypothèse que si l'influence de nos connaissances n'atteint pas les niveaux de traitement les plus bas, le seuil de détection auditive ne doit pas différer entre les stimulations phonologiques (mots et pseudo-mots) et non phonologiques (sons complexes) et/ou entre les stimulations lexicales (mots) et non lexicales (pseudo-mots). En revanche, si les connaissances phonologiques et/ou lexicales atteignent les niveaux les plus bas de traitement, le seuil de détection auditive doit être amélioré pour les stimulations phonologiques (mots et pseudo-mots) par rapport aux stimulations non phonologiques (sons complexes) et/ou pour les stimulations lexicales (mots) par rapport aux stimulations non lexicales (pseudo-mots). Dans les deux études complémentaires que nous proposons ci-dessous, nous avons manipulé deux paramètres expérimentaux afin de connaître leur influence sur la détection : le type d'égalisation acoustique (dB-SPL ou dB-A) et l'attention (présentation aléatoire ou par blocs, présence ou non d'une deuxième tâche après la tâche de détection). Aucune hypothèse *a priori* ne peut être émise par rapport à l'influence de ces deux paramètres expérimentaux car aucune étude, à notre connaissance, n'a examiné la détection des sons de parole lexicaux, comparativement aux sons de parole non lexicaux ou aux sons complexes sans composante phonologique.

## **4.1 Étude I : Performances de détection des mots, pseudo-mots et sons complexes**

L'influence de la nature de la stimulation sur le premier niveau de traitement que représente la détection a été montrée lors de la présentation de stimulations visuelles. Merikle et Reingold (1990) ont brièvement présenté des mots dans une condition et des non-mots dans une autre condition chez les mêmes participants. La tâche des participants était d'effectuer une détection de stimulations visuelles par une procédure oui/non avant de réaliser une tâche de reconnaissance. Les résultats montrent que les participants obtiennent une meilleure détection pour les mots que pour les non-mots. Alors que Merikle et Reingold (1990) observent un effet de supériorité de détection des mots en modalité visuelle, aucune étude en modalité

auditive n'a testé ce résultat. Nous proposons donc à travers l'Étude I d'examiner une possible répliation en modalité auditive des résultats obtenus en modalité visuelle. Pour cela, trois conditions de stimulations (des mots, des pseudo-mots et des sons complexes sans composante phonologique) seront présentées par blocs séparés à différents niveaux d'intensité sonore, allant de l'inaudible (à -10 dB-SPL) au clairement audible (à 35 dB-SPL).

#### 4.1.1 Méthodologie

**Participants.** 30 participants (âge moyen de 22.28 ans  $\pm$  3.84, 24 femmes) de langue maternelle française et droitiers (quotient de latéralité moyen = 93.94  $\pm$  6.71) selon le test "Edinburgh Handedness Inventory" (Oldfield, 1971), ont été inclus dans cette expérience. Aucun participant n'a rapporté de problèmes auditifs ou neurologiques. Ils ont été sélectionnés sur la base de leur audition qui a été testée grâce à une audiométrie tonale réalisée pour chaque oreille : leur perte auditive est inférieure à 15 dB pour les fréquences de sons purs compris entre 250 et 8000 Hz (ANSI, 2004).

**Matériel.** Hormis le fait que les stimulations soient égalisées en dB-SPL, le matériel et l'appareillage utilisés au cours de cette étude sont similaires à ceux utilisés dans l'Étude II.

**Procédure.** Chaque participant a passé trois blocs expérimentaux comprenant 100 stimulations sonores. L'ordre de présentation des blocs a été aléatoirement distribué de façon à ce qu'un tiers des participants entendent les sons complexes, les pseudo-mots puis les mots, un deuxième tiers entendent les pseudo-mots, les mots puis les sons complexes, et un dernier tiers entendent les mots, les sons complexes puis les pseudo-mots. Les participants étaient installés dans une salle insonorisée. À l'apparition de la croix de fixation était présentée une stimulation auditive. Les participants réalisaient une tâche de détection par une procédure oui/non et devaient répondre le plus rapidement et le plus correctement possible. Un nouvel essai commençait 300 ms après la réponse motrice à la tâche de détection. Les stimulations étaient présentées à un niveau aléatoire d'intensité, allant de -10 à 35 dB-SPL par pas de 5 dB. Dix stimulations étaient présentées par condition. Aucune stimulation n'était

répétée pour un même participant et toutes les stimulations ont été présentées à chaque niveau d'intensité entre les participants.

## 4.1.2 Résultats

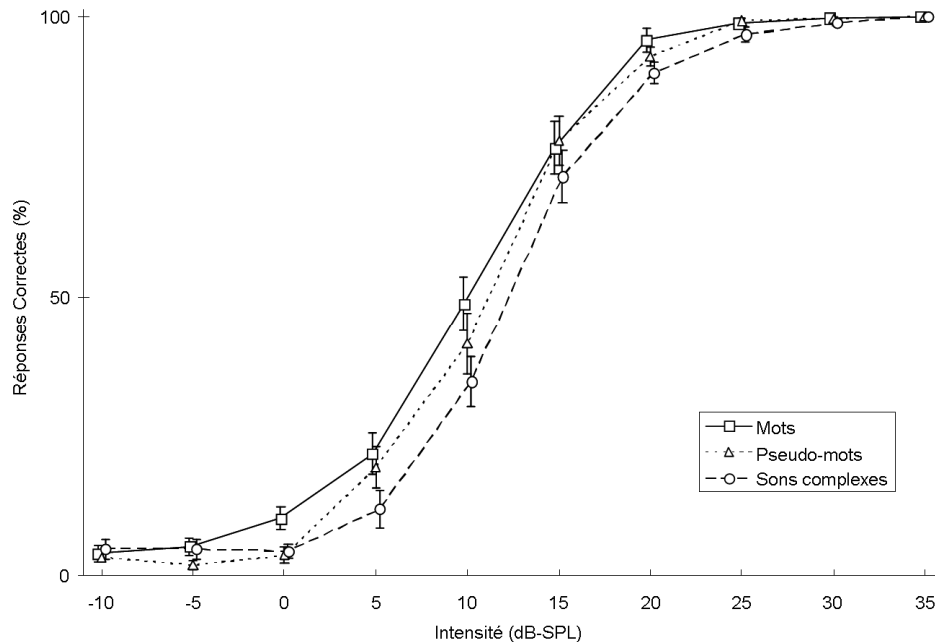


Figure 31 : Le pourcentage de réponses correctes à la tâche de détection est meilleur pour les mots que pour les pseudo-mots, et est meilleur pour les pseudo-mots que pour les sons complexes.

Une analyse de variance (ANOVA) a été effectuée sur les performances de réponses correctes selon la nature de la Stimulation (mots, pseudo-mots, sons complexes) et le niveau d'Intensité (10 intensités, de -10 à 35 dB-SPL). L'ANOVA montre un effet principal de la nature de la Stimulation ( $F(2,60) = 10.22, p < .001$ ), un effet principal du niveau d'Intensité ( $F(8,240) = 435.96, p < .001$ ) ainsi qu'une interaction significative entre ces deux facteurs ( $F(16,480) = 3.11, p < .001$ ). Les analyses *post-hocs* (*Least Significant Difference*, LSD Fisher) révèlent que les performances de détection sont meilleures pour les mots que pour les pseudo-mots ( $ps = .029$ ) entre 0 et 20 dB-SPL, pour les pseudo-mots que pour les sons complexes ( $ps = .027$ ) entre 5 et 25 dB-SPL et pour les mots que pour les sons complexes ( $ps < .001$ ) entre 0 et 25 dB-SPL.

### 4.1.3 Discussion

Les résultats de l'Étude I montrent que la détection auditive est influencée par la nature de la stimulation. Pour un même niveau d'intensité sonore, les participants détectent plus facilement des mots que des pseudo-mots et détectent également plus facilement des pseudo-mots que des sons complexes sans composante phonologique. Les connaissances lexicales et phonologiques que les participants possèdent sur les stimulations auditives semblent alors faciliter leur détection. Cette étude réplique le résultat obtenu en modalité visuelle (Merikle & Reingold, 1990) et l'étend en suggérant un effet phonologique. Cependant, ces résultats sont à prendre avec précaution dans la mesure où d'autres explications pourraient être à l'origine de ce résultat.

Premièrement, une explication en termes de différences énergétiques entre les stimulations pourrait être avancée. Dans la mesure où les stimulations ont été équilibrées en terme d'intensité, de durée, d'enveloppe temporelle, de spectre moyen et/ou de phonèmes, les mots et les pseudo-mots diffèrent seulement par leurs aspects lexicaux et les stimulations de parole et de non parole (les sons complexes) diffèrent seulement par leurs aspects lexicaux et phonologiques (pour une discussion détaillée, voir page 123).

Deuxièmement, il est possible que l'égalisation choisie n'harmonise pas correctement les différences énergétiques entre les stimulations. En effet, l'égalisation en dB-SPL choisie n'est peut-être pas la mieux adaptée pour rendre compte de la sonie des stimulations complexes, tels que les sons de parole (Kewley-Port, 1991). Une pondération en dB-A est généralement utilisée lorsque les sons de parole sont présentés à de faibles niveaux d'intensité. Le choix d'une égalisation en dB-A, reflétant les fonctions de transfert exercées par l'oreille interne et moyenne, sera utilisé dans l'Étude II afin de minimiser les différences de sonie des stimulations auditives.

Troisièmement, des différences attentionnelles pourraient expliquer nos résultats. En effet, le choix de la présentation des stimulations par blocs pourrait induire plus de monotonie dans les séquences de sons complexes comparativement aux séquences de mots par exemple. Au cours des expériences 2 et 3 de l'étude de Merikle et Reingold (1990), la détection des mots et les non-mots a été testée dans des

blocs distincts. Lors de l'expérience 4 de Merikle et Reingold (1990), les mots et les non-mots ont été présentés aléatoirement lors d'une même tâche de détection de présence ou d'absence de stimulations, réduisant ainsi la contribution des effets d'attention probables de leurs expériences 2 et 3. Les mêmes résultats sont obtenus pour l'expérience 4 que pour les expériences 2 et 3 : les participants obtiennent de meilleures performances en ce qui concerne la détection des mots plutôt que celle des non-mots, suggérant que la présentation par blocs n'a pas d'influence. Toutefois, dans cette étude, une seconde tâche de catégorisation était toujours présente après la tâche de détection, ce qui suggère que le niveau d'attention devait être maintenu à travers les expériences. Ceci n'est probablement pas le cas dans notre Étude I. Pour pallier à cette critique, une présentation aléatoire des stimulations sera réalisée dans l'Étude II. De plus, afin de tester l'influence de l'attention sur la détection, une seconde tâche sera (Étude II, Expérience 1) ou non (Étude II, Expérience 2) présentée à la suite de la tâche de détection.

## **4.2 Étude II : « Facilitated auditory detection for speech »**

L'Étude II a pour objectif d'examiner si les effets des connaissances phonologiques et lexicales sur la détection sont encore observés lorsque deux paramètres expérimentaux sont manipulés : l'attention du participant et le type d'égalisation des stimulations. Pour cela, les stimulations ne seront plus présentées par blocs mais de façon aléatoire dans une même condition expérimentale. Une seconde tâche sera ou non incluse après la tâche de détection afin de contrôler l'influence des effets d'attention sur la détection. Deux types d'égalisation seront également utilisés (dB-SPL et dB-A) afin de tester leur influence sur la détection. Enfin, les séquences de stimulations comprendront des silences afin de pouvoir mesurer les performances des participants à l'aide d'une approche objective (voir page 46). Les résultats mettent notamment en avant un effet de supériorité de détection des stimulations linguistiques, et un effet de supériorité de détection des stimulations lexicales seulement lorsque les participants sont engagés dans une tâche de reconnaissance ultérieure.

Cette étude a été présentée entièrement ou en parties lors des Journées Jeunes Chercheurs en Audition, Acoustique musicale et Signal audio, à Lyon en octobre 2006, lors des XVth Conference of the European Society for Cognitive Psychology (ESCOP) à Marseille en septembre 2007, lors des Journées fondatrices Perception Sonore, à Lyon en janvier 2007, lors du Forum des Sciences Cognitives à Paris en avril 2007, et lors des Rencontres Scientifiques de l'École Doctorale en mai 2009. Elle a également fait l'objet d'une communication orale au Laboratoire de Sciences Cognitives et Psychoacoustique à Paris en mai 2007, sur invitation du Dr. Sid Kouider. À ce jour, elle fait l'objet d'un article en cours de révision (*Journal of Experimental Psychology : Human Performance and Perception*).



# FACILITATED AUDITORY DETECTION FOR SPEECH

Carine Signoret<sup>1</sup>, Barbara Tillmann<sup>1</sup>, Etienne Gaudrain<sup>1, 2, 3</sup>,  
Nicolas Grimault<sup>1</sup> & Fabien Perrin<sup>1</sup>

1. UMR 5020 “Neurosciences Sensorielles, Comportement, Cognition”, CNRS -  
Université Claude Bernard Lyon 1, Lyon, France
2. Centre for the Neural Basis of Hearing, Department of Physiology, Development  
and Neuroscience, University of Cambridge, Cambridge, United-Kingdom
3. MRC Cognition and Brain Sciences Unit, Cambridge, United-Kingdom

## **Abstract**

Knowledge facilitates higher cognitive functions, such as word processing. Our study tested the influence of phonological and lexical processes on auditory detection. Words, pseudo-words and complex non-phonological sounds, energetically matched as closely as possible, were presented at a range of levels from sub-threshold to clearly audible. Participants performed a detection task followed (Experiment 1) or not (Experiment 2) by a two alternative forced-choice recognition task. In both experiments, phonological stimuli (words and pseudo-words) were better detected than non-phonological stimuli (complex sounds), presented close to auditory threshold. This finding suggests an advantage of speech for signal detection. Moreover, an additional advantage of words over pseudo-words was observed in Experiment 1, suggesting that lexical knowledge could also improve auditory detection when listeners have to recognize the stimulus. The results of Experiment 1 further suggest correct recognition of words in the absence of detection with a subjective threshold approach, but not with an objective approach derived from signal detection theory.

*Keywords:* speech detection effect, auditory threshold

## 4.2.1 Introduction

The influence of linguistic knowledge on perception has been demonstrated in numerous studies using stimuli in context or in isolation. The priming effect is one of the main sources of evidence for the influence of linguistic knowledge on perception in context (Neely, 1977, 1991; McNamara, 2005). In the visual priming paradigm, two stimuli are presented: a prime (a word) is followed by a target (a word or a pseudo-word). Participants have to perform a lexical decision task on the target (i.e., they have to decide whether the target stimulus is a word or a pseudo-word). The priming effect refers to higher accuracy and shorter response times for targets that are related to the prime (semantically, orthographically, phonologically or by repetition) than for targets that are unrelated. For instance, participants decide more accurately and faster that the character string *cat* is a word when the preceding stimulus is *dog* (semantically related) than when the preceding stimulus is *mug* or a pseudo-word like *peb* (unrelated). The priming effect has also been observed for auditory stimuli (Church & Schacter, 1994 ; Deutsch & Bentin, 1994 ; Orfanidou, Marslen-Wilson, & Davis, 2006 ; Pilotti & Beyer, 2002 ; Ratcliff, Allbritton, & McKoon, 1997 ; Schacter & Church, 1992 ; Slowiaczek, Nusbaum, & Pisoni, 1987). Representations of words in memory can also be activated by similar-sounding non-words (used as primes), leading to shorter lexical decision times for semantically related targets than for unrelated targets (Connine, Blasko, & Titone, 1993). Behavioral experiments have shown that listeners' lexical knowledge can influence the perceptual restoration of a phoneme. Warren and colleagues have studied situations in which a speech sound is replaced by a non-speech sound (Warren, 1970, 1984 ; Warren & Obusek, 1971 ; Warren & Sherman, 1974). Participants were not aware of the missing phoneme and could not specify the location of the non-speech sound in the sentence they had just heard. Their lexical knowledge filled in the missing speech sound.

The influence of lexical as well as of phonological knowledge has been shown for the processing of stimuli presented in isolation. In the seminal study of Cattell (1886), words and non-words were visually presented for a short duration (5-10 ms) and participants had to report as many letters as they could. The author has reported a Word Superiority Effect (WSE): target letters are better reported, identified or recognized when they are part of a word than when they are part of either a pseudo-

word (an orthographically legal pronounceable letter string) or a non-word (an orthographically illegal unpronounceable letter string) (see also Grainger, Bouttevin, Truc, Bastien, & Ziegler, 2003; Grainger & Jacobs, 1994; McClelland, 1976; Reicher, 1969; Wheeler, 1970). This finding suggests that short-term memory limitations can be more easily overcome for words than for pseudo-words or non-words because fewer items have to be remembered within a word thanks to the lexical knowledge, which helps to reconstruct the word and thus to find the letters. Similarly, a Pseudo-Word Superiority Effect (PWSE) has been reported (Baron & Thurston, 1973 ; Grainger & Jacobs, 1994 ; McClelland, 1976 ; McClelland & Johnston, 1977). This effect refers to facilitated perception thanks to phonological features of the stimulus: target letters are better identified when they are part of a pseudo-word than when they are part of a non-word (for example Maris, 2002; McClelland & Rumelhart, 1981). Facilitation was also observed on a more local level (i.e., level of one character): letters are better recognized than pseudo-letters or symbols (Burgund et al., 2009 ; Mason, 1982 ; Reingold & Jolicoeur, 1993 ; Tydgat & Grainger, 2009).

Two classes of models have been proposed to explain facilitation due to higher-level, knowledge-related influences. The first class of word recognition models in the auditory modality suggests a bottom-up analysis of the acoustic signal that starts with the beginning of the word, thus these models emphasize the importance of the word onset in auditory processing (the Activation-Verification Model of Paap, Newsome, McDonald, & Schvaneveldt, 1982; the Dual Read Out Model of Grainger & Jacobs, 1994; see also Norris, McQueen, & Cutler, 2000). The information carried by the word onset activates a group of candidate words and, because of the sequential nature of speech, the analysis of the word continues until only one word matches the signal (recognition point) and becomes selected (Marslen-Wilson, 1987 ; Norris, 1994). Perceptual processing is seen as a strictly autonomous, bottom-up process, with the influence of lexical and semantic contextual information arising only at a later decision stage. The second class of models explains lexical and phonological facilitations by top-down processing. McClelland and Rumelhart (1981) developed the Interactive Activation Model (IAM) for visual word recognition. In this model, the WSE, the PWSE and the semantic priming effects can be explained by top-down feedback from the word representation level to the letter representation level. Word units activate the representation of their component letters, thus facilitating letter

perception for words but not for non-words. Pseudo-words activate units for words that are consistent with most of the activated letters, and feedback from the activated word units strengthens the activation of the letters in the pseudo-words. Based on the IAM, McClelland and Elman (1986) developed a model for the auditory modality: the interactive TRACE model of speech perception. This model combines bottom-up information and top-down feedback from the lexical level down to the phonemic level. More recently, McClelland, Mirman and Holt (2006) further proposed that lexical influences can affect pre-lexical mechanisms that are triggering compensation, adaptation and retuning of phonological processes.

Previous research have shown that linguistic knowledge improves higher-level processing (such as letter identification or lexical discrimination) and the two models above have been proposed to explain these facilitatory effects (McClelland & Elman, 1986 ; McClelland et al., 2006). In contrast, only two studies have investigated the possible influence of lexical or phonological knowledge *on earlier stages of sensory processing*. In the visual modality, Doyle and Leach (1988) and Merikle and Reingold (1990) have reported an advantage of words over non-words for detection, which has been called the Word Detection Effect (WDE). In the study of Doyle and Leach (1988), participants had to detect a stimulus (a word or a non-word) that was briefly displayed on the screen. The results suggest that words were more readily detected than non-words. However, the two sets of stimuli were not matched for the number of letters. Any difference in detection might thus be attributed to a difference in physical properties between the two sets of stimuli. In contrast, Merikle and Reingold (1990) used words and non-words that were matched for the number of letters. In this better controlled study, a WDE remained: when the onset asynchrony between the visual target and a following mask was so short that participants had difficulties in detecting the target stimulus, words were more readily detected than non-words. To our knowledge, no study has investigated this kind of facilitation in the auditory modality. In particular, no study has investigated the influence of phonological knowledge on detection, i.e. whether pseudo-words are better detected than non-speech stimuli. The data collected in the visual modality do not allow us to predict findings for the auditory modality since visual and auditory nervous pathways are different, at least from the sensory organs to the primary sensory cortices. Thus, it remains to be shown

whether a WDE or more generally, a Speech Detection Effect (SDE) can also be observed in the auditory modality.

Our present study investigated the influence of phonological and lexical knowledge on auditory detection and perception, using three types of stimuli: words, pseudo-words and non-phonological complex sounds. In order to minimize differences in sensory processing or bottom-up facilitation, we matched the energetic properties of the three stimulus types as closely as possible (i.e., in terms of loudness, duration, temporal envelope, average spectrum). In addition, the words and the pseudo-words were composed of the same phonemes to minimize phonological differences. Participants performed a detection task followed (Experiment 1) or not (Experiment 2) by a two alternative forced-choice (2AFC) recognition task. We hypothesized that if top-down influences do not reach the lowest levels of event processing, the detection thresholds should not differ between phonological (words and pseudo-words) and non-phonological (complex sounds) stimuli and/or between lexical (words) and non-lexical (pseudo-words) stimuli. In contrast, if phonological and/or lexical top-down influences reach the lowest level of event processing, auditory detection should be improved for phonological over non-phonological stimuli and/or for lexical over non-lexical stimuli.

### **4.2.2 Experiment 1**

The aim of Experiment 1 was to investigate whether phonological and lexical knowledge influences auditory detection, which is one of the first steps of sound processing. Non-phonological complex sounds, pseudo-words and words were randomly presented at levels that ranged from inaudible to audible. For each trial, participants had to detect the presence or absence of a stimulus. We compared auditory detection performance for stimuli that were energetically matched as closely as possible. With this experimental material, any difference in detection performance would suggest top-down influences of phonological and/or lexical knowledge.

The detection task was followed by a 2AFC recognition task. The aim of this second task was twofold. First, it ensured that attention attributed to the stimulus was, as far as possible, constant for all stimuli. Second, this task allowed us to investigate dissociation between auditory detection and recognition. The recognition of an

auditory stimulus in absence of its detection is still a matter of debate. While dissociation between detection and higher level processing has been observed for the visual modality with different experimental and theoretical backgrounds (e.g. Dehaene et al., 1998; Merikle & Eyal M. Reingold, 1990; Naccache & Dehaene, 2001; E M Reingold & Merikle, 1988), few studies have investigated these effects for the auditory modality. In experiments carried out by Shipley (1961, cited in Shipley, 1965), participants had to decide whether a masked pure tone was presented or not and then to choose, between two tones, which one has been presented in the detection task. Following a “no” response, listeners did not discriminate which of the signals had been presented. In a very similar study, Lindner (1968) told participants that a recognition judgment following a “no” detection response was possible (this was not specified by Shipley) and observed the opposite result pattern: when participants did not detect the tone, the proportion of correct recognition responses was above chance. To our knowledge, no study has previously investigated a potential dissociation between detection and recognition of speech and non-speech stimuli without using a masking procedure. As done in previous studies using visual materials (Fisk & Haase, 2005 ; Haase & Fisk, 2004 ; Holender & Duscherer, 2004 ; Reingold, 2004a ; Snodgrass, Bernat, & Shevrin, 2004a, 2004b), Experiment 1 investigated this issue with two current approaches. The subjective threshold approach supposes a dissociation between detection and recognition such that under stimulus conditions in which participants do not report awareness of the stimuli, they can nonetheless perform above chance on perceptual discrimination tasks (e.g. Cheesman & Merikle, 1984, 1986; Merikle & Cheesman, 1986). The objective threshold approach is based on an index of sensitivity on the awareness variable (i.e., performance is at chance with a direct measure of detection) that is used as an indicator of null awareness (e.g. Greenwald, Klinger, & Schuh, 1995; Snodgrass, Shevrin, & Kopka, 1993). With the advent of signal detection theory (SDT, Macmillan & Creelman, 2005), it became accepted that the subjective threshold might reflect response criteria applied to a single, conscious process, rather than delimiting the conscious/unconscious boundary. Even if SDT has the advantage to exclude influences linked to bias, Macmillan (1986) and Merikle and Reingold (Merikle & Reingold, 1990 ; Reingold & Merikle, 1988) admitted that demonstrating true null sensitivity on direct measures is very difficult. Since no real consensus has been found to establish dissociation between detection

and recognition, both models were applied to our data, and then compared, as it has been recently proposed by Fisk and Haase (2005).

## **a. Method**

### ***Participants***

19 students of the University of Lyon (mean age  $21.2 \pm 2.1$  years, 14 females), right-handed on the “Edinburgh Handedness Inventory” (Oldfield, 1971), were included in Experiment 1. All were French native speakers and did not report any hearing problems or history of neurological disease. They were selected on the basis of their low auditory detection thresholds, notably they were included only if their pure tone thresholds (as described in American National Standards Institute [ANSI] S3-21, 2004) were below 15 dB-HL for frequencies between 250 Hz to 8000 Hz. All participants provided written informed consent to the study, which was conducted in accordance with the guidelines of the Declaration of Helsinki, and approved by the local Ethics Committee (CPRB Léon Bérard, n°05/026).

### ***Materials***

Three types of stimuli were used: words, pseudo-words and complex sounds. Words were selected from a French database (Lexique 2, New, Pallier, Brysbaert, & Ferrand, 2004). They were common singular monosyllabic nouns and contained 2-7 letters and 2-5 phonemes. All words had a frequency of occurrence higher than 1 per million occurrences in books as well as in movies (subtitles, New, Brysbaert, Veronis, & Pallier, 2007), and were uttered by the same female speaker. A list of pseudo-words was generated from words, by mixing the phonemes: all phonemes of the words were used. Pseudo-words could be pronounced but did not have any meaning<sup>1</sup>. The number of letters and phonemes were matched between words and pseudo-words. The pseudo-words were uttered by the same female speaker as the words.

---

<sup>1</sup> In a pretest, five other participants (mean age  $26 \pm 2.1$  years, 2 women) evaluated the phonological similarities of the pseudo-words to words. They had to judge if the pronounced pseudo-words sounded like a word, and if this was the case, they had to write down the corresponding word. All pseudo-words for which words have been indicted by at least two participants were eliminated. In a second part of the pretest, the participants judged the strength of semantic associations of pairs of words on a 5-point scale (from 0 = no association to 5 = very strong association). Word pairs with scores inferior to 2 were used in the second task of the Experiment 1 (i.e., recognition task) as distractors.

The complex sounds were created from the words and pseudo-words using the algorithm *Fonds sonores* (Hoen et al., 2007 ; Perrin & Grimault, 2005). This method is similar, at least in its principles, to other methods successfully used in neuroimaging studies (e.g. Davis & Johnsrude, 2003; Giraud et al., 2004). First, starting from a word or a pseudo-word, the overall phase spectrum was randomized while the overall magnitude spectrum of the phonological stimulus was preserved. Second, the slow temporal envelope (below 60 Hz) of the phonological stimulus was applied on the resulting signal. Consequently, the onsets and offsets of the complex sounds were matched to those of the original stimuli. This transformation roughly preserved the average spectral content and preserved the slow time course of the amplitude. The excitation patterns (Moore & Glasberg, 1987) evoked by the stimuli and computed before and after transformation are almost superimposable. However, due to the phase spectrum randomization, these stimuli sound like different variations of noise, so they are not recognized as speech and the perceived pitch is relatively weak<sup>2</sup>. A diagram of the algorithm is available at <http://olfac.univ-lyon1.fr/unite/equipe-02/FondsSonores.html> and sound samples of all stimuli categories are available at [http://olfac.univ-lyon1.fr/unite/equipe-02/sound\\_examples\\_Signoret.html](http://olfac.univ-lyon1.fr/unite/equipe-02/sound_examples_Signoret.html). These complex sounds were used rather than temporally reversed speech to avoid preserving phonological characteristics, hence providing a stronger contrast with speech material. Speech segments that are steady-state, like vowels, are largely unaffected by time reversal. As a consequence, reversed speech is generally identified as speech whereas the complex sounds used in the present experiment were not. In summary, these complex sounds had the same overall energetic properties as the words and pseudo-words – the average spectrum, and the slow temporal variations – while not being recognized as speech.

The average duration of words, pseudo-words and complex sounds were not significantly different ( $F(2,459) = 2.58, p = .08$ ) and were respectively 521.5 ms ( $SD = 115.5$  ms), 539.2 ms ( $SD = 87.8$  ms) and 546.6 ms ( $SD = 92.1$  ms). To reduce differences in the perceived loudness, all stimuli were equalized to the same dB-A

---

<sup>2</sup> The pitch strength estimated according to a method similar to that of Ives and Patterson (2008), indicated that non-speech stimuli (complex sounds) had a pitch only 15% less salient ( $t(2020) = 34.67, p < .001$ ) than the speech stimuli (words and pseudo-words). There was no difference in pitch strength between words and pseudo-words ( $t(1009) = .23, p = .39$ ).



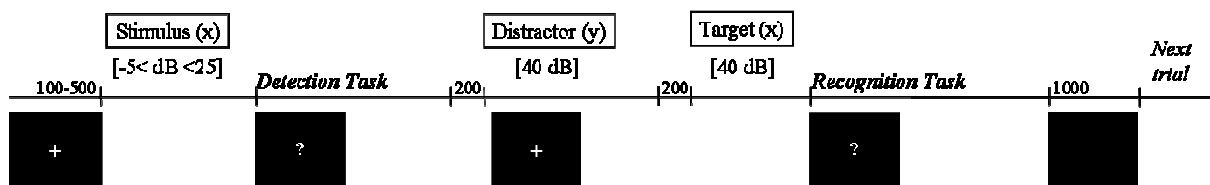
level. The A-weighting roughly mimics the external and middle ear transfer functions, therefore the stimuli are equalized in the energy that reaches the inner ear.

### *Apparatus*

Words and pseudo-words were recorded (32 bits, 44.1 kHz) using a Røde NT1 microphone, a Behringer Ultragain preamplifier, and a VxPocket V2 Digigram soundcard. The mean level of presentation was calibrated (ANSI, 1995) to reach 80 dB-A in a standard artificial ear (Larson Davis AEC101 and 824). All stimuli were played through a soundcard (Creative Soundblaster Audigy 2) followed by an analog attenuator (TDT PA4, one for each channel) that applied a fixed 40 dB attenuation. This attenuation was analog rather than digital to prevent acoustical distortion at low levels of presentation. All stimuli were binaurally presented to participants through comfortable headphones (Sennheiser HD 250 Linear II) connected to a headphone buffer (TDT HB6).

### *Design and Procedure*

For each participant, 504 trials were presented in random order using Presentation 9.7 software. Within each trial, three stimuli from the same category were presented (see Figure 1).



*Figure 1.* After displaying a fixation cross, a stimulus (word, pseudo-word or complex sound) or a silence was presented (from -5 to 25 dB-A) and participants had to perform a detection task by pressing yes or no answer keys. Two hundred milliseconds after the response, two stimuli separated by 200 ms were presented at audible level (one being the same as the stimulus presented in the detection task - the target- and the other being a distractor of the same category, presented in a random order). Participants had to choose in a 2AFC recognition task the stimulus that was the same as the stimulus presented in the detection task by pressing one of two answer keys (first or second stimulus).

The first stimulus was a word, a pseudo-word or a complex sound in 30.6 % of trials respectively, and in 8.2 % of trials there was no stimulus (i.e., a silence). A digital attenuation was randomly applied to the first stimulus from 15 to 45 dB by steps of 3 dB to reach levels of presentation from -5 to +25 dB-A. Overall, 14 words,

14 pseudo-words and 14 complex sounds were presented for each level in random order along with 42 silences. Participants were told that the stimulus was sometimes replaced by a silence. Each stimulus was presented only once to a participant (i.e., at one given level of presentation). Across participants, each stimulus was presented at a different presentation level. Participants had to decide whether they detected or not an auditory stimulus (detection task) by pressing *yes* or *no* answer keys, whose position was counterbalanced across participants.

Two hundred milliseconds after the response to the detection task, the second and third stimuli were presented at an audible level (+40 dB-A). The inter-stimulus-interval between the two stimuli was 200 ms. One of the two stimuli was the same as the detection stimulus (repetition relationship) and was randomly and equally presented in the first or second interval over stimuli and participants. The other stimulus was a distractor of the same category (154 words, 154 pseudo-words and 154 complex sounds, not presented in the detection task, were presented as distractors) and it appeared only once in the experimental session. For words and pseudo-words, the distractor was neither phonologically nor semantically related to the first stimulus (as evaluated by a pretest<sup>1</sup>). The number of letters and phonemes was the same for the items within each pair of words or pseudo-words. For the trials where the first stimulus was a silence, a pair of words, pseudo-words or complex sounds was randomly selected (in total, 14 pairs of words, 14 pairs of pseudo-words and 14 pairs of complex sounds were presented). After having listened to the pair of stimuli, the participants had to choose whether the previously heard stimulus was similar to the first or to the second stimulus of the pair (2AFC recognition task). They were asked to respond as quickly as possible (but no timeout was imposed) even if they had indicated that they had not heard anything in the detection task. The next trial appeared 1000 ms after the participants' response. A visual fixation cross appeared 100 to 500 ms before the onset of the first stimulus and remained on the screen until its offset. Participants heard six blocks of 84 trials in a randomized order. Short breaks were imposed between the six blocks. The duration of the entire experiment was about one hour.

## b. Results

### *Detection task*

Overall percentages of correct responses for silence (i.e., correct rejections) were high (mean = 97.62 %,  $SD = .70\%$ ). From the proportion of ‘yes’ responses  $Pr(\text{yes})$ , a measure of detectability ( $d'_D$ ) and criterion ( $k$ ) were calculated for each participant as proposed in the SDT (Macmillan & Creelman, 2005). The constraints of the experimental design imposed that only one false-alarm rate ( $FA$ ) was collected for each participant, i.e. common to all types of stimulus and all stimulus levels. Therefore, the criterion  $k$  calculated here was equal to  $-z(FA)$ , which represents the overall response bias (Macmillan & Creelman, 2005, p.116).  $d'_D$  was analyzed with a two-way analysis of variance (ANOVAs) with Type of Stimulus (words/pseudo-words/complex sounds) and Stimulus Level (11 levels, from 1 to 11, i.e., from -5 to +25 dB-A with 3 dB steps) as within-participant factors. An alpha level of .05 after Greenhouse-Geisser correction was used for all statistical tests. Differences between stimulus types can occur on the slope of a psychometric curve, but not at lowest and highest presentation levels for which performance of all stimuli can be minimal (0-5%) or maximal (95-100%) (i.e., floor and ceiling effects). For our data, a floor effect was observed between levels 1 and 3 (no significant differences were observed between levels 1 and 2, 2 and 3, 1 and 3,  $ps > .07$ ), and a ceiling effect appeared between levels 9 and 11 (no significant differences were observed between 9 and 10, 10 and 11 and 9 and 11,  $ps > .37$ ). This suggested that performance between the three types of stimulus could only differ between levels 4 to 8, i.e., on the slope of the psychometric curve. Based on this a priori hypothesis, a significant interaction was analyzed with planned comparisons (local contrasts) to test for stimulus type differences, thus excluding the influences of ceiling and floor effects for the analysis.

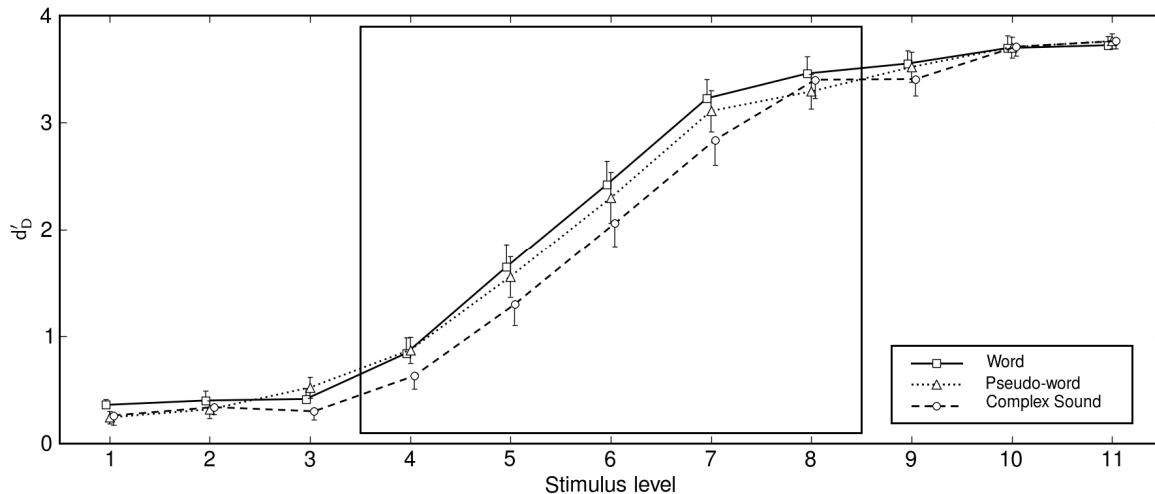


Figure 2. Analyses with signal detection theory:  $d'_D$  presented as a function of Stimulus Level (1 to 11, i.e., -5 to +25 dB-A ) for each Type of Stimulus (word/pseudo-word/complex sound). The error bars show the standard error.

For the analysis of  $d'_D$  (see Figure 2), the main effect of the Type of Stimulus ( $F(2,36) = 29.93, p < .001$ ), the main effect of the Stimulus Level ( $F(10,180) = 1817.3, p < .001$ ) and the interaction between these two factors ( $F(20,360) = 2.26, p < .001$ ) were significant. For levels 4 to 8 (i.e., from +4 to +16 dB-A), planned comparisons revealed that words were better detected than pseudo-words ( $F(1,18) = 7.01, p = .016$ ) and than complex sounds ( $F(1,18) = 59.80, p < .001$ ), and that pseudo-words were better detected than complex sounds ( $F(1,18) = 17.28, p < .001$ ). On average, the criterion  $k$  was equal to 1.99 ( $SD = .32$ ).

The  $d'_D$  measure represents the detectability of a stimulus by accounting for participant's tendency to respond "yes" or "no". In addition, performance can be broken-up into two measures: the *absolute sensitivity* and the *absolute detectability*. The *absolute sensitivity* is captured by the slope of the psychometric function and is an unbiased measure of sensitivity. The *absolute detectability* is captured by the horizontal position of the psychometric function, and represents the detection bias related to each type of stimulus. These measures were estimated by fitting a cumulative Gaussian on the percent-correct detection data, using the maximum-likelihood method (as suggested in Macmillan & Creelman, 2005). These two measures were analyzed by two ANOVAs with Type of Stimulus as within-participant factor. For the *absolute sensitivity*, this analysis revealed no significant effect ( $F(2,36) = .51, p = .61$ ) while for the *absolute detectability*, the effect of Type

of Stimulus was significant ( $F(2,36) = 24.76, p < .001$ ): words and pseudo-words yielded better *absolute detectability* than complex sounds ( $ps < .001$ ).

### **Recognition task**

Performance (% of correct recognition) was analyzed with a two-way ANOVA on the Type of Stimulus (words/pseudo-words/complex sounds) and the Stimulus Level (11 levels numbered 1 to 11, ranging from -5 to +25 dB-A with 3 dB steps) as within-participant factor. An alpha level of .05 after Greenhouse-Geisser correction was used for all statistical tests.

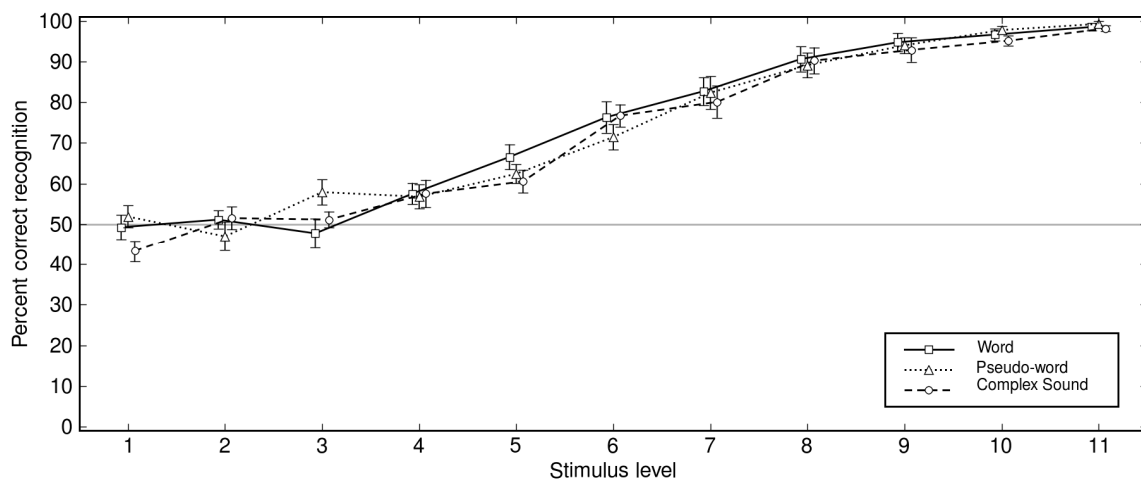


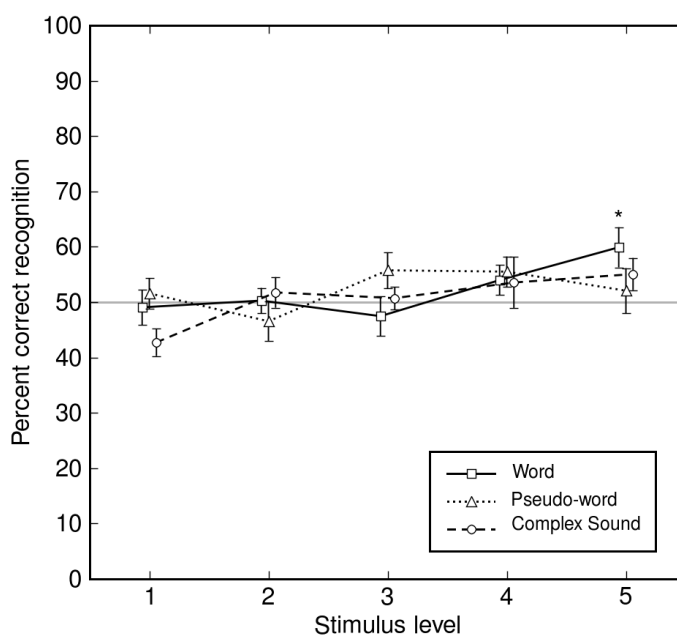
Figure 3. Percent correct in the recognition task presented as a function of Stimulus Level (1 to 11, i.e., -5 to +25 dB-A) for each Type of Stimulus (word/pseudo-word/complex sound). The error bars show the standard error. The dotted line represents chance level.

The ANOVA showed a significant main effect of the Stimulus Level on recognition performance ( $F(10,180) = 110.77, p < .001$ ): correct recognitions were more numerous for high levels than for low levels (see Figure 3). There was no significant effect of the Type of Stimulus ( $F(2,36) = .95, p = .40$ ) and no significant interaction between the Stimulus Level and the Type of Stimulus ( $F(20,360) = 1.56, p = .06$ ). No significant correlation was observed between detection and recognition performance ( $r(17) = .004; p = .96$  for words;  $r(17) = .014; p = .86$  for pseudo-words;  $r(17) = .020; p = .80$  for complex sounds). Moreover, the target position in the recognition interval had no influence on recognition for the three sets of stimuli ( $t(18) = .67, p = .51$  for words;  $t(18) = .51, p = .62$  for pseudo-words;  $t(18) = .47, p = .64$  for complex sounds).

### ***Recognition without detection***

Dissociation between detection and recognition was analyzed with a subjective threshold approach (e.g. Merikle & Cheesman, 1986) and an objective threshold approach (e.g. Greenwald et al., 1995).

The subjective threshold approach tests whether correct recognition performance exceeded chance level (50%). For trials in which participants did not detect the first stimulus (as indicated by “no” responses – misses – in the detection task), analyses were performed for levels of presentation at which all participants had a minimum of 15% of misses per level (as in Fisk & Haase, 2005), i.e., for levels 1 to 5 (see Figure 4). When participants had responded “no” in the detection task, the recognition performance was significantly above chance only at level 5 (i.e., +7 dB-A;  $t(18) = 2.29$ ,  $p < .05$ ). At this level, the percentage of correct recognition was above chance only for words ( $t(16) = 2.65$ ,  $p = .017$ ), but not for pseudo-words ( $t(18) = .49$ ,  $p = .62$ ) or complex sounds ( $t(18) = 1.69$ ,  $p = .11$ ).



*Figure 4.* Percent correct in the recognition task presented as a function of Stimulus Level (1 to 5, i.e., -5 to +7 dB-A) and Type of Stimulus (word/pseudo-word/complex sound) for trials in which participants had a “no” response in the detection task. The grey line indicates the chance level and the error bars show the standard error. Point marked with an asterisk is significantly different from chance level.

With the objective threshold approach, recognition was modeled with methods based on the SDT (Macmillan & Creelman, 2005). Recognition sensitivity was expressed as  $d'_R$ , calculated with the  $\sqrt{2}$  correction given that two response choices were possible (Macmillan & Creelman, 2005). The  $d'_R$  and  $d'_D$  variables were compared at each stimulus level using paired  $t$ -tests (two-tailed):  $d'_D$  was always significantly greater than  $d'_R$  ( $p > .05$ ). To assess the possibility of recognition without detection, the value of  $d'_R$  when  $d'_D$  is close to zero needs to be evaluated (Figure 5). Since the distribution of individual  $d'_D$  was not centered on zero, the value of  $d'_R$  when  $d'_D = 0$  was extrapolated using a linear regression (Greenwald et al., 1995). This analysis was conducted for levels for which individual  $d'_D$  were distributed near zero (i.e., levels 1 to 3) as the contribution of better detected conditions would bias the regression (Miller, 2000). A vertical-intercept greater than zero indicates recognition without detection. A slope greater than zero indicates a correlation between recognition and detection performances. The intercepts from the regression lines were never significantly greater than zero for words, pseudo-words and complex sounds ( $ps > .74$ ). The slopes of the regressions were never significantly greater than zero for words, pseudo-words and complex sounds ( $ps > .56$ ).

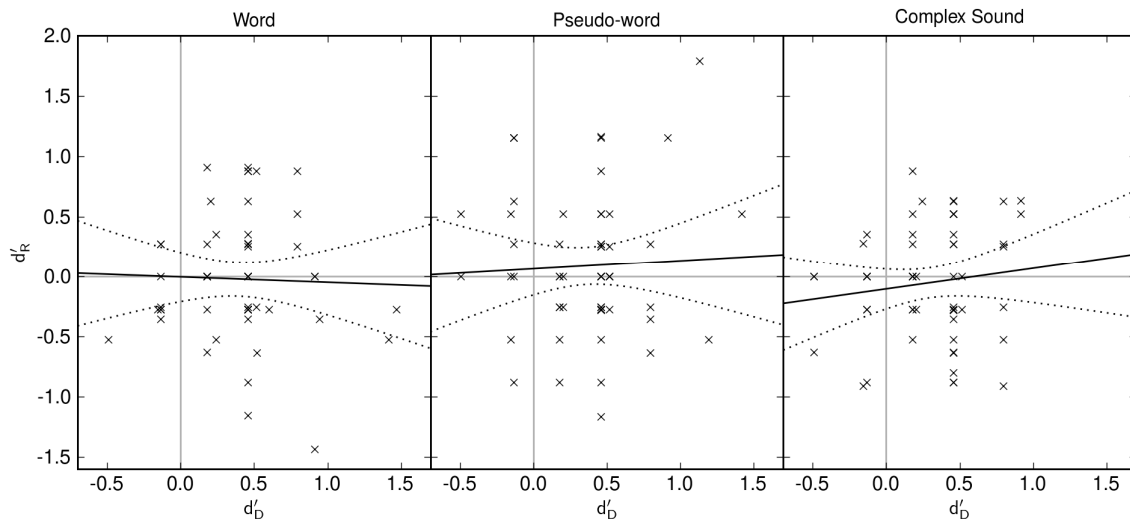


Figure 5. Linear regression between  $d'_R$  and  $d'_D$  for each Type of Stimulus. The linear regression is the black line and confidence interval is represented by the dotted lines.

## c. Discussion

### *Speech and word detection effects*

Experiment 1 showed that, close to auditory threshold, phonological stimuli (words and pseudo-words) were better detected than non-phonological stimuli (complex sounds). This was revealed by detectability differences between stimuli from +4 to +16 dB-A. In addition, better detection of meaningful phonological stimuli (words) over meaningless ones (pseudo-words and complex sounds) has been observed. These findings suggest that when auditory stimuli are difficult to detect, (1) phonological knowledge facilitates the detection of stimuli, i.e., suggesting the existence of a speech detection effect (SDE) in the auditory modality and it further shows that (2) lexical knowledge facilitates the auditory detection, i.e., a word detection effect (WDE). However, this latter effect was not as strong as the SDE in the  $d'_d$  analysis (difference in  $d'_d$  between words and pseudo-words was .10 against .23 between speech and non-speech).

The detectability differences between the stimulus types do not seem to be related to a difference in the slope of the psychometric function, but to a difference in horizontal shift along the stimulus level axis. This means that the observed differences on  $d'_d$  were due to differences in *absolute detectability*. Under the hypothesis that the internal criterion is the same within the experimental session, a difference in *absolute detectability* suggests that the type of stimulus modulates the amount of internal noise and its effect on sensory representations. For our study, the difference in *absolute detectability* cannot be explained by systematic energetic differences between the items of the three stimulus types. The stimuli were carefully built so that the energetic features matched as closely as possible between the three categories. Indeed, intensity, duration, temporal envelope and spectrum, as well as phonemes for words and pseudo-words, were on average as similar as possible across the sets of stimuli. Nevertheless, although equalization in dB-A has been shown to be relatively efficient to reduce the variability in thresholds for vowels (Kewley-Port, 1991), equalization in dB-A is based on a rudimentary model of loudness. This model reflects the action of the external and middle ear and acts on the average spectrum of the sounds. This procedure might be relatively well-suited for vowels, which are steady-state, but might be less suited for words, pseudo-words or complex sounds that contain transient



components making the average spectrum less reliable for energy measurement. To experimentally assess the robustness of the facilitatory effects to loudness differences, a different equalization scheme, where stimuli are equalized in the energy required to produce the sound, was added in Experiment 2. With this second equalization scheme, all loudness differences that were due to the specific nature of the stimulus were preserved and could contribute to detection performance.

### ***Word superiority effect***

In the 2AFC recognition task, the data of Experiment 1 did not show a word superiority effect (WSE), which had been observed for letter recognition (e.g. Cattell, 1886; Reicher, 1969). At first surprising, the absence of WSE might be explained by the specificity of our experimental paradigm. In contrast to studies that demonstrated a WSE, the recognition task of our study was both time-delayed and disrupted by another task (the detection task). The temporal decay of WSE and its sensitivity to task interference has not been studied yet, but it has been shown that linguistic facilitations are time-limited (e.g. Neely, 1977). Further studies without an interfering detection task would be necessary to investigate the presence or absence of WSE in the 2AFC recognition task.

### ***Dissociation between detection and recognition***

The second task of Experiment 1 also allowed us to explore a potential dissociation between detection and recognition of speech and non-speech stimuli. The subjective approach suggests that participants could recognize auditory stimuli even when they had previously said they had not detected any sound in this trial. This is demonstrated by recognition performance above chance level for words at 7 dB-A. Previously, Merikle and Reingold (1990, their Experiments 2 and 3) observed a dissociation between detection and recognition for visual stimuli using the subjective threshold approach. When participants did not detect the stimulus, words were recognized whereas non-words were not. The authors interpreted this finding as unconscious processing and concluded that only familiar stimuli can be perceived unconsciously. Our study further suggests that (i) recognition of familiar stimuli without subjective detection can also be observed in the auditory modality and (ii) unfamiliar stimuli cannot be “perceived unconsciously” even if they had phonetic structure. Future experiments should investigate whether lexical decision could be

observed without subjective auditory detection, as it has been previously reported in the visual modality by Merikle and Reingold (1990). In their study, participants have to perform a detection task immediately followed by a categorization task. With this experimental design, the retrieval environment is not disrupted by the presentation of two stimulus alternatives. If unconscious perception is a general phenomenon, we could predict that comparable patterns of results should be found with a lexical decision task in the auditory modality, as it has been previously reported in the visual modality (Merikle & Reingold, 1990 - Experiment 4).

In contrast to the subjective approach, the objective approach derived from the SDT, does not suggest a dissociation between detection and recognition for any of the stimulus types. The recognition performance was never greater than the detection performance at lower levels of detection (from -5 to 1 dB-A), suggesting that detection decision does not necessarily entail recognition.

Non-concordant results between the subjective and objective approaches have been previously reported in the visual modality by Fisk and Haase (2005) using a dissociation paradigm in which detection performance was compared to identification performance. The absence of dissociation using an objective approach, in their study as in the present one, might be explained by the use of direct tasks. Indeed, in the visual modality, unconscious perception effects using the objective approach have been observed both with direct and indirect tasks (e.g. van Selst & Merikle, 1993; Snodgrass et al., 1993). It can therefore not be excluded that an indirect task might be capable of showing sensitivity to unconscious perception under comparable display parameters. Thus, future studies should address this issue in the auditory modality using indirect tasks, such as the priming paradigm.

### **4.2.3 Experiment 2**

The aim of Experiment 2 was to investigate both the influence of the subsequent recognition task and of the A-weighting equalization on detection performance. The experimental design and the stimuli were as described in Experiment 1, except that the second task was eliminated and two equalization methods were used. If detection does not differ between the three types of stimuli in Experiment 2, this would suggest

that the second task in Experiment 1 had modified participants' strategies for the detection task.

To further assess that the differences observed in Experiment 1 was not a consequence of equalization scheme, all stimuli were equalized once with A-weighting and once with flat-weighting. In contrast to A-weighting, which equalizes stimuli with regard to a simple model of human auditory perception based on the external and middle ear transfer functions, flat-weighting does not apply any weighting to the spectrum, equalizing the physical energy required to produce the sound. The flat-weighting was thus not used as a better equalization scheme, but to test if the A-weighting used previously could have affected the observed results.

## **a. Method**

### ***Participants***

20 other students of the University of Lyon (age  $22.9 \pm 3.7$  years, 16 females) participated in Experiment 2. They were selected with the same criteria as described in Experiment 1.

### ***Materials***

The stimuli described in Experiment 1 were used. Using the original (non-weighted) spectrum, the stimuli were equalized in dB-A, i.e., using the A-weighted, and in dB-SPL, i.e., using flat-weighting. In an attempt to reduce the length of the study, and thus effects due to participants' fatigue, the range of levels of presentation used in Experiment 2 was restricted to levels at which detection was difficult in Experiment 1 (i.e., between levels 4 to 8). For the new sets of stimuli, the average duration of words, pseudo-words and complex sounds were not significantly different ( $F(2,358) = .73$ ;  $p = .48$ ) and were respectively 527.2 ms ( $SD = 103.7$  ms), 542.6 ms ( $SD = 82.1$  ms) and 531.3 ms ( $SD = 93.5$  ms).

### ***Apparatus***

The same apparatus as in Experiment 1 was used. The mean level of presentation of the stimuli equalized in dB-A was calibrated (ANSI S3.7, 1995) to reach 80 dB-A in a standard artificial ear (Larson Davis AEC101 and 824). The

stimuli equalized with the flat-weighting scheme produced a level of presentation of 84 dB-SPL. Because dB-A and dB-SPL are different in nature, it was not possible to equalize the two sets of stimuli to have the same energy level.

### ***Design and Procedure***

For each participant, 396 trials were randomly presented. One trial was equivalent to one stimulus. The stimulus was a word, a pseudo-word or a complex sound (30.3 % of trials respectively), and in 9.1 % of trials there was no stimulus (i.e., a silence). The auditory stimuli were digitally attenuated in order to randomly obtain one of five selected levels of sound presentation: from 0 to +20 dB-A or to +4 to +24 dB-SPL with steps of 5 dB. Overall, 12 words, 12 pseudo-words and 12 complex sounds (resulting in 24 words, 24 pseudo-words and 24 complex sounds per dB-level, thus a total of 120 words, 120 pseudo-words, 120 complex sounds) were presented for each stimulus level and for each equalization in random order along with 36 silences. Participants were told that the stimulus was sometimes replaced by a silence. The order of the stimuli was randomized between participants. Each stimulus was presented only once to a participant and dB-A and dB-SPL versions of a same stimulus were never both presented to the same participant. Overall, each stimulus was presented at all presentation levels, but at only one level to any given participant. After the stimulus, participants had to decide if they had detected an auditory item in a detection task by pressing *yes* or *no* answer keys, whose positions were counterbalanced across participants.

The next trial occurred 500 ms after the response of the participants. A fixation cross appeared 100-500 ms before the presentation of the stimulus and remained until its end. Participants heard three blocks of 132 trials and short breaks were imposed between the three blocks. The duration of the Experiment was about 25 minutes.

### **b. Hypotheses**

As explained in the Apparatus section, dB-A and dB-SPL represent different units of dB-measure and cannot be equated for energy. The energy levels and the loudness corresponding to a given level of attenuation were actually different for the two equalization schemes. Therefore, the response curves for the two equalization schemes may not be centered on the same attenuation value, corresponding to a main

effect of Equalization on  $d'_D$  and on the *absolute detectability* (hypothesis 1). However, the purpose of using two equalization schemes was to assess whether the effect of Type of Stimulus was preserved when the equalization scheme is changed: the main effect of Type of Stimulus should be significant for the two equalization schemes individually. More precisely, the observation of a WDE and/or a SDE should indicate that these effects were not due to the recognition task in Experiment 1 and the robustness of these two effects would be demonstrated if they can be observed for each equalization scheme (hypothesis 2a and 2b for WDE and SDE respectively). If the effect of Type of Stimulus is an horizontal shift of the psychometric function like in Experiment 1, the main effect of Type of Stimulus should be significant on *absolute detectability*, but not on *absolute sensitivity* (hypothesis 3). Moreover, in that case, an interaction between Equalization and Type of Stimulus will be observed on  $d'_D$  if this effect differs between the two equalization schemes. As dB-A values are closer to loudness than dB-SPL, the variability of detectability is also expected to be smaller, at each attenuation level, for dB-A than for dB-SPL. Since the variability of detectability affects the slopes of the psychometric functions, the equalization scheme difference should result in an interaction between Equalization and Stimulus Level for  $d'_D$ , and in a main effect of Equalization on *absolute sensitivity* (hypothesis 4). A change in slope between the two equalization schemes can also yield an interaction between Type of Stimulus and Equalization on  $d'_D$ . The absence of interaction between Type of Stimulus and Equalization on *absolute sensitivity* and *absolute detectability* would support consistency for the WDE and/or SDE (hypothesis 5). Such an interaction on absolute sensitivity would also correspond to a three way interaction on  $d'_D$ .

### c. Results

Overall percentages of correct responses to silence (i.e., correct rejections) were high (mean = 96.11 %,  $SD = .74$  %). From the proportion of 'yes' responses  $Pr(\text{yes})$ , measures of detectability ( $d'_D$ ) and criterion ( $k$ ) were calculated for each participant as defined by the SDT (Macmillan & Creelman, 2005).  $d'_D$  was analyzed with a three-way ANOVA with Type of Stimulus (words vs. pseudo-words vs. complex sounds), Stimulus Level (5 levels, from 0 to +20 dB-A and from +4 to +24 dB-SPL, with 5 dB

steps) and Equalization (dB-A vs. dB-SPL) as within-participant factors. An alpha level of .05 was used after Greenhouse-Geisser correction for all statistical tests. Differences between stimulus types can occur on the slope of a psychometric curve, but not at lowest and highest presentation levels for which performance of all stimuli can be minimal (0-5%) or maximal (95-100%) (i.e., floor and ceiling effects). For our data, a ceiling effect appeared between levels 4 and 5 ( $p = .97$ ). This suggested that performance between the three types of stimulus could only differ between levels 1 to 3, i.e., on the slope of the psychometric curve. Based on this a priori hypothesis, planned comparisons (local contrasts) were carried out on these levels for significant interaction between Stimulus Level and Type of Stimulus or Equalization. For significant interaction between Type of Stimulus and Equalization, post-hoc bilateral paired  $t$ -tests, with Tukey correction (Howell, 1998), were performed.

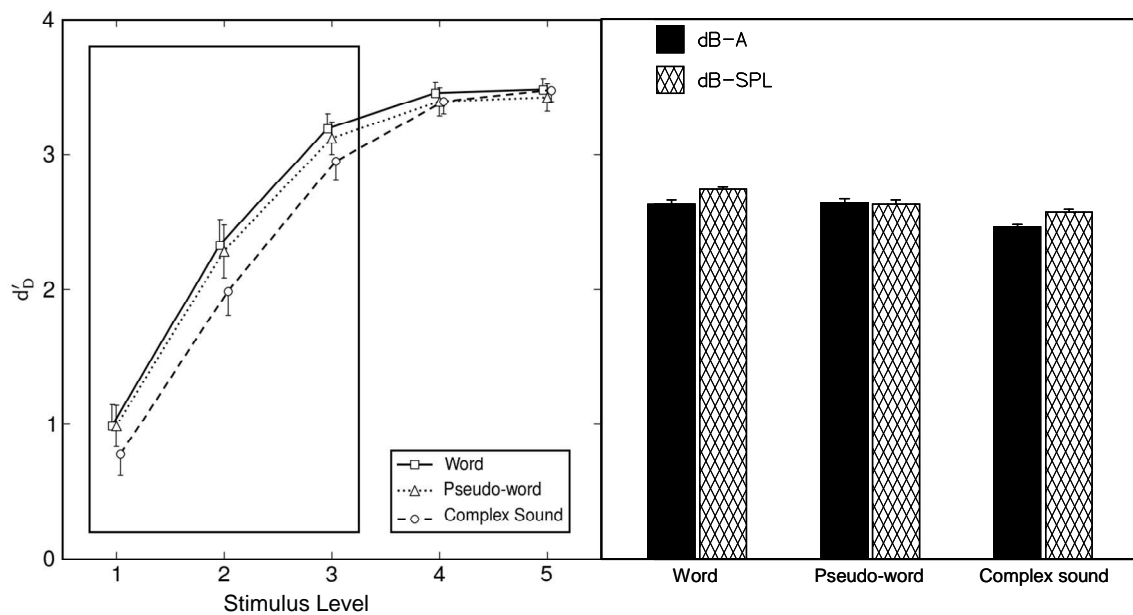


Figure 6. Analysis with signal detection theory. *Left panel:*  $d'_D$  presented as a function of Stimulus Level (1 to 5; 0 to +20 dB-A and +4 to +24 dB-SPL), Type of Stimulus (word/pseudo-word/complex sound) and Equalization. The error bars are the standard errors. *Right panel:*  $d'_D$  presented as a function of Equalization (dB-A and dB-SPL) and Type of Stimulus (word/pseudo-word/complex sound). The error bars are the standard errors.

For  $d'_D$  analysis (see Figure 6), the main effect of the Type of Stimulus ( $F(2,38) = 18.17, p < .001$ ), the main effect of the Stimulus Level ( $F(4,76) = 218.15, p < .001$ ) and the interaction between the Stimulus Level and the Type of Stimulus

( $F(8,152) = 2.67, p = .009$ ) were significant. For levels 1 to 3 (i.e., 0 to +10 dB-A or +4 to +14 dB-SPL), planned comparison revealed that words were detected better than complex sounds ( $F(1,19) = 21.43, p < .001$ ) and pseudo-words were detected better than complex sounds ( $F(1,19) = 26.79, p < .001$ ), consistent with hypothesis 2b. However, no difference was observed between words and pseudo-words ( $F(1,19) = .7, p = .41$ ), inconsistent with hypothesis 2a. The ANOVA also revealed a significant effect of Equalization ( $F(1,19) = 7.41, p = .014$ ) and a significant interaction between the Stimulus Level and the Equalization ( $F(4,76) = 5.43, p < .001$ ), consistent with hypotheses 1 and 4, as also confirmed by the analyses on *absolute sensitivity* and *detectability* presented below. Planned comparisons revealed that participants obtained better detection performance ( $F(1,19) = 16.70, p < .001$ ) in dB-SPL than in dB-A at level 1 only (i.e., 0 dB-A or +4 dB-SPL). Finally, the interaction between the Equalization and the Type of Stimulus was significant ( $F(2,38) = 3.49, p = .04$ ), consistent with hypothesis 3 or 4. Post-hoc analysis revealed that participants obtained better detection performance in dB-SPL than in dB-A for complex sounds only ( $p = .03$ ), but not for words ( $p = .07$ ) or pseudo-words ( $p = .99$ ). Finally, the interaction between Equalization, Stimulus Level and Type of Stimulus (hypothesis 5) was not significant ( $F(8,152) = 1.36, p = .22$ ).

The criterion  $k$  is equal to 1.81 ( $SD = .32$ ). When compared to the criterion observed in Experiment 1, a t-test (two-tailed) revealed that the difference was not significant ( $t(37) = 1.81, p = .08$ ).

For the *absolute sensitivity* (related to the slope of the psychometric function, see Experiment 1), the ANOVA revealed only a significant main effect of Equalization ( $F(1,19) = 6.25, p = .022$ ) consistent with hypothesis 4. For the *absolute detectability* (related to the horizontal position of the psychometric function), the ANOVA revealed a significant main effect of Type of Stimulus ( $F(2,38) = 5.17, p = .01$ ) consistent with hypothesis 3. The speech stimuli (words and pseudo-words) were both better detected than the complex sounds ( $ps < .03$ ). A significant main effect of Equalization ( $F(1,19) = 4.80, p = .04$ ) was also observed for the *absolute detectability*: the stimuli equalized in dB-SPL were better detected than those equalized in dB-A, consistent with hypothesis 1. No interaction was observed between Equalization and Type of Stimulus, for the *absolute sensitivity*

( $F(2,38) = .12, p = .89$ ) nor for the *absolute detectability* ( $F(2,38) = .18, p = .84$ ), consistent with hypothesis 5.

Finally, separate ANOVAs on  $d'_D$  for the two Equalization schemes were also performed and revealed a significant effect of Type of Stimulus in both cases ( $F(2,38) = 13.21, p < .0001$  for dB-A, and  $F(2,38) = 10.21, p < .0001$  for dB-SPL), consistent with hypothesis 2. Post-hocs revealed that both words and pseudo-words were better detected than complex sounds ( $p < .001$  and  $p = .002$  respectively for dB-A and dB-SPL), consistent with hypothesis 2b (SDE) but not 2a (WDE).

#### **d. Discussion**

Experiment 2 showed that, near auditory threshold (i.e., between 0 and +10 dB-A or between +4 and +14 dB-SPL), detection performance was better for speech stimuli (words and pseudo-words) than for non-speech stimuli (complex sounds). This result strongly suggests that when auditory stimuli were difficult to detect, listeners' knowledge facilitated the detection of phonological sounds over meaningless non-phonological sounds. In contrast to Experiment 1, detection performance did not differ between words and pseudo-words.

The greater *absolute sensitivity* observed for dB-A than for dB-SPL equalization suggests that the dispersion of loudness across the stimuli was narrower in dB-A than in dB-SPL equalization. The effect of equalization on  $d'_D$  and *absolute detectability* indicates that the stimuli equalized in dB-SPL were perceived louder than those equalized in dB-A. This is likely due to the difficulty to match the two equalization schemes to produce equal average loudness. Most importantly, for these two measures (*absolute sensitivity* and *absolute detectability*), no interaction was observed between the equalization types and the stimulus types. As also confirmed by the two separate ANOVAs, the effect of stimulus type was observed for both equalization schemes. These findings thus suggest that the effect of phonological knowledge on detection does not depend on the equalization method, i.e. it is not influenced by small changes in loudness. Consequently, the observed effects can more confidently be interpreted as an influence of knowledge rather than differences due to energetic characteristics of the stimuli.



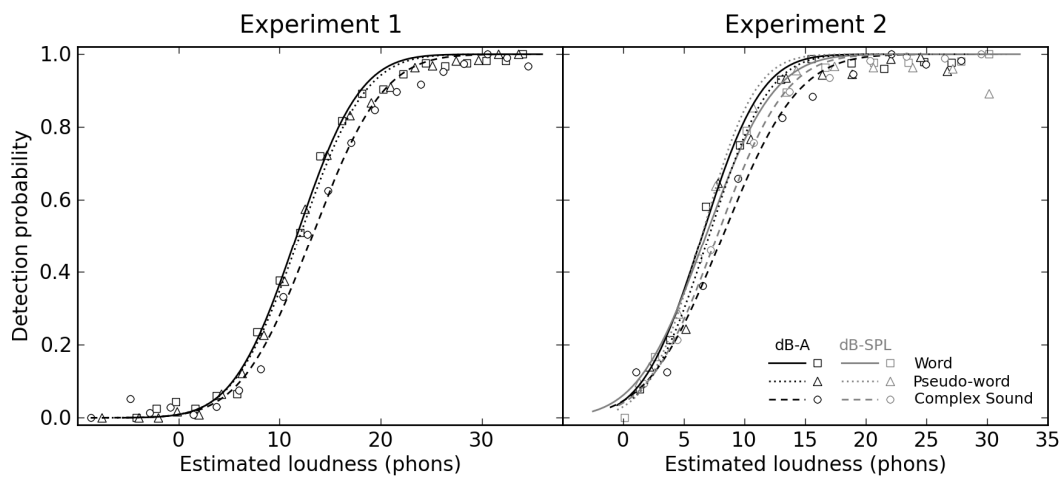
In Experiments 1 and 2, detection performance differed for a given level. For example, at 0 dB-A, the percentage of detection is about 5 % in Experiment 1, but about 15 % in Experiment 2 (for stimuli equalized in dB-A). This effect is probably due to differences in the ranges of presentation levels used in the two experiments: Experiment 2 restricted the investigation to a smaller range of stimulus levels than did Experiment 1, leading to smaller contrasts between the highest and the lowest level in Experiment 2. The increased ability to detect sounds in a small range of levels is consistent with the study of Luce and Green (1978) that showed better detection performance when the range spanned over 10 to 20 dB than when the range was larger (more than 20 dB, as in Experiment 1 of our study). In consequence, a ceiling effect was observed. This ceiling effect can explain the interaction observed on  $d'_D$  between Equalization and Type of Stimulus. Moreover, the absence of such an interaction on *absolute sensitivity* and *detectability* is consistent with this explanation.

#### 4.2.4 Loudness model

Although the stimuli used in Experiment 1 and 2 have been designed to be energetically as similar as possible, the differences remaining between the stimulus types could be crucial for detection. To be able to conclude that the observed detection differences are not due to physical differences, detection probability based solely on the energetic properties of the sounds must be evaluated. Since human listeners normally have some knowledge of a language, a top-down influence of knowledge — if it exists — would affect their detection performance. Therefore, detection probability based solely on energetic properties cannot be evaluated by humans, but can be evaluated by models that would be oblivious of any language.

Detection probability depends on loudness, even near threshold (Buus, M $\ddot{u}$ sch, & Florentine, 1998). In the study of Kewley-Port (1991) on vowel detectability, the best predictor was a loudness estimate based on excitation pattern, i.e. estimating the energetic level in the auditory nerve. However, the vowels used in this experiment were steady-state whereas the stimuli used in our study were fluctuating in time. Glasberg and Moore (2002) proposed a model, also based on excitation patterns, but specifically designed to handle time-varying sounds. In this model, the instantaneous loudness, available every millisecond, is derived from a specific loudness pattern,

itself derived from the excitation pattern. This instantaneous loudness is then integrated over time to produce a short-term loudness. This short-term loudness can also be integrated over longer time periods to produce a long-term loudness. This model successfully predicts the loudness of amplitude modulated sounds. Although this loudness model performs temporal integration, it is also solely based on the energetic level in the auditory nerve. If detection of auditory stimuli is only based on the energetic level induced in the auditory nerve, the main effect of Type of Stimulus should disappear when detection probability is expressed as a function of estimated loudness.



*Figure 7.* Detection probability as a function of estimated loudness for the two experiments. The curves are cumulative Gaussian fitted to the raw data. The symbols show the detection probability averaged on different loudness intervals uniformly distributed on the loudness range.

To be detected, a stimulus does not need to have a high loudness during the entire sound (> 500 ms), but only during a short period of time. Consequently, the estimated loudness for a stimulus used in our study is the maximum over time of the short-term loudness calculated with the model of Glasberg and Moore (2002). The measured performances in the two experiments are plotted as functions of the estimated loudness in Figure 7.

Because each stimulus has a specific loudness and was presented at different levels to the different participants, the data were sparsely represented and a linear mixed model had to be used to evaluate the effects. Such a model with Estimated loudness and Type of Stimulus (and Equalization for Experiment 2) as fixed effects and Subject as random intercept was fitted to the data. The analysis of variance on this

model revealed a main effect of Type of Stimulus for both experiments ( $F(2,8754) = 20.46$ ,  $p < .001$  in Experiment 1;  $F(2,7209) = 16.79$ ,  $p < .001$  in Experiment 2) as well as a word detection effect in Experiment 1 ( $t(8754) = -2.34$ ,  $p = .02$ ). Also, in agreement with the predictions made for the Experiment 2, the effect of Equalization was not significant in this representation ( $F(1,7209) = .02$ ,  $p = .90$ ).

The fact that the main effect of Type of Stimulus was still observed in both experiments indicates that the energetic properties accounted for by the loudness model cannot predict the observed effect, and therefore indicates that other properties of the stimuli have influenced the detection probability. This model estimates the loudness based on the energy evoked by a sound in the auditory nerve, thus corresponding to early stages of auditory processing. The present result indicates that features extracted from the sound at later stages of auditory processing also influence loudness and hence detectability. One feature that could be considered to explain the SDE is pitch. Indeed, the pitch sensation elicited by the complex sounds was weaker<sup>2</sup> than that of the words and pseudo-words. Pitch perception is elicited by a specific spectral or temporal coherent pattern (see Plack, Oxenham, Fay, & Popper, 2005, for a review) and thus requires more advanced sound processing than the evaluation of energetic level (He & Trainor, 2009). This hypothesis could be assessed by using unvoiced speech instead of voiced words and pseudo-words as used in the current study. Unvoiced speech sounds similar to whispered speech, but is artificially generated from recorded voiced speech by removing the periodicity. However, the pitch strength of the words was the same as that of the pseudo-words. Consequently, the WDE observed in Experiment 1 cannot be explained by pitch strength differences. This suggests that the WDE – and consequently the SDE – are due to other features such as lexical and phonological content as discussed in the next section.

#### **4.2.5 General Discussion**

The main aim of the present study was to investigate whether phonological and lexical knowledge could facilitate lower processing levels, such as auditory signal detection. Words, pseudo-words and complex sounds, energetically matched as closely as possible, were presented from inaudible to audible levels, and participants performed a detection task followed (Experiment 1) or not (Experiment 2) by a 2AFC

recognition task. Experiments 1 and 2 showed a Speech Detection Effect (SDE): near the auditory threshold, phonological stimuli (words and pseudo-words) were better detected than non-phonological ones (complex sounds). In addition, in Experiment 1, where participants were also engaged in a second task (i.e., recognition task), phonologically meaningful stimuli (words) were better detected than phonologically meaningless ones (pseudo-words). This suggests that the recognition task may have had an effect on auditory detection and allowed showing a Word Detection Effect (WDE). However, this WDE was not as a robust phenomenon as the SDE and was not replicated in the second experiment where the recognition task was removed. To our knowledge, such cognitive facilitation effects on auditory detection has not been previously reported and thus further investigations should be conducted to specifically assess this effect.

#### **a. Speech Detection Effect**

A novel feature of our study was to ensure that the three types of stimuli (words, pseudo-words and complex sounds) were energetically matched as closely as possible in order to ensure that differences observed between the three types of stimuli were not due to energetic features only. Analyses estimating the loudness of the stimuli on the basis of these purely energetic features, showed that loudness differences cannot explain the difference in detectability between speech and non-speech stimuli. Only extracted features such as phonological content are likely to explain these results. The influence of knowledge on auditory detection is in agreement with the study of Merikle and Reingold (1990), conducted in the visual modality. Indeed, using a visual subliminal detection paradigm, they showed that words were more readily detected than non-words, i.e. they showed a WDE. Our present study extended their results to the auditory modality and thus suggests that linguistic knowledge could facilitate lower level tasks for the two modalities. Our finding can be integrated into auditory models that simulate processing facilitation thanks to perceivers' knowledge. McClelland and Elman (1986) proposed to explain linguistic facilitations with a cognitive model in which lexical and semantic processes can influence lower-level acoustic and phonetic processes. Samuel and colleagues (Kraljic & Samuel, 2005 ; Rapp & Samuel, 2002 ; Samuel, 1997 ; Samuel & Pitt, 2003) have shown that linguistic-based facilitation can affect phonemic level of word processing. Our study

extends these previous findings and suggests that phonological knowledge also operates on the first level of auditory processing. Indeed, the ability to detect phonological stimuli, which might carry lexical information, more easily than non-phonological ones might be important for human communication. It could be helpful to react quickly when a speech stimulus, which presents a social interest, emerges in our environment.

## **b. Recognition without awareness?**

In Experiment 1, detection measure was chosen as the indicator of conscious awareness. According to Merikle and Reingold (Merikle & Reingold, 1990 ; Reingold & Merikle, 1988), a valid measure should meet the criteria of being exhaustive and exclusive. While it seems difficult to meet the exclusiveness criterion (see discussion in Reingold, 2004), it is important to use an exhaustive measure of conscious perceptual information (Merikle & Reingold, 1990 ; Reingold & Merikle, 1988). Some authors used discrimination tasks as indicators of awareness (as for example in Greenwald et al., 1995; Kouider & Dupoux, 2004), but discrimination tasks cannot rule out the possibility that partial or weak stimulus awareness explains indirect effects (a stimulus, which is not discriminated, might be consciously detected). In contrast, detection tasks asking for presence versus absence have great sensitivity. As suggested by Snodgrass, Bernat and Shevrin (2004a), this task appears to be the best candidate (as compared to identification or semantic classification) for meeting the exhaustiveness criterion and thus for measuring awareness.

The present study is the first experiment to investigate a potential auditory dissociation between detection and higher levels of processing in an ecological presentation, i.e., with a reduction of the level of presentation of the stimuli and without signal distortion or masking. Previously, Urban (1992) cited studies that used decreased levels of presentation to modify auditory detectability. However, no conclusion could be drawn from these studies because they used different protocols and reported contradictory results. With masking studies, Shipley (1965) did not observe any dissociation between detection and recognition for masked tones, whereas Lindner (1968) did observe such a dissociation when signaling to the participants that recognition was possible even without detection. More recently, using time-compressed and masked primes, Kouider and Dupoux (2005) suggested

dissociation between categorization and semantic processing for speech sounds. When participants were not able to categorize a stimulus (but when they might have been probably able to detect it), an auditory repetition priming effect was observed. However, the use of time-compressed masked primes, adapted from the visual subliminal priming research, strongly distorted the auditory stimuli (while visual stimuli remained the same in visual subliminal priming), and is far from any ecological presentation.

Experiment 1 did not yield unambiguous evidence for auditory unconscious perception. Indeed, subjective and objective threshold approaches did not lead to the same conclusions: while the first approach suggests that words were recognized when participants did not detect the stimulus, the second one did not show dissociation between detection and recognition for all types of stimulus. One may suggest that these findings are consistent with the possibility that participants need at least some degree of stimulus awareness to perform a correct recognition response following an absence of stimulus detection. However, Merikle and Reingold (1990) have argued that the qualitative difference between the “detect” (words and non-words recognition was observed after stimulus detection) and the “non-detect” state (only word recognition was observed after an absence of stimulus detection) supports the validity of the subjective measure of conscious awareness. This would suggest that auditory dissociation could be observed only between the subjective detection threshold and the objective detection threshold, as it was previously reported (Cheesman & Merikle, 1984, 1986).

### **c. Conclusion**

This study is the first to investigate the detection of speech and non-speech sounds in the auditory modality. The influences of knowledge suggested by the SDE and the WDE indicate that all levels of auditory processing, from early encoding in the auditory nerve to linguistic processing, are involved in this detection task. Consequently, this method could be used to explore the relevant features that are unconsciously extracted from the sounds to construct conscious percepts. In particular, this method could be used to identify the missing links between low-level auditory models (e.g. Bleeck, Ives, & Patterson, 2004) and higher-level sound perception models (e.g. Kiebel, von Kriegstein, Daunizeau, & Friston, 2009). This

would help explaining how higher-level tasks, such as recognition, are influenced by unconscious lower-level processing when the stimuli are familiar, as it has been suggested in the current study for subjective unconscious perception of words, and similarly in the study of Merikle and Reingold (1990) for the visual modality. Future research will have to assess this effect in order to confirm that only stimuli containing relevant information can be unconsciously perceived at the processing level required by the task.

## Acknowledgements

The work described in this paper was supported by a doctoral grant from the Ministère de l'Éducation Nationale et de la Recherche of France, the grant ACI "Junior Research Team", the grant ANR "PICS" and a grant of the UK-Medical Research Council (G9900369). The authors wish to thank Samuel Garcia for providing technical assistance, Alexis Hervais-Adelman and Neil A. Macmillan for their helpful comments on this work, and Brian C.J. Moore for suggesting to apply his loudness model to the stimuli.

## References

- American National Standard Institute. (1995). ANSI S3.7-1995 (R2003), Methods for coupler calibration of earphones. New-York.
- American National Standard Institute. (2004). ANSI S3.6-2004, Specifications for audiometers. New-York.
- Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology*, 4(2), 207-228. doi:10.1016/0010-0285(73)90012-1
- Bleeck, S., Ives, T., & Patterson, R. D. (2004). Aim-mat: The Auditory Image Model in MATLAB. *Acta Acustica united with Acustica*, 90(4), 781-787.
- Burgund, E. D., Guo, Y., & Aurbach, E. L. (2009). Priming for letters and pseudoletters in mid-fusiform cortex: examining letter selectivity and case invariance. *Experimental Brain Research*, 193(4), 591-601. doi:10.1007/s00221-008-1661-9
- Buus, S., Müsch, H., & Florentine, M. (1998). On loudness at threshold. *The Journal of the Acoustical Society of America*, 104(1), 399-410.
- Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 11, 63-65.

- Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, 36(4), 387-395.
- Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, 40(4), 343-367.
- Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 521-533.
- Connine, C., Blasko, D., & Titone, D. (1993). Do the Beginnings of Spoken Words Have a Special Status in Auditory Word Recognition? *Journal of Memory and Language*, 32(2), 193-210. doi:10.1006/jmla.1993.1011
- Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience*, 23(8), 3423-3431.
- Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F., et al. (1998). Imaging unconscious semantic priming. *Nature*, 395(6702), 597-600. doi:10.1038/26967
- Deutsch, A., & Bentin, S. (1994). Attention mechanisms mediate the syntactic priming effect in auditory word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 595-607.
- Doyle, J., & Leach, C. (1988). Word superiority in signal detection: Barely a glimpse, yet reading nonetheless. *Cognitive Psychology*, 20(3), 283-318. doi:10.1016/0010-0285(88)90001-1
- Fisk, G. D., & Haase, S. J. (2005). Unconscious perception or not? An evaluation of detection and discrimination as indicators of awareness. *The American Journal of Psychology*, 118(2), 183-212.
- Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., & Kleinschmidt, A. (2004). Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cerebral Cortex*, 14(3), 247-255.
- Glasberg, B. R., & Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331-342.
- Grainger, J., Bouttevin, S., Truc, C., Bastien, M., & Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: a comparison of dyslexic and normal readers. *Brain and Language*, 87(3), 432-440.
- Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1158-1176. doi:10.1037/0096-1523.20.6.1158



- Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible (“subliminal”) stimuli: dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, 124(1), 22-42.
- Haase, S. J., & Fisk, G. D. (2004). Valid distinctions between conscious and unconscious perception? *Perception & Psychophysics*, 66(5), 868-871; discussion 888-895.
- He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *The Journal of Neuroscience*, 29(24), 7718-8822. doi:10.1523/JNEUROSCI.0157-09.2009
- Hoen, M., Meunier, F., Grataloup, C., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., et al. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Communication*, 49(12), 905-916. doi:10.1016/j.specom.2007.05.008
- Holender, D., & Duscherer, K. (2004). Unconscious perception: the need for a paradigm shift. *Perception & Psychophysics*, 66(5), 872-881; discussion 888-895.
- Howell, D. (1998). *Méthodes statistiques en sciences humaines*. Paris: De Boeck Université.
- Ives, D. T., & Patterson, R. D. (2008). Pitch strength decreases as F0 and harmonic resolution increase in complex tones composed exclusively of high harmonics. *The Journal of the Acoustical Society of America*, 123(5), 2670-2679. doi:10.1121/1.2890737
- Kewley-Port, D. (1991). Detection thresholds for isolated vowels. *The Journal of the Acoustical Society of America*, 89(2), 820-829.
- Kiebel, S. J., von Kriegstein, K., Daunizeau, J., & Friston, K. J. (2009). Recognizing sequences of sequences. *PLoS Computational Biology*, 5(8), e1000464. doi:10.1371/journal.pcbi.1000464
- Kouider, S., & Dupoux, E. (2004). Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological Science*, 15(2), 75-81.
- Kouider, S., & Dupoux, E. (2005). Subliminal speech priming. *Psychological Science*, 16(8), 617-625. doi:10.1111/j.1467-9280.2005.01584.x
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51(2), 141-178. doi:10.1016/j.cogpsych.2005.05.001
- Lindner, W. A. (1968). Recognition performance as a function of detection criterion in a simultaneous detection-recognition task. *The Journal of the Acoustical Society of America*, 44(1), 204-211.
- Luce, R. D., & Green, D. M. (1978). Two tests of a neural attention hypothesis for auditory psychophysics. *Perception & Psychophysics*, 23(5), 363-371.
- Macmillan, N. A. (1986). The psychophysics of subliminal perception. *Behavioral and Brain Sciences*, 9(01), 38-39. doi:10.1017/S0140525X00021427
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory : a user’s guide* (2<sup>nd</sup> ed.). Mahwah N.J.: Lawrence Erlbaum Associates.

- Maris, E. (2002). The role of orthographic and phonological codes in the word and the pseudoword superiority effect: an analysis by means of multinomial processing tree models. *Journal of Experimental Psychology: Human Perception and Performance*, 28(6), 1409-1431.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Mason, M. (1982). Recognition time for letters and nonletters: effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance*, 8(5), 724-738.
- McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 2(1), 80-91.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.
- McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics*, 22(3), 249-261.
- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10(8), 363-369. doi:10.1016/j.tics.2006.06.007
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375-407. doi:10.1037/0033-295X.88.5.375
- Merikle, P. M., & Cheesman, J. (1986). Consciousness is a "subjective" state. *Behavioral and Brain Sciences*, 9(01), 42. doi:10.1017/S0140525X00021452
- Merikle, P. M., & Reingold, E. M. (1990). Recognition and lexical decision without detection: Unconscious perception? *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 574-583. doi:10.1037/0096-1523.16.3.574
- Miller, J. (2000). Measurement error in subliminal perception experiments: Simulation analyses of two regression methods. *Journal of Experimental Psychology: Human Perception and Performance*, 26(4), 1461-1477. doi:10.1037/0096-1523.26.4.1461
- Moore, B. C., & Glasberg, B. R. (1987). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing Research*, 28(2-3), 209-225.
- Naccache, L., & Dehaene, S. (2001). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, 11(10), 966-974.

- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106(3), 226-254. doi:10.1037/0096-3445.106.3.226
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28(4), 661-677. doi:10.1017/S014271640707035X
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behavior Research Methods, Instruments & Computers*, 36(3), 516-524.
- Norris, D. (1994). A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1212-1232. doi:10.1037/0096-1523.20.6.1212
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23(3), 299-325.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Orfanidou, E., Marslen-Wilson, W. D., & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of Cognitive Neuroscience*, 18(8), 1237-1252. doi:10.1162/jocn.2006.18.8.1237
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: the word-superiority effect. *Psychological Review*, 89(5), 573-594.
- Perrin, F., & Grimault, N. (2005). Fonds Sonores (Version 1.0) [Sound samples]. Retrieved April 7, 2010, from <http://olfac.univ-lyon1.fr/unite/equipe-02/FondsSonores.html>
- Pilotti, M., & Beyer, T. (2002). Perceptual and lexical components of auditory repetition priming in young and older adults. *Memory & Cognition*, 30(2), 226-236.
- Plack, C. J., Oxenham, A. J., Fay, R., & Popper, A. N. (Eds.). (2005). Pitch: neural coding and perception. *Springer Handbook of Auditory Research*. New York: Springer Science Business Media Inc.
- Rapp, D. N., & Samuel, A. G. (2002). A reason to rhyme: phonological and semantic influences on lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 564-571.
- Ratcliff, R., Allbritton, D., & McKoon, G. (1997). Bias in auditory priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(1), 143-152.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81(2), 275-280.
- Reingold, E. M. (2004). Unconscious perception and the classic dissociation paradigm: a new angle? *Perception & Psychophysics*, 66(5), 882-887; discussion 888-895.

- Reingold, E. M., & Jolicoeur, P. (1993). Perceptual versus postperceptual mediation of visual context effects: evidence from the letter-superiority effect. *Perception & Psychophysics*, 53(2), 166-178.
- Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, 44(6), 563-575.
- Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, 32(2), 97-127. doi:10.1006/cogp.1997.0646
- Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, 48(2), 416-434. doi:10.1016/S0749-596X(02)00514-4
- Schacter, D. L., & Church, B. A. (1992). Auditory priming: implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 915-930.
- van Selst, M., & Merikle, P. M. (1993). Perception below the Objective Threshold? *Consciousness and Cognition*, 2(3), 194-203. doi:10.1006/ccog.1993.1018
- Shipley, E. (1965). Detection and recognition: Experiments and choice models. *Journal of Mathematical Psychology*, 2(2), 277-311. doi:10.1016/0022-2496(65)90006-4
- Shipley, E. F. (1961). *Detection and Recognition with Uncertainty* (PhD Thesis). University of Pennsylvania.
- Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(1), 64-75.
- Snodgrass, M., Bernat, E., & Shevrin, H. (2004a). Unconscious perception at the objective detection threshold exists. *Perception & Psychophysics*, 66(5), 888-895.
- Snodgrass, M., Bernat, E., & Shevrin, H. (2004b). Unconscious perception: A model-based approach to method and evidence. *Perception & Psychophysics*, 66(5), 846-867.
- Snodgrass, M., Shevrin, H., & Kopka, M. (1993). The Mediation of Intentional Judgments by Unconscious Perceptions: The Influences of Task Strategy, Task Preference, Word Meaning, and Motivation. *Consciousness and Cognition*, 2(3), 169-193. doi:10.1006/ccog.1993.1017
- Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), 480-498. doi:10.1037/a0013027
- Urban, M. J. (1992). Auditory subliminal stimulation: a re-examination. *Perceptual and Motor Skills*, 74(2), 515-541.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167(917), 392-393.
- Warren, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin*, 96(2), 371-383.

- Warren, R. M., & Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, 9, 358-362.
- Warren, R. M., & Sherman, G. L. (1974). Phonemic restorations based on subsequent context. *Perception & Psychophysics*, 16(1), 150-156.
- Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, 1(1), 59-85.  
doi:10.1016/0010-0285(70)90005-8

## 4.3 Bilan

Ces deux premières études se proposaient de déterminer si les connaissances de l'auditeur pouvaient avoir un effet précoce sur les premiers niveaux de traitement auditif, en particulier sur la détection des événements sonores complexes. Des mots, pseudo-mots et sons complexes, égalisés le plus finement possible au niveau énergétique ont été présentés à des niveaux d'intensité allant de tout à fait inaudible à clairement audible. Un contrôle attentif des caractéristiques physiques a permis d'écarter l'hypothèse selon laquelle la détection reposerait simplement sur les caractéristiques énergétiques des stimulations. Un effet de supériorité des stimulations phonologiques a été mis en évidence : les performances des participants étaient meilleures lorsqu'il s'agissait de détecter une stimulation de parole plutôt qu'une stimulation complexe non phonologique. De plus, lorsque les participants étaient engagés dans une tâche demandant un niveau de traitement de plus haut niveau (comme la reconnaissance) ou lorsqu'ils connaissaient par avance la nature de la stimulation (présentée par blocs), alors un effet de supériorité des mots était observé : les performances des participants étaient meilleures lorsqu'il s'agissait de détecter des mots plutôt que des pseudo-mots. Ainsi, il est proposé que les connaissances de l'auditeur, qu'elles soient lexicales ou phonologiques, facilitent la détection des stimulations auditives. Concrètement, pour un même niveau d'excitation énergétique au niveau du nerf auditif, un auditeur détecterait plus facilement les stimulations qui ont des composantes phonologiques. Si l'attention de l'auditeur sur les stimulations est accrue, alors il détecte plus facilement les stimulations qui ont des composantes lexicales, probablement parce que les mots représentent des stimulations auditives pertinentes pour lesquelles l'auditeur a des connaissances précises. Nous démontrons ainsi pour la première fois en modalité auditive que les connaissances stockées en mémoire ont un effet *top-down* sur le plus bas niveau de traitement qu'est la détection. Un récent modèle de compréhension des sons de parole est compatible avec un tel résultat (Kiebel, von Kriegstein, Daunizeau, & Friston, 2009). Dans ce modèle fonctionnel, l'effet *top-down* des connaissances a été simulé, ce qui conduit à des résultats compatibles avec ceux obtenus dans des tâches de reconnaissance dans des expériences comportementales.

La deuxième étude apporte aussi des résultats concernant la perception non consciente des stimulations auditives. Dans un paradigme de dissociation entre détection et reconnaissance, l'approche subjective met en évidence une reconnaissance sans détection pour les mots et non pas pour les pseudo-mots ou les sons complexes. Ce résultat est en accord avec des données obtenues en modalité visuelle (Merikle & Reingold, 1990) qui proposent que la reconnaissance sans détection ne s'observe que pour les stimulations les plus familières. En revanche, l'approche objective ne montre aucune reconnaissance possible sans détection. Ce résultat, déjà reporté pour des études en modalité visuelle (Fisk & Haase, 2005), suggère que le traitement inconscient se réalise pour des stimulations dont le niveau de perception se situe entre le seuil objectif et le seuil subjectif (Merikle & Cheesman, 1986). Par ailleurs, les conditions d'observation d'une telle dissociation étaient plutôt défavorables dans la mesure où le participant devait d'abord répondre à la tâche de détection avant de réaliser la tâche de reconnaissance : il y avait donc un délai entre la présentation de la stimulation et la mesure de perception. Par conséquent, il est nécessaire d'explorer dans de meilleures conditions expérimentales la perception auditive subliminale avant de conclure. Notamment, il sera fondamental d'explorer les différentes pistes qui ont été empruntées dans les études visuelles puisque des résultats différents ont été obtenus en fonction de la mesure de perception (directe/indirecte) et de la mesure du niveau de conscience (détection, catégorisation, etc.).

## Chapitre 5

# Explorations des processus activés à la suite d'une perception auditive non consciente

L'influence des stimulations perçues inconsciemment sur la perception ultérieure a été montrée en modalité visuelle dans de nombreuses études (Haase & Fisk, 2004 ; Holender & Duscherer, 2004 ; Reingold, 2004a, 2004b ; Snodgrass, Bernat, & Shevrin, 2004a, 2004b). En revanche peu d'études ont été réalisées en modalité auditive, et particulièrement pour les sons de parole (Kouider & Dupoux, 2005 ; Urban, 1992). Ce manque d'intérêt pour les sons de parole peut certainement s'expliquer par le fait que les stimulations auditives sont des stimulations complexes, dont les caractéristiques physiques sont difficiles à manipuler. Urban (1992) a d'ailleurs recensé plusieurs études qui n'ont rapporté aucun résultat significatif suite à la présentation subliminale de stimulations auditives lexicales. Deux études ont par ailleurs examiné la capacité de reconnaissance en l'absence de détection mais en utilisant des sons purs (Lindner, 1968 ; Shipley, 1965). De ces deux études ressortent des observations contradictoires : l'une montre une dissociation (Lindner, 1968) et l'autre n'observe aucune dissociation (Shipley, 1965) entre des processus de détection et de reconnaissance. La principale différence entre ces deux études réside dans le fait que les participants avaient pour consigne (Lindner, 1968) ou non (Shipley, 1965) qu'une reconnaissance était possible sans détection.

Au cours de l'Étude II, des signes de traitement lexical suite à une perception auditive subliminale ont été mis en évidence dans un paradigme de dissociation entre détection et reconnaissance. L'approche psychoacoustique quant à la création et la présentation des stimulations auditives utilisées nous a permis de contrôler de façon rigoureuse les caractéristiques énergétiques des stimulations. Grâce à une approche subjective, une dissociation a été observée entre les processus de détection et les



processus de reconnaissance des sons lexicaux de parole. Ces différences ne peuvent pas être attribuées aux caractéristiques énergétiques des stimulations mais plutôt aux caractéristiques lexicales des stimulations. Ce résultat suggèrerait alors que la présentation subliminale des stimulations auditives peut permettre des traitements cognitifs, à la condition que ces stimulations auditives contiennent une information lexicale.

Les études présentées dans le second axe de recherche se proposent d'explorer, de façon plus détaillée, dans quelles conditions expérimentales une perception auditive non consciente peut être observée, et quelles en sont les bases neuronales. Dans les études présentées ci-dessous, nous avons choisi de ne pas masquer énergétiquement ou informationnellement les stimulations, ni de les compresser temporellement (comme cela a été proposé par Kouider et Dupoux, 2004), mais de réduire l'intensité de présentation des stimulations : dans ces conditions, le bruit de fond physiologique est vu comme un masque (voir Étude IV, Discussion page 166). La première étude présentée (Étude III) nous a permis de tester quels sont les seuils de perception des stimulations auditives obtenus lors de différentes tâches. Par analogie à d'autres études réalisées en modalité visuelle (Kouider & Dehaene, 2007), nous avons choisi un niveau d'intensité pour lequel les participants n'étaient pas capables de catégoriser la stimulation auditive présentée, mais la détectaient partiellement. De plus, l'exploration des traitements cognitifs réalisés à la suite de la présentation subliminale de stimulations auditives a été étudiée avec une tâche indirecte (contrairement à l'Étude II), à savoir un paradigme d'amorçage (Étude IV). Enfin, les corrélats neuronaux associés à une perception auditive consciente et non consciente ont été explorés grâce une étude réalisée en EEG et deux techniques d'analyses complémentaires : les potentiels évoqués et la décomposition temporo-fréquentielle (Étude V).

## **5.1 Étude III : Les différentes mesures du niveau de conscience**

Le problème empirique lié à l'investigation de la perception non consciente est la mesure du niveau de conscience qui permet de déterminer précisément ce que les participants ne sont pas capables de réaliser consciemment suite à la présentation

d'une stimulation. La présentation de la stimulation peut être qualifiée de subliminale quand les participants échouent à la tâche demandée, par exemple à une tâche de détection ou de catégorisation. Ainsi, en fonction de la tâche demandée, la mesure obtenue peut faire référence à différents niveaux de conscience. Différentes tâches mesurant la conscience des participants ont été proposées dans la littérature (voir Section 2.4.2b) et il n'existe aucun consensus pour dire qu'une tâche est meilleure qu'une autre (voir par exemple, Sandberg et al., 2010). L'approche objective dans l'analyse des performances des participants, a été modélisée grâce à la théorie de détection du signal (TDS, Macmillan & Creelman, 2005). Elle est devenue courante pour mesurer le niveau de conscience, l'indice  $d'$  étant considéré comme le critère de sensibilité le plus fiable (Draine & Greenwald, 1998) permettant de discriminer deux stimulations. L'indice  $d'$  permet d'obtenir des performances de discrimination entre deux conditions. Lorsque l'une des deux conditions est apparentée à une absence de stimulation alors que l'autre condition contient la présentation d'une stimulation, la mesure de sensibilité  $d'$  peut être considérée comme un critère objectif de détection.

Dans l'Étude III, nous proposons d'évaluer les seuils de perception à différentes tâches : une tâche de détection (Condition 1), une tâche de catégorisation de parole, entre des mots et des sons complexes d'une part (Condition 2) et entre des pseudo-mots et des sons complexes d'autre part (Condition 3), et enfin une tâche de décision lexicale, entre des mots et des pseudo-mots (Condition 4). Cette étude permettra de connaître les niveaux d'intensités sonores correspondant à un niveau de conscience (1) pour lequel les participants ne sont capables ni de détecter ni de catégoriser une stimulation auditive, (2) pour lequel les participants sont capables de détecter mais pas de discriminer un son de parole, et (3) pour lequel les participants sont capables de détecter mais pas de catégoriser une stimulation lexicale.

### 5.1.1 Méthodologie

**Participants.** 8 participants (âge moyen de 23.72 ans  $\pm$  1.17, 5 femmes) de langue maternelle française et droitiers (quotient de latéralité = 76.55  $\pm$  6.60) selon le test "Edinburgh Handedness Inventory" (Oldfield, 1971), ont été inclus dans cette expérience. Aucun participant ne reporte de problèmes auditifs ou neurologiques. Ils ont été sélectionnés sur la base de leur audition qui a été testée grâce à une

audiométrie tonale réalisée pour chaque oreille : leur perte auditive est inférieure à 15 dB pour les fréquences de sons purs compris entre 250 et 8000 Hz (ANSI, 2004).

**Matériel.** Le matériel et l'appareillage utilisés au cours de cette étude sont les mêmes que ceux employés lors de l'Étude II.

**Procédure.** Chaque participant a réalisé une expérience comportant quatre conditions. Dans chacune des conditions, les stimulations étaient présentées à cinq niveaux différents d'intensité compris entre 0 et 20 dB-A, séparés par des intervalles de 5 dB. Chacune des stimulations n'a été présentée qu'une seule fois par participant à travers toutes les conditions. Dans la première condition, la tâche des participants consistait à effectuer une détection de stimulation auditive par une procédure oui/non. Des mots, des pseudo-mots, des sons complexes ou des silences ont été présentés de façon aléatoire (de manière analogue à l'Expérience 2 de l'Étude II). Dans les deuxième et troisième conditions, la tâche des participants consistait à effectuer une catégorisation de parole : les stimulations étaient soit des sons complexes et des mots (Condition 2), soit des sons complexes et des pseudo-mots (Condition 3). Dans la quatrième condition, la tâche des participants consistait à réaliser une catégorisation lexicale entre mots et pseudo-mots (Condition 4). L'ordre des tâches était fixe pour tous les participants pour que les consignes de catégorisation n'influencent pas les performances de la tâche de détection.

Un total de 360 mots, 360 pseudo-mots et 360 sons complexes ont été utilisés à travers toutes les conditions. 36 silences ont été ajoutés de façon aléatoire pour la Condition 1. Une croix de fixation, d'une durée variable entre 100 et 500 ms, apparaissait à l'écran. Une stimulation sonore était présentée (ou un silence dans la Condition 1). À la fin de la stimulation, la croix de fixation disparaissait et une réponse motrice à la tâche était demandée aux participants. Un nouvel essai commençait 500 ms après la réponse motrice. La durée de l'expérience était d'environ une heure et des pauses étaient imposées entre chaque bloc expérimental.

## 5.1.2 Résultats

La mesure objective de discriminabilité a été calculée en  $d'$  selon la théorie de la détection du signal : ont été considérés comme des fausses-alarmes les silences dans la Condition 1, les sons complexes dans les Conditions 2 et 3, et les pseudo-mots dans

la Condition 4 (Figure 32). Les résultats ont été analysés à l'aide d'une analyse de variance (ANOVA) selon la Condition (1, 2, 3 ou 4) et le niveau d'Intensité (5 intensités, de 0 à 20 dB-A). L'ANOVA montre un effet principal de la Condition ( $F(3,21) = 50.37, p < .001$ ) : les analyses *post-hocs* (*Least Significant Difference*, LSD Fisher) révèlent que les performances étaient meilleures pour la condition 1 (détection) que pour la condition 2 (catégorisation de mots vs. sons complexes) ( $p < .001$ ), la condition 3 (catégorisation de pseudo-mots vs. sons complexes) ( $p < .001$ ) et la condition 4 (catégorisation de mots vs. pseudo-mots) ( $p < .001$ ). Les performances entre les conditions 2 et 3 n'étaient pas différentes ( $p = .676$ ). Un effet principal du niveau d'Intensité ( $F(4,28) = 117.46, p < .001$ ) ainsi qu'une interaction significative entre le niveau d'Intensité et la Condition ( $F(8,56) = 9.14, p < .001$ ) étaient aussi observés : les analyses *post-hocs* (LSD Fisher) révèlent que les performances sont également meilleures pour la condition 2 et pour la condition 3 que pour la condition 4, mais seulement au-dessus de 5 dB ( $ps < .001$  et  $ps < .001$ , respectivement).

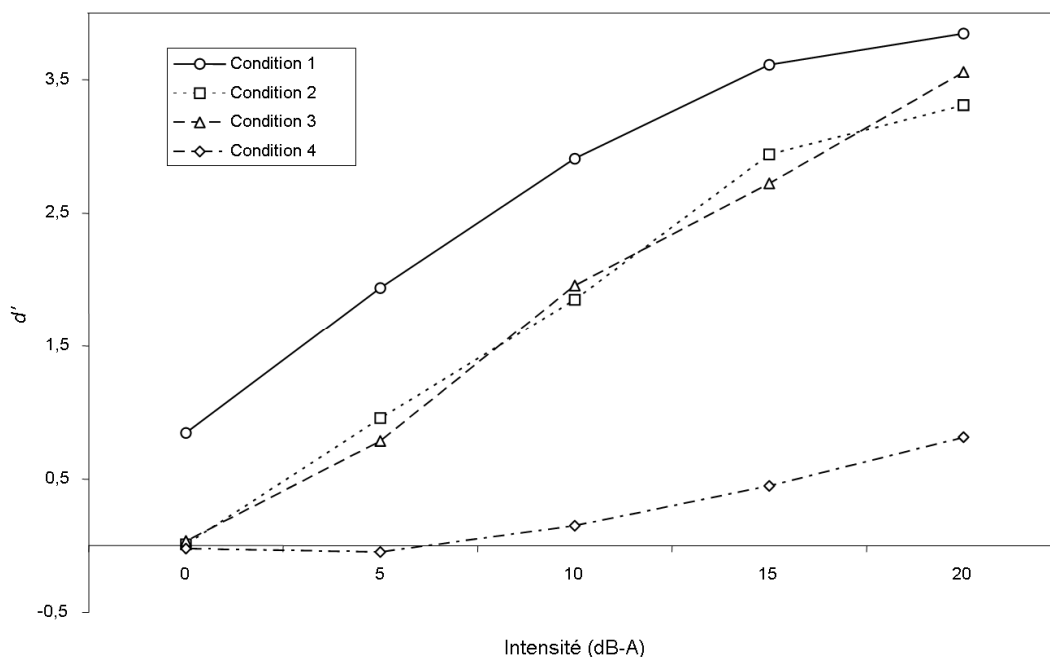


Figure 32 : Taux de discrimination en  $d'$  entre une stimulation auditive et un silence (Condition 1), un mot et un son complexe (Condition 2), un pseudo-mot et un son complexe (Condition 3) et entre un mot et un pseudo-mot (Condition 4).

Pour chaque condition, les pourcentages de réponses correctes (Figure 32) ont été analysés à l'aide d'une ANOVA selon la nature de la Stimulation (mots, pseudo-

mots et sons complexes en Condition 1, mots et sons complexes en Condition 2, pseudo-mots et sons complexes en Condition 3 et mots et pseudo-mots en Condition 4) et l'Intensité (5 intensités, de 0 à 20 dB-A). Les ANOVAs ont mis en évidence une interaction entre la nature de la Stimulation et l'Intensité pour toutes les conditions (Condition 1 :  $F(8,56) = 4.63$ ,  $p < .001$ ; Condition 2 :  $F(4,28) = 17.33$ ,  $p < .001$ ; Condition 3 :  $F(4,28) = 11.97$ ,  $p < .001$ ; Condition 4 :  $F(4,28) = 10.94$ ,  $p < .001$ ). Pour la Condition 1, les analyses *post-hocs* (LSD Fisher) ont révélé qu'entre 5 et 10 dB-A les performances de détection étaient meilleures pour les mots comparativement aux sons complexes ( $ps < .003$ ) et pour les pseudo-mots comparativement aux sons complexes ( $ps < .033$ ). Les performances de détection ne sont pas différentes entre les mots et les pseudo-mots ( $ps > .353$ ). Les analyses *post-hocs* (LSD Fisher) montrent que les performances de catégorisation étaient meilleures pour les mots comparativement aux sons complexes (Condition 2) entre 5 et 10 dB-A ( $ps < .005$ ), pour les pseudo-mots comparativement aux sons complexes (Condition 3) entre 10 et 15 dB-A ( $ps < .003$ ) et pour les mots comparativement aux pseudo-mots (Condition 4) entre 10 et 20 dB-A ( $ps < .003$ ). À 0 dB-A, les performances de catégorisation étaient significativement inversées pour les conditions 2, 3 et 4 ( $ps < .022$ ).

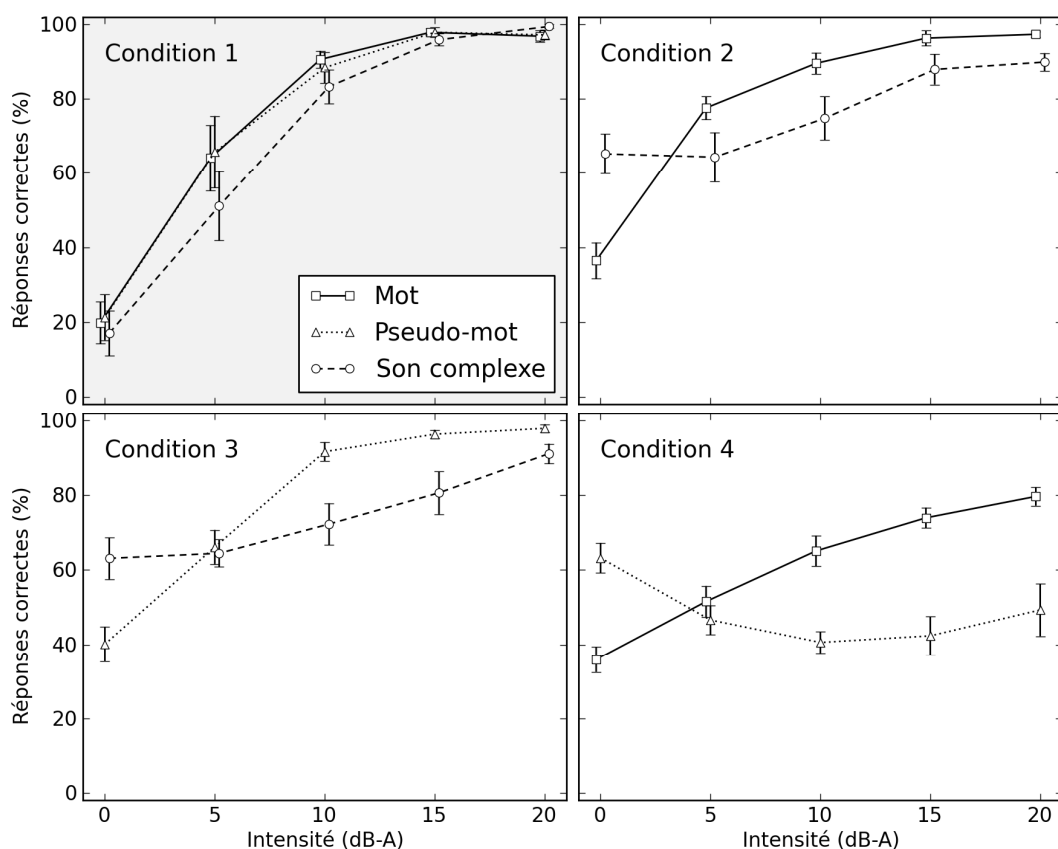


Figure 33 : Pourcentages de réponses correctes aux différentes tâches de détection et catégorisation. La condition 1 compare les performances de détection entre mots, pseudo-mots et sons complexes. La condition 2 compare les performances de catégorisation de parole entre mots et sons complexes, la condition 3 compare les performances de catégorisation de parole entre pseudo-mots et sons complexes et la condition 4 compare les performances de catégorisation de parole entre mots et pseudo-mots. Les mots sont représentés par des carrés, les pseudo-mots par des triangles et les sons complexes par des ronds.

Le biais a été calculé pour chaque condition (Table 4). Pour la tâche de détection (Condition 1), les participants montrent un biais de réponse pour la réponse « non » jusqu'à 10 dB-A. Pour les tâches de catégorisation, les participants montrent un biais de réponse en faveur des sons complexes (Condition 2), un biais de réponse en faveur des pseudo-mots (Condition 3) et un biais de réponse en faveur des mots (Condition 4) est observé à partir de 5 dB-A.

Table 4 : Biais de réponses observé en fonction de l'intensité pour chaque condition. Les valeurs en italiques correspondent aux niveaux d'intensité pour lesquels des différences de performances de catégorisation sont obtenues.

	<b>Intensité : 0</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>20</b>
<b>Condition 1</b>	1,520	0,828	0,350	-0,072	-0,106
<b>Condition 2</b>	0,455	<i>-0,162</i>	<i>-0,260</i>	-0,245	-0,265
<b>Condition 3</b>	0,357	-0,068	<i>-0,386</i>	<i>-0,419</i>	-0,239
<b>Condition 4</b>	0,417	-0,016	<i>-0,312</i>	<i>-0,488</i>	<i>-0,503</i>

### 5.1.3 Discussion

Pour un même niveau d'intensité, les performances de discrimination des participants varient en fonction de la tâche demandée. En accord avec les données de la littérature (par exemple Snodgrass, Bernat, & Shevrin, 2004a), les performances obtenues à la tâche de détection sont meilleures que celles obtenues aux deux tâches de catégorisation de la parole, qui sont elles-mêmes meilleures que celles obtenues à la tâche de catégorisation lexicale. Plus la tâche requiert un haut niveau de traitement cognitif, plus il est difficile de discriminer entre deux signaux (ou un signal et un silence) de même intensité. La mesure du niveau de conscience est donc reliée au choix de la tâche mais aussi du niveau d'intensité pour lequel la stimulation est présentée. La variation du niveau de conscience peut alors s'effectuer en fonction de la variation du niveau d'intensité et de la tâche demandée. Il serait alors possible de montrer si la stimulation objectivement déterminée comme étant non perçue consciemment à un niveau de détection, ou de catégorisation de parole ou de catégorisation lexicale peut influencer la perception d'une stimulation suivante. Autrement dit, si le contexte auditif non perçu consciemment, peut influencer la perception auditive consciente et de quelle manière. Comme la tâche utilisée pour déterminer le niveau de conscience peut influencer le niveau de traitement observé dans l'investigation de la perception non consciente, ce niveau pourrait-il également influencer le niveau de traitement réalisé à partir de la stimulation subliminale ?

Plusieurs études ont mis en évidence un effet de supériorité des mots (WSE) et de supériorité des pseudo-mots (PWSE) est mis en évidence dans des tâches de décision lexicale ou de catégorisation de parole (Grainger et al., 2003 ; Maris, 2002). Nous retrouvons cet effet dans l'Étude III : les participants obtiennent de meilleures performances de catégorisation (de parole ou lexicale) pour les mots comparativement aux pseudo-mots et aux sons complexes, et pour les pseudo-mots comparativement

aux sons complexes. Cet effet n'avait pas été obtenu lors de la tâche de reconnaissance de l'Expérience 1 de l'Étude II. Nous avons expliqué l'absence d'effet par le fait que la tâche de reconnaissance était à la fois retardée dans le temps et perturbée par une autre tâche (la tâche de détection). Par les présents résultats, nous pouvons confirmer cette explication : le WSE ainsi que le PWSE sont observés lorsqu'ils ne sont pas perturbés par une autre tâche ou trop retardés dans le temps. Toutefois, il est possible aussi que les WSE et PWSE observés soient dus à des biais de réponses. En effet, une stratégie conservatrice a été observée à 0 dB-A, alors qu'une stratégie libérale a été observée entre 5 et 20 dB-A. Cependant, l'observation d'une stratégie libérale ne suffit pas à expliquer un WSE ou un PWSE car pour les intensités ne montrant pas de WSE ou de PWSE, la stratégie libérale est toujours observée (par exemple, à l'intensité 20 de la Condition 2). Enfin si ces effets sont bien réels dans notre étude, il est important de souligner qu'ils ne pourraient pas être expliqués par des différences énergétiques, compte tenu des contrôles psychoacoustiques rigoureux qui ont été réalisés (voir Étude II, page 106).

## **5.2 Étude IV : « *Subliminal semantic priming in speech* »**

Nos études précédentes ont mis en évidence des indices de perception subjective non consciente seulement pour les stimulations auditives lexicales par l'intermédiaire d'une tâche directe. Ce résultat ne suggère pas qu'un traitement sémantique inconscient est possible mais qu'une stimulation lexicale non détectée peut être reconnue, de façon supérieure au hasard, par répétition. Même si c'est toujours un sujet de controverse (par exemple, Greenwald et al., 1995 ; Reingold & Merikle, 1988), il est possible que des signes de perception inconsciente de plus haut niveau pourraient être observés en employant une mesure indirecte mais aussi en utilisant une mesure du niveau de conscience moins stricte, comme la discrimination plutôt que la détection. C'est d'ailleurs dans ces conditions expérimentales que se sont placés Kouider et Dupoux (2005) pour tenter de mettre en évidence un amorçage auditif sémantique subliminal. Par analogie au système visuel, les auteurs ont utilisé des amorces qui étaient compressées dans le temps et masquées. Dans ces conditions, seul un amorçage de répétition a pu être mis en évidence. Il est possible que l'absence



d'effet sémantique soit due à leur manipulation des stimulations. En effet, les systèmes visuels et auditifs diffèrent considérablement dans leur gestion des aspects temporels et il est donc difficile de transposer un paradigme visuel à la modalité auditive sans prendre en considération cet aspect temporel.

L'Étude IV se propose d'étudier l'amorçage sémantique subliminal en audition en présentant des stimulations auditives non compressées ou déformée, mais en diminuant leur niveau d'intensité. Les résultats de l'Étude III nous ont permis de définir l'intensité pour laquelle les participants sont incapables de discriminer les mots des pseudo-mots.

Cette étude a fait l'objet d'une communication orale au *Medical Research Council (MRC) – Cognition and Brain Sciences Unit (CBSU)*, sur invitation du Dr. Étienne Gaudrain. À ce jour, elle fait l'objet d'un article en préparation, prochainement soumis à une revue à comité international de lecture. Elle sera également prochainement présentée au *14<sup>th</sup> Annual Meeting of the Association of Scientific Studies on Consciousness*, à Toronto (Canada) en juin 2010.

# SUBLIMINAL SEMANTIC PRIMING IN SPEECH

Jérôme Daltrozzo\*, Carine Signoret\*, Barbara Tillmann, & Fabien Perrin

CNRS UMR 5020, Claude Bernard Lyon University, France

\* These authors contributed equally to this work

## Abstract

Numerous studies have reported subliminal repetition and semantic priming in the visual modality. In contrast, only one study has reported an auditory repetition priming effect when the prime was time compressed and masked. In the present study, prime awareness was manipulated through a reduction of intensity level. These prime words were followed by semantically related, unrelated, and repeated target words (presented without intensity reduction), on which participants performed a lexical decision task (LDT) on the target. Participants' level of awareness of the prime was definite with a categorization task showing chance performance ( $d' = 0.08$ ). Participants with slower response times at the LDT showed semantic priming (faster response times for semantically related compared to unrelated targets) and negative repetition priming (slower response times for repeated compared to semantically related targets). This is the first report of unconscious semantic priming in the auditory modality.

## 5.2.1 Introduction

In how far stimuli are processed unconsciously has been a topic of considerable debate. While objective signs of semantic processing have been observed in reduced and altered states of consciousness (e.g. in sleep, see for a review Perrin, 2004 ; and in coma, see for a review Vanhaudenhuyse et al., 2008), it is less clear whether such mechanisms also arise under a waking, subliminal condition. This research faces methodological and theoretical questions, such as the demonstration of the absence of consciousness or how to measure unconscious effects (see for instance the special

issue in *Perception and Psychophysics*, 66 (5), 2004). A frequently used method for studying subliminal processing is the visual masked priming paradigm. Following the seminal study of Marcel (1983), Humphreys and collaborators (e.g. Evett & Humphreys, 1981) and Forster and collaborators (e.g. Forster & Davis, 1984), have developed this paradigm: the prime word is presented for a short time (usually less than 50 ms) and is surrounded by a forward and/or a backward visual mask. In these conditions, visual priming effects have been found at orthographic and morphologic levels (e.g. Ferrand & Grainger, 1992 ; Forster & Davis, 1984). Other studies showing phonological and semantic effects with this paradigm have received a lot of criticisms, notably because most of them did not provide an index of prime awareness (for reviews, see Van den Bussche, Van den Noortgate, & Reynvoet, 2009 ; Kouider & Dehaene, 2007). Using a measure of awareness, a response-window procedure, and a regression method, (Greenwald, Draine, & Abrams, 1996) showed that invisible primes could influence semantic processing (see also Dehaene et al., 1998 ; Dell'Acqua & Grainger, 1999 ; Draine & Greenwald, 1998 ; Kiefer & Brendel, 2006).

Subliminal perception has been considerably less investigated in the auditory modality. Similarly to studies investigating the visual modality, masking techniques (e.g. white noise or music masks) have been used to reduce the prime awareness. Unfortunately, most of these studies reported contradictory results (see for a review, Urban, 1992) allowing no clear conclusions. More recently, using masked and time-compressed primes, Kouider and Dupoux (2005) succeeded in showing a subliminal repetition priming effect. However, they did not observe subliminal semantic priming.

In the auditory modality, the scarceness of data and the absence of data showing subliminal semantic priming may be explained by the characteristics of the auditory domain, notably the temporal dynamics of stimulus presentation and the length of perceptual processes are very different from the visual domain. First, while a visual word can be fully presented without distortion in a very short time, time constrains speech intelligibility (Beattie, 1986) and speech production. Indeed, spoken words unfold over time, whereas printed words are available all at once. Second, perceptual processes are longer in the auditory compared to the visual domain (e.g. Anderson & Holcomb, 1995 ; Holcomb & Neville, 1990 ; Leinonen et al., 2009). Indeed, longer reaction time and N400 latency were obtained in semantic paradigms for short as compared to long inter-stimuli interval (ISI), this effect increasing in the auditory

modality (Anderson & Holcomb, 1995). This suggests that, in an auditory priming paradigm, participants need more time to process the prime and develop expectations for the target. Hence, in agreement with other (phonological or conceptual) priming studies, slow responders would show facilitation whereas fast responders would develop a lack of priming (Dupoux, 1993 ; Humphreys, 1983 ; Humphreys & Quinlan, 1988).

Transferring the visual subliminal priming paradigm to the auditory domain needs to take into account the physical differences between visual and auditory stimuli as well as the physiological differences between the two modalities. In the present study, we investigated auditory subliminal repetition and semantic priming, using primes presented at a low intensity of 10 dB-A. Using sounds that were not distorted but reduced in intensity, the internal noise could be considered to serve as a mask. Following the priming experiment, a prime categorization test (word/pseudo-word) was presented in order to check that participants were unable to categorize the prime. We expected facilitation only in the slow responders because at a short ISI, classically used in subliminal priming paradigm (e.g. Greenwald et al., 1996), longer perceptual processes are induced (see above Anderson & Holcomb, 1995).

## **5.2.2 Methods**

### **a. Participants**

Forty-five volunteers were tested: sixteen in a pre-experiment (12 females,  $22.1 \pm 0.4$  years) and 29 others in the main experiment. All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971), French native speakers, and did not report any hearing problems or history of neurological disease. They were paid for their participation and provided written informed consent to the study, which was conducted in accordance with the guidelines of the Declaration of Helsinki, and approved by the local Ethics Committee (CPP Lyon Sud-Est II, n°07/006). Participants were included in the main experiment only if their pure tone auditory threshold was below 15 dB-HL for frequencies between 250 Hz to 8000 Hz (American National Standard Institute, 2004). In the main experiment, a priming task, in which the participant performed a lexical decision task (LDT), was followed by two post-tests. Five participants were excluded from the main experiment because

of their poor performance at the LDT task and/or because of their performance being two standard deviations (SD) below the group performance during the post-tests (see details at the end of the Procedure section). Thus, 24 participants (16 females,  $21.5 \pm 0.3$  years) were included in the analysis of the main experiment.

## **b. Stimuli**

Two types of stimuli were used: words and pseudo-words. One-hundred-sixty words were selected from a French database (Lexique 2, New, Pallier, Ferrand, & Matos, 2001; New, Pallier, Brysbaert, & Ferrand, 2004; New, Brysbaert, Veronis, & Pallier, 2007). They were common singular monosyllabic nouns and contained 2-7 letters and 2-5 phonemes (e.g. “sable” [sand], “vache”[cow]). All words had a frequency of occurrence higher than 1 per million occurrences in books as well as in movies (subtitles). A list of pseudo-words was created from words, by mixing the phonemes: all phonemes of the words were used. Pseudo-words could be pronounced but did not have any meaning<sup>1</sup>. The number of phonemes was matched between words and pseudo-words. The average duration of words and pseudo-words were not significantly different and were respectively 521 ms (SD = 115 ms) and 539 ms (SD = 88 ms). To reduce differences in the perceived loudness, all stimuli were equalized to reach the same dB-A level (A-weighting roughly mimics the external and middle ear transfer functions; Kewley-Port, 1991).

Words and pseudo-words were uttered by the same female speaker and were recorded (32 bits, 44.1 kHz) using a Røde NT1 microphone, a Behringer Ultragain preamplifier, and a VxPocket V2 Digigram soundcard. The mean level of presentation was calibrated (American National Standard Institute, 1995) to reach 80 dB-A in a standard artificial ear (Larson Davis AEC101 and 824). All stimuli were played through a soundcard (Creative Soundblaster Audigy 2). A null (for primes and targets of the pretest and for targets of the main Experiment) or moderate digital attenuation (35 dB for primes of the main Experiment) was combined with an analog fixed attenuation (TDT PA4, one for each channel). This attenuation was analog rather than digital to prevent acoustical distortion at low levels of presentation. All stimuli were

---

<sup>1</sup> In a pretest, five other participants (2 Females,  $26 \pm 2.1$  years) assessed the phonological similarities between pseudo-words and words. They had to judge if the pronounced pseudo-words sounded like a word, and if this was the case, they had to write down the corresponding word. All pseudo-words for which words were proposed by at least two participants were excluded.

binaurally presented to participants through comfortable headphones (Sennheiser HD 250 Linear II) connected to a headphone buffer (TDT HB6).

### **c. Procedure**

#### ***Pretest***

The aim of this experiment was to show repetition priming and semantic priming for our experimental material and timing of presentation. Sixteen participants were presented a prime word followed by a word or a pseudo-word target at a comfortable hearing level (60 dB-A). They had to perform a LDT on the target as fast and accurately as possible. One hundred twenty prime-target pairs were presented in random order: sixty with a word target (20 semantically related to the prime, 20 semantically unrelated to the prime, and 20 repeated) and sixty with a pseudo-word target. All word stimuli were used across participants for each of the conditions. But for each participant, none of the words were repeated, except inside the pairs of the repeated condition. A fixation cross was displayed at the center of a monitor screen while the prime word was played in the headphone (i.e., the presentation duration of the prime and the fixation cross were equal). The target was presented 50 ms after the end of the prime (and the fixation cross) presentation. After the participant's lexical decision (left or right button press, sides counterbalanced across participants), participant had to press the space bar to start a new trial. The new trial started 200 ms after the key press.

#### ***Experiment***

In the priming experiment, twenty-nine participants performed a target lexical decision task in a subliminal priming paradigm, followed by two post-tests: a prime categorization task and a prime detection task. The prime was presented at a low intensity level (10 dB-A) and the target at a comfortable hearing level (45 dB-A). A 10 dB-A intensity level was chosen because previous experiments in our team suggested an absence of conscious categorization at this level (Signoret et al., in revision). The methodology was exactly the same than that of the pre-experiment, except for the level of presentation and except that the participants were not told about the presence of the prime.

The awareness of the prime was estimated with a prime categorization test, as it has been previously done (e.g. Kouider & Dupoux, 2005). A word ( $n = 60$ ) or a pseudo-word ( $n = 60$ ), presented at a low intensity level (10 dB-A, as in the priming phase of the main experiment), was randomly presented at the same time as the fixation cross. Participants were told to decide whether a word or a pseudo-word was presented during the fixation cross and to give their response (left or right button press, sides counterbalanced across participants) as accurately as possible and as soon as the fixation cross disappeared. Participants were told that the task was difficult and that they should not be discouraged by the difficulty. Fifty milliseconds after the participant's response, a second stimulus (a word or a pseudo-word), on which no task was requested, was presented at the same hearing level than that in the main experiment (45 dB-A, in order to keep the same intensity context between the main experiment and the prime categorization test, (Luce & Green, 1978). None of the stimuli were repeated between the main experiment and this post-test, but all stimuli presented in the post-test were used in the priming task across participants.

In order to eliminating results associated to the statistical deviance of some participants, we aimed to explore subliminal auditory perception in a homogenous group, in which participants were (1) able to perform the LDT in the main experiment and (2) unable to categorize the prime, but nevertheless (3) could detect it. To test for prime detection, participants performed a prime detection test for which a word ( $n = 60$ ) or a silence ( $n = 60$ ) was randomly presented at the same time as a fixation cross. To avoid participants searching for words, the prime detection task was performed before the prime categorization task. Analysis of the prime detection task showed an average accuracy of  $87.4 \pm 1.7\%$  and a  $d'$  sensitivity ( $d'_D$ ) of  $2.7 \pm 0.1$ . Five participants were excluded from the analysis of the main experiment because: (1) their performance at the LDT in the priming task was below to the mean accuracy minus two standard deviations, and/or (2) their categorization performance of the prime was two SD above the mean  $d'_C$  (at chance level, i.e. close to zero) at the prime categorization test, and/or (3) their detection of the prime was two SD below the mean  $d'_D$  at the prime detection test.

### 5.2.3 Results

Accuracy and correct RTs were analyzed by repeated ANOVAs with Relatedness (3 levels: related, unrelated, and repeated word pairs) as within-subjects factor and Rapidity (slow responders/fast responders) as between-subjects factor. Participants were separated into slow and fast responders with a median split. All reported  $p$  values were adjusted with the Greenhouse–Geisser correction for nonsphericity, when appropriate. The Least Significant Difference (LSD) test of Fisher was applied for post hoc comparisons as the Relatedness factor had only three levels (Howell, 1998).

#### a. Pretest

Participants performed the LDT on the target word with an averaged accuracy of  $96.2 \pm 0.7\%$ . There was a main effect of Relatedness [ $F(2,28) = 6.90$ ;  $p = .01$ ] with a better accuracy for semantically related ( $98.7 \pm 0.6\%$ ) and repeated word pairs ( $97.5 \pm 0.6\%$ ) compared to unrelated pairs ( $92.5 \pm 2.2\%$ ) ( $ps < .01$ ). There was no main effect of Rapidity (slow/fast) [ $F(1,14) = 1.68$ ;  $p = .22$ ] and no interaction with Relatedness [ $F(2,28) = 2.31$ ;  $p = .12$ ]. A main effect of Relatedness was also found with RTs [ $F(2,28) = 21.84$ ;  $p < .001$ ] with faster responses to repeated targets ( $921 \pm 38$  ms) compared to semantically related ( $1005 \pm 38$  ms) ( $p = .003$ ) and unrelated targets ( $1092 \pm 32$  ms) ( $p < .001$ ). Responses were also faster to semantically related compared to unrelated pairs ( $p = .002$ ). The main effect of Rapidity [ $F(1,14) = 28.85$ ;  $p < .001$ ] confirmed that the slow and fast responders' performance were significantly different but no interaction with Relatedness [ $F(2,28) = 0.41$ ;  $p = .67$ ]. (*Figure 1*). These results indicated that our material, when both primes and targets are presented at a comfortable hearing level (60 dB-A), elicited semantic and repetition priming effects.



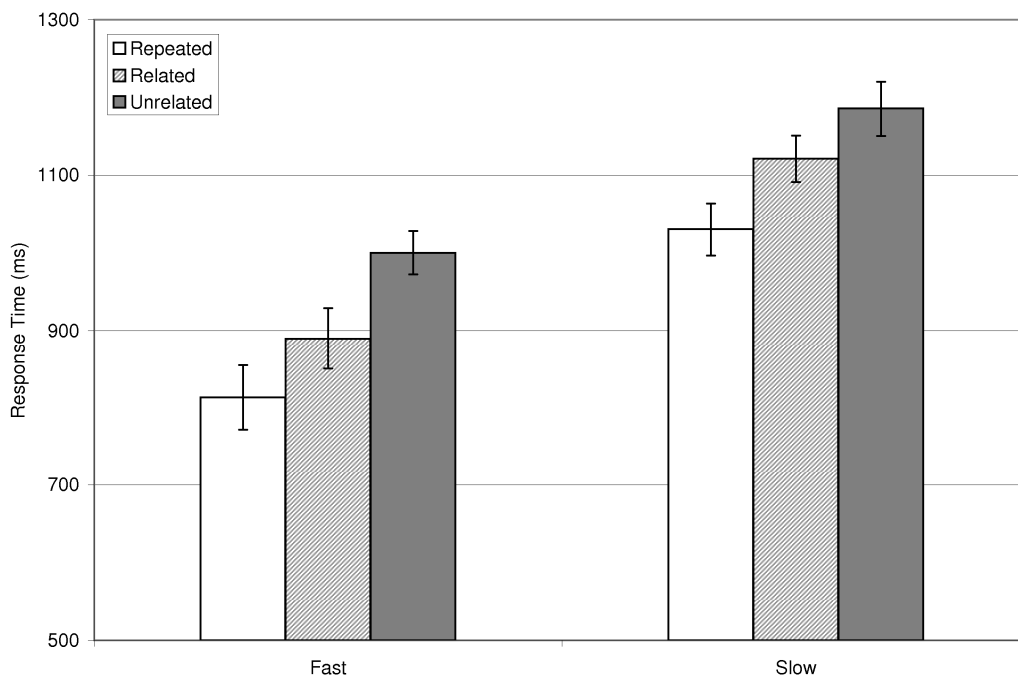


Figure 1. Response Time (Unit: milliseconds; with SD) to the lexical decision task in the Pre-experiment with semantically unrelated, semantically related, and repeated word pairs in slow ( $N=12$ ) and fast responders ( $N=12$ ). All differences are significant.

## b. Experiment

In the priming task, the participants' accuracy at the LDT was  $93.9 \pm 0.7\%$ . The accuracy did not differ between repeated ( $93.1 \pm 1.6\%$ ), semantically related ( $94.8 \pm 1.7\%$ ), and unrelated word pairs ( $93.5 \pm 1.1\%$ ) for both slow and fast groups [Relatedness:  $F(2,44) = 1.21$ ,  $p = .31$ ; Relatedness  $\times$  Rapidity:  $F(2,44) = 0.28$ ,  $p = .76$ ]. There was no main effect of Rapidity [ $F(1,22) = 1.97$ ;  $p = .17$ ]. For RTs (Figure 2), the performance of the fast responders ( $N = 12$ ) contrasted with those of the slow responders ( $N = 12$ ) as indicated by an interaction between Rapidity (slow/fast) and Relatedness [ $F(2,44) = 3.54$ ;  $p = .04$ ]. This interaction indicated that for the slower participants only, responses were faster to semantically related ( $1105 \pm 29$  ms) compared to unrelated ( $1153 \pm 28$  ms) ( $p = .034$ ) and repeated pairs ( $1166 \pm 28$  ms) ( $p = .008$ ). RTs did not differ between unrelated and repeated pairs ( $p = .549$ ). No significant differences were observed for the fast responders ( $ps > .470$ ). The main effect of Rapidity [ $F(1,22) = 19.8$ ;  $p < .001$ ] confirmed that the slow and fast responders' performance were significantly different.

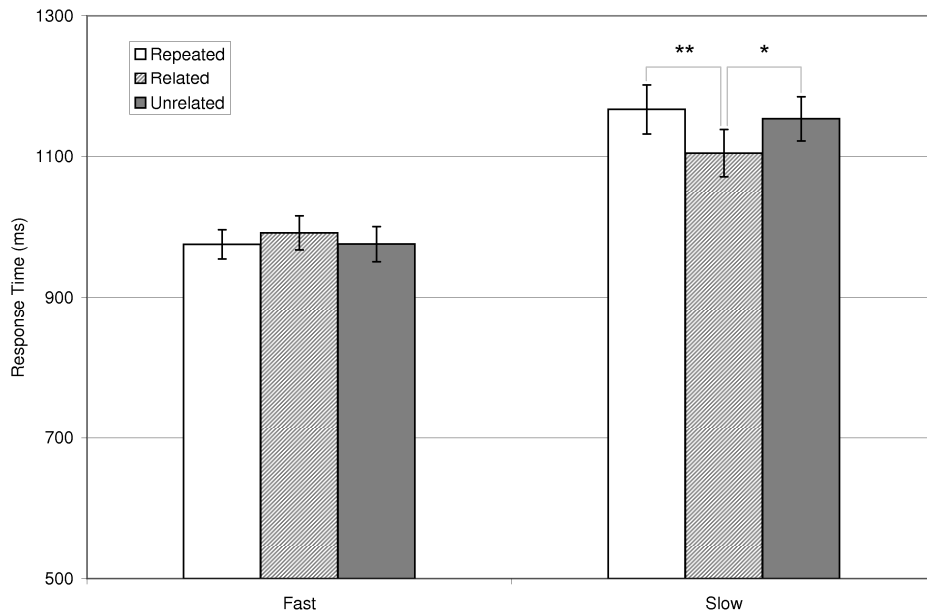


Figure 2. Response Time (Unit: milliseconds; with SD) to the lexical decision task in the main Experiment with semantically unrelated, semantically related, and repeated word pairs in slow (N=12) and fast responders (N=12). ( $p < .05$  \* ;  $p < .01$  \*\*).

In the post-test prime categorization task, average accuracy was  $52.0 \pm 1.2\%$  (chance level: 50%) with  $d'_C = .10 \pm .04$  (slow group:  $51.6 \pm 1.6\%$  with  $d'_C = .08 \pm .05$ ; fast group:  $52.4 \pm 1.8\%$  with  $d'_C = .12 \pm .06$ ).  $d'_C$  were not significantly greater than zero for the slow group ( $t(11) = 1.58, p = .14$ ) and the fast group ( $t(11) = 1.94, p = .08$ ), and did not differ significantly between the fast and slow groups ( $t(22) = .51, p = .61$ ), thus suggesting that participants were unable to categorize the prime.

Even if the estimation of the y-intercept of the direct measure (i.e. the prime categorization task) to the indirect measure (i.e. from the lexical decision task on the target of the priming task of the main experiment) (Greenwald, Klinger, & Schuh, 1995) is not necessary when  $d'$  was close to zero (e.g., Dell'Acqua & Grainger, 1999), we applied a regression procedure to ensure that priming effects were not induced by participants showing the greater sensitivity. The results show that the regression line obtained for the priming effects (1) between related and unrelated pairs and (2) and between related and repeated pairs. The first line regression was  $y = -120x + 29$  with an y-intercept that was significantly above zero, ( $t(22) = 2.77, p = .010$ ). The second line regression was  $y = -80.2x + 31.0$  with an y-intercept that was significantly above

zero, ( $t(22) = 2.70, p = .012$ ). As the effect of the indirect measure is observed when the direct measure is equal to zero, the regressions lines confirmed priming effects, for the entire group, when participants did not categorize the stimulus.

## 5.2.4 Discussion

By reducing intensity levels of primes, we showed speech priming effects in the absence of awareness (as measured by a prime categorization task). Participants were faster to discriminate target words that followed semantically related words than target words that followed unrelated words. Moreover, they showed faster responses for semantically related words than for repeated words. These two priming effects were observed for slow responders but not for fast responders.

### a. Subliminal semantic priming in speech

Our study is the first one to report subliminal semantic priming in the auditory modality. Participants responded faster to target words that followed semantically related words than to target words that followed semantically unrelated words. Even in the visual domain, only a few studies that have controlled properly the level of prime awareness, have reported unconscious semantic priming effects (for a review see Kouider & Dehaene, 2007). Dell'Acqua and Grainger (1999) showed a 30 ms faster word categorization (between words referring to an artifact or a natural thing) when the unconsciously perceived picture prime belonged to the same semantic category compared to a different semantic category. They further replicated their findings with target pictures and a picture naming task, showing a 22 ms unconscious facilitation between semantically related and unrelated pictures. Kiefer and Brendel (2006) reported unconscious semantic priming with visual word pairs and a LDT. They found a semantic priming effect of about 30 ms (see Figure 5 in Kiefer & Brendel, 2006) and the observed semantic priming was further confirmed with Event-Related Potentials. Interestingly, the size of the semantic priming we reported here (48 ms) is larger than what was found in these studies. This difference might be linked to the difference in modality as larger semantic priming has been observed for the auditory and the visual modality for conscious presentations (Poulin-Charronat & Tillmann, in revision). In addition this difference might be also explained by a

different procedure for the reduction of the prime awareness. Indeed, in these two studies (Dell'Acqua & Grainger, 1999 ; Kiefer & Brendel, 2006), the visual prime was shortly presented between a forward and backward mask. Although the visual prime is not physically degraded, the presentation of these two masks in close temporal conjunction may result in a single percept of a degraded prime (Scheerer, 1973). Possibly, a physical or a perceptual degradation of the prime might reduce the activation of its representation in the mental lexicon and yield weaker priming. By contrast, the primes of our present study were neither degraded physically nor at a perceptual level (the only degradation might result from the internal noise of the perceiver).

In the auditory modality, subliminal semantic priming has never been reported. The absence of semantic priming in the study of Kouider and Dupoux (2005) might be explained by the fact that they used time-compressed and masked primes and/or that they did not split their result pattern into low and fast responders. It is possible that unconscious semantic priming was present in their study but reduced by the time distortion of the prime. Indeed, Beattie (1986) has shown a high decrease of intelligibility performances with 60% time-compressed speech, thus preventing any auditory semantic analysis. Furthermore, as normal speech perception usually does not require the perception of time compressed words, it is likely that this kind of stimuli has a poor representation in the mental lexicon or leads to a poor access to the mental lexicon, and thus intrinsically evokes only weak semantic priming.

## **b. Subliminal repetition priming in speech**

Participants responded faster to target words that followed semantically related words than to target words that followed repeated words. This RTs difference between related and repeated pairs can be referred as a negative priming effect, comparable to the effect that has been reported for compatible trials (e.g. Eimer & Schlaghecken, 1998 ; Milliken, Joordens, Merikle, & Seiffert, 1998), notably when the awareness of the prime was reduced (Klapp & Hinkley, 2002). In our study, this negative priming effect is observed for repeated words, when the awareness of the prime is reduced. Note that the same experimental material showed the classical repetition effect (with strongest facilitation for the repeated target) under awareness (in our pre-test).

An explanation of this negative repetition priming is that inhibitory mechanisms could affect the RTs. Eimer and Schlaghecken (1998) proposed inhibitory mechanisms to account for a negative compatibility effect (NCE) that had been observed for pairs of visual arrows that were repeated. The reason of the emergence of this inhibitory mechanism remains under debate (Bowman, Schlaghecken, & Eimer, 2006 ; Lleras & Enns, 2004 ; Mattler, 2007 ; Schlaghecken & Eimer, 2006 ; Verleger, Jaśkowski, Aydemir, van der Lubbe, & Groen, 2004). A potential interpretation of inhibitory mechanisms observed for NCE resides in a two-system interpretation (Eimer & Schlaghecken, 2002 ; Klapp & Hinkley, 2002): the unconscious and the conscious processing are separated in two different neural systems. This interpretation is in line with findings of neuroimaging studies (Milner & Goodale, 1995), it has been proposed that the unconscious and the conscious processing are separated in two different neural systems.

An alternative or additional explanation for the delayed responses to repeated compared to related pairs would be a backward mechanism (Kahan, 2000). Kahan's theory of retrospective prime clarification suggests that the target is retrospectively compared to the prime. As a function of what is relevant for the participant, i.e. depending on the task that needs to be performed, the comparison would be faster when the target is expected after the prime compared to pairs of words where the target is unexpected after the prime. When participants are engaged in a LDT, the semantic code becomes relevant. When participants are presented a prime that is not consciously perceived, the semantic related words are activated, but the prime (also called the center in the spreading activation model, Collins & Loftus, 1975) is inhibited. Thus, participants' responses become delayed more strongly for repeated as compared to related pairs. When the prime is consciously perceived, the repetition priming effect was so strong that the inhibitory mechanisms are compensated. Consequently, the fact that negative priming was observed for repeated as compared to related words pairs would be the sign that the prime was unconsciously processed at a lexical level.

In contrast to the study of Kouider and Dupoux (2005), which showed positive repetition priming in the auditory modality, we found negative repetition priming. Possibly, the above-mentioned inhibitory and backward comparison mechanisms that might lead to inverse repetition priming were attenuated in the study of Kouider and

Dupoux (2005) because of the prime degradation and/or the surrounding mask pattern. It could be hypothesized that the repetition priming reported by Kouider and Dupoux (2005) resulted from low level acoustic cues that matched between the repeated stimuli (despite the time-compression) rather than higher level cognitive processing, i.e., at a semantic or conceptual levels. With the present method of prime awareness reduction, i.e., using a low sound intensity level, we did not directly perform a degradation of the prime and did not use additional masking sounds. In the low intensity presentation, the “internal noise”, i.e. the random variability in neural responses to sensory stimuli (Green & Swets, 1974), served to at least partly mask the prime.

### **c. Rapidity effect**

Subliminal priming effects were not observed for all participants but, as expected, only in the group of the slow responders. This could be explained by the fact that slow responders had more time than the fast responders to achieve matching processes between the prime and the target (Humphreys, 1983), when they were not aware of the prime. Indeed differences between slow and fast responders were not observed when participants were aware of the prime with both groups showing semantic priming effects. In the auditory modality, (Anderson & Holcomb, 1995) reported longer RTs with shorter ISI. Therefore, the absence of RTs differences between slow and fast responders when the prime was consciously categorized seems first surprising. An explanation of this result could be that even if the ISI was short in this study (50 ms), the stimulus onset asynchrony (SOA) was long on average (600 ms). If in the visual modality, the RTs were related to the ISI, in the auditory modality, the SOA might have more influence on the RTs. This explanation is also in concordance with the fact that differences between fast and slow responders were observed when the prime was unconsciously perceived: NCE being observable only when primes and targets are separated by longer SOA (about 150 ms in visual modality, Klapp & Hinkley, 2002).

#### **d. Measures of Awareness**

As done by Kouider and Dupoux (2005) in the auditory domain and others in the visual domain (e.g. Abrams & Greenwald, 2000 ; Dell'Acqua & Grainger, 1999 ; Draine & Greenwald, 1998 ; Greenwald et al., 1996 ; Haase & Fisk, 2001 ; Kiefer & Brendel, 2006 ; Kouider & Dupoux, 2004), we controlled the awareness of the prime with a categorization task. In the present study, and probably in previous studies using categorization as measure of awareness, priming effects were observed while primes remained detectable (visible or audible). This finding integrates into the debate about what is the best measure of awareness. According to Reingold and Merikle (1988) and Merikle and Reingold (1990), a valid measure should meet the criteria of being exhaustive and exclusive. While it seems difficult to meet the exclusiveness criterion (see discussion in Reingold, 2004), it is important to use an exhaustive measure of conscious perceptual information (Merikle & Reingold, 1990 ; Reingold & Merikle, 1988). However, discrimination tasks cannot rule out the possibility that partial or weak stimulus awareness might explain indirect effects (a stimulus, which is not discriminated, might be consciously detected). In contrast, detection tasks asking for presence versus absence have great sensitivity. As suggested by Snodgrass, Bernat and Shevrin (2004a), this task appears to be the best candidate (as compared to identification or semantic classification) for meeting the exhaustiveness criterion and thus for measuring awareness. Thus, it would be interesting in future studies to investigate whether speech priming effects could be also observed when the prime is inaudible, i.e. not detected.

#### **Acknowledgments**

The work was supported by the French National Agency for Research (ANR 2007).

#### **References**

- Abrams, & Greenwald, A. G. (2000). Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychological Science*, *11*(2), 118-124.
- American National Standard Institute. (1995). *ANSI S3.7-1995 (R2003), Methods for coupler calibration of earphones*. New-York.

- American National Standard Institute. (2004). *ANSI S3.6-2004, Specifications for audiometers*. New-York.
- Anderson, J. E., & Holcomb, P. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: an event-related brain potential study. *Psychophysiology*, *32*(2), 177-190.
- Beattie, R. C. (1986). Normal Intelligibility Functions for the Auditec Cid W-22 Test At 30% and 60% Time-Compression. *Otology & Neurotology*, *7*(1), 60-64.
- Bowman, H., Schlaghecken, F., & Eimer, M. (2006). A neural network model of inhibitory processes in subliminal priming. *Visual Cognition*, *13*(4), 401-480. doi:10.1080/13506280444000823
- van den Bussche, E., van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychological Bulletin*, *135*(3), 452-477. doi:10.1037/a0015329
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428. doi:10.1037/0033-295X.82.6.407
- Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F., et al. (1998). Imaging unconscious semantic priming. *Nature*, *395*(6702), 597-600. doi:10.1038/26967
- Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition*, *73*(1), B1-B15.
- Draine, & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology. General*, *127*(3), 286-303.
- Dupoux, E. (1993). The time course of prelexical processing: The syllabic hypothesis revisited. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing : the Second Sperlonga Meeting* (pp. 81-114). Hillsdale, NJ: Psychology Press, L. Erlbaum.
- Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1737-1747.
- Eimer, M., & Schlaghecken, F. (2002). Links between conscious awareness and response inhibition: evidence from masked priming. *Psychonomic Bulletin & Review*, *9*(3), 514-520.
- Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *33*(4), 325-350. doi:10.1080/14640748108400797
- Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *45*(3), 353-372. doi:10.1080/02724989208250619



- Fisk, G. D., & Haase, S. J. (2005). Unconscious perception or not? An evaluation of detection and discrimination as indicators of awareness. *The American Journal of Psychology*, *118*(2), 183-212.
- Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680-698. doi:10.1037/0278-7393.10.4.680
- Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington N.Y.: R.E. Krieger Pub. Co.
- Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, *273*(5282), 1699-1702.
- Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible (“subliminal”) stimuli: dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, *124*(1), 22-42.
- Haase, S. J., & Fisk, G. (2001). Confidence in word detection predicts word identification: implications for an unconscious perception paradigm. *The American Journal of Psychology*, *114*(3), 439-468.
- Holcomb, P., & Neville, H. (1990). Auditory and Visual Semantic Priming in Lexical Decision: A Comparison Using Event-related Brain Potentials. *Language and Cognitive Processes*, *5*(4), 281-312. doi:10.1080/01690969008407065
- Howell, D. (1998). *Méthodes statistiques en sciences humaines*. Paris: De Boeck Université.
- Humphreys, G. W. (1983). Reference frames and shape perception. *Cognitive Psychology*, *15*(2), 151-196.
- Humphreys, G. W., & Quinlan, P. T. (1988). Priming effects between two-dimensional shapes. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 203-220.
- Kahan, T. A. (2000). Negative priming from masked words: retrospective prime clarification of center-surround inhibition? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1392-1410.
- Kiefer, M., & Brendel, D. (2006). Attentional modulation of unconscious “automatic” processes: evidence from event-related potentials in a masked priming paradigm. *Journal of Cognitive Neuroscience*, *18*(2), 184-198. doi:10.1162/089892906775783688
- Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: unconscious inhibition influences reaction time and response selection. *Journal of Experimental Psychology: General*, *131*(2), 255-269.
- Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London: Series B-Biological Sciences*, *362*(1481), 857-875. doi:10.1098/rstb.2007.2093

- Kouider, S., & Dupoux, E. (2004). Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological Science, 15*(2), 75-81.
- Kouider, S., & Dupoux, E. (2005). Subliminal speech priming. *Psychological Science, 16*(8), 617-625. doi:10.1111/j.1467-9280.2005.01584.x
- Leinonen, A., Grönholm-Nyman, P., Järvenpää, M., Söderholm, C., Lappi, O., Laine, M., & Krause, C. M. (2009). Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: evidence from a morphologically rich language. *Brain Research, 1275*, 54-66. doi:10.1016/j.brainres.2009.03.057
- Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General, 133*(4), 475-493. doi:10.1037/0096-3445.133.4.475
- Luce, R. D., & Green, D. M. (1978). Two tests of a neural attention hypothesis for auditory psychophysics. *Perception & Psychophysics, 23*(5), 363-371.
- Marcel, A. J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology, 15*(2), 238-300. doi:10.1016/0010-0285(83)90010-5
- Mattler, U. (2007). Inverse target- and cue-priming effects of masked stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 83-102. doi:10.1037/0096-1523.33.1.83
- Merikle, P. M., & Reingold, E. M. (1990). Recognition and lexical decision without detection: Unconscious perception? *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 574-583. doi:10.1037/0096-1523.16.3.574
- Milliken, B., Joordens, S., Merikle, P. M., & Seiffert, A. E. (1998). Selective attention: a reevaluation of the implications of negative priming. *Psychological Review, 105*(2), 203-229.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. New-York, N.Y.: Oxford University Press.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics, 28*(4), 661-677. doi:10.1017/S014271640707035X
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behavior Research Methods, Instruments & Computers, 36*(3), 516-524.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE™ / A lexical database for contemporary french : LEXIQUE™. *L'Année Psychologique, 101*(3-4), 447-462.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.
- Perrin, F. (2004). Auditory event-related potentials studies of information processing during human sleep. *Psychologica Belgica, 44*(1/2), 43-57.

- Poulin-Charronnat, B., & Tillmann, B. (n.d.). The effect of word repetition is overruled by sentence integration.
- Reingold, E. M. (2004). Unconscious perception: assumptions and interpretive difficulties. *Consciousness and Cognition, 13*(1), 117-122. doi:10.1016/j.concog.2003.11.002
- Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics, 44*(6), 563-575.
- Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. I. Review. *Psychologische Forschung, 36*(1), 71-93.
- Schlaghecken, F., & Eimer, M. (2006). Active masks and active inhibition: a comment on Lleras and Enns (2004) and on Verleger, Jaskowski, Aydemir, van der Lubbe, and Groen (2004). *Journal of Experimental Psychology: General, 135*(3), 484-494. doi:10.1037/0096-3445.135.3.484
- Snodgrass, M., Bernat, E., & Shevrin, H. (2004). Unconscious perception at the objective detection threshold exists. *Perception & Psychophysics, 66*(5), 888-895.
- Urban, M. J. (1992). Auditory subliminal stimulation: a re-examination. *Perceptual and Motor Skills, 74*(2), 515-541.
- Vanhaudenhuyse, A., Laureys, S., & Perrin, F. (2008). Cognitive event-related potentials in comatose and post-comatose states. *Neurocritical Care, 8*(2), 262-270. doi:10.1007/s12028-007-9016-0
- Verleger, R., Jaśkowski, P., Aydemir, A., van der Lubbe, R. H. J., & Groen, M. (2004). Qualitative differences between conscious and nonconscious processing? On inverse priming induced by masked arrows. *Journal of Experimental Psychology: General, 133*(4), 494-515. doi:10.1037/0096-3445.133.4.494

### 5.3 Étude V : « Electrical brain dissociation for consciously and unconsciously categorized auditory stimuli »

Dans l'Étude IV, un traitement sémantique non conscient a été observé pour la première fois en modalité auditive. En réduisant l'intensité de la présentation sonore de stimulations de telle sorte que les participants ne pouvaient pas discriminer la stimulation, des effets d'amorçage négatif de répétition et d'amorçage sémantique ont été obtenus.

Dans l'Étude V, nous proposons d'étudier les bases neuronales sous-jacentes à la perception non consciente des stimulations auditives. Sur la base des résultats obtenus pour l'Étude III et l'Étude IV, nous avons présentés des stimulations auditives (mots, pseudo-mots et sons complexes) à deux niveaux d'intensité différente : la première entraînant une catégorisation consciente, l'autre n'entraînant pas de catégorisation consciente. La technique de l'EEG a été utilisée dans le but d'étudier les modifications de l'activité électrique cérébrale consécutive à la perception consciente ou non consciente des stimulations auditives. Les signaux EEG étaient enregistrés pendant que les participants entendaient et catégorisaient des stimulations auditives (mots/pseudo-mots/sons/silences) ; leurs performances de catégorisation étant au niveau du hasard lorsque les stimulations étaient présentées au niveau d'intensité le plus faible. Compte tenu des hypothèses existantes dans la littérature sur rôle des oscillations et des synchronies dans la représentation consciente des stimulations (voir page 90), nous avons porté une attention particulière aux modifications fréquentielles induites par les stimulations.

Cet article a été présenté entièrement ou en partie lors du *11th Annual Meeting of the Association of Scientific Studies on Consciousness*, à Las Vegas (États-Unis) en juin 2007, lors du Colloque des Jeunes Chercheurs en Sciences Cognitives (CJC-SC), à Lyon en mai 2007 et lors des *Journées Fondatrices Perception Sonore(JPS)* à Lyon en janvier 2007.

Il est actuellement soumis à une revue à comité international de lecture (*The Journal of Neuroscience*) et sera prochainement présenté au *14<sup>th</sup> Annual Meeting of*

*the Association of Scientific Studies on Consciousness*, à Toronto (Canada) en juin 2010.

# **ELECTRICAL BRAIN DISSOCIATION FOR CONSCIOUSLY AND UNCONSCIOUSLY CATEGORIZED AUDITORY STIMULI**

Carine Signoret<sup>1</sup>, Etienne Gaudrain<sup>1,2</sup> and Fabien Perrin<sup>1</sup>

1. UCBL - CNRS UMR 5020 Neurosciences Sensorielles, Comportement, Cognition, Lyon,

France

2. MRC Cognition and Brain Sciences Unit, Cambridge, United-Kingdom

Keywords: Speech, Consciousness, EEG, Beta, event-related, Auditory

## **Acknowledgements**

We thank Olivier Bertrand for helpful comments in the preparation of this study.

## **Abstract**

Although a number of studies have investigated the neural correlates of visual conscious perception, little is known regarding auditory conscious perception. In the present study, we investigated electroencephalographic (evoked and induced) responses to behaviorally categorized or uncategorized words, pseudo-words and complex sounds as well as responses to silences. Our results suggest that, like silences, uncategorized stimuli were not perceived because they neither evoked potentials nor theta activity. Interestingly, uncategorized stimuli also displayed neural mechanisms similar to consciously categorized stimuli (notably a decrease in the beta band), suggesting that uncategorized stimuli were partially processed. Hemispheric differences in the beta bands generated by words and non-lexical sounds (pseudo-words and complex sounds), differences that occurred with both categorized and

uncategorized stimuli, confirm that stimuli were behaviorally uncategorized but unconsciously processed — probably at the lexical level.

### 5.3.1 Introduction

The neural correlates of conscious perception are still enigmatic, but an increasing number of studies (primarily investigating the visual modality) suggest that conscious perception depends on the activation of specific structures and/or their interactions. From an anatomical perspective, conscious perception is either associated with enhanced activity in early sensory areas (Grill-Spector, Kushnir, Hendler, & Malach, 2000 ; Pins & Ffytche, 2003) or with a widely distributed frontoparietal network that may include the primary sensory cortex (Marois, Yi, & Chun, 2004 ; Naccache & Dehaene, 2001). From a dynamic perspective, conscious perception results from the neural synchronization of structures involved in perception, ranging from local coordination within structures to long-distance interactions between structures (Crick & Koch, 2003 ; Engel & Singer, 2001 ; Tononi, Srinivasan, Russell, & Edelman, 1998). Some studies have argued for a strong correlation between the increase local or global of gamma synchronizations and the conscious representation of the stimuli (Meador et al., 2002 ; Melloni et al., 2007 ; Rodriguez et al., 1999 ; Wyart & Tallon-Baudry, 2008). Other studies have also shown that reduced or increased alpha and/or beta activities are correlated with the conscious perception of a stimulus (Gaillard et al., 2009 ; Dan Glauer & Scherer, 2008 ; Luo et al., 2005 ; Palva et al., 2005). These results suggest that the conscious representation of our environment is probably the result of multiple interactions between structures that could operate at different frequency bands (Varela, Lachaux, Rodriguez, & Martinerie, 2001).

While many studies have investigated the neural correlates of visual conscious perception, little is known regarding auditory conscious perception. This lack of interest is likely caused by the difficulty in demonstrating unconscious auditory perception. Indeed, studies employing dichotic-listening procedures or masking techniques have failed to produce satisfactory experimental situations and concordant results (Holender, 1986b ; Lindner, 1968 ; Shipley, 1965 ; Urban, 1992). To our knowledge, only two studies have observed behavioral signs of unconscious speech processing. Using attenuated, time-compressed primes surrounded by masks, Kouider

and Dupoux (2005) found an auditory repetition priming effect when participants were unable to categorize a stimulus. In a similar study using attenuated (but not deformed) speech and non-speech stimuli, Signoret et al. (unpublished observations) observed recognition performance above chance for words when no detection was reported.

In the present study, we investigate the electroencephalographic correlates of auditory conscious representation. To this purpose, silences, words, pseudo-words and complex sounds were used. The words, pseudo-words and complex sounds could be presented at two levels, one in which the participants could behaviorally categorize them and one in which they could not. This experimental design allows us to identify not only the cerebral responses that are associated with the participant's subjective perception (i.e., her/his conscious percept) but also those that are associated with the objective presence of sound (i.e., the processing of a stimulus independently of its conscious representation). We also use this experimental design to investigate the neural differences between stimuli that differ in their phonological or lexical contents — or both — when the stimuli are consciously categorized or uncategorized.

### **5.3.2 Materials and Methods**

#### **a. Participants**

Fourteen students (mean age  $23 \pm 3.9$  years, 5 women), right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) were included in the study. All were native French speakers and did not report any hearing problems or any history of neurological disease. Participants were selected on the basis of their low auditory detection thresholds (ANSI S3.21, 2004). The study was approved by the local Ethics Committee (CPP Sud Est II, n° 2007-010-2).

#### **b. Stimuli**

Three types of stimuli were used: words, pseudo-words and complex sounds. The words were selected from a French database (Lexique 2; New et al., 2004, 2007) and were common, singular, monosyllabic nouns of 2-5 phonemes. All words had a frequency of occurrence higher than 1 per million. A list of pseudo-words



(pronounceable letter strings judged to be meaningless by 5 other participants) was generated from words by mixing their phonemes. Words and pseudo-words were spoken by the same female speaker and recorded separately (32 bits, 44100 Hz). The non-phonological, complex sounds were created from the words and pseudo-words using the Fonds sonores algorithm (Perrin & Grimault, 2005); the phase spectrum of words and pseudo-words were randomized, but their average spectral content and temporal envelope (up to 60 Hz) were preserved (for more details see <http://olfac.univ-lyon1.fr/unite/equipe-02/FondsSonores.html>). The average durations of the words ( $527.22 \pm 103.66$  ms), pseudo-words ( $542.60 \pm 82.09$  ms) and complex sounds ( $531.33 \pm 93.53$  ms) were not significantly different. All stimuli were equalized to the same dB-A level. The A-weighting was chosen because it roughly mimics the external and middle ear transfer functions.

In a pretest, 8 other right-handed participants ( $23.72 \pm 1.17$  years old, 5 women) heard the same stimuli randomly presented at 5 levels in both the inaudible and audible range (0 to 20 dB-A). The participants had to successively perform a detection task (presence/absence), a speech-categorization task (speech/non-speech) and a lexical-categorization task (words/pseudo-words). Significantly better performances were observed in the detection task relative to the speech-categorization task ( $p = .019$ ) and in the speech-categorization task relative to the lexical-categorization task ( $p = .000$ ). At 5 dB-A, the participants could detect the stimuli 51.50% of the time ( $d' = 1.94$ ), but they could not consciously categorize them ( $d' = .12$ , not significantly different from zero,  $t_{(7)} = .74$ ,  $p = .48$ ). At 20 dB-A, the participants could detect the stimuli 86.88% of the time ( $d' = 3.29$ ), and they could categorize them 65.63% of the time ( $d' = .90$ , significantly different from zero,  $t_{(7)} = 3.57$ ,  $p = .009$ ).

Based on the results of the pretest, stimuli were presented at 5 dB-A (labeled *uncategorized* stimuli) and 40 dB-A (labeled *categorized* stimuli).

### **c. Experimental Procedure**

Each participant was seated in a sound-attenuated, electromagnetically shielded booth while the EEG was recorded. A white fixation cross was presented in the middle of the screen 100 to 500 ms (by 100 ms steps) before the presentation of silence or a stimulus (word, pseudo-word or complex sound) and until its offset. The

participants were instructed to keep their eyes open and focus on the white cross. Participants had to categorize the stimulations as words, pseudo-words, complex sounds or silences. They were not informed that some stimuli were presented at low levels. In 10% of the trials, a question mark replaced the fixation cross after the stimulus offset and, in these cases, the participants had to provide a motor response on the keyboard as quickly and as accurately as possible (all conditions were equally and randomly represented). Before a new trial started, participants were shown a black screen for 1200 ms to allow them to blink. Each participant was presented with 180 stimuli from each category and condition as well as 180 silences, forming a total of 1260 stimuli that were randomly presented in 4 blocks. Half of the non-silence stimuli were presented in the *uncategorized* condition (5 dB-A) whereas the other half were presented in the *categorized* condition (40 dB-A).

#### **d. EEG recordings**

Electroencephalographic (EEG) signals from 60 Ag-AgCl electrodes (positioned according to the International 10-20 system) referenced to the nose, horizontal electrooculograms (EOG) from bipolar electrodes positioned at the outer canthi of both eyes and vertical EOGs from bipolar electrodes positioned below and above the left eye were amplified using the Brain Quick SD64 Micromed system and sampled at 512 Hz (16 bits) using an analog bandpass filter of .05–128 Hz. A ground electrode was placed between the FPz sites and the impedance at all electrodes was kept below 5 k $\Omega$ .

#### **e. Data analysis**

All analyses were performed using custom scripts written in Matlab (The MathWorks) and the Eeglab toolbox (Delorme & Makeig, 2004). Signals containing non-stereotypical artifacts, including high-amplitude, high-frequency muscle noise and electrode cable movements, were rejected. Stereotypical artifacts such as eye movements and eye blinks were corrected with an extended Infomax independent component analysis (Lee, Girolami, & Sejnowski, 1999) implemented in Eeglab. In this analysis, the data were broken into 60 component activations per subject, component activations representing non-brain artifacts were removed (based on a

visual inspection of their scalp topographies, time courses, and frequency spectra) and EEG data were reconstructed from the remaining component activations (Delorme & Makeig, 2004). The data from one subject were excluded from the analysis because of large artifacts in the signal.

### ***Event-Related Potentials (ERPs)***

The artifact-free signal was segmented from - 200 to + 1000 ms around the stimulus onsets and the baseline, which was corrected from 200 ms before the stimulus onset. For each electrode, grand average waveforms were computed based on the nature of the stimulation (silence, word, pseudo-word, complex sound) and/or the stimulation processing (categorized, uncategorized). Four potentials were evoked by the categorized stimulations (see Figure 1a): N1 (50–200 ms), P2 (200–450 ms), N2 (450–700 ms) and P3 (700–1000 ms). For individual ERPs, the mean amplitudes and maximal peak latencies in each time window were analyzed using two-way analyses of variance (ANOVAs) with repeated measures on Condition (categorized stimuli, uncategorized stimuli, silences) and Electrode Position (Fz, Cz, Pz). To test the laterality effects of speech and/or the lexical features of the stimuli, the mean amplitudes and maximal peak latencies of individual ERPs for categorized stimuli were also analyzed using three-way ANOVAs with repeated measures on Stimulation (word, pseudo-word, complex sound), Lateral Electrode Position (left, right) and Antero-posterior Electrode Position (anterior, posterior). For this purpose, four quadrants (of 11 electrodes) were defined: a right anterior site, a left anterior site, a right posterior site and a left posterior site. If  $p < .05$  after applying the Greenhouse-Geisser correction, the interaction was considered significant; *post-hoc* tests (Fisher's LSD) were applied to significant interactions.

### ***Event-Related Spectral Perturbations (ERSPs)***

The artifact-free signal was segmented from - 1000 to + 2000 ms around the stimulus onsets and the time-frequency was decomposed (trial-by-trial) using Hann-windowed sinusoidal wavelets of 5 cycles at 4 Hz rising linearly to 12 cycles at 48 Hz (i.e., a window length of 1395 ms). For each electrode, the signals were then averaged based on the nature of the stimulation (word, pseudo-word, complex sound) and/or the stimulation processing (categorized, uncategorized). ERSPs were observed in 5 visible spectro-temporal windows relative to the baseline (see Figure 2): theta band

( $\theta$ : 4–7 Hz) between 0 and 500 ms, alpha band ( $\alpha$ : 8–12 Hz) between 0 and 500 ms, beta band ( $\beta$ : 15–30 Hz) between 150 and 400 ms, low-gamma band between 0 and 100 ms (early gamma,  $\gamma_a$  between 30–42 Hz) and low-gamma band between 150 and 450 ms (late gamma,  $\gamma_b$  between 30–48 Hz). The average power of each spectro-temporal window was analyzed using two-way ANOVAs with repeated measures on Condition (categorized stimuli, uncategorized stimuli, silence) and Electrode Position (Fz, Cz, Pz). Previous studies have observed lateralized alpha (8–12 Hz), low-beta (15–20 Hz) and/or high-beta (20–30 Hz) power differences over the frontal and/or temporal regions for linguistic and non-linguistic stimuli (Krause et al., 2006 ; Luo et al., 2005 ; Lutzenberger, Pulvermüller, & Birbaumer, 1994 ; Palva et al., 2002 ; Pulvermüller, Preissl, Lutzenberger, & Birbaumer, 1996 ; Shahin et al., 2009). Therefore, alpha, low-beta and high-beta windows for categorized and uncategorized stimuli were also analyzed using four-way ANOVAs with repeated measures on Condition (categorized stimuli, uncategorized stimuli), Stimulation (words, pseudo-words, complex sounds), Lateral Electrode Position (left, right) and Fronto-temporal Electrode Position (frontal, temporal). For this purpose, four electrodes were chosen: F4 (right frontal), F3 (left frontal), TP8 (right temporal) and TP7 (left temporal). If  $p < .05$  after applying the Greenhouse-Geisser correction, the interaction was considered significant; *post-hoc* tests (Fisher’s LSD) were applied to significant interactions.

### 5.3.3 Results

#### a. Behavioral data

In accordance with the pretest, the behavioral analyses showed that silences were categorized as “silence” in 89.06% of the cases, that stimuli presented at 40 dB-A were correctly categorized in 85.74% of the cases, and that stimuli presented at 5 dB-A were not correctly categorized in 83.67% of the cases. At 5 dB-A, a correct categorization of 16.33% corresponds to the chance level. When participants detected a stimulus (in 42.88% of the cases they did not categorize the 5 dB-A stimuli as silence), they had a 1 out of 3 chance (i.e., 14.29%) of finding the correct response (word, pseudo-word or complex sound). The participants’ performance was not different from the chance level for words, pseudo-words and complex sounds (*t*-tests

against chance level were  $t_{(13)} = -.98, p = .34, \text{NS}$ ;  $t_{(13)} = .50, p = .63, \text{NS}$ ;  $t_{(13)} = 1.21, p = .25, \text{NS}$ , respectively). Thus, the stimuli presented at 5 dB-A were sometimes detected but never consciously categorized.

## **b. Event-related potentials (ERPs)**

Classical auditory ERPs (*Figure 1a*) were observed after the categorized stimuli but not after the silences or the uncategorized stimuli (main effect: Condition on N1,  $F_{(2,24)} = 6.69, p = .005$ ; on P2,  $F_{(2,24)} = 12.12, p < .001$ ; on N2,  $F_{(2,24)} = 29.27, p < .001$ ; and on P3,  $F_{(2,24)} = 11.57, p < .001$ ). A restricted Cz scalp distribution was observed for N1, P2 and N2 for categorized stimuli (interaction Condition  $\times$  Electrode Position: N1,  $F_{(4,48)} = 6.38, p < .001$ ; P2,  $F_{(4,48)} = 25.19, p < .001$  and N2,  $F_{(4,48)} = 9.81, p < .001$ ; *post-hoc* comparisons of categorized stimuli: Cz – Fz,  $p = .004$ ; Cz – Pz,  $p < .001$ ).

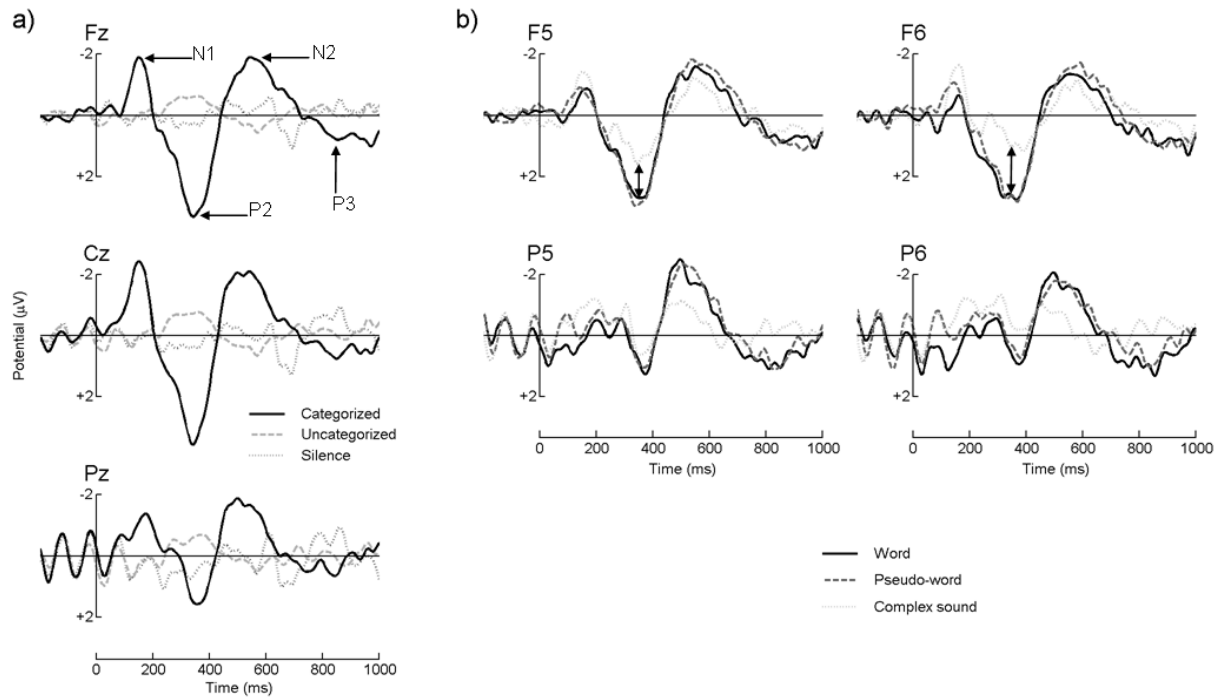


Figure 1. a) Event-Related Potentials (N1, P2, N2 and P3) are presented as a function of Condition (categorized stimuli, uncategorized stimuli and silences) for each Electrode (Fz, Cz, Pz). The black curve represents the categorized stimuli; the grey, dashed curve represents the uncategorized stimuli; and the dotted curve represents the silences. Only the categorized stimuli presented classical auditory ERPs. b) Event-Related Potentials for categorized stimuli are presented as a function of Stimulation (words, pseudo-words and complex sounds) from one electrode in each quadrant (F5, F6, P5, P6). The black curve represents the categorized words; the grey, dashed curve represents the categorized pseudo-words and the dotted curve represents the categorized complex sounds. Higher N2 and P2 amplitudes were observed for complex sounds in the left hemisphere.

For categorized stimuli (Figure 1b), N1 was greater for complex sounds than for pseudo-words and was greater for pseudo-words than for words in posterior sites (interaction Stimulation  $\times$  Antero-posterior Electrode Position:  $F_{(2,24)} = 8.56, p = .002$ ; *post-hoc* comparisons for posterior site: complex sounds – pseudo-words,  $p = .005$ ; pseudo-words – words,  $p < .001$ ; complex sounds – words,  $p < .001$ ). P2, N2 and P3 were higher for speech sounds than non-speech sounds with a maximal amplitude at anterior sites for P2 and N2 and at posterior sites for P3 (main effect: Stimulation, P2,  $F_{(2,24)} = 4.13, p = .029$ ; N2,  $F_{(2,24)} = 10.05, p < .001$ ; interaction: Stimulation  $\times$  Antero-posterior Electrode Position, P2,  $F_{(2,24)} = 7.63, p = .003$ , N2,  $F_{(2,24)} = 5.46, p = .011$ ; P3,  $F_{(2,24)} = 13.60, p < .001$ ; *post-hoc* comparisons: anterior to posterior, words – complex sounds,  $ps < .008$ ; pseudo-words – complex sounds,  $ps < .001$ ; words – pseudo-words,  $p > .312$ , NS for P2,  $p > .051$ , NS for N2 and  $p > .103$ , NS for P3).

More importantly, P2 and N2 were greater in the left than in the right hemisphere for complex sounds, and N2 was greater in the left than in the right hemisphere for pseudo-words (interaction: Stimulation  $\times$  Lateral Electrode Position, P2,  $F_{(2,24)} = 4.58$ ,  $p = .021$ ; N2,  $F_{(2,24)} = 6.23$ ,  $p = .007$ ; *post hoc* comparisons: left – right, complex sounds,  $ps < .001$ ; pseudo-words, P2,  $p = .189$ , NS and N2,  $p = .012$ ; words, P2,  $p = .988$ , NS and N2,  $p = .597$ , NS). No latency effect was observed for categorized stimuli.

### c. Event-related Spectral Perturbations (ERSPs)

The results suggest: (i) an early (0 to 100 ms) gamma power increase for silences relative to categorized and uncategorized stimuli (main effect: Condition,  $F_{(2,24)} = 8.63$ ,  $p = .002$ ; *post hoc* comparisons: categorized – silence,  $p = .003$ ; uncategorized – silence,  $p < .001$ ; categorized – uncategorized,  $p = .569$ , NS); (ii) a theta power increase for categorized relative to uncategorized stimuli and silences (main effect: Condition,  $F_{(2,24)} = 49.56$ ,  $p < .001$ ; *post hoc* comparisons: categorized – silence,  $p < .001$ ; uncategorized – silence,  $p = .718$ , NS; categorized – uncategorized,  $p < .001$ ) with maximal differences at Cz (interaction: Condition  $\times$  Electrode Position,  $F_{(4,48)} = 5.51$ ,  $p < .001$ ); (iii) a greater alpha power decrease for uncategorized stimuli and silences than for categorized stimuli at Fz and Cz sites (interaction: Condition  $\times$  Electrode Position,  $F_{(4,48)} = 5.60$ ,  $p < .001$ ; *post hoc* comparisons: Fz and Cz, categorized – silence,  $ps < .001$ ; uncategorized – silence,  $p = .507$ , NS and  $p = .709$ , NS; categorized – uncategorized,  $ps = .001$ ) and (iv) a greater beta power decrease for categorized than for uncategorized stimuli and a greater decrease at Pz only for uncategorized stimuli than for silences (interaction: Condition  $\times$  Electrode Position,  $F_{(4,48)} = 3.17$ ,  $p = .022$ ; *post hoc* comparisons: categorized – uncategorized,  $ps < .017$ ; Fz and Pz, categorized – silences,  $ps < .030$ ; Pz, uncategorized – silences,  $p = .004$ ). Neither the main effect of Condition nor any interaction was significant in late gamma (*Figure 2*).

ERSPs contain both phase-locked (evoked) and non-phase locked (induced) oscillatory brain responses. To distinguish phase-locked activity from ERSPs, the phase coherence between trials was analyzed (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). In our study, a trial-by-trial phase-coherence analysis from 4 to 48 Hz showed phase-locked activity only for the theta band after the categorized stimuli.

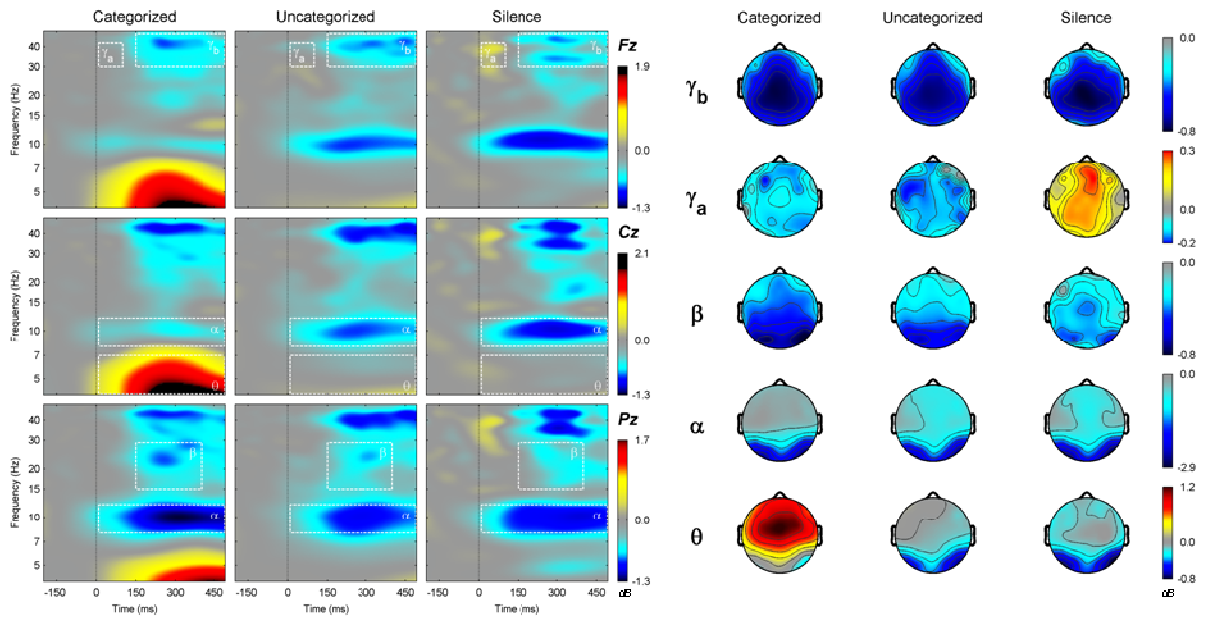
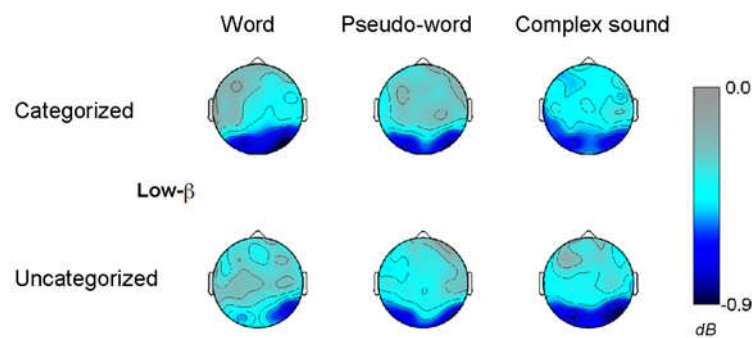


Figure 2. Left panel: Event-Related Spectral Perturbations as a function of Condition (categorized stimuli, uncategorized stimuli and silences) at Fz, Cz and Pz sites. Right panel: Activity maps (average of rectangles in the left panel) for each frequency band (theta, alpha, beta, early gamma and late gamma).

For both categorized and uncategorized stimuli (Figure 3), there was a greater power decrease in low-beta for pseudo-words and complex sounds relative to words; this decrease was greater in the left hemisphere (interaction: Stimulation  $\times$  Lateral Electrode Position,  $F_{(2,24)} = 3.49$ ,  $p = .047$ ; *post-hoc* comparisons: left, words – pseudo-words,  $p = .016$ ; words – complex sounds,  $p = .002$ ; pseudo-words – complex sounds,  $p = .356$ , NS). There was also a greater power decrease in beta activity in the right than in the left hemisphere for words only (*post-hoc* comparisons: words, left – right,  $p = .007$ ). The high-beta activity revealed a greater decrease for categorized than for uncategorized stimuli only in the right hemisphere (interaction: Condition  $\times$  Lateral Electrode Position,  $F_{(1,12)} = 5.00$ ,  $p = .045$ ; *post-hoc* comparisons: right, categorized – uncategorized,  $p = .002$ ). The alpha band decrease was greater for uncategorized than for categorized stimuli in frontal sites (interaction: Condition  $\times$  Fronto-temporal Electrode Position,  $F_{(1,12)} = 7.16$ ,  $p = .020$ ; *post-hoc* comparisons: categorized and uncategorized, frontal – temporal,  $ps < .001$ ). Finally, the alpha power decrease was greater at the right temporal site (main effect: Fronto-temporal Electrode Position,  $F_{(1,12)} = 13.42$ ,  $p = .003$ ; Lateral Electrode Position,  $F_{(1,12)} = 6.30$ ,  $p = .027$ ; interaction: Fronto-temporal Electrode Position  $\times$  Lateral Electrode Position,



$F_{(1,12)} = 7.67$ ,  $p = .017$ ; *post hoc* comparisons: frontal – temporal,  $ps = .0013$ ; temporal, left – right,  $p < .001$ ).



*Figure 3.* Maps of Event-Related Spectral Perturbations for categorized and uncategorized stimuli as a function of Stimulation (words, pseudo-words and complex sounds) of the beta frequency band.

### 5.3.4 Discussion

This study shows different electroencephalographic responses to behaviorally categorized stimuli, uncategorized stimuli, and silences. The results suggest that uncategorized stimuli, like silences, were not processed (they neither evoked potentials nor theta activity). However, the results also show that uncategorized stimuli share neural mechanisms with consciously categorized stimuli (notably a decrease in the beta band), thus suggesting that they were partially processed. In the beta band, hemispheric differences between words and non-lexical sounds (pseudo-words and complex sounds) occurred with both categorized and uncategorized stimuli, confirm that the stimuli were behaviorally uncategorized but were unconsciously processed — probably at the lexical level.

Like silences, behaviorally uncategorized stimuli generated limited cerebral variations. Classical auditory N1-P2-N2-P3 components were found for categorized stimuli but not for uncategorized stimuli or silences. This result is not surprising because it has been shown that ERP amplitude decreases when intensity is reduced (Hillyard, Squires, Bauer, & Lindsay, 1971 ; Polich, Ehlers, Otis, Mandell, & Bloom, 1986) to the psychophysical threshold or near it (Musiek, Froke, & Weihing, 2005 ; Suzuki, Yamamoto, Taguchi, & Sakabe, 1976). These findings suggest that, in our study, the behaviorally uncategorized stimuli (5 dB-A) were presented below the

psychophysical auditory threshold. Our study also showed that the absence of an evoked response was associated with variations in low-frequency bands like theta and alpha. This result is reminiscent of studies demonstrating that phase-locking is involved in the generation of ERPs. For example, Klimesch et al. (2007) have shown that ERPs reflect the superposition of an increase in the evoked theta band over a decrease in the alpha band (within one second post-stimulus). This finding is concordant with the theta band increase observed for consciously categorized stimuli and in agreement with the inter-trial phase-coherence analysis of the present study, showing that this activity was phase-locked to the stimulus (as ERPs). Thus, as previously shown, the theta activity is likely associated with the sensory coding of the stimulus (see review in Buzsáki, 2006). In contrast to the theta band, alpha band variations seem less linked to phase-locked responses; a posterior alpha decrease was recorded after all stimuli and silences, and a greater fronto-central alpha decrease was observed for uncategorized stimuli and silences. These effects are in accordance with previous studies showing an association between alpha desynchronization and attentional or ‘alertness’ processes (see review in Klimesch et al., 2006). In our study, the posterior alpha decrease likely indicates that participants had a constant level of alertness during the task (Klimesch, Doppelmayr, Russegger, & Pachinger, 1996), and the frontal decrease is probably linked to the attended stimulus (Yamagishi et al., 2003). Supporting this hypothesis, the combined frequency of ‘inaudible’ stimuli at 5 dB-A and silences was higher than the frequency of audible stimuli at 40 dB-A.

Although the response evoked by behaviorally uncategorized stimuli was similar to the one evoked by silences, they also shared properties with the responses observed for consciously categorized stimuli. This similarity suggests that the behaviorally uncategorized stimuli were partially processed. An early, transient increase in the gamma band was observed only for silences but was followed by a decrease in the gamma band for all stimuli and silences. It has been previously shown that an early gamma band increase followed by a decrease in induced gamma activity reflects processes related to attention (Bertrand & Tallon-Baudry, 2000 ; Debener, Herrmann, Kranczioch, Gembris, & Engel, 2003 ; Fell, Hinrichs, & Röschke, 1997 ; Herrmann & Knight, 2001 ; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1997 ; Tiitinen et al., 1993). In our study, the gamma band decrease reflects that participants had a constant level of alertness during the task for all conditions (a proposal

corroborated by the posterior alpha desynchronization). However, it is more difficult to explain the early gamma synchronization in response to silences. The participants were trying to categorize a stimulus that could begin at any time between 100 and 500 ms after the appearance of the cross. The longer the participants waited without detecting anything, the more likely it was that a silence was being presented; the participants may have inferred this probability. The small early gamma activity may indicate that the participants were performing the task (i.e., categorizing the trial as silence) using this inferred probability. The absence of early gamma activity for categorized stimuli could be explained by the fact that, in order to correctly perform the task, the participants required more information about the stimulation and were thus waiting for the end of the stimulus presentation. Interestingly, the absence of early gamma activity for uncategorized stimuli could indicate that these stimulations were not processed as silences (in contrast to the results observed for ERPs and low-frequency bands) and that participants were waiting — probably at an unconscious level — for more information regarding the stimulus. In sum, this interpretation suggests that the participants inferred the probability of silences based on the conscious, but also on the unconscious, perception of stimuli presentation. Variations in the beta frequency band (probably the high-beta band) also suggest that stimuli (both behaviorally categorized and uncategorized) and silences were processed differently. Indeed, a higher beta decrease was observed for stimuli than for silences in posterior regions; the biggest decrease was observed for consciously categorized stimuli. A study by Palva et al. (2005) observed a similar effect between somatosensory-perceived and unperceived stimuli. However, it was not possible in their study to know whether unperceived stimuli were unconsciously perceived because they did not compare their data to a condition without stimuli. Using silences, we showed that the beta decrease is modulated by both the objective presence of the stimulus and its conscious perception.

Analyses that distinguish between types of stimuli provide important information regarding the implications of beta oscillations not only in conscious/unconscious perception but also in lexical processing. Several studies have reported that oscillations around 20-30 Hz are involved in the processing of speech and/or lexical sounds. For example, Palva et al. (2002) observed an increase in evoked responses to non-speech sounds in the right hemisphere relative to speech sounds.

Most interestingly, Pulvermüller et al. (1996) observed an induced spectral depression in the left hemisphere after pseudo-words but not after words. In our study, the low-beta decrease was significantly greater for non-lexical stimuli (pseudo-words and complex sounds) than for lexical stimuli (words); further, this decrease was observed in the left temporal hemisphere but not in the right one. This result again confirms the lexical specialization of the left hemisphere, a specialization that has recently been demonstrated using neuroimaging techniques (Binder et al., 2000, 1997) and ERPs (Ziegler et al., 1997). In our study, hemispheric differences were also observed in the results from ERP analysis. We observed responses in the left hemisphere that were greater for speech than for non-speech stimuli near the anterior (P2, N2) and posterior (P3) regions. Finally, our data showed hemispheric differences in beta activity between stimuli, differences that occurred for both consciously categorized and uncategorized stimuli. This finding suggests that behaviorally uncategorized stimuli elicited neural responses similar to consciously categorized stimuli i.e., that behaviorally uncategorized stimuli were lexically processed at an unconscious level.

While many studies have investigated the neural basis of auditory unconscious processing under conditions of altered or reduced awareness (see Perrin, 2004, for a review of studies during sleep and Vanhaudenhuyse et al., 2008, for a review of studies during post-comatose states), only one study has examined this processing during wakefulness. Recently, Kouider et al. (2010) investigated the cerebral mechanisms of unconscious processing using pairs of repeated or non-repeated speech words in conjunction with time-compressed and masked primes. While the authors found that subliminal word repetition reduced the activity in the left superior temporal gyrus and the insula, they did not explain why there were different neural responses involved in conscious and unconscious speech processing. The present study suggests that the conscious categorization of sounds is not the result of an extended synchronization of structures within a restricted gamma band (as has often been reported) but instead is the result of concomitant neuronal activations (reflected in both evoked and induced responses). While the unconscious processing of auditory stimuli was associated with a decrease in the beta band (with hemispheric variations depending on the lexical status of the stimulus), conscious auditory representation was associated with an extended decrease in the beta band and the evocation of classical auditory potentials (associated with theta increase). On a physiological level, a

decrease in oscillatory activity — particularly in the alpha and beta frequency ranges — may correlate with brain activation (Pfurtscheller & Aranibar, 1977). In addition, simulation studies have revealed that the beta frequency has characteristics compatible with long-range interactions, whereas the gamma frequency band seems to be optimal for local processing (Bibbig et al., 2002 ; Kopell, Ermentrout, Whittington, & Traub, 2000). If this division holds true, beta desynchronization would be a prime candidate for mediating interactions within the widely distributed network that determines the conscious representation of stimuli (Varela et al., 2001). In the present study, the smaller decrease in beta activity for behaviorally uncategorized stimuli relative to the consciously categorized stimuli suggests less efficient, long-distance interactions between structures. Future studies using intracerebral electrodes would be able to confirm this hypothesis.

## References

- American National Standard Institute (2004) ANSI S3.6-2004, Specifications for audiometers. New-York.
- Bertrand O, Tallon-Baudry C (2000) Oscillatory gamma activity in humans: a possible role for object representation. *Int. J. Psychophysiol* 38:211-223
- Bibbig A, Traub RD, Whittington MA (2002) Long-range synchronization of gamma and beta oscillations and the plasticity of excitatory and inhibitory synapses: a network model. *J. Neurophysiol* 88:1634-1654
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10:512-528
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T (1997) Human brain language areas identified by functional magnetic resonance imaging. *J. Neurosci* 17:353-362
- Buzsáki G (2006) *Rhythms of the brain*. Oxford, UK: Oxford University Press.
- Crick F, Koch C (2003) A framework for consciousness. *Nat. Neurosci* 6:119-126
- Dan Gläuser ES, Scherer KR (2008) Neuronal processes involved in subjective feeling emergence: oscillatory activity during an emotional monitoring task. *Brain Topogr* 20:224-231
- Debener S, Herrmann CS, Kranczioch C, Gembris D, Engel AK (2003) Top-down attentional processing enhances auditory evoked gamma band activity. *Neuroreport* 14:683-686

- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134:9-21
- Engel AK, Singer W (2001) Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci* 5:16-25
- Fell J, Hinrichs H, Röschke J (1997) Time course of human 40 Hz EEG activity accompanying P3 responses in an auditory oddball paradigm. *Neurosci. Lett* 235:121-124
- Gaillard R, Dehaene S, Adam C, Clémenceau S, Hasboun D, Baulac M, Cohen L, Naccache L (2009) Converging intracranial markers of conscious access. *PLoS Biol* 7:e61
- Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. *Nat. Neurosci* 3:837-843
- Herrmann CS, Knight RT (2001) Mechanisms of human attention: event-related potentials and oscillations. *Neurosci. Biobehav. R* 25:465-476
- Hillyard SA, Squires KC, Bauer JW, Lindsay PH (1971) Evoked potential correlates of auditory signal detection. *Science* 172:1357-1360
- Holender D (1986) Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behav. Brain Sci* 9:1
- Klimesch W, Doppelmayr M, Russegger H, Pachinger T (1996) Theta band power in the human scalp EEG and the encoding of new information. *Neuroreport* 7:1235-1240
- Klimesch W, Hanslmayr S, Sauseng P, Gruber W, Brozinsky CJ, Kroll NEA, Yonelinas AP, Doppelmayr M (2006) Oscillatory EEG correlates of episodic trace decay. *Cereb. Cortex* 16:280-290
- Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev* 53:63-88
- Kopell N, Ermentrout GB, Whittington MA, Traub RD (2000) Gamma rhythms and beta rhythms have different synchronization properties. *P. Natl. Acad. Sci. USA* 97:1867-1872
- Kouider S, Dupoux E (2005) Subliminal speech priming. *Psychol. Sci* 16:617-625
- Kouider S, de Gardelle V, Dehaene S, Dupoux E, Pallier C (2010) Cerebral bases of subliminal speech priming. *NeuroImage* 49:922-929
- Krause C.M., Gronholm P., Leinonen A., Laine M., Sakkinen A.-L., Soderholm C. (2006) Modality matters: The effects of stimulus modality on the 4- to 30-Hz brain electric oscillations during a lexical decision task *Brain Res.* 1110(1): 182-192
- Lee TW, Girolami M, Sejnowski TJ (1999) Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput* 11:417-441

- Lindner WA (1968) Recognition performance as a function of detection criterion in a simultaneous detection-recognition task. *J. Acoust. Soc. Am* 44:204-211
- Luo H, Husain FT, Horwitz B, Poeppel D (2005) Discrimination and categorization of speech and non-speech sounds in an MEG delayed-match-to-sample study. *NeuroImage* 28:59-71
- Lutzenberger W, Pulvermüller F, Birbaumer N (1994) Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans. *Neurosci. Lett* 176:115-118
- Marois R, Yi D, Chun MM (2004) The neural fate of consciously perceived and missed events in the attentional blink. *Neuron* 41:465-472
- Meador KJ, Ray PG, Echaz JR, Loring DW, Vachtsevanos GJ (2002) Gamma coherence and conscious perception. *Neurology* 59:847-854
- Melloni L, Molina C, Pena M, Torres D, Singer W, Rodriguez E (2007) Synchronization of neural activity across cortical areas correlates with conscious perception. *J. Neurosci* 27:2858-2865
- Musiek FE, Froke R, Weihing J (2005) The auditory P300 at or near threshold. *J. Am. Acad. Audiol* 16:698-707
- Naccache L, Dehaene S (2001) The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cereb. Cortex* 11:966-974
- New B, Brysbaert M, Veronis J, Pallier C (2007) The use of film subtitles to estimate word frequencies. *Appl. Psycholing* 28:661-677
- New B, Pallier C, Brysbaert M, Ferrand L (2004) Lexique 2: a new French lexical database. *Behav. Res. Meth. Ins. C.* 36:516-524
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97-113
- Palva S, Linkenkaer-Hansen K, Näätänen R, Palva JM (2005) Early neural correlates of conscious somatosensory perception. *J. Neurosci* 25:5248-5258
- Palva S, Palva JM, Shtyrov Y, Kujala T, Ilmoniemi RJ, Kaila K, Näätänen R (2002) Distinct gamma-band evoked responses to speech and non-speech sounds in humans. *J. Neurosci* 22:RC211(1-5)
- Perrin F (2004) Auditory event-related potentials studies of information processing during human sleep. *Psychol. Belg* 44:43-57
- Perrin F, Grimault N (2005) Fonds Sonores (Version 1.0) [Sound samples]. Available at: <http://olfac.univ-lyon1.fr/unite/equipe-02/FondsSonores.html> [Accessed April 7, 2010].
- Pfurtscheller G, Aranibar A (1977) Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroen. Clin. Neuro* 42:817-826

- Pins D, Ffytche D (2003) The neural correlates of conscious vision. *Cereb. Cortex* 13:461-474
- Polich J, Ehlers CL, Otis S, Mandell AJ, Bloom FE (1986) P300 latency reflects the degree of cognitive decline in dementing illness. *Electroen. Clin. Neuro* 63:138-144
- Pulvermüller F, Preissl H, Lutzenberger W, Birbaumer N (1996) Brain rhythms of language: nouns versus verbs. *Eur. J. Neurosci* 8:937-941
- Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ (1999) Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397:430-433
- Shahin AJ, Picton TW, Miller LM (2009) Brain oscillations during semantic evaluation of speech. *Brain Cogn* 70:259-266
- Shipley E (1965) Detection and recognition: Experiments and choice models. *J. Math. Psychol* 2:277-311
- Suzuki T, Yamamoto K, Taguchi K, Sakabe N (1976) Reliability and validity of late vertex-evoked response audiometry. *Audiology* 15:357-369
- Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J. Neurosci* 16:4240-4249
- Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1997) Oscillatory gamma -Band (30-70 Hz) Activity Induced by a Visual Search Task in Humans. *J. Neurosci* 17:722-734
- Tiitinen H, Sinkkonen J, Reinikainen K, Alho K, Lavikainen J, Näätänen R (1993) Selective attention enhances the auditory 40-Hz transient response in humans. *Nature* 364:59-60
- Tononi G, Srinivasan R, Russell DP, Edelman GM (1998) Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *P. Natl. Acad. Sci. USA* 95:3198-3203
- Urban MJ (1992) Auditory subliminal stimulation: a re-examination. *Percept. Motor Skill* 74:515-541
- Vanhaudenhuyse A, Laureys S, Perrin F (2008) Cognitive event-related potentials in comatose and post-comatose states. *Neurocrit. Care* 8:262-270
- Varela F, Lachaux JP, Rodriguez E, Martinerie J (2001) The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci* 2:229-239
- Wyart V, Tallon-Baudry C (2008) Neural dissociation between visual awareness and spatial attention. *J. Neurosci* 28:2667-2679
- Yamagishi N, Callan DE, Goda N, Anderson SJ, Yoshida Y, Kawato M (2003) Attentional modulation of oscillatory activity in human visual cortex. *NeuroImage* 20:98-113
- Ziegler JC, Besson M, Jacobs AM, Nazir TA, Carr TH (1997) Word, Pseudoword, and Nonword Processing: A Multitask Comparison Using Event-Related Brain Potentials. *J. Cogn. Neurosci* 9:758-775



# Chapitre 6

## Conclusions et perspectives

L'acte de percevoir permet à l'individu de se représenter son environnement par l'intermédiaire de ses cinq sens, et ainsi d'interagir avec celui-ci. La perception des stimulations sensorielles présentes dans l'environnement est sans cesse influencée par des mécanismes dont l'individu n'a pas nécessairement conscience. De nombreuses études ont mis en évidence le fait que, lorsqu'une stimulation sensorielle est perçue, des connaissances relatives à la stimulation sont activées, ce qui faciliterait le traitement de cette stimulation. De plus, l'activation des connaissances faciliterait la perception ultérieure des stimulations sensorielles comme le reflètent les effets d'amorçage : une stimulation reliée (par une relation sémantique par exemple) au contexte qui la précède est plus rapidement traitée qu'une stimulation qui n'est pas reliée au contexte. Alors que les influences des connaissances ont été observées sur les hauts niveaux de traitement, comme la catégorisation ou la reconnaissance, seules quelques études suggèrent une influence sur les bas niveaux de traitement, comme la détection. Par ailleurs, s'il a été largement démontré qu'un contexte consciemment perçu activait des connaissances lexicales qui influençaient la perception ultérieure (comme dans le cas de l'amorçage sémantique), il n'est pas unanimement admis qu'une stimulation présentée sous le seuil de perception consciente puisse aussi activer des connaissances lexicales et engendrer des processus sémantiques inconscients. En effet, la majorité des données provient de la modalité visuelle, et peu d'études ont abordé ces phénomènes sous l'angle de la modalité auditive. L'objectif de ce travail doctoral a été d'explorer le fonctionnement de ces deux mécanismes et de rechercher, quasiment pour la première fois, si des mécanismes comparables existaient aussi pour la modalité auditive.

## 6.1 Influences des caractéristiques énergétiques des signaux

Dans l'ensemble des études de ce travail doctoral, nous devions nous assurer que les résultats obtenus ne pouvaient pas résulter d'un traitement *bottom-up* lié aux caractéristiques énergétiques des stimulations auditives. Dans cet objectif, il a été nécessaire de contrôler aussi précisément que possible les caractéristiques physiques et linguistiques des stimulations utilisées.

Les mécanismes impliqués dans les stades précoces de traitement des stimulations sensorielles sont difficiles à contrôler, tant au niveau du traitement des paramètres physiques qu'au niveau du traitement des caractéristiques linguistiques. Afin de s'assurer que seules les variables étudiées (aspects lexicaux et/ou phonologiques) différencient les stimulations utilisées, la création des stimulations auditives a été réalisée sous contrôle des paramètres psychoacoustique et psycholinguistique. Les caractéristiques énergétiques des trois types de stimulations (mots, pseudo-mots et sons complexes sans composantes phonologiques) ont été égalisées le plus finement possible en termes de durée, d'intensité, d'enveloppe temporelle et de spectre fréquentiel moyen et, entre les mots et les pseudo-mots, en termes de cooccurrence de phonèmes, de nombres de phonèmes et de lettres. Pour s'assurer que la méthode d'égalisation d'intensité entre les trois catégories de stimulations n'induit pas d'effet *bottom-up*, les deux modèles les plus courants d'égalisation énergétique ont été testés (dB-A et dB-SPL). Quel que soit le modèle d'égalisation des stimulations, les résultats demeurent les mêmes. De plus, une modélisation de la sonie des stimulations auditives a été appliquée *a posteriori* sur nos stimulations (Glasberg & Moore, 2002). Cette modélisation montre que les stimulations ayant la sonie la plus importante (c'est-à-dire les sons complexes) sont par ailleurs les stimulations les moins détectées. Bien que cette modélisation soit basée sur un modèle physiologique du traitement des signaux sonores, elle ne permet pas de prédire les résultats obtenus. Toutefois, le modèle de sonie de Glasberg et Moore (2002) étant le plus adapté aux sons complexes, les futures études utilisant des sons complexes devraient égaliser la sonie des stimulations lors de la création du matériel.

Lors des enregistrements des sons de parole, les variations de saillance de la hauteur fondamentale ont été minimisées par une prononciation monotone des mots et des pseudo-mots. Bien que la saillance de la hauteur fondamentale des sons complexes dépendent de la hauteur fondamentale des signaux originaux, une analyse *a posteriori* a révélé que des différences minimales perduraient entre les stimulations de parole (mots et pseudo-mots) et les sons complexes (différence estimée à 15% - voir Étude II, note de bas de page 116 -). Cette différence de saillance de la hauteur fondamentale pourrait expliquer une partie des résultats mais ne permet pas d'expliquer tous les résultats observés car aucune différence de saillance de la hauteur fondamentale n'est observée entre les mots et les pseudo-mots. Par conséquent, les résultats présentés dans les sections suivantes de la discussion générale ne peuvent pas être mis en relation avec un effet *bottom-up* dans le traitement acoustique des stimulations. Cependant, pour s'affranchir des effets de saillance de la hauteur fondamentale, des signaux sonores n'ayant pas de hauteur fondamentale, pourraient être utilisés lors de futures expériences. Il serait intéressant par exemple d'utiliser des sons de parole chuchotée qui ne sont pas des sons voisés (c'est-à-dire sans vibration des cordes vocales) et n'induisent donc pas de fréquence fondamentale. Il serait également possible de contrôler la fréquence fondamentale en utilisant de la parole sinusoïdale (*sine-wave-speech*, Remez, Rubin, Pisoni, & Carrell, 1981). Avec de tels signaux, il serait possible de contrôler plus précisément les différents paramètres acoustiques qui ont été égalisés pour les études du présent travail doctoral (spectre, durée, amplitude, enveloppe temporelle), ainsi que la fréquence fondamentale. Différentes stimulations pourraient être construites, dont un ou plusieurs paramètres varieraient, ce qui permettrait d'explorer l'influence de chacun de ces paramètres sur la détection. De telles explorations psychoacoustiques n'ont jamais été réalisées dans le domaine de la perception des sons de parole et pourraient aider à définir clairement l'importance des contraintes acoustiques, souvent ignorées dans les études rapportées en psychologie cognitive ou en neurosciences étudiant la perception auditive (excepté par exemple dans les travaux de Giraud et al., 2004).

## 6.2 Traitements cognitifs des sons de parole

Les études présentées dans ce travail doctoral apportent des données concernant les processus participant à l'analyse des traits acoustiques, phonologiques et sémantiques des stimulations. Dans l'Étude V, les différences électroencéphalographiques observées sur les composantes N1, N2-P3 (voir Figure 1, page 184) et l'activité dans la bande de fréquence bêta (voir Figure 3, page 188) pourraient refléter des mécanismes neuronaux différentiels pour les trois catégories de stimulations (mots, pseudo-mots et sons complexes). Puisque les trois catégories de stimulations ont été égalisées le plus finement possible d'un point de vue énergétique, les différences observées peuvent être mises en relation avec les caractéristiques phonologiques et/ou sémantiques des stimulations.

La gradation d'amplitude de la composante N1 (plus ample pour les sons complexes que pour les pseudo-mots, et plus ample pour les pseudo-mots que pour les mots) peut être mise en relation avec la sonie de la stimulation et/ou avec les influences des connaissances sur la détection (décrite dans l'Étude II et l'Étude III). Premièrement, la modélisation de sonie (Glasberg et Moore, 2002), appliquée aux stimulations créées pour ce travail de thèse, suggérerait que les sons complexes avaient une sonie plus importante que les pseudo-mots ou les mots. La modulation d'amplitude de la composante N1 pourrait donc refléter la différence d'intensité des signaux acoustiques, ce qui serait en accord avec les études qui ont rapporté la sensibilité de cette composante électroencéphalographique à l'intensité de la stimulation (par exemple, Musiek et al., 2005). Deuxièmement, la modulation de la composante N1 pourrait refléter une facilitation du traitement des stimulations pour lesquelles nous avons des connaissances, faisant probablement intervenir des processus *top-down* sur les cortex sensoriels primaires (dans la mesure où la composante N1 reflète en partie l'activation des structures primaires). Récemment, Coch et Mitra (2010) ont montré une modulation similaire de la composante P150 : son amplitude était plus grande en réponse à des non-mots qu'en réponse à des pseudo-mots ou des mots. Les auteurs ont suggéré que cette modulation serait le reflet d'une influence *top-down* des connaissances orthographiques sur les premiers niveaux de traitement du signal visuel. De façon analogue, la modulation de la composante N1 observée dans la présente Étude V pourrait refléter une influence *top-down* des

connaissances phonologiques et/ou sémantiques sur les premiers niveaux de traitement du signal acoustique.

Les différences observées pour le complexe N2-P3 et dans la bande de fréquence bêta pourraient refléter des mécanismes de perception plus tardifs. La diminution de l'amplitude du N2-P3 observée pour les sons complexes, notamment dans les régions antérieures droites qui reflètent l'activation des cortex associatifs, pourrait correspondre à une absence d'analyse phonologique des sons complexes. La diminution de l'activité bêta pour les stimulations non lexicales (pseudo-mots et sons complexes) dans l'hémisphère droit, pourrait pour sa part refléter l'absence d'analyse sémantique associée au traitement des sons complexes et des pseudo-mots. Cette dernière observation est concordante avec une étude de Pulvermüller et ses collaborateurs (1996) qui montrait une diminution de l'activité oscillatoire autour de 20-30 Hz pour les stimulations verbales non lexicales. Les auteurs suggéraient alors que les activités neuronales de hautes fréquences (autour de 20-30 Hz) avaient un rôle prépondérant dans les processus lexicaux.

Les différences hémisphériques observées pour le complexe N2-P3 et pour l'activité bêta entre les trois catégories de stimulations sont concordantes avec les études de neuropsychologie concernant les patients aphasiques, les études de neuro-imagerie et les études d'électroencéphalographie qui suggèrent le rôle dominant de l'hémisphère gauche dans les analyses linguistiques, notamment phonologiques et sémantiques (par exemple, Binder et al., 2000 ; Ziegler et al., 1997). Au-delà des différences linguistiques des stimulations, les variations hémisphériques pourraient aussi refléter le fait que les sons complexes étaient composés de fréquences dont la phase a été aléatoirement mélangée, c'est-à-dire dont la cohérence spectrale était moins probable que celle des mots ou des pseudo-mots. En effet, il a été suggéré que l'hémisphère gauche serait davantage impliqué dans le traitement des aspects temporels et que l'hémisphère droit serait davantage impliqué dans le traitement des aspects spectraux (Zatorre, Belin, & Penhune, 2002). Cette sensibilité fonctionnelle des hémisphères pourrait alors se refléter par une activation différentielle des deux hémisphères se retrouvant sur les potentiels évoqués et les réponses induites dans la bande de fréquence bêta.

Enfin, l'Étude V a apporté des renseignements concernant les activités neuronales associées aux processus attentionnels indépendants de la nature de la

stimulation ou de la tâche. En effet, des baisses similaires d'activités dans les bandes de fréquence alpha et gamma ont été observées en réponse aux trois catégories de stimulations, reflétant un niveau constant de vigilance à l'apparition des stimulations (par exemple Debener et al., 2003 ; Klimesch, Doppelmayr, Russegger, Pachinger, & Schwaiger, 1998).

## **6.3 Influences des connaissances sur la perception auditive**

Les études de ce travail doctoral apportent des données qui montrent que les connaissances sémantiques et phonologiques influencent les hauts niveaux de traitement et qui suggèrent que les connaissances exerceraient également une influence sur les bas niveaux de traitement.

### **6.3.1 Influences des connaissances sur les traitements de hauts niveaux**

Le travail de thèse a montré l'influence des connaissances sur les hauts niveaux de traitement, appelée effets de supériorité des mots (WSE) ou des pseudo-mots (PWSE), (pour une revue, voir Grainger, 2008) même si les différentes études n'ont pas directement porté sur ces influences.

La première expérience de l'Étude II, dans laquelle une tâche de reconnaissance suivait immédiatement la tâche de détection d'une stimulation auditive, ne rapportait ni un WSE ni un PWSE. Cette absence d'effets pourrait être expliquée par le fait que la présentation de la stimulation était séparée de la tâche de reconnaissance par une tâche de détection. L'ajout de la tâche de détection induisait un délai entre la présentation de la stimulation et la réalisation de la tâche qui était probablement responsable de l'absence de WSE ou de PWSE. En effet, les facilitations linguistiques sont connues pour être limitées dans le temps (par exemple Neely, 1977). L'hypothèse d'une influence de délai a été validée lors de l'Étude III, dans laquelle le même matériel a été utilisé excepté le fait que la tâche impliquant un traitement de haut niveau (tâche de catégorisation) succédait directement à la présentation de la stimulation. L'hypothèse d'une influence de délai a été validée lors de l'Étude III qui

a montré les effets de WSE et PWSE. Dans cette étude, le même matériel que dans l'Étude II était utilisé mais le délai entre la présentation de la stimulation et la réalisation de la tâche était réduit (les participants devaient répondre le plus rapidement possible), et le paradigme ne comprenait pas une tâche interposée (absence de la tâche de détection).

Une étude très récente, réalisée en modalité visuelle (Coch & Mitra, 2010), a exploré les bases neuronales associées aux WSE et PWSE. Les auteurs rapportent des influences *top-down* des connaissances linguistiques sur les premiers niveaux de traitement (voir Section 6.2) et sur les niveaux de traitement plus tardifs. Ces influences se refléteraient respectivement par la composante P150, qui apparaît plus ample en réponse aux non-mots qu'en réponse aux mots ou aux pseudo-mots, et par les composantes N300 et N400, qui apparaissent plus amples pour les mots et les pseudo-mots que pour les non-mots. Il serait intéressant d'étudier si ces résultats sont reproductibles dans la modalité auditive avec des stimulations qui sont contrôlées acoustiquement. L'Étude V, réalisée en EEG, contient un résultat qui étend en modalité auditive les résultats reportés dans l'étude de Coch et Mitra (2010). L'analyse en potentiels évoqués sur la composante N1 montre que les sons complexes ont une amplitude maximale significativement supérieure à celle des pseudo-mots et des mots. Des différences entre le traitement des mots et des pseudo-mots ont également été révélées dès la première centaine de millisecondes. En effet, l'activité N1 (50-200 ms) est plus ample pour les pseudo-mots que pour les mots au niveau des régions postérieures (P5-P6). Ces résultats sont en accord avec ceux reporté par Coch et Mitra (2010) en modalité visuelle. Ces différences précoces d'activité pourraient alors être mises en relation directe avec une activation *top-down* des connaissances.

### **6.3.2 Influences des connaissances sur les traitements de bas niveaux**

L'Étude I et l'Étude II ont montré que lorsque la détection devient difficile autour du seuil de perception, un effet de supériorité de détection des sons de parole (*Speech Detection Effect*, SDE) et un effet de supériorité de détection des mots (*Word Detection Effect*, WDE) apparaissaient. Pour un même niveau d'intensité sonore (égalisation réalisée a priori) ou pour un même niveau de sonie (égalisation estimée a

posteriori), les performances de détection des sons de parole étaient meilleures que celles des sons complexes (SDE), et les performances de détection des mots étaient meilleures que celles des pseudo-mots et des sons complexes (WDE). En d'autres termes, les connaissances phonologiques et sémantiques amélioreraient la détection auditive.

Les résultats suggèrent que l'effet dû aux connaissances lexicales (WDE) interviendrait seulement lorsque l'individu est engagé dans une tâche demandant une représentation du signal à un plus haut niveau et/ou que le niveau d'attention est plus élevé que celui requis lors de la réalisation d'une tâche de détection. En effet, le WDE a été observé au cours de l'Étude I lorsque les participants réalisaient la tâche de détection et que les stimulations étaient regroupées selon leur nature dans des blocs expérimentaux séparés. Il a également été observé au cours de l'Étude II, lorsque les participants étaient engagés dans une expérience de détection incluant une seconde tâche de reconnaissance. En revanche, le WDE n'était pas observé lorsque les participants réalisaient seulement la tâche de détection et que les stimulations étaient aléatoirement mélangées dans des mêmes blocs.

Des résultats analogues ont également pu être observés lors de l'Étude V qui ne s'intéressait pas directement aux WDE et SDE. Au cours de cette étude, les participants devaient catégoriser des silences mais aussi des sons complexes, des pseudo-mots et des mots présentés selon deux intensités sonores différentes (l'une à 40 dB-A, permettant la catégorisation, et l'autre à 5 dB-A, ne permettant pas la catégorisation). Lorsque les participants répondaient qu'une stimulation à 5 dB-A était présente (c'est-à-dire répondaient « mot », « pseudo-mot » ou « sons », mais pas « silence »), leurs performances de catégorisation étaient au hasard. Toutefois, les participants répondaient plus souvent qu'une stimulation à 5 dB-A était présente lorsqu'il s'agissait d'un mot ou d'un pseudo-mot que lorsqu'il s'agissait d'un son complexe. Ces résultats supplémentaires (Annexe 2, page 216) suggèrent alors indirectement que les sons de parole (mots et pseudo-mots) étaient mieux détectés que les sons complexes et que les mots étaient mieux détectés que les pseudo-mots, à un niveau d'intensité qui ne permettait pas la catégorisation (les performances de catégorisation étaient au niveau du hasard). Cette étude permet donc de confirmer les SDE et WDE observés pour l'Étude I et pour l'Étude II lorsque les participants sont



engagés dans tâche qui demande un traitement de haut niveau comme la catégorisation.

Les résultats de l'Étude V ont montré que les premiers niveaux de traitement (probablement réalisés au niveau des cortex sensoriels) étaient probablement influencés de façon *top-down* par les connaissances lors des processus de catégorisation : l'amplitude de la composante N1 diminuerait lorsque des connaissances relatives à la stimulation seraient stockées (voir Section 6.1). Il serait intéressant de rechercher si de telles modulations électroencéphalographiques sont observées au cours d'une tâche de détection et d'explorer si l'influence des connaissances peut être effective sur des plus bas niveaux neuronaux, par exemple aux niveaux sous-corticaux. En effet, nous pourrions faire l'hypothèse que les connaissances de l'auditeur moduleraient le traitement de l'information auditive encore plus tôt, c'est-à-dire aux différents niveaux du tronc cérébral, grâce aux projections corticales descendantes qui empruntent les voies efférentes (voir Section 1.2.4). Il serait par exemple intéressant d'enregistrer l'activité du tronc cérébral (en utilisant la réponse d'adoption de fréquence, *frequency following response*, FFR). Les réponses enregistrées permettraient de tester l'influence *top-down* des connaissances au niveau des voies efférentes (pour une revue, voir Krishnan & Gandour, 2009). Ce type d'investigation a été réalisé en modalité auditive en utilisant des sons de musique. De récents travaux ont montré que les connaissances musicales de l'auditeur influenceraient le traitement de saillance de la hauteur fondamentale pour la discrimination de la hauteur au niveau comportemental (Marmel, 2009), mais aussi au niveau électroencéphalographique (Marmel, 2009), et enfin au niveau du tronc cérébral (Marmel, Parbery-Clark, Skoe, & Kraus, 2009).

L'influence des connaissances lexicales sur la détection pourrait également être étudiée en utilisant des sons de parole sinusoïdale, ou plus généralement de la parole dégradée manipulant certaines propriétés acoustiques. En effet, la première fois que des signaux de parole sinusoïdale sont entendus, l'auditeur est incapable de détecter un quelconque contenu sémantique et/ou phonologique. Il perçoit plutôt un signal robotique sans aucun sens (semblable aux sons d'antennes de télévision). Lorsque l'auditeur est informé du contenu lexical de la stimulation, il devient alors tout à fait capable de l'identifier grâce à ses connaissances (Vouloumanos, Kiehl, Werker, & Liddle, 2001). Une expérience en EEG dans laquelle les participants seraient exposés

une première fois aux sons de parole sinusoïdale sans avoir été préalablement informés du contenu lexical, puis seconde fois après avoir été informé du contenu lexical, pourrait être réalisée. Ainsi, en comparant l'activité électrique engendrée par la première présentation des sons de parole sinusoïdale à celle engendrée par la seconde présentation des mêmes signaux, seule l'activité liée à l'implication des connaissances lexicales devrait émerger.

## **6.4 Influence de la perception auditive inconsciente sur la perception ultérieure**

Dans ce travail de thèse, des perceptions auditives non conscientes ont pu être observées aussi bien comportementalement par l'utilisation de mesures subjectives et objectives qu'électroencéphalographiquement par l'exploration des activités électriques cérébrales. Des signes de perception non consciente ont été illustrés lors de l'Étude II par un effet de répétition observé pour les mots seulement. Dans un paradigme de dissociation, les participants qui ne détectaient subjectivement pas les mots présentés à une faible intensité dans un premier temps étaient capables de les reconnaître dans un second temps. Lors de l'Étude IV, un effet d'amorçage sémantique mais aussi un effet d'amorçage négatif de répétition étaient observés suite à la présentation auditive subliminale de stimulations lexicales. Enfin lors de l'Étude V, une activation spécifique aux stimulations lexicales non catégorisées consciemment dans la bande de fréquence bêta a permis de mettre en évidence que les stimulations lexicales non catégorisées consciemment pouvaient cependant accéder à un niveau de représentation lexical inconscient.

### **6.4.1 Conditions expérimentales utilisées**

La principale difficulté dans l'étude de la perception non consciente est de définir quels sont les paramètres permettant de considérer un traitement comme non conscient. Dans ce travail de thèse, nous avons visé à apporter des éléments de réponse à cette question encore débattue en définissant précisément les conditions expérimentales dans lesquelles ces résultats ont été obtenus.

La seule étude qui a utilisé un paradigme d'amorçage inconscient auditif (Kouider & Dupoux, 2005) se base sur une analogie avec la technique de présentation subliminale utilisée en modalité visuelle et utilise une technique de masquage. Les systèmes auditifs et visuels fonctionnant différemment, il s'avère que l'utilisation de masques auditifs ne remplit pas la même fonction en audition qu'en vision. Alors que la persistance rétinienne joue un rôle déterminant lors du masquage visuel (Enns, 2002), le masquage auditif ne bénéficie pas de persistance auditive semblable (Glasberg, Moore, & Bacon, 1987). Entourer une stimulation de masques auditifs ne réduit donc pas les informations énergétiques des stimulations au cours du temps. En revanche, un masque présenté de façon concomitante à la stimulation peut réduire certaines informations (énergétiques et/ou informationnelles) de la stimulation. Par conséquent, aucun masquage ni aucune compression temporelle n'ont été utilisés dans les études présentées dans ce travail doctoral. Nous avons cependant choisi de réduire l'intensité sonore pour que le bruit interne du système auditif des participants "masque" de façon énergétique les stimulations. Par ailleurs, l'analogie qui peut être faite entre une présentation visuelle brève et une compression temporelle des stimulations auditives peut se révéler problématique dans la mesure où le traitement temporel séquentiel des stimulations auditives est une composante inhérente à la modalité auditive. Par conséquent, nous avons également fait le choix de ne pas réaliser de compression temporelle des stimulations auditives. Au cours de ce travail doctoral, nous avons donc manipulé l'intensité des sons pour que les conditions expérimentales d'observation d'une perception non consciente correspondent soit au fait que les participants soient incapables d'effectuer une catégorisation lexicale (Étude IV et Étude V), soit au fait qu'ils soient incapables de réaliser une détection subjective (Étude II) de la stimulation auditive présentée.

## **6.4.2 Perception non consciente des stimulations lexicales**

### **a. Analyse de bas niveau**

L'Étude II et l'Étude IV ont montré que la perception non consciente des stimulations sensorielles s'observait seulement pour des stimulations lexicales. La présentation des pseudo-mots et des sons complexes n'a pas entraîné de perceptions

non conscientes, même à un bas niveau de traitement c'est-à-dire dans des situations de répétition de la stimulation.

Une interprétation possible de ce résultat est que seules les stimulations familières ayant une représentation lexicale pourraient être traitées non consciemment. La multitude d'événements existants dans notre environnement, non seulement auditifs, mais également visuels, somesthésiques et chimiques, ne permet pas à l'individu de créer une représentation consciente de chaque événement. L'analyse non consciente des stimulations permettrait à l'individu de traiter beaucoup d'informations sans solliciter autant de ressources attentionnelles pour chaque événement. Seules les stimulations lexicales pour lesquelles des connaissances pertinentes sont stockées en mémoire, pourraient être analysées inconsciemment. Ce mécanisme de traitement non conscient des stimulations s'avèrerait utile dans le sens où toutes les stimulations pertinentes pour l'individu présentes dans notre environnement ne peuvent pas être traitées de façon consciente. Ainsi, la stimulation pertinente traitée non consciemment aurait une importance particulière dans le contexte où se trouve l'individu. Nous pouvons alors supposer que le traitement non conscient de la stimulation pertinente puisse influencer le comportement de l'individu dans le but de réorienter son attention vers la stimulation perçue inconsciemment.

De futures études pourraient nous aider à comprendre ce qui, au-delà des aspects lexicaux, est vraiment pertinent dans la stimulation pour qu'elle soit perçue inconsciemment. Il serait intéressant de tester notamment deux caractéristiques : la charge émotionnelle et la signification. Une stimulation ayant une composante émotionnelle pourrait-elle davantage être traitée inconsciemment par rapport à une stimulation sans composante émotionnelle ? Est-ce qu'une stimulation non verbale mais contenant une information pertinente pourrait être perçue inconsciemment ? Comme nous l'avons vu dans ce travail de thèse, il est possible de comparer des sons complexes tout en contrôlant finement leurs caractéristiques énergétiques et acoustiques. Les sons de l'environnement pourraient être testés dans un paradigme d'amorçage inconscient similaire à celui utilisé lors de l'Étude IV. Certains sons environnementaux contenant ou non une information pertinente (comme une sirène d'alarme ou un coup de klaxon qui sont des stimulations auditives permettant de signaler un danger) pourraient être utilisés comme amorces non conscientes, et nous pourrions tester si la présentation d'une cible associée à l'amorce est catégorisée plus

rapidement ou plus justement quand l'amorce est pertinente que lorsqu'elle ne l'est pas. La part spécifique de la composante émotionnelle impliquée dans le traitement non conscient serait également intéressante à étudier. Pour cela, des stimulations auditives contenant une information émotionnelle mais aucune information lexicale, comme les pseudo-mots à prosodie émotionnelle variable créés par Banse et Scherer (1996), pourraient également être utilisées dans un paradigme d'amorçage auditif inconscient.

## **b. Traitement sémantique**

Pour la première fois en modalité auditive, l'Étude IV et l'Étude V ont permis de mettre en évidence un traitement sémantique des stimulations lexicales non perçues consciemment. Dans l'Étude IV, un effet d'amorçage sémantique a été observé et les résultats de l'Étude V ont permis d'identifier une activité bêta, latéralisée à gauche, en relation avec le traitement des stimulations lexicales non perçues consciemment. L'effet d'amorçage négatif de répétition observé dans l'Étude IV confirmerait l'activation sémantique inconsciente dans la mesure où cet effet révèle l'activation sémantique mais aussi l'inhibition centrale liée au traitement l'amorce (voir Discussion de l'Étude IV, page 167).

Bien qu'un traitement sémantique non conscient ait été rapporté plusieurs fois dans la littérature en modalité visuelle (pour une méta-analyse, voir Van den Bussche et al., 2009), ce résultat n'a jamais été mis en évidence en modalité auditive. La seule étude mettant en évidence un traitement auditif non conscient a révélé un effet de répétition, mais pas un effet d'amorçage sémantique (Kouider & Dupoux, 2005). Dans cette étude utilisant un paradigme d'amorçage inconscient, les stimulations lexicales utilisées comme amorces ont subi une compression temporelle et étaient entourées de masques auditifs. La stimulation cible, dont l'intensité était augmentée de 15 dB, était présentée immédiatement après la fin de la stimulation amorce. Alors que les participants n'étaient pas capables de réaliser une tâche de catégorisation de parole sur l'amorce, de meilleures performances à la tâche de décision lexicale de la cible sont obtenues lorsque l'amorce et la cible ont une relation de répétition. Bien que les amorces ne soient pas consciemment catégorisées ( $d' < .25$  et significativement supérieur à zéro dans une tâche de discrimination de parole ou lexicale), elles sont probablement détectées. Tandis que l'effet d'amorçage de

répétition mis en évidence par Kouider et Dupoux (2005) est interprété comme un traitement auditif non conscient, nous ne pouvons alors pas exclure que le traitement de la cible soit influencé par un amorçage sensoriel conscient.

L'Étude V montre une variation de la puissance bêta dans l'hémisphère gauche entre des stimulations lexicales et des stimulations non lexicales catégorisées consciemment. Une variation similaire a été retrouvée pour les stimulations non catégorisées consciemment, suggérant que ces dernières pouvaient atteindre un niveau de traitement lexical. Ces données sont les premières à mettre en évidence des modifications neuronales en lien avec la perception auditive inconsciente. Précédemment, certaines études se sont intéressées aux corrélats électroencéphalographiques de la perception consciente des stimulations visuelles ((et une étude pour la modalité somesthésique, voir Palva et al., 2005) et suggéraient que la perception consciente serait en lien avec des modifications neuronales oscillatoires et à une augmentation de synchronie dans certaines bandes de fréquence (par exemple, Melloni et al., 2007). Ces phénomènes oscillatoires corroborent l'hypothèse selon laquelle la conscience d'une stimulation émergerait de la synchronisation longue distance des structures cérébrales participant au traitement de l'information (par exemple Crick et Lock, 1990). Dans l'Étude V, il n'est pas possible de savoir si les structures cérébrales qui participent au traitement de l'information auditive consciente se synchronisent. Toutefois, les activités observées dans la bande de fréquence bêta sont également celles qui sont le plus souvent impliquées dans les phénomènes de synchronisations longue distance (par exemple, Bibbig et al., 2002 ; Gaillard et al., 2009 ; Gross et al., 2004). Afin de confirmer cette hypothèse, il serait intéressant dans de futures études réalisées en EEG de profondeur et/ou en MEG, de rechercher si effectivement les structures dont les neurones oscillent, pendant le traitement lexical inconscient des stimulations, se synchronisent.

### **6.4.3 Niveau de conscience**

La façon de mesurer les performances des participants (par une approche subjective ou objective) conditionne la façon d'interpréter les résultats et par conséquent de mettre en évidence, ou non, un traitement non conscient. Les résultats de l'Étude II mettent en évidence, par une mesure de seuil subjectif, un traitement auditif sans détection. L'analyse des mêmes données par une mesure de seuil objectif

ne le permet pas. Cependant, il est possible d'observer, par une mesure de seuil objectif, un traitement auditif non conscient sans catégorisation (Étude III, Étude IV et Étude V).

La mise en évidence de la perception non consciente repose alors sur sa façon de la mesurer. Deux paramètres sont importants dans la définition du traitement non conscient : la mesure effectuée et la tâche demandée. Les différentes mesures réalisées dans notre travail doctoral permettent de définir différents niveaux de perception à partir de ces deux paramètres (Figure 34). Si nous pourrions considérer qu'au-dessus du seuil objectif de discrimination, le participant perçoit consciemment, nos données suggèrent qu'aucune perception inconsciente n'est possible au-dessous du seuil subjectif de détection. Ceci n'est pas en accord avec toutes les études qui se sont intéressées à la perception non consciente (par exemple certains auteurs comme Snodgrass et collaborateurs considèrent qu'une perception inconsciente existe sous le seuil subjectif de détection), mais reflète plutôt les données de nos études.

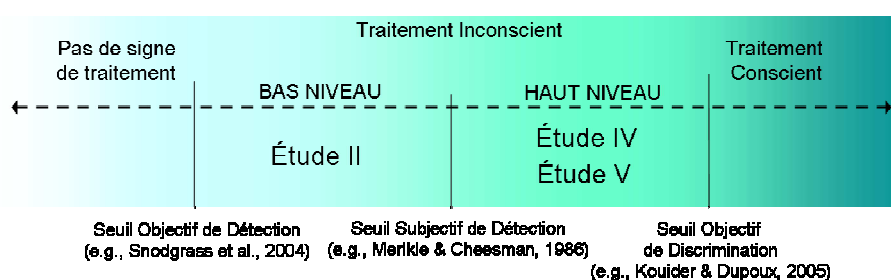


Figure 34 : Continuum du traitement de l'information selon la tâche demandée (de bas ou haut niveau) et la mesure effectuée (objective ou subjective).

Un traitement non conscient de bas niveau (répétition) est observé au cours de l'Étude II dans laquelle le niveau de conscience des participants se situe entre le seuil objectif et le seuil subjectif de détection. De plus, un traitement non conscient de haut niveau (accès lexical) est observé au cours de l'Étude IV et de l'Étude V dans lesquelles le niveau de conscience des participants se situe entre le seuil subjectif de détection et le seuil objectif de discrimination.

Alors qu'il paraît probable que le traitement sémantique non conscient soit clairement établi si la mesure du niveau de conscience a été réalisée avec une tâche demandant un haut niveau de traitement, il est beaucoup moins clair qu'il soit possible d'observer le même phénomène si la mesure du niveau de conscience a été réalisée avec une tâche demandant un bas niveau de traitement (pour une revue, voir Kouider

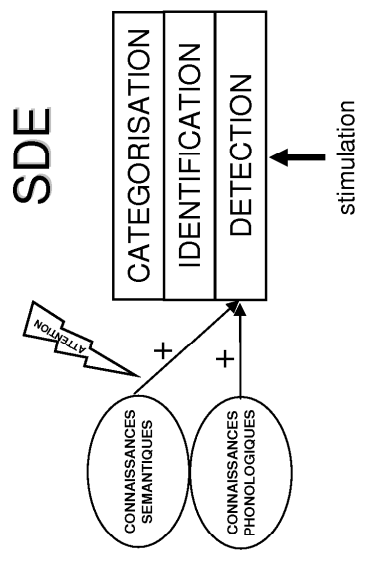
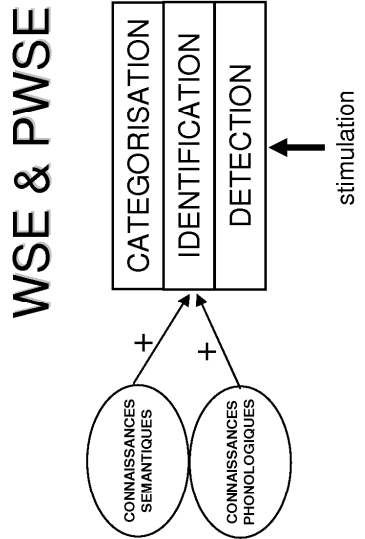
& Dupoux, 2007). Peu d'études ont effectivement rapporté des effets d'amorçage sémantique non conscients au niveau du seuil objectif de détection et une interprétation non sémantique a été proposée pour la plupart des études qui ont défendus de tels effets (Abrams & Greenwald, 2000). Cette interprétation suggère que les résultats d'amorçage sémantiques puissent être expliqués par des associations directes de type « stimulus-réponse », prédites par la théorie de la spécification motrice directe. Les auteurs préconisent alors de ne jamais répéter les stimulations chez un même participant.

Bien que nous pensions qu'il serait difficile d'observer des effets d'amorçage sémantique non conscients dans la modalité auditive au-dessous du seuil subjectif de détection, cette question pourrait être abordée dans une étude similaire à l'Étude IV, hormis que le niveau d'intensité de l'amorce serait plus faible, correspondant à un niveau de conscience pour lequel les participants ne peuvent pas réaliser de détection objective.

## **6.5 Conclusions**

Alors qu'il est établi que les connaissances influencent les traitements de haut niveau des stimulations, ce travail de thèse a permis de poursuivre cette réflexion et de mettre en évidence que les connaissances pouvaient également avoir une influence sur le plus bas niveau de traitement que représente la détection. De plus, l'activation des connaissances suite à la perception d'une stimulation sensorielle a été mise en évidence même lorsque la stimulation est perçue non consciemment (Figure 35).





## Amorçage subliminal

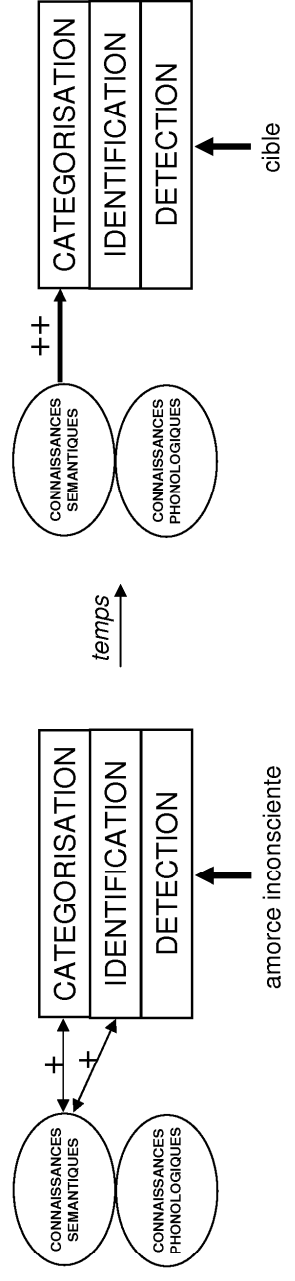


Figure 35 : Résumé des différents effets observés au cours du travail doctoral. Chacun d'entre eux montre des interactions entre les processus perceptifs (détection, identification, catégorisation sémantique) et les connaissances (phonologiques et sémantiques).

Notre travail de thèse a permis de mettre en évidence un certain nombre d'interactions entre les processus perceptifs (détection, identification, catégorisation sémantique) et les connaissances (phonologiques et sémantiques), qui sont éventuellement modulées par les processus attentionnels (Figure 35). Ces interactions sont à la base de processus de facilitation (ou d'inhibition dans le cas particulier de l'amorçage négatif de répétition, voir page 208) du traitement de l'information sensorielle (*Speech Detection Effect*, *Word* et *Pseudo-Word Superiority Effects*). Un contexte congruent pourrait potentialiser ces facilitations, même s'il est perçu inconsciemment (amorçage subliminal).

Parmi les nombreuses voies de recherche qui pourraient compléter notre modélisation temporaire des phénomènes explorés dans ce travail doctoral, il serait intéressant de rechercher maintenant s'il existe d'autres interactions entre les connaissances et les mécanismes perceptifs. Il est probable par exemple que l'activation des connaissances par la présentation subliminale d'un contexte puisse améliorer la détection, en d'autres termes qu'une amorce inconsciente augmenterait le *Speech Detection Effect* observé hors contexte. Enfin, il serait également intéressant de rechercher si d'autres connaissances (phonologiques, acoustiques) pourraient faciliter les processus perceptifs après un contexte perçu inconsciemment, ou si seules les connaissances lexicales peuvent être activées inconsciemment.

# ANNEXE 1

## Procédure pour audiométrie tonale

Le participant est assis dans une salle insonorisée et place le casque sur ses deux oreilles. Chaque oreille est testée séparément, en commençant par l'oreille droite puis en continuant par l'oreille gauche. Le participant est exposé à une première procédure de familiarisation avec la tâche. L'expérimentateur présente à l'oreille droite une stimulation sonore qui est un son pur d'une fréquence de 1000 Hz et d'une intensité de 60 dB (rendant la perception aisée). Le participant lève la main dès qu'il entend la stimulation sonore dans le casque. L'expérimentateur observe la réponse motrice du participant par l'intermédiaire d'une caméra. Ensuite, tout en restant à la même fréquence, l'expérimentateur diminue l'intensité de la présentation de la stimulation sonore par pas de 5 dB-SPL jusqu'à n'obtenir plus aucune réponse motrice du participant. Après une absence de réponse, l'expérimentateur diminue à nouveau l'intensité de la présentation de la stimulation sonore de 10 dB-SPL. Après chaque absence de réponse, l'expérimentateur remonte l'intensité de la présentation de la stimulation sonore par pas de 5 dB-SPL jusqu'à obtenir à nouveau une réponse motrice du participant. La première stimulation à nouveau perçue détermine le seuil de détection auditive pour la fréquence du son présenté (soit 1 000 Hz dans la procédure de familiarisation). L'expérimentateur renouvelle l'opération pour toutes les fréquences à tester. D'une manière générale, les fréquences suivantes sont testées dans l'ordre : 1 000 Hz, 2 000 Hz, 4 000 Hz, 6 000 Hz, 8 000 Hz, 125 Hz, 250 Hz et 500 Hz. L'exploration est effectuée en présentant des sons d'une durée d'une à deux secondes. L'intervalle entre les présentations successives doit être varié, mais ne doit pas être plus court que la durée de présentation des stimulations sonores. Pour chaque fréquence, le seuil de détection est testé trois fois de suite et validé si le participant répond au moins deux fois à la même intensité de présentation. Les résultats sont reportés sur une grille appelée audiogramme (Figure 36) présentant l'intensité en ordonnée et la fréquence de la stimulation sonore présentée en abscisse. Une audiométrie tonale est réalisée pour chaque oreille.

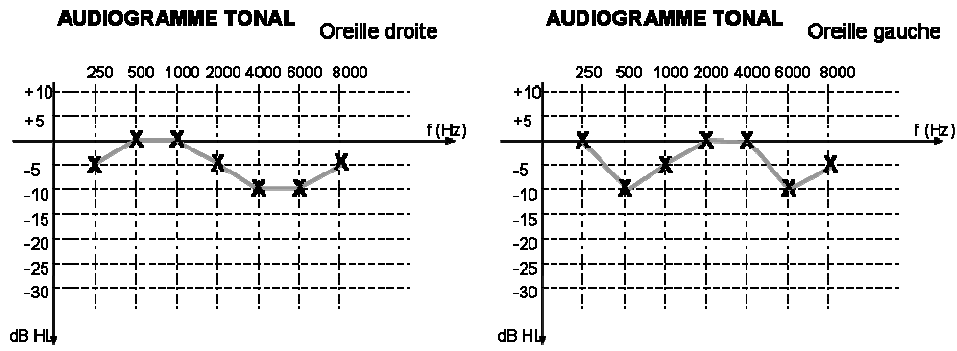


Figure 36 : Exemple d'audiogramme tonal enregistré chez un participant normo-entendant. Ne sont reportés que les seuils de perception (c'est-à-dire la plus faible intensité perçue pour chaque fréquence). Un audiogramme est réalisé pour chaque oreille. Chez un participant de moins de 30 ans, une audition normale dans le silence correspond à des seuils inférieurs à 20 dB entre 250 Hz et 8 000 Hz pour les deux oreilles.

## ANNEXE 2

### Résultats supplémentaires de l'Étude V

Les réponses comportementales de l'Étude V ont été analysées pour les stimulations non catégorisées consciemment (présentées à 5 dB-A) selon la perception subjective des participants (Figure 37). Lorsqu'une stimulation est présentée à 5 dB-A, les participants doivent catégoriser cette stimulation comme étant soit un mot, soit un pseudo-mot, soit un son complexe, soit un silence. Sont considérées comme détectées, les stimulations présentées à 5 dB-A pour lesquelles le participant n'a pas répondu "silence" mais "mot", "pseudo-mot" ou "son complexe" indifféremment. Sont considérées comme non détectées, les stimulations présentées à 5 dB-A pour lesquelles le participant a répondu "silence". Une analyse de variance a montré un effet principal de la Stimulation (mots, pseudo-mots, sons complexes) ( $F(2,26) = 7.83$  ;  $p = .002$ ) : les analyses *post-hocs* (LSD Fischer) montrent une meilleure détection des mots comparativement aux pseudo-mots ( $p = .044$ ) et aux sons complexes ( $p = .002$ ), et une meilleure détection des sons de parole comparativement aux sons complexes ( $p = .006$ ). Par ailleurs la détection des mots et des pseudo-mots est au niveau du hasard (c'est-à-dire 50% ;  $p = .862$  et  $p = .187$  respectivement) alors que la détection de sons complexes est inférieure au hasard ( $p = .031$ ).

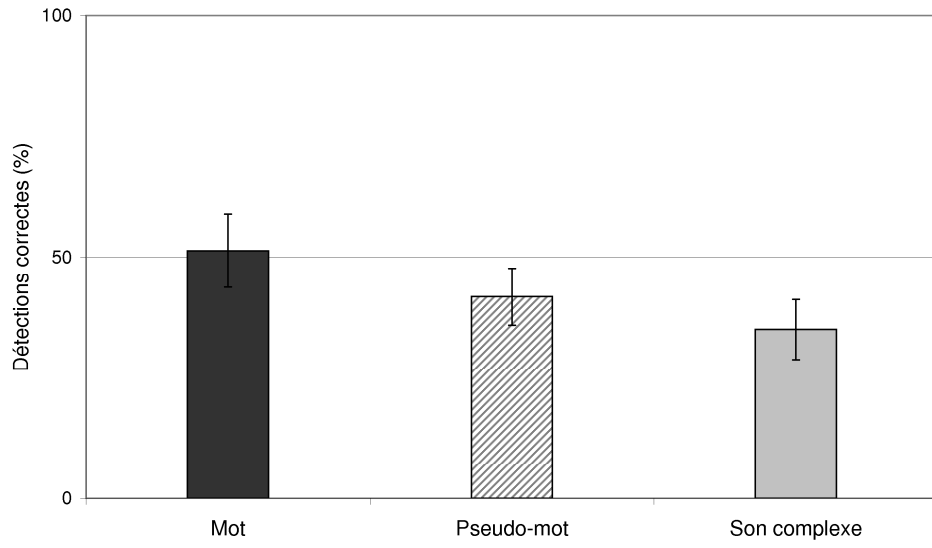


Figure 37 : Pourcentage de détections correctes à la tâche de catégorisation : les mots sont plus souvent catégorisés comme “mot”, “pseudo-mot” ou “son complexe” que les pseudo-mots ou les sons complexes.



# LISTES DES FIGURES

Figure 1 : Représentation d'une onde de vibration de forme sinusoïdale. Quatre cycles sont représentés, bien que cette onde peut se répéter à l'infini. L'amplitude instantanée est donnée par l'expression $A \sin(2\pi f t + \varphi)$ , où $t$ désigne le temps, $f$ la fréquence, $A$ l'amplitude maximale et $\varphi$ la phase initiale. La phase se mesure en degrés (notés °) ou radians. Un cycle complet correspond à $360^\circ$ ou $2\pi$ radians. (Adapté de Moore, 2003). .....	22
Figure 2 : Représentations acoustiques du mot « stage ». a) Amplitude globale de toutes les fréquences variant au cours du temps. b) Enveloppe temporelle représentant l'amplitude générale du son. c) Sonagramme obtenu par la transformée de Fourier. Les niveaux de gris indiquent l'amplitude de chaque fréquence au cours du temps. Les fréquences harmoniques correspondent aux traits noirs horizontaux les plus foncés. La distance observée entre deux fréquences harmoniques correspond à la valeur de la fréquence fondamentale (ici, 217 Hz). d) Amplitude donnée pour chaque fréquence d'un son complexe. e) Amplitude moyenne pour chaque fréquence sur une partie d'un son complexe. La courbe définie par les amplitudes maximales de chaque fréquence correspond à l'enveloppe spectrale. La distance observée entre deux pics fréquentiels correspond à la valeur de la fréquence fondamentale (ici, 217 Hz). .....	24
Figure 3 : Fréquences sonores (en Hertz) perçues par l'oreille humaine en fonction de l'intensité (en dB-SPL). La même sensation de sonie est ressentie de manière égale tout au long de chaque courbe d'isotonie représentée. La zone orange correspond à la zone conversationnelle dans laquelle les sons sont les mieux perçus. Les sons situés en dessous du seuil d'audition ne sont pas perçus par l'oreille humaine. La ligne rouge représente la pondération A, la ligne verte représente la pondération B et la ligne bleue représente la pondération C. (Adapté de Fletcher & Munson, 1933).....	29
Figure 4 : Le pseudo-mot « gucre » a été utilisé dans ces représentations. <i>En haut</i> : Forme de l'onde temporelle d'un son complexe. <i>En bas</i> : Sonie mesurée au cours du temps pour un son complexe. La courbe bleue représente la sonie instantanée, la courbe verte la sonie à court-terme et la courbe rouge la sonie à long-terme. ....	30
Figure 5 : Schéma de l'oreille externe, moyenne et interne, chez l'humain. (Adapté de Chittka & Brockmann, 2005).....	33
Figure 6 : Présentation schématique des voies auditives ascendantes, de la cochlée jusqu'au cortex auditif primaire (adaptée de Bonnet & Chancier, 1999). Indication de plusieurs relais de ces voies. Les fibres auditives de chaque oreille se projettent dans le cortex gauche et dans le cortex droit.....	35
Figure 7 : Potentiels évoqués de latence précoce (I à VI) de moyenne latence (N0 à Nb) et de latence tardive (P1 à N2). Ces composantes sont considérées comme sensorielles ou obligatoires car elles apparaissent à la suite de toutes stimulations auditives. (Adapté de Picton et al., 1974). .....	36
Figure 8 : a) Vue latérale du cortex cérébral gauche situant le sillon temporal supérieur (STS), le gyrus temporal supérieur (GTS), la scissure de Sylvius et laissant apparaître l'insula à l'aide des écarteurs (schéma adapté de Purves et al., 2004) b) Vue de dessus du plan temporal supérieur situant l'aire auditive primaire (A1) au niveau du gyrus de Heschl (GH). Les aires auditives associatives se situent dans la partie latérale du gyrus de Heschl, dans les planum polaire (PP) et temporale (PT) et dans le gyrus temporal supérieur (GTS) (schéma adapté du cerveau humain standard du Montréal Neurological Institut et Bidet-Caulet, 2007).....	37



Figure 9 : Représentation schématique de la localisation anatomique des générateurs des potentiels évoqués de latence moyenne (N/P16 pouvant correspondre au Na et N/P30 pouvant correspondre aux Pa et Nb observés en EEG) et tardifs (N/P50 pouvant correspondre au P1a et N/P60-75 pouvant correspondre au P1b et N/P100 pouvant correspondre au N1 observés en EEG) dans le gyrus de Heschl et le planum temporal (Liégeois-Chauvel et al., 1994).....	38
Figure 10 : a) Localisation du gyrus de Heschl au niveau du sillon temporal supérieur. b) organisation tonotopique du gyrus de Heschl en surface, sur une coupe latérale et sur une coupe horizontale (d'après Humphries et al., 2010).....	39
Figure 11 : Voies ascendantes (en noir) et descendantes (en rouge) du système auditif. Une première voie efférente descend du cortex au colliculus inférieur, puis au complexe olivaire supérieur pour ensuite atteindre la cochlée. Une seconde voie efférente se projette directement du cortex aux noyaux cochléaires qui vont ensuite atteindre le complexe olivaire supérieur avant de redescendre au niveau de la cochlée. Seule la voie descendante olivo-cochléaire est étudiée de manière non-invasive chez l'humain (d'après Collet, 2005). .....	40
Figure 12 : La distribution des réponses des participants se répartit sur l'axe $x$ suivant une distribution normale de moyenne $\mu$ et de même variance $\sigma$ . La moyenne de la distribution normale lorsqu'il y a une absence de signal est notée $\mu_N$ et $\mu_S$ lorsque le signal est présent. La distance entre les moyennes des deux distributions représente la valeur $d' \sigma$ . Le biais est noté $c$ et est lié au seuil interne $t$ du participant. En deçà du seuil interne, les participants répondent « pas de signal » et au-delà du seuil interne, les participants répondent « signal ».....	48
Figure 13 : Enregistrement de l'activité oscillatoire spontanée à la surface du scalp dans une région frontale (1) et dans une région occipitale (2). Apparition du rythme alpha observé lorsque l'individu éveillé ferme les yeux. Apparition du rythme bêta (et disparition du rythme alpha) observé lorsque l'individu éveillé ouvre les yeux. Le fait que l'activité alpha disparaisse lorsque les yeux sont ouverts a été nommé réaction d'arrêt (d'après Berger, 1930).....	50
Figure 14 À gauche : La notion de dipôle permet de comprendre que les neurones créent une activité électrique post-synaptique (en rouge) ainsi qu'une activité magnétique (en vert). À droite : L'activité électrique générée par les neurones Les neurones orientés parallèlement à la surface du scalp (comme les cellules pyramidales du cortex), créent des dipôles de courant radiaux lorsqu'ils sont situés sur un gyrus ou des dipôles tangentiels lorsqu'ils sont situés sur une scissure. ....	52
Figure 15 a) Deux secondes d'enregistrement brut de signal EEG. b) Ce signal est composé de plusieurs activités fréquentielles que l'on peut identifier grâce à une transformée de Fourier, et isoler par filtrage (à droite).....	54
Figure 16 : a) Observation sur Cz des potentiels évoqués N100, P200, P300 et N400 suite à la présentation d'une stimulation auditive rare (en gris) comparativement à la présentation d'une stimulation auditive fréquente (en noir) pour laquelle seule les potentiels évoqués N100, P200 et N200 sont observés. b) Suite à la présentation d'un mot sémantiquement incongru (en noir), observation sur Pz du potentiel évoqué N400 qui n'apparaît pas pour le mot sémantiquement congruent (d'après Vanhauzenhuyse, Laureys, & Perrin, 2008). .....	57
Figure 17 : Représentation schématique des réponses évoquées et induites par une stimulation après moyennage classique des signaux pour l'étude des réponses gamma (Tallon-Baudry & Bertrand, 1999). A) Les essais EEG présentent de faibles réponses gamma dont la latence d'apparition est calée après l'apparition de la stimulation (en bleu) ainsi que des réponses gamma plus amples dont la latence d'apparition est variable (en vert). B) Les méthodes classiques de moyennage mènent à la mise évidence d'un potentiel évoqué classique. C) Représentation de la réponse évoquée avec le temps en abscisse et la fréquence en ordonnée. L'échelle de couleur code la variation de puissance corrigée	

en ligne de base (pré-stimulation). D) Analyse temps-fréquence réalisée sur chaque essai individuellement. E) le moyennage entre les essais permet de révéler la réponse gamma induite.....	59
Figure 18 : Reconstruction 3D des cerveaux des patients ayant un déficit de catégorisation (d'après Damasio et al., 1996). a) Régions concernant les déficits de catégorisation des personnes célèbres, b) Régions concernant les déficits de catégorisation des animaux, c) Régions concernant les déficits de catégorisation des outils, d) Régions concernant les déficits de catégorisation des personnes célèbres, des animaux et des outils. e) Schéma récapitulatif des régions concernées par les déficits de catégorisation des personnes (en rouge), des animaux (en bleu) et des outils (en vert). .....	65
Figure 19 : Représentation du modèle de la parole proposé par Pisoni et ses collègues (Pisoni, 1978 ; Pisoni & Sawusch, 1975). .....	67
Figure 20 : Représentation du modèle d'Ellis et Young pour la reconnaissance des mots. Ce modèle postule que les mots soient reconnus grâce à la mise en œuvre de plusieurs systèmes. ....	70
Figure 21 : Représentation schématique du modèle autonome proposé par Norris, McQueen et Cutler (2000). Les flèches rouges indiquent le sens d'activation entre les différents niveaux alors que les flèches bleues indiquent une inhibition à un seul niveau (adapté de McClelland, Mirman, & Holt, 2006).....	72
Figure 22 : Localisation de l'aire de Broca au niveau du gyrus frontal extérieur et de l'aire de Wernicke (comprenant le planum temporale) au niveau du gyrus temporal supérieur. ....	73
Figure 23 : Activités cérébrales issues d'une TEP (Zatorre et al., 1992). a) Activation de régions cérébrales situées à gauche lors de l'analyse phonétique. b) Activation de régions cérébrales situées à droite lors de l'analyse de la hauteur. ....	74
Figure 24 Résumé des activations observées pour les stimulations de parole versus les stimulations de non parole. Les résultats des études de Zatorre et collaborateurs (1992) sont reportées en bleu clair, les résultats des études de Démonet et collaborateurs (1992) sont reportées en bleu marine, les résultats des études de Binder et collaborateurs (1997) sont reportées en jaune et les résultats des études de Binder et collaborateurs (2000) sont reportées en rouge (d'après Binder et al., 2000). ....	75
Figure 25 : Interactions prédites entre les traitements phonétiques (STG, gyrus temporal supérieur), lexicaux (SMG, gyrus supramarginal) et décisionnels (IFG, gyrus frontal inférieur) pour une explication top-down de l'effet Ganong, représentée schématiquement sur une reconstruction de surface corticale (d'après Gow et al., 2008). ....	78
Figure 26 : Représentation schématique du modèle TRACE proposé par McClelland et Elman en 1986 pour rendre compte de la reconnaissance des sons de parole. Les liens activateurs entre les niveaux sont en rouge, les liens inhibiteurs sont en bleu et les liens représentant une influence rétrograde sont en vert. ....	81
Figure 27 : Réponses N1-P2-N2 et P3 observées chez des participants contrôles ainsi que chez des patients dont l'état de conscience est altéré : dans des états de conscience minimale et dans des états végétatifs (d'après Perrin et al., 2006). ....	87
Figure 28 : Paradigme de présentation subliminale masquée. À chaque essai, un premier masque constitué de chaînes de caractères précède la stimulation amorcée qui est flashée durant quelques dizaines de millisecondes (entre 29 et 41 ms selon les études). En présentant un second masque immédiatement après l'amorce, on supprime la perception consciente de l'amorce (d'après Ferrand, 2001). ....	89
Figure 29 : Synchronisation de l'activité gamma entre les différentes électrodes de scalp pour les mots perçus consciemment (en haut) et les mots non perçus consciemment (en bas). ....	92

Figure 30 : Représentation schématique du fonctionnement de l'accès à la conscience dans un modèle neuronal d'espace de travail global (d'après Deheane et al., 2006). .....	94
Figure 31 : Le pourcentage de réponses correctes à la tâche de détection est meilleur pour les mots que pour les pseudo-mots, et est meilleur pour les pseudo-mots que pour les sons complexes. ....	104
<i>Figure 1.</i> After displaying a fixation cross, a stimulus (word, pseudo-word or complex sound) or a silence was presented (from -5 to 25 dB-A) and participants had to perform a detection task by pressing yes or no answer keys. Two hundred milliseconds after the response, two stimuli separated by 200 ms were presented at audible level (one being the same as the stimulus presented in the detection task -the target- and the other being a distractor of the same category, presented in a random order). Participants had to choose in a 2AFC recognition task the stimulus that was the same as the stimulus presented in the detection task by pressing one of two answer keys (first or second stimulus).....	116
<i>Figure 2.</i> Analyses with signal detection theory: $d'_D$ presented as a function of Stimulus Level (1 to 11, i.e., -5 to +25 dB-A ) for each Type of Stimulus (word/pseudo-word/complex sound). The error bars show the standard error.....	119
<i>Figure 3.</i> Percent correct in the recognition task presented as a function of Stimulus Level (1 to 11, i.e., -5 to +25 dB-A) for each Type of Stimulus (word/pseudo-word/complex sound). The error bars show the standard error. The dotted line represents chance level. ....	120
<i>Figure 4.</i> Percent correct in the recognition task presented as a function of Stimulus Level (1 to 5, i.e., -5 to +7 dB-A) and Type of Stimulus (word/pseudo-word/complex sound) for trials in which participants had a “no” response in the detection task. The grey line indicates the chance level and the error bars show the standard error. Point marked with an asterisk is significantly different from chance level. ....	121
<i>Figure 5.</i> Linear regression between $d'_R$ and $d'_D$ for each Type of Stimulus. The linear regression is the black line and confidence interval is represented by the dotted lines.....	122
<i>Figure 6.</i> Analysis with signal detection theory. <i>Left panel:</i> $d'_D$ presented as a function of Stimulus Level (1 to 5; 0 to +20 dB-A and +4 to +24 dB-SPL), Type of Stimulus (word/pseudo-word/complex sound) and Equalization. The error bars are the standard errors. <i>Right panel:</i> $d'_D$ presented as a function of Equalization (dB-A and dB-SPL) and Type of Stimulus (word/pseudo-word/complex sound). The error bars are the standard errors.....	129
<i>Figure 7.</i> Detection probability as a function of estimated loudness for the two experiments. The curves are cumulative Gaussian fitted to the raw data. The symbols show the detection probability averaged on different loudness intervals uniformly distributed on the loudness range.....	133
Figure 32 : Taux de discrimination en $d'$ entre une stimulation auditive et un silence (Condition 1), un mot et un son complexe (Condition 2), un pseudo-mot et un son complexe (Condition 3) et entre un mot et un pseudo-mot (Condition 4).....	151
Figure 33 : Pourcentages de réponses correctes aux différentes tâches de détection et catégorisation. La condition 1 compare les performances de détection entre mots, pseudo-mots et sons complexes. La condition 2 compare les performances de catégorisation de parole entre mots et sons complexes, la condition 3 compare les performances de catégorisation de parole entre pseudo-mots et sons complexes et la condition 4 compare les performances de catégorisation de parole entre mots et pseudo-mots. Les mots sont représentés par des carrés, les pseudo-mots par des triangles et les sons complexes par des ronds. ....	153

<i>Figure 1.</i> Response Time (Unit: milliseconds; with SD) to the lexical decision task in the Pre-experiment with semantically unrelated, semantically related, and repeated word pairs in slow (N=12) and fast responders (N=12). All differences are significant.....	164
<i>Figure 2.</i> Response Time (Unit: milliseconds; with SD) to the lexical decision task in the main Experiment with semantically unrelated, semantically related, and repeated word pairs in slow (N=12) and fast responders (N=12). ( $p < .05 *$ ; $p < .01 **$ ).....	165
<i>Figure 1. a)</i> Event-Related Potentials (N1, P2, N2 and P3) are presented as a function of Condition (categorized stimuli, uncategorized stimuli and silences) for each Electrode (Fz, Cz, Pz). The black curve represents the categorized stimuli; the grey, dashed curve represents the uncategorized stimuli; and the dotted curve represents the silences. Only the categorized stimuli presented classical auditory ERPs. <i>b)</i> Event-Related Potentials for categorized stimuli are presented as a function of Stimulation (words, pseudo-words and complex sounds) from one electrode in each quadrant (F5, F6, P5, P6). The black curve represents the categorized words; the grey, dashed curve represents the categorized pseudo-words and the dotted curve represents the categorized complex sounds. Higher N2 and P2 amplitudes were observed for complex sounds in the left hemisphere. ....	185
<i>Figure 2. Left panel:</i> Event-Related Spectral Perturbations as a function of Condition (categorized stimuli, uncategorized stimuli and silences) at Fz, Cz and Pz sites. <i>Right panel:</i> Activity maps (average of rectangles in the left panel) for each frequency band (theta, alpha, beta, early gamma and late gamma). ....	187
<i>Figure 3.</i> Maps of Event-Related Spectral Perturbations for categorized and uncategorized stimuli as a function of Stimulation (words, pseudo-words and complex sounds) of the beta frequency band. ....	188
Figure 34 : Continuum du traitement de l'information selon la tâche demandée (de bas ou haut niveau) et la mesure effectuée (objective ou subjective).....	210
Figure 35 : Résumé des différents effets observés au cours du travail doctoral. Chacun d'entre eux montre des interactions entre les processus perceptifs (détection, identification, catégorisation sémantique) et les connaissances (phonologiques et sémantiques).....	213
Figure 36 : Exemple d'audiogramme tonal enregistré chez un participant normo-entendant. Ne sont reportés que les seuils de perception (c'est-à-dire la plus faible intensité perçue pour chaque fréquence). Un audiogramme est réalisé pour chaque oreille. Chez un participant de moins de 30 ans, une audition normale dans le silence correspond à des seuils inférieurs à 20 dB entre 250 Hz et 8 000 Hz pour les deux oreilles. ....	215
Figure 37 : Pourcentage de détections correctes à la tâche de catégorisation : les mots sont plus souvent catégorisés comme "mot", "pseudo-mot" ou "son complexe" que les pseudo-mots ou les sons complexes. ....	217

# BIBLIOGRAPHIE

- Abrams, R. L., & Greenwald, A. G. (2000). Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychological Science, 11*(2), 118-124.
- Adams, P. A., & Adams, J. K. (1960). Confidence in the recognition and reproduction of words difficult to spell. *The American Journal of Psychology, 73*, 544-552.
- Adrian, E. D. (1942). Olfactory reactions in the brain of the hedgehog. *The Journal of Physiology - London, 100*(4), 459-473.
- Akhoun, I. (2008). *Analyse de l'encodage des signaux de parole dans le tronc cerebral* (Thèse de Doctorat). Lyon: Université Claude Bernard Lyon 1.
- American National Standard Institute. (2004). *ANSI S3.6-2004, Specifications for audiometers*. New-York.
- Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: an event-related brain potential study. *Psychophysiology, 32*(2), 177-190.
- ANSI. (1995). *S3.7-1995 (R2003), Methods for coupler calibration of earphones*. New-York: American National Standard Institute.
- ANSI. (2004). *S3.21-2004, Methods for Manual Pure-Tone Threshold Audiometry*. New-York: American National Standard Institute.
- Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *NeuroImage, 22*(1), 401-408. doi:10.1016/j.neuroimage.2004.01.014
- Avan, P., & Bonfils, P. (1992). Analysis of possible interactions of an attentional task with cochlear micromechanics. *Hearing Research, 57*(2), 269-275.
- Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267. doi:10.1037/0096-1523.19.2.250
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636. doi:10.1037/0022-3514.70.3.614
- Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology, 4*(2), 207-228. doi:10.1016/0010-0285(73)90012-1
- Bastiaansen, M. C. M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research, 159*, 179-196. doi:10.1016/S0079-6123(06)59012-0
- Bastuji, H., Perrin, F., & García-Larrea, L. (2002). Semantic analysis of auditory input during sleep: studies with event related potentials. *International Journal of Psychophysiology, 46*(3), 243-255.
- Beattie, R. C. (1986). Normal Intelligibility Functions for the Auditec Cid W-22 Test At 30% and 60% Time-Compression. *Otology & Neurotology, 7*(1), 60-64.
- Berger, H. (1930). Electroencephalogram of humans. *Journal Fur Psychologie Und Neurologie, 40*, 160-179.

- Bertrand, O., & Tallon-Baudry, C. (2000). Oscillatory gamma activity in humans: a possible role for object representation. *International Journal of Psychophysiology*, 38(3), 211-223.
- Bibbig, A., Traub, R. D., & Whittington, M. A. (2002). Long-range synchronization of gamma and beta oscillations and the plasticity of excitatory and inhibitory synapses: a network model. *Journal of Neurophysiology*, 88(4), 1634-1654.
- Bidet-Caulet, A. (2007). *Mécanismes neurophysiologiques de la perception de flux sonores chez l'Homme*: Effets des contextes acoustiques et attentionnels (Thèse de Doctorat). Lyon: Université Claude Bernard Lyon 1.
- Bidet-Caulet, A., & Bertrand, O. (2005). Dynamics of a Temporo-Fronto-Parietal Network during Sustained Spatial or Spectral Auditory Processing. *Journal of Cognitive Neuroscience*, 17(11), 1691-1703. doi:10.1162/089892905774589244
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10(5), 512-528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17(1), 353-362.
- Bleeck, S., Ives, T., & Patterson, R. D. (2004). Aim-mat: The Auditory Image Model in MATLAB. *Acta Acustica united with Acustica*, 90(4), 781-787.
- Bonnet, C., & Chancier, N. (1999). Bases physiologiques des traitements sensoriels et moteurs. Dans R. Ghiglione & J.-F. Richard (Éd.), *Cours de psychologie. 3. Champs et théories* (3<sup>e</sup> éd., p. 185-207). Paris: Dunod.
- Bonte, M., Valente, G., & Formisano, E. (2009). Dynamic and task-dependent encoding of speech and voice by phase reorganization of cortical oscillations. *The Journal of Neuroscience*, 29(6), 1699-1706. doi:10.1523/JNEUROSCI.3694-08.2009
- Bowman, H., Schlaghecken, F., & Eimer, M. (2006). A neural network model of inhibitory processes in subliminal priming. *Visual Cognition*, 13(4), 401-480. doi:10.1080/13506280444000823
- Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research*, 9(3), 227-238.
- Burgund, E. D., Guo, Y., & Aurbach, E. L. (2009). Priming for letters and pseudoletters in mid-fusiform cortex: examining letter selectivity and case invariance. *Experimental Brain Research*, 193(4), 591-601. doi:10.1007/s00221-008-1661-9
- Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychological Bulletin*, 135(3), 452-477. doi:10.1037/a0015329
- Buus, S., Müsch, H., & Florentine, M. (1998). On loudness at threshold. *The Journal of the Acoustical Society of America*, 104(1), 399-410.
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford, UK: Oxford University Press.
- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304(5679), 1926-1929. doi:10.1126/science.1099745

- Caton, R. M. D. (1875). Electrical Currents of the Brain. *The Journal of Nervous and Mental Disease*, 2(4), 610.
- Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 11, 63-65.
- Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, 36(4), 387-395.
- Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, 40(4), 343-367.
- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, 25(5), 975-979. doi:10.1121/1.1907229
- Chittka, L., & Brockmann, A. (2005). Perception Space—The Final Frontier. *PLoS Biology*, 3(4), e137. doi:10.1371/journal.pbio.0030137
- Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 521-533.
- Coch, D., & Mitra, P. (2010). Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Research*, 1329, 159-174. doi:10.1016/j.brainres.2010.02.084
- Collet, L. (2005). *Asymétrie de l'appareil auditif périphérique*. Cours de DEA de Neuropsychologie, Université Claude Bernard Lyon 1.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428. doi:10.1037/0033-295X.82.6.407
- Combescure, P. (1981). Vingt listes de dix phrases phonétiquement équilibrées. *Revue d'Acoustique*, 56, 34-38.
- Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the Beginnings of Spoken Words Have a Special Status in Auditory Word Recognition? *Journal of Memory and Language*, 32(2), 193-210. doi:10.1006/jmla.1993.1011
- Connine, C. M., Titone, D., Deelman, T., & Blasko, D. (1997). Similarity Mapping in Spoken Word Recognition. *Journal of Memory and Language*, 37(4), 463-480. doi:10.1006/jmla.1997.2535
- Crick, F., & Koch, C. (1990). Some reflections on visual awareness. *Cold Spring Harbor Symposia on Quantitative Biology*, 55, 953-962.
- Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, 6(2), 119-126. doi:10.1038/nn0203-119
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113-121. doi:10.1037/0096-1523.14.1.113
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380(6574), 499-505. doi:10.1038/380499a0
- Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience*, 23(8), 3423-3431.
- Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing Research*, 229(1-2), 132-147. doi:10.1016/j.heares.2007.01.014

- Debener, S., Herrmann, C. S., Kranczioch, C., Gembris, D., & Engel, A. K. (2003). Top-down attentional processing enhances auditory evoked gamma band activity. *Neuroreport*, *14*(5), 683-686. doi:10.1097/01.wnr.0000064987.96259.5c
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*(1-2), 1-37.
- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204-211. doi:10.1016/j.tics.2006.03.007
- Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F., et al. (1998). Imaging unconscious semantic priming. *Nature*, *395*(6702), 597-600. doi:10.1038/26967
- Delattre, P. (1958). Les indices acoustiques de la parole: Premier rapport. *Phonetica*, *2*(3-4), 226-251.
- Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition*, *73*(1), B1-B15.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21. doi:10.1016/j.jneumeth.2003.10.009
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115* ( Pt 6), 1753-1768.
- Desmedt, J. E. (1980). P300 in serial tasks: an essential post-decision closure mechanism. *Progress in Brain Research*, *54*, 682-686.
- Deutsch, A., & Bentin, S. (1994). Attention mechanisms mediate the syntactic priming effect in auditory word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 595-607.
- Doyle, J., & Leach, C. (1988). Word superiority in signal detection: Barely a glimpse, yet reading nonetheless. *Cognitive Psychology*, *20*(3), 283-318. doi:10.1016/0010-0285(88)90001-1
- Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology. General*, *127*(3), 286-303.
- Dupoux, E. (1993). The time course of prelexical processing: The syllabic hypothesis revisited. Dans G. T. M. Altmann & R. Shillcock (Éd.), *Cognitive models of speech processing : the Second Sperlonga Meeting* (p. 81-114). Hillsdale, NJ: Psychology Press, L. Erlbaum.
- Edwards, E., Soltani, M., Kim, W., Dalal, S. S., Nagarajan, S. S., Berger, M. S., & Knight, R. T. (2009). Comparison of Time-Frequency Responses and the Event-Related Potential to Auditory Speech Stimuli in Human Cortex. *Journal of Neurophysiology*, *102*(1), 377-386. doi:10.1152/jn.90954.2008
- Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1737-1747.
- Eimer, M., & Schlaghecken, F. (2002). Links between conscious awareness and response inhibition: evidence from masked priming. *Psychonomic Bulletin & Review*, *9*(3), 514-520.



- Ellis, A. W. (1984). Introduction to Byrom Bramwell's (1897) case of word meaning deafness. *Cognitive Neuropsychology*, *1*(3), 245. doi:10.1080/02643298408252025
- Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology*. Hillsdale, NJ: L. Erlbaum Associates.
- Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5*(1), 16-25.
- Engel, A. K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and Cognition*, *8*(2), 128-151. doi:10.1006/ccog.1999.0389
- Enns, J. T. (2002). Visual binding in the standing wave illusion. *Psychonomic Bulletin & Review*, *9*(3), 489-496.
- Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *33*(4), 325-350. doi:10.1080/14640748108400797
- Fell, J., Hinrichs, H., & Röschke, J. (1997). Time course of human 40 Hz EEG activity accompanying P3 responses in an auditory oddball paradigm. *Neuroscience Letters*, *235*(3), 121-124.
- Ferrand, L. (2001). *Cognition et lecture*. Paris: De Boeck Université.
- Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *45*(3), 353-372. doi:10.1080/02724989208250619
- Fisk, G. D., & Haase, S. J. (2005). Unconscious perception or not? An evaluation of detection and discrimination as indicators of awareness. *The American Journal of Psychology*, *118*(2), 183-212.
- Fletcher, H., & Munson, W. A. (1933). Loudness, Its Definition, Measurement and Calculation. *The Journal of the Acoustical Society of America*, *5*(2), 82-108. doi:10.1121/1.1915637
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627-635. doi:10.1016/S0022-5371(73)80042-8
- Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680-698. doi:10.1037/0278-7393.10.4.680
- Fournier, J. E. (1951). *Audiométrie vocale*. Paris: Maloine.
- Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: an experimental framework. *Trends in Cognitive Sciences*, *3*(3), 105-114.
- Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L., et al. (2009). Converging intracranial markers of conscious access. *PLoS Biology*, *7*(3), e61. doi:10.1371/journal.pbio.1000061
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 110-125.

- Geschwind, N. (1965). Disconnexion syndromes in animals and man. I. *Brain*, 88(2), 237-294.
- Giard, M. H., Collet, L., Bouchet, P., & Pernier, J. (1994). Auditory selective attention in the human cochlea. *Brain Research*, 633(1-2), 353-356.
- Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., & Kleinschmidt, A. (2004). Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cerebral Cortex*, 14(3), 247-255.
- Glasberg, B. R., & Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5), 331-342.
- Glasberg, B. R., Moore, B. C., & Bacon, S. P. (1987). Gap detection and masking in hearing-impaired and normal-hearing subjects. *The Journal of the Acoustical Society of America*, 81(5), 1546-1556.
- Dan Glaser, E. S., & Scherer, K. R. (2008). Neuronal processes involved in subjective feeling emergence: oscillatory activity during an emotional monitoring task. *Brain Topography*, 20(4), 224-231. doi:10.1007/s10548-008-0048-3
- Gow, D. W., Segawa, J. A., Ahlfors, S. P., & Lin, F.-H. (2008). Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates. *NeuroImage*, 43(3), 614-623. doi:10.1016/j.neuroimage.2008.07.027
- Grabowski, T. J., & Damasio, A. R. (2000). Investigating Language with Functional Neuroimaging. Dans A. W. Toga & J. C. Mazziotta (Éd.), *Brain Mapping: The Systems* (p. 425-461). San Diego: Academic Press.
- Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, 23(1), 1. doi:10.1080/01690960701578013
- Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1158-1176. doi:10.1037/0096-1523.20.6.1158
- Grainger, J., Bouttevin, S., Truc, C., Bastien, M., & Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: a comparison of dyslexic and normal readers. *Brain and Language*, 87(3), 432-440.
- Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: the neighborhood frequency effect. *Perception & Psychophysics*, 45(3), 189-195.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, 8(2), 145-146. doi:10.1038/nn1392
- Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C. M., & Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topography*, 14(2), 131-137.
- Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington N.Y.: R.E. Krieger Pub. Co.

- Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, 273(5282), 1699-1702.
- Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible (« subliminal ») stimuli: dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, 124(1), 22-42.
- Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3(8), 837-843. doi:10.1038/77754
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, 28(4), 267-283.
- Grosjean, F., & Frauenfelder, U. H. (1997). *A guide to spoken word recognition paradigms*. Hove, UK: Psychology Press.
- Grosjean, F., & Gee, J. P. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25(1-2), 135-155.
- Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., & Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 101(35), 13050-13055. doi:10.1073/pnas.0404944101
- Haase, S. J., & Fisk, G. D. (2001). Confidence in word detection predicts word identification: implications for an unconscious perception paradigm. *The American Journal of Psychology*, 114(3), 439-468.
- Haase, S. J., & Fisk, G. D. (2004). Valid distinctions between conscious and unconscious perception? *Perception & Psychophysics*, 66(5), 868-871; discussion 888-895.
- Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, 49(3-4), 277-290.
- He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *The Journal of Neuroscience*, 29(24), 7718-8822. doi:10.1523/JNEUROSCI.0157-09.2009
- Herrmann, C. S., & Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience and Biobehavioral Reviews*, 25(6), 465-476.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, 182(108), 177-180.
- Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, 172(990), 1357-1360.
- Hoen, M., Meunier, F., Grataloup, C., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., et al. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Communication*, 49(12), 905-916. doi:10.1016/j.specom.2007.05.008
- 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(4), 281-312.  
doi:10.1080/01690969008407065
- Holender, D. (1986a). Conceptual, experimental, and theoretical indeterminacies in research on semantic activation without conscious identification. *Behavioral and Brain Sciences*, 9(01), 50. doi:10.1017/S0140525X00021543
- Holender, D. (1986b). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, 9(01), 1. doi:10.1017/S0140525X00021269
- Holender, D., & Duscherer, K. (2004). Unconscious perception: the need for a paradigm shift. *Perception & Psychophysics*, 66(5), 872-881; discussion 888-895.
- Howard, M. A., Volkov, I. O., Mirsky, R., Garell, P. C., Noh, M. D., Granner, M., Damasio, H., et al. (2000). Auditory cortex on the human posterior superior temporal gyrus. *The Journal of Comparative Neurology*, 416(1), 79-92.
- Howell, D. (1998). *Méthodes statistiques en sciences humaines*. Paris: De Boeck Université.
- Huffman, R. F., & Henson, O. W. (1990). The descending auditory pathway and acousticomotor systems: connections with the inferior colliculus. *Brain Research Reviews*, 15(3), 295-323.
- Humphreys, G. W. (1983). Reference frames and shape perception. *Cognitive Psychology*, 15(2), 151-196.
- Humphreys, G. W., & Quinlan, P. T. (1988). Priming effects between two-dimensional shapes. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 203-220.
- Humphries, C., Liebenthal, E., & Binder, J. R. (2010). Tonotopic organization of human auditory cortex. *NeuroImage*, 50(3), 1202-1211. doi:10.1016/j.neuroimage.2010.01.046
- Ives, D. T., & Patterson, R. D. (2008). Pitch strength decreases as F0 and harmonic resolution increase in complex tones composed exclusively of high harmonics. *The Journal of the Acoustical Society of America*, 123(5), 2670-2679. doi:10.1121/1.2890737
- Johnsrude, I. S., Zatorre, R. J., Milner, B. A., & Evans, A. C. (1997). Left-hemisphere specialization for the processing of acoustic transients. *Neuroreport*, 8(7), 1761-1765.
- Kaas, J. H., & Hackett, T. A. (1999). « What » and « where » processing in auditory cortex. *Nature Neuroscience*, 2(12), 1045-1047. doi:10.1038/15967
- Kahan, T. A. (2000). Negative priming from masked words: retrospective prime clarification of center-surround inhibition? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1392-1410.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kewley-Port, D. (1991). Detection thresholds for isolated vowels. *The Journal of the Acoustical Society of America*, 89(2), 820-829.
- Khalfa, S. (1999). *Caractérisation des asymétries fonctionnelles de l'appareil auditif périphérique chez l'humain : relations avec la spécialisation hémisphérique ?* (Thèse de Doctorat). Université Claude Bernard Lyon 1.

- Khalifa, S., Bougeard, R., Morand, N., Veuillet, E., Isnard, J., Guenot, M., Ryvlin, P., et al. (2001). Evidence of peripheral auditory activity modulation by the auditory cortex in humans. *Neuroscience*, *104*(2), 347-358.
- Kiebel, S. J., von Kriegstein, K., Daunizeau, J., & Friston, K. J. (2009). Recognizing sequences of sequences. *PLoS Computational Biology*, *5*(8), e1000464. doi:10.1371/journal.pcbi.1000464
- Kiefer, M., & Brendel, D. (2006). Attentional modulation of unconscious « automatic » processes: evidence from event-related potentials in a masked priming paradigm. *Journal of Cognitive Neuroscience*, *18*(2), 184-198. doi:10.1162/089892906775783688
- Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: unconscious inhibition influences reaction time and response selection. *Journal of Experimental Psychology: General*, *131*(2), 255-269.
- Klimesch, W., Doppelmayr, M., Russegger, H., & Pachinger, T. (1996). Theta band power in the human scalp EEG and the encoding of new information. *Neuroreport*, *7*(7), 1235-1240.
- Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters*, *244*(2), 73-76.
- Klimesch, W., Hanslmayr, S., Sauseng, P., Gruber, W., Brozinsky, C. J., Kroll, N. E. A., Yonelinas, A. P., et al. (2006). Oscillatory EEG correlates of episodic trace decay. *Cerebral Cortex*, *16*(2), 280-290. doi:10.1093/cercor/bhi107
- Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63-88. doi:10.1016/j.brainresrev.2006.06.003
- Kopell, N., Ermentrout, G. B., Whittington, M. A., & Traub, R. D. (2000). Gamma rhythms and beta rhythms have different synchronization properties. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(4), 1867-1872.
- Kotchoubey, B., & Lang, S. (2001). Event-related potentials in an auditory semantic oddball task in humans. *Neuroscience Letters*, *310*(2-3), 93-96. doi:10.1016/S0304-3940(01)02057-2
- Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London: Series B-Biological Sciences*, *362*(1481), 857-875. doi:10.1098/rstb.2007.2093
- Kouider, S., & Dupoux, E. (2004). Partial awareness creates the « illusion » of subliminal semantic priming. *Psychological Science*, *15*(2), 75-81.
- Kouider, S., & Dupoux, E. (2005). Subliminal speech priming. *Psychological Science*, *16*(8), 617-625. doi:10.1111/j.1467-9280.2005.01584.x
- Kouider, S., de Gardelle, V., Dehaene, S., Dupoux, E., & Pallier, C. (2010). Cerebral bases of subliminal speech priming. *NeuroImage*, *49*(1), 922-929. doi:10.1016/j.neuroimage.2009.08.043
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141-178. doi:10.1016/j.cogpsych.2005.05.001

- Krause, C. M., Grönholm, P., Leinonen, A., Laine, M., Säkkinen, A.-L., & Söderholm, C. (2006). Modality matters: The effects of stimulus modality on the 4- to 30-Hz brain electric oscillations during a lexical decision task. *Brain Research, 1110*(1), 182-192. doi:10.1016/j.brainres.2006.06.100
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language, 110*(3), 135-148. doi:10.1016/j.bandl.2009.03.005
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25*(1), 161-168. doi:10.1016/j.cogbrainres.2005.05.004
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463-470. doi:10.1016/S1364-6613(00)01560-6
- Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*(2), 99-116.
- Lafon, J.-C. (1967). Auditory integration of verbal message. *Journal of the Neurological Sciences, 5*(2), 197.
- Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation, 11*(2), 417-441.
- Leibniz, G. W. (1886). *Nouveaux essais sur l'entendement humain*. C. Delagrave.
- Leinonen, A., Grönholm-Nyman, P., Järvenpää, M., Söderholm, C., Lappi, O., Laine, M., & Krause, C. M. (2009). Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: evidence from a morphologically rich language. *Brain Research, 1275*, 54-66. doi:10.1016/j.brainres.2009.03.057
- Liégeois-Chauvel, C., Laguitton, V., Badier, J. M., Schwartz, D., & Chauvel, P. (1995). Étude des mécanismes corticaux de la perception auditive chez l'homme: apport des potentiels intracérébraux et des champs magnétiques évoqués par les stimulations auditives. *Revue Neurologique, 151*(8-9), 495-504.
- Liégeois-Chauvel, C., Musolino, A., & Chauvel, P. (1991). Localization of the primary auditory area in man. *Brain, 114A*(1), 139-153.
- Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92*(3), 204-214.
- van Linden, S., Stekelenburg, J. J., Tuomainen, J., & Vroomen, J. (2007). Lexical effects on auditory speech perception: an electrophysiological study. *Neuroscience Letters, 420*(1), 49-52. doi:10.1016/j.neulet.2007.04.006
- Lindner, W. A. (1968). Recognition performance as a function of detection criterion in a simultaneous detection-recognition task. *The Journal of the Acoustical Society of America, 44*(1), 204-211.
- Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General, 133*(4), 475-493. doi:10.1037/0096-3445.133.4.475

- Luce, R. D. (1959). *Individual choice behavior*. Oxford, UK: Wiley.
- Luce, R. D., & Green, D. M. (1978). Two tests of a neural attention hypothesis for auditory psychophysics. *Perception & Psychophysics*, 23(5), 363-371.
- Luo, H., Husain, F. T., Horwitz, B., & Poeppel, D. (2005). Discrimination and categorization of speech and non-speech sounds in an MEG delayed-match-to-sample study. *NeuroImage*, 28(1), 59-71. doi:10.1016/j.neuroimage.2005.05.040
- Lutzenberger, W., Pulvermüller, F., & Birbaumer, N. (1994). Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans. *Neuroscience Letters*, 176(1), 115-118.
- Macmillan, N. A. (1986). The psychophysics of subliminal perception. *Behavioral and Brain Sciences*, 9(01), 38-39. doi:10.1017/S0140525X00021427
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: a user's guide* (2<sup>e</sup> éd.). Mahwah N.J.: Lawrence Erlbaum Associates.
- Marcel, A. J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, 15(2), 238-300. doi:10.1016/0010-0285(83)90010-5
- Maris, E. (2002). The role of orthographic and phonological codes in the word and the pseudoword superiority effect: an analysis by means of multinomial processing tree models. *Journal of Experimental Psychology: Human Perception and Performance*, 28(6), 1409-1431.
- Marmel, F. (2009). *Influence du contexte mélodique tonal sur la perception de sons musicaux : approches comportementales et neurophysiologiques* (Thèse de Doctorat). Lyon: Université Lumière Lyon 2.
- Marmel, F., Parbery-Clark, A., Skoe, E., & Kraus, N. (2009, octobre 16). *Abstract Musical Knowledge Influences Auditory Encoding In The Brainstem*. Poster présenté à Tucker-Davis Symposium on Advances and Perspectives in Auditory Neurophysiology (APAN VII), Chicago, IL.
- Marois, R., Yi, D.-J., & Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron*, 41(3), 465-472.
- Marslen-Wilson, W. D. (1984). Function and process in spoken word-recognition. Dans H. Bouma & D. G. Bouwhuis (Éd.), *Attention and performance X: control of language processes*. Hillsdale, NJ: L. Erlbaum Associates.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. Dans G. T. M. Altmann (Éd.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (p. 148-172). Cambridge, MA: MIT Press.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10(1), 29-63. doi:10.1016/0010-0285(78)90018-X
- Mason, M. (1982). Recognition time for letters and nonletters: effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance*, 8(5), 724-738.

- Mattler, U. (2007). Inverse target- and cue-priming effects of masked stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1), 83-102. doi:10.1037/0096-1523.33.1.83
- McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology. Human Perception and Performance*, 6(2), 265-276.
- McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 2(1), 80-91.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.
- McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics*, 22(3), 249-261.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375-407. doi:10.1037/0033-295X.88.5.375
- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10(8), 363-369. doi:10.1016/j.tics.2006.06.007
- McNamara, T. P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review*, 99(4), 650-662. doi:10.1037/0033-295X.99.4.650
- McNamara, T. P. (2005). *Semantic priming : perspectives from memory and word recognition*. New York: Psychology Press.
- Meador, K. J., Ray, P. G., Echaz, J. R., Loring, D. W., & Vachtsevanos, G. J. (2002). Gamma coherence and conscious perception. *Neurology*, 59(6), 847-854.
- Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *The Journal of Neuroscience*, 27(11), 2858-2865. doi:10.1523/JNEUROSCI.4623-06.2007
- Meric, C., & Collet, L. (1994). Attention and otoacoustic emissions: a review. *Neuroscience and Biobehavioral Reviews*, 18(2), 215-222.
- Merikle, P. M., & Cheesman, J. (1986). Consciousness is a « subjective » state. *Behavioral and Brain Sciences*, 9(01), 42. doi:10.1017/S0140525X00021452
- Merikle, P. M., & Reingold, E. M. (1990). Recognition and lexical decision without detection: Unconscious perception? *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 574-583. doi:10.1037/0096-1523.16.3.574
- Meunier, F. (1997). *Morphologie et Traitement du Langage Parlé* (Thèse de Doctorat). Paris: Université René Descartes Paris V.
- Micheyl, C., Carbonnel, O., & Collet, L. (1995). Medial Olivocochlear System and Loudness Adaptation - Differences Between Musicians and Non-Musicians. *Brain and Cognition*, 29(2), 127-136. doi:10.1006/brcg.1995.1272
- Miller, J. (2000). Measurement error in subliminal perception experiments: Simulation analyses of two regression methods. *Journal of Experimental*



- Psychology: Human Perception and Performance*, 26(4), 1461-1477.  
doi:10.1037/0096-1523.26.4.1461
- Milliken, B., Joordens, S., Merikle, P. M., & Seiffert, A. E. (1998). Selective attention: a reevaluation of the implications of negative priming. *Psychological Review*, 105(2), 203-229.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. New-York, N.Y.: Oxford University Press.
- Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5<sup>e</sup> éd.). Amsterdam: Academic Press.
- Moore, B. C. J., & Glasberg, B. R. (1987). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. *Hearing Research*, 28(2-3), 209-225.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11(1), 56.  
doi:10.1080/17470215908416289
- Moulin, A., Collet, L., & Duclaux, R. (1993). Contralateral auditory stimulation alters acoustic distortion products in humans. *Hearing Research*, 65(1-2), 193-210.
- Musiek, F. E., Froke, R., & Weihing, J. (2005). The auditory P300 at or near threshold. *Journal of the American Academy of Audiology*, 16(9), 698-707.
- Myers, E. B., & Blumstein, S. E. (2008). The neural bases of the lexical effect: an fMRI investigation. *Cerebral Cortex*, 18(2), 278-288.  
doi:10.1093/cercor/bhm053
- Näätänen, R. (1990). The role of attention in auditory information-processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13(2), 201-232.
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313-329.
- Naccache, L., & Dehaene, S. (2001). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, 11(10), 966-974.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106(3), 226-254. doi:10.1037/0096-3445.106.3.226
- Neely, J. H. (1991). Semantic priming effects in visual words recognition: a selective review of current findings and theories. Dans D. Besner & G. W. Humphreys (Éd.), *Basic processes in reading: visual word recognition* (p. 264-336). Hillsdale N.J.: L. Erlbaum Associates.
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28(04), 661-677.  
doi:10.1017/S014271640707035X
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behavior Research Methods, Instruments & Computers*, 36(3), 516-524.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE<sup>TM</sup> / A lexical database for

- contemporary french : LEXIQUE™. *L'Année Psychologique*, 101(3-4), 447-462.
- Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, 75(6), 522-536. doi:10.1037/h0026699
- Norris, D. (1994). A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1212-1232. doi:10.1037/0096-1523.20.6.1212
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23(3), 299-325.
- Nunez, P. L. (1989). Estimation of large scale neocortical source activity with EEG surface Laplacians. *Brain Topography*, 2(1-2), 141-154.
- Obusek, C. J., & Warren, R. M. (1973). A comparison of speech perception in senile and well-preserved aged by means of the verbal transformation effect. *Journal of Gerontology*, 28(2), 184-188.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Orfanidou, E., Marslen-Wilson, W. D., & Davis, M. H. (2006). Neural response suppression predicts repetition priming of spoken words and pseudowords. *Journal of Cognitive Neuroscience*, 18(8), 1237-1252. doi:10.1162/jocn.2006.18.8.1237
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785-806. doi:10.1016/0749-596X(92)90039-Z
- Owen, A. M., & Coleman, M. R. (2008). Using neuroimaging to detect awareness in disorders of consciousness. *Functional Neurology*, 23(4), 189-194.
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: the word-superiority effect. *Psychological Review*, 89(5), 573-594.
- Palva, S., Linkenkaer-Hansen, K., Näätänen, R., & Palva, J. M. (2005). Early neural correlates of conscious somatosensory perception. *The Journal of Neuroscience*, 25(21), 5248-5258. doi:10.1523/JNEUROSCI.0141-05.2005
- Palva, S., Palva, J. M., Shtyrov, Y., Kujala, T., Ilmoniemi, R. J., Kaila, K., & Näätänen, R. (2002). Distinct gamma-band evoked responses to speech and non-speech sounds in humans. *The Journal of Neuroscience*, 22(4), RC211(1-5).
- Pascual-Marqui, R. D., Lehmann, D., Koenig, T., Kochi, K., Merlo, M. C. G., Hell, D., & Koukkou, M. (1999). Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. *Psychiatry Research-Neuroimaging*, 90(3), 169-179. doi:10.1016/S0925-4927(99)00013-X
- Perrin, F. (2004). Auditory event-related potentials studies of information processing during human sleep. *Psychologica Belgica*, 44(1/2), 43-57.
- Perrin, F., & Grimault, N. (2005). Fonds Sonores (Version 1.0) [Sound samples]. Consulté avril 7, 2010, de <http://olfac.univ-lyon1.fr/unite/equipe-02/FondsSonores.html>

- Perrin, F., Bastuji, H., & García-Larrea, L. (2002). Detection of verbal discordances during sleep. *Neuroreport*, *13*(10), 1345-1349.
- Perrin, F., Schnakers, C., Schabus, M., Degueldre, C., Goldman, S., Brédart, S., Faymonville, M.-E., et al. (2006). Brain response to one's own name in vegetative state, minimally conscious state, and locked-in syndrome. *Archives of Neurology*, *63*(4), 562-569. doi:10.1001/archneur.63.4.562
- Perrot, X., Ryvlin, P., Isnard, J., Guénot, M., Catenoix, H., Fischer, C., Mauguière, F., et al. (2006). Evidence for corticofugal modulation of peripheral auditory activity in humans. *Cerebral Cortex*, *16*(7), 941-948. doi:10.1093/cercor/bhj035
- Pfurtscheller, G., & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalography and Clinical Neurophysiology*, *42*(6), 817-826.
- Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II. Effects of attention. *Electroencephalography and Clinical Neurophysiology*, *36*(2), 191-199.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, *37*(2), 127-152.
- Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography and Clinical Neurophysiology*, *36*, 179-190. doi:10.1016/0013-4694(74)90155-2
- Pilotti, M., & Beyer, T. (2002). Perceptual and lexical components of auditory repetition priming in young and older adults. *Memory & Cognition*, *30*(2), 226-236.
- Pins, D., & Ffytche, D. (2003). The neural correlates of conscious vision. *Cerebral Cortex*, *13*(5), 461-474.
- Pisoni, D. B. (1978). Speech perception. Dans W. K. Estes (Éd.), *Handbook of learning and cognitive processes* (Vol. 6, p. 167-233). Hillsdale, NJ: L. Erlbaum Associates.
- Pisoni, D. B., & Sawusch, J. R. (1975). Some Stages of Processing in Speech Perception. Dans A. Cohen & S. G. Nooteboom (Éd.), *Structure and process in speech perception* (p. 16-35). New-York, NY: Springer-Verlag.
- Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 699-725.
- Plack, C. J., Oxenham, A. J., Fay, R., & Popper, A. N. (Éd.). (2005). *Pitch: neural coding and perception*. Springer Handbook of Auditory Research. New York: Springer Science Business Media Inc.
- Polich, J., Ehlers, C. L., Otis, S., Mandell, A. J., & Bloom, F. E. (1986). P300 latency reflects the degree of cognitive decline in dementing illness. *Electroencephalography and Clinical Neurophysiology*, *63*(2), 138-144.
- Poulin-Charronnat, B., & Tillmann, B. (s. d.). The effect of word repetition is overruled by sentenceintegration.

- Pressnitzer, daniel, Patterson, R. D., & Krumbholz, K. (2001). The lower limit of melodic pitch. *The Journal of the Acoustical Society of America*, *109*(5 Pt 1), 2074-2084.
- Price, C. J. (2000). The anatomy of language: contributions from functional neuroimaging. *Journal of Anatomy*, *197 Pt 3*, 335-359.
- 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., Preissl, H., Lutzenberger, W., & Birbaumer, N. (1996). Brain rhythms of language: nouns versus verbs. *The European Journal of Neuroscience*, *8*(5), 937-941.
- Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2004). *Neurosciences*. (J.-M. Coquery, Trad.) (2<sup>e</sup> éd.). Paris: De Boeck Université.
- Rapp, D. N., & Samuel, A. G. (2002). A reason to rhyme: phonological and semantic influences on lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 564-571.
- Ratcliff, R., Allbritton, D., & McKoon, G. (1997). Bias in auditory priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(1), 143-152.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275-280.
- Reingold, E. M. (2004a). Unconscious perception and the classic dissociation paradigm: a new angle? *Perception & Psychophysics*, *66*(5), 882-887; discussion 888-895.
- Reingold, E. M. (2004b). Unconscious perception: assumptions and interpretive difficulties. *Consciousness and Cognition*, *13*(1), 117-122. doi:10.1016/j.concog.2003.11.002
- Reingold, E. M., & Jolicoeur, P. (1993). Perceptual versus postperceptual mediation of visual context effects: evidence from the letter-superiority effect. *Perception & Psychophysics*, *53*(2), 166-178.
- Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*(6), 563-575.
- Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., & Lang, J. M. (1994). On the perceptual organization of speech. *Psychological Review*, *101*(1), 129-156.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science (New York, N.Y.)*, *212*(4497), 947-949.
- Renault, B., Ragot, R., Lesevre, N., & Remond, A. (1982). Onset and offset of brain events as indices of mental chronometry. *Science*, *215*(4538), 1413-1415.
- Ritter, W., Simson, R., Vaughan, H. G., & Friedman, D. (1979). A brain event related to the making of a sensory discrimination. *Science*, *203*(4387), 1358-1361. doi:10.1126/science.424760
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature*, *397*(6718), 430-433. doi:10.1038/17120

- Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, *125*(1), 28-51. doi:10.1037/0096-3445.125.1.28
- Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, *32*(2), 97-127. doi:10.1006/cogp.1997.0646
- Samuel, A. G., & Pitt, M. A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language*, *48*(2), 416-434. doi:10.1016/S0749-596X(02)00514-4
- Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069-1078. doi:10.1016/j.concog.2009.12.013
- Schacter, D. L., & Church, B. A. (1992). Auditory priming: implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 915-930.
- Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. I. Review. *Psychologische Forschung*, *36*(1), 71-93.
- Schlaghecken, F., & Eimer, M. (2006). Active masks and active inhibition: a comment on Lleras and Enns (2004) and on Verleger, Jaskowski, Aydemir, van der Lubbe, and Groen (2004). *Journal of Experimental Psychology: General*, *135*(3), 484-494. doi:10.1037/0096-3445.135.3.484
- van Selst, M., & Merikle, P. M. (1993). Perception below the Objective Threshold? *Consciousness and Cognition*, *2*(3), 194-203. doi:10.1006/ccog.1993.1018
- Shahin, A. J., Picton, T. W., & Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain and Cognition*, *70*(3), 259-266. doi:10.1016/j.bandc.2009.02.008
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.
- Shipley, E. (1965). Detection and recognition: Experiments and choice models. *Journal of Mathematical Psychology*, *2*(2), 277-311. doi:10.1016/0022-2496(65)90006-4
- Shipley, E. F. (1961). *Detection and Recognition with Uncertainty* (PhD Thesis). University of Pennsylvania.
- Sidis, B. (1898). The psychology of suggestion. *Science*, *8*(188), 162-163. doi:10.1126/science.8.188.162
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, *18*, 555-586. doi:10.1146/annurev.ne.18.030195.003011
- Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 64-75.
- Snodgrass, M., Bernat, E., & Shevrin, H. (2004a). Unconscious perception at the objective detection threshold exists. *Perception & Psychophysics*, *66*(5), 888-895.
- Snodgrass, M., Bernat, E., & Shevrin, H. (2004b). Unconscious perception: A model-based approach to method and evidence. *Perception & Psychophysics*, *66*(5), 846-867.

- Snodgrass, M., Shevrin, H., & Kopka, M. (1993). The Mediation of Intentional Judgments by Unconscious Perceptions: The Influences of Task Strategy, Task Preference, Word Meaning, and Motivation. *Consciousness and Cognition*, 2(3), 169-193. doi:10.1006/ccog.1993.1017
- Snoeren, N. D., Hallé, P. A., & Segui, J. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics*, 34(2), 241-268. doi:10.1016/j.wocn.2005.06.001
- Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S., Brewer, C. C., et al. (1999). Language dominance in neurologically normal and epilepsy subjects: a functional MRI study. *Brain*, 122 ( Pt 11), 2033-2046.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643-662. doi:10.1037/h0054651
- Summerfield, C., Jack, A. I., & Burgess, A. P. (2002). Induced gamma activity is associated with conscious awareness of pattern masked nouns. *International Journal of Psychophysiology*, 44(2), 93-100.
- Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, 150(3700), 1187-1188. doi:10.1126/science.150.3700.1187
- Suzuki, T., Yamamoto, K., Taguchi, K., & Sakabe, N. (1976). Reliability and validity of late vertex-evoked response audiometry. *Audiology: Official Organ of the International Society of Audiology*, 15(5), 357-369.
- Taft, M., & Hambly, G. (1986). Exploring the Cohort Model of spoken word recognition. *Cognition*, 22(3), 259-282.
- Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, 3(4), 151-162.
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience*, 16(13), 4240-4249.
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory gamma -Band (30-70 Hz) Activity Induced by a Visual Search Task in Humans. *The Journal of Neuroscience*, 17(2), 722-734.
- Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature*, 364(6432), 59-60. doi:10.1038/364059a0
- Tillmann, B., Hoch, L., & Marmel, F. (sous presse). Influence du contexte sur le traitement en musique et en langage. Dans J. Morais & R. Kolinsky (Éd.), *Musique, Langage, Émotion*. Rennes: Presses Universitaires de Rennes.
- Tononi, G., Srinivasan, R., Russell, D. P., & Edelman, G. M. (1998). Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proceedings of the National Academy of Sciences of the United States of America*, 95(6), 3198-3203.
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12(4), 242. doi:10.1080/17470216008416732
- Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, 20, 12-16.

- Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 480-498. doi:10.1037/a0013027
- Urban, M. J. (1992). Auditory subliminal stimulation: a re-examination. *Perceptual and Motor Skills*, *74*(2), 515-541.
- Le Van Quyen, M., & Bragin, A. (2007). Analysis of dynamic brain oscillations: methodological advances. *Trends in Neurosciences*, *30*(7), 365-373. doi:10.1016/j.tins.2007.05.006
- Vanhauzenhuysse, A., Laureys, S., & Perrin, F. (2008). Cognitive event-related potentials in comatose and post-comatose states. *Neurocritical Care*, *8*(2), 262-270. doi:10.1007/s12028-007-9016-0
- Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(4), 229-239. doi:10.1038/35067550
- Vaughan, H. G., & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalography and Clinical Neurophysiology*, *28*(4), 360-367.
- Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.-F., & Pernier, J. (1995). Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis. *Electroencephalography and Clinical Neurophysiology*, *96*(2), 143-156. doi:10.1016/0168-5597(94)00242-7
- Verleger, R. (1988). Event-Related Potentials and Cognition: A Critique of the Context Updating Hypothesis and an Alternative Interpretation of P3. *Behavioral and Brain Sciences*, *11*(03), 343-356. doi:10.1017/S0140525X00058015
- Verleger, R., Jaśkowski, P., Aydemir, A., van der Lubbe, R. H. J., & Groen, M. (2004). Qualitative differences between conscious and nonconscious processing? On inverse priming induced by masked arrows. *Journal of Experimental Psychology: General*, *133*(4), 494-515. doi:10.1037/0096-3445.133.4.494
- Veuille, E., Collet, L., & Duclaux, R. (1991). Effect of contralateral acoustic stimulation on active cochlear micromechanical properties in human subjects: dependence on stimulus variables. *Journal of Neurophysiology*, *65*(3), 724-735.
- Vicary, J. (1957). Subliminal Svengali? *Sponsor*, *42*.
- Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience*, *13*(7), 994-1005. doi:10.1162/089892901753165890
- Warr, W. B., Guinan, J. J. J., & White, J. S. (1986). Organization of the efferent fibers: the lateral and medial olivocochlear systems. Dans R. A. Altschuler, R. P. Bobbin, & D. W. Hoffman (Éd.), *Neurobiology of hearing: the cochlea* (p. 333-348). New-York, N.Y.: Raven Press.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*(917), 392-393.
- Warren, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin*, *96*(2), 371-383.

- Warren, R. M., & Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, *9*, 358-362.
- Warren, R. M., & Sherman, G. L. (1974). Phonemic restorations based on subsequent context. *Perception & Psychophysics*, *16*(1), 150-156.
- Warren, R. M., & Warren, R. P. (1970). Auditory illusions and confusions. *Scientific American*, *223*(6), 30-36.
- Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain: A Journal of Neurology*, *97*(4), 709-728.
- Whaley, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 143-154. doi:10.1016/S0022-5371(78)90110-X
- Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*(1), 59-85. doi:10.1016/0010-0285(70)90005-8
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420-422. doi:10.1038/nn1872
- Wyart, V., & Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. *The Journal of Neuroscience*, *28*(10), 2667-2679. doi:10.1523/JNEUROSCI.4748-07.2008
- Yamagishi, N., Callan, D. E., Goda, N., Anderson, S. J., Yoshida, Y., & Kawato, M. (2003). Attentional modulation of oscillatory activity in human visual cortex. *NeuroImage*, *20*(1), 98-113.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*(10), 946-953.
- Zatorre, R. J., & Binder, J. R. (2000). Functional and Structural Imaging of the Human Auditory System. Dans A. W. Toga & J. C. Mazziotta (Éd.), *Brain Mapping: The Systems* (p. 365-402). San Diego: Academic Press.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, *6*(1), 37-46.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*(5058), 846-849.
- Ziegler, J. C., Besson, M., Jacobs, A. M., Nazir, T. A., & Carr, T. H. (1997). Word, Pseudoword, and Nonword Processing: A Multitask Comparison Using Event-Related Brain Potentials. *Journal of Cognitive Neuroscience*, *9*(6), 758-775. doi:10.1162/jocn.1997.9.6.758