

**Université Paris 5 - René Descartes
UFR DE PSYCHOLOGIE**

N° attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE PARIS 5

Discipline : Psychologie Cognitive

Présentée et soutenue publiquement par

Natalie Dominique SNOEREN

**VARIATIONS PHONOLOGIQUES EN PRODUCTION
ET PERCEPTION DE LA PAROLE :
LE PHENOMENE DE L'ASSIMILATION**

Sous la direction de

Juan SEGUI

DECEMBRE 2005

Membres du Jury

M. Ulrich FRAUENFELDER
M. Gareth GASKELL
M. Christian LORENZI
Mme Monique RADEAU
M. Juan SEGUI

Variations phonologiques en production et perception de la parole : le phénomène de l'assimilation en français

Résumé Ce travail de thèse porte sur une variation phonologique fréquemment observée en français, l'assimilation de voisement. La première série d'expériences a eu pour objectif d'étudier de façon approfondie la production des mots assimilés et de caractériser les propriétés acoustico-phonétiques des obstruents en position finale ayant subi une assimilation. Nous avons proposé un nouvel indice acoustique, doté d'une validité perceptive, qui permet de mesurer le degré d'assimilation entre deux mots. Les mesures acoustiques ont montré que l'assimilation est souvent un processus gradué plutôt que dichotomique. Qui plus est, le degré d'assimilation est modulé en fonction du voisement sous-jacent. Les obstruents non voisés subissent plus fréquemment des assimilations complètes que les obstruents voisés. Cette propriété phonétique a été opérationnalisée dans une deuxième série d'expériences d'amorçage intermodal de répétition qui avait pour but d'étudier le traitement perceptif des mots plus ou moins fortement assimilés. Les résultats de ces expériences ont montré que le rôle du contexte droit, responsable de l'assimilation, varie en fonction du degré d'assimilation. Le traitement perceptif des mots complètement assimilés est facilité en présence du contexte droit, alors que la présence des traces acoustiques suffirait pour accéder à la forme sous-jacente des mots partiellement assimilés. Enfin, l'hypothèse de la présence des traces acoustiques dans les mots assimilés était confirmée dans une troisième série d'expériences en amorçage sémantique. Les résultats ont montré que la présentation auditive des mots assimilés donne lieu à l'activation du réseau lexico-sémantique de leurs formes sous-jacentes. Dans l'ensemble, ces travaux s'inscrivent dans un cadre à la fois méthodologique et théorique selon lequel la production et la perception de la parole sont envisagées comme correspondant à deux versants du même système de traitement.

Mots clefs : variations phonologiques, assimilation de voisement, production et perception de la parole.

Phonological variations in speech production and perception : The phenomenon of assimilation in French.

Abstract The present PhD thesis provides an in-depth study of a phonological variation that is frequently encountered in French, namely voice assimilation. The goal of the first series of experiments was to study the production of assimilated words and to provide a acoustico-phonetic description of word-final assimilated obstruents. A perceptually relevant acoustic cue was proposed that measures degrees of assimilation. Acoustic measurements showed that voice assimilation is often a graded, rather than a categorical phonetic process. Moreover, degrees of assimilation varied as a function of underlying voicing. Voiceless obstruents underwent more frequently complete assimilation than voiced obstruents. This phonetic property was used in the second series of experiments to study perceptual processing of completely and partially assimilated words. The results showed that the role of the right context, responsible for assimilation, varies as a function of the degree of assimilation. It was found that perceptual processing of completely assimilated segments was facilitated in the presence of the right context, whereas the presence of « acoustic traces » sufficed to access partially assimilated segments. The hypothesis of the presence of acoustic traces in assimilated words was confirmed in the third series of experiments using semantic priming. The results of these experiments revealed that auditory presentation of assimilated words does not affect the activation of the lexico-semantic network of their underlying forms. To sum up, the present study corresponds to a methodological and theoretical framework according to which speech production and perception are considered as two parts of the same underlying processing system.

Key words: phonological variations, voice assimilation, speech perception, speech production.

Discipline : Psychologie Cognitive (Psycholinguistique)
UFR de Psychologie, Université Paris 5 - René Descartes
71, Avenue Edouard Vaillant
92774 Boulogne-Billancourt

REMERCIEMENTS

J'ai l'honneur d'avoir eu comme Directeur de thèse Juan Segui, dont la rigueur scientifique et les commentaires précieux n'ont fait qu'affûter l'argumentation de mes propos. Je voudrais ici lui exprimer ma plus profonde reconnaissance pour sa grande disponibilité, son aide et savoir-faire. Cette thèse n'aurait jamais vu le jour sans ses encouragements et sa confiance en moi.

Je tiens à remercier Ludovic Ferrand, Christian Lorenzi, Nicole Bacri, Pierre Hallé, Willy Serniclaes, Shigeko Shinohara, Catherine-Marie Longtin, Mirjam Ernestus, Gareth Gaskell, Celine Chang et Boris New pour leurs encouragements, commentaires et questions qui m'ont permis d'améliorer ce travail. Je dois aussi beaucoup à Christophe Pallier et Harald Baayen pour leur aide en analyses statistiques.

Je sais gré à Ulrich Frauenfelder, Gareth Gaskell, Christian Lorenzi et Monique Radeau, d'avoir accepté d'être membre de mon jury.

Durant ma thèse, j'ai consacré une partie non négligeable de mon temps à l'enseignement en Psychologie Expérimentale (niveau Licence). Je voudrais ainsi profiter de l'occasion pour remercier chaleureusement les enseignants en Psychologie Expérimentale à Paris 5 avec lesquels j'ai pu travailler dans d'excellentes conditions : Pierre Marquer, Agnès Charvillat, Véronique Dubois, Karine Doré-Mazars, Dorine Vergilino-Perez, Jean-Didier Bagot et Boris New. Un merci très particulier doit être adressé à tou(te)s les étudiant(e)s, fréquemment recruté(e)s dans mes propres cours, d'avoir participé à mes expériences.

Je tiens également à adresser mes remerciements aux secrétaires et ingénieurs-techniciens du LPE et de l'Ecole Doctorale à l'Institut de Psychologie (Paris 5) pour leur gentillesse, aide et disponibilité : Françoise Levesque, Jasmine Drouillac, Flore Nguentat, Jean-Jacques Boislaroussie, Sylvain Hauptert, Franck Bourry et Laurent Savin.

Un grand merci à Véronique, Catherine-Marie et Géraud pour les corrections de dernière minute et d'avoir allégé ce texte de nombreuses coquilles.

Je remercie vivement Véronique, Catherine-Marie, Katia, Alicia, Marlies, Thérèse, Vicky, Malika, Marie, Fred, Ludivine, Paul, David et Ed d'avoir partagé de nombreux moments agréables au sein du LPE, les déjeuners animés, et toutes sortes d'activités de *procrastination* (quelqu'un a enfin trouvé la traduction en français ?).

Lisa, to quote a small note I found on my desk a while ago ;-)
you are magic. Thank you for being such a lovely friend. Hélène, j'étais si contente que tu m'aies côtoyé quotidiennement, (enfin, à part les matinées ☺). Merci beaucoup pour ton amitié et ton soutien moral, toujours en douceur. Karen, I think it's about time we hit that bottle of nice and peaty malt to drink to research and friendship, don't you agree ? Florian, nous savons bien que parfois c'est difficile de faire de la recherche tout en encaissant les sacrifices qui vont avec... *Augen zu und durch*, merci pour tes conseils et ton amitié ! Enfin, je remercie Marythé et Christian Bayada pour m'avoir fait connaître l'hospitalité du sud de la France.

Avec une franchise toute hollandaise :

Mams, jij bent nog steeds mijn allergrootste bron van inspiratie. Van jou heb ik de werkelijke betekenis van het woord « doorzettingsvermogen » geleerd, en daarmee veel kunnen bereiken. Nadja, dank je wel voor al je begrip, luisterend oor en aanmoedigingen. Het was immens fijn mijn *PhD Blues* met jou te kunnen delen (en te kunnen spelen, nietwaar Sjoerd ?) Alexander, heel veel dank voor je steun en interesse. We hadden het onlangs nog over het bereiken van doelen. Ik heb met dit proefschrift inmiddels een doel bereikt, maar vergeet niet dat *de weg ernaar toe* het allerbelangrijkst was ! Heel veel dank gaat tenslotte uit naar Ton Dijkstra, met wie mijn grote taalpsychologie-avontuur inmiddels een aantal jaar geleden is begonnen aan het NICI in Nijmegen, voor zijn steun en toeverlaat.

Le sourire de talisman (souviens-toi) a dorénavant pris la forme d'une certaine bague...

Ma dette la plus grande, je l'ai envers Denis. Ta présence et ton soutien tout au long de ces années a été d'une valeur inestimable pour moi et je t'en remercie de tout cœur.

*Ce n'est pas avec des « représentations » ou avec une pensée
que je communique d'abord, mais avec un sujet parlant,
avec un certain style d'être et avec le « monde » qu'il vise.*

Maurice Merleau-Ponty, *Phénoménologie de la Perception* (1945).

TABLE DES MATIERES

TABLE DES MATIERES.....	6
INTRODUCTION.....	10
CHAPITRE I. BASES THEORIQUES DE L'ASSIMILATION ET ETUDE PSYCHOLINGUISTIQUE	12
1. DESCRIPTIONS LINGUISTIQUES DE L'ASSIMILATION.....	12
1.1. <i>Définitions de l'assimilation</i>	13
1.1.1. Phonologies linéaire et non linéaire.....	16
1.1.2. L'assimilation selon la phonologie articulatoire□recouvrement des gestes.....	18
1.1.3. Données en phonétique expérimentale.....	20
1.1.4. Psycholinguistique de la parole assimilée□Production et Perception.....	21
2. MODELES PSYCHOLINGUISTIQUES ET RECHERCHES EMPIRIQUES SUR LES VARIATIONS PHONOLOGIQUES.	23
2.1. <i>Modèles théoriques en psycholinguistique</i>	23
2.1.1. Le rôle de la variabilité et problèmes de modélisation.....	25
2.2. <i>Etude psycholinguistique des variations phonologiques</i>	26
2.2.1. Théorie de la sous-spécification et le modèle FUL.....	28
3. L'ASSIMILATION□CADRES THEORIQUES ET RECHERCHES EXPERIMENTALES.....	29
3.1. <i>L'Hypothèse du Mécanisme d'Inférence Régressive</i>	30
3.2. <i>L'Hypothèse des Indices Acoustiques d'Assimilation et le Mécanisme Perceptif d'Alignement des Traits</i>	31
3.3. <i>Conclusion</i>	33
CHAPITRE II. LA PRODUCTION DES CONSONNES OBSTRUENTES ASSIMILEES :	
PARAMETRES ACOUSTICO-PHONETIQUES ET FACTEURS LEXICAUX.....	35
1. INTRODUCTION.....	35
1.1. <i>Paramètres acoustiques du voisement</i>	36
1.2. <i>Le rôle des facteurs lexicaux et pragmatiques sur les variations phonologiques</i>	38
1.2.1. L'influence de la fréquence des mots.....	39
1.2.2. L'influence des facteurs pragmatiques.....	39
1.2.3. Le rôle du voisinage phonologique.....	40
1.3. <i>La présente recherche</i>	41
2. EXPERIENCE 1A□PRODUCTION DE MOTS EN CONTEXTE ASSIMILATOIRE ET NON ASSIMILATOIRE.....	42
2.1. <i>Matériel</i>	42
2.2. <i>Locuteurs</i>	43
2.3. <i>Procédure</i>	43
2.4. <i>Analyses acoustiques</i>	44
2.5. <i>Résultats acoustiques</i>	46
3. EXPERIENCE 1B□CATEGORISATION PERCEPTIVE DES MOTS ASSIMILES ET NON ASSIMILES.....	50
3.1. <i>Matériel expérimental</i>	50
3.2. <i>Sujets</i>	51

3.3.	<i>Procédure</i>	51
3.4.	<i>Résultats</i>	52
4.	EXPERIENCE 2 □ PRODUCTION DE MOTS EN CONTEXTE ASSIMILATOIRE	54
4.1.	<i>Matériel expérimental</i>	54
4.2.	<i>Locuteurs</i>	54
4.3.	<i>Procédure</i>	54
4.4.	<i>Résultats acoustiques</i>	55
4.5.	<i>Discussion des Expériences 1-2</i>	55
5.	EXPERIENCE 3A □ PRODUCTION DE MOTS DANS DES CONTEXTES ASSIMILATOIRES ET NON ASSIMILATOIRES □ ROLE DE L' AMBIGUITE ET DU VOISINAGE PHONOLOGIQUE	58
5.1.	<i>Matériel expérimental</i>	59
5.2.	<i>Locuteurs</i>	60
5.3.	<i>Procédure</i>	60
5.4.	<i>Résultats acoustiques</i>	61
6.	EXPERIENCE 3B □ CATEGORISATION PHONETIQUE DE MOTS ASSIMILES.....	64
6.1.	<i>Matériel expérimental</i>	64
6.2.	<i>Sujets</i>	65
6.3.	<i>Procédure</i>	65
6.4.	<i>Résultats</i>	66
6.5.	<i>Discussion des Expériences 3a-b</i>	66
7.	DISCUSSION GENERALE	67
CHAPITRE III. LE TRAITEMENT PERCEPTIF DES SEGMENTS PLUS OU MOINS FORTEMENT ASSIMILES : TRACES ACOUSTIQUES ET ROLE DU CONTEXTE DROIT		73
1.	INTRODUCTION	73
1.1.	<i>Données empiriques en faveur de l'Hypothèse de l'Inférence Régressive</i>	73
1.2.	<i>Données empiriques en faveur de l'Hypothèse des Indices Acoustiques d'Assimilation</i>	76
1.3.	<i>Hypothèses alternatives ou complémentaires</i> □.....	78
1.4.	<i>La présente recherche</i>	79
2.	EXPERIENCE 4 □ CATEGORISATION DES PHONEMES ASSIMILES	82
2.1.	<i>Matériel expérimental</i>	82
2.2.	<i>Sujets</i>	83
2.3.	<i>Procédure</i>	83
2.4.	<i>Résultats et Discussion</i>	84
3.	EXPERIENCE 5 □ AMORÇAGE SANS CONTEXTE DROIT	85
3.1.	<i>Matériel expérimental</i>	85
3.2.	<i>Sujets</i>	85
3.3.	<i>Procédure</i>	86
3.4.	<i>Résultats</i>	86
3.4.1.	Analyse sur les temps de réponse.....	88
3.4.2.	Analyse sur les erreurs.....	89

3.5.	<i>Discussion</i>	89
4.	EXPERIENCE 6 □ AMORÇAGE AVEC CONTEXTE DROIT	91
4.1.	<i>Matériel expérimental</i>	91
4.2.	<i>Sujets</i>	91
4.3.	<i>Procédure</i>	91
4.4.	<i>Résultats</i>	92
4.4.1.	Analyses sur les temps de réponse	93
4.4.2.	Analyses sur les erreurs	94
4.4.3.	Analyse combinée Expériences 5 et 6.....	94
4.5.	<i>Discussion</i>	96
5.	DISCUSSION GENERALE	98
CHAPITRE IV. REPRESENTATION LEXICALE DES SEGMENTS ASSIMILES ET ACTIVATION DU RESEAU LEXICO-SEMANTIQUE.....		105
1.	INTRODUCTION	105
1.1.	<i>Activation lexicale et proximité phonologique</i>	105
1.2.	<i>Assimilation et ambiguïté lexicale</i>	111
1.3.	<i>La présente recherche</i>	112
2.	EXPERIENCE 7 □ AMORÇAGE DES FORMES CANONIQUES ET ASSIMILEES	114
2.1.	<i>Matériel</i>	114
2.1.1.	Pré-test □ Technique d'association libre	116
2.2.	<i>Sujets</i>	116
2.3.	<i>Procédure</i>	117
2.4.	<i>Résultats</i>	117
2.4.1.	Analyse sur les temps de réponse.....	118
2.5.	<i>Discussion</i>	119
3.	EXPERIENCE 8 □ AMORÇAGE DES FORMES CANONIQUES DERIVEES	121
3.1.	<i>Matériel</i>	121
3.2.	<i>Procédure</i>	121
3.3.	<i>Sujets</i>	121
3.4.	<i>Résultats & Discussion</i>	121
4.	DISCUSSION GENERALE	122
CHAPITRE V. DISCUSSION GENERALE.....		128
1.	RESUME DES RESULTATS PRINCIPAUX	128
2.	IMPLICATIONS DES RESULTATS POUR LES THEORIES PSYCHOLINGUISTIQUES	131
3.	LIGNES DIRECTIVES POUR DE FUTURES RECHERCHES	134
3.1.	<i>La modélisation des variations phonologiques</i>	134
3.1.1.	Modélisation de l'assimilation en anglais et en français	135
3.1.2.	Modélisation d'autres types de variations phonologiques.....	135
3.2.	<i>Assimilation et morphologie</i>	136
3.3.	<i>Vers une perspective d'intégration de la production et perception de la parole</i>	137

CONCLUSION.....	140
ANNEXE 1. MATERIEL EXPERIMENTAL DES EXPERIENCES 1-3.....	151
ANNEXE 2. PRE-TEST JUGEMENTS SUBJECTIFS (EXPERIENCE 3).....	155
ANNEXE 3. CONSIGNE PRODUCTION DE LA PAROLE (EXPERIENCES 2-3A).....	156
ANNEXE 4. CONSIGNE POUR LA CATEGORISATION PHONETIQUE (EXPERIENCE 1B).....	157
ANNEXE 5. CONSIGNE CATEGORISATION DES PHONEMES (EXPERIENCE 3B-4).....	158
ANNEXE 6. MATERIEL EXPERIMENTAL DES EXPERIENCES 4-6.....	159
ANNEXE 7. CONSIGNE DECISION LEXICALE AVEC AMORÇAGE DE REPETITION (EXPERIENCES 5 ET 6).....	162
ANNEXE 8. POST-TEST DE RECONNAISSANCE (EXPERIENCES 5-6).....	163
ANNEXE 9. PRE-TEST D'ASSOCIATION LIBRE (EXPERIENCES 7 ET 8).....	165
ANNEXE 10. MATERIEL EXPERIMENTAL (EXPERIENCES 7 ET 8).....	166
ANNEXE 11. CONSIGNE DECISION LEXICALE AVEC AMORÇAGE SEMANTIQUE (EXPERIENCES 7 ET 8).....	171
ANNEXE 12. QUESTIONNAIRE SUR LES CONNAISSANCES LINGUISTIQUES.....	172
ANNEXE 13. ANALYSE SUPPLEMENTAIRE DE REGRESSION MULTIPLE (LME).....	175

INTRODUCTION

Tout travail empirique en psycholinguistique s'appuie sur des observations faites à partir d'expériences. Ces expériences sont conduites dans un boîtier expérimental dans lequel se trouve le sujet, assis devant un écran d'ordinateur. Généralement, le sujet doit faire des décisions lexicales sur un matériel linguistique comportant, la plupart du temps, des mots isolés ou encore des phrases. S'il est indéniable que ce genre d'expérimentation a fait progresser énormément nos connaissances sur le comportement linguistique humain, il n'est pas anodin de se rappeler que le langage, apanage de l'homme, est avant tout un moyen de communication entre les personnes dans un monde où la *parole est déjà instituée*, comme disait Merleau-Ponty. De par le fait que l'homme est à la fois auditeur et locuteur des sons de parole, il convient de prendre en compte l'ensemble des processus mis en œuvre dans la production et la perception de la parole. Tel est le cap que nous nous sommes fixés dans la présente recherche.

L'un des principaux défis pour le développement des systèmes automatiques de traitement de la parole durant les dernières décennies, a été le problème de la variabilité. Un mot n'est jamais prononcé deux fois de la même façon et la parole humaine contient de nombreuses variations, qu'elles soient de nature accidentelle ou systématique. Si l'élaboration du traitement automatique s'est heurtée au problème de la variabilité dans le signal de parole, l'homme ne semble pas pour autant gêné dans sa compréhension du langage en présence de cette variabilité. En effet, le système cognitif dispose de cette remarquable capacité de traiter les variations afin que l'on puisse comprendre les sons de parole malgré leurs variations de réalisation. Les mécanismes cognitifs mis en jeu lors de telles procédures restent encore mal compris, d'où l'intérêt de conduire des études psycholinguistiques sur les variations systématiques.

L'objet de la présente recherche est d'étudier en profondeur une variation systématique fréquemment observée en langue française, l'assimilation de voisement. Dans le chapitre 1, nous présenterons d'abord une description linguistique et psycholinguistique nécessaire à la compréhension du phénomène de l'assimilation en français. Nous donnerons également une revue de la littérature, en insistant tout particulièrement sur les études psycholinguistiques,

essentiellement conduites en anglais, portant sur les conséquences perceptives de l'assimilation de lieu. Le champ expérimental de cette thèse est organisé en trois chapitres. Le chapitre 2 a pour objectif de fournir une description acoustico-phonétique appropriée du phénomène d'assimilation en français afin de déterminer s'il s'agit de processus phonétique discret ou bien gradué. Dans le chapitre 3, nous nous intéressons au rôle du contexte droit dans le traitement perceptif des segments assimilés en prenant en compte des caractéristiques phonétiques de l'assimilation en français, étudiées au chapitre 2. Le chapitre 4 vise à approfondir l'étude du traitement perceptif de la parole assimilée, en examinant l'activation du réseau lexico-sémantique des formes assimilées. Enfin, dans le chapitre 5, nous donnerons un résumé de nos principaux résultats, et discuterons de leurs implications pour l'étude psycholinguistique des variations phonologiques de réalisation des segments de parole.

CHAPITRE I.

BASES THEORIQUES DE L'ASSIMILATION ET ETUDE PSYCHOLINGUISTIQUE

1. DESCRIPTIONS LINGUISTIQUES DE L'ASSIMILATION

La perception auditive de la parole dépend à la fois de la clarté du signal acoustique et de notre capacité cognitive à traiter ce signal. Or, les mécanismes cognitifs mis en jeu pour déchiffrer le signal de parole restent encore mal compris. Ceci est majoritairement dû à deux caractéristiques qui sont intrinsèquement liées à la parole, à savoir la continuité et la variabilité. En effet, contrairement à l'écriture, le signal de parole est le plus souvent continu, en particulier à l'intérieur d'une unité prosodique qui peut regrouper plusieurs mots. Les mots ne sont donc pas des entités facilement isolables, car il n'existe pas d'indices acoustiques clairs qui marquent leurs frontières. Cette question, connue comme la problématique de la segmentation lexicale, a été largement abordée dans les travaux psycholinguistiques ayant pour but de déterminer l'unité de traitement perceptif permettant aux auditeurs de segmenter les mots (e.g., Mehler, Dommergues, Frauenfelder, & Segui, 1981 ; Cutler, Mehler, Norris, & Segui, 1983 ; McQueen, Norris, & Cutler, 1994 ; Gow & Gordon, 1995 ; Gaskell, Spinelli, & Meunier, 2003; Banel & Bacri, 1997 ; Isel & Bacri, 1999 ; Mattys, 2004).

La variabilité de réalisation des sons constitue une autre particularité importante de la parole. La parole peut se manifester de façon très variable pour diverses raisons. Parmi les variations possibles, il faut distinguer celles qui résultent des erreurs de production, ou encore celles liées aux différences entre locuteurs, comme l'âge, le sexe, l'accent etc. Ce type de variations, dit « arbitraire », n'est pas lié aux principes de l'organisation phonologique de la langue. Les variations de type « systématique » sont en revanche motivées par les règles phonologiques spécifiques à la langue en question. L'étude de ces variations est intéressante dans la mesure où elle peut nous renseigner sur les façons dont l'auditeur a accès à son lexique mental afin de reconnaître un mot malgré ses variations de réalisation.

L'objectif de cette thèse est d'examiner la production et la perception d'une variation phonologique fréquemment observée dans la langue française : l'assimilation de voisement régressive qui se produit *entre* les mots.

L'assimilation peut avoir lieu lorsque deux consonnes de voisement différent sont en contact. La première consonne adopte le voisement de la seconde (assimilation régressive). En français, les occlusives sonores (/b/, /d/, /g/) deviennent sourdes quand elles sont suivies par des consonnes sourdes (/p/, /t/, /k/) et *vice versa*. Par exemple, le mot français avec pour consonne finale une occlusive non voisée *soute* est prononcé *soude* [sud] quand il est suivi par un mot ayant une consonne initiale voisée (e.g., *la soute ne sent pas si mauvais*).

Un objectif important en psycholinguistique est de rendre compte des procédures d'accès au lexique au cours desquelles l'entrée sensorielle est comparée à la représentation lexicale stockée dans le lexique mental. Or, tout segment transformé par un processus phonologique peut provoquer une ambiguïté phonologique : l'identité sous-jacente de ce segment ne pourra pas être dérivée directement des propriétés phonétiques présentes dans la forme de surface du signal. La question est alors de savoir si ces variations de surface systématiques perturbent ou non l'accès au lexique. Autrement dit, quelles sont les conséquences perceptives des variations présentes dans le signal de parole ? Afin de pouvoir répondre à cette question, nous pensons qu'il faut tout d'abord s'interroger sur les propriétés linguistiques et acoustiques des segments assimilés, d'où l'intérêt d'une description adéquate de l'assimilation, fournie par la phonologie et la phonétique expérimentale.

1.1. Définitions de l'assimilation

Généralement, les phonologues définissent l'assimilation comme le transfert d'une caractéristique – ou trait phonétique – d'un son sur un son voisin. Quand deux sons de parole sont en contact, l'un communique à l'autre un de ses traits, partiellement ou totalement. Deux sons contigus tendent ainsi à acquérir un ou plusieurs traits communs (Calliope, 1989 ; Carton, 1974). On entend ici par trait la propriété (articulatoire, acoustique, perceptive) la plus basique qui permet de distinguer entre les phonèmes de la langue. Nous ferons référence principalement aux traits de nature articulatoire.

L'assimilation de lieu d'articulation, fréquemment étudiée en anglais et en allemand, est définie comme une forme de neutralisation de contraste dans la mesure où elle réduit le nombre potentiel de porteurs d'indices de deux traits à un trait dans une séquence de consonnes (Steriade, 2000). En anglais, l'assimilation de place se produit quand une consonne coronale finale (e.g., /d/, /t/, /n/) change en prenant le lieu d'articulation du segment suivant. Par exemple, le mot *wicked* en anglais peut être réalisé comme [wikib] dans l'énoncé *a wicked pranck* ; la consonne coronale /d/ de l'adjectif *wicked* peut se labialiser en /b/

puisqu'elle est suivie par la consonne labiale /p/ de *pranck* (cf. Gaskell & Marslen-Wilson, 1996). Les linguistes postulent que l'assimilation a lieu toutes les fois où un son se rapproche d'un autre son, quant à son point ou mode d'articulation, jusqu'à manifester avec lui certains traits de ressemblance pouvant éventuellement aboutir à une complète identité (Thomas, Bouquiaux, & Cloarec-Heiss, 1976).

L'assimilation de voisement existe dans un grand nombre de langues et correspond à la propagation du trait de voisement, qui peut être [+voisé] ou [-voisé], à un segment adjacent. Cette transformation peut impliquer un *voisement* aussi bien qu'un *dévoisement* du segment d'origine. L'assimilation peut concerner des suites consonantiques : elle peut toucher un groupe de deux occlusives qui diffèrent par leur voisement, séparées ou non par une frontière de morphème ou de mot (pour une typologie exhaustive des différentes formes que peut prendre l'assimilation, voir Wetzels & Mascaro, 2001).

En ce qui concerne l'assimilation de voisement en français, les obstruents sonores (qui incluent les consonnes occlusives et fricatives) peuvent devenir sourdes quand elles sont suivies par les obstruents sourdes et *vice versa*¹. L'assimilation peut se produire au sein des mots ou bien entre les mots. Il a été suggéré par les linguistes que cette particularité peut avoir des conséquences sur le degré de réalisation de l'assimilation. A titre illustratif, Rigault (1967) a observé qu'une assimilation complète de voisement s'observe au sein d'un mot lorsque les deux consonnes sont en contact direct, comme c'est le cas pour *médecin* prononcé [mɛt̥sɛ̃] ou la coronale /d/ s'est dévoisée au contact de la fricative non voisée /s/².

¹ Les nasales (/m/,/n/) et les liquides (/l/,/r/) ne provoquent pas d'assimilation en français, mais les liquides peuvent parfois subir une assimilation.

² Les linguistes ne s'accordent pas sur la question de savoir si, dans le cas de *médecin* prononcé [mɛt̥sɛ̃], il s'agit vraiment d'une assimilation. D'un point de vue synchronique, des arguments ont été mis en avant pour dire que la forme lexicale de *médecin* est stockée sous forme de /mɛt̥sɛ̃/ avec la coronale non voisée, n'ayant subi aucune assimilation de voisement. En revanche, au niveau morphophonémique, *médecin* comporte la coronale voisée /d/, étant donné que les autres membres de la même famille morphologique (e.g., *médicament* ou *médical*), se prononcent avec un /d/. Ainsi, l'assimilation au sein des mots pourrait se produire entre plusieurs niveaux de représentation, par le biais d'une règle de transformation qui détermine, selon les cas, les prononciations alternatives entre les /t/ et les /d/ au sein d'une famille morphologique. Il est également à noter que la voyelle orale /e/ précédant la coronale dans *médecin* est prononcée plus fréquemment /ɛ/ que /e/, alors qu'elle devrait plutôt être prononcée /ɛ/ quand elle se trouve en syllabe initiale fermée /mɛd/ dans /mɛd. sɛ̃/.

L'assimilation partielle aurait lieu quand le contact entre deux consonnes de différente sonorité résulte de la chute d'un schwa à l'intérieur d'un mot ou bien de la rencontre fortuite de deux mots différents. Rigault (1967) a étudié l'assimilation de voisement de la coronale voisée /d/ sous l'influence de la fricative sonore /s/. Il a trouvé que la consonne /d/ sous forme assimilée présentait les mêmes caractéristiques acoustiques et perceptives que la consonne /t/ dans la séquence /t+s/ mais différait radicalement de la consonne /d/ dans la séquence /d+z/. Il a conclu que la distinction entre 'fortis' vs. 'lenis', telle que proposée par d'autres (cf. Grammont, 1933 ; Van Dommelen, 1984), ainsi que la distinction entre [tense] vs. [lax], (cf. Jessen, 2001) n'a pas lieu d'être. Il n'y avait aucune différence acoustico-perceptive entre le /t/ sous forme assimilée de la consonne sourde /d/ dans le mot *médecin* et le /t/ dans *hauteserve*.

En français, l'assimilation de voisement est dans la plupart des cas de nature régressive, puisque dans une séquence de deux consonnes occlusives de voisement différent, c'est la seconde qui propage sa valeur de voisement sur la première. C'est par exemple le cas pour *cheval* et *cheveux* qui sont prononcés [ʒvaɫ] et [ʒvø] en Savoie et en Suisse romande (Thomas et al., 1976). Cependant il existe -selon les locuteurs- des cas d'assimilation progressive (cf. *cheval* et *cheveux* fréquemment prononcés [ʃfaɫ] et [ʃfø] en Ile-de-France). Ces auteurs font également remarquer qu'il serait plus juste de parler d'assimilation prolongative et anticipative au lieu d'assimilation progressive ou régressive, dans la mesure où un son peut être réalisé par anticipation de la mise en place des organes pour la réalisation tout en maintenant la position des autres organes. Notons enfin qu'il existe également des assimilations doubles, un son étant alors assimilé par deux sons identiques qui l'entourent. Ainsi dans l'exemple de *maintenant* [mɛ̃tnã], le phonème /t/ peut être nasalisé, étant entouré par une voyelle et une consonne nasales, ce qui résulte en [mɛ̃nnã].

L'assimilation a suscité l'intérêt des linguistes et psycholinguistes, car elle nécessite la mise en relation entre, d'une part, les règles phonologiques qui sont discrètes par définition et, d'autre part, la manifestation phonétique qui est très souvent variable et continue. Le défi consiste précisément à proposer une théorie qui réconcilie les descriptions linguistiques et physiques de la parole (Lindblom, 1989). Dans les sections suivantes, nous discuterons

Ceci pourrait être considéré comme une extrapolation au niveau morphophonémique dans la mesure où *médicament* et *médical* sont prononcés avec /e/ plutôt que /ɛ/.

d'abord des approches en phonologie et en phonétique expérimentale qui ont cherché à rendre compte des variations phonologiques avant d'examiner les différentes études en psycholinguistique.

1.1.1. *Phonologies linéaire et non linéaire*

Toute règle phonologique peut être considérée comme une expression formelle ayant pour but de décrire les changements dans les représentations phonologiques des mots (Gussenhoven & Jacobs, 1998). En 1968, Chomsky et Halle ont publié une œuvre d'une importance considérable pour les théories phonologiques, *The Sound Pattern of English* (désormais SPE), dans laquelle ils proposent la notation formalisée des règles phonologiques. Dans la formalisation proposée par Chomsky et Halle (1968), les représentations phonologiques sont définies à partir de matrices contenant des « traits » à valeur binaire. Les traits peuvent avoir des valeurs [+] ou [-] pour spécifier les caractéristiques des sons (par exemple, la consonne nasale /n/ peut avoir la matrice suivante : [-syl, +cons, +son, +voice, -spread, -constr, -cont, +nas, -lat, +ant, -distr]. Ainsi, la valeur [+] indique la présence d'une caractéristique, la valeur [-] indique son absence. Dans l'annotation de SPE, le trait de voisement [voice] est de nature binaire ([± voice]), étant donné qu'il ne peut prendre que deux valeurs. (Certains autres traits sont soit présents, soit absents dans la représentation phonologique, et sont dénommés *univalents* par les phonologues, par exemple le trait spécifiant la place labiale [lab].) Le traitement de la parole doit faire face à une complexité énorme dans le sens où le répertoire phonémique est estimé entre dix et 100 sons à partir d'un ensemble de seulement 12 traits distinctifs selon les langues (Jakobson & Halle, 1956 ; Jakobson, Fant & Halle, 1963). Or, nous savons par ailleurs que le lexique mental peut contenir entre 10000-100000 mots (selon la définition de ce qu'est un mot). L'espace lexico-phonologique est en conséquence très dense, car il contient beaucoup de mots qui partagent les mêmes sons, tout en désignant des choses différentes (e.g., la présence des homophones comme *sein* vs. *saint*).

La linéarité constitue l'une des caractéristiques principales de la formalisation SPE. Cependant, certains phonologues ont noté que cette sorte de formalisation pose un certain nombre de problèmes pour rendre compte des variations phonologiques qui peuvent se produire dans une langue (cf. Gussenhoven & Jacobs, 1998). A titre illustratif, considérons la règle de voisement régressif en néerlandais, illustrée dans (1), qui stipule que les consonnes occlusives deviennent des consonnes voisées quand elles sont suivies par [b,d]. Selon

l'annotation SPE, il est alors concevable de stipuler une seconde règle qui reflète pourtant un processus phonologique non plausible, comme le montre l'exemple illustré dans (2) :

(1) *voisement régressif* [-son] □ [+voice] / __ (##) [-son, -cont, +voice]

(2) *règle non plausible* [+son] □ [-voice] / __ (##) [-son, -cont, +voice]

La règle 2 qui découle de l'annotation SPE donne lieu à un processus phonologique en néerlandais en vertu duquel les segments se dévoisent quand ils sont suivis par des obstruents voisés ! Cet exemple illustre bien que ce type de règles *linéaires* n'est pas toujours capable de fournir une description réaliste des changements phonologiques.

Dans la théorie de la Géométrie des Traits (*Feature Geometry*, cf. Clements, 1985 ; Clements & Hume, 1995 ; Clements, 1993), les phonologues ont tenté de proposer une solution à ce problème. Dans cette théorie, les traits ne sont plus organisés à un seul niveau comme dans SPE, mais sont organisés de façon hiérarchique dans une structure arborescente de traits (*feature tree*). Dans cette structure, les traits et leurs groupements sont représentés dans les nœuds. Un nœud peut faire partie d'une ou de plusieurs structures arborescentes. Cette façon de représenter les traits semble particulièrement bien appropriée pour rendre compte des assimilations par rapport aux théories linéaires comme SPE, qui considèrent toute modification phonologique comme un processus de remplacement d'un trait par un autre.

Toutefois, il serait très complexe si chaque règle phonologique devait être illustrée à l'aide des structures arborescentes tri-dimensionnelles. Afin de rendre la description plus parcimonieuse, Clements (1985) a proposé une adaptation qui consiste à séparer deux aspects qui sont impliqués dans le changement phonologique : (a) l'identification d'un changement structurel (*Structural Change*, SC) et (b) la description structurelle (*Structural Description*, SD). Le changement structurel stipule quel nœud est propagé vers le nœud approprié dans le segment adjacent. Afin d'exprimer l'assimilation, il faut préciser s'il s'agit soit d'une association de droite-à-gauche (i.e. assimilation régressive), soit d'une association de gauche-à-droite (i.e. assimilation progressive). La description structurelle indique quelles sont les propriétés phonologiques nécessaires que doit avoir tout segment susceptible de subir une modification. La figure 1.1 montre un exemple de la règle d'assimilation coronale en anglais. La règle stipule que le nœud coronal (avec tous les traits qui lui appartiennent) soit propagé à un nœud coronal non continuant à gauche. Il est à noter que la théorie de la géométrie des traits a contribué à l'élaboration d'autres théories en phonologie, notamment les théories en phonologie articulatoire et la théorie de la sous-spécification, dont nous parlerons ultérieurement.

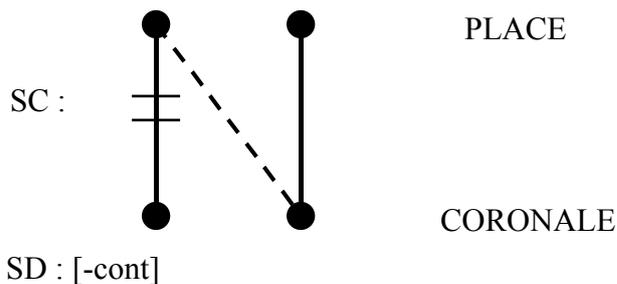


Figure 1.1. Règle d'assimilation coronale en anglais : « propagation du nœud coronal (avec tous les traits qu'il domine) à une coronale non continue se trouvant à gauche » (d'après Gussenhoven & Jacobs, 1998).

1.1.2. *L'assimilation selon la phonologie articulatoire : recouvrement des gestes*

Suite à la phonologie non linéaire (cf. Clements, 1985), un nouveau programme d'étude de phonologie articulatoire a été élaboré par les chercheurs aux laboratoires Haskins à la fin des années 1980 (cf. Browman & Goldstein 1990a, 1990b ; voir aussi Goldstein & Fowler, 2003 pour un aperçu récent de la théorie). La phonologie articulatoire avait pour but principal de combiner les descriptions linguistiques (issues de la phonologie non linéaire) et les descriptions physiques de la parole. D'après cette approche, la structure phonologique n'est pas constituée d'un ensemble de règles comme c'est le cas dans les théories classiques en phonologie, mais formulée à partir d'une interaction dynamique entre les dimensions articulatoires, acoustiques, et perceptives. Sans pour autant négliger l'importance d'autres dimensions, ce courant en phonologie se focalise avant tout sur la dimension articulatoire puisque beaucoup de phénomènes phonologiques peuvent être expliqués en prenant en compte les propriétés articulatoires des sons.

En phonologie articulatoire, toute structure linguistique est définie en termes de mouvements articulatoires en coordination, appelés « gestes »³. Les gestes sont relativement abstraits, de nature discrète, et ils sont composés des variables pertinentes de l'appareil vocal. A titre

³ Il est à noter que le terme de *geste* a été employé par le philosophe français Merleau-Ponty dans le chapitre traitant de la parole dans *Phénoménologie de la Perception* (1945, pp. 214) : *La parole est un véritable geste et elle contient son sens comme le geste contient le sien. [...] La parole est un geste et sa signification un monde.* Si la notion de geste n'est ici pas identique à celle employée en phonologie articulatoire, nous sommes d'avis que, tout au moins sur le plan théorique, ce courant théorique est compatible avec le cadre phénoménologique, auquel nous reviendrons dans la Discussion Générale.

d'exemple, le geste de la fermeture des lèvres est composé de trois variables de l'appareil vocal : (1) le mâchoire, (2) la lèvre inférieure et (3) la lèvre supérieure. Deux combinaisons des gestes peuvent donner lieu à des contrastes potentiels, si elles contiennent au moins un geste de constriction distinct. Par exemple, les mots anglais *hill*, *sill* et *pill* sont en contraste les uns avec les autres par la combinaison des organes de constriction au début des mots. *Hill* commence avec le geste du larynx, *sill* ajoute le geste de pointe de la langue au geste du larynx, alors que *pill* ajoute le geste de la lèvre. Ces trois gestes (larynx, pointe de la langue, et lèvres) sont combinés dans l'articulation du mot *spill* (cf. Goldstein & Fowler, 2003).

Les trajectoires articulatoires correspondant aux gestes sont générées par une description dynamique à un niveau supplémentaire d'abstraction. Ainsi, tout mouvement articulatoire, de nature graduée et continue, est analysé sous la forme d'un certain nombre de gestes discrets. Si les gestes sont par leur nature des unités discrètes et invariantes dans les différents contextes, ils sont dotés des propriétés liées à l'espace et au temps. Par conséquent, il peut y avoir un recouvrement entre plusieurs gestes dans l'espace temporel. C'est précisément ce recouvrement entre gestes qui rend compte d'un certain nombre de variations phonologiques, telles que la suppression des segments, la réduction des schwas, ou encore l'assimilation. A titre d'exemple, considérons la séquence en anglais qui contient une assimilation de place donnée dans (3) :

(3) /hundred pounds/ [hʌndrəd' paʊndz] → [hʌndrəb' paʊndz]

En parole naturelle, la consonne coronale /d/ peut prendre la place d'articulation labiale, induite par la consonne labiale /p/ dans *pounds*. D'après la phonologie articulatoire, la suppression (tout au moins perceptive) de la consonne est due à un chevauchement entre deux gestes qui se trouvent au même niveau articulatoire (*articulatory tier*). Ainsi, le geste qui correspond à la fermeture bilabiale peut se chevaucher avec le geste alvéolaire précédent, ce qui rend ce dernier inaudible par la suite. Le recouvrement de voisement sur le début de fermeture bilabiale donne la transcription /bp/. En phonologie articulatoire, l'assimilation est donc considérée comme un processus de recouvrement des gestes articulatoires (*gestural overlap*), plutôt qu'un transfert de caractéristiques entre deux sons contigus.

Si le processus d'assimilation des sons correspond à un recouvrement des gestes, plutôt qu'à un remplacement d'un geste par un autre, il devrait alors être possible d'observer des preuves articulatoires pour des gestes « cachés » (*hidden gestures*). Browman & Goldstein (1990a) ont analysé des séquences telles que *perfect memory*, prononcé [pə' r fɛk' mɛm...], où il y a

délétion de la consonne coronale /t/. A l'aide des rayons X, ces auteurs ont pu observer que le geste alvéolaire correspondant à /t/ était visible, alors que il était invisible sur le spectrogramme. Qui plus est, ce phonème était inaudible d'un point de vue perceptif. Ainsi, le recouvrement entre les gestes est à l'origine des assimilations et des délétions consonantiques, mais il laisse néanmoins des « traces articulatoires » dans le signal. Ces observations sont alors en accord avec de nombreux travaux conduits en phonétique expérimentale et en psycholinguistique comme nous le verrons par la suite.

1.1.3. *Données en phonétique expérimentale*

Si l'assimilation est souvent décrite en termes d'un processus complet, les travaux en phonétique expérimentale, utilisant la technique d'électropalatographie (EPG, Wright & Kerswill, 1989 ; Nolan, 1992) ont fourni des preuves articulatoires indiquant que la transformation d'un son à un autre son est rarement complète mais plutôt graduée. C'est-à-dire que le segment assimilé contient des traces subtiles concernant le son d'origine et le son dérivé par la transformation.

A titre d'illustration, Nolan (1992) a recueilli des données électropalatographiques pour étudier l'assimilation de place en anglais. Cette technique consiste à placer dans la bouche du sujet un palais artificiel contenant des électrodes. Les sujets devaient prononcer des items insérés dans un contexte phrastique assimilatoire (e.g., *late calls*) ou dans un contexte neutre (*make calls*). L'ordinateur enregistre tout contact de la langue avec les électrodes du palais artificiel, ce qui permet de comparer les degrés d'accomplissement de l'occlusion alvéolaire entre la condition test et la condition contrôle. Les résultats obtenus ont montré que, d'un point de vue articulatoire, l'assimilation de place en anglais est un processus gradué, qui va d'une occlusion complète de l'arcade alvéolaire à une ouverture proche de celle observée pour la production de séquences non alvéolaires. Ces données suggèrent donc qu'il peut y avoir des traces subtiles de la prononciation sous-jacente d'un segment, bien que ce segment puisse, par ailleurs, être perçu comme étant assimilé à 100 %. Wright et Kerswill (1989) ont demandé à des phonéticiens expérimentés d'écouter des séquences contenant des assimilations de place. Les phonéticiens avaient pour tâche de catégoriser les stimuli, utilisant une classification sur une échelle de 1 à 4 allant de (1) « entièrement alvéolaire » (i.e. pas d'assimilation) ; (2) « partiellement alvéolaire » (assimilation partielle) ; (3) « zéro alvéolaire » (forte assimilation) ; et enfin (4) « non alvéolaire » (stimuli avec pour forme sous-jacente des vélaires). Les résultats de cette catégorisation donnaient lieu à un recouvrement important entre les distributions de réponses, ce qui suggérait qu'il y avait un degré élevé d'ambiguïté généré par

les assimilations dans les stimuli fortement assimilés de type « zéro alvéolaire ». Holst et Nolan (1995) ont montré que l'assimilation des fricatives en position finale peut conduire à la production des segments avec des enveloppes de friction qui contiennent une sorte de mélange des caractéristiques de lieu appartenant à la fois à la place alvéolaire et à la place palato-alvéolaire. Cependant, ces auteurs notent que l'assimilation des consonnes coronales peut parfois donner lieu à la création d'une enveloppe spectrale qui est compatible avec le lieu d'articulation non coronal. Ceci suggère que, en dépit de sa nature graduée, la modification par assimilation peut conduire à un changement absolu (voir aussi Solé, 2003).

Les études en phonétique expérimentale ont donc permis de démontrer que l'assimilation de place implique un processus phonétique gradué plutôt que dichotomique, ce qui est conforme avec de nombreuses données expérimentales obtenues sur l'assimilation et d'autres types de variations phonologiques, comme en anglais (cf. Barry, 1985 ; Gow & Hussami, 1999 ; Gow 2002, 2003), en allemand (cf. Kuzla 2003 ; Dinssen & Charles-Luce, 1984 ; Peter & Crawford, 1989), en russe (Burton & Robblee, 1997), en néerlandais (Warner et al., 2004 ; Ernestus, 2000 ; Ernestus & Baayen, sous presse) ou encore en catalan (Solé, 2003). Les segments qui ont subi une modification contiennent apparemment des indices acoustiques résiduels et cette observation a des implications importantes pour leur traitement perceptif.

1.1.4. Psycholinguistique de la parole assimilée : Production et Perception

Il a été fréquemment observé par les phonologues que l'assimilation peut donner lieu à des profils asymétriques, dans le sens où certains traits peuvent adopter les caractéristiques d'autres traits alors que le cas inverse ne se produit jamais. Par exemple, nous avons déjà vu plus haut que, en ce qui concerne l'assimilation de place en anglais, une consonne coronale peut adopter le lieu labial alors que l'inverse ne se produit pas. Blumstein (1986) cite le profil asymétrique dans la palatalisation de [k] dans l'environnement de [i], fréquemment observé en hongrois⁴. L'observation qu'il existe des asymétries dans les processus d'assimilation a conduit certains phonologues à attribuer un rôle important à la perception lors de la

⁴ Quand la consonne vélaire [k] est suivie par une voyelle antérieure non arrondie comme [i], le [k] peut devenir [c] par processus de palatalisation, donc [ki] → [ci]. En revanche, quand la vélaire est suivie par une voyelle postérieure comme [u], une telle palatalisation n'est pas possible étant donné que les caractéristiques spectrales de [k] et [c] sont différentes dans un environnement vocalique [u]. Les conditions nécessaires au processus d'assimilation basées sur les propriétés acoustiques partagées entre le son original et le son modifié ne sont alors pas remplies.

production de la parole (Fujimura, Macchi, & Streeter, 1978 ; voir aussi Steriade, 2000). Ces idées ont donné lieu à des cadres théoriques plus élaborés, comme le modèle H&H (*Hyper- & Hypoarticulation Model*), proposé par Lindblom (1990). D'après ce modèle, la production de parole est en quelque sorte « guidée » par la perception. Le locuteur serait en mesure de contrôler les modifications phonologiques pendant qu'il calcule la distance perceptive entre la forme canonique et la forme de surface. La parole est alors produite dans un continuum allant d'une « sous-articulation » (hypo-articulation) à une « sur-articulation » (hyper-articulation). C'est-à-dire que, d'un côté, les lois du système articulatoire cherchent à obtenir une économie motrice, comme un principe du moindre effort ou de coût énergétique (cf. Kohler, 1990), et de l'autre côté, le locuteur semble être guidé par un modèle contenant les impératifs perceptifs exigés pour l'intelligibilité de la parole, une idée qui a été reprise par certains phonologues (cf. Steriade, 2000 ; Shinohara, à paraître). Dans tous les cas, les développements récents en phonologie et en phonétique expérimentale ont clairement montré que ces domaines d'études ont le même objectif que l'étude psycholinguistique, à savoir l'élaboration d'un modèle intégré du locuteur-auditeur permettant de rendre compte des capacités de perception et de production du langage (cf. Segui & Ferrand, 2000).

En résumé, nous avons abordé ici plusieurs approches qui ont cherché à donner des descriptions appropriées de l'assimilation en termes des propriétés linguistiques et physiques. En phonologie classique, les sons de parole sont définis à partir de matrices de traits phonétiques. Toute variation phonologique peut alors être expliquée grâce à un ensemble de règles à caractère linéaire, mais il s'est avéré que cette formalisation n'est pas suffisante dans la mesure où elle pouvait générer des processus phonologiques non plausibles. Les tenants de la phonologie non linéaire (ou autosegmentale) ont alors proposé que les traits phonétiques soient organisés de façon hiérarchique sous forme d'une structure arborescente. Tout processus d'assimilation peut ainsi être considéré comme une sorte de « transfert » d'un trait ou d'un groupe de traits.

Si la tradition en phonologie veut que l'assimilation soit décrite comme s'agissant d'une transformation complète d'un son en un autre son, les données acoustiques ont permis d'avancer l'idée que, d'un point de vue phonétique, l'assimilation constitue un processus rarement complet mais plutôt gradué. D'après l'approche articulatoire, la structure phonologique est représentée en termes de mouvements articulatoires coordonnés ou gestes. Dans cette optique, les variations phonologiques sont décrites en termes de recouvrement de gestes, et non en termes de transfert de traits d'un son à un autre son. Enfin, des données de la

phonétique expérimentale ont permis de mettre en évidence que les segments assimilés contiennent des traces acoustiques fines qui appartiennent au son d'origine et au son dérivé par assimilation.

2. MODELES PSYCHOLINGUISTIQUES ET RECHERCHES EMPIRIQUES SUR LES VARIATIONS PHONOLOGIQUES

2.1. Modèles théoriques en psycholinguistique

Tout processus psychologique implique la conversion d'une sorte d'information en une autre. Dans la reconnaissance des mots, il s'agit de convertir des sons en une séquence de mots qui sont reconnus par l'auditeur. Ce processus de reconnaissance impose forcément une direction dans l'envoi des informations, à savoir des sons aux mots ; cette direction est nécessaire avant que tout processus de reconnaissance puisse avoir lieu. La question reste de savoir si des informations provenant des niveaux supérieurs affectent le traitement des niveaux inférieurs. Cette problématique constitue un sujet de débat majeur en psycholinguistique depuis fort longtemps.

Dans la littérature consacrée à la perception de la parole, on distingue généralement deux types de modèles : les modèles dits « interactifs » (cf. le modèle TRACE, McClelland & Elman, 1986) vs. les modèles dits « autonomes » (cf. le modèle MERGE, Norris McQueen et Cutler, 2000). Par modèle interactif nous entendons des modèles qui postulent que les différents niveaux de traitement peuvent échanger des informations qui les rendent ainsi interdépendants. Dans les modèles autonomes, chaque niveau de traitement procède de manière indépendante et ne peut échanger des informations avec les autres niveaux que par leurs outputs correspondants. Il n'existe donc pas de véritable interaction entre les différents niveaux. Les modèles interactifs postulent des traitements descendants (*top-down processing*) et ascendants (*bottom-up processing*), alors que les modèles autonomes supposent généralement que le traitement se fait de manière strictement ascendante (*bottom-up processing*). A titre d'exemple de ce débat, nous pouvons citer les interprétations proposées pour rendre compte de l'identification, ou catégorisation phonémique.

Il a été généralement admis que cette identification phonémique peut être influencée par les informations provenant du niveau lexical (cf. Warren, 1970, Ganong, 1980 ; Samuel, 1981). A titre d'exemple, quand l'auditeur entend un item comme /?ape/, où /?/ désigne un phonème

ambigu qui se trouve quelque part entre le /t/ et /d/, il est clair que l'auditeur est biaisé vers l'interprétation de /t/ pour que le signal d'entrée soit compatible avec le mot anglais *tape*. Ce genre d'observation s'accorde avec l'hypothèse que le traitement phonémique est affecté par les informations lexicales (*top-down processing*). Cependant, d'autres auteurs dont Norris, McQueen et Cutler (2000) ont rejeté cette interprétation de nature interactive en soulignant que toute influence lexicale sur l'identification phonémique résulte d'une intégration tardive des informations provenant des niveaux phonémiques et lexicaux autonomes (voir également, Burton, Baum, & Blumstein, 1989 ; Frauenfelder, Segui, & Dijkstra, 1990 ; mais voir Mirman, McClelland, & Holt, sous presse). L'interprétation de cet effet reste encore d'actualité (voir Samuel, 2001, pour une critique du modèle MERGE)⁵.

A un niveau plus général, le débat sur la nature autonome vs. interactive du traitement des différents niveaux linguistiques n'a pas encore été tranché, loin s'en faut, mais l'opposition entre ces deux cadres théoriques ne semble plus très appropriée, comme le notent Boland et Cutler (1996). En effet, à l'intérieur même des différents modèles, on peut trouver des aspects autonomes et interactifs de traitement. Ainsi, par exemple, le modèle Cohorte de Marslen-Wilson, qui est considéré généralement comme un modèle « interactif », est autonome quant à la création des candidats lexicaux de la cohorte initiale. Ceci résulte de l'application du principe de la priorité du signal. En revanche, le modèle TRACE est interactif par la présence des connexions descendantes qui peuvent directement affecter la génération de la cohorte initiale des candidats lexicaux.

A l'heure actuelle, les psycholinguistes cherchent à réconcilier les deux cadres théoriques qui les a si longtemps divisés (cf. Boland et Cutler, 1996). Quand deux courants opposés dominent les grands débats théoriques, la vérité n'est pas une mais plurielle. Ainsi, la nature du traitement de la parole n'est, selon toute probabilité, ni exclusivement autonome, ni exclusivement interactive, mais elle se situe quelque part entre ces deux types d'architecture. Norris, McQueen et Cutler (2003), tenants des modèles de type autonome, ont récemment proposé que le *feedback* lexical pourrait jouer un rôle dans l'apprentissage phonémique, alors qu'il n'est toujours pas nécessaire lors du traitement perceptif en temps réel. Ces auteurs ont utilisé un réseau simple de type *feed-forward* pour illustrer que le traitement en temps réel

⁵ Comme l'a noté Samuel (2001), MERGE dispose d'un niveau d'entrée phonémique (*phoneme input level*), d'un niveau de sortie phonémique (*phoneme output level*) et un niveau lexical (*lexical level*). Il y a des connexions descendantes allant du niveau lexical au niveau de sortie phonémique, ce qui est quelque peu surprenant pour un modèle qui se veut strictement autonome.

fonctionne de façon ascendante alors que l'apprentissage, procède, quant à lui, de manière descendante (voir aussi Magnuson, McMurray, Tanenhaus, & Aslin, 2003 ; McQueen, 2003). Comme le soulignent Bowers et Davis (2004), même si cette nouvelle approche soulève un certain nombre de problèmes, elle a eu le mérite de poser en termes nouveaux le débat sur la nature autonome vs. interactive du traitement de la parole.

2.1.1. *Le rôle de la variabilité et problèmes de modélisation*

Norris et collègues (2000) ont beaucoup insisté sur le fait que le système perceptif humain semble très sensible aux variations acoustiques subtiles présentes dans le signal, ce qui a été confirmé par de nombreuses études de psycholinguistique (cf. Warren & Marslen-Wilson, 1987; Andruski, Blumstein, & Burton, 1994 ; Marslen-Wilson & Warren, 1994 ; Marslen-Wilson, Moss, & Van Halen, 1996 ; Utman, Blumstein, & Burton 2000). Dans les modèles classiques d'activation interactive, la variabilité présente dans le signal a été considérée comme du bruit (cf. McClelland & Rumelhart, 1981). Cependant, les modèles connexionnistes plus récents de la perception de la parole présupposent que tous les contrastes phonétiques présents à un moment donné dans le signal de parole peuvent être mis à profit dans la reconnaissance auditive des mots. Ainsi, toute variation motivée par le contexte phonologique (*lawful variation*) est interprétée dans le modèle TRACE non pas comme du bruit, mais comme une source potentielle d'information, pouvant faciliter le traitement (McClelland & Elman, 1986 ; Elman et McClelland, 1986). De même, dans les premières versions de ces modèles, les connexions entre les nœuds étaient fixes (McClelland & Rumelhart, 1981). Or, les connexions fixes posent un problème dans le sens où il y n'existe pas une relation invariante entre les valeurs de traits selon leurs différents contextes de réalisation. Dans TRACE, les nœuds phonémiques disposent des connexions libres qui peuvent moduler la force des interconnexions entre les nœuds traits et les nœuds phonémiques dans les tranches temporelles adjacentes. Ce dispositif de modulation-par-pondération (*weight-modulation scheme*) peut non seulement exciter ou inhiber les nœuds, mais également déterminer la force d'excitation ou d'inhibition exercée sur ceux-ci. Cette propriété computationnelle rend le modèle alors plus sensible aux effets contextuels dans le traitement de la variabilité systématique, et représente donc un avantage par rapport aux modèles classiques d'activation interactive.

Plus récemment, des réseaux de neurones récurrents simples (Elman, 1990 ; voir aussi Elman, sous presse) ont été employés dans le but de simuler le décours temporel du traitement de l'information phonologique dans la perception de la parole et la reconnaissance des mots (cf.

Gaskell, Hare, & Marslen-Wilson ; Gaskell 2003). L'utilisation de ce type de modélisation s'est avérée très fructueuse, et nous y reviendrons à plusieurs reprises.

2.2. Etude psycholinguistique des variations phonologiques

Reconnaître la parole implique un processus d'accès au lexique durant lequel l'entrée sensorielle acoustique est mise en relation avec les représentations lexicales stockées en mémoire. Or, la présence de segments ayant subi des variations génère une ambiguïté dans le sens où l'identité sous-jacente du mot stimulus ne peut pas être directement extraite à partir des propriétés phonétiques présentes dans sa forme de surface. Tout modèle théorique sur la perception de la parole devrait alors expliquer de quelle manière le système perceptif est organisé afin d'assurer l'accès au lexique en dépit de la variation phonologique. Si la plupart des études psycholinguistiques se sont consacrées au phénomène d'assimilation, quelques recherches récentes ont également examiné d'autres types de variations phonologiques, comme les variations allophoniques (McClennan, Luce, & Charles-Luce, 2003 ; Sumner & Samuel, 2005). Ces travaux ont permis aux chercheurs d'aborder, en particulier, des questions importantes liées à la nature des représentations lexicales. Avant d'aborder le cas particulier des interprétations avancées à propos de l'assimilation, nous examinerons ici très brièvement quelques recherches consacrées à d'autres sortes de variations phonologiques.

A titre d'exemple, on peut citer l'étude effectuée par McClennan et al. (2003). Ces auteurs ont examiné le statut de la variation allophonique du son battu [R] (*flap*) en anglais. Ce son battu représente la forme neutralisée et allophonique des consonnes intervocaliques /t/ et /d/. Un mot tel que [aRɒm] contenant un son battu est ambigu dans le sens où il peut donner lieu à deux interprétations : *atom* ou *Adam*. Des stimuli ayant des consonnes intervocaliques alvéolaires et non alvéolaires (*bacon*) ont été sélectionnés et présentés dans deux blocs expérimentaux. Tantôt les mots stimuli servaient d'amorce, tantôt ils servaient de cible. La tâche utilisée était celle de *shadowing*, pendant laquelle les mots stimuli sont présentés auditivement aux participants, qui doivent les répéter aussi rapidement que possible. Les résultats obtenus ont montré que, en ce qui concerne les stimuli « alvéolaires », les amorces hypoarticulées (i.e. celles contenant les sons battus) et hyperarticulées (e.g., /atom/) ont amorcé les cibles hypo- et hyperarticulées de façon aussi efficace. Quant aux stimuli non alvéolaires (*bacon*), les temps de production ont révélé que l'amorçage était le plus efficace quand l'amorce et la cible partageaient le même style articulatoire. Ces résultats suggèrent donc que, en présence d'une ambiguïté (i.e. les stimuli alvéolaires), des représentations intermédiaires étaient activées alors que, en absence de toute ambiguïté (i.e. les stimuli non

alvéolaires), seules les formes de surface permettent un accès efficace aux représentations lexicales correspondantes. Les auteurs ont suggéré que le degré d'activation des représentations intermédiaires peut également être modulé par la « profondeur » de traitement lexical. En effet, dans la tâche de décision lexicale dite « facile », dans laquelle étaient insérés des non-mots qui ne respectent pas les règles phonotactiques en anglais (e.g., *thushtugde*), les mots hypoarticulés n'ont pas amorcé leurs contreparties hyperarticulées, alors qu'un tel effet d'amorçage était observé dans une tâche de décision lexicale dite « difficile », dans laquelle des non-mots très semblables aux mots réels étaient insérés. Les auteurs ont avancé l'idée que l'accomplissement de cette dernière tâche nécessite une analyse lexicale plus approfondie. Cette complexité supplémentaire de la tâche rallonge le temps de traitement et permet ainsi aux représentations intermédiaires d'être activées.

Pour leur part, Sumner et Samuel (2005) ont étudié les conséquences perceptives d'une variation systématique qui consiste à transformer la consonne coronale finale /t/ en anglais. À part sa réalisation canonique, cette consonne peut être réalisée sous forme coarticulée glottalisée, ou, encore, sous forme glottalisée. Afin de savoir si les différentes formes de cette consonne perturbent ou non l'accès au lexique, ces auteurs ont conduit une expérience en utilisant une tâche d'amorçage sémantique dans laquelle les trois variantes étaient présentées comme amorce (e.g., *flute*) à une cible sémantiquement reliée (e.g., MUSIC). Les résultats ont montré que les trois variantes ont donné lieu à des effets d'amorçage sémantique similaires. En revanche, dans une tâche d'amorçage à long terme, il y avait un avantage perceptif pour la forme canonique.

Le fait que les variations systématiques observées dans une langue soient motivées par les règles phonologiques qui lui sont propres a conduit certains auteurs à proposer que le traitement perceptif des variations est modulé par les mécanismes spécifiques à la langue en question. En ce qui concerne l'assimilation de lieu d'articulation en anglais par exemple, il est présumé que l'auditeur de langue maternelle anglaise prend appui sur le contexte droit afin d'inférer la forme sous-jacente du segment assimilé, car il a des « connaissances implicites » des règles phonologiques de sa langue et sait donc qu'une consonne coronale peut prendre la forme labiale dans le contexte labial (alors qu'une labiale ne prend jamais la forme coronale). On retrouve l'idée de la spécificité à la langue dans les théories qui se prononcent sur la nature des représentations, comme c'est le cas de la théorie de la sous-spécification.

2.2.1. *Théorie de la sous-spécification et le modèle FUL*

Dans leur étude sur le traitement perceptif du trait de nasalité vocalique en anglais et bengali, Lahiri et Marlsen-Wilson (1991) ont proposé une solution représentationnelle qui est censée être spécifique à la langue. Suivant la théorie phonologique de sous-spécification (Archangeli, 1988 ; Kiparsky, 1985), l'idée proposée est la suivante : certains traits phonologiques sont sous-spécifiés dans le lexique mental, ce qui veut dire qu'ils restent neutres par rapport à certaines dimensions dans la représentation sous-jacente. Il convient d'introduire ici la notion phonologique de la marque (*markedness*). Précédemment, il a été précisé que, selon SPE, chaque trait dispose de la valeur [+] ou [-]. Or, dans la théorie de la sous-spécification, une de ces valeurs est considérée comme étant la valeur dite marquée. Il en résulte que l'autre valeur non marquée est la valeur par défaut. La notion de la marque s'avère pertinente dans le sens où elle permet d'expliquer par exemple la fréquence d'occurrence de certains sons qui sont spécifiés dans une langue donnée. A l'inverse, la valeur non marquée correspond à des propriétés phonologiques plus générales. De ce fait, la sous-spécification des sons dans le lexique mental dépend à la fois à des contraintes spécifiques à la langue et à des contraintes universelles. Lahiri et Marslen-Wilson (1991) ont fourni des données empiriques qui sont en accord avec les propositions de la théorie de sous-spécification. Ces auteurs ont examiné le cas de la nasalité vocalique, un trait qui n'est pas contrastif en anglais (et donc sous-spécifié) alors qu'il est contrastif en bengali. Les résultats d'une tâche de dévoilement progressif (*gating*, cf. Grosjean, 1980) ont montré que toute trace de nasalité perçue était associée à la présence d'une consonne nasale par les auditeurs anglais. En revanche, les auditeurs bengalais associaient toute trace de nasalité à la présence des voyelles (même si des voyelles orales peuvent subir une assimilation nasale dans la présence des consonnes nasales).

L'approche de sous-spécification revêt peut-être sa forme la plus radicale avec le modèle FUL (*Full Underspecified Lexicon*), proposé par Lahiri et Reetz (2002)⁶. Suivant les principes de la sous-spécification, le modèle FUL postule que certains traits ne sont pas spécifiés dans le lexique mental. Le système perceptif analyse le signal de parole à la recherche des traits

⁶ Dans le cadre de la discussion sur les mécanismes cognitifs mis en jeu dans la perception de la parole assimilée, nous avons choisi de présenter le modèle FUL ici comme étant représentatif pour les modèles qui postulent l'existence des mécanismes de traitement perceptifs spécifiques à la langue. Si la discussion portait sur la nature des représentations des formes lexicales, il est évident que le modèle FUL s'inscrit dans le cadre abstractionniste, par opposition aux modèles de type épisodique / à exemplaires (cf. Klatt, 1979 ; Goldinger, 1998 ; voir aussi Nguyen, sous presse ; Hawkins, 2003).

acoustiques qui sont directement convertis en traits phonologiques. Ainsi, les traits phonologiques sont appariés avec les représentations stockées dans le lexique mental par une conversion triple : *match* (concordance), *mismatch* (disconcordance), ou *no mismatch* (absence de toute discordance, ce qui n'équivaut pas à « concordance »). D'après le modèle FUL, l'auditeur est parfois insensible aux variations présentées par le lieu d'articulation de certaines consonnes finales, dans la mesure où celles-ci ne sont pas spécifiées pour certains traits. A titre illustratif, dans la séquence *where can Mister Bean be ?*, la consonne coronale dans le nom propre *Bean* peut devenir une labiale en présence du contexte labial dans *be*, et être ainsi prononcée [bim]. A l'inverse, les règles phonologiques en anglais ne permettent pas la transformation d'une consonne labiale en coronale (*lame duck* □ *[len]⁷). Il est alors supposé qu'en anglais (et aussi en allemand), la place coronale est sous-spécifiée dans le lexique, ce qui donne lieu à une certaine tolérance vis-à-vis de la forme de surface : la forme coronale peut se transformer en labiale et cela ne gêne en rien le système perceptif. En revanche, la place labiale est spécifiée dans le lexique anglais, ce qui rend tout changement de lieu d'articulation impossible : la labiale ne se transforme jamais en coronale. En résumé, la sous-spécification est utile dans la reconnaissance de la parole, car elle permettrait aux auditeurs de pouvoir tolérer certaines variations qui sont motivées par les règles phonologiques de leur langue. On retrouve l'idée selon laquelle les mécanismes de traitement de la parole assimilée sont guidés par les règles spécifiques à la langue dans les travaux de Gaskell et collègues, comme nous allons le voir dans la section suivante.

3. L'ASSIMILATION □ CADRES THEORIQUES ET RECHERCHES EXPERIMENTALES

L'assimilation est représentative de certains phénomènes de parole continue qui donnent lieu à des variations systématiques et qui posent un problème aux modèles actuels de perception de la parole. C'est la raison pour laquelle de nombreux psycholinguistes se sont intéressés à ce phénomène depuis maintenant une dizaine d'années. Dans tous ces travaux, les chercheurs se sont focalisés sur les conséquences perceptives liées au traitement des segments assimilés. Dans la littérature, deux types de mécanismes ont été proposés pour rendre compte du traitement perceptif : (1) l'hypothèse du mécanisme d'inférence régressive et (2) l'hypothèse de l'utilisation des indices acoustiques et l'alignement des traits.

⁷ En phonologie, la présence de l'astérisque désigne des formes lexicales non existantes.

3.1. L'Hypothèse du Mécanisme d'Inférence Régressive

Nous avons vu plus haut que l'assimilation d'un son est toujours déclenchée par le son suivant dans le cas d'assimilation régressive, et par le son précédent dans le cas d'assimilation progressive. En ce qui concerne l'assimilation régressive de lieu d'articulation en anglais ou en allemand, le son assimilé peut prendre la place d'articulation du son suivant. Etant donné que c'est le contexte phonologique suivant qui déclenche le processus d'assimilation, certains auteurs ont proposé que les auditeurs prennent appui sur ce contexte par inférence régressive afin de dériver l'identité sous-jacente du segment assimilé (cf. Gaskell & Marslen-Wilson, 1996 ; 1998 ; Coenen, Zwitserlood, & Bölte, 2001, Mitterer & Blomert, 2003, mais voir Gumnior, Zwitserlood, & Bölte, 2005).

Gaskell, Hare, et Marslen-Wilson (1995) ont implémenté le mécanisme d'inférence régressive dans un modèle de type connexionniste basé sur les réseaux récurrents simples (*simple recurrent networks*, désormais SRN, voir Plunkett & Elman, 1997). Les SRN sont dotés d'un algorithme de propagation de retour (*back-propagation*), permettant de modifier les connexions entre nœuds par comparaison entre la sortie du réseau et l'ensemble des valeurs obtenues par entraînement. Le signal d'entrée dans ce genre de modèles est constitué des traits phonologiques. A l'entrée, tout segment assimilé (i.e. [lim] *bacon*) passe au niveau phonologique, et est perçu tel quel (e.g., /m/) par le modèle. En revanche, au niveau phonologique, une couche dite cachée (*hidden layer*) connectée à un réseau de neurones simple, peut affecter l'activation des nœuds correspondant aux traits phonologiques et corriger ensuite l'assimilation (i.e. le /m/ dans [lim] est interprété comme étant la forme de surface de /n/). Ainsi, l'interprétation du segment n est influencée par le segment $n+1$, ce qui rend le modèle particulièrement approprié pour rendre compte des effets contextuels. Le modèle peut être entraîné à apparier les différentes variations dans la forme de surface à la forme sous-jacente. Les régularités statistiques du signal de parole sont extraites afin de pouvoir généraliser sur les entrées nouvelles, n'ayant jamais été présentées auparavant au modèle.

Plusieurs auteurs ont obtenu des données empiriques qui sont accord avec l'approche d'inférence régressive. Gaskell & Marslen-Wilson (1996) ont utilisé la tâche de décision lexicale avec amorçage intermodal (*cross-modal priming*) et le paradigme de détection de phonèmes (1998) afin de montrer que l'accès au lexique peut être perturbé si la variation phonologique n'était pas motivée par le contexte droit. A titre illustratif, l'amorce [lim] ne donnait pas lieu à une activation de la cible LEAN si le mot était présenté dans la séquence

leam gammon, où le contexte vélaire (/g/ de *gammon*) ne peut pas induire une assimilation labiale en /m/ de la coronale finale /n/ dans *lean*. Des conclusions similaires sur le rôle du contexte droit ont été formulés par d'autres (Coenen, Zwitserlood et Bólte, 2001 ; Mitterer & Blomert, 2003). C'est à partir de ces études, qui seront examinées plus en détail dans le chapitre 3, que l'on a proposé l'hypothèse selon laquelle le système de traitement perceptif est doté d'un mécanisme d'*inférence régressive*, permettant de ré-analyser un segment assimilé en fonction de son contexte phonologique droit.

3.2. L'Hypothèse des Indices Acoustiques d'Assimilation et le Mécanisme Perceptif d'Alignement des Traits

L'hypothèse du mécanisme d'inférence régressive formulée par Gaskell et collègues a été remise en question dans une série de travaux essentiellement conduits par David Gow (2001, 2002, 2003, 2004). Cet auteur a montré en particulier qu'un segment est rarement, voire jamais, complètement assimilé, ce qui est en accord avec les données provenant de la phonétique expérimentale (présentées dans la section 1.1.3). Dans la mesure où il existe des degrés intermédiaires de transformation d'un son à un autre son, le segment assimilé contiendrait suffisamment d'indices acoustiques fins correspondant à la forme sous-jacente. Par conséquent, l'auditeur n'aurait pas besoin du contexte phonologique droit pour traiter l'assimilation : les indices acoustiques seraient directement exploités par l'auditeur pour inférer la forme sous-jacente du segment altéré (dans certains cas il peut même « anticiper » le segment suivant le segment assimilé). A l'appui de son hypothèse, Gow (2002) a montré que, quand les auditeurs entendent un mot comme *right* ayant subi une assimilation labiale et prononcé [raip], ils détectent plus rapidement le phonème labial suivant. D'après Gow, cette facilitation est le résultat de l'association du trait labial dans le segment assimilé, et le trait labial du segment suivant. Dans la même étude, Gow a également mis en évidence que la présentation de la forme assimilée de *right* amorce la cible RIGHT, mais non pas la forme de surface apparente RIPE. Ce résultat est probablement dû à l'association des traces de coronalité, présentes dans le signal et récupérées par l'auditeur, avec le segment coronal final dans *right*. Ces données remettent en cause l'existence même d'assimilations complètes. Il semblerait qu'un segment ne soit rarement voire jamais complètement assimilé, mais qu'il existe au contraire des degrés intermédiaires de transformation d'un son à un autre.

Afin de rendre compte des données empiriques, Gow a proposé la théorie d'alignement des traits (*Feature Alignment Theory*), selon laquelle des traits acoustiques temporellement distribués sont regroupés et réintégrés afin de former des représentations unifiées. Ces

représentations unifiées sont alignées avec les positions segmentales correspondantes. La théorie de l'alignement des traits est illustrée dans la Figure 1.2.

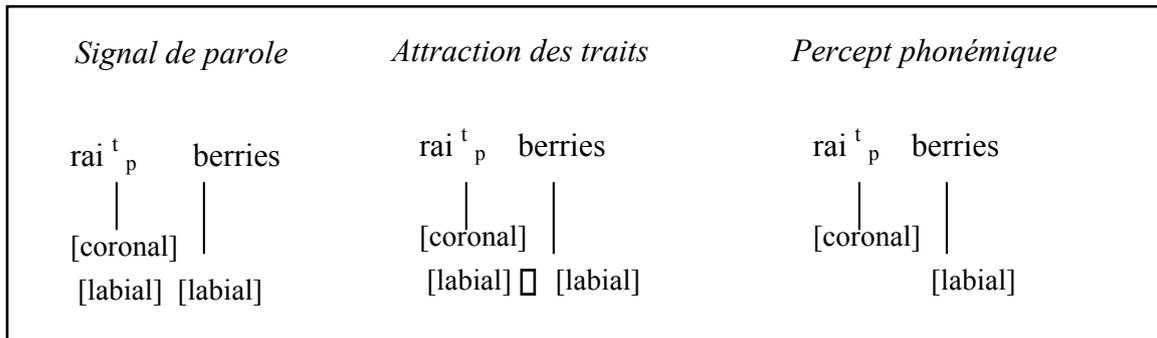


Figure 1.2. Illustration de la théorie d'alignement des traits (d'après Gow & Zoll, 2002).

La consonne finale coronale /t/ peut subir une assimilation labiale quand elle est insérée dans la séquence *right* [raip] *berries*. Cependant, étant donné la nature incomplète de l'assimilation, cette consonne contient en fait deux places d'articulation au lieu d'une seule : celle qui correspond à la place coronale et celle qui correspond à la place labiale. D'après la théorie de l'alignement des traits, les indices acoustiques correspondant à la place labiale dans le segment assimilé peuvent être associés à la place labiale du contexte suivant dans *berries*. Ceci laisse des traces articulatoires en faveur de la place coronale, permettant ainsi aux auditeurs d'accéder à *right* et non pas à *ripe*⁸.

La notion d'alignement et groupement des traits est compatible avec l'observation selon laquelle de multiples indices sont intégrés au sein du même trait (cf. Best, Morrongo, &

⁸ En fait, dans la théorie d'alignement des traits, un changement assimilatoire donne lieu à des segments dits complexes, avec deux places d'articulation (Gow & Zoll, 2002). Le phénomène des segments « complexes », i.e. des segments à articulation complexe n'ayant pas nécessairement subi de processus d'assimilation, est connu des phonologues (Gussenhoven & Jacobs, 1998 ; voir aussi Jacobs, 1993, pour des exemples en linguistique diachronique). On peut distinguer trois types de segments complexes : (1) des segments à lieu complexe (*complex place segments*), ces segments contiennent plusieurs nœuds articulatoires (cf. les sons à double articulation comme le son labial-coronal [pt] qui peut se produire dans certaines langues) ; (2) des segments à contour de mode (*manner contour segments*), ces segments comportent une séquence avec des occurrences différentes d'un même mode d'articulation (cf. des segments comme [mb], qui contiennent la valeur [+nas] et [-nas] à la fois) ; et enfin (3) les articulations secondaires, c'est-à-dire des segments qui contiennent deux spécifications, avec une place indiquant la location du mode d'articulation, et une seconde spécification indiquant simultanément une articulation vocalique (cf. [p^w]).

Robson, 1981 ; Repp, 1982). Gow a suggéré qu'un segment transformé par assimilation de place contient deux places d'articulation. D'après les principes perceptifs généraux (cf. Bregman, 1990), des groupements d'items saillants d'un point de vue perceptif peuvent attirer des éléments des groupements plus faibles. Des indices forts de place labiale associés à la consonne labiale en position initiale (e.g., /b/ dans *berries*) sont censés alors attirer les indices plus faibles de labialité dans le segment assimilé précédent (e.g., [rait]). Selon Gow, l'alignement et le regroupement perceptif des traits font partie des mécanismes perceptifs généraux, et en conséquence leur mise en œuvre ne dépend pas des expériences linguistiques préalables de l'auditeur (Gow & Im, 2004).

Gow (2003) a insisté sur le fait que la théorie d'alignement des traits peut rendre compte des effets régressifs et progressifs obtenus dans ses études. Les indices correspondant à la place labiale dans la forme assimilée [raip] permet d'une part de désambigüiser la place du segment assimilé (laissant seulement des indices de coronalité), et d'autre part de renforcer la perception de la consonne labiale /b/ du contexte suivant *berries* (i.e. les effets de facilitation progressifs obtenus dans la tâche de détection des phonèmes).

Il est important de souligner que, dans les toutes expériences menées par Gow, l'assimilation était toujours naturelle et jamais produite de façon intentionnelle ou délibérée. Le locuteur qui produisait les stimuli n'était pas au courant des règles d'assimilation. Or, dans les expériences menées par Gaskell et collègues, les stimuli étaient délibérément produits avec assimilation. D'après Gow (2001), les contradictions obtenues dans les études en perception peuvent, en partie, être expliquées en prenant en compte la nature acoustico-phonétique du segment assimilé. Dans le cas d'assimilation complète, l'auditeur s'appuie exclusivement sur le contexte suivant le segment assimilé afin de dériver sa forme sous-jacente. En revanche, dans le cas d'assimilation partielle, le segment assimilé contient suffisamment d'informations acoustiques sur la forme sous-jacente pour permettre son identification sans devoir recourir au contexte suivant. L'hypothèse selon laquelle les mécanismes mis en œuvre lors de traitement perceptifs varient en fonction du degré d'assimilation a récemment été validée à l'aide des simulations de type connexionniste (Gaskell, 2003).

3.3. Conclusion

L'extraordinaire flexibilité et subtilité du système de production et de perception du langage parlé permet généralement au sujet humain de faire face à la variabilité de réalisation du signal de parole. Comprendre une telle capacité constitue un objectif majeur en

psycholinguistique et les recherches que nous présentons dans cette thèse s'inscrivent directement dans cette problématique.

Nous avons pu constater que les études portant sur l'assimilation de lieu d'articulation ont permis d'avancer deux types de mécanismes mis en jeu lors du traitement perceptif de la parole assimilée. Certains auteurs ont proposé que l'auditeur s'appuie exclusivement sur le contexte suivant le segment assimilé afin de dériver sa forme sous-jacente. Tandis que d'autres ont suggéré que les segments assimilés contiennent suffisamment de détails acoustico-phonétiques sur le phonème sous-jacent pour permettre son identification sans avoir recours au contexte. L'étude empirique sur l'assimilation de voisement en français présentée dans cette thèse a pour principal objectif de démontrer que les deux mécanismes cognitifs proposés, à savoir l'inférence régressive et l'utilisation des indices acoustiques présents dans le segment assimilé, sont en fait compatibles dès lors que l'on prend en considération les caractéristiques acoustiques fines du matériel utilisé et, dans le cas qui nous intéresse, le degré d'assimilation des segments. Avant que nous puissions aborder le traitement perceptif de la parole assimilée, il est d'abord nécessaire de fournir une description acoustico-phonétique adéquate de l'assimilation de voisement en français, ce qui est l'objectif principal du chapitre 2.

CHAPITRE II.

LA PRODUCTION DES CONSONNES OBSTRUENTES

ASSIMILEES : PARAMETRES ACOUSTICO-PHONETIQUES ET FACTEURS LEXICAUX⁹

1. INTRODUCTION

Le système de traitement de la parole doit faire face à deux exigences à première vue contradictoires. D'une part, il doit être en mesure de distinguer les mots qui diffèrent par un ou plusieurs traits minimaux (*soude* diffère de *soute* par rapport au trait de voisement [\pm voice]). D'autre part, il doit faire face à des modifications phonétiques dans les formes de surface qui neutralisent ces distinctions subtiles (*soute* peut devenir *soude* dans un contexte phonologique assimilatoire). Toute approche théorique sur le traitement de la variabilité dans le signal de parole devrait prendre en compte ces deux exigences.

Nous avons déjà vu dans le chapitre précédent qu'il existe plusieurs points de vue concernant la nature même des variations phonologiques et que ces différentes conceptions peuvent avoir des conséquences pour les différentes interprétations concernant leur traitement perceptif. Si nous récapitulons très brièvement les données expérimentales sur le traitement perceptif de l'assimilation de place, nous trouvons d'un côté les tenants de l'approche en vertu de laquelle le traitement des variations est régi par des règles phonologiques (cf. Gaskell & Marslen-Wilson, 1996, 1998). De l'autre côté, nous trouvons ceux qui maintiennent que le traitement perceptif des variations phonologiques est guidé par des mécanismes généraux de groupements perceptifs (cf. Gow, 2002, 2003). Même si ce débat n'a pas encore été complètement tranché, il n'en reste pas moins qu'au moins une partie des résultats divergents dans la littérature est liée au fait que dans les études en perception de la parole, les propriétés acoustiques n'ont pas toujours été prises en considération. En effet, les mots stimuli utilisés par Gaskell et Marslen-Wilson (1996, 1998) ont été produits avec assimilation délibérée, alors que ceux employés par David Gow étaient issus de la parole naturelle. D'après Gow, il est

⁹ Les expériences rapportées dans ce chapitre ont fait l'objet d'un article dans *Journal of Phonetics* (Snoeren, Hallé, & Segui, sous presse), et d'un article apparu dans *Proceedings of the 15th ICPHS* (Snoeren & Segui, 2003).

alors fort probable que les effets divergents obtenus dans la littérature soient en fait dus à l'utilisation des stimuli contenant différents degrés d'assimilation.

Des données acoustiques et articulatoires fournies par la phonétique expérimentale ont permis de montrer que l'assimilation est rarement complète, mais au contraire très souvent partielle. C'est-à-dire qu'il existe des degrés intermédiaires de transformation d'un son à un autre son (Hardcastle & Roach, 1979 ; Kohler, 1976 ; Barry, 1985 ; Ernestus, 2000; Holst & Nolan, 1995 ; Kerswill, 1985 ; Nolan, 1992 ; Wright & Kerswill, 1989 ; voir aussi Browman & Goldstein, 1990, pour des données en phonologie articulatoire). A titre illustratif, dans l'étude sur l'assimilation de place en anglais conduite par Nolan (1992), les participants devaient prononcer des items insérés dans un contexte phrastique assimilatoire (e.g., *late calls*) ou dans un contexte neutre (*make calls*). Les résultats ont montré que, d'un point de vue articulatoire, l'assimilation de place en anglais est un processus gradué plutôt que catégorique, allant d'une occlusion complète de l'arcade alvéolaire à une ouverture proche de celle observée pour la production de séquences non alvéolaires. Ces données suggèrent qu'il peut y avoir des traces subtiles de la prononciation sous-jacente d'un segment (qui par ailleurs peut être perçu comme étant assimilé à 100 %).

1.1. Paramètres acoustiques du voisement

L'objectif principal de ce chapitre est de mieux appréhender la nature acoustico-phonétique de l'assimilation de voisement en français et d'en fournir une description adéquate. Nous pensons qu'une description appropriée de cette variation est nécessaire afin de mieux connaître les conséquences perceptives (voir aussi Fowler, 1987). Même s'il existe relativement peu d'études expérimentales sur l'assimilation de voisement en français, l'implémentation phonétique du trait de voisement a fait l'objet de nombreux travaux (voir, e.g., Lisker & Abrahamson, 1964 ; Ohde, 1983 ; van den Berg, 1986, 1989 ; Kohler, 1985 ; Ernestus, 2000 ; Jessen, 2002). En ce qui concerne l'implémentation du voisement en français, il convient de citer ici les travaux peu connus de van Dommelen (1983, 1985 ; voir aussi Kohler, 1985). Cet auteur a examiné la perception du voisement dans des obstruents en français et il a fait varier deux paramètres : la durée de la vibration des cordes vocales (désormais ViCV) au sein de l'occlusion et le bruit d'explosion dans les mots stimuli *rate* et *rade*, extraits de leur contexte phrastique *C'est une rate / rade* . Ensuite, ces mots stimuli étaient présentés auditivement à des participants qui avaient pour tâche de dire s'ils entendaient un /t/ ou bien un /d/. Il a été montré que la ViCV était le paramètre acoustique le plus fiable permettant aux auditeurs de distinguer l'obstruente non voisée t/ de sa contrepartie

voisée /d/. Dans la deuxième expérience, les paramètres acoustiques suivants étaient étudiés : la durée de l'occlusion, ainsi que la durée de la voyelle précédant la consonne obstruente. De nouveau, il s'est avéré que ViCV était l'indice pertinent. Qui plus est, la fiabilité perceptive de cet indice était renforcée par les durées plus courtes de l'occlusion. Cela veut dire que pour la distinction perceptive du voisement en français, ces deux indices ne peuvent pas être évalués séparément. Finalement, dans ces expériences, la durée de la voyelle était également un paramètre utilisé par l'auditeur afin de distinguer /t/ et /d/ dans la mesure où le /d/ était perçu plus fréquemment à des durées vocaliques longues. Jusqu'ici, les paramètres acoustiques étaient étudiés pour les mots en position finale, mais sont-ils aussi efficaces quand il s'agit des mots en position non finale ?

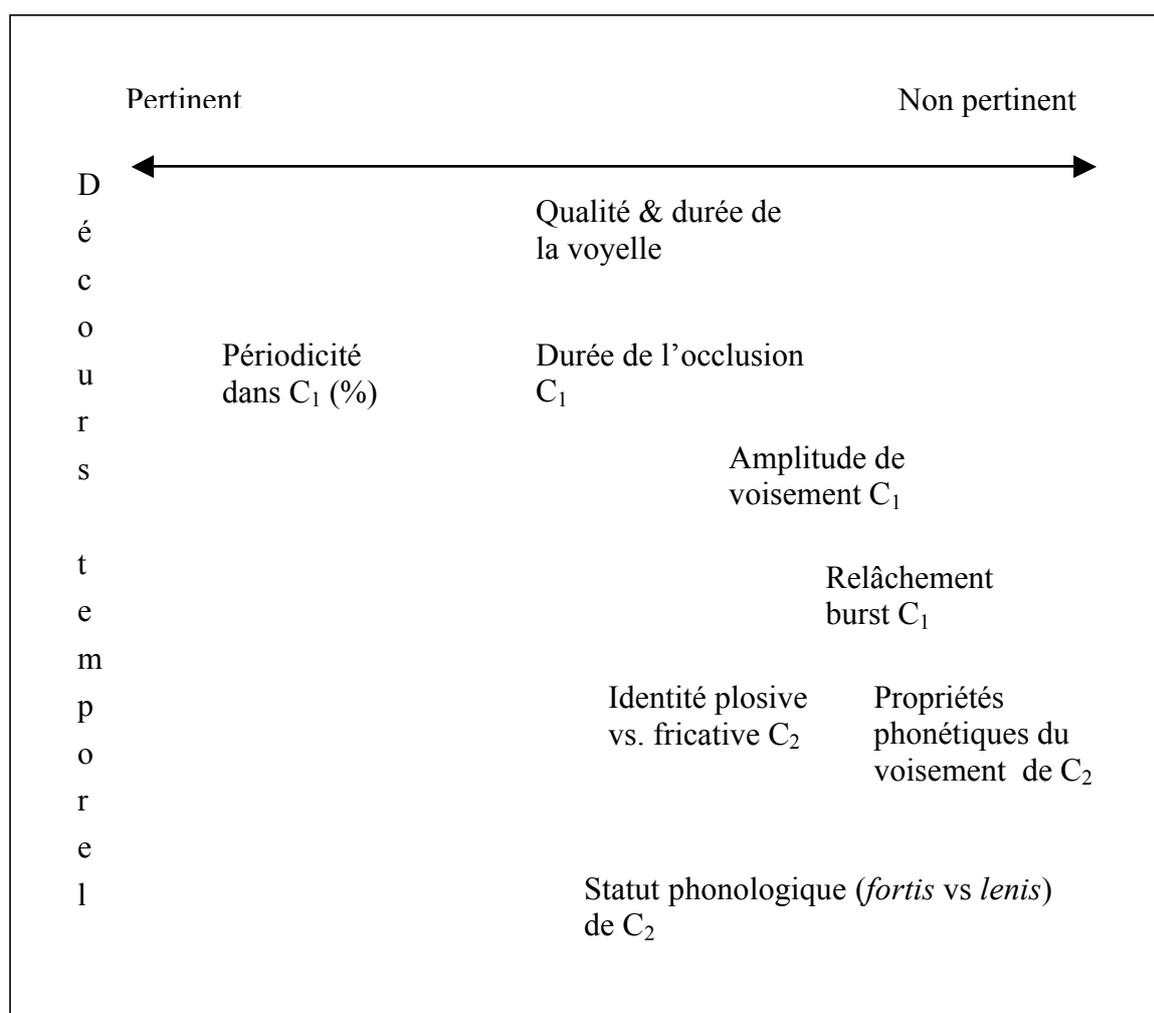


Figure 2.1. Modèle d'estimation qualitative de la contribution relative de plusieurs indices de l'identification de C₁ dans une séquence C₀VC₁ # C₂ (d'après van Dommelen, 1985).

Telle est la question qui motive les expériences dans lesquelles les paramètres acoustiques étaient étudiés pour des mots *rate / rade* insérés dans des contextes *C'est une rate/rade noire*.

Le paramètre ViCV était de nouveau un indice fiable de voisement, tout comme la durée de l'occlusion présente dans la consonne. De plus, la qualité de la voyelle précédant la consonne critique était également importante dans la mesure où une occlusion non voisée ne conduit pas à la perception du /t/ à 100% si celui-ci est suivi par une voyelle extraite d'un contexte voisé (comme dans *rade*). Inversement, une occlusion voisée ne conduit pas à la perception du /d/ à 100% si celui-ci est précédé par une voyelle extraite d'un contexte non voisé (*rate*). L'ensemble des expériences a donné lieu à un modèle, illustré dans Figure 2.1, qui présente la contribution des indices acoustiques à l'identification de C_1 dans une séquence de type $C_0VC_1 \# C_2$, en fonction de la pertinence perceptive et du décours temporel du mode de réalisation de la séquence. Nous pouvons constater que, d'après ce modèle, un rôle primordial est accordé à la périodicité de voisement présente dans C_1 , alors que les autres indices jouent un rôle secondaire. Il est à noter que dans ces travaux, les paramètres acoustiques du voisement ont été systématiquement variés, alors que dans nos expériences, nous ne manipulons jamais directement les paramètres acoustiques impliqués dans le voisement. Notre objectif ici est d'évaluer le degré d'assimilation dans les séquences $C_0VC_1 \# C_2$, à partir des paramètres acoustiques qui sont susceptibles d'être impliqués dans la perception du voisement.

1.2. Le rôle des facteurs lexicaux et pragmatiques sur les variations phonologiques

En linguistique diachronique, il a souvent été avancé que les changements phonologiques affectent tous les mots de manière semblable pourvu qu'ils connaissent le même environnement phonétique. Cependant, le linguiste allemand Schuchardt (1885, cité dans Bybee, 2002) a été l'un des premiers à faire observer que les mots ne changent pas à la même vitesse. Ceci n'est pas dû aux différences régionales, comme il est souvent prétendu par les linguistes, mais à d'autres facteurs. Labov (1994) a proposé qu'il existe deux sortes de changements phonologiques : le changement phonologique régulier « de bas niveau » qui est gradué et motivé par des règles phonétiques et le changement de « diffusion lexicale » qui est, quant à lui, conditionné par les propriétés du lexique mental. Il alors a été suggéré que ce dernier type de changement phonologique est modulé par des informations cognitives de plus haut niveau, telles que la fréquence d'usage des mots, leurs propriétés morphologiques et/ou orthographiques, ou encore les intentions communicatives du locuteur.

1.2.1. *L'influence de la fréquence des mots*

Il existe de nombreuses données qui s'accordent avec l'idée selon laquelle les variations phonologiques peuvent être modulées par des facteurs cognitifs de plus haut niveau. Par exemple, Bybee (2000, 2002) a étudié le cas de déletion des obstruents finales /t/ et /d/ en anglais et a trouvé que les mots de haute fréquence étaient plus affectés par la déletion que ceux de basse fréquence (voir aussi Jurafsky, Bell, Gregory, & Raymond, 2001). Whalen (1991, 1992) a trouvé que les mots de basse fréquence ont des durées plus longues, c'est-à-dire que ces mots-là sont produits à un débit de parole plus lent que les mots de haute fréquence. Cet auteur avait expressément utilisé des homophones (e.g., *right / rite*) donc cette différence ne pouvait pas être attribuée au nombre de phonèmes. Ces travaux mettent donc en évidence, de la part du locuteur, un ajustement de ses productions aux propriétés des items à produire afin d'assurer l'intelligibilité de son discours.

1.2.2. *L'influence des facteurs pragmatiques*

Des recherches conduites par Port et Crawford (1989) ont montré que l'importance de la neutralisation dans le dévoisement final (*Final Devoicing*) en allemand dépend de la situation communicative dans laquelle se trouve le locuteur. Dans une première condition, ces auteurs ont demandé aux participants de lire à voix haute une liste de mots isolés. Dans une seconde condition, les participants avaient pour tâche de lire à voix haute des phrases pour lesquelles le contraste neutralisé était rendu explicite (cf. *Ich habe Rat (Rad) gesagt, nicht Rad (rat)*). Ces phrases étaient lues devant un assistant qui notait les mots test produits par les locuteurs. Ce dispositif avait pour but d'éliciter l'intention du locuteur de produire les mots de façon distincte (isolés vs. en contexte). Il a été observé que dans la deuxième condition, on obtenait davantage de formes incomplètes de neutralisation que dans la première condition (la condition de liste de mots isolés). Les auteurs ont conclu que le dévoisement final était modulé par le contexte communicatif dans lequel se trouve le locuteur. Le besoin d'éviter une confusion entre le mot voulu et un mot homophone dans la condition de lecture isolée de mots, induit des productions qui préservent le voisement sous-jacent.

Dans le même ordre d'idées, Gafos (sous presse) a également étudié le domaine de la neutralisation du dévoisement final en allemand et avance l'hypothèse que les locuteurs sont en mesure de faire varier la force de la neutralisation avec plus ou moins de contraste selon leurs intentions communicatives. Un modèle dynamique non linéaire a été employé pour modéliser l'intervention des facteurs phonétiques de bas niveau (i.e. l'implémentation

phonétique de nature graduée de la règle du dévoisement final) et des facteurs cognitifs de plus haut niveau (i.e. les intentions communicatives du locuteur).

Ces travaux sont à rapprocher de ceux conduits par Fowler et Housum (1987) qui montrent que les locuteurs « atténuent » en quelque sorte leur production des mots dits anciens, i.e. déjà produits dans un discours par rapport aux mots « nouveaux », i.e. des mots qui apparaissent pour la première fois dans une conversation. Les valeurs des trois paramètres acoustiques (i.e., la durée du mot, les valeurs F0 présentes dans la voyelle accentuée et son amplitude) étaient systématiquement plus élevées pour les mots « nouveaux » que pour les mots « anciens » ce qui indique une articulation plus soignée pour les premiers. Ceci a eu des conséquences perceptives : quand ces mots étaient présentés auditivement à des participants, il a été trouvé que les mots « anciens » sont en effet perçus comme étant moins intelligibles que les mots « nouveaux » quand ils sont présentés en isolation, mais probablement pas en contexte. Finalement, ces auteurs ont observé que les auditeurs sont capables d'identifier les mots comme étant « nouveaux » ou « anciens » et qu'ils peuvent même utiliser cette information, afin de faciliter l'intégration des éléments linguistiques reliés dans un discours. Ainsi, ces résultats suggèrent que les locuteurs réduisent la production des mots « anciens » car a) il est plus difficile de produire des articulations soignées et b) à partir d'informations acoustiques présentes dans le signal, les auditeurs peuvent déduire si le mot apparaît pour la première fois dans un discours ou s'il a déjà été employé auparavant.

1.2.3. Le rôle du voisinage phonologique

Quels que soient les critères pour définir le voisinage phonologique (cf. Bailey & Hahn, 2001), il semble bien établi qu'en reconnaissance auditive de la parole, le voisinage phonologique affecte le traitement lexical. Des études ont montré que les mots avec de nombreux voisins sont traités plus lentement que les mots avec peu de voisins dans des tâches comme la décision lexicale auditive ou de la reconnaissance (cf. Luce & Pisoni, 1998 ; Vitevitch & Luce, 1999). Ces résultats n'ont rien d'étonnant si l'on considère la reconnaissance des mots comme un processus de discrimination du mot cible parmi d'autres mots : moins il y a de mots similaires, plus facile sera la discrimination. En revanche, en production de la parole, le voisinage phonologique semble faciliter le traitement, c'est-à-dire que les performances en production sont moins erronées et plus rapides. Par exemple, en utilisant la technique de dénomination des images, Vitevitch (2002) a observé que les images correspondant à des mots issus des voisinages de forte densité étaient dénommées plus rapidement que celles correspondant à des mots à voisinage de faible densité. Il semblerait

donc qu'une forte densité de voisinage gêne systématiquement la reconnaissance des mots, alors qu'elle faciliterait la production. (Voir Dell et Gordon, 2003, pour une modélisation de ces effets).

1.3. La présente recherche

Les études mentionnées ci-dessus suggèrent que l'implémentation phonétique des formes de neutralisation sont modulées par des facteurs lexicaux. Dans ce travail, nous avons examiné l'influence de deux facteurs lexicaux sur la production des segments assimilés : a) l'ambiguïté lexicale et b) le voisinage phonologique.

Le rôle de l'ambiguïté lexicale sur la perception de la parole assimilée a fait l'objet de plusieurs études en anglais (cf. Gaskell & Marslen-Wilson, 2001 ; Gow, 2002 ; 2003), mais à notre connaissance, aucune étude n'a encore évalué son influence sur la production des segments assimilés. Dans certains cas, l'assimilation peut conduire à une ambiguïté lexicale comme c'est le cas de la forme *run* dans *A quick run / rum picks you up* (Gaskell & Marslen-Wilson, 2001). Dans les expériences rapportées dans ce chapitre, cette forme d'ambiguïté par assimilation a été manipulée. Nous utiliserons le terme « mot ambigu par assimilation » (désormais MAA) pour désigner un mot appartenant à une paire minimale de deux mots par voisement en position finale. Par exemple le mot *soute* [sut] est en contraste avec *soude* [sud]. Nous appellerons « mot non ambigu » (désormais MNA) un mot qui n'appartient pas à une paire minimale de deux mots par voisement final. Ainsi, mot *frites* [frɪt] est considéré comme un mot non ambigu car sa contrepartie voisée *frides* [frɪd] ne correspond pas à un mot en français.

Le deuxième facteur lexical auquel nous nous sommes intéressé ici, concerne le voisinage phonologique. Le voisinage phonologique est un facteur lexical qui fait référence au nombre de mots qui sont phonologiquement proches du mot cible. D'après la définition classique donnée par Luce, Pisoni, et Goldinger (1990), deux mots sont considérés comme des voisins s'ils diffèrent par seulement un phonème (par substitution, addition, ou encore délétion). A titre d'exemple, le mot *cat* a comme voisins *cast*, *rat*, et *at* parmi ses 35 voisins, alors que le mot *dog* n'a que 8 voisins. Pour examiner le rôle du voisinage phonologique dans la production de la parole assimilée, nous employons dans la présente recherche une définition limitée du voisinage phonologique. Selon cette définition, seuls les mots qui diffèrent du mot

cible *après substitution de la consonne finale* sont des voisins phonologiques du mot cible. Cette définition du voisinage est liée au fait que seul le dernier phonème du mot soit susceptible d'être modifié par l'assimilation. A titre d'exemple, le mot *jupe* possède, après substitution de la consonne finale, trois voisins : *juge*, *jute*, et *jure*.

En résumé, dans ce chapitre nous avons tout d'abord cherché à fournir une description acoustico-phonétique appropriée de l'assimilation de voisement en français. En second lieu, nous avons voulu étudier le rôle de deux facteurs lexicaux sur la production de l'assimilation afin de tester l'hypothèse suivante : la production de la parole devrait réduire l'importance des variations systématiques induites par l'assimilation, si celles-ci donnent lieu à un risque de confusion lexicale. Ceci est en particulier le cas pour les mots ambigus par assimilation de voisement : le mot *soute* peut effectivement être confondu avec son compétiteur *soude* par l'auditeur si sa consonne finale est totalement assimilée et donc voisée. Dans la même optique, la production devrait résister à n'importe quel type de variation, quand le mot cible à produire est difficilement reconnaissable de par sa nature même. Ceci est le souvent le cas pour les mots ayant beaucoup de voisins ou compétiteurs lexicaux, ou encore des compétiteurs de plus haute fréquence (cf. Grainger, O'Regan, Jacobs, & Segui, 1989). Dans les deux premières expériences de la série présentée ci-dessous, seul le facteur « ambiguïté par assimilation » a été manipulé. Dans l'Expérience 3, les facteurs « ambiguïté par assimilation » et « voisinage phonologique » ont été manipulés.

2. EXPERIENCE 1A □ PRODUCTION DE MOTS EN CONTEXTE ASSIMILATOIRE ET NON ASSIMILATOIRE

2.1. Matériel

Nous avons utilisé deux types de mots stimuli : des mots de type MAA (mots ambigus par assimilation) et MNA (mots non ambigus). Tous les stimuli étaient des mots monosyllabiques avec une obstruente en position finale. Il est à noter que les consonnes fricatives en position finale sont moins sujettes à des influences contextuelles que les obstruents, ce qui a été rapporté pour les formes de réduction et d'assimilation de place (cf. Kohler, 1995 ; Jun, 1995). Cette observation s'applique également pour l'assimilation de voisement en français (Duez, 1995) sauf dans le cas de quelques exceptions comme dans *je* suivi par [s] dans *je ne sais pas* > [ʃsɛpa], [ʃ:sɛpa] (cf. Tranel, 1983, et voir aussi Duez, 2001 qui a rapporté des

cas d'assimilation de voisement pour des consonnes fricatives). Chaque MAA fait partie d'une paire minimale comme (*soute*, *soude*) ou encore (*rate*, *rade*). Quatorze paires ont été sélectionnées dont les membres sont appariés au niveau de la fréquence (d'après la base de données BRULEX, Content, Mousty, & Radeau, 1990¹⁰). Au total il y avait donc 28 mots ambigus par assimilation. Trente-huit mots non ambigus ont été également choisis avec une fréquence moyenne proche de celle des mots ambigus. Chaque stimulus, MAA et MNA, était inséré dans un contexte phrastique, pouvant donner lieu ou non à un changement assimilatoire. Dans le cas des MAA, les deux membres d'une paire minimale étaient insérés dans le même contexte phrastique, comme dans « La *soute* (*soude*) pue vraiment fort » vs. « La *soute* (*soude*) ne sent pas si mauvais ». Les phrases en elles-mêmes étaient donc ambiguës par assimilation. Il y avait quatre phrases par paire ambiguë, (dorénavant quadruplet). La sélection initiale contenait un ensemble de 132 phrases (14 x 4 + 38 x 2). Dans un pré-test douze locuteurs de langue maternelle française ont eu pour tâche de juger la plausibilité sémantique et syntaxique de chaque phrase sur une échelle allant de 1 (non plausible) à 5 (très plausible). Cinq quadruplets contenant des MAA et 20 phrases contenant des MNA ont été retenus avec des jugements au-dessus de 3.8 (voir annexe 1 pour la liste complète du matériel expérimental). Les paires contenant les MAA étaient constituées de cinq mots avec pour consonne finale une obstruente voisée et leur cinq contreparties non voisées (fréquence moyenne 27.3 et 46.8 o.p.m. respectivement, d'après la base de données Lexique, New, Pallier, Ferrand & Matos, 2001). Parmi les 10 mots non ambigus, il y avait 7 mots avec pour consonne finale une obstruente non voisée et trois avec pour consonne finale une obstruente voisée. La fréquence moyenne de ces mots était de 27.5.

2.2. Locuteurs

Deux hommes et deux femmes ont participé à l'expérience. Ils avaient tous pour langue maternelle le français et ils avaient vécu dans la région parisienne pendant au moins une dizaine d'années. Au moment du déroulement de l'expérience, ils travaillaient tous à l'Institut de Psychologie de l'Université René Descartes (Paris 5).

2.3. Procédure

Les quatre locuteurs avaient pour tâche de lire chaque phrase de façon fluide et à un débit de parole normal. Ils se sont d'abord entraînés une ou deux fois avant l'enregistrement. Les

¹⁰ Lors de la sélection du matériel, la base de données *Lexique* n'était pas encore disponible en ligne.

locuteurs avaient pour consigne de ne pas laisser de silence entre les mots au sein de chaque phrase, alors qu'il pouvait y en avoir entre les phrases. Ils n'étaient pas au courant des objectifs de l'expérience. Les enregistrements se déroulaient dans une cabine insonorisée, en utilisant un microphone Sennheiser et un magnétophone Tascam DAT. Les productions étaient ensuite transférées à l'ordinateur (fréquence d'échantillonnage : 20 kHz) afin de les soumettre à des analyses acoustiques.

2.4. Analyses acoustiques

La présence d'une vibration des cordes vocales (désormais ViCV) est probablement l'indice acoustique le plus fiable permettant de mesurer le voisement en français et dans d'autres langues (Lisker & Abramson, 1964). Contrairement à ce que leur nom indique, les cordes vocales ne sont pas des « cordes » proprement dites, mais de petits muscles en forme de lèvres, attachés à la paroi fixe du larynx vers l'avant et aux deux aryténoïdes vers l'arrière. Pour une respiration normale, les cordes vocales sont écartées alors que pour la phonation, elles sont rapprochées, avec la glotte fermée (Thomas, Bouquiaux, & Cloarec-Heiss, 1976). D'autres indices de voisement ont également été identifiés (cf. Lisker & Abramson 1971). En anglais, les plosives non voisées (i.e., aspirées) en position initiale syllabique sont associées à des valeurs de fréquence fondamentale (F0) plus élevées à l'offset de voisement par rapport à des plosives voisées (i.e., non aspirées, Ohde, 1984 ; Jessen, 2001 ; voir aussi Stevens, Keyser, & Kawasaki, 1986, qui proposent le trait « [+stiff vocal folds] » pour indiquer des valeurs élevées de F0 dans la voyelle et qui s'ajoute au trait [-voice] pour les plosives non voisées en anglais). De plus, dans cette même langue les voyelles ont des durées plus longues quand elles sont suivies par des plosives voisées que par des plosives non voisées (Peterson & Lehiste, 1960 ; Raphael, 1972 ; Umeda, 1975). Quant à la durée d'occlusion, c'est l'inverse qui se produit : les durées d'occlusion sont typiquement plus longues devant des plosives non voisées que devant des voisées.

En français, les transitions de formants, l'intensité du relâchement, les valeurs de F0 à l'offset de voisement et la durée vocalique ont tous été proposés comme étant des paramètres acoustiques qui contribuent à la distinction de voisement (Duez, 1995 ; O'Shaughnessy, 1981 ; Saerens, Serniclaes, & Beeckmans, 1989 ; Wajskop, 1979 ; van Dommelen, 1983, 1985). Dans cette étude, nous avons fait les mesures acoustiques suivantes : (1) valeurs de F0 des voyelles précédant l'occlusion, (2) durée de la voyelle précédant l'obstruente, (3) durée de l'occlusion, et (4) durée de la portion voisée de l'occlusion (cf. van Dommelen, 1983 ; Warner, Sereno, Jongman, & Kemps, 2004, pour des mesures similaires). Dans nos mesures,

le début (*onset*) de l'occlusion est pris comme étant la fin (*offset*) acoustique de la voyelle précédente, ce qui est signalé par une baisse abrupte de l'amplitude et par la terminaison de la structure de formants. La fin de l'occlusion est identifiée à partir du début de l'explosion (*burst*). De plus, la fin de la portion voisée au sein de l'occlusion est identifiée à partir du point où aucun *pulse glottal* n'est observé (voir la Figure 2.2 pour une illustration). Afin d'établir laquelle de ces mesures acoustiques est la plus fiable pour distinguer les obstruents voisés des non voisés, leur puissance prédictive était comparée pour les obstruents en position finale des mots insérés dans des contextes phrastiques non assimilatoires, i.e. sous forme de réalisation canonique.

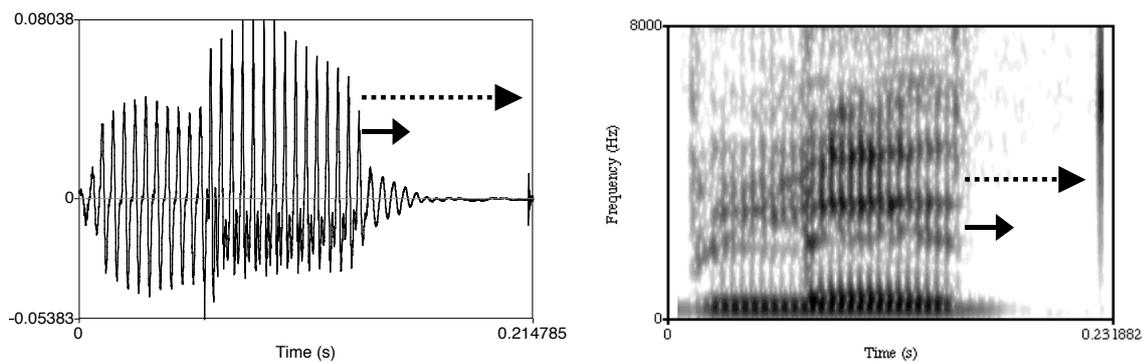


Figure 2.2. Le signal de parole et le spectrogramme correspondant d'une version du mot « blette » partiellement assimilé. La flèche pleine représente la portion voisée au sein de l'occlusion, la flèche pointillée représente l'occlusion.

Parfois l'obstruente en position finale ne montrait pas de relâchement quand elle était suivie par une autre obstruente, d'autant plus quand il s'agissait de deux consonnes homorganiques (comme dans *le sud de la France*), ce qui résultait en une séquence de consonnes géminées. Dans le cas des consonnes homorganiques, il ne se produit non seulement pas de changement de point d'articulation entre la première et la seconde consonne, mais pas non plus de changement de mode. La succession étant composée de deux consonnes strictement identiques, la désocclusion de la première consonne et l'occlusion de la seconde ne sont pas réalisées. L'absence de tout relâchement est observée dans 4% et 12% des cas pour les obstruents non voisés et voisés respectivement. Dans ces cas-là, la durée de l'occlusion était estimée en prenant 50% de la durée totale de l'occlusion. En outre, nous avons observé l'insertion d'une voyelle épenthétique dans 3% de toutes les productions.

2.5. Résultats acoustiques

Un résumé des mesures acoustiques sur les mots extraits des contextes phrastiques non assimilatoires est présenté dans le Tableau 2.1. Les résultats sont donnés en fonction de l'ambiguïté lexicale et du voisement sous-jacent. Comme les MAA font partie des paires minimales concernant le voisement en position finale, ces items se prêtent davantage à la comparaison des mesures acoustiques dans un environnement phonétique strictement contrôlé. Ceci n'est pas le cas pour les mots non ambigus. Pour le premier set de mots stimuli, les mesures sont en lien avec les conséquences connues pour le voisement des obstruents en position finale : les voyelles sont plus longues de 12% quand elles sont précédées par des obstruents voisés que par des obstruents non voisés ; les durées de l'occlusion sont quant à elles plus courtes devant des obstruents voisés que non voisés (12%). Toutefois, ces différences ne sont pas significatives.

		<i>Offset voyelle</i>	<i>Durée</i>	<i>Durée</i>	<i>Degré</i>
		<i>F0</i>	<i>vocalique</i>	<i>occlusion</i>	<i>voisement</i>
<i>MAA</i>	Non voisés	198 (67)	73 (26)	60 (17)	26 (22)
	Voisés	191 (66)	82 (22)	53 (15)	95 (22)
	□ (%)	-4%	+12%	-12%	+265%
<i>MNA</i>	Non voisés	181 (71)	61 (18)	60(16)	32 (25)
	Voisés	162 (59)	40 (18)	54 (14)	100 (0)
	□ (%)	-11%	-35%	-10%	+213%
<i>Tous les items</i>	Non voisés	188 (69)	66 (22)	60(16)	30 (24)
	Voisés	180 (64)	66 (29)	53 (14)	97 (18)
	□ (%)	-4%	0	-12%	+223%

Tableau 2.1. Valeurs moyennes des mesures acoustiques (écart-types entre parenthèses) pour les consonnes finales de 20 mots extraits des contextes non assimilatoires, produits par les quatre locuteurs. Les résultats sont présentés en fonction de l'ambiguïté potentielle (MAA vs. MNA) et du voisement sous-jacent (voisé vs. non voisé).

En ce qui concerne les valeurs F0 au début de la voyelle, il y avait une légère tendance à avoir des valeurs plus basses de F0 devant les obstruents voisés (~4%), mais cette tendance était loin d'être significative. Les mêmes tendances sont obtenues pour les mots non ambigus, sauf en ce qui concerne la durée vocalique : la durée des voyelles précédant les obstruents non voisés était en moyenne plus élevée que celle des voyelles précédant les obstruents voisés.

Toutefois, comme nous pouvons le constater à partir du Tableau 2.1, il y avait une quantité d'environ 30% de ViCV présente dans les obstruents non voisés (extraites des contextes non assimilatoires). Aussi étrange que cela puisse paraître au premier abord, il existe une explication plausible pour la présence de ViCV dans les obstruents non voisés. Le début de la réalisation de ViCV a une nature dynamique et quelque peu différée par rapport aux résonances de l'appareil vocal reliées à la voyelle précédente. C'est la raison pour laquelle les cordes vocales peuvent encore continuer de vibrer pendant quelque temps dans la portion initiale de l'occlusion¹¹.

Dans ce qui suit, nous appellerons « degré de voisement » la durée relative de la portion voisée au sein de l'occlusion. Ceci amène à une dérivation directe du degré de voisement à un degré d'assimilation de voisement (suivant Nolan, 1992, exprimé en pourcentage) : c'est le pourcentage de la portion voisée de l'occlusion dans les obstruents non voisés ('100 x « degré de voisement »') ou bien le pourcentage de la portion non voisée de l'occlusion dans les obstruents voisés ('100 x (1- « degré de voisement »)'). Les analyses rapportées dans les expériences suivantes sont basées sur ces indices de degré d'assimilation.

Le degré d'assimilation, comme étant défini ci-dessus, était calculé pour chaque obstruente en position finale pour les mots critiques insérés dans des contextes phrastiques assimilatoires. Une différence importante était observée entre les MAA et MNA, 25.2% et 77.8% respectivement, $t(78) = 7.59$, $p < .00001$. Néanmoins, les stimuli non ambigus n'étaient pas balancés par rapport au voisement sous-jacent. Il y avait plus de mots avec pour consonne finale une obstruente non voisée que voisée. Cela signifie que le matériel utilisé ne permet pas d'évaluer l'effet de l'ambiguïté lexicale indépendamment du voisement des obstruents en position finale. La Figure 2.3 montre la distribution du degré d'assimilation pour un ensemble de 80 stimuli. Ces résultats sont présentés en fonction du voisement sous-jacent (12 obstruents non voisés, 8 obstruents voisés, produites par quatre locuteurs). Dans la Figure

¹¹ Le mot *trac* était produit avec une consonne finale complètement voisée [g] par deux locuteurs (degré de voisement = 100%). Il y avait une variation non négligeable concernant le degré de voisement pour les consonnes non voisés à travers les 4 locuteurs : L1 : [32%-63%], L2 : [0-41%], L3 : [0-30%], L4 : [31%-51%]. Ces différences peuvent en partie être expliquées en termes de style d'articulation et de débit de parole. Il est également à noter que nos locuteurs étaient originaires de régions différentes de France (même s'ils avaient vécu en région parisienne depuis une dizaine d'années). Quant aux consonnes voisés, elles étaient toutes produites avec un degré de voisement de 100%, à l'exception d'un item *bled* prononcé par un locuteur comme [blɛt] (V-ratio = 0%).

3, cinq niveaux de degré d'assimilation sont indiqués : le niveau d'assimilation négligeable [0-20%], trois niveaux intermédiaires d'assimilation [20-40%], [40-60%], et [60-80%], et finalement le niveau d'assimilation (quasi-) complète [80-100%].

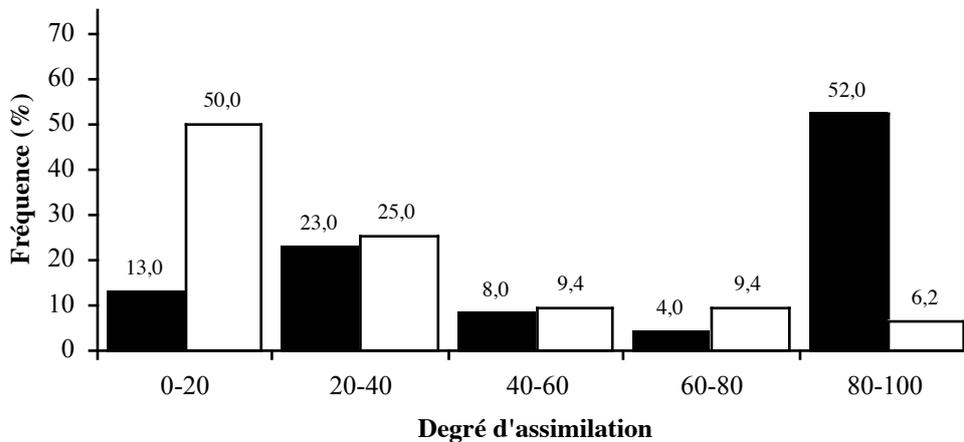


Figure 2.3. Distribution du degré d'assimilation pour les consonnes obstruantes non voisées (barres noires) et voisées (barres blanches) position finale, extraites des contextes assimilatoires. La distribution est calculée à partir de 80 stimuli (4 locuteurs x 20 mots stimuli).

Ces données révèlent deux choses sur la nature de l'assimilation de voisement en français. Premièrement, si la distribution fait apparaître deux points culminants sur les points finaux du continuum du degré d'assimilation, un pourcentage important (i.e., 39%) des items ayant des valeurs intermédiaires d'assimilation (dans l'intervalle [20-80%]). Autrement dit, l'assimilation en français ne constitue pas toujours un processus phonétique catégorique ou « tout-ou-rien », mais plutôt gradué. Deuxièmement, nous pouvons observer une asymétrie saisissante entre les consonnes obstruantes non voisées et voisées : en moyenne, le degré d'assimilation est de 67.7% pour les non voisées alors qu'il était de 27.3% pour les voisées, $t(78) = 4.96, p < .0001$ ¹². Cela veut dire que les obstruantes non voisées sont plus fortement assimilées que leurs contreparties voisées (cf. Kohler, 1979, pour des observations similaires en allemand).

¹² Dans certains cas, le voisement semblait se poursuivre au sein de l'occlusion, s'arrêtait pendant un instant pour continuer ensuite avant la fin de l'occlusion. Dans ces cas-là, l'intervalle secondaire de voisement n'était pas pris en compte dans les analyses quantitatives, ce qui amenait à une sous-estimation négligeable du degré de voisement. Si nous avons choisi de prendre en compte le second intervalle de voisement, l'asymétrie entre voisés et non voisés serait encore plus nette.

Les résultats présentés plus haut sont basés sur l'indice original de voisement que nous avons proposé, fournissant une mesure d'assimilation pour consonnes obstruantes non voisées et voisées. Bien que d'autres paramètres acoustiques puissent être mis en avant, la durée relative de voisement à l'intérieur de l'occlusion était de loin l'indice acoustique le plus fiable du degré de voisement (cf. van Dommelen, 1983). La durée de l'occlusion et de la voyelle se sont fréquemment avérées comme étant des indices fiables de voisement, au moins dans les cas de neutralisation incomplète de voisement, pour les langues telles que le néerlandais ou encore l'allemand (Ernestus & Baayen, sous presse ; Jessen, 2001 ; Warner, et al. 2004). Dans nos données, la pertinence de ces paramètres est seulement observée pour les MAA, i.e. pour les mots figurant dans des contextes phonétiques contrôlés (e.g., *soude* / *soute*) et encore l'effet de ces paramètres manque de robustesse. Ceci pourrait être dû aux différences intrinsèques liées à l'implémentation phonétique du trait [\pm voice] en français par rapport à d'autres langues. Quant à notre indice, celui-ci est capable de capturer les degrés intermédiaires de voisement entre complètement voisé et complètement non voisé, quel que soit le contexte phonétique (l'indice s'avère conforme pour les mesures de voisement des MAA et MNA). Si l'indice montre une certaine cohérence d'un point de vue acoustico-phonétique, il est essentiel de savoir si tel est également le cas d'un point de vue perceptif. L'asymétrie considérable entre voisés et non voisés justifie davantage un test perceptif : est-ce que cette asymétrie reflète une véritable asymétrie concernant le sens de la direction de l'assimilation de voisement, ou s'agit-il d'un biais asymétrique dans la mesure acoustique que nous avons utilisée permettant de distinguer les segments voisés des non voisés ? Si l'asymétrie a une réalité perceptive, les segments non voisés devraient être perçus comme étant des voisés plus fréquemment que les segments voisés ne sont perçus comme des non voisés. Cette observation attendue reflèterait ainsi une assimilation plus importante pour les obstruantes non voisées que pour les voisées.

L'objectif de l'Expérience 1b était de tester la pertinence perceptive de l'indice d'assimilation que nous avons employé jusqu'ici, pour les degrés extrêmes et intermédiaires d'assimilation de voisement. Les mots stimuli utilisés dans l'Expérience 1a étaient extraits de leur contexte phrastique et ensuite présentés en isolation à des participants français qui avaient pour tâche de catégoriser la consonne en position finale.

3. EXPERIENCE 1B □ CATEGORISATION PERCEPTIVE DES MOTS ASSIMILES ET NON ASSIMILES

3.1. Matériel expérimental

Tous les mots stimuli prononcés pour l'Expérience 1a étaient de nouveau utilisés dans l'Expérience 1b. Dix mots ambigus par assimilation et dix mots non ambigus étaient insérés dans des contextes assimilatoires (i.e. donnant lieu à une assimilation de voisement de la consonne en position finale du mot critique) et non assimilatoires, d'où un nombre de 40 mots stimuli (20 mots x 2 contextes). Il y avait 4 locuteurs dans la première expérience, donc il y avait 160 mots stimuli au total. Chaque mot critique était extrait de son contexte phrastique. Nous avons pris soin d'éviter des « clicks » audibles au début ou à la fin du mot. Pour les mots avec pour consonne initiale une consonne obstruente non voisée, le point de départ d'extraction était déterminé en plein milieu du silence d'occlusion de la plosive. Dans tous les autres cas de figure (i.e. consonnes sonorantes, fricatives, plosives orales voisées), le point de départ était déterminé comme étant le point de stabilité spectrale soit de la consonne initiale en entier (donnant lieu à une troncation) soit comme étant le « murmure » voisé dans l'occlusion. Le point final de l'extraction constituait dans tous les cas la fin du relâchement de l'explosion (à un passage zéro le plus proche). Le signal de parole extrait était ensuite atténué par une rampe d'atténuation de 8ms. L'intensité acoustique était égalisée à travers stimuli et locuteurs.

Nous avons construit deux listes expérimentales, par une attribution aléatoire des contextes assimilatoires et non assimilatoires à une des deux listes. Les deux versions de chaque MAA (issues de deux contextes phrastiques), étaient attribuées à l'une des deux listes expérimentales et les deux versions de leur contrepartie voisée (ou non voisée) étaient attribuées à l'autre liste restante. La répartition du matériel expérimental est illustrée dans le Tableau 2.2.

		<i>Liste 1</i>		<i>Liste 2</i>		
<i>Type</i>	<i>Test</i>	<i>Contexte</i>	<i>Exemple</i>	<i>Test</i>	<i>Contexte</i>	<i>Exemple</i>
<i>MAA</i>	[-voice]	[-voice]	soute pue	[+voice]	[+voice]	soude ne
	[-voice]	[+voice]	soute ne	[+voice]	[-voice]	soude pue
	[+voice]	[-voice]	vide ne	[-voice]	[-voice]	vite sans
	[+voice]	[+voice]	vide sans	[-voice]	[+voice]	vite ne
<i>MNA</i>	[-voice]	[-voice]	frites très	[-voice]	[+voice]	frites de
	[+voice]	[-voice]	tube passe	[+voice]	[+voice]	tube de

Tableau 2.2. Répartition du matériel expérimental sur deux listes contrebalancées (Expérience 1b). [\pm voice] fait référence au voisement sous-jacent des consonnes critiques et consonnes suivantes.

3.2. Sujets

Vingt-deux étudiants / doctorants (sept hommes et 15 femmes, âge moyen = 25 ans) de l'Institut de Psychologie de l'Université René Descartes (Paris 5) ont participé à l'expérience. Tous les sujets avaient pour langue maternelle le français et ils avaient tous une audition correcte et une vision correcte (ou corrigée).

3.3. Procédure

Les sujets étaient assis devant l'écran d'ordinateur dans une salle silencieuse et ils ou elles ont reçu une consigne écrite et des instructions orales. Ensuite, les mots stimuli étaient présentés auditivement par écouteurs (Sennheiser) à un niveau d'écoute confortable. Les sujets avaient pour tâche de catégoriser la consonne en position finale du stimulus qu'ils entendaient et ils devaient pour cela choisir parmi un ensemble de huit consonnes proposées sur l'écran d'ordinateur (/p, t, k, b, d, g, s, z/). A chaque essai, ils devaient également donner un jugement subjectif de confiance sur une échelle allant de 1 (peu confiant) à 5 (très confiant). Il y avait deux phases d'entraînement qui précédaient la phase test proprement dite. Pendant la première phase d'entraînement, les participants devaient catégoriser la consonne en position finale des mots n'ayant subi aucun changement phonologique (prononcés en isolation). Lors de cette phase, les sujets ont reçu un feedback indiquant s'ils avaient ou non répondu correctement. Aucun jugement subjectif n'était demandé. Pendant la deuxième phase d'entraînement, seuls des non-mots étaient présentés afin de décourager l'emploi de stratégies lexicales pendant la

catégorisation ; lors de cette phase les participants n'ont plus reçu de feedback mais ils devaient juger leurs réponse sur une échelle. Après l'expérience, les participants ont rempli un questionnaire portant sur leurs connaissances linguistiques (i.e., de langues étrangères). La durée de l'expérience était environ de 20 minutes.

3.4. Résultats

Dans ce qui suit, nous désignerons par « congruentes » les réponses des participants qui correspondent au voisement *sous-jacent* de la consonne cible, quel que soit le lieu d'articulation. Le lieu d'articulation, mais non pas le mode, était parfois mal perçu. Le taux d'erreurs en moyenne était de 10.8%. Toutefois, si par exemple la consonne assimilée ayant pour forme sous-jacente un /p/ donnait lieu à une réponse /t/, nous avons quand même pris en compte cette réponse par son caractère non voisé « congruent » avec le caractère non voisé sous-jacent du /p/. Les résultats peuvent être résumés, pour chacun des mots stimuli, par un pourcentage de réponses congruentes, c'est-à-dire le pourcentage des participants ayant rapporté le voisement *sous-jacent* de la consonne cible dans le mot stimulus. L'ensemble des résultats sont présentés en fonction du type de voisement sous-jacent (voisé vs. non voisé), type de mot (MAA vs. MNA) et type de contexte (assimilatoire vs. non assimilatoire) dans le Tableau 2.3.

Les pourcentages de réponses congruentes étaient mis en relation avec les degrés d'assimilation calculés dans l'Expérience 1a. Si cet indice acoustique est doté d'une pertinence perceptive, alors nous devrions nous attendre à ce qu'un degré d'assimilation élevé induise un pourcentage plus bas de réponses « congruentes » qu'un degré d'assimilation faible. En d'autres termes, nous devrions observer une corrélation négative entre l'indice acoustique d'assimilation et la réponse perceptive du voisement sous-jacent, pourvu qu'il y ait suffisamment de variation dans le degré d'assimilation pour qu'une telle corrélation ait un sens statistique. Ce dernier critère a été pris en compte dans la sélection des stimuli extraits du contexte assimilatoire. Dans cet ensemble, nous avons observé une corrélation négative et statistiquement fortement significative entre l'indice acoustique indiquant le degré d'assimilation et le pourcentage de réponses « congruentes », $r(78) = -0.83$, $p < .0001$. Ce niveau élevé de corrélation nous permet ainsi d'établir une réalité psychologique de l'indice acoustique du degré d'assimilation, qui était utilisé afin de mesurer l'assimilation des consonnes obstruées en position finale.

Type	Voisement	Contexte		Nombre d'items
		~ Assimilatoire	Assimilatoire	
MAA	[-voice]	83.2	70	20
	[+voice]	93.5	90.9	20
MNA	[-voice]	84.2	14.3	28
	[+voice]	76.4	62.9	12
Moyenne pondérée		85.1	54.7	

Tableau 2.3. Pourcentages de réponses « congruentes » (i.e., réponses consistantes avec le voisement sous-jacent) selon le type de mots (MAA vs. MNA), type de contexte (non assimilatoire vs. assimilatoire) et type de voisement sous-jacent (non voisés [-voice] vs. voisés [+voice]).

Comme nous pouvons constater dans le Tableau 2.3, l'ensemble des résultats est cohérent avec le fait que les jugements perceptifs entretiennent une corrélation négative avec l'indice acoustique du degré d'assimilation, et reflètent ainsi les mesures acoustiques rapportées dans l'Expérience 1a. Le pourcentage de réponses « congruentes » est systématiquement plus élevé pour les items extraits des contextes non assimilatoires que des contextes assimilatoires. Seules les données sur les mots assimilés sont soumises à des analyses plus approfondies, c'est-à-dire les données obtenues dans les contextes assimilatoires. Les MAA sont associés à des pourcentages de réponses congruentes plus élevés que les MNA (80.5% vs. 28.9%, $t(78) = 7.02$, $p < .0001$). De même, les segments voisés sous forme canonique ont donné des pourcentages de réponses congruentes plus élevés que les segments non voisés sous forme canonique (80.4% vs. 37.5%, $t(78) = 5.20$, $p < .0001$) ce qui est compatible avec l'observation d'un degré d'assimilation plus élevé pour les non voisés par rapport aux voisés (cf. Expérience 1a, 67.7% vs 27.3% respectivement). Le pourcentage le plus bas de réponses congruentes (14.1%) était observé pour les consonnes non voisées en position finale des mots non ambigus.

Les résultats pour la catégorisation phonétique sont compatibles avec les données acoustiques concernant l'asymétrie en terme d'ampleur de l'assimilation entre les non voisés et voisés. Ces résultats portent à croire que l'indice acoustique du degré d'assimilation peut rendre compte de la perception des segments assimilés par (dé)voisement. Toutefois, il est important de noter qu'une partie du matériel utilisé dans les Expériences 1a-b n'était pas équilibré par rapport au voisement sous-jacent en position finale : il y avait sept mots non voisés et trois mots voisés dans les mots non ambigus (en revanche, le matériel ambigu était contrôlé quant au voisement sous-jacent par défaut.) Qui plus est, les consonnes finales des mots non

ambigus n'étaient pas équilibrées en termes de place d'articulation de la consonne subséquente, ni du type de voyelle précédant la consonne finale, deux facteurs qui sont censés affecter la durée de la vibration des cordes vocales (voir Ohala, 1983). Le contexte donnant lieu à l'assimilation n'était pas un facteur maintenu constant, et comprenait un mélange de consonnes plosives, fricatives et nasales. C'est pour toutes ces raisons que nous avons conduit une expérience de production supplémentaire dans laquelle le matériel était strictement équilibré par rapport au voisement sous-jacent, au lieu d'articulation des consonnes subissant une assimilation, et au contexte phonétique donnant lieu à l'assimilation. De plus, cette fois-ci tous les locuteurs étaient originaires de la région parisienne.

4. EXPERIENCE 2 □ PRODUCTION DE MOTS EN CONTEXTE ASSIMILATOIRE

4.1. Matériel expérimental

Deux ensembles de 20 mots monosyllabiques chacun ont été sélectionnés: la consonne obstruente en position finale était voisée dans un jeu de mots et non voisée dans l'autre. Les mots voisés et non voisés étaient équilibrés par rapport à la fréquence (38.71 o.p.m. pour les mots voisés, 39.23 o.p.m. pour les non voisés). Cette fois-ci, les mots étaient insérés non pas dans des phrases entières (cf. Expérience 1) mais dans des syntagmes nominaux de type Det-Nom-Adj (e.g., *une jupe grise*). Le contexte était toujours assimilatoire, i.e., la consonne critique et la consonne suivante était de voisement différent. Dans chaque ensemble de 20 syntagmes, la moitié contenait un contexte fricatif droit, alors que l'autre moitié contenait un contexte plosif droit (contexte voisé : /v,ʒ/ ; contexte non voisé /f,s /). Quarante mots de remplissage ont été rajoutés au matériel expérimental (voir matériel en annexe 1).

4.2. Locuteurs

Cinq hommes et cinq femmes, tous étudiants à l'Institut de Psychologie à l'Université René Descartes ont participé à l'expérience. Ils étaient de langue maternelle française, ainsi que de la région parisienne.

4.3. Procédure

L'expérience se déroulait dans une chambre insonorisée. Quarante-vingts syntagmes nominaux étaient présentés sur des feuilles et les participants avaient pour tâche de prononcer les syntagmes de façon continue et fluide, sans laisser de pause entre les mots au sein d'un syntagme. Les participants n'étaient pas au courant du but de l'expérience. Les

enregistrements ont été effectués sur un magnétophone DAT et ensuite transférés sur ordinateur pour des analyses acoustiques.

4.4. Résultats acoustiques

La Figure 2.4 montre les distributions des degrés d'assimilation observés pour les consonnes en position finale suivies par une consonne plosive (distribution à gauche) ou par une fricative (distribution à droite). Ces distributions montrent de nouveau que l'assimilation de voisement ne correspond pas toujours à un phénomène discret, mais qu'il s'agit plutôt d'un processus gradué. En outre, nous n'avons pas trouvé de différence significative entre les degrés d'assimilation provoqués par les consonnes plosives ou fricatives. Enfin, l'asymétrie observée dans l'expérience précédente entre les consonnes non voisées et voisées s'avère confirmée : le degré d'assimilation était significativement plus important pour les consonnes non voisées (degré moyen de 75%) que pour les consonnes voisées (degré moyen de 29%), $t(19) 9.34, p < .001$.

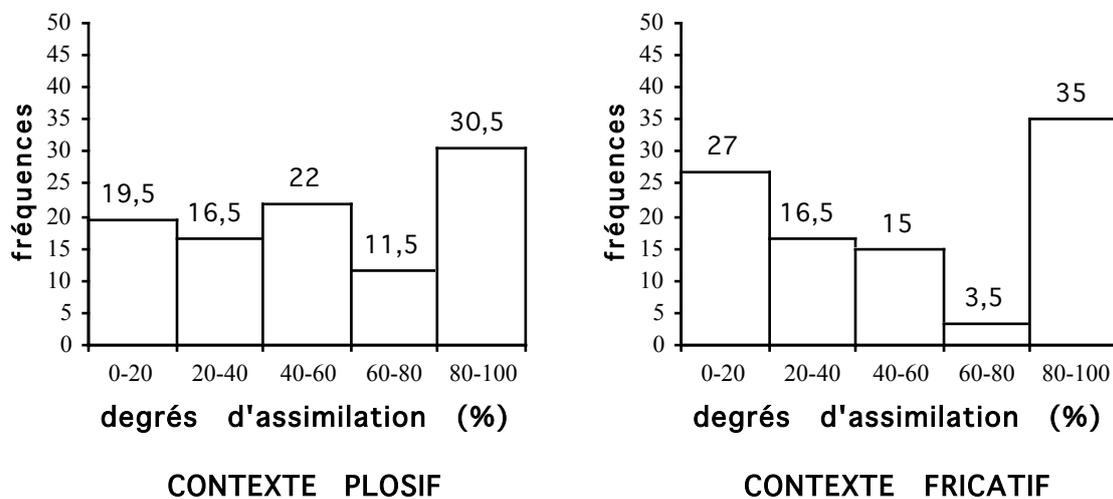


Figure 2.4. Distribution des degrés d'assimilation des consonnes terminales suivies par des plosives (à gauche) ou par des fricatives (à droite).

4.5. Discussion des Expériences 1-2

Les données acoustiques et perceptives des Expériences 1a et 1b ont clairement suggéré que la durée relative de voisement au sein de l'occlusion fournit une mesure assez robuste permettant d'évaluer le degré de voisement pour les consonnes obstruentes en position finale en français. Qui plus est, les jugements perceptifs ont montré que cet indice dispose d'une certaine pertinence perceptive. Dans le matériel que nous avons employé dans nos expériences, d'autres paramètres acoustiques tels que les durées vocaliques ou d'occlusion

n'ont pas permis de mesurer le voisement de façon fiable. Il se peut qu'un nombre important d'items et/ou de locuteurs soit nécessaire afin que ces indices puissent diagnostiquer les différences phonétiques de voisement, comme l'avaient suggéré Warner et al. (2004). Il se peut également que la base acoustico-phonétique principale du trait de voisement [voice] soit avant tout représentée par la présence de la vibration des cordes vocales, alors que les autres indices acoustiques seraient secondaires (van Dommelen, 1985 ; voir aussi van den Berg, 1986, 1989, pour des conclusions similaires sur la production et perception du voisement en néerlandais).

Après avoir proposé un indice fiable pour mesurer le degré d'assimilation des consonnes obstruantes en position finale, les données acoustiques dans l'Expérience 1a ont permis de mettre en évidence deux aspects de l'assimilation de voisement en français. Premièrement l'assimilation n'est pas toujours un processus complet. Plutôt, il s'agit d'un processus gradué dans la mesure où il existe des degrés intermédiaires de voisement dans les segments assimilés. De telles données empiriques posent un défi aux théories phonologiques selon lesquelles l'assimilation doit être considérée comme un processus « tout-ou-rien » (cf. Rigault, 1967 ; Casagrande, 1984). Ces résultats sont plutôt en accord avec d'autres données, mettant en avant la nature graduée de l'assimilation de place en anglais (cf. Nolan, 1992 ; Gow, 2002, 2003).

Un deuxième résultat, plutôt inattendu, émergeait des données d'assimilation : les obstruantes non voisées, telle que le /t/ dans *frites*, sont plus fréquemment et complètement assimilées que les obstruantes voisées, telle que /b/ dans *tube*. Cette asymétrie est aussi observée dans les jugements perceptifs. Une expérience supplémentaire de production (Expérience 2) a permis de démontrer que cette asymétrie persistait entre les segments non voisés et voisés avec un matériel linguistique plus équilibré au niveau du voisement sous-jacent et lieu d'articulation de la consonne critique. En outre, le degré d'assimilation n'était pas modulé par le contexte donnant lieu au changement : les consonnes plosives et fricatives ont provoqué des degrés d'assimilation similaires (cf. van Dommelen, 1985).

Le troisième résultat principal concerne l'effet de l'ambiguïté lexicale. Nous avons observé un degré d'assimilation plus élevé pour les mots non ambigus tel que *jupe* (77.8%) par rapport aux mots ambigus par assimilation tels que *soude* et *soute* (25.2%). Il est tentant de supposer que les locuteurs cherchent à préserver le voisement sous-jacent si un changement de voisement pouvait déclencher une confusion lexicale. En effet, ce serait une illustration de l'adaptation de production à la perception anticipée, telle que proposé par Lindblom (1990).

Pourtant, il faut d'abord exclure la possibilité que l'asymétrie entre MAA et MNA ne relève pas de la méthodologie expérimentale. En effet, nous avons déjà mentionné qu'il y avait plus de consonnes non voisées que voisées pour ce qui est du matériel non ambigu. Ceci pourrait en partie expliquer la différence entre ambigus et non ambigus étant donné que les non voisés se voient plus fréquemment que les voisés ne se dévoient. De plus, dans l'Expérience 1a, les locuteurs se sont entraînés sur le matériel expérimental avant l'enregistrement et ils ont dû lire des quadruplets qui consistaient en deux membres d'une paire minimale (*soute / soude*) insérés dans deux phrases identiques. Même si nous avons pris soin de séparer autant que possible les deux phrases identiques contenant un des deux membres de la paire minimale dans la présentation des stimuli, il se pourrait que les locuteurs aient pris conscience de la présence des items lexicaux ambigus par assimilation. Cette conscience aurait pu induire une articulation plus soignée des items ambigus, d'où un degré moins élevé d'assimilation. Il semble donc approprié de conduire une nouvelle expérience dans laquelle les conditions d'enregistrement soient plus contrôlées afin d'évaluer plus précisément l'influence de l'ambiguïté lexicale sur la production des segments assimilés.

La possibilité que la production de la parole assimilée soit modulée par les connaissances implicites sur la perception mérite la réalisation d'une expérience à part entière, ce qui constitue la motivation de l'Expérience 3. L'hypothèse que nous voudrions tester est la suivante : la production de la parole devrait réduire l'importance des variations contextuelles si ces variations donnent lieu à un risque de confusion lexicale. Ceci est en particulier le cas pour les mots ambigus par assimilation de voisement. Le mot *soute* peut effectivement être confondu avec son compétiteur *soude* par l'auditeur si sa consonne finale est totalement assimilée et donc voisée. Dans la même optique, la production devrait résister à n'importe quel type de variation, quand le mot cible à produire est difficilement reconnaissable par sa nature même. Ceci est le souvent le cas pour les mots ayant beaucoup de compétiteurs lexicaux, ou encore des compétiteurs de plus haute fréquence. Dans l'Expérience 3, nous avons voulu examiner le rôle potentiel du voisinage phonologique sur la production des obstruents en position finale. Le type de voisins phonologiques qui nous intéresse dans cette étude est le nombre de voisins obtenu après substitution de la consonne finale, car soumise à l'assimilation et non pas tous les mots obtenus après un changement phonémique (cf. Luce, Pisoni, & Goldinger, 1990). A titre d'exemple, le mot *jupe* possède trois compétiteurs au sens étroit de voisinage phonologique : *juge*, *jure*, et *jute*. Notons que jusqu'ici le mot *jupe* était pourtant considéré comme étant non ambigu par rapport au voisement final car *jube* ne

constitue pas un mot en français. La production des mots tels que *jupe* permet donc d'évaluer l'influence de la compétition lexicale, tout en écartant le risque de confusion avec l'ambiguïté dû au voisement de la consonne en position finale. Il a été démontré par ailleurs que le nombre de voisins (ou densité) phonologiques peut affecter la reconnaissance auditive de la parole. Par exemple, Luce et Pisoni (1998) ont employé le paradigme d'identification perceptive dans lequel les sujets avaient pour tâche d'identifier des mots dans du bruit. Les mots à faible densité phonologique étaient reconnus plus rapidement que les mots à forte densité. Si le voisinage phonologique influence la production de la même façon que la perception de la parole, nous devrions nous attendre à ce que les mots à forte densité phonologique soient articulés avec plus de soin que les mots à faible densité afin de compenser la difficulté qu'il peut y avoir en perception. Ainsi, nous devrions observer un degré d'assimilation plus bas pour les mots à forte densité que pour les mots à faible densité phonologique.

5. EXPERIENCE 3A □ PRODUCTION DE MOTS DANS DES CONTEXTES ASSIMILATOIRES ET NON ASSIMILATOIRES □ ROLE DE L'AMBIGUITE ET DU VOISINAGE PHONOLOGIQUE

L'Expérience 3 était conçue pour étudier de manière plus approfondie les effets déjà trouvés dans l'Expérience 1. Un matériel linguistique sélectionné est similaire à celui utilisé dans l'expérience précédente. Toutefois, il a été conçu de façon à mieux contrôler les facteurs en jeux. Premièrement, un nombre égal de mots non voisés et voisés a été choisi dans le matériel non ambigu afin de réexaminer l'asymétrie entre les obstruents non voisés et voisés. Deuxièmement, le rôle potentiel de l'ambiguïté lexicale sera de nouveau évalué. Nous avons voulu éviter que les participants aient conscience de la présence des items phonologiquement similaires. Cette fois-ci, les participants se sont familiarisés avec des phrases de remplissage mais pas avec les phrases expérimentales comme dans l'Expérience 1a. Les mots critiques étaient insérés dans les contextes phrastiques assimilatoires afin que chaque mot critique apparaisse une seule fois dans le matériel expérimental. Finalement, une population de locuteurs plus homogène a été sélectionnée par rapport à l'âge et au dialecte régional.

L'Expérience 3 était également conçue afin de considérer le rôle de deux facteurs lexicaux qui influent potentiellement sur l'articulation des mots afin d'éviter toute confusion lexicale : l'ambiguïté lexicale (comme dans l'Expérience 1) et la densité du voisinage phonologique, les

deux facteurs se focalisant sur la consonne terminale des mots critiques. Ces deux facteurs étaient manipulés de façon orthogonale.

5.1. Matériel expérimental

Comme dans l'Expérience 1, tous les mots étaient monosyllabiques et avaient pour consonne finale une obstruente. Dans un premier temps, nous avons choisi 18 paires minimales (par rapport au voisement de la consonne finale). Ces paires formaient 36 mots ambigus par assimilation qui étaient appariées autant que possible du point de vue de leur fréquence d'occurrence (5.21 et 8.74 pour les mots non voisés et voisés respectivement). Comme la fréquence subjective a parfois été considérée comme un meilleur prédicteur d'accès au lexique que les fréquences objectives issues des bases de données (cf. Gordon, 1985), nous avons demandé à 15 participants français –étudiants de première année en Psychologie– de juger la fréquence des 36 mots potentiellement ambigus et 36 mots de remplissage sur une échelle de 1-à-5. Ces étudiants n'ont pas participé aux expériences antérieures. Douze paires dont les membres étaient jugés proches en termes de fréquence subjective, ont été retenues. Les jugements de fréquence moyens étaient 3.11 et 3.06 pour les mots non voisés et voisés respectivement (la différence n'était pas significative, $t(22) = 0.57$, n.s.). Les fréquences objectives correspondantes (d'après la base de données Lexique), étaient de 6.7 et 7.2 respectivement. Vingt-quatre mots non ambigus étaient ensuite sélectionnés dans la même gamme de fréquence (8.43). La liste avec les 48 mots stimuli, dont la moitié était ambiguë par assimilation et l'autre moitié ne l'était pas, est présentée dans l'annexe 1. Le deuxième facteur lexical, le voisinage phonologique, était manipulé de manière orthogonale afin que les trois niveaux de densité soient croisés avec l'ambiguïté lexicale et le type de voisement sous-jacent. Le voisinage phonologique pouvant être défini selon des critères différents (Bailey & Hahn, 2001 ; Luce & Pisoni, 1998 ; Luce, Pisoni, & Goldinger, 1990), nous nous limitons ici à une définition quelque peu approximative comme le nombre de mots dérivés du mot cible par substitution phonémique. De plus, comme nous avons pour objectif de tester l'influence du voisinage sur la production des consonnes en position finale, la définition pertinente du voisinage phonologique semble être l'ensemble de compétiteurs du mot cible après substitution de sa consonne finale. A titre illustratif, notre définition du voisinage n'inclut pas le mot *jappe* [ʒap] comme étant un compétiteur de mot cible *jupe* [ʒyp], car la différence phonémique entre ces deux mots est déterminée par la voyelle et non par la consonne finale. En prenant en compte cette définition stricte du voisinage, la densité pour les mots stimuli

était vérifiée dans la base de données *Vocolex* (Dufour, Peereman, Pallier, & Radeau, 2002). Les 48 mots test étaient distribués à travers trois niveaux de densité phonologique : faible (2-7 voisins), moyen (8-11 voisins), et élevé (12-17 voisins). La densité phonologique était croisée avec l'ambiguïté lexicale (2 niveaux) et type de voisement sous-jacent (2 niveaux) afin d'avoir un nombre égal de mots stimuli dans chaque cellule factorielle. Le plan factoriel, avec des exemples des mots stimuli, est illustré dans le Tableau 2.4.

<i>Ambiguïté</i>	<i>Voisement</i>	<i>Densité phonologique</i>		
		[2-7]	[8-11]	[12-17]
<i>MNA</i>	[-voice]	gratte	bec	bac
	[+voice]	grade	bègue	bague
<i>MNA</i>	[-voice]	floc	bouc	botte
	[+voice]	vogue	digue	rabe

Tableau 2.4. Exemples pour les différents types de mots utilisés dans l'Expérience 3.

Quarante-huit mots de remplissage, tous appariés en termes de fréquence et de voisinage phonologique avec les mots test, ont été rajoutés au matériel expérimental. Les mots test et les mots de remplissage ont ensuite été insérés dans des contextes phrastiques de type *J'ai dit... vingt fois* ou encore *J'ai dit...cent fois*. Les mots test figuraient toujours dans des contextes assimilatoires ; une consonne finale voisée était toujours suivie d'une consonne fricative non voisée (/s/ dans *cent*), une consonne non voisée était suivie d'une consonne fricative voisée (/v/ dans *vingt*). Nous avons voulu insérer les mots test expressément dans des contextes phrastiques sémantiquement neutres afin de réduire la probabilité que les locuteurs prennent conscience de la présence des paires minimales et adoptent des stratégies de locution.

5.2. Locuteurs

Huit étudiants (quatre hommes et quatre femmes, âge moyen = 25 ans), tous étudiants à l'Institut de Psychologie, ont participé à l'expérience. Ils avaient tous le français comme langue maternelle et pour dialecte le parisien (région Ile-de-France).

5.3. Procédure

Les locuteurs avaient pour consigne écrite de lire à voix haute des phrases test en français. Ils n'étaient pas au courant des objectifs de l'expérience. Dans la consigne, nous avons insisté sur le fait de prononcer les phrases à un débit de parole normal mais fluide, sans laisser de pauses entre les mots. Avant l'enregistrement, nous avons présenté auditivement quatre phrases

prononcées par un locuteur masculin non naïf afin d'illustrer le type de locution voulue. Les locuteurs se sont d'abord « entraînés » sur dix phrases ne faisant partie du matériel expérimental et ils recevaient un feedback par rapport au débit de parole. Nous avons effectué l'enregistrement à l'aide du même dispositif que dans l'Expérience 1. L'enregistrement (fréquence d'échantillonnage = 44 kHz) a ensuite été transférée sur ordinateur. A la fin de l'expérience, nous avons demandé aux participants s'il ou elle avait remarqué quelque chose de particulier sur les phrases test. Naturellement, tous nos locuteurs se sont aperçus que les contextes phrastiques étaient toutes identiques, mais personne n'avait remarqué la présence des paires minimales (e.g., *soude / soute*).

5.4. Résultats acoustiques

Pour chaque item, la durée relative du voisement au sein de l'occlusion dans la consonne finale était mesurée. Cette proportion était convertie en degré d'assimilation exprimé en pourcentage comme dans l'Expérience 1a. La Figure 2.6 montre la distribution des degrés d'assimilation pour les MAA et MNA, alors que le Tableau 2.5 présente les degrés d'assimilation en fonction du plan factoriel expérimental. Nous pouvons constater qu'un nombre important d'occurrences appartient à des degrés intermédiaires d'assimilation (44.5% d'occurrences sont associés à la gamme d'assimilation [20-80%]). Généralement il y avait autant d'assimilation dans cette expérience que dans l'Expérience 1 (52.2% vs. 54.7% respectivement).

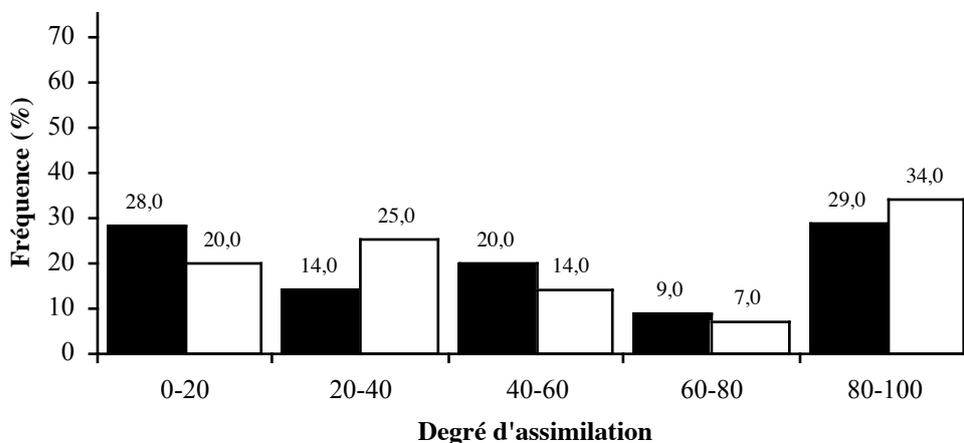


Figure 2.6. Distribution du degré d'assimilation pour les MAA (barres noires) et MNA (barres blanches, (Expérience 3a).

<i>Type</i>	<i>Voisement</i>	<i>Densité de voisinage phonologique</i>			
		<i>[2-7]</i>	<i>[8-11]</i>	<i>[12-17]</i>	<i>Moyenne</i>
<i>MAA</i>	<i>[-voice]</i>	75.0	76.6	61.5	71.1
	<i>[+voice]</i>	23.5	30.7	35.0	29.7
	<i>Moyenne</i>	49.3	53.6	48.3	50.4
<i>MNA</i>	<i>[-voice]</i>	82.2	73.7	80.7	78.9
	<i>[+voice]</i>	30.7	28.5	29.1	29.4
	<i>Moyenne</i>	56.5	51.1	54.9	54.1

Tableau 2.5. Degrés d'assimilation (%) selon le type de mot (MAA vs. MNA), nombre de voisins phonologiques (2-7, 8-11, ou 12-17 voisins), et voisement sous-jacent (non voisé vs. voisé).

Il a été suggéré que la force d'assimilation peut varier en fonction du lieu d'articulation des segments critiques (cf. Umeda, 1977). C'est la raison pour laquelle nous avons aussi considéré les degrés d'assimilation en fonction du lieu d'articulation et le voisement des consonnes obstruantes. Ceci est indiqué dans le Tableau 2.6 : comme nous pouvons le constater, le degré d'assimilation était systématiquement plus élevé pour les segments non voisés que voisés, quel que soit leur lieu d'articulation (voir aussi van Dommelen, 1985).

	<i>Nombre</i>	<i>Non voisés</i>	<i>Voisés</i>
<i>Place dentale</i>	8	81	32
<i>Place vélaire</i>	6	80	23
<i>Place labiale</i>	10	68	29

Tableau 2.6. Le degré d'assimilation (%) selon le lieu d'articulation (dental, vélaire ou labial) des segments non voisés et voisés en position finale.

Une analyse de variance (par items et par locuteurs) était effectuée sur les données avec pour facteurs principaux Voisinage phonologique à trois modalités (voisinage à faible densité, à densité moyen et à forte densité), Ambiguïté à deux modalités (MAA vs. MNA), et type de Voisement sous-jacent à deux modalités (non voisé vs. voisé). Ni le facteur Voisinage, ni le facteur Ambiguïté ne montraient un effet significatif. L'effet de Voisement sous-jacent était

très significatif, $F_1(1,7) = 32.55, p < .001, F_2(1, 36) = 172.6, p < .0001$ ¹³, ce qui reflète un degré d'assimilation plus important pour les obstruents non voisés que voisés. Nous avons également observé une interaction significative entre les facteurs Voisement et Ambiguïté dans l'analyse par locuteurs $F_1(1,7) = 6.94, p < .05$, mais pas dans l'analyse par items, $F_1(1,36) = 1.39, n.s.$ Cette interaction significative dans l'analyse par sujets, reflète une asymétrie plus marquée entre segments non voisés et voisés pour les mots non ambigus que pour les mots ambigus, bien que l'effet de Voisement sous-jacent soit observé pour les deux types d'items ($p < .001$). Il n'y avait pas d'autre effet ou interaction qui atteigne le seuil de significativité.

Les résultats suggèrent que les MAA étaient assimilés moins complètement que les MNA (degrés moyens de 50% et 54% respectivement). Quand nous observons les données de plus près (cf. Figure 6), des différences subtiles de degrés d'assimilation apparaissent pour les deux types d'items. Il y avait plus de mots ambigus que non ambigus qui n'ont pas ou peu subi d'assimilation (cf. gamme de [0-20%]) : 28% vs. 20% respectivement, $t(7) = 2.5, p < .05$. De façon similaire, il y avait plus de mots non ambigus qu'ambigus ayant subi une assimilation (quasi-)complète (cf. gamme de [80-100%]) : 34% vs. 29% respectivement, alors que cette différence ne s'est pas révélée significative. Ces tendances indiquent que les MAA résistent plus à l'assimilation que les MNA, bien que la différence entre les deux types de mots soit bien plus faible que celle obtenue dans l'Expérience 1a.

Comme nous l'avons indiqué, il n'y avait pas d'effet du facteur Voisinage phonologique, quand sa définition correspond à la densité ou nombre de compétiteurs lexicaux après substitution de la consonne finale du mot critique. Nous avons alors pris en considération une mesure de voisinage phonologique qui pourrait mieux rendre compte de l'influence des voisins sur le traitement lexical du mot cible. Cette mesure est définie comme étant le nombre de voisins phonologiques plus fréquents que le mot cible ou « nombre de voisins dangereux » (désormais NVD). Des études sur la reconnaissance visuelle des mots ont montré la pertinence de cette mesure et son rôle lors du traitement visuel et auditif des mots (Grainger, O'Regan, Jacobs, & Segui, 1989 ; Luce, 1986 ; Meunier & Segui, 1999). Nous avons pris cette nouvelle définition du voisinage phonologique pour ré-examiner l'effet du voisinage sur la production des segments assimilés. Pour les mots non voisés, ayant présenté la gamme la

¹³ Toutes les analyses statistiques inférentielles dans cette thèse ont été effectuées avec le logiciel R (R Foundation for Statistical Computing, <http://www.R-project.org>).

plus étendue de degrés d'assimilation, le NVD a corrélé négativement avec le degré d'assimilation, $r(22) = -0.51, p < .05$.

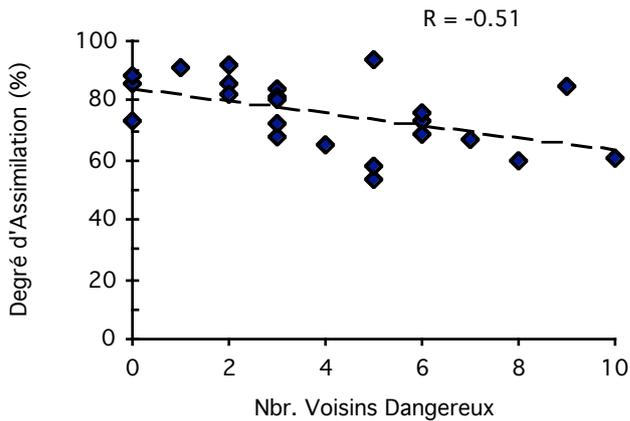


Figure 2.7. Graphique de corrélation entre le degré d'assimilation en ordonnée et le nombre de voisins dangereux en abscisse pour les segments non voisés.

La Figure 2.7 correspond la graphique de corrélation où le degré d'assimilation des segments non voisés est présenté en fonction du NVD.

Pour résumer, nous avons trouvé quelques indications montrant que les deux facteurs lexicaux d'ambiguïté lexicale et de voisinage phonologique (défini en termes de Nombre de Voisins Dangereux, NVD) tendent à inhiber la force avec laquelle les mots sont assimilés, probablement par le fait que ces facteurs lexicaux induisent une articulation plus soignée afin d'éviter toute confusion lexicale.

L'objectif de l'Expérience 3b était de confirmer la pertinence perceptive de l'indice acoustique que nous avons utilisé, tout comme nous l'avons fait dans l'Expérience 1b. Nous avons décidé d'utiliser les enregistrements d'un locuteur et de sélectionner un ensemble de mots qui présentait la gamme la plus étendue des degrés d'assimilation, pour que la comparaison entre les données acoustiques et perceptives ne soit pas limitée aux intervalles étroits des degrés d'assimilation. Tout comme dans l'Expérience 1b, les participants ont dû catégoriser les consonnes finales des mots extraits de leur contexte phrastique.

6. EXPERIENCE 3B □ CATEGORISATION PHONETIQUE DE MOTS ASSIMILES

6.1. Matériel expérimental

Les productions d'un locuteur masculin (L1) ont été sélectionnées, car elles étaient les plus représentatives des productions de l'ensemble des locuteurs ($n = 8$) de l'Expérience 3a. Le

degré moyen d'assimilation de cet individu était de 49.8% , alors que la moyenne du groupe de locuteurs était de 52.3% ; le pourcentage des profils extrêmes d'assimilation (toute absence d'assimilation 0%, ou assimilation complète 100%) était de 52.1% pour L1 alors que la moyenne du groupe était de 54.7%. Parmi les 48 mots test qui avaient été produits par L1, 24 mots étaient sélectionnés dont les degrés d'assimilation couvraient tout le spectre d'assimilation de 0% à 100%. En effet, deux tiers des items retenus avaient un degré d'assimilation appartenant à l'intervalle intermédiaire d'assimilation de [20-80%]. La moitié du matériel correspondait à des MAA (degré moyen de 45.4%), alors que l'autre moitié correspondait à des MNA (degré moyen de 49.3%) avec des distributions d'assimilation similaires pour les deux types d'items. Nous avons rajouté au matériel expérimental douze items de remplissage ainsi que 10 items qui figuraient dans la phase d'entraînement. Tous les 46 items étaient extraits de leur contexte phrastique suivant les mêmes procédures que celles de l'Expérience 1b.

6.2. Sujets

Vingt étudiants à l'Institut de Psychologie, Université René Descartes (deux hommes et 18 femmes, âge moyen = 24 ans) ont participé à la tâche de catégorisation phonétique. Tous les participants avaient pour langue maternelle le français et aucun d'entre eux n'avait indiqué des problèmes auditifs et/ou visuels. Aucun de ces participants n'avait pris part aux expériences précédentes ou aux pré-tests.

6.3. Procédure

Les participants étaient testés individuellement dans une salle silencieuse. Les items leur étaient présentés auditivement à un niveau d'écoute confortable à l'aide d'une paire d'écouteurs Sennheiser. L'expérience était mise en place avec le logiciel Praat (Boersma & Weenink, 1992-2004). La consigne indiquait aux participants qu'ils allaient entendre des signaux de parole finissant par une consonne et que leur tâche était de décider pour chaque stimulus, quelle consonne ils croyaient avoir entendu, en choisissant celle-ci parmi les consonnes /p, t, k, b, d, g, s, z/ qui étaient présentées visuellement sur l'écran d'ordinateur. La réponse était effectuée en cliquant avec la souris sur la consonne correspondante. Pour chaque stimulus, les participants devaient donner un jugement de confiance sur une échelle de 1-à-5. La phase test consistait en deux blocs, chacun contenant les 24 mots cibles ainsi que les 12 mots de remplissage. Chaque item test était présenté deux fois. Dix items monosyllabiques étaient présentés dans la phase d'entraînement. Après l'expérience, les participants devaient

remplir un questionnaire portant sur leurs connaissances linguistiques. La durée de l'expérience était environ de 15 minutes.

6.4. Résultats

Tout comme dans l'Expérience 1b, les réponses sont prises en compte comme étant « congruente » quand elles sont compatibles avec le voisement sous-jacent (le lieu d'articulation était parfois mal perçu comme dans Expérience 1b, le taux d'erreur moyen était de 5.8%.) Nous observons de nouveau qu'il existe une corrélation négative entre le degré d'assimilation et le pourcentage de réponses congruentes, $r(22) = -0.73$, $p < .000$, ce qui est en lien avec l'idée qu'un degré plus élevé d'assimilation induit un pourcentage plus bas de réponses qui correspondent au voisement sous-jacent. La corrélation est représentée dans la Figure 2.8.

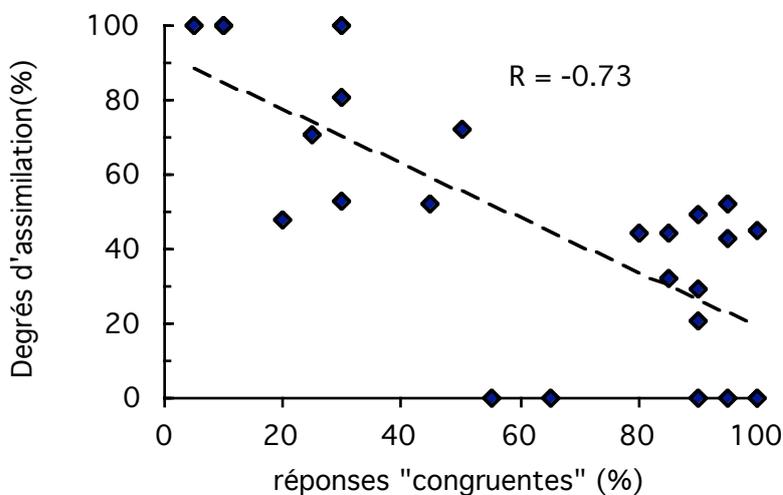


Figure 2.8. Graphique de corrélation entre le degré d'assimilation et le pourcentage de réponses congruentes.

6.5. Discussion des Expériences 3a-b

Les résultats de l'Expérience 3a ont confirmé les effets principaux de l'Expérience 1a. Premièrement, l'assimilation de voisement en français implique souvent une transformation graduée d'un son à un autre son. Deuxièmement, il existe un profil asymétrique d'assimilation dans le sens où les obstruents non voisés sont plus fréquemment et plus complètement assimilés que les obstruents voisés. Enfin, la forte corrélation négative entre l'indice acoustique d'assimilation et les jugements perceptifs est obtenue dans les deux expériences. Etant donné sa réalité perceptive, ceci nous amène à conclure que cet indice acoustique

constitue une mesure fiable afin d'évaluer le degré d'assimilation des consonnes obstruantes en position finale.

L'autre objectif de l'Expérience 3a était d'examiner de manière mieux contrôlée l'influence potentielle des facteurs lexicaux sur la production de la parole assimilée. Contrairement aux résultats de l'Expérience 1, nous n'avons trouvé que des différences très subtiles entre les mots ambigus et non ambigus dans la distribution des degrés d'assimilation. Le voisinage phonologique affectait la force d'assimilation tout au moins pour les segments non voisés. Ceci était observé en prenant en compte le nombre de voisins phonologiques plus fréquents que le mot cible, et non pas la densité phonologique tout court. La production d'un mot cible est donc influencée par la présence des voisins phonologiques « dangereux » et ceci résulte en une inhibition dans le degré d'assimilation. Des études ont montré que des mesures similaires de voisinage phonologique étaient capables d'évaluer leur influence sur la reconnaissance auditive et visuelle des mots (Luce, 1986 ; Meunier & Segui, 1999 ; Grainger, O'Regan, Jacobs, & Segui, 1989).

7. DISCUSSION GENERALE

Dans ce chapitre, notre objectif a été d'étudier l'assimilation régressive de voisement des consonnes obstruantes en position finale. Nous avons présenté une description acoustico-phonétique de l'assimilation et ensuite nous avons examiné l'influence potentielle de deux facteurs lexicaux sur la production des obstruantes assimilées. Les mesures acoustiques ont montré que la durée relative de voisement, i.e. la portion voisée au sein de l'occlusion totale, était une signature fiable du voisement pour distinguer les obstruantes non voisées et voisées. A partir de cette mesure de voisement, un indice d'assimilation a été dérivé. La pertinence perceptive de cet indice acoustique a été évalué à l'aide d'une tâche de catégorisation phonétique. Dans cette tâche, les participants ont dû catégoriser les consonnes en position finale des mots qui étaient extraits de leur contexte phrastique. Une corrélation négative importante a été obtenue entre les jugements phonétiques des participants et le degré d'assimilation tel que donné par notre indice acoustique. Nous sommes portés à croire que ceci justifie le choix de ce paramètre acoustique pour l'étude quantitative de l'assimilation de voisement en français.

Les résultats principaux peuvent être résumés comme suit. Premièrement, l'assimilation de voisement ne constitue pas toujours un processus discret dans la mesure où une consonne obstruante en position finale est transformée d'un son à un autre, mais il s'agit plutôt d'une

transformation graduée. En effet, nous avons observé qu'il peut y avoir des degrés intermédiaires d'assimilation. Cette observation est compatible avec des données acoustico-phonétiques des variations similaires dans d'autres langues (i.e., assimilation de place en anglais : Barry, 1985 ; Gow, 2002, 2003; Gow & Hussami, 1999 ; Nolan, 1992 ; Wright & Kerswill, 1989 ; assimilation de voisement en allemand : Kuzla, 2003). Ces résultats sont également en accord avec les études sur la neutralisation incomplète du dévoisement final (*Final Devoicing*) dans d'autres langues (Dinssen & Charles-Luce, 1984 ; Ernestus & Baayen, sous presse ; Port & Crawford, 1989 ; Warner et al., 2004).

Le deuxième résultat principal est que le voisement sous-jacent module la force de l'assimilation (voir aussi Kohler, 1979). Nous avons systématiquement trouvé que les obstruents non voisés étaient plus fréquemment et plus complètement « voisés » que les obstruents voisés ne se dévoisent. Les données perceptives des Expériences 1b-3b ont suggéré que le profil asymétrique acoustique entre non voisés et voisés reproduit une asymétrie perceptive : après avoir été extraites des contextes phrastiques assimilatoires, les consonnes non voisés étaient fréquemment perçues comme des voisés tandis que les consonnes voisés étaient rarement perçues comme des non voisés.

Dans un article sur l'implémentation phonétique des traits [voice] et [tense], Jessen (2001a, 2001b) a noté que l'assimilation régressive de voisement est souvent observée dans les langues qui emploient le trait [voice], comme le français. Il a été proposé que l'assimilation de voisement d'une consonne non voisée suivie par une consonne voisée peut être considérée comme une sorte de stratégie d'expansion de la durée de l'occlusion voisée présente dans le second membre de la séquence C_1C_2 , ce qui résulte en une amélioration perceptive du voisement. Dans le même ordre d'idées, le profil asymétrique entre voisés et non voisés est également rapporté par Saerens et al. (1989) qui ont examiné la perception du trait de voisement des consonnes plosives en français. Dans une de leurs expériences, les participants avaient à identifier les plosives extraites de la parole continue. Les auteurs ont trouvé que le pourcentage d'identification *correcte* de voisement, c'est-à-dire correspondant au voisement sous-jacent, était plus élevé pour les consonnes voisés que non voisés. Ce résultat suggère que les plosives voisés ont subi moins d'assimilation que les non voisés. Par conséquent, les auditeurs ont eu plus facilement accès au voisement sous-jacent des voisés.

Dans le cadre du programme de phonologie articulatoire, Browman et Goldstein (1990) ont expliqué l'absence de tout dévoisement en termes d'une réduction d'ampleur des gestes de fermeture et d'ouverture glottale. Si l'ampleur de l'ouverture est suffisamment réduite, il se

peut que le dévoisement ne puisse même pas avoir lieu, comme c'était le cas pour les obstruents intervocaliques en japonais.

Le profil asymétrique d'assimilation pourrait également être expliqué en prenant en compte les fréquences lexicales. En français, il existe plus de mots monosyllabiques avec pour consonne finale une obstruente non voisée que voisée (d'après la base de données *Lexique* : 2556 vs. 1091, respectivement), ce qui implique que les consonnes non voisées représentent le cas non marqué. Ainsi, le profil asymétrique d'assimilation pourrait être considéré en termes de « préservation » ou encore de résistance contre toute variation des formes marquées (i.e., segments voisés) plutôt que non marquées (i.e., segments non voisés). Cette approche est également conciliable avec l'approche dite statistique du traitement langagier, selon laquelle les locuteurs utiliseraient de façon implicite leurs connaissances sur la fréquence d'occurrence des formes lexicales (Bybee, 1995, 2002 ; Hay, Pierrehumbert, & Beckman, 2003 ; Booij, 1995 ; Jurafsky et al., 2001 ; Bergen, en préparation ; voir aussi Ernestus & Baayen, 2003). Une contribution précoce à ce débat est l'observation faite par Schuchardt (1885, cité dans Bybee, 2002) que les mots de haute fréquence sont affectés par des changements phonologiques plutôt que les mots de basse fréquence. Dans tous les cas, ce profil asymétrique d'assimilation indique que le voisement sous-jacent constitue, tout au moins en français, un facteur phonétique robuste qui module le degré d'assimilation des obstruents en position finale.

Le troisième résultat concerne le rôle des facteurs lexicaux sur la production de la parole assimilée. Dans l'introduction de ce chapitre, nous avons mentionné des études ayant montré que la neutralisation phonétique semble parfois être modulée par des sources d'information de plus haut niveau. En effet, nos données suggèrent que toute confusion lexicale potentielle induite par ambiguïté lexicale par assimilation de voisement ou par l'existence des voisins phonologiques « dangereux », peut influencer la production des segments assimilés. L'expérience 1a mettait en évidence un effet d'ambiguïté potentielle. Il y avait une différence nette entre les profils d'assimilation des MAA (*soute / soude*), dans la mesure où ils étaient produits avec bien moins d'assimilation que les MNA. Cependant, comme nous l'avons indiqué précédemment, le matériel expérimental n'était pas équilibré par rapport au voisement sous-jacent. Par conséquent, nous n'avons pas mesuré l'influence de l'ambiguïté lexicale indépendamment du voisement des consonnes finales. Qui plus est, la différence observée entre mots ambigus et non ambigus était liée au fait que les locuteurs dans l'expérience 1a ont probablement pu prendre conscience de la présence des paires minimales. En

conséquence, il est possible qu'ils aient articulé les mots potentiellement ambigus avec plus de soin (i.e., avec moins d'assimilation) que les mots non ambigus. En effet, dans l'Expérience 1, les locuteurs se sont directement entraînés sur les phrases expérimentales afin de pouvoir les prononcer de façon aussi fluide que possible lors de l'enregistrement. Les deux membres de chaque paire minimale apparaissaient dans deux contextes phrastiques. Dans l'Expérience 3, il y avait d'abord une phase d'entraînement avant que l'enregistrement proprement dit se soit déroulé. Cette fois-ci les mots stimuli étaient insérés seulement dans des contextes assimilatoires, ainsi chaque membre d'une paire minimale figurait une fois dans le matériel expérimental. Contrairement à l'Expérience 1, des phrases courtes et sémantiquement neutres étaient sélectionnées, ce qui visait à prévenir que les locuteurs ne détectent pas la similarité phonologique entre les mots stimuli de chaque paire. Dans ces conditions plus strictes d'enregistrement, l'ambiguïté lexicale cessait d'affecter la production des segments assimilés de façon dramatique, bien que les deux distributions aient montré quelques différences subtiles entre les deux types de mots. La présence des voisins phonologiques plus fréquents que le mot cible (ou les voisins « dangereux ») semblait affecter quelque peu l'assimilation d'une consonne en position finale. Toutefois, la densité de voisinage (qui ne prend pas en compte les fréquences relatives d'un mot et de ses voisins), n'avait pas d'impact sur le degré d'assimilation. Ces résultats donnent à croire que des effets résiduels de confusion lexicale peuvent émerger, même si le locuteur n'a pas réellement pris conscience de la présence des mots ambigus par assimilation (Expérience 3). Est-ce que ces effets reflètent une intention, non consciente éventuellement, de la part du locuteur d'éviter toute confusion lexicale potentielle ? Nous sommes plutôt portés à croire que ces effets sont de nature automatique et irrépressible et révèlent l'interdépendance stricte entre les processus de production et de perception d'accès au lexique mental. Dans le cas de l'effet de voisinage phonologique, il est peu probable que les locuteurs puissent adapter de façon intentionnelle leur production afin d'éviter toute confusion induite par les voisins phonologiques du mot cible à produire. La marque d'une articulation plus soignée, ou *hyper-articulation*, pourrait être attachée intrinsèquement à la représentation mentale des mots qui appartiennent à des voisinages phonologiques plus ou moins denses, et ceci devrait être implémenté en production tout comme en perception.

Des études en perception de la parole ont révélé un rôle assez important de voisinage sur le traitement lexical, pourvu que les mesures de voisinage prennent en compte les fréquences des voisins (Grainger & Segui, 1990). A titre d'exemple, nous pouvons citer une étude sur le

traitement des mots à dérivation de suffixes, menée par Meunier et Segui (1999). Ces auteurs ont employé une mesure de compétition au sein des familles morphologiques, qui est réminiscente de la mesure que nous avons proposée dans l'Expérience 2, le nombre de voisins « dangereux ». La compétition morphologique d'un mot donné était définie comme le nombre de mots à suffixe au sein d'une même famille morphologique qui sont plus fréquents que le mot cible. Il a été trouvé que les temps de décision étaient plus longs pour les mots avec beaucoup de tels voisins par rapport aux mots avec peu de voisins. Luce (1986) a employé une tâche d'identification dans laquelle des mots monosyllabiques parlés étaient présentés dans du bruit. Il a trouvé que les auditeurs avaient plus de mal à identifier des mots à forte densité phonologique que ceux à faible densité. Dans l'ensemble, ces études montrent que la compétition lexicale (et morphologique) de voisinage peut affecter l'accès au lexique, ce qui suggère que les mots avec beaucoup de compétiteurs lexicaux requièrent une articulation plus soignée ou une saillance acoustique majeure afin d'être reconnus par l'auditeur. Il est tentant de spéculer qu'en conséquence, ces mots sont produits de façon plus articulée, car après tout, chaque locuteur est aussi un auditeur. Les résultats de l'Expérience 3 ont fourni quelques données expérimentales en faveur de cette hypothèse.

En résumé, nous avons abordé la question de l'implémentation phonétique de l'assimilation de voisement des consonnes obstruantes en position finale en français. Les données acoustiques ont montré que l'assimilation est fréquemment (mais pas toujours) incomplète. Ce processus peut être évalué de façon précise à l'aide des analyses quantitatives obtenues avec un indice acoustique doté d'une certaine pertinence perceptive. Ces mesures ont permis de démontrer que la force d'assimilation dépend du voisement sous-jacent des consonnes soumises à cette variation phonologique dans le sens où les consonnes non voisées subissent plus d'assimilation que les consonnes voisées, quelles que soient les intentions communicatives du locuteur. Si les locuteurs sont peu concernés par des intentions communicatives, des facteurs lexicaux semblent toutefois influencer sur la production des segments assimilés (Expérience 2). Bien entendu, davantage de recherches sont nécessaires afin d'apporter plus de confirmation au rôle des facteurs lexicaux sur la production de la parole assimilée.

La connaissance qu'apportent les données acoustico-phonétiques sur la parole assimilée dont nous rendons compte dans ce chapitre, peut s'avérer utile pour l'étude du traitement perceptif (cf. Fowler, 1987). En guise d'introduction de cette thèse, nous avons souligné l'apport des travaux conduits en phonologie et en phonétique expérimentale pour l'étude

psycholinguistique des variations phonologiques. Une meilleure connaissance de l'implémentation phonétique de l'assimilation de voisement peut révéler des aspects fondamentaux du traitement perceptif de la parole assimilée. L'objectif majeur du chapitre suivant sera d'examiner les mécanismes perceptifs impliqués dans le traitement des segments plus ou moins fortement assimilés en français. Etant donné l'asymétrie d'assimilation que nous avons observée entre obstruents non voisés et voisés, cette caractéristique sera utilisée afin d'opérationnaliser l'hypothèse selon la sensibilité contextuelle varie en fonction de l'ampleur de l'assimilation.

CHAPITRE III.

LE TRAITEMENT PERCEPTIF DES SEGMENTS PLUS OU MOINS FORTEMENT ASSIMILÉS : TRACES ACOUSTIQUES ET RÔLE DU CONTEXTE DROIT

1. INTRODUCTION

Depuis une décennie, de nombreux chercheurs en psycholinguistique se sont intéressés aux mécanismes perceptifs impliqués dans le traitement de l'assimilation de lieu d'articulation (cf. Otake, Yoneyama, Cutler, van der Lugt, 1996 ; Gaskell & Marslen-Wilson, 1996, 1998 ; Coenen, Zwitserlood, & Bölte, 2001 ; Mitterer & Blomert, 2003 ; Gumnior, Zwitserlood, & Bölte, 2005 ; Gow, 2001, 2002, 2003, 2004 ; Weber, 2001, 2002). Même si les chercheurs ne s'accordent pas sur son rôle exact, il est clair que toutes ces études ont mis en avant le rôle du contexte phonologique qui donne lieu à l'assimilation dans le traitement perceptif des segments assimilés. Dans ce chapitre, nous examinerons les deux principales hypothèses avancées à propos des mécanismes impliqués dans la perception de la parole assimilée avant de présenter nos propres travaux sur ce point. Comme nous l'avons déjà indiqué dans le chapitre 1, ces deux hypothèses correspondent à l'influence régressive et à la présence des traces acoustiques d'assimilation. Dans ce qui suit, nous présenterons les données empiriques favorables à l'une et à l'autre de ces hypothèses.

1.1. Données empiriques en faveur de l'Hypothèse de l'Inférence Régressive

L'hypothèse en vertu de laquelle le contexte phonologique est employé par l'auditeur pour inférer la forme sous-jacente du segment assimilé a été examinée dans une série de travaux sur l'assimilation de lieu d'articulation en anglais conduites par Gaskell et collègues (Gaskell 1994 ; Gaskell, Hare, & Marslen-Wilson, 1995 ; Gaskell & Marslen-Wilson, 1996, 1998). Dans l'étude publiée en 1996, Gaskell et Marslen-Wilson ont employé le paradigme de décision lexicale avec amorçage intermodal (*cross-modal priming*). Parmi les amorces auditives, il y avait des mots assimilés (e.g., *leam* dérivé de *lean*) et leurs formes canoniques (*lean*). Dans la première expérience, la cible était présentée visuellement à la fin de l'amorce auditive ; le contexte phonologique qui justifie l'assimilation n'était donc pas disponible aux sujets. Ils ont trouvé que l'ampleur des effets d'amorçage était comparable pour les formes

assimilées et canoniques. Autrement dit, *leam* amorçait LEAN tout autant que *lean*. Dans la deuxième expérience, le contexte droit était rendu disponible aux sujets. Ces auteurs ont observé que les effets de facilitation n'étaient plus observés quand la forme assimilée était insérée dans un contexte non viable, c'est-à-dire, dans un contexte qui ne justifie par l'assimilation labiale (e.g., *leam gammon*). Ces résultats suggèrent que des mots prononcés d'une façon non canonique ne perturbent pas l'accès au lexique, pourvu qu'ils apparaissent dans un contexte phonologique viable, c'est-à-dire compatible avec le changement assimilatoire. Ceci a conduit Gaskell et Marslen-Wilson à interpréter leurs résultats en termes d'un mécanisme cognitif d'*inférence régressive* qui sous-tendrait le traitement perceptif des segments assimilés. Les auditeurs prendraient appui sur le contexte suivant le segment assimilé afin de pouvoir inférer son identité sous-jacente.

Dans l'étude publiée en 1998, Gaskell et Marslen-Wilson ont cherché à déterminer le locus du mécanisme d'inférence régressive. Si l'inférence régressive phonologique est de nature lexicale, alors le mécanisme devrait seulement compenser les variations phonologiques présentes dans les mots et non dans les non-mots. En revanche, si l'inférence phonologique est de nature pré-lexicale, alors ce mécanisme devrait également s'appliquer aux non-mots. Afin de tester ces prédictions, ces auteurs ont employé le paradigme de détection de phonèmes. Le sujet a pour tâche de détecter un phonème (qui lui a été présenté visuellement pendant 1500 ms) à l'intérieur des mots stimuli présentés auditivement. Il donne sa réponse en appuyant sur une touche, dès qu'il entend le phonème cible. Dans la première expérience, les sujets devaient détecter le phonème sous-jacent /t/ dans les mots insérés dans les énoncés donnant lieu ou non à l'assimilation (e.g., *freight* [f r a y p] *bearer* vs. *freight* [f r a y p] *carrier*) ou dans les non-mots insérés dans les mêmes contextes (cf. *preight* [p r a y p] *bearer* vs. *preight* [p r a y p] *carrier*). Les temps de détection étaient plus courts pour les phonèmes assimilés insérés dans des contextes viables, c'est-à-dire compatibles avec les changements, que pour les phonèmes insérés dans les contextes non viables. La même tendance était obtenue pour les non-mots, mais les effets étaient plus faibles, ce qui suggère que l'inférence régressive n'est pas un mécanisme exclusivement lexical. Dans une seconde expérience, les sujets devaient détecter les phonèmes suivant les segments assimilés, comme le /b/ dans l'énoncé *freight bearer*. Les résultats confirmaient ceux de la première expérience : les temps de détection pour les mots et non-mots étaient plus courts quand le contexte motivait

l'assimilation. L'inférence phonologique semble donc être un mécanisme pré-lexical mais elle est modulée par des facteurs lexicaux.

D'autres études ont apporté des données expérimentales qui s'accordent avec celles obtenues par Gaskell et collègues. Ainsi, Coenen, Zwitserlood et Bölte (2001) ont étudié l'assimilation régressive et progressive en allemand en utilisant le même paradigme que Gaskell et Marslen-Wilson (1996). Ces auteurs ont trouvé que l'effet d'amorçage était plus important pour les formes canoniques (*kannst du das Wort mal deutlicher aussprechen*) que pour les formes assimilées (...*Worp mal*...). Quand les mots assimilés étaient présentés dans un contexte non viable (...*Worp kurz*...), ces auteurs n'ont pas observé d'effet d'amorçage (mais voir Gumnior et al., 2005). Qui plus est, ils n'ont pas observé d'effet d'amorçage quand les mots étaient présentés en isolation, ce qui renforce l'hypothèse du rôle crucial du contexte droit.

Mitterer et Blomert (2003) ont examiné le cas de l'assimilation au sein des mots en néerlandais. Ces auteurs ont comparé le traitement perceptif de deux paires de stimuli : (1) *tuinbank* – *tuimbank* où la nasale /n/ est devenue une labiale /m/, en raison de l'assimilation de lieu (la condition viable) ; et (2) *tuinstoel* – *tuimstoel* où le changement de la consonne nasale en labiale n'est pas compatible avec le contexte droit (condition non viable). Mitterer et Blomert ont employé la tâche d'identification des mots (*word identification task*) qui consiste à présenter auditivement les stimuli et à demander aux sujets d'indiquer s'ils entendaient *tuin* ou *tuim* en appuyant sur une des deux touches correspondantes. Les résultats ont révélé que les performances étaient similaires pour les conditions non viables (*tuimstoel*) et contrôle (où le mot *tuin* était présenté en isolation). En revanche, les sujets produisaient de nombreuses « erreurs » dans la condition viable (*tuimbank*) et rapportaient fréquemment avoir entendu *tuin* au lieu de *tuim*. Ces résultats sont donc en lien avec l'approche d'inférence régressive dans la mesure où les auditeurs recourent au contexte suivant le segment assimilé, i.e., /b/, afin de compenser l'assimilation de lieu dans /m/ et d'accéder à la forme sous-jacente /n/. De plus, dans cette expérience, les auteurs ont enregistré les ERP. Les données électrophysiologiques obtenues montraient que seulement les changements phonologiques viables provoquent des ondes de négativité (*mismatch negativity*), ce qui exclut l'influence de tout processus attentionnel et/ou décisionnel.

Le rôle du contexte a également été étudié dans le paradigme de l'identification phonémique par van Dommelen (1983). Dans une étude sur les paramètres acoustiques du voisement en français (mentionnée dans le chapitre précédent), cet auteur a conduit une série d'expériences, dont le but était de déterminer si la présence du contexte pouvait influencer la perception du

voisement de la consonne précédente. La génération des stimuli consistait en l'extraction du mot *rate* du contexte phrastique non voisé (*une rate plate*) et du contexte voisé (*une rade grise*). En outre, les durées de portions voisées dans les consonnes critiques (/t/ ou /d/) et contextuelle (/g/) l'occlusion étaient variées. Les résultats ont montré que la perception /t/-/d/ était avant tout déterminée par le degré de voisement appartenant à la consonne C₁ et non par les divers degrés de voisement manipulés dans la consonne contextuelle C₂ (/g/). En revanche, les réponses /t/ ont diminué seulement lorsque la consonne /g/ était complètement voisée, c'est-à-dire à 100%. L'auteur fait l'observation importante que même en absence de tout voisement (par assimilation progressive, par exemple dans *rate grise*), une vélaire voisée /g/ devenue non voisée reste *phonologiquement* voisée, c'est-à-dire que dans ce cas-là, l'auditeur associe le trait [+voice] à la consonne vélaire /g/, en dépit de sa réalisation phonétique non voisée. Au lieu d'évoquer explicitement le rôle du contexte droit, cet auteur préfère parler des « attentes linguistiques » (*linguistic expectancy*) ou encore des connaissances implicites des phénomènes d'assimilation (*unconscious knowledge of assimilatory phenomena*) de l'auditeur qui peuvent contribuer à la perception du voisement. Ainsi, la consonne contextuelle dans *plate* ou *grise* est utilisée dans le traitement du voisement présent dans l'occlusion de /t/-/d/. Dans la perception d'une séquence consonantique C₁#C₂, l'estimation perceptive du voisement dans l'occlusion de C₁ dépend, entre autres facteurs, du statut phonémique de C₂. Si C₂ est voisée, tout voisement présente dans C₁ peut alors être décodé comme étant originaire de C₂ (qui a été dévoisée). Certes, ce type de raisonnement est analogue à celui de la théorie d'alignement des traits (cf. Gow, 2003, voir chapitre 1), où les traits qui appartiennent à deux consonnes qui se suivent sont regroupés par des mécanismes perceptifs. Cependant, van Dommelen insiste sur la nécessité de la présence de C₂ avant que toute décision concernant le statut de C₁ ne puisse être prise.

1.2. Données empiriques en faveur de l'Hypothèse des Indices Acoustiques d'Assimilation

Les études citées plus haut suggèrent que les segments assimilés ne sont pas évalués en isolation, mais au contraire en fonction du contexte phonologique suivant. Toutefois, des études utilisant le paradigme de détection de phonèmes ont révélé des résultats quelque peu divergents. Comme mentionné précédemment, Gaskell et Marslen-Wilson (1998) ont obtenu des effets relativement faibles quant à la reconnaissance du segment suivant celui qui était assimilé (e.g., détection du /b/ dans *freight bearer*). En contraste, Gow (2001) a obtenu des effets importants d'anticipation en utilisant la même technique. Les temps de détection étaient

significativement plus courts pour les phonèmes-cibles suivant les consonnes assimilées générées dans des contextes viables, c'est-à-dire compatibles avec le changement phonologique (e.g., *ten* [təm] *buns*), que pour les cibles suivant des consonnes canoniques (e.g., *ten* [tɛn] *buns*), elles-mêmes plus courtes que les cibles suivant les consonnes assimilées générées dans des contextes non viables (e.g., *ten* [tɛŋ] *buns*).

Otake, Yoneyama, Cutler et van der Lugt (1996) ont examiné des effets de contexte régressifs et progressifs dans la perception de l'assimilation de lieu nasal chez les auditeurs japonais et néerlandais. Ces auteurs ont trouvé que les consonnes nasales en position finale, i.e. des nasales moraïques, étaient détectées plus rapidement que les nasales non moraïques par les participants japonais. Le lieu d'articulation des nasales n'a pas affecté les temps de réponses. Qui plus est, les participants japonais étaient également capables d'utiliser la réalisation phonétique d'une nasale moraïque afin d'anticiper le phonème suivant (voir aussi Quené, van Rossum, & van Wijck, 1998, pour des résultats similaires en néerlandais). Néanmoins, comme l'a indiqué Gaskell (2003), les changements de lieu produits en position finale syllabique ne sont jamais contrastifs en japonais. Ceci implique que tout changement phonologique peut correspondre à une seule forme canonique, sans nécessité d'évaluation contextuelle. De plus, les temps de détection des auditeurs néerlandais qui ont également effectué la tâche en japonais n'ont pas reflété des effets d'anticipation d'assimilation, alors que le changement de lieu d'articulation est contrastif en néerlandais.

Plus récemment, Gow (2003) a employé l'amorçage formel et la technique de détection des phonèmes afin d'examiner quel type d'information est extrait et interprété par les auditeurs quand ils sont confrontés à des segments assimilés qui correspondent à un segment potentiellement ambigu. Dans une première expérience dans laquelle l'amorçage formel était utilisé, la présentation comme stimuli des mots tels que *cone*, voisin phonologique du mot assimilé *comb* (après insertion dans *cone bent*), a donné lieu à des effets d'amorçage pour le mot cible CONE et non pas pour COMB. Une deuxième expérience de détection de phonèmes a montré un traitement facilitateur pour les phonèmes suivant les segments assimilés. Les résultats combinés de ces deux expériences s'accordent avec l'hypothèse selon laquelle un segment assimilé contient en fait des indices acoustiques correspondant à deux lieux d'articulation : celle du segment assimilé et celle du segment donnant lieu à l'assimilation. Par conséquent, la présence de ces indices dans les des segments assimilés peuvent faciliter la détection du phonème suivant. Dans une troisième expérience, la technique de combinaison

des segments acoustiques (*cross-splicing*) était employée afin de remplacer le contexte labial induisant une labialisation du segment précédent par un contexte coronal. Le contexte labial *bent*, qui induisait une assimilation labiale de la consonne coronale /n/, était cette fois-ci remplacé par le mot *dents*, dont l'attaque contient une place coronale. Les résultats d'amorçage ont révélé que les auditeurs avaient accès à COMB et non pas à CONE. Ces résultats montrent que les indices acoustico-phonétiques présents dans le segment assimilé (i.e., par rapport à la place labiale) sont préservés et utilisés par les auditeurs, en dépit d'un contexte phonologique coronal, incompatible avec ce changement¹⁴.

1.3. Hypothèses alternatives ou complémentaires ?

Les résultats divergents dans la littérature que nous venons d'examiner poussent à se demander quel est le rôle exact de la sensibilité contextuelle dans le traitement perceptif des segments assimilés. Les effets de « viabilité contextuelle » obtenus par Gaskell et Marslen-Wilson (1996, 1998) et d'autres (e.g., Coenen et al., 2001) ont été expliqués en supposant que les segments assimilés sont évalués exclusivement par des mécanismes d'inférence régressive. Selon cette approche, le contexte phonologique est utilisé de façon régressive par l'auditeur afin d'évaluer les possibilités de changement phonologique dans la forme de surface. L'inférence régressive phonologique pose des contraintes qui permettent la dérivation des formes sous-jacentes en rapport avec le contexte subséquent. Nous avons vu dans l'introduction que Gow (2001, 2002, 2003) a cherché à fournir une explication perceptuelle plus générale afin de rendre compte des effets de contexte. Il a proposé que les interprétations divergentes concernant le rôle du contexte droit dans le traitement perceptif des mots assimilés étaient probablement dues aux caractéristiques acoustico-phonétiques des stimuli utilisés dans les différentes études. Dans les chapitres précédents, nous avons déjà amplement

¹⁴ L'objectif principal de la troisième expérience était de déterminer si les effets obtenus sont le résultat de la mise en œuvre des mécanismes perceptifs généraux, hypothèse soutenue par l'auteur, ou des mécanismes gouvernés par les connaissances implicites, spécifiques à la langue (cf. Gaskell et collègues). En effet, il semble difficile de réconcilier les résultats de la troisième expérience avec cette dernière approche, mais nous tenons à signaler que les stimuli utilisés étaient créés à partir de la technique de *cross-splicing*. Étant donné leur nature artificielle, on se demande alors si dans ce cas-là, on peut réellement parler des stimuli « linguistiques » proprement dits. Plus précisément, il est difficile d'envisager la mise en œuvre des mécanismes spécifiques à la langue pour un traitement perceptif qui requiert une analyse des stimuli *non linguistiques*. Ainsi, les résultats constituent, à notre sens, un argument plutôt faible pour mettre en cause l'existence de tels mécanismes comme le fait entendre Gow.

discuté des études en phonétique expérimentale ayant démontré que l'assimilation en anglais, mais aussi dans d'autres langues, coïncide rarement avec une transformation complète d'un son à un autre son, mais qu'il s'agit plutôt d'une modification graduée dans la forme de surface (e.g., Nolan, 1992 ; Warner et al., 2004). D'après Gow, une modification graduée du segment (partiellement) assimilé implique que celui-ci contient des indices sur les deux lieux d'articulation. Ces informations détaillées sont employées par l'auditeur afin de (a) dériver sa forme sous-jacente et éventuellement de (b) prédire les propriétés du segment suivant, sans nécessité de s'appuyer sur le contexte phonologique droit. Gow a proposé que la reconnaissance d'un segment assimilé puisse être considérée comme un problème d'alignement des traits (*feature parsing*). Selon ce type de traitement perceptif, des valeurs de traits acoustiques temporellement distribués sont groupés ensemble et intégrés dans des représentations au niveau des traits, à leur tour alignés avec les positions segmentales. Les auditeurs ayant entendu des consonnes coronales /n/ produites sous forme labiale peuvent récupérer des preuves acoustico-phonétiques fines de lieu labial et coronal. Par conséquent, les auditeurs sont capables de reconnaître la forme voulue par le locuteur (e.g., *cone* [ko:n] et non *comb* [ko:m]) malgré l'assimilation, car ils ont aligné avec succès la place coronale récupérée dans le /n/ en position finale de *cone* et la place labiale dans /m/ provoqué par le contexte *bent* (cf. Gow, 2003).

Gow a employé des segments assimilés de façon naturelle (i.e. une production fluide, sans que le locuteur soit au courant des règles d'assimilation), alors que Gaskell et collègues ont utilisé des stimuli délibérément assimilés dans leurs expériences. Comme l'avait suggéré Gow (2001), il se peut que les stimuli qui sont réalisés avec des degrés divers d'assimilation donnent lieu à des effets de sensibilité contextuelle différents, d'où la divergence empirique obtenue dans la littérature. Cette hypothèse a récemment été validée à l'aide des modélisations de type connexionnistes effectuées par Gaskell (2003), auxquelles nous reviendrons dans la discussion générale.

1.4. La présente recherche

Les expériences rapportées dans ce chapitre ont pour objectif de tester l'hypothèse selon laquelle la sensibilité contextuelle lors la perception des segments assimilés varie en fonction de leur degré d'assimilation. Plus précisément, le rôle du contexte droit dans le traitement des segments assimilés devrait varier en fonction de la force d'assimilation. Les segments fortement assimilés donnent lieu à une incertitude que l'auditeur est supposé résoudre

s'appuyant sur le contexte suivant afin d'inférer le voisement sous-jacent. Le traitement perceptif des segments fortement assimilés devrait être facilité par la présence du contexte suivant (cf. le mécanisme d'inférence régressive). En revanche, pour les segments faiblement assimilés, la présence du contexte suivant ne devrait pas influencer le traitement car il est supposé que des « traces »¹⁵ acoustiques sont présentes dans le signal et permettent à l'auditeur de dériver le voisement sous-jacent du segment assimilé. Afin de vérifier cette hypothèse, nous avons conduit des expériences sur l'assimilation de voisement en français.

Dans le chapitre 2, nous avons mis en évidence deux points sur cette variation phonologique. Primo, d'un point de vue acoustico-phonétique, l'assimilation de voisement entre les mots en français implique une transformation graduée. Secundo, la force avec laquelle des plosives en position finale sont assimilées semble être modulée par des facteurs phonétiques et lexicaux (quoique de façon moins importante). Plus précisément, il a été observé que les consonnes plosives non voisées (e.g., /p/ dans *jupe*) sont plus fréquemment et complètement assimilées que les plosives voisées (e.g., /b/ dans *robe*). Il semblerait donc que, pour le français, le trait de voisement soit une caractéristique pertinente afin d'opérationnaliser le degré d'assimilation. S'il s'avère que les segments non voisés subissent des assimilations fortes, nous devrions nous attendre à ce que la présence du contexte facilite le traitement perceptif de ces segments. Si les segments voisés subissent des assimilations faibles, nous ne devrions pas observer une telle influence du contexte droit car les traces acoustiques présentes dans ces signaux sont supposées être employées par l'auditeur afin d'accéder à leur forme sous-jacente. La procédure que nous avons utilisée dans cette étude est celle de l'amorçage intermodal de répétition (voir Gaskell & Marslen-Wilson, 1996). Des phrases nominales étaient présentées auditivement aux sujets et contenaient une amorce représentant soit la forme phonologiquement assimilée, soit la forme canonique. A la fin de l'amorce auditive, la cible visuelle, qui correspondait toujours à la forme canonique non altérée (e.g., JUPE), était présentée. Les participants devaient faire une décision lexicale sur la cible. Le paradigme

¹⁵ La notion de « traces » est en effet réminiscente des théories sur la structure syntaxiques développées par Noam Chomsky dans les années 1970 (voir Jackendoff, 2002, pour une discussion approfondie). Dans ces théories, les traces (*trace of movement*) permettent que toutes les interprétations sémantiques peuvent avoir lieu en termes de structure syntaxique de surface, étant donné que les positions des syntagmes présentes dans la structure syntaxique sous-jacente peuvent être identifiées à partir des traces, laissées dans la structure de surface. De façon analogue, nous proposons ici que l'identité sous-jacente d'un segment assimilé pourrait être accessible à partir des traces acoustiques, laissées dans la forme de surface.

d'amorçage intermodal a été utilisé parce qu'il a été montré que les expériences utilisant ce type de paradigme peuvent mettre en évidence des effets de facilitation quand l'amorce et la cible sont phonologiquement proches. Par exemple, il a été montré que les amorces qui riment avec les cibles (e.g., *fist* et *list* en anglais) peuvent faciliter le traitement des cibles (e.g., Norris, McQueen, & Cutler, 2002). Pourtant, si les amorces et les cibles sont présentées visuellement, la présentation des amorces qui riment avec les cibles n'ont pas facilité leur traitement (e.g., Radeau, Morais, & Segui, 1994). Il a donc été suggéré que l'amorçage intramodal reflète un traitement prélexical, plutôt que lexical (cf. Marslen-Wilson et Warren, 1994 ; Utman, Blumstein, & Burton, 2000). Gaskell et Marslen-Wilson (1996) ont observé des effets d'amorçage forts pour les amorces canoniques [l in] et assimilées [l im] sur le traitement visuel de la cible (LEAN), avec peu de preuves pour un *mismatch* pour les formes altérées. L'Expérience 5 était conçue dans le but de s'assurer que les formes canoniques et assimilées donneraient lieu à des effets d'amorçage. Dans cette expérience, le contexte auditif est constitué par un déterminant et un nom (e.g., *une jupe*). Dans cette condition, le contexte droit responsable de l'assimilation n'est pas disponible aux sujets. Dans l'Expérience 6, l'amorce auditive comporte le contexte droit afin d'évaluer son rôle sur le traitement perceptif des segments assimilés (e.g., *une jupe grise*). Dans les deux expériences, le mot cible (e.g., JUPE) est présenté visuellement à la fin du nom amorce auditif (cf. Tableau 3.1).

	<i>Expérience 5</i>	<i>Expérience 6</i>
Cible visuelle	JUPE	JUPE
Contexte auditif	/une jupe/	/une jupe grise/
Présentation visuelle cible (JUPE) □	/une jupe/ □	/une jupe □ grise/

Tableau 3.1. Présentation des stimuli dans les Expériences 5 et 6.

Nous avons pu constater dans le chapitre 2 que les degrés d'assimilation peuvent varier selon les locuteurs. Nous avons donc d'abord effectué un pré-test (Expérience 4) afin de nous assurer que les stimuli utilisés dans les expériences couvraient un continuum d'assimilation d'un point de vue perceptif.

2. EXPERIENCE 4 □ CATEGORISATION DES PHONEMES ASSIMILES

Un pré-test était nécessaire afin d'évaluer les degrés d'assimilation des mots stimuli utilisés dans les deux expériences principales. Nous avons voulu nous assurer que les stimuli couvraient un continuum de degrés d'assimilation. Dans le chapitre précédent, nous avons déjà montré que les mots non voisés sont plus fréquemment et complètement assimilés que les mots voisés. Afin de garantir que les décisions des participants n'étaient pas influencées par les contextes donnant lieu aux assimilations, seuls les mots critiques étaient présentés auditivement (i.e. sans contexte droit).

2.1. Matériel expérimental

Trente-six mots non voisés et 36 mots voisés, tous monosyllabiques, ont été sélectionnés avec des fréquences moyennes logarithmiques de 1,44 et de 1,12 respectivement. Les fréquences ont été transformées en valeurs logarithmiques afin de réduire l'étendue de distributions des fréquences (cf. Baayen, Tweedie, Schreuder, 2002). En outre, il a été démontré que les temps de réponse en décision lexicale sont corrélés au logarithme de leur fréquence (voir e.g., Rubenstein & Pollack, 1963). Il y avait douze mots par catégorie de plosives en position finale /p, t, k, b, d, g/. Les mots étaient toujours non ambigus dans la mesure où tout changement induit par assimilation a donné lieu à un non-mot en français (e.g., la forme assimilée de *jupe* [ʒyb] ne représente pas un mot en français). Les 72 mots ont tous été contrôlés en termes de propriétés phonologiques et insérés dans des syntagmes nominaux de type Det-NOM-Adj dont la plausibilité sémantique a été jugée par trois locuteurs natifs du français. Les contextes donnaient lieu à des assimilations (*une jupe* [ʒyb] *grise*) ou pas (*une jupe* [ʒyp] *serrée*). Le contexte phonémique induisant une assimilation n'était jamais identique à la consonne assimilée afin de pouvoir éviter la création des consonnes géminées. A part les mots stimuli, nous avons rajouté 72 mots contrôles non reliés aux cibles visuelles afin d'obtenir un temps de réaction à partir duquel nous avons calculé les effets d'amorçage. Ces mots contrôles étaient également insérés dans des contextes (cf. Annexe 6 pour la liste complète du matériel expérimental).

Les 216 mots stimuli (72 mots insérés dans des contextes non assimilatoires, 72 mots identiques insérés dans des contextes assimilatoires et 72 mots dans des contextes neutres) ont été produits par un locuteur masculin de la région parisienne. Nous avons demandé à notre locuteur de prononcer les syntagmes de façon continue et fluide, sans laisser de silence entre

les mots au sein de chaque syntagme, (il pouvait toutefois laisser des pauses entre les séquences). Le locuteur était au courant des règles d'assimilation pour les obstruents en français et nous lui avons demandé de produire délibérément des assimilations là où il y avait lieu d'en avoir. La production était directement enregistrée et digitalisée (20 kHz) sur ordinateur à l'aide du logiciel PRAAT (Boersma & Weenink, 2004). La fin du mot amorce était identifiée par l'analyse des spectrogrammes avec le même logiciel (voir section 2.4, chapitre 2, pour plus de détails sur les mesures acoustiques).

2.2. Sujets

Vingt étudiants de l'Institut de Psychologie de l'Université René Descartes (Paris 5) ont participé à cette expérience. Tous les sujets étaient de langue maternelle française et ils avaient tous une audition correcte et une vision correcte ou corrigée.

2.3. Procédure

L'expérience a été effectuée avec le logiciel PRAAT (Boersma & Weenink, 2004) sur un ordinateur PC. Les participants étaient assis devant l'écran dans une salle silencieuse. Les instructions étaient données par écrit et oralement. Tous les stimuli étaient ensuite présentés auditivement à l'aide d'écouteurs (Sennheiser). Pendant la phase d'entraînement, les participants devaient catégoriser la consonne en position finale d'items non-mots qui respectent les règles phonotactiques du français. Nous avons choisi de présenter des non-mots afin de décourager l'utilisation des connaissances lexicales. Les participants devaient répondre en cliquant à l'aide de la souris sur une des six cases présentées visuellement, qui représentaient chacune des plosives /b, d, g, p, t, k/. La phase test consistait en une catégorisation forcée à deux alternatives (*two forced-alternative choice task*) de la consonne finale du mot test. Par exemple, chaque sujet avait pour tâche de choisir entre /b/ et /p/ après avoir entendu le mot *jupe* extrait du contexte assimilatoire *une jupe grise*. De plus, les participants devaient fournir un jugement subjectif sur le choix phonémique effectué sous forme d'une note sur une échelle de confiance de 1 – à – 5. Ce jugement indiquait si la correspondance entre le son entendu et la réponse était de très bonne (5) à très médiocre (1). L'expérience a duré environ 20 minutes.

2.4. Résultats et Discussion

Les résultats pour la tâche de catégorisation phonétique sont résumés dans le Tableau 3.2. Les taux de réponses (%) des participants correspondent à la forme de surface (i.e. la forme assimilée).

	[-voisé]			[+voisé]		
	/p/	/t/	/k/	/b/	/d/	/g/
<i>% surface</i>	66 (26.8)	95 (5)	95 (6)	70 (18.6)	62 (29.4)	46 (33.7)
<i>confiance</i>	3.3 (0.5)	4.2 (0.4)	4.1 (0.4)	3.5 (0.5)	3.6 (0.5)	3.6 (0.4)

Tableau 3.2. Taux de réponses (%) et jugements subjectifs (sur une échelle de 1-5, écart types entre parenthèses) correspondant à la forme de surface pour la tâche de catégorisation phonétique.

Comme nous pouvons le constater à partir du Tableau 3.2, les mots contenant les consonnes non voisées présentées sous forme assimilée ont donné lieu à des réponses voisées dans 85 % des cas en moyenne. En revanche, les mots contenant des consonnes voisées présentées sous forme assimilée ont donné lieu à des réponses non voisées dans 59 % des cas en moyenne. Ces résultats confirment que les auditeurs ont eu plus facilement accès à la forme sous-jacente des consonnes voisées qu'à celles des consonnes non voisées quand ces formes étaient extraites des contextes assimilatoires.

En résumé, quand les mots assimilés avaient pour consonne terminale une plosive non voisée, ces consonnes ont été très majoritairement perçues par les auditeurs comme étant des voisées. En revanche, quand les assimilés avaient pour consonne finale une plosive voisée, ces consonnes ont été perçues tantôt comme des voisées, tantôt comme des non voisées. Ces résultats perceptifs sont compatibles avec les données acoustiques que nous avons présentées en chapitre 2. De nouveau, d'un point de vue perceptif (et acoustique), il y a une asymétrie au niveau des degrés d'assimilation entre les segments non voisés et voisés. Cette observation, qui confirme nos résultats antérieurs, nous permet d'opérationnaliser le degré d'assimilation par le voisement sous-jacent. De façon approximative, les plosives non voisées (/p, t, k/) correspondront aux segments fortement assimilés. Quant aux plosives voisées (/b, d, g/), elles correspondront aux segments partiellement assimilés.

3. EXPERIENCE 5 □ AMORÇAGE SANS CONTEXTE DROIT

3.1. Matériel expérimental

Dans l'Expérience 4, les séquences auditives ont été présentées de façon à ce que le contexte phonologique donnant lieu à l'assimilation soit rendu indisponible aux participants. Il y avait deux conditions où les paires amorce-cible étaient reliées : une condition dans laquelle l'amorce reliée représentait le mode de réalisation canonique (e.g., *une jupe* [ynʒyp] – JUPE), et une condition dans laquelle l'amorce reliée contenait la consonne finale réalisée sous forme assimilée (e.g., *une jube* [ynʒyb] – JUPE). Les deux conditions étaient comparées à une condition contrôle dans laquelle l'amorce et la cible n'avaient aucune relation formelle (e.g., *une chaise* – JUPE). De plus, 144 paires amorce-cible de remplissage appariées en terme de fréquence avec le matériel expérimental étaient rajoutées. Les paires ont été diversifiées afin d'éviter la mise en place de stratégies de la part des sujets. Cent-huit paires étaient accompagnées d'une cible non-mot. Deux tiers de ces non-mots étaient précédés par une amorce orthographiquement reliée (e.g., *une bière* – BIEVE) et un tiers était précédé par des amorces non reliées (e.g., *une nymphe* – REUX). Les mots amorces qui étaient suivis par les items de remplissage étaient toujours des mots français, enregistrés et extraits dans les mêmes conditions que les items expérimentaux. Les paires amorce-cible reliées étaient incluses afin de décourager les participants d'associer la relation phonologique à des mots cibles réels (cf. Lukatela, Eaton, Sabadini, & Turvey, 2004). En outre, nous avons pris soin de faire en sorte que les non-mots ne correspondent pas à des homophones hétérographiques (e.g., *heure* – HEURT), car il a été montré par ailleurs que ces items peuvent faciliter le traitement des cibles en l'absence de tout contexte biaisé (cf. Grainger, Van Kang, & Segui, 2001). Nous avons construit trois listes qui étaient contrebalancées entre trois groupes de participants, de façon à ce que chaque participant ne voit qu'une fois la même cible tout en étant soumis aux trois conditions d'amorçage (i.e. condition canonique, assimilée et contrôle). Le nombre total des items expérimentaux était de 216. En outre, dix items d'entraînement ont été rajoutés, tout comme deux items de remplissage après un bloc d'items pour permettre aux participants de se familiariser à la tâche.

3.2. Sujets

Soixante-et-un étudiants en Psychologie (quatre hommes et 57 femmes, âge moyen = 23 ans) ont participé à cette expérience. Tous les participants étaient de langue maternelle française et

avaient une audition normale et une vision correcte (ou corrigée). Aucun participant n'avait participé au pré-test.

3.3. Procédure

Les participants étaient testés individuellement dans une salle sombre et silencieuse. Avant le déroulement de l'expérience proprement dite, les participants devaient remplir un questionnaire sur leurs connaissances linguistiques. Ils étaient également prévenus que, à la fin de l'expérience, ils seraient soumis à un test de reconnaissance sur les items auditifs. La tâche de décision lexicale avec technique d'amorçage intermodal était similaire à celle utilisée par Gaskell et Marslen-Wilson (1996). L'expérience consistait en trois blocs expérimentaux : un premier bloc d'entraînement pendant lequel 10 paires amorce-cible étaient présentées. Après le bloc d'entraînement, le participant pouvait encore poser des questions. L'expérience principale était composée de deux blocs de 110 items chacun. Les amorces auditives étaient présentées à l'aide d'écouteurs (Sennheiser). A la fin du mot amorce, le mot cible apparaissait sur l'écran d'ordinateur pendant 3000 ms ou jusqu'à ce que le sujet donne sa réponse. Chaque sujet avait pour tâche de dire si la suite de caractères en lettres majuscules figurant sur l'écran correspondait à un mot français ou non. La cible était en Arial 14 points, de couleur noire sur un fond d'écran blanc, afin que la présentation sur écran soit la plus proche d'une situation de lecture réelle. L'expérience a été effectuée avec le logiciel DMDX (Forster & Forster, 2003) sur un ordinateur PC. Les temps de réaction étaient calculés à partir du moment de l'apparition de la cible visuelle. Les réponses ont été enregistrées sur une manette de jeu Logitech Wingman. A la fin de la tâche expérimentale, nous avons donné aux participants une liste de 30 mots dont la moitié correspondait aux mots présentés comme amorces durant l'expérience et 15 mots de remplissage. Les participants avaient pour tâche d'encercler les mots qui leur semblaient familiers. La durée moyenne de l'expérience était de 30 minutes.

3.4. Résultats

Les temps de réponse d'une participante ont été exclus en raison d'un taux d'erreur sur les cibles test (20%) et parce qu'elle était locutrice native du créole haïtien (une langue proche du français, mais dispose d'un répertoire phonémique quelque peu différent). Les résultats de deux participants supplémentaires ont été éliminés en raison d'une moyenne de temps de réponse trop élevée (au-dessus de 850 ms). Nous avons ainsi conservé les résultats de 59 sujets. En outre, les réponses erronées (5,1%) ainsi que les temps de réponses individuels au-dessus de 1200 ms (1,4% des données) ont été éliminés.

Après ces exclusions, les temps de réponse moyens pour les conditions canoniques, assimilation, et contrôle étaient 500 ms, 551 ms, et 592 ms respectivement. Les résultats sont résumés dans le Tableau 3.3. Les effets d’amorçage pour les deux types d’amorces sont représentés dans la Figure 3.1. Ces résultats sont donnés en fonction du voisement sous-jacent. Les temps de réponse, ainsi que les taux d’erreur, ont été soumis à des analyses de variance par sujet et par item avec les facteurs principaux suivants : Nature de l’amorce N_3 à trois modalités (amorce canonique, assimilée, et contrôle) et Voisement sous-jacent V_2 à deux modalités (voisé vs. non voisé). Les facteurs Liste (dans les analyses par item) ainsi que l’assignation à une liste (dans l’analyse par sujet) n’ont pas eu d’effet significatif ($F_s < 1$) et n’ont pas interagi avec les autres facteurs.

		<i>Canonique</i>	<i>Assimilé</i>	<i>Contrôle</i>
<i>[-voisé]</i>	TRs	486 (80)	542 (88)	578 (82)
	% erreurs	1.87	3.88	4.45
<i>[+voisé]</i>	TRs	515 (80)	560 (96)	606 (84)
	% erreurs	3.74	7.04	9.63

Tableau 3.3. Temps de réponse (TRs, écart-types entre parenthèses) et taux d’erreurs (%) pour chaque condition d’amorçage dans l’Expérience 5 (en absence contexte droit). Les résultats sont présentés en fonction des segments non voisés [-voisé] et voisés [+voisé].

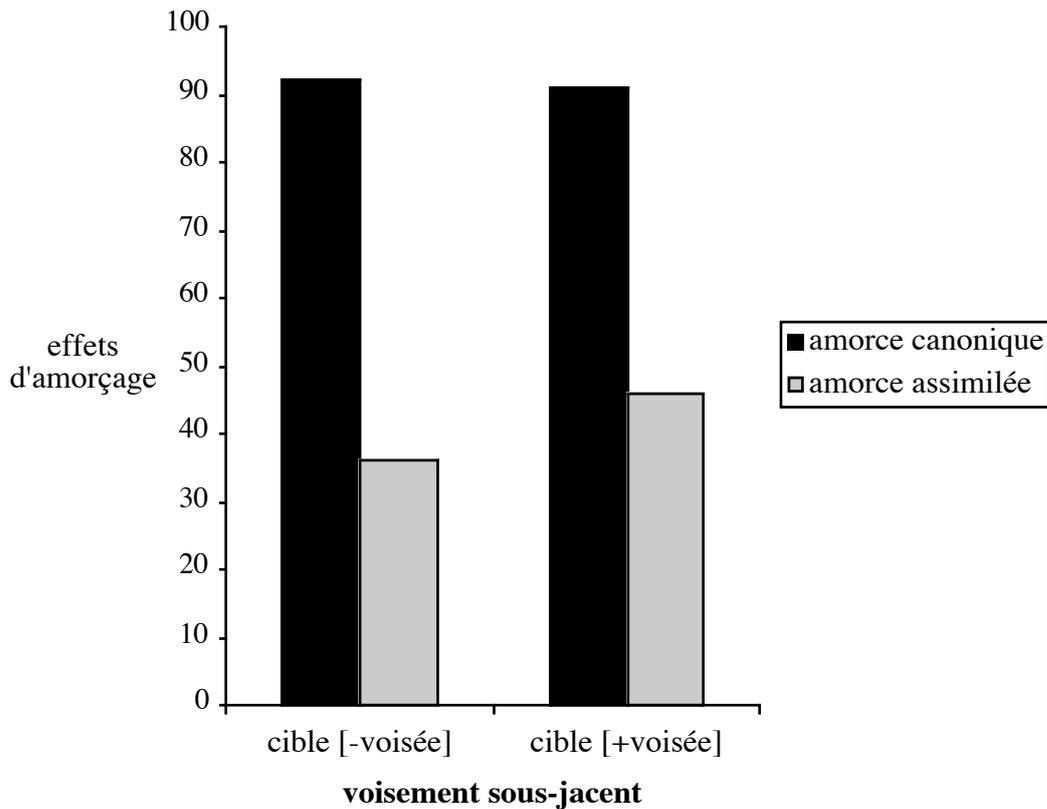


Figure 3.1. Effets d’amorçage pour les amorces non voisées et voisées obtenus dans l’Expérience 5 (en absence du contexte droit).

3.4.1. Analyse sur les temps de réponse

Le facteur principal Nature de l’amorce N_3 a eu un effet significatif, $F_1(2, 114) = 99.056$, $p < .0001$; $F_2(2, 140) = 110.19$, $p < .001$. Le type de voisement a également eu un effet significatif, $F_1(1,57) = 60.95$, $p < .0001$; $F_2(1, 70) = 6.41$, $p < .05$. Cet effet implique que les temps de réponses sont plus courts pour les cibles non voisées que pour les cibles voisées. Toutefois, l’interaction entre la Nature de l’amorce et le type de Voisement n’est pas significative, ni par sujets $F_1(2,114) = 0.64$, $p = 0.52$, ni par items $F_2(2, 140) = 0.69$, $p = 0.5$. De ce fait, nous avons effectué des comparaisons spécifiques de type « canonique » vs. « assimilé » et « assimilé » vs. « contrôle ». Toutes les comparaisons spécifiques selon la nature de l’amorce canonique, assimilé et contrôle se sont avérées significatives (cf. Tableau 3.4).

<i>Comparaisons spécifiques</i>	<i>[-voisé]</i>	<i>[+voisé]</i>
Canonique vs. assimilée	$t(70) = -5.25, p < .0001$	$t(70) = -3.22, p < .01$
Canonique vs. contrôle	$t(70) = -8.96, p < .0001$	$t(70) = -6.83, p < .0001$
Assimilée vs. contrôle	$t(70) = -3.003, p < .01$	$t(70) = -3.15, p < .01$

Tableau 3.4. Expérience 5 : Tests de t pour les comparaisons spécifiques entre les différentes conditions d'amorçage.

3.4.2. Analyse sur les erreurs

Tout comme pour les analyses sur les temps de réponses, deux analyses de variance ont été réalisées : une par participants et une par items. Le taux d'erreurs moyen était de 5.1%. Il est à noter que les taux d'erreurs sont systématiquement plus élevés pour les cibles voisées que pour les cibles non voisées. L'effet du facteur Nature de l'amorçage N_3 est significatif, $F_1(2, 114) = 7.88, p < .0001$; $F_2(2, 140) = 9.78, p < .0001$, tout comme l'effet de Voisement sous-jacent V_2 , $F_1(1, 57) = 20.26, p < .0001$; $F_2(1, 70) = 5.42, p < .05$. L'interaction entre la Nature de l'amorce et Voisement n'est pas significative, ni par sujet $F_1(2, 114) = 1.47, p = 0.23$, ni par items $F_2(2, 140) = 1.38, p = 0.26$.

3.5. Discussion

Le principal résultat de cette expérience est que les amorces canoniques et assimilées donnent lieu à un profil d'amorçage différent. L'effet d'amorçage est plus important quand le mot cible est précédé par l'audition de sa forme canonique (*une jupe* [ynʒyp] – JUPE) que quand il est précédé par celle de sa forme assimilée (*une jube* [ynʒyb] – JUPE). Cette différence d'amorçage entre les amorces canoniques et assimilées est observée aussi bien pour les mots fortement assimilés (i.e. avec consonne finale non voisée) que pour les mots faiblement assimilés (i.e. avec consonne finale voisée). Ce résultat ne s'accorde pas avec celui obtenu par Gaskell & Marslen-Wilson (1996). En effet, ces auteurs ont trouvé que la forme assimilée [l i m] du mot *lean* amorçait la cible LEAN autant que sa forme canonique [l i n]. Toutefois, il est important de noter que dans leur étude, l'amorce critique était précédée par une phrase qui était biaisée sémantiquement vers l'interprétation canonique (e.g., *We have a house full of fussy eaters. Sandra will only eat LEAN bacon*). La présence d'un tel contexte favorisait sans doute l'interprétation de la forme canonique. Ceci aurait pu annuler tout effet de distorsion ralentissant les temps de réponse sur la cible, d'où une tolérance du système de traitement

pour ces distorsions phonologiques. Utilisant le même paradigme expérimental, nos résultats ont clairement montré qu'en absence de tout biais sémantique, le système perceptif fait une distinction entre les formes canoniques et les formes assimilées dans la mesure où les effets d'amorçage ne sont pas identiques pour les deux types d'amorces.

Nos résultats sont en accord avec ceux obtenus par Gumnior, Zwitterlood et Bölte (2005). Ces auteurs ont récemment examiné l'assimilation de lieu d'articulation *au sein* des mots composés en allemand, tels que *Bahngleis*. Ils ont mesuré les effets d'amorçage pour les mots cibles (e.g., BAHN) quand ceux-ci étaient précédés soit par les formes canoniques (e.g., *Bahngleis*), soit par les formes assimilées appropriées résultant des contextes où le changement assimilatoire est compatible avec le phonème donnant lieu à l'assimilation (e.g., *bahngleis*, la coronale prend le lieu vélaire dû à la présence de la consonne vélaire /g/), ou encore par des formes assimilées non appropriées (e.g., *bahmgleis*, la coronale prend le lieu labial en présence du contexte vélaire). Il n'y avait pas de différences significatives entre les effets d'amorçage provoqués par les formes assimilées appropriées et non appropriées (i.e. *bahngleis* a amorcé BAHN tout autant que *bahmgleis*). En revanche, il y avait une différence significative entre les effets d'amorçage provoqué par les formes canoniques (*Bahngleis*) et les formes assimilées appropriées (*bahngleis*, Expérience 2). Ces résultats s'accordent avec les prédictions du modèle FUL (cf. chapitre 1, section 2.2.1), selon lequel la viabilité du contexte phonologique ne module pas le traitement perceptif des segments assimilés. Cependant, comme le notent les auteurs, il y avait une différence significative entre les effets d'amorçage provoqués par les formes canoniques et assimilées, ce qui est problématique pour ce modèle. Le modèle FUL ne semble pas pouvoir rendre compte de nos résultats, car cette théorie, s'inspirant largement des théories phonologiques, considère tout changement par assimilation comme étant de nature dichotomique. En tout état de cause, nos expériences n'ont pas permis de tester directement les prédictions de la théorie de sous-spécification, car il aurait fallu insérer une condition supplémentaire avec des contextes « non viables », c'est-à-dire des contextes qui ne motivent pas de changements par assimilation (e.g., *robe [rop] grise*, cf. Gaskell & Marslen-Wilson, 1996, 1998; Coenen et al., 2001 ; Gumnior et al., 2005).

La différence d'amorçage entre les amorces canoniques et assimilées était analogue pour les mots ayant une consonne finale non voisée (i.e. fortement assimilée) et voisée (i.e. partiellement assimilée). Compte tenu du fait que ces types d'items diffèrent sur de nombreuses dimensions, l'interprétation de ce résultat reste ouverte.

L'Expérience 5 a permis de montrer qu'en absence du contexte droit, les mots réalisés sous leur forme canonique donnent lieu à un effet d'amorçage plus important que leur réalisation sous forme assimilée. Nous avons conçu l'Expérience 6 afin d'évaluer plus précisément l'influence du contexte droit sur le traitement des mots ayant différents degrés d'assimilation. Nous avons proposé le voisement comme l'un des facteurs principaux permettant d'opérationnaliser les degrés de d'assimilation. De façon approximative, les segments non voisés correspondent à des assimilations fortes, alors que les voisés correspondent à ces segments partiellement ou peu assimilés. S'il s'avère que les auditeurs prennent appui sur le contexte droit afin de récupérer la forme sous-jacente du segment assimilé, nous devrions observer un effet de contexte droit plus important pour les mots fortement assimilés (i.e. avec une consonne finale non voisée) que pour les mots faiblement assimilés (i.e. avec une consonne finale voisée).

4. EXPERIENCE 6 □ AMORÇAGE AVEC CONTEXTE DROIT

4.1. Matériel expérimental

Le matériel expérimental était strictement identique à l'Expérience 5, mais cette fois-ci l'amorce entière était présentée auditivement aux participants (e.g., condition canonique, *une jupe serrée* [ynʒypsɛrɛ] – JUPE ; condition assimilée, *une jupe grise* [ynʒybgriz] – JUPE ; condition contrôle, *un acte parfait* [œnaktparfɛ] – JUPE). De nouveau, la cible visuelle était présentée à la fin de l'amorce contenant le segment critique. La répartition des mots de remplissage était identique à celle utilisée dans l'Expérience 5.

4.2. Sujets

Soixante-deux étudiants en Psychologie (11 hommes et 51 femmes, âge moyen = 23 ans) ont participé à la tâche de décision lexicale avec technique d'amorçage intermodal. Tous les participants étaient de langue maternelle française et avaient une audition et vision correcte (ou corrigée). Aucun participant n'avait participé aux deux expériences précédentes.

4.3. Procédure

La procédure était identique à celle de l'Expérience 5, à part que la séquence auditive, constituant l'amorce, était présentée entièrement (e.g., *une jupe serrée* [ynʒypsɛrɛ] – JUPE), rendant ainsi le contexte disponible aux auditeurs. La cible visuelle (e.g., JUPE) était

présentée à la fin du mot contenant le segment critique, c'est-à-dire au milieu du syntagme nominal.

4.4. Résultats

Les résultats de quatre participants ont été exclus des analyses en raison des temps de réponse moyen au-dessus de 800 ms. De plus, les réponses erronées (4,24%) ainsi que les réponses aux items individuels au-dessus de 1200 ms ont été exclues de toute analyse ultérieure (0,48%).

Les temps de réponse et les taux d'erreur sont résumés dans le Tableau 3.5. Après les exclusions, les temps de réponses moyens étaient 506 ms pour la condition canonique, 532 ms pour la condition assimilée, et 583 ms pour la condition contrôle. Les effets d'amorçage pour les amorces non voisées et voisées sont présentés dans la Figure 3.2.

		<i>Canonique</i>	<i>Assimilé</i>	<i>Contrôle</i>
<i>[-voisé]</i>	TRs	495 (80)	513 (69)	577 (75)
	% erreurs	1.72	1.87	3.88
<i>[+voisé]</i>	TRs	518 (73)	551 (71)	589 (81)
	% erreurs	4.02	4.17	9.77

Tableau 3.5. Temps de réponse (TRs, écart types entre parenthèses) et taux d'erreurs (%) pour chaque condition d'amorçage dans l'Expérience 6 (amorçage avec contexte droit). Les résultats sont présentés en fonction des segments non voisés [-voisé] et voisés [+voisé].

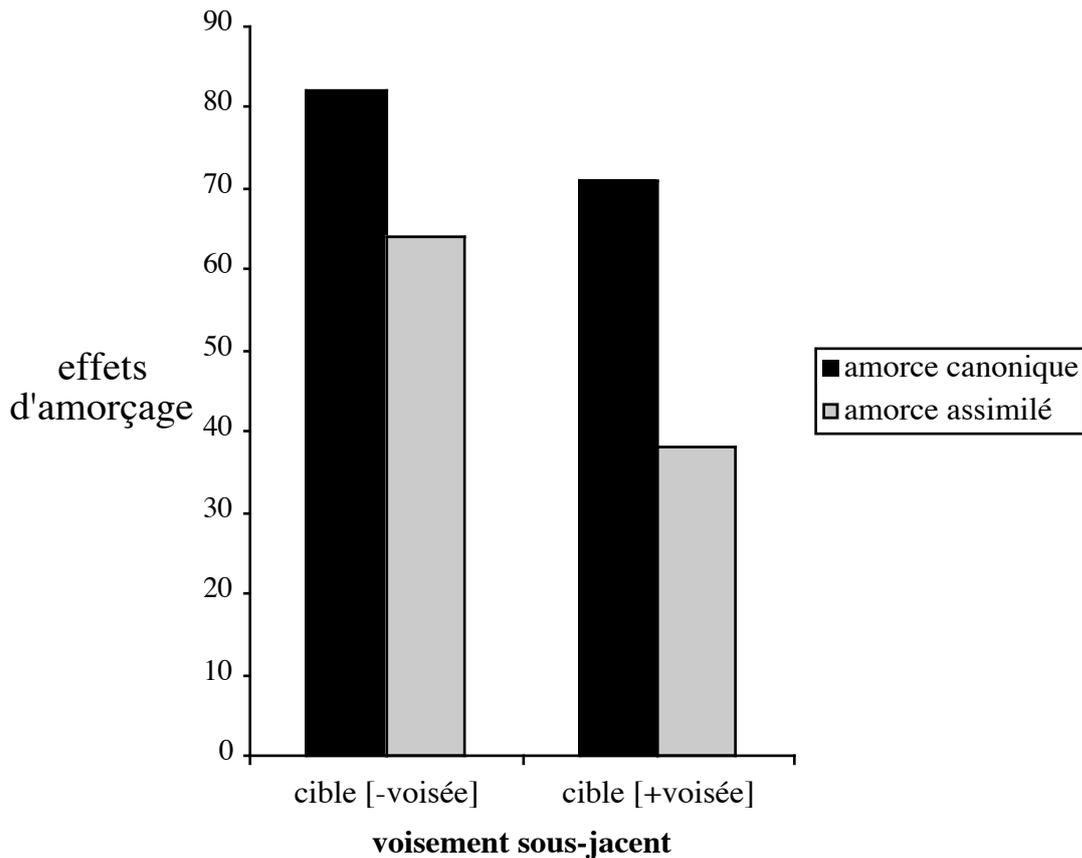


Figure 3.2. Effets d’amorçage pour les amorces non voisées et voisées obtenus dans l’Expérience 6.

4.4.1. Analyses sur les temps de réponse

Les temps de réponse ainsi que les taux d’erreur ont été soumis à des analyses de variance par sujet et par item avec les facteurs principaux suivants : Nature de l’amorce N_3 à trois modalités (canonique, assimilé, contrôle) et type de Voisement sous-jacent V_2 à deux modalités (amorces voisées vs. non voisées).

La nature de l’amorce N_3 a eu un effet significatif, $F_1(2,114) = 84.68, p < .0001$; $F_2(2,140) = 97.92, p < .001$. Le type de voisement V_2 a également eu un effet significatif, $F_1(1,57) = 52.73, p < .001$; $F_2(1, 70) = 5.64, p < .05$. Les temps de réponses ont été plus courts pour les cibles non voisées que pour les cibles voisées. Ce qui est à noter, c’est que contrairement à l’Expérience 5, l’interaction entre la relation d’amorçage et le type de voisement s’est avérée significative par sujets $F_1(2,114) = 4.31, p < .05$, (mais non pas par items, $F_2(2, 140) = 1.74, p = 0.18$). Les résultats obtenus dans cette expérience confirment que les amorces correspondant à la réalisation canonique de la cible déterminent un effet d’amorçage plus important que les amorces assimilées. Toutefois, contrairement aux résultats obtenus dans l’Expérience 5, les effets d’amorçage varient selon la nature du voisement sous-jacent. Ceci indique que

l'introduction du contexte droit a affecté de manière différente les mots fortement et partiellement assimilés.

4.4.2. *Analyses sur les erreurs*

Deux analyses de variance ont été réalisées, par participants et par items, pour évaluer l'influence des facteurs principaux sur les taux d'erreurs (4.24%). L'effet de relation d'amorçage s'est avéré de nouveau significatif, $F_1(2, 114) = 10.52, p < .0001$; $F_2(2, 140) = 12.19, p < .0001$, tout comme l'effet de voisement, $F_1(1, 57) = 28.13, p < .0001$; $F_2(1, 70) = 3.66, p < .05$. Tout comme pour l'analyse sur les temps de réponses, l'analyse sur les erreurs a révélé une interaction entre relation d'amorçage et voisement par participants $F_1(2, 114) = 3.42, p < .05$, alors que dans l'analyse par items n'effleure que le seuil de significativité, $F_2(2, 140) = 2.61, p = .07$.

4.4.3. *Analyse combinée Expériences 5 et 6*

Afin d'étudier le rôle du contexte selon le degré d'assimilation des mots amorces, nous avons conduit une analyse combinée des résultats obtenus dans les Expériences 5 et 6. Cette analyse intègre comme facteur inter-groupes le Contexte droit C_2 à deux modalités (« présence » vs. « absence ») et comme pour facteurs intra groupes Nature de l'amorce N_3 à trois modalités (« canonique », « assimilé » et « contrôle ») et Voisement sous-jacent de la cible V_2 à deux modalités (« segment non voisé » vs. « segment voisé »).

Nous avons obtenu un effet significatif pour le facteur Nature de l'amorce N_3 $F_1(2,228) = 180,89, p < .0001$; $F_2(2, 140) = 97,92, p < .0001$, ainsi que pour le Voisement sous-jacent V_2 , $F_1(1, 114) = 113.46, p < .0001$; $F_2(1,70) = 5.64, p < .05$. Le facteur Contexte ne s'est pas révélé significatif, ni par sujets $F_1(1, 114) = 0.34, p = 0.56$, ni par items $F_2(1, 70) = 0.80, p = 0.37$. Cette absence d'effet montre qu'au niveau global, les temps de réponse sont analogues dans les Expériences 5 et 6. L'interaction entre les trois facteurs principaux Contexte C_2 , Nature de l'amorce N_3 et type de Voisement V_2 est significative par sujets, $F_1(2, 228) = 3.64, p < .05$, et marginalement significative par items, $F_2(2, 140) = 3.34, p = 0.10$.

Afin de mieux saisir le sens de cette interaction, nous avons effectué des analyses de variance séparées pour les cibles voisées et non voisées. En ce qui concerne les cibles non voisées, on observe une interaction significative entre Contexte C_2 et Nature de l'amorce N_3 , par sujets $F_1(2, 228) = 4.23, p < .05$, mais non par items, $F_1(2,70) = 0.35, p = 0.70$. Ce résultat indique que la présence du contexte droit affecte le traitement perceptif des mots non voisés, i.e. fortement assimilés, tandis qu'il ne modifie guère le traitement des amorces canoniques.

L'interaction entre les deux facteurs est présentée dans la Figure 3.3. Les effets d'amorçage sont indiqués en ordonnée, alors que le type d'amorce (canonique vs. assimilée) est affiché en abscisse. Le facteur Contexte est présenté dans le corps du graphique. Comme on peut le voir dans ce graphique, les amorces assimilées non voisées ont donné lieu à un effet d'amorçage plus important en présence d'un contexte droit qu'en absence de ce contexte. En revanche, quand les mêmes mots sont réalisés sous leur forme canonique, l'effet d'amorçage est comparable en présence ou en absence du contexte droit. Ce résultat confirme le rôle important joué par le contexte droit pour le traitement des mots fortement assimilés.

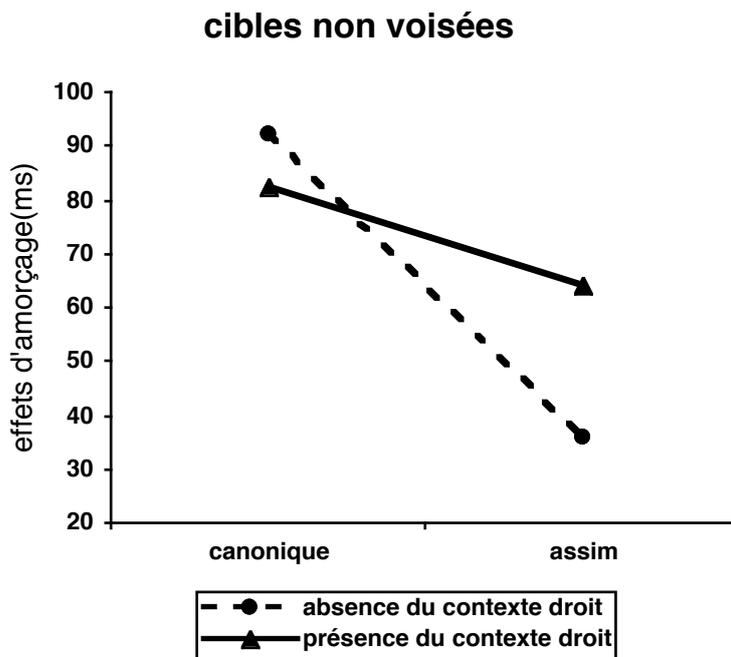


Figure 3.3. Effets d'amorçage pour les segments non voisés, i.e. fortement assimilés. Une interaction significative est observée entre les deux facteurs principaux Nature de l'amorce et Contexte droit.

Quant aux cibles voisées, aucune interaction entre le facteur Nature de l'amorce et le facteur Contexte droit n'a été observée, ni par sujets $F_1(2, 228) = 1.44, p = .24$, ni par items, $F_2(2, 70) = 0.11, p = .89$. Ainsi, l'absence d'interaction suggère que la présence du contexte droit n'a pas affecté le traitement des mots voisés, i.e. faiblement assimilés par référence à celui observé pour les amorces canoniques (cf. Figure 3.4).

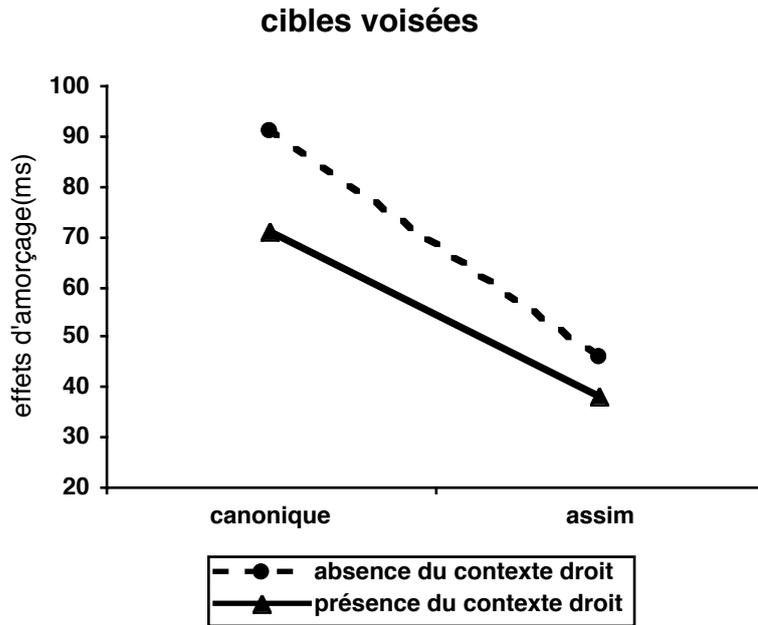


Figure 3.4. Effets d'amorçage observés pour les segments voisés, i.e. faiblement assimilés. L'interaction entre les deux facteurs principaux Nature de l'amorce et Contexte droit n'est plus significative.

Même si les amorces assimilées donnent lieu à un effet d'amorçage plus faible que celui observé pour les amorces canoniques, ces différences de traitement ne varient pas en fonction de la présence ou de l'absence du contexte droit.

4.5. Discussion

Dans l'Expérience 6, les amorces auditives étaient composées des phrases nominales afin de rendre le contexte droit disponible aux participants. Nous avons observé que, dans de telles conditions, la présence du contexte droit a significativement facilité le traitement des cibles non voisées quand celles-ci étaient précédées par des amorces fortement assimilées (e.g., *jupe* devenu [ʒy \underline{b}]). En revanche, la présence du contexte droit n'a guère modifié le traitement des cibles voisées précédées par des amorces partiellement assimilées (e.g., *robe* devenu [rɔ \underline{p}]) mais à un moindre degré. L'analyse combinée des Expériences 5 et 6 a mis en évidence une interaction entre les trois facteurs principaux Contexte, Nature d'amorçage et Type de voisement. Cette interaction justifiait deux analyses séparées. L'une n'incluait que les temps réponses aux cibles non voisées, alors que l'autre n'incluait que ceux aux cibles voisées. Rappelons que nous avons opérationnalisé le degré d'assimilation par le voisement sous-jacent. Ainsi, les segments non voisés (/p, t, k/) correspondent à des segments fortement assimilés alors que les segments voisés (/b, d, g/) correspondent à des segments partiellement

ou peu assimilés. En ce qui concerne les segments non voisés, une interaction significative entre les facteurs Nature de l'amorce et Contexte a été observée, alors que l'analyse comportant les segments voisés n'a pas montré une telle interaction. Ces résultats suggèrent que la présence du contexte droit a affecté le traitement des cibles non voisées, i.e., précédées par des amorces fortement assimilées, alors qu'elle n'a pas affecté le traitement des cibles voisées, i.e. précédées par des amorces peu assimilées.

Nos résultats sont en accord avec les prédictions du modèle probabiliste de Gaskell (2003), selon lequel le rôle du contexte dans le traitement des segments assimilés varie en fonction de leur degré d'assimilation. Les effets d'amorçage ont montré que le traitement perceptif est facilité en présence du contexte droit dans le cas des assimilations des plosives non voisées. Nos études acoustiques et perceptives préalables ont mis en évidence que les obstruents non voisés subissent fréquemment des assimilations complètes. Cette facilitation peut être expliquée en termes d'inférence régressive. Quand un segment a subi une assimilation quasi-complète, l'auditeur aurait recours au contexte suivant afin de dériver son identité sous-jacente (cf. Gaskell & Marslen-Wilson, 1996, 1998). En revanche, le contexte droit joue un rôle nettement moins important quand les cibles sont précédées par des amorces voisées. Préalablement, nous avons mis en évidence que les plosives voisées subissent moins d'assimilation que leurs contreparties non voisées. Dans le cas des segments partiellement ou peu assimilés, la présence du contexte droit est moins nécessaire, car il est supposé que la présence des « traces » acoustiques dans le signal permettrait aux auditeurs d'accéder à la forme sous-jacente de ces segments.

Afin d'effectuer les analyses de variance classiques, nous avons opérationnalisé le degré d'assimilation par le facteur Voisement sous-jacent V_2 à deux modalités (non voisé vs. voisé). Il est important de revenir sur le fait que cette factorisation ne reflète pas pour autant la réalité phonétique, car nous avons constaté dans le chapitre précédent que l'assimilation de voisement est un phénomène gradué plutôt que dichotomique. En effet, même si les plosives non voisées subissent très fréquemment des assimilations (quasi-complètes), ceci n'est pas toujours le cas. A l'inverse, il arrive que les plosives voisées subissent parfois des assimilations complètes. Nous avons alors jugé pertinent de conduire des analyses supplémentaires de régression multiple, qui ne nécessitent pas une telle dichotomisation

(Baayen, Tweedie, & Schreuder, 2002; Baayen, 2004)¹⁶. Dans ces analyses, les degrés d'assimilation pour tous les items expérimentaux présentés dans la condition « assimilation » sont pris en considération et insérés dans le modèle statistique. Les résultats confirment ceux des analyses de variance classiques. La présence du contexte droit joue un rôle sur les temps de réponse aux mots cibles et ceci d'autant plus que les mots amorces sont fortement assimilés.

5. DISCUSSION GENERALE

Dans ce chapitre, nous avons voulu étudier les conséquences perceptives du traitement des segments ayant subi des degrés d'assimilation variés. La littérature sur l'assimilation de lieu d'articulation, notamment en anglais, a donné lieu à des interprétations divergentes concernant le rôle du contexte droit. Gaskell et collègues (Gaskell & Marslen-Wilson, 1996, 1998) ont surtout insisté sur le rôle de ce contexte qui permettrait aux auditeurs d'inférer l'identité sous-jacente du segment assimilé. Un mécanisme cognitif d'inférence régressive sous-tendrait le traitement perceptif des segments assimilés. D'autres auteurs, comme Gow (2001, 2002, 2003), ont fait remarquer que le contexte n'est pas nécessaire pour accéder à la forme canonique du segment assimilé, car ces segments ne sont jamais complètement transformés. En effet, les travaux en phonétique expérimentale que nous avons passés en revue dans le chapitre 1 ont révélé que l'assimilation représente souvent un processus phonétique gradué. En conséquence, le signal de parole contient des indices acoustico-phonétiques fins qui peuvent être employés par l'auditeur afin de dériver la forme canonique du segment assimilé.

L'objectif principal que nous nous sommes posé dans ce chapitre était de tester l'hypothèse selon laquelle l'influence du contexte droit dans le traitement perceptif des segments assimilés varie en fonction de la force ou ampleur de l'assimilation. Précédemment, nous avons démontré que l'assimilation de voisement en français implique souvent un processus phonétique gradué. Qui plus est, dans le chapitre 2, nous avons pu constater que la force d'assimilation des consonnes en position finale est modulée par le voisement sous-jacent.

¹⁶ Nous avons conduit les analyses de régression multiple en collaboration avec le Prof. R. Harald Baayen (*Interfaculty Research Unit for Speech*, Radboud Universiteit Nijmegen, et *Max Planck Institute for Psycholinguistics*, Nijmegen, Pays-Bas). Pour plus de détails sur ces analyses, le lecteur intéressé peut se référer à l'Annexe 13.

Bien évidemment, il nous a été impossible de demander à notre locuteur de prononcer les mots tout en contrôlant leur degré d'assimilation. Pour des raisons méthodologiques évidentes, nous avons indirectement opérationnalisé les différents degrés d'assimilation en utilisant les mots voisés et non voisés du fait qu'il existe une asymétrie d'assimilation entre ceux-ci. Nous avons considéré en première approximation que les mots non voisés sont fortement, voire complètement assimilés. De même, nous avons considéré que les mots voisés sont faiblement ou peu assimilés. Cependant, nous sommes bien conscients du fait que cette opérationnalisation est une méthode approximative, ne permettant peut-être pas d'appréhender le traitement de l'assimilation dans toute sa complexité. Aussi, il peut parfois arriver que les plosives non voisées ne subissent guère d'assimilation alors que les plosives voisées sont assimilées complètement, (ce qui est d'ailleurs reflété dans les jugements perceptifs du pré-test, voir section 2.5)

Les effets d'amorçage observés dans l'Expérience 5 ont permis de montrer que les effets d'amorçage sont plus importants quand les mots cibles sont précédés par leur forme canonique que par leur forme assimilée. L'effet d'amorçage est comparable pour les cibles voisées et non voisées. Les résultats de l'Expérience 6 ont mis en évidence une interaction entre le type d'amorce et le type de cible (voisée vs. non voisée). La présence de cette interaction dans l'Expérience 6 nous a conduit à effectuer des analyses séparées pour les cibles non voisées et voisées. Ces analyses confirment l'hypothèse selon laquelle la présence du contexte affecte le traitement des mots non voisés tandis qu'elle ne modifie pas celui des mots voisés.

Soit un mot fortement assimilé tel que *jupe*, prononcé [ʒyb]. L'obstruente finale de ce mot va bénéficier de la présence du contexte droit (e.g., *grise*), car c'est à partir du segment suivant que l'auditeur français sait implicitement que les obstruents non voisées peuvent devenir voisées en présence du contexte voisé, ce qui est illustré dans Figure 3.5. Il est à noter que l'inférence régressive n'a de sens qu'en faisant appel aux connaissances phonologiques implicites¹⁷ que l'auditeur est supposé mettre en œuvre pour la résolution perceptive des segments assimilés. Par conséquent, ce type de mécanisme perceptif est spécifique à la langue en question.

¹⁷ Voir à cet égard van Dommelen (1983), qui parle explicitement du rôle des « attentes linguistiques » (*linguistic expectancy*) de l'auditeur, quand il évoque le rôle du phonème contextuel (cf. /p/ dans *plate* ou /g/ dans *grise*) dans le traitement perceptif du voisement de /t-/d/ dans *rate* / *rade*.

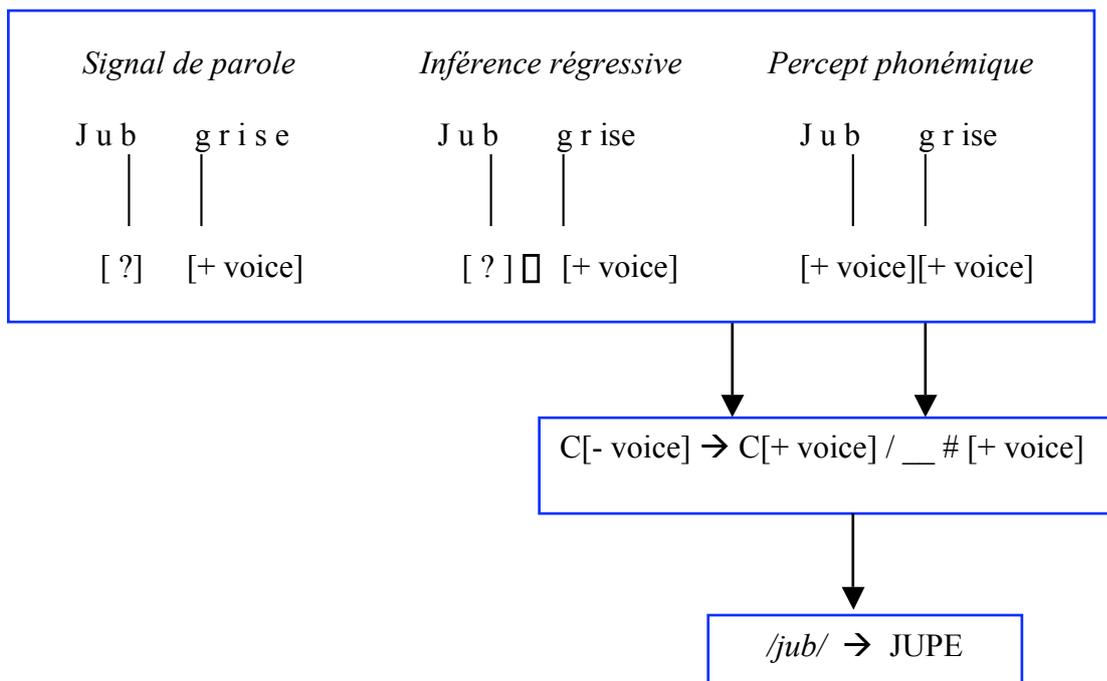


Figure 3.5. Mécanisme d'inférence régressive sous forme schématisée pour un mot complètement assimilé tel que *jupe* prononcé [Zub].

La présence des traces acoustiques dans les mots partiellement assimilés permettrait à l'auditeur d'avoir accès à leur forme sous-jacente. Nous avons vu dans l'introduction que Gow (2003, voir aussi Gow & Zoll, 2002) a avancé la théorie d'alignement des traits pour rendre compte du traitement perceptif de l'assimilation de lieu d'articulation en anglais. On pourrait envisager des mécanismes similaires pour le traitement perceptif de l'assimilation de voisement (ce qui est illustré dans la Figure 3.6).

Il semblerait donc que différents types de processus perceptifs soient à l'origine du traitement des segments plus ou moins fortement assimilés. Le fait que le type de processus perceptif varie selon que les obstruents soient voisés ou non résulte directement des degrés différents d'assimilation observés pour ces deux types d'obstruents.

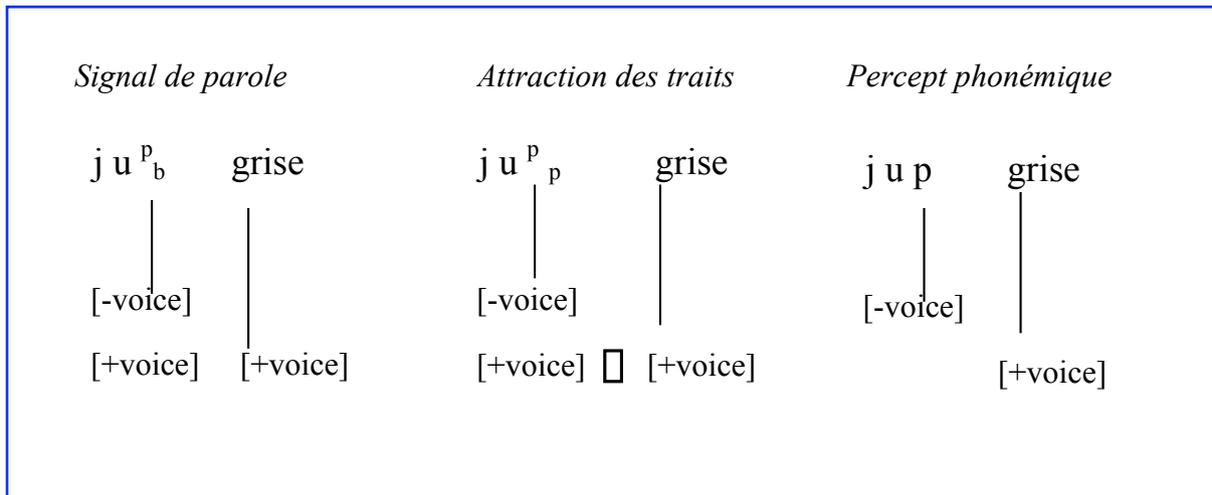


Figure 3.6. Mécanisme perceptif d'alignement des traits pour le segment partiellement assimilé /p/ dans *jupe*. Par l'attraction du trait [+voice] par le contexte droit /g/, la forme sous-jacente non voisée /p/ dans *jupe* devient accessible.

L'inférence régressive est un mécanisme plausible pour rendre compte du traitement perceptif des mots fortement assimilés, tandis que la présence des traces de voisement dans les mots partiellement assimilés devrait suffire pour son identification, sans faire appel à d'autres mécanismes perceptifs¹⁸. Toutefois, étant donné les différences intrinsèques entre les deux types d'assimilation¹⁹ et l'asymétrie d'assimilation observée entre obstruents voisées²⁰ et non voisées, ceci reste encore à confirmer.

¹⁸ Dans les travaux de van Dommelen (1983), on peut trouver une sorte de réconciliation entre les deux types de mécanismes avancés par Gaskell et collègues d'une part, et Gow d'autre part. Cet auteur, qui a étudié le voisement en français, évoque également la possibilité que les auditeurs utilisent leurs connaissances phonologiques implicites ou « attentes linguistiques » lorsqu'ils utilisent les informations présentes dans le contexte droit lors du traitement perceptif du voisement. L'estimation perceptive du voisement durant l'occlusion de C₁ dépendrait, parmi d'autres facteurs, du statut phonémique de C₂ dans la perception d'une séquence consonantique de type C₁ # C₂. La suite du raisonnement rappelle la théorie d'alignement des traits (proposée par Gow), en soulignant que, en présence de la consonne voisée C₂, toute trace acoustique de voisement présente dans C₁ doit être associée à C₂ (qui a été dévoisée). Pourtant, d'après Gow, la présence du contexte droit n'est jamais nécessaire, alors que van Dommelen insiste sur sa nécessité pour le traitement perceptif quand il dit que ([...] a decision as to the identity of C1 cannot be taken until the phonemic status of C2 has been ascertained, van Dommelen, 1983 : 22, nos italiques).

¹⁹ Bien qu'ils aient directement comparé l'assimilation de voisement (en hongrois) et de lieu d'articulation (en anglais), Gow et Im (2004) ne se prononcent pas sur les conséquences de la théorie d'alignement pour l'assimilation de voisement. En revanche, ces auteurs se penchent sur les différences acoustiques entre ces deux types d'assimilation. Étant donné la nature périodique de voisement, les indices de voisement s'étendent sur un

Ces résultats principaux confirment notre hypothèse de départ selon laquelle le rôle du contexte lors du traitement perceptif de l'assimilation est en effet modulé par la force avec laquelle les phonèmes ont subi une transformation, et permettent de valider les prédictions du modèle probabiliste élaboré par Gaskell (2003), inspiré largement de son ancien modèle (Gaskell et al., 1995). Cette fois-ci, le SRN est limité aux niveaux des traits de façon à pouvoir focaliser sur les effets temporels dans la perception de l'assimilation de lieu d'articulation. L'architecture du SRN est illustrée dans la Figure 3.7.

Après entraînement, le modèle était capable d'interpréter une séquence telle que *lean* (*bacon*) comme une représentation de surface de la forme canonique *lean* (*bacon*). Afin d'évaluer les effets de contexte, le SRN dispose de trois ensembles de nœuds de sortie, qui représentent les formes sous-jacentes des segments d'entrée actuel, précédent et suivant. Un effet régressif est reflété par un changement d'évaluation d'un phonème une fois que le contexte suivant est présenté au réseau, alors qu'un effet progressif apparaît dans les nœuds correspondant au segment suivant. Ce qui distingue le modèle actuel de sa version précédente, c'est la prise en compte de la nature graduée de l'assimilation. C'est-à-dire qu'un continuum de 0-100% d'assimilation a été utilisé dans les simulations. Dans l'ancien SRN (Gaskell et al., 1995), deux cas d'assimilation avaient été implémentés : a) aucune assimilation (représentée comme « $C^+ = 1, C^- = 0$ », où C^+ désigne le trait coronal, et C^- désigne le trait non coronal) et b) l'assimilation complète (représentée comme « $C^+ = 0, C^- = 1$ »). Dans le nouveau modèle, des cas intermédiaires d'assimilation ont été incorporés en réduisant la valeur du trait coronal et en augmentant la valeur du trait non coronal (e.g., 40% d'assimilation était représenté comme « $C^+ = 0.6, C^- = 0.4$ »). Les résultats de ces simulations ont mis en évidence les deux types d'effets observés dans la littérature. Premièrement, quand l'assimilation est complète, l'inférence du segment sous-jacent exige l'utilisation du contexte droit. Ce résultat s'accorde avec les observations de Gaskell et Marslen-Wilson, 1996, 1998, qui ont conduit ces auteurs à proposer le mécanisme d'inférence régressive.

intervalle plus long que les indices de lieu (qui sont souvent identifiés à un seul point dans le temps, comme la dernière période de la hauteur (*pitch*) précédant l'occlusion dans une occlusive en position syllabique finale). Etant donné les différences physiques entre l'assimilation de voisement et de lieu, les indices de voisement doivent être intégrés dans une fenêtre temporelle plus longue que les indices de place. Des fenêtres d'intégration plus larges donnent lieu à plus d'incertitude quant à l'alignement correct des indices. Il en résulte que les auditeurs feront plus d'erreurs d'analyse impliquant le voisement que la place d'articulation.

²⁰ Il paraît difficile d'envisager l'attraction du trait [-voisé], traduisant l'absence du voisement, dans le cas d'assimilation de dévoisement des mots voisés tels que *robe*.

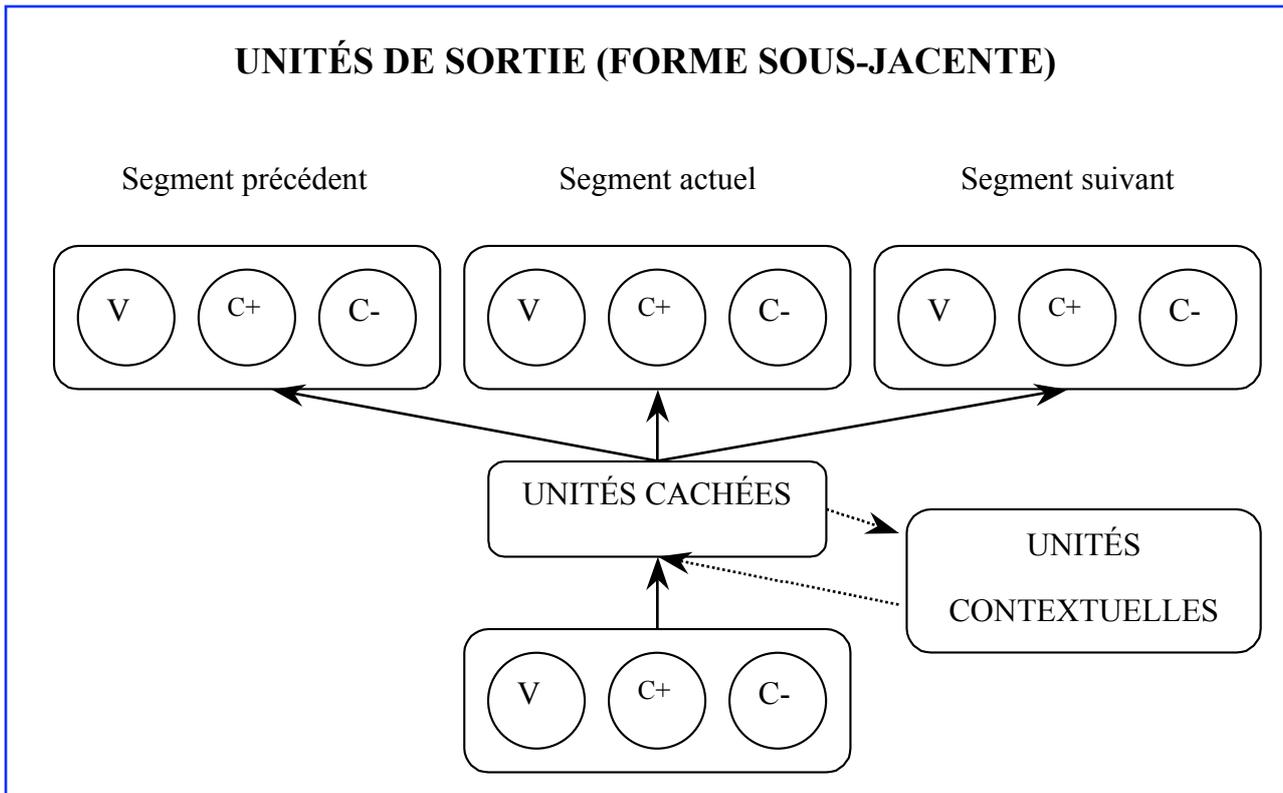


Figure 3.7. Architecture du réseau employé par Gaskell (2003). Les unités d'entrée envoient des informations à une couche d'unités cachées récupérant les informations sur des cycles temporels successifs, grâce à la présence des liens récurrents (représentés par des flèches pointillées). Les unités cachées sont entraînées pour être projetées sur des représentations des segments précédent, actuel, et suivant à la sortie. Les segments à l'entrée et à la sortie sont représentés sous forme de voyelles (V), consonnes coronales (C+) ou consonnes non coronales (C-).

Deuxièmement, quand l'assimilation est partielle, les indices acoustiques de surface doivent permettre d'inférer le phonème sous-jacent. Une conséquence de ce dernier résultat est que le modèle est en mesure de prédire les propriétés phonétiques du segment qui suit le segment assimilé. C'est en effet ce type de résultat que Gow (2001) a obtenu dans ses expériences de détection de phonèmes : les temps de détection du phonème suivant le phonème assimilé est réduit si cette assimilation était partielle. Ceci est compatible avec l'idée que le segment partiellement assimilé contient en fait des informations sur les deux lieux d'articulation. Les données empiriques suggèrent que le traitement de l'assimilation de lieu est basé sur une évaluation régressive ou progressive en fonction de la force d'assimilation. Un modèle probabiliste tel celui présenté par Gaskell peut donc bien rendre compte des données expérimentales obtenues sur l'assimilation de lieu d'articulation en anglais. Nos résultats constituent les premières données empiriques à l'appui des prédictions de ce modèle. Nous avons démontré que les deux types de mécanismes cognitifs proposés, à savoir l'inférence

régressive par le contexte droit et la présence des traces acoustiques dans le signal, sont en fait compatibles dès lors que l'on prend en considération les caractéristiques acoustiques fines du matériel linguistique et, dans le cas qui nous intéresse, le degré d'assimilation des segments de parole.

CHAPITRE IV.

REPRESENTATION LEXICALE DES SEGMENTS ASSIMILES ET ACTIVATION DU RESEAU LEXICO-SEMANTIQUE

1. INTRODUCTION

Dans le chapitre 2, nous avons pu mettre en évidence que la production de la parole assimilée est parfois modulée par des facteurs phonétiques et lexicaux, tels que le voisement sous-jacent ou encore la présence d'un compétiteur lexical potentiel. Le chapitre 3 avait pour objectif d'étudier les conséquences perceptives de la parole assimilée et nous avons montré que le rôle du contexte droit lors du traitement perceptif varie en fonction de la force de l'assimilation. Plus un segment est assimilé, plus l'auditeur prend appui sur le contexte suivant afin d'inférer son identité sous-jacente. En revanche, moins le segment est assimilé, moins l'auditeur aurait recours au contexte, car la présence des traces acoustiques dans le signal assimilé lui permettrait d'accéder à sa forme canonique.

Dans ce chapitre, nous poursuivrons nos recherches sur le traitement perceptif de la parole assimilée en nous focalisant sur l'activation du réseau lexico-sémantique des formes sous-jacentes. L'objectif principal est d'établir si la présence de traces acoustiques dans la forme assimilée, correspondant au mot sous-jacent, induit lors de la présentation une activation de son réseau lexico-sémantique. L'étude du traitement des mots ambigus par assimilation (MAA, cf. chapitre 2) semble particulièrement appropriée pour soulever cette question car il est présumé que la présentation d'un mot ambigu par assimilation tel que *soute* active deux formes lexicales (et leurs réseaux lexico-sémantiques): celle qui correspond à la forme sous-jacente *soute* et celle qui correspond à la forme de surface *soude*.

1.1. Activation lexicale et proximité phonologique

Comme nous l'avons déjà souligné auparavant, certains auteurs ont proposé qu'en dépit des variations phonologiques présentes dans le signal de parole, des indices acoustiques fins sont préservés et peuvent être utilisés par l'auditeur lors du traitement perceptif de la parole (Marslen-Wilson & Warren, 1987 ; Warren & Marslen-Wilson, 1988 ; Streeter & Nigro, 1979 ; Whalen, 1991). Si le traitement perceptif est très sensible à la présence des détails acoustiques fins, les représentations lexicales des mots stockées en mémoire sont-elles pour

autant activées par la présentation de stimuli contenant des variations de réalisation par rapport à leur forme canonique de réalisation ? En s'appuyant sur le paradigme d'amorçage sémantique intermodal, Connine et al. (1993) ont cherché à savoir si la modification des phonèmes initiaux des mots perturbe ou non l'accès au lexique. Des non-mots ont été créés en changeant la consonne initiale de mots, soit d'un ou de deux traits phonétiques (e.g., *nanner* dérivé à partir de *manner*), soit d'au moins quatre traits phonétiques (e.g., *zanner* à partir de *manner*). Les résultats ont montré que les non-mots dont le phonème initial différait d'un ou deux traits phonétiques ont facilité les temps de réponses des cibles sémantiquement associées aux mots d'origine. En revanche, il n'y avait pas d'effet d'amorçage pour la présentation des non-mots dont le phonème initial différait d'au moins quatre traits du phonème d'origine. Qui plus est, les effets d'amorçage n'étaient pas modulés par la position du phonème modifié, ni par la longueur des items. C'est-à-dire que les effets d'amorçage étaient comparables pour les non-mots bi-syllabiques et tri-syllabiques, contenant des phonèmes modifiés en position initiale et intermédiaire (ce qui remet en question le statut privilégié des phonèmes initiaux dans le traitement de la parole). Ces expériences ont eu le mérite de démontrer que la tolérance du système perceptif vis-à-vis des variations dans le signal de parole a toutefois ses limites.

Dans la même ligne de travaux, Marslen-Wilson, Moss et van Halen (1996) ont effectué des recherches ayant pour but de déterminer si la distance phonologique peut moduler l'accès au lexique en s'appuyant sur la technique d'amorçage intra-modal (auditif-auditif) avec des mots stimuli en néerlandais. Trois variables ont été manipulées : le type d'amorce, la distance phonologique et la présence ou absence d'un mot compétiteur. En ce qui concerne la première variable, le mot cible était soit précédé par un mot amorce sémantiquement relié (« *tonijn – VIS* », signifiant « thon – POISSON »), soit par un non-mot amorce partageant la rime avec le mot source en changeant la consonne initiale (e.g., *ponijn – VIS*), soit par une amorce contrôle non reliée à la cible (*wekker – VIS*). Pour ce qui est de la distance phonologique, la consonne initiale du non-mot amorce différait du mot source soit d'un trait phonétique (e.g., *ponijn* vs. *tonijn*), soit par au moins deux traits phonétiques (e.g., *nonnik* vs. *tonic*). Enfin, la distance phonologique covariait avec la variable d'environnement compétiteur ; le mot source avait un mot compétiteur (qui ne servait d'ailleurs jamais comme mot amorce) qui était soit proche soit éloigné phonologiquement du non-mot amorce (cf. *tonijn / konijn* [« lapin »] □ *ponijn* vs. triplet *tostie / hostie* [« hostie »] □ *postie*). Cette covariation permettait aux auteurs d'une part d'établir si la distance phonologique entre le mot source et ses versions dérivées

perturbe l'accès au lexique, et d'autre part de savoir si l'accès au lexique est modulé par la présence des mots compétiteurs proches dans l'espace lexical. Les résultats de cette première expérience ont permis de mettre en évidence que les consonnes initiales altérées influencent fortement l'accès au lexique. Les mots sources qui servaient d'amorce donnaient lieu aux temps de réponse les plus rapides. Si les non-mots amorces facilitaient également les temps de réponse, il y avait une différence négligeable entre les effets d'amorçage suscités par les non-mots amorces phonologiquement « proches » et non-mots « éloignés » des mots sources, ce qui signifie que tous les non-mots amorces étaient suffisamment éloignées des mots sources d'un point de vue perceptif.

Afin de déterminer plus précisément à partir de quel moment une déviation dans le signal est traitée par le système comme une discordance, une deuxième expérience a été conduite par Marslen-Wilson et al. (1996), cette fois-ci en anglais et en utilisant la technique d'amorçage intermodal. Les amorces pouvaient différer non seulement en terme d'un trait phonétique (cf. Expérience 1), mais pouvaient le faire aussi en terme du voisement en manipulant la durée de VOT présente dans la consonne initiale. Ceci permet de créer des stimuli ambigus. A titre d'exemple, dans la première condition la cible était précédée par un mot amorce associé (*plank* – *WOOD*) ; dans une deuxième condition, elle était précédée par un mot amorce ambigu avec un degré de voisement intermédiaire entre /p/ et /b/, (*p/blank* – *WOOD*) et enfin, dans la troisième condition la cible était précédée par un mot amorce compétiteur (*blank* – *WOOD*). De plus, ce premier groupe de mots stimuli était contrasté avec un second groupe en l'absence de tout mot compétiteur (*task* – *JOB* ; *t/dask* – *JOB* ; *dask* – *JOB* ; *dask* est un non-mot en anglais). Les résultats ont mis en évidence une interaction entre type d'amorce et présence (ou absence) du compétiteur. Pour les items sans compétiteur, l'amorce source (e.g., *task*) et l'amorce ambiguë (*t/dask*) ont facilité de façon similaire les temps de réponse à la cible par rapport à l'amorce contrôle et par rapport à l'amorce partageant sa rime avec le mot source (*dask*). En revanche, en ce qui concerne les items avec compétiteur, seule l'amorce source (e.g., *plank*) a donné lieu à un effet d'amorçage facilitateur alors que l'amorce ambiguë (e.g., *p/blank*) et l'amorce partageant la rime n'ont pas facilité le traitement de la cible. Ces données montrent donc que les amorces ambiguës sont traitées par le système perceptif comme les mots sources quand elles sont présentées dans un environnement lexical approprié. Andruski et collègues (1994) ont examiné le rôle des informations acoustico-phonétiques lors de l'accès au lexique, en utilisant la technique d'amorçage sémantique. Plus précisément, ces auteurs ont cherché à élucider la question de savoir si l'activation lexicale est fonction du

degré de compatibilité entre le signal acoustique et la représentation lexicale stockée dans le lexique mental. Dans ces expériences, la manipulation phonétique consistait à faire varier la durée de VOT dans les plosives en début de mots. Ainsi, trois catégories de plosives ont été construites : la forme canonique, la forme altérée dans laquelle un tiers de VOT était éliminé (désormais stimuli à VOT_{-1/3}) et enfin la forme altérée dans laquelle deux tiers de VOT étaient éliminés (stimuli à VOT_{-2/3}). Les résultats préliminaires de la tâche de catégorisation phonétique ont permis de démontrer une sensibilité perceptive aux manipulations effectuées de la part des auditeurs : ils catégorisaient les consonnes initiales canoniques et les consonnes à VOT_{-1/3} de façon similaire alors que les stimuli à VOT_{-2/3} donnaient lieu à des temps de catégorisation plus longs. Ensuite, les stimuli étaient présentés dans une expérience de décision lexicale avec la technique d'amorçage sémantique selon quatre conditions : (1) canonique (*peace* – WAR), (2) amorce à VOT_{-1/3} (*peace*_{1/3} – WAR), et (3) amorce à VOT_{-2/3} (*peace*_{2/3} – WAR), et enfin (4) une condition contrôle (*hip* – WAR). De plus, l'ambiguïté lexicale potentielle était également manipulée : pour la moitié des mots stimuli, la manipulation phonétique concernant le VOT donnait lieu à un autre mot anglais (e.g., *pet* vs. *bet*), alors que pour l'autre moitié la manipulation donnait lieu à un non-mot (e.g., *peace* vs. *beace*). Ces auteurs ont également fait varier l'intervalle entre l'amorce et la cible (*Inter Stimulus Interval*, désormais ISI). A un ISI de 250 ms, il y avait un effet d'amorçage sémantique classique dans la mesure où les paires amorce-cible reliées donnaient des temps de décision plus rapides que les paires non reliées. Qui plus est, la présence d'un compétiteur lexical par voisement (cf. *pet* vs. *bet*) était suffisante pour ralentir les temps de réponse de façon générale (voir aussi Milberg, Blumstein, & Dworetzky, 1988). Les temps de réponse n'étaient cependant pas affectés par les manipulations liées au VOT. En revanche, à un ISI de 50 ms les résultats de la première expérience étaient non seulement répliqués mais les temps de décision lexicale étaient aussi plus longs pour les cibles précédées par des amorces à VOT_{-2/3} que par des amorces canoniques ou à VOT_{-1/3}²¹. Ces résultats suggèrent que les modifications acoustiques de bas niveau peuvent perturber l'accès au lexique quand l'intervalle entre l'amorce et la cible est très court.

Récemment, Misiurski, Blumstein, Rissman et Berman (2004) ont abordé l'étude du rôle des informations acoustico-phonétiques dans la compétition lexicale sous un angle légèrement différent. Si la présence d'un compétiteur peut provoquer un ralentissement du traitement

²¹ Van Alphen & McQueen, soumis, qui n'ont pas trouvé d'effets d'amorçage gradués modulés par les durées de VOT lors de l'accès au lexique en néerlandais.

lexical, il est alors pertinent de savoir si l'activation partielle d'une forme lexicale induit également l'activation du réseau lexico-sémantique de son compétiteur lexical. Dans une tâche de décision lexicale ils ont utilisé la technique d'amorçage sémantique intermodal, avec trois conditions expérimentales. Dans la condition « canonique », l'amorce était sémantiquement reliée à la cible (*dime* – PENNY). Dans la deuxième condition, l'amorce est phonétiquement manipulée en réduisant la durée du VOT de deux tiers. Cette manipulation donnait ainsi lieu au compétiteur lexical du mot canonique présentée dans la condition canonique (e.g., *t*ime* (dérivé de *dime*) – PENNY). Enfin, la troisième condition contrôle (e.g., *nose* – PENNY) servait de ligne de base à partir de laquelle l'effet d'amorçage était évalué. Les résultats ont mis en évidence un effet d'amorçage pour les paires amorces – cibles canoniques (*dime* – PENNY), ainsi que pour les paires amorces – cibles altérées (*t*ime* – PENNY). Toutefois, cet effet était plus faible pour les paires amorces-cibles altérées, et ceci pour les deux intervalles inter-stimuli (ISI à 50 et 250 ms). Le fait que ce profil de résultats ne soit pas observé pour les amorces altérées suivies par des non-mots ont conduit Misiurski et al. (2004) à conclure que la réduction des effets d'amorçage provoquée par les amorces modifiées n'est pas due à la manipulation phonétique *per se*, mais plutôt à l'activation partielle du réseau lexico-sémantique du mot compétiteur²².

Suite aux travaux d'Andruski et al. (1994), Van Alphen (2004) a utilisé les deux techniques d'amorçage intermodal associatif et de répétition pour étudier le rôle de l'indice de prévoisement sur la reconnaissance des mots en néerlandais. Cet auteur a fait varier les périodes de VOT, indice acoustique qui permet la distinction perceptive entre consonnes voisées et non voisées en position initiale. Il y avait 0, 6 ou 12 périodes de prévoisement, correspondant à différentes durées selon la place d'articulation coronale ou labiale de la consonne concernée. Les résultats de ces expériences ont montré qu'il y avait un effet d'amorçage sémantique, i.e. les temps de réponse étaient plus courts pour des paires amorce-cible sémantiquement associées telles que *bloem* (fleur) – ROOS (rose) par rapport aux paires non reliées (cf. *bloem* – BAAN). Plus important, cet effet d'amorçage n'était pas modulé par les périodes de prévoisement, alors que les résultats d'un post-test de discrimination ont

²² Notons cependant qu'afin de pouvoir valider une telle hypothèse, une condition supplémentaire aurait été nécessaire qui consiste à utiliser le mot compétiteur par manipulation phonétique comme mot amorce sémantiquement relié à la cible, cf. la paire *t*ime* – WATCH. Dans leur ensemble, les résultats sont donc en accord avec l'idée selon laquelle l'activation lexicale est graduée plutôt que dichotomique et qu'elle est parfois modulée par des facteurs acoustico-phonétiques de très bas niveau (cf. McQueen, Dahan, & Cutler 2003).

permis de démontrer que les participants pouvaient discriminer les différentes périodes de prévoisement. Il est important de noter que les résultats étaient similaires quelle que soit la technique d'amorçage utilisée (amorçage sémantique, cf. *bloem* – ROOS vs. amorçage de répétition, cf. *bloem* – BLOEM). Même en absence de tout prévoisement dans l'amorce (cf. *bloem* prononcé [pɫum]), un effet d'amorçage facilitateur fût observé, (i.e., [pɫum] a amorçé ROOS tout autant que *bloem* [bɫum]). C'est-à-dire qu'en absence de tout compétiteur lexical, le système perceptif tolère des déviations dans le signal de parole ce qui provoque les effets d'amorçage associatif identiques à ceux provoqué par les mots canoniques.

Van Alphen & McQueen (soumis) ont cherché à examiner l'influence de l'environnement compétiteur lexical sur la variation du prévoisement. De nouveau, trois périodes de VOT étaient utilisées. Dans une tâche de décision lexicale avec la technique d'amorçage de répétition, les participants étaient soumis à des présentations de mots et non-mots avec des consonnes plosives voisées. Certains de ces mots et non-mots avaient un compétiteur lexical généré par le voisement (e.g., *beer* (bière) - *peer* (poire) vs. *brins* – *prins* (prince)), alors que d'autres mots n'avaient pas de compétiteur (e.g., *blauw* (bleu) *plauw* vs. *blem* - *plem*). Les mots cibles étaient précédés par des mots amorces avec trois (0, 6, ou 12) périodes de prévoisement dans la consonne initiale. En plus de ces trois modalités de prévoisement, il y avait des amorces avec pour consonne initiale une plosive non voisée, et une condition d'amorçage non relié. Les prédictions étaient les suivantes : les effets d'amorçage devraient être les mêmes pour les amorces avec 6 ou 12 périodes de prévoisement, étant donné que la *durée* exacte de prévoisement n'est pas essentielle à la perception du voisement des plosives initiales. En revanche, une différence d'effets d'amorçage devrait être observée entre les amorces ayant zéro ou six périodes de prévoisement puisque la suppression du prévoisement enlèverait l'indice principal de voisement. Les résultats ont mis en évidence des effets classiques d'amorçage sémantique dans le sens où les amorces identiques aux cibles ont facilité leur traitement. Cependant, il n'y avait pas de différence significative entre les mots amorces avec zéro prévoisement et 6 périodes de prévoisement (pourtant, les résultats d'un post-test de discrimination phonémique avait indiqué que les auditeurs sont capables de distinguer ces deux types de prévoisement). En revanche, quand le mot amorce avait un compétiteur lexical, il y avait un effet de l'absence du prévoisement. A titre illustratif, les temps de réponse étaient plus courts pour les paires amorce-cible *beer* (sans prévoisement) – PEER par rapport aux paires *beer* (6 périodes de prévoisement) – PEER. De façon similaire, quand les cibles non-mots tels que *brins* étaient précédés par des amorces voisées sans

prévoisement (i.e., *brins* sans prévoisement), les réponses (« non » cette fois-ci) étaient plus lentes par rapport aux mêmes mots cibles précédées par des amorces voisées avec prévoisement. Une tâche d'identification suivant la décision lexicale avait permis de montrer que les plosives voisées sans prévoisement d'un point de vue acoustique, sont toutefois identifiées comme étant des plosives voisées. Donc les candidats voisés (*beer*) reçoivent plus d'activation d'emblée, même si tout prévoisement a été enlevé, et de fait sont traités comme si leurs compétiteurs non voisés (*peer*) n'étaient jamais activés dès le départ. Dans leur ensemble, les résultats obtenus par Van Alphen et McQueen (soumis) suggèrent que la durée *per se* (6 vs. 12 périodes) de prévoisement n'affecte pas l'accès au lexique, contrairement à ce que d'autres avaient pu observer en anglais (cf. Andruski et al., 1994 ; Utman et al., 2000 ; Misiurski et al., 2004). Il semblerait donc que la *valeur informative* de cet indice acoustique varie selon les langues : si l'indice acoustique a surtout une valeur *quantitative* en anglais dans le sens où les auditeurs anglais sont sensibles à la durée du VOT, il a une valeur *qualitative* en néerlandais dans le sens où les auditeurs sont sensibles à la présence du VOT dans le signal et non pas à sa durée. Si les détails acoustico-phonétiques peuvent affecter l'accès au lexique, de tels effets dépendent également du fait que ces informations sont pertinentes à la distinction lexicale, c'est-à-dire si elles ont une valeur informative. Des comparaisons inter-langues ont permis de démontrer que la variation d'un indice acoustique comme le VOT influence l'accès au lexique en fonction de la valeur informationnelle de l'indice dans une langue donnée.

En résumé, les études examinées plus haut ont mis en avant deux points importants. Premièrement, le système perceptif s'avère très sensible aux variations acoustico-phonétiques fines présentes dans le signal de parole. Par exemple, l'ampleur des effets d'amorçage semble varier en fonction de la distance phonologique entre le mot source et le non-mot dérivé. Deuxièmement, la présence des compétiteurs lexicaux est un facteur qui est pris en compte par le système perceptif. Les candidats lexicaux qui sont activés suite à la présentation acoustique d'un mot stimulus entrent en compétition, ce qui ralentit le temps de traitement de ce mot. Dans ce chapitre, nous avons eu pour objectif d'apporter de nouveaux éléments de réponses en étendant ce champ d'étude à la parole assimilée en français.

1.2. Assimilation et ambiguïté lexicale

Comme cela a déjà été mentionné dans le chapitre 2, l'assimilation peut parfois donner lieu à une ambiguïté lexicale potentielle, ce qui est le cas pour le mot anglais *run* qui peut donner le mot *rum* après avoir subi une assimilation de place labiale dans la séquence *a quick run picks*

you up. Plusieurs études ont examiné le rôle de ce type d'ambiguïté lexicale en utilisant la technique d'amorçage intermodal. Par exemple, Gow (2002) a présenté des stimuli assimilés, issus de parole naturelle, et il a montré que les auditeurs ont pu accéder à la forme sous-jacente des mots tels que *right*, même si ces mots étaient insérés dans des contextes phrastiques assimilatoires (e.g., *right berries*) où la coronale /t/ en position finale subit une assimilation de place labiale. Qui plus est, dans un pré-test les mots stimuli étaient sélectionnés afin de représenter des mots bien assimilés (e.g., *ripe* [raip]). En fait, d'après Gow, un segment est rarement assimilé à 100%. La présence des indices acoustiques fins dans le segment permettrait généralement à l'auditeur d'avoir accès à sa forme sous-jacente.

Gaskell et Marslen-Wilson (2001) ont également étudié le traitement perceptif des items ambigus par assimilation (*run/rum*) et ils ont trouvé que dans des contextes sémantiquement neutres, seule la représentation lexicale de la forme de surface (i.e., *rum* dérivé de *run*) était activée. Par contre, un contexte phrastique contenant un biais sémantique favorisant une interprétation de la forme d'origine (*run*), combinée avec le contexte phonologique approprié a permis l'activation de la forme sous-jacente, ce qui est conforme avec les résultats d'autres études sur l'ambiguïté lexicale (voir e.g., Tabossi, 1988 ; Duffy, Morris & Rayner, 1988).

1.3. La présente recherche

Dans les expériences rapportées dans ce chapitre, nous avons utilisé la technique d'amorçage associatif (Meyer & Schvaneveldt, 1971 ; Neely, 1991 ; Ferrand & New, 2004). C'est-à-dire que les paires amorces-cibles ont été sélectionnées à l'aide de la technique d'association libre qui permet de recueillir sur une population homogène des normes de réponses associatives. Pour chaque mot-stimulus présenté, un ensemble de mots-réponses distincts est évoqué dont chacun est donné par un certain pourcentage de la population. Ensuite, ces paires ont été présentées dans une tâche de décision lexicale avec la technique d'amorçage.

Nos expériences ont pour objectif d'examiner dans quelle mesure la parole assimilée affecte le traitement des mots ayant des compétiteurs lexicaux. Plus précisément, il s'agit d'établir dans quelle mesure l'activation d'un mot « compétiteur », généré par assimilation (par exemple *bague* [bag] dérivé de *bac* [bak] assimilé)²³ va donner lieu à l'activation du réseau lexico-sémantique du mot d'origine *bac*. Autrement dit, la présentation du mot amorce

²³ Afin de faciliter la lecture, la forme sous-jacente d'un mot est dorénavant indiquée en italiques (*bac*) alors que son mode de réalisation est indiqué entre crochets [bak].

assimilé [bag] amorce-t-elle la cible EXAMEN comme peut le faire la forme canonique *bac*, qui est sémantiquement reliée à cette cible ? Afin de répondre à cette question, des mots français étaient sélectionnés qui donnent lieu à d'autres mots après assimilation (MAA, cf. la paire *bac / bague*) et des mots non ambigus (MNA, cf. la paire *jupe / jube*, voir aussi chapitre 2). Ces mots stimuli étaient insérés dans des contextes phrastiques donnant lieu ou non à l'assimilation. Les mots stimuli ont ensuite été extraits de leur contexte phrastique et présentés auditivement à nos sujets. Dans le paradigme d'amorçage associatif intermodal, les amorces auditives sont soit associées sémantiquement aux cibles visuelles, soit elles ne le sont pas.

Dans un premier temps, nous avons cherché à mettre en place une expérience dans laquelle les participants devaient recevoir les trois conditions expérimentales : « canonique » (e.g., *soute* [sɔt] – BAGAGE), « assimilée » (e.g., *soute* [sɔd] – BAGAGE) et enfin une condition « contrôle » (e.g., *rite* [rit] – BAGAGE). Toutefois, étant donné le nombre très restreint des items expérimentaux disponibles (12 MAA et en conséquence 12 MNA), nous avons choisi d'associer les conditions de réalisation canonique ou assimilée du mot amorce à deux sous-groupes de sujets. Le facteur « mode de réalisation » est donc un facteur inter groupes. Un groupe de sujets recevait les amorces réalisées sous la forme canonique (e.g., *soute* prononcé [sɔt]), alors que l'autre groupe recevait les amorces réalisées sous la forme assimilée (e.g., *soute* prononcé [sɔd]). Au sein de chaque groupe, deux sous-groupes de sujets ont été utilisés afin de contrebalancer la présentation des paires associées (e.g., *soute* – BAGAGE) et non reliées (e.g., *rite* – BAGAGE). Le plan expérimental comportait ainsi un facteur principal inter groupes : « Mode de réalisation » M₂ à deux modalités (canonique vs. assimilé), et deux facteurs principaux intra groupes, à savoir « Type de mots » T₂ à deux modalités (MAA vs. MNA) et « Relation sémantique » R₂ à deux modalités (paires amorces-cibles sémantiquement reliées vs. paires non reliées).

Nos prédictions étaient les suivantes. Premièrement, nous nous attendons à un effet de relation d'amorçage classique dans le sens où les cibles précédées par des amorces associés devraient donner lieu à des temps de réaction plus courts que les paires amorces-cibles non associées. Deuxièmement, l'effet d'amorçage devrait être modulé par la présence d'un mot compétiteur par assimilation (Andruski et al., 1994 ; Misiurski et al., 2004 ; Marslen-Wilson et al., 1996). Dans le cas des MNA, la présentation de l'item non-mot assimilé (e.g., [ʒyb] dérivé de *jupe*), doit activer les mots associés du mot d'origine *jupe* tels que *robe* ou *veste*, étant donné que

c'est la forme phonologique la plus proche. En revanche, dans le cas des MAA, la présentation de la forme assimilée (e.g., [sʊd] dérivé de *soute*), étant elle-même un mot français, pourrait activer davantage ses propres associés (comme par exemple *chlore* étant sémantiquement associé à *soude*), plutôt que ceux correspondant au mot d'origine (cf. *bagage* étant associé à *soute*). Il est présupposé que les formes assimilées des MAA vont donner lieu à un effet d'amorçage plus faible que les MNA, car la forme de surface des premiers constitue un mot et sa présentation devrait activer son propre réseau lexico-sémantique. Donc, l'effet de relation sémantique pour la forme sous-jacente devrait être plus important pour les MNA que pour les MAA en raison du fait que ces derniers provoqueraient une interférence entre les deux réseaux lexico-sémantiques. Autrement dit, nous devrions observer une interaction entre les deux facteurs Type de mots et Relation sémantique dans la condition « assimilation » alors qu'il ne devrait pas y avoir une telle interaction pour la condition « canonique ».

2. EXPERIENCE 7 □ AMORÇAGE DES FORMES CANONIQUES ET ASSIMILEES

2.1. Matériel

Vingt-quatre mots « source » ont été sélectionnés : 12 MAA (e.g., *soute* qui peut devenir *soude* [sʊd] après avoir subi une assimilation) et 12 MNA (e.g., *jupe* qui peut devenir /jube/ [ʒyp] qui ne correspond pas à un mot français). Les propriétés lexicales des mots stimuli utilisés dans les expériences de ce chapitre, sont récapitulées dans le Tableau 4.1.

Type	Fréq (log)	VPhon	Vphon –CF	NVD	JS Associés	Fréq (log)
MAA (<i>soute</i>)	6.7 (0.59)	24.18	9.67	4.63	41.3%	34.95 (1.31)
MNA (<i>jupe</i>)	31.2 (1.42)	16.75	6.17	0.92	36.6%	90.38 (1.7)

Tableau 4.1. Propriétés lexicales des mots stimuli (MAA vs. MNA), et leurs associés utilisés dans les Expérience 7. Sont indiquées les valeurs suivantes : la fréquence moyenne d'usage (cf. *Lexique*, valeurs logarithmiques entre parenthèses) ; le nombre total des voisins phonologiques (Vphon) ; le nombre de voisins phonologiques après substitution de la consonne finale (Vphon-CV), le nombre de voisins plus fréquents après substitution de la consonne finale (NVD, Nombre de Voisins Dangereux) ; les moyennes des jugements subjectifs des mots associés (JS associés, voir texte) et leurs fréquences moyennes (valeurs logarithmiques).

Tous les 24 mots « source » étaient monosyllabiques et avaient toujours pour consonne finale une obstruente (vélaire, dentale ou labiale) non voisée. Rappelons que les consonnes non voisées sont plus fréquemment et complètement assimilées que les consonnes voisées (cf.

chapitre 2, section 5.4 ; voir aussi chapitre 3). Tous les mots étaient ensuite insérés dans des syntagmes nominaux de type Det-Nom-Adj. Les contextes donnaient lieu à des assimilations (e.g., *une soute [sud] bondée), ou non (e.g., *une soute [sut] pleine, voir l'annexe 10 pour toutes les phrases stimuli). La consonne donnant lieu à une assimilation ne correspondait jamais à la consonne assimilée pour éviter la création des consonnes géminées. En outre, 48 mots amorce de remplissage étaient sélectionnés dont la moitié avaient des associés sémantiques (e.g., crabe – MER, ces associations ont été sélectionnées à partir des normes d'associations en français, Ferrand & Alario 1998, avec une moyenne de 0.27 de jugements subjectifs), alors que l'autre moitié ne l'était pas (larve – LOISIR). Enfin, pour les besoins de la tâche de décision lexicale, 72 cibles non-mots ont été sélectionnées à l'aide du logiciel WordGen (Duyck, Desmet, Verbeke & Brysbaert, sous presse). Soixante-douze mots monosyllabiques servant comme amorce dans la tâche expérimentale ont été choisis, dont les fréquences étaient appariées avec les mots stimuli. Pour un tiers de ces paires « mots amorces – non-mots cibles », le non-mot cible était pseudo-relié au mot amorce, c'est-à-dire que le non-mot était phonologiquement proche, par substitution d'une consonne à une position quelconque, du mot « source » lui-même relié à l'amorce (e.g., ciel – ESLACE, où le non-mot ESLACE est phonologiquement proche du mot source ESPACE, sémantiquement relié à l'amorce *ciel*). Nous avons voulu éviter ainsi des stratégies mises en place par les participants comme l'association systématique de la relation sémantique aux mots cibles. Les mots de remplissage servant d'amorces ont ensuite été insérés dans des contextes neutres mais sémantiquement plausibles.**

Tous les stimuli étaient prononcés par une locutrice originaire de la région parisienne dans une chambre insonorisée. Nous lui avons demandé de prononcer les syntagmes de façon continue et fluide, sans laisser de silence entre les mots au sein de chaque syntagme (pouvant toutefois laisser des pauses entre les séquences). La locutrice était informée de l'objectif de l'expérience et nous lui avons demandé de produire délibérément des assimilations là où il y avait lieu d'en avoir. La production était directement enregistrée et digitalisée (20 kHz) sur ordinateur à l'aide du logiciel PRAAT (Boersma & Weenink, 1999). La fin du mot amorce était identifiée par l'analyse des spectrogrammes avec le même logiciel. (Nous avons demandé à une personne ayant pour langue maternelle le français et experte en perception auditive, d'écouter tous les stimuli pour assurer que les mots étaient bien assimilés ou pas.)

Pour résumer, le matériel expérimental était donc constitué de 72 paires *mots amorces - mots cibles* (12 amorces de type MAA + cibles ; 12 amorces de type MNA + cibles ; 48 paires

amorce-cible de remplissage dont la moitié était sémantiquement reliée, alors que l'autre moitié était non reliée) et de 72 paires *mots* amorces- *non-mots* cibles (dont un tiers de paires était pseudo-reliées, e.g., ciel – ESLACE).

2.1.1. *Pré-test : Technique d'association libre*

Afin de sélectionner des cibles sémantiquement associées aux 24 mots amorces test, nous avons utilisé la technique d'association libre simple (Ferrand & Alario, 1998). Cette technique est le plus souvent employée pour recueillir sur une population nombreuse des normes de réponses associatives, c'est-à-dire pour chaque mot-stimulus présenté, un ensemble de mots-réponses distincts évoqué dont chacun est donné par un certain pourcentage de la population.

Cinquante-huit participants, tou(te)s étudiant(e)s en troisième année de Psychologie à l'Université René Descartes (Paris 5), ont pris part au pré-test. La totalité de 48 associants (24 mots test et 24 mots de remplissage) était présentée sur une feuille en quatre colonnes. Suivant Ferrand et Alario (1998), la consigne invitait les participants : 1 / à écrire le premier mot qui leur venait à l'esprit en face de l'associant, 2 / à ne donner qu'une seule réponse, 3 / à suivre l'ordre des associants, 4 / à ne sauter aucun mot, 5 / à aller le plus vite possible, 6 / à écrire lisiblement, 7 / à rester concentré et silencieux. Dans cette technique, la liberté de choix du participant est illimitée, puisqu'il peut répondre par n'importe quel mot du vocabulaire de sa langue. De plus, l'association est dite simple dans le sens où le sujet ne doit donner qu'une seule réponse. La présentation des mots était visuelle ; la réponse écrite ; la passation de l'épreuve collective ; le temps de réponse libre. Au total, la passation du pré-test durait environ 10 minutes.

Généralement, nous avons observé que les réponses à un mot stimulus présente une gamme très étendue de « fréquence culturelle », allant des réponses banales données par un grand nombre de participants aux réponses totalement originales (données par un seul participant). Nous constatons que la plupart des mots donnait lieu à une unanimité dans les choix des mots associés. En moyenne, les MAA donnaient lieu à une réponse lexicale unanime dans 41.3% des cas alors que le pourcentage moyen concernant les MNA était de 36.6%, (voir annexe 10 pour les résultats de tous les mots expérimentaux).

2.2. Sujets

Trente-huit sujets (34 femmes et deux hommes, âge moyen = 24 ans), pour la plupart étudiant(e)s de troisième année à l'Institut de Psychologie de l'Université René Descartes, ont pris part à l'expérience correspondant à la condition de mode de réalisation « canonique ».

Trente-six étudiant(e)s en deuxième année en Psychologie (six hommes et 30 femmes, âge moyen = 21 ans), ont pris part à l'expérience correspondant à la condition de Mode de réalisation « assimilée ». Ils / elles avaient tou(te)s une audition correcte et une vision correcte ou corrigée.

2.3. Procédure

Les participants étaient testés individuellement dans une salle sombre et silencieuse. Avant le déroulement de l'expérience proprement dite, les participants devaient remplir un questionnaire qui portait sur leurs connaissances linguistiques. L'expérience consistait en trois blocs expérimentaux : un premier bloc d'entraînement pendant lequel 10 paires amorce-cible étaient présentées. Après le bloc d'entraînement, le participant pouvait encore poser des questions. L'expérience principale était composée de deux blocs de 74 items chacun, après un bloc expérimental le déroulement s'arrêtait automatiquement afin que les participants puissent prendre une pause de quelques minutes. L'expérience était conçue de telle sorte qu'avant et après l'arrêt au milieu il y avait deux mots de remplissage.

Les amorces auditives étaient présentées par écouteurs (Sennheiser). A la fin du mot amorce, le mot cible apparaissait sur l'écran d'ordinateur pendant 3000 ms ou jusqu'à ce que le sujet donne sa réponse. Chaque sujet avait pour tâche de dire si la suite de caractères en lettres majuscules figurant sur l'écran correspondait à des mots français ou non. Les participants ont répondu avec leur main dominante. Le mot cible était en Arial 14 points, de couleur noire sur un fond d'écran blanc afin de rapprocher la situation expérimentale à une situation de lecture réelle. L'expérience a été effectuée avec le logiciel DMDX (Forster & Forster, 2003) sur un ordinateur PC. Les temps de réaction étaient calculés à partir du moment de l'apparition de la cible visuelle. Les réponses ont été enregistrées sur une manette de jeu Logitech Wingman. La durée moyenne de l'expérience était de 15 à 20 minutes.

2.4. Résultats

Les données de deux participants de la condition « canonique » ont été écartées des analyses statistiques ultérieures. Des erreurs étaient survenues dans le fichier des données d'un participant, et un autre participant avait pour langue maternelle l'allemand (nombre de participants par liste expérimentale = 18). En revanche les données de tous les participants de la sous-expérience « assimilation » ont été retenues (nombre par liste expérimentale = 18).

En outre, les réponses erronées (3.19 % et 4.25 % pour les conditions « canonique » et « assimilation » respectivement) ont été exclues des analyses, ainsi que les temps de réponse

en dessous de 300 ms et au-dessus de 1200 ms (1.72 % et 1.27 % pour les conditions « canonique » et « assimilation » respectivement). Les temps de réponse moyens (ms) et les taux d'erreurs (%) sont présentés en fonction des facteurs et leurs modalités dans le Tableau 4.2. Etant donné les taux d'erreurs très faibles, nous n'avons effectué une analyse de variance que sur les temps de réponse avec pour facteurs principaux intra sujets Type de mot T₂ à deux modalités (i.e., MAA vs. MNA) et Relation sémantique R₂ entre l'amorce et la cible à deux modalités (sémantiquement relié vs. non relié), et pour facteur principal inter sujets Mode de réalisation M₂ à deux modalités (« canonique » vs. « assimilation »). (Le facteur secondaire de contrebalancement Liste expérimentale dans les analyses par item ainsi que l'assignation à une liste (dans l'analyse par sujet) n'ont pas eu d'effet significatif ($F_s < 1$) et n'ont pas interagi avec les autres facteurs).

	<i>Canonique</i>			<i>Assimilation</i>		
	<i>Relié</i>	<i>Non relié</i>	<i>Amorçage</i>	<i>Relié</i>	<i>Non relié</i>	<i>Amorçage</i>
	soute-BAGAGE	rite-BAGAGE		([sud]-BAGAGE)	([rid]-BAGAGE)	
<i>MAA</i>	500 (61)	511 (83)	+11	492 (61)	515 (87)	+23
<i>Err.(%)</i>	0.35	0.83		0.95	0.71	
<i>MNA</i>	<i>Relié</i>	<i>Non relié</i>		<i>Relié</i>	<i>Non relié</i>	
	jupe-ROBE	troupe-ROBE		[zyb]-ROBE	[trub]-ROBE	
	505 (71)	521 (86)	+16	493 (62)	511 (76)	+18
<i>Err.(%)</i>	0.83	1.18		1.06	1.54	

Tableau 4.2. Expérience 7 : temps de réponse moyens (ms, écart types entre parenthèses) pour les cibles en fonction de l'ambiguïté lexicale (MAA vs. MNA) et relation sémantique (paire amorce-cible reliée vs. non reliée).

2.4.1. Analyse sur les temps de réponse²⁴

Comme on peut le voir dans le Tableau 4.2, les temps de réponse au mot cible sont plus courts lorsque ce mot est précédé par des amorces associées que lorsqu'il est précédé par des amorces non reliées. Les différences numériques entre paires reliées et non reliées (condition canonique : 503 ms vs. 516 ms ; condition assimilation : 493 ms vs. 513 ms) montrent qu'il

²⁴ Etant donné le taux d'erreurs très bas, nous ne rapportons pas ici les analyses de variance sur cette variable dépendante.

s'agit d'un effet assez faible, ce qui est généralement le cas dans la littérature (cf. Marslen-Wilson & Zhou 1993 ; Moss, Tyler, Ostrin & Marslen-Wilson, 1995). L'effet de Relation sémantique est observé, quel que soit le mode de réalisation. Cet effet R_2 est significatif par sujets $F_1(1, 70) = 8.9, p < .01$; et par items $F_2(1,22) = 5.48, p < .05^{25}$. En revanche, ni le facteur Type de mots T_2 (MAA vs MNA), ni le facteur Mode de réalisation M_2 (canonique vs. assimilation) ne se sont avérés significatif, ni par sujets, ni par items. Aucune interaction significative n'était observée entre les trois facteurs principaux .

2.5. Discussion

Les résultats de l'expérience d'amorçage sémantique ont révélé un effet classique de relation sémantique : les temps de réponse aux mots cibles étaient significativement plus courts lorsqu'ils étaient précédés par des mots amorces sémantiquement associés (bac – EXAMEN) que lorsqu'ils étaient précédés par des mots amorces non reliés (jupe – EXAMEN). Contrairement à notre hypothèse de travail, cet effet d'amorçage sémantique, n'était pas modulé par la présence d'un mot compétiteur par assimilation. C'est-à-dire que l'effet d'amorçage était analogue pour les mots ambigus par assimilation (cf. *bac* avec pour compétiteur lexical *bague* par assimilation) et pour les mots non ambigus, i.e. des mots n'ayant pas de tel compétiteur généré par assimilation (e.g., *jupe*).

Nous nous attendions à une différence d'effet d'amorçage entre les mots ayant des compétiteurs lexicaux générés par assimilation et ceux qui n'en ont pas. Il était supposé que les mots dérivés après assimilation (*soude* dérivé de *soute*) allaient activer le réseau lexico-sémantique correspondant à la forme assimilée alors qu'une telle activation ne pourrait avoir lieu pour les formes assimilées des mots non ambigus, car cette forme (e.g., *jube*) ne correspond pas à un mot de la langue. Des études ont montré que la présence des compétiteurs lexicaux provoque une activation partielle de leur réseau lexico-sémantique, ce qui diminue, d'après certains auteurs, l'ampleur des effets d'amorçage pour les mots canoniques (cf. Andruski et al., 1994 ; Misiurski et al., 2004). Or, nos résultats n'ont pas révélé une telle interaction, ce qui suggère que la forme assimilée *soude* [sud] dérivée du mot d'origine *soute*, active toutefois le réseau lexico-sémantique propre à ce mot source. Il semblerait donc

²⁵ Nous avons effectué des t-tests au sein de chaque condition (« canonique » et « assimilation ») entre les modalités du facteur Relation R_2 (paires reliées vs. non reliées). Pour ce qui est de la condition « canonique », l'effet de Relation n'était pas significatif ($p = 0.29$) alors que pour la condition « assimilation », l'effet ne faisait qu'effleurer le seuil de significativité ($p = 0.09$).

que la désignation « ambigu par assimilation » ne soit pas très appropriée dans la mesure où ces mots-là ne sont pas traités différemment des mots non ambigus en ce qui concerne l'activation du réseau lexico-sémantique de leur forme sous-jacente. On observe un effet d'amorçage facilitateur, quel que soit le type d'amorce (MAA ou MNA) et quelle que soit le type de mode de réalisation (canonique ou assimilée). Cependant, il est à noter que l'absence de différence entre les deux conditions *canonique* et *assimilation* devrait être interprétée avec prudence dans la mesure où deux groupes de sujets différents ont participé à ces deux conditions. Si ces résultats sont en désaccord avec certaines études rapportées dans la littérature, ils s'accordent toutefois avec l'hypothèse que nous avons d'ores et déjà proposée dans les chapitres précédents. En dépit de toute variation dans la forme de surface, il semblerait que les segments assimilés en français contiennent suffisamment de traces acoustiques permettant à l'auditeur d'avoir accès à sa forme sous-jacente.

Afin de valider l'interprétation d'après laquelle l'activation du mot source est due à la présence des indices acoustiques dans la forme assimilée, une expérience contrôle s'imposait. Dans l'expérience contrôle, les formes de surfaces dérivées des mots assimilés sont produites et présentées sous leur réalisation canonique. A titre illustratif, le mot dérivé par assimilation du mot canonique *soute* est maintenant réalisé sous la forme canonique de *soude* [sʊd], c'est-à-dire sans aucune trace acoustique d'assimilation de (dé-)voisement. Ces mots amorces sont suivis par les mêmes cibles que celles de l'Expérience 7, (cf. *soude* est suivi par la cible BAGAGE, qui est sémantiquement reliée à *soute*). Si la forme assimilée de *soute* prononcée [sʊd] et la forme canonique *soude* sont traitées de façon similaire par les auditeurs du seul fait qu'ils sont proches phonologiquement, alors le mot *soude* réalisé de façon canonique devrait amorcer BAGAGE tout autant que la forme assimilée [sʊd]. Si, au contraire, nous n'observons pas d'effet d'amorçage pour la paire *soude* [sʊd] – BAGAGE, alors on pourra en conclure que la forme assimilée de *soute* [sʊd] et la forme canonique (*soude* [sʊd]) sont traitées différemment d'un point de vue perceptif. L'effet d'amorçage sémantique obtenu pour la paire *soute* [sʊd] – BAGAGE (cf. Expérience 7) ne pourrait être dû qu'à la présence des traces acoustiques, exploitées par l'auditeur lors de l'accès au lexique. Bien entendu, ce point ne considère que les amorces MAA, car la réalisation canonique correspondant à la forme de surface des mots MNA (e.g., *jube* [ʒyb]) ne correspond pas à un mot de la langue française.

3. EXPERIENCE 8 □ AMORÇAGE DES FORMES CANONIQUES DERIVEES

3.1. Matériel

Douze équivalents canoniques des formes assimilées des MAA utilisées dans l'expérience précédente étaient insérés dans des contextes neutres et sémantiquement plausibles. A titre illustratif, l'équivalent canonique de la forme assimilée du mot *bac*, *bague* [bag] était inséré dans le contexte *une bague métallique* où l'assimilation n'avait pas lieu d'être (cf. Annexe 10). La plausibilité sémantique de ces phrases était vérifiée par trois personnes ayant pour langue maternelle le français. Ces douze phrases étaient produites par la même locutrice dans l'Expérience 7, avec le même dispositif d'enregistrement (voir la description du matériel expérimental, section 2.1). A part ces modifications, le reste du matériel était identique à celui que nous avons utilisé dans l'Expérience 7. La liste expérimentale comportait également des 48 mots de remplissage dont la moitié était sémantiquement associée, alors que l'autre ne l'était pas. Ces mots servaient de ligne de base afin de confirmer la présence d'un effet de relation sémantique.

3.2. Procédure

Nous avons utilisé le même paradigme d'amorçage sémantique que pour l'Expérience 7 (voir section 2.3).

3.3. Sujets

Trente-cinq sujets (5 hommes et 30 femmes, âge moyen = 22 ans), étudiant(es) en deuxième année à l'Institut de Psychologie, ont pris part à l'expérience. Tou(te)s les participant(e)s étaient de langue maternelle française et aucun d'entre eux / elles n'ont rapporté des problèmes auditifs ou visuels.

3.4. Résultats & Discussion

Les données d'un seul participant, qui avait pour langue maternelle le portugais, ont été écartées des analyses statistiques. Après cette exclusion, le nombre de participants par Liste (cf. contrebalancement) était de 17. En outre, les réponses erronées (2.33 %) ainsi que les temps de réponse en dessous de 300 ms et au-dessus de 1200 ms (0.86%) ont été exclus des analyses. Les facteurs Liste (dans les analyses par item) ainsi que l'assignation à une liste (dans l'analyse par sujet) n'ont pas eu d'effet significatif ($F_s < 1$) et n'ont pas interagi avec les autres facteurs.

En ce qui concerne les items expérimentaux, les résultats ont montré que les amorces pseudo-reliées, i.e. les équivalents canoniques des formes assimilées (*soude*), n'ont pas facilité le traitement des cibles (BAGAGE) par rapport aux amorces non reliées (*ride* – BAGAGE), avec des temps de réponses moyens de 530 ms et 525 ms respectivement, $t(22) = 0.39$, $p = 0.70$. En revanche, on observe un clair effet de facilitation pour les paires contrôles, le temps de réponse moyen pour les paires reliées étaient de 524 ms, alors que celui pour les paires non reliées était de 553 ms. Cet effet de facilitation sémantique 29 ms était significatif, $t(46) = -2.046$, $p < .05$.

Les résultats montrent que les équivalents canoniques des formes assimilées (*soude*) ne facilitent pas le traitement des cibles (BAGAGE), contrairement aux effets observés pour formes assimilées ([sʊd]). Cette expérience contrôle avait pour but principal de montrer que les formes assimilées donnant lieu à d'autres mots, (cf. *soute* donnant [sʊd]), ne sont pas traitées comme les formes canoniques de ces mots dérivés (*soude*). Alors que le mot amorce *soute* prononcé [sʊd] par assimilation avait significativement facilité le traitement de la cible sémantiquement associée BAGAGE, sa réalisation sous la forme canonique *soude* n'a pas facilité le traitement de cette même cible.

Dans l'ensemble, nos résultats suggèrent que premièrement, la présentation des formes assimilées ne perturbe pas l'activation du réseau lexico-sémantique de leurs formes sous-jacentes. Deuxièmement, ceci est probablement dû à la présence des traces acoustiques exploitées par l'auditeur, comme nous l'avons suggéré dans les deux chapitres précédents. Quand les formes de surface générées par assimilation sont elles-mêmes réalisées sous leur forme canonique, l'effet facilitateur d'amorçage associatif n'est plus observé (i.e., la forme assimilée [sʊd], dérivée de *soute* amorce BAGAGE, alors que la forme canonique *soude* [sʊd] n'a pas facilité le traitement de BAGAGE).

4. DISCUSSION GENERALE

Ce chapitre avait pour objectif principal d'étudier les conséquences perceptives de la parole assimilée, mais sous un angle quelque peu différent du chapitre précédent. Si le chapitre trois était majoritairement consacré à l'étude du rôle du contexte phonologique des segments plus ou moins fortement assimilés, nous avons ici cherché à savoir si l'activation du réseau lexico-sémantique d'un mot est affectée, lorsque celui-ci est réalisé sous forme assimilée. A cette fin,

nous avons sélectionné des mots ambigus par assimilation (MAA, *soute/soude* [sʊd]) et des mots non ambigus (MNA, *jupe/jube* [ʒyb]). Les deux types de mots étaient insérés dans des contextes phrastiques donnant lieu ou non à l'assimilation. Ces mots ont été utilisés comme des amorces dans une tâche de décision lexicale avec une technique d'amorçage sémantique intermodal. Les cibles associées ou non associées aux mots amorces étaient sélectionnées à partir des résultats d'un pré-test d'association libre.

Les résultats de l'Expérience 7 ont mis en évidence un effet général d'amorçage sémantique. C'est-à-dire que les mots cibles donnaient lieu à des temps de réponse plus courts lorsqu'ils étaient précédés par des mots amorces sémantiquement associés que lorsqu'ils étaient précédés par des amorces non associées. Cet effet d'amorçage était obtenu *quel que soit le type de mot* (MAA ou MNA), ou de mode de réalisation (« canonique » vs. « assimilation »). A titre d'exemple, *bac* prononcé *bague* [bag] amorçait EXAMEN tout autant que *bac* prononcé [bak]. De même, *jupe* prononcé [zyb] amorçait ROBE tout autant que *jupe* [ʒyp].

L'Expérience 8 a permis de démontrer que les effets d'amorçage concernant les MAA, i.e les mots ayant un compétiteur lexical généré par l'assimilation (*bac/bague* ; *soude/soute*), ne sont pas dus à une sorte de tolérance générale pour des mots phonologiquement proches, puisque la présentation auditive du mot *bague* prononcée sous la forme canonique [bag] n'a pas facilité le traitement de la cible EXAMEN. Cela implique que le système de traitement perceptif fait clairement une distinction entre deux formes de réalisation de [bag], celle qui correspond à la forme assimilée de *bac* et contient des traces acoustiques de la forme sous-jacente, et celle qui correspond à la forme canonique de *bague*.

Nos résultats en français ne s'accordent pas avec les résultats en anglais et en néerlandais, cités dans l'introduction de ce chapitre, dans la mesure où ces études ont montré que les effets d'amorçage peuvent être modulés par la présence d'un mot compétiteur (cf. Marslen-Wilson, Moss, & Van Halen 1996 ; Andruski et al., 1994 ; Misiurski, et al. 2004 ; Van Alphen & McQueen, soumis). Comment peut-on rendre compte du fait que dans toutes ces études, la présence d'un mot compétiteur ait influencé l'accès au lexique, alors que nos résultats en français n'ont pas révélé un tel effet ? Une explication tentative peut être liée à la nature de la variation phonologique étudiée. Il n'est pas inutile de rappeler que dans toutes les études citées précédemment, les stimuli test utilisés étaient créés à partir d'une manipulation du signal acoustique, qui consiste en l'occurrence à faire varier la durée des périodes de VOT

dans la consonne initiale. Or, dans notre étude, la variation phonologique dans les mots stimuli n'a jamais été manipulée de façon expérimentale, mais elle est le résultat d'une règle phonologique d'assimilation, spécifique à la langue française. Dans le chapitre 2, nous avons déjà pu constater qu'en français, l'assimilation est un processus gradué plutôt que catégorique. Nous avons suggéré que des traces acoustiques appartenant à la forme sous-jacente, sont très souvent préservées dans le segment assimilé. Les données d'amorçage portent à croire que ces traces sont exploitées par les auditeurs dans une tâche perceptive d'amorçage sémantique. La forme assimilée de *bac* réalisée comme [baɡ], contient néanmoins des traces acoustiques appartenant à [bak], conforme aux « intentions » communicatives du locuteur (cf. Fowler & Houssum, 1987), dans le sens où il a voulu prononcer *bac* et non pas *bague*. L'auditeur, étant très sensible à la présence des traces acoustiques de voisement, a pu accéder à la forme sous-jacente du mot assimilé. De fait, même les formes fortement assimilées donnent lieu à un effet facilitateur pour les cibles associées aux mots sous-jacents.

Comme le font remarquer Burton, Baum et Blumstein (1989), il est généralement admis que l'influence des informations de haut niveau, comme les informations lexicales, sur la perception de la parole est fonction du degré d'incertitude créé par le stimulus. Ainsi, si la valeur informative des indices acoustiques présents dans le signal de parole du mot ne suffit pas pour son identification, d'autres sources d'information de plus haut niveau peuvent alors être utilisées. À part le degré d'incertitude, l'émergence des effets lexicaux sur la perception de la parole peut aussi être liée à la présence des indices en conflit (*conflicting cues*, voir aussi Marslen-Wilson et Warren, 1994), comme c'est parfois le cas pour les stimuli générés superficiellement par manipulation phonétique d'un indice acoustique (e.g., la durée du VOT). À titre d'exemple, Burton et al. (1989) ont manipulé la durée du VOT en anglais, afin d'examiner plus précisément le rôle des paramètres acoustiques sur les effets lexicaux observés en catégorisation phonétique. Ces auteurs ont créé artificiellement deux continua de stimuli tests en manipulant le VOT : un continuum était le résultat des variations dans la durée de VOT allant du mot *duke* au non-mot *tuke*, alors que l'autre continuum était le résultat des variations allant du non-mot *doot* au mot *toot*. Les résultats de la catégorisation phonétique montraient un effet lexical, généralement observé dans la littérature (cf. Ganong, 1980 ; Samuel, 1981). Il y avait un changement abrupt dans l'identification de la frontière phonétique en fonction du statut lexical du stimulus. En revanche, quand les stimuli ressemblaient plus à la parole naturelle, l'effet lexical n'était plus observé, ce qui a été montré

dans une seconde expérience. Burton et al. (1989) ont fait covarier la durée VOT et l'amplitude, ce qui avait pour but de rapprocher les valeurs des paramètres acoustiques de la parole naturelle tout en évitant des indices acoustiques conflictuels. Dans ce cas, les auteurs ont observé que le locus de la frontière phonétique était identique pour les deux continua *duke-tuke* et *doot-toot* (pour les temps de réponse longs ainsi que courts). Dans leur ensemble, ces résultats suggèrent donc que le rôle des informations lexicales ne devrait pas être surestimé. Elles peuvent effectivement influencer la perception de la parole, mais à condition qu'il y ait beaucoup d'incertitude au sens où les stimuli test ressemblent peu à la parole naturelle.

Pour revenir à nos expériences, il semble donc que la présence des traces acoustiques dans les formes assimilées des MAA soit suffisante pour activer le réseau lexico-sémantique de leurs formes sous-jacente. Le niveau d'activation de la forme sous-jacente (e.g., *soute*) ne semble pas être affecté par l'activation éventuelle du réseau de la forme de surface²⁶.

Ces données suggèrent alors que l'unité de traitement de l'information, extraite à partir du signal acoustique, présente un caractère riche et gradué (McQueen, Dahan, & Cutler, 2003). Des études montrant que les variations sous-phonémiques interviennent lors de l'accès au lexique, semblent remettre en question l'idée selon laquelle l'activation lexicale serait un processus discret, calculé sur la base de nombre de phonèmes correspondant au signal d'entrée (cf. Streeter & Nigro, 1979 ; Marslen-Wilson & Warren, 1994). En conséquence, ces données ne sont pas en accord avec les modèles comme TRACE ou Shortlist²⁷, qui ont implémenté un niveau pré-lexical avec pour unité de traitement le phonème. En revanche, selon d'autres modèles tels que LAFS (*Lexical Access From Spectra*, Klatt, 1979) ou les modèles de type épisodique / à exemplaires, le signal acoustique est directement projeté sur le niveau lexical. Dans le modèle LAFS, chaque mot est représenté sous la forme d'une

²⁶ Nos résultats ne permettent pas de conclure que la représentation sous-jacente des MAA est privilégiée à la forme de surface, car nous n'avons pas directement comparé le niveau d'activation entre ces deux formes. Il aurait fallu insérer une condition supplémentaire dans laquelle l'amorce assimilée est suivie par une cible sémantiquement reliée à la forme de surface (e.g., [sud] - CHLORE). Cependant, une telle condition supplémentaire soulève le problème de comparaison entre les temps de réponse aux cibles différentes (e.g., BAGAGE vs. CHLORE), dont toutes les propriétés devraient pouvoir être contrôlées.

²⁷ Norris, McQueen, et Cutler (2000) ont toutefois remarqué que l'existence d'un niveau pré-lexical avec le phonème comme unité de base peut quand même rendre compte des effets sous-phonémiques en postulant une nouvelle distinction entre informations segmentales et supra-segmentales au sein même du niveau phonémique.

séquence de spectres acoustiques et le lexique mental est assimilé à une sorte de treillis comportant toutes les séquences de spectres possibles associées à toutes les combinaisons lexicales possibles. D'après les modèles de type épisodique tel que MINERVA (Goldinger, 1998), chaque expérience linguistique peut créer une trace indépendante en mémoire, dans laquelle est intégrée toute la structure acoustique détaillée. Il est présumé qu'à la présentation auditive d'un mot, une sonde à l'image de ce mot (*analog probe*) est mise en lien avec toutes les traces en parallèle.

Nos données sont plus compatibles avec celles obtenues par Gow (2002), qui avait utilisé la technique d'amorçage intermodal de répétition avec des mots stimuli en anglais. Cet auteur a montré que la parole assimilée n'est pas pour autant perçue comme ambiguë dans une tâche en temps réel (e.g., la décision lexicale avec technique d'amorçage), même si elle peut créer une ambiguïté dans une tâche en temps différé (e.g., catégorisation phonémique). Dans ces expériences, les mots stimuli étaient insérés dans des phrases donnant lieu à une assimilation de place (e.g., *this time she tried to get the right [raɪp] berries for her pies*) qui ensuite ont été présentées auditivement (en entier) aux participants. Les cibles visuelles étaient affichées à la fin de l'amorce critique. Les résultats ont montré que la forme assimilée a amorcé la cible RIGHT et non pas la cible RIPE, même si les auditeurs avaient préférentiellement jugé *ripe* comme étant la forme la plus proche de la forme assimilée dans un pré-test en temps différé.

Spinelli, McQueen, et Cutler (2003) se sont également intéressés au rôle de l'ambiguïté lexicale en évoquant un autre type de variation systématique en français, à savoir la liaison. Ces auteurs ont examiné l'activation des voyelles et consonnes en position initiale (e.g., *rognon* et *oignon*) insérées dans des phrases comme *C'est le dernier oignon*. Dans ce contexte, la consonne finale [R] de *dernier* est produite et resyllabifiée avec la syllabe suivante, ce qui rend la phrase ambiguë car interprétable comme *C'est le dernier rognon*. Cependant, des mesures acoustiques avaient révélé qu'il y avait des différences significatives entre les durées des consonnes, selon les intentions communicatives du locuteur, i.e., la durée de la consonne [R] était plus longue dans *dernier rognon* que dans *dernier oignon*. Quand ces mots étaient présentés dans une tâche d'amorçage intermodal, seules les réponses correspondant aux intentions du locuteur ont facilité le traitement. Même en cas d'ambiguïté où l'information est compatible avec un /r/, la différence de durée semble avoir influencé l'accès lexical, ce qui indique une sensibilité perceptive aux informations acoustiques fines (voir aussi Gow & Gordon, 1995).

Pour résumer, l'objectif principal de ce chapitre était d'examiner l'activation du réseau lexico-sémantique des formes assimilées, en utilisant la technique d'amorçage sémantique. Nous n'avons pas observé, comme certains auteurs, un effet d'ambiguïté lexicale et nous avons cherché à expliquer cette absence d'effet lexical en insistant sur la nature même de la variation phonologique en question. Si la variation phonologique consiste en une manipulation expérimentale, des effets lexicaux peuvent influencer le traitement lexical. En revanche, nos données suggèrent qu'il est bien plus difficile de mettre en évidence de tels effets, tant en production qu'en perception de la parole, quand il s'agit d'une variation phonologique motivée par les règles phonologiques de la langue en question, comme c'est le cas de l'assimilation de voisement en français. Si cette explication s'avère correcte, il devrait alors être possible de trouver un effet lexical en manipulant l'indice d'assimilation de voisement de manière systématique.

CHAPITRE V.

DISCUSSION GENERALE

1. RESUME DES RESULTATS PRINCIPAUX

L'objectif majeur de cette thèse était de mieux comprendre la production et la perception d'une variation phonologique fréquemment observée en français : l'assimilation de voisement. L'étude des variations systématiques, motivées par les règles phonologiques de la langue, a fait l'objet de très nombreux travaux en psycholinguistique, dont l'objectif est de mieux comprendre la contribution de l'auditeur dans la détermination de ce qui est perçu.

La revue de la littérature expérimentale sur l'assimilation a montré qu'il est essentiel de prendre en compte les propriétés acoustico-phonétiques de la variation en question pour mieux comprendre les caractéristiques de leur traitement perceptif. La première série d'expériences (rapportée dans chapitre 2), a eu pour objectif d'étudier de façon approfondie la production des mots assimilés, afin de caractériser l'assimilation d'un point de vue acoustico-phonétique. Des locuteurs français ont prononcé des mots insérés dans des contextes phrastiques donnant lieu ou non à une assimilation de voisement. Il y avait soit des « mots ambigus par assimilation » (MAA), soit des « mots non ambigus » (MNA). Nos résultats ont révélé que l'indice acoustique le plus fiable de l'assimilation de voisement était la portion voisée au sein de l'occlusion. A partir de cette mesure acoustique, nous avons proposé un nouvel indice permettant d'estimer le degré d'assimilation d'une consonne en position finale. Afin d'évaluer la validité perceptive de cet indice acoustique, nous avons mis en œuvre une tâche de catégorisation phonétique. Les mots stimuli étaient extraits de leur contexte phrastique et présentés auditivement aux participants dont la tâche était de catégoriser la consonne finale. Une corrélation négative a été observée entre les jugements perceptifs et le degré d'assimilation (mesuré à partir de l'indice proposé).

D'un point de vue acoustico-phonétique, les résultats de cette étude ont révélé deux points importants. Premièrement, l'assimilation de voisement n'implique pas forcément un changement phonologique catégorique. Cette variation phonologique est plutôt de nature graduée : il existe des degrés intermédiaires de transformation d'un son en un autre. Deuxièmement, il s'est avéré que les mots non voisés se voient plus facilement et complètement que les mots voisés ne se dévoient.

Enfin, nous avons également observé que le degré d'assimilation des segments peut parfois être modulé par des facteurs lexicaux comme l'ambiguïté lexicale et le nombre des voisins lexicaux plus fréquents. Ceci suggère que les mots sont produits avec plus de soin (i.e. moins de variation) dès qu'il existe un risque de confusion lexicale.

Après avoir fourni une description des caractéristiques acoustico-phonétiques de l'assimilation de voisement entre les mots, nous étions en mesure d'étudier plus précisément les conséquences perceptives de l'assimilation dans une deuxième série d'expériences d'amorçage intermodal, rapportée dans le chapitre 3. L'objectif principal de ces expériences était de démontrer que les deux mécanismes perceptifs proposés dans la littérature, à savoir, l'inférence régressive et l'exploitation des traces acoustiques du trait de voisement sous-jacent présentes dans le signal, sont en fait compatibles dès lors que l'on prend en considération le degré d'assimilation des segments.

Dans une première expérience d'amorçage de répétition, l'amorce auditive était constituée uniquement du mot stimulus et de son déterminant (e.g., *une jupe*), c'est-à-dire que le contexte droit donnant lieu à l'assimilation n'était pas rendu disponible aux sujets. Dans une seconde expérience, l'amorce auditive comportait également le contexte droit (e.g., *une jupe grise*). Les résultats de ces expériences ont montré que les temps de traitement étaient significativement plus longs quand l'amorce était un mot assimilé par rapport à la condition de réalisation canonique. Plus important, l'interaction attendue entre la présence du contexte droit et le voisement sous-jacent a été observée. Plus précisément, les données obtenues montrent que seuls les mots non voisés, i.e. les mots ayant subi de fortes assimilations, ont bénéficié de la présence du contexte droit. Il semblerait donc que deux mécanismes perceptifs soient mis en jeu dans le traitement perceptif des segments assimilés, selon que ceux-ci sont plus ou moins fortement assimilés. Un mot fortement assimilé tel que *jupe*, prononcé [ʒyb] va bénéficier de la présence du contexte droit (e.g., *grise*), car c'est à partir du segment suivant que l'auditeur infère régressivement le voisement présent dans la forme de surface. Il est présumé que l'auditeur français sait implicitement que les obstruents non voisés peuvent devenir voisés dans le contexte voisé. En revanche, le traitement perceptif des mots partiellement assimilés, (tels que *robe* prononcé [rɔp]) ne ferait, quant à lui, pas recours au contexte droit. Pour ce type de mots, il est présumé que les traces acoustiques de voisement sont présentes dans le signal de parole. Grâce à la présence de ces traces, l'auditeur aurait accès à la forme sous-jacente du segment partiellement assimilé.

Ce dernier point est confirmé par les résultats rapportés dans le chapitre 4. Dans ces expériences, nous avons examiné l'activation du réseau lexico-sémantique lors de la présentation des mots assimilés. Il a été montré dans la littérature que les représentations des mots peuvent parfois être activées, malgré le fait que les stimuli utilisés ne correspondent pas à leur forme de réalisation canonique. Par exemple, des études ont montré que des items non-mots, dérivés des mots par le changement de la consonne initiale, peuvent faciliter le traitement des cibles qui sont reliées sémantiquement aux mots d'origine (cf. Connine, Blasko, & Titone, 1993). Toutefois, l'ampleur des effets d'amorçage semble varier en fonction de la *distance phonologique* entre le mot canonique et le non-mot dérivé (Andruski, Blumstein, & Burton, 1994 ; Misiurski, Blumstein, Risan, & Berman, 2004). Les expériences que nous avons conduites ont eu pour but d'établir si l'activation d'un mot « compétiteur », généré par assimilation (par exemple [bag] dérivé de *bac* [bak] assimilé), peut gêner l'activation du réseau lexico-sémantique du mot d'origine *bac*. Pour faire cela, nous avons employé la technique d'amorçage sémantique intermodal. Nous avons de nouveau utilisé des mots de type MAA (e.g., la paire *soute-soude*), ainsi que des mots de type MNA (e.g., la paire *jupe-jube*). Les deux types de mots étaient insérés dans des contextes donnant lieu ou non à l'assimilation. Ensuite, les mots stimuli ont été extraits de leur contexte phrastique et présentés auditivement à nos sujets. Nous nous attendions à ce que l'effet de facilitation sémantique soit modulé par l'ambiguïté lexicale et donc à une différence des effets d'amorçage entre les MAA et MNA. Dans le cas des MNA, la présentation du non-mot assimilé (e.g., [ʒyb] dérivé de *jupe*) doit activer les mots associés du mot d'origine, tels que *robe* ou *veste*, étant donné que c'est la forme phonologique la plus proche du mot d'origine. En revanche, dans le cas des MAA, la présentation de la forme assimilée, étant elle-même un mot français (e.g., [sʊd] dérive de *soute*), pourrait activer davantage ses propres associés. Donc, l'effet de relation sémantique pour la forme sous-jacente devrait être plus important pour les MNA que pour les MAA.

Contrairement à nos prédictions, les effets d'amorçage sémantique sont analogues pour les deux types de mots. Les amorces MAA ont amorcé les mots cibles tout autant que les amorces MNA. Une expérience contrôle a permis de démontrer que l'on n'observe plus d'effet de facilitation sur la cible (e.g., EXAMEN), quand les équivalents canoniques des formes assimilées sont présentées (e.g., *bague* [bag]). C'est donc bel et bien la forme assimilée du mot d'origine qui est responsable de l'effet préalablement observé. Ces résultats montrent que

de par la présence de traces acoustiques dans le signal, la présentation des formes assimilées active le réseau lexico-sémantique du mot d'origine.

Force est de constater que l'influence des informations lexicales de plus haut niveau s'est avérée relativement limitée dans nos expériences. Si la présence d'un compétiteur lexical généré par assimilation (e.g., *soude / soute*) a peu affecté la production des segments assimilés, elle n'a pas affecté le traitement perceptif en amorçage sémantique. L'absence de tout effet d'ambiguïté lexicale est compatible avec l'idée selon laquelle les traces acoustiques, présentes dans le signal, permettent justement de désambiguïser le segment soi-disant ambigu. Cela ne veut pas pour autant dire que les formes assimilées et canoniques sont traitées de façon identique, comme cela a été montré dans nos expériences d'amorçage de répétition (chapitre 3). Certains auteurs ont mis en évidence que la présence des compétiteurs phonologiquement proches affecte le traitement lexical (cf. Andruski et al., 1994; Misiurski et al., 2004). Nous avons alors suggéré que la nature même de la variation en question (i.e. une manipulation du signal vs. une variation phonologique naturelle) peut être à l'origine de ces données divergentes.

2. IMPLICATIONS DES RESULTATS POUR LES THEORIES PSYCHOLINGUISTIQUES

Dans l'ensemble, nos résultats sont en accord avec l'idée principale selon laquelle le signal acoustique contient des informations détaillées qui sont accessibles à l'auditeur. Ces informations sont utilisées par l'auditeur lors des procédures d'accès au lexique. Comme nous l'avons mentionné dans l'introduction, les données acoustiques sur l'assimilation de lieu d'articulation en anglais avaient montré que cette variation est rarement (voire jamais) de nature complète, mais plutôt de nature graduée. De par la nature graduée même des segments assimilés, il avait été proposé que des mécanismes perceptifs généraux soient à l'œuvre dans le traitement de l'assimilation de place en anglais (cf. Gow 2003 ; Gow & Zoll, 2002). Par exemple, un segment assimilé comme *right* (prononcé [raip]) contient des indices acoustiques correspondant au lieu d'articulation coronale et au lieu labial induit par le segment suivant (e.g., /b/ dans *berries*). Le processus d'alignement et de regroupement des traits fait en sorte que tout indice de labialité présent dans *right* soit attirée par le segment suivant, laissant la place à l'indice de coronalité dans le signal. C'est donc l'attraction et l'alignement des traits qui permet à l'auditeur d'accéder ainsi à la forme sous-jacente.

Les données perceptives rapportées dans cette thèse constituent les premières données empiriques ayant montré que le rôle du contexte droit était modulé par le degré d'assimilation des segments de parole. En effet, nos expériences avaient pour objectif d'examiner les conséquences perceptives de l'assimilation dans la parole naturelle. A l'aide de notre indice d'assimilation, nous avons pu quantifier le voisement. A partir de nos mesures acoustiques, nous avons observé une asymétrie de voisement entre obstruents non voisés et voisés, ce qui nous a permis d'opérationnaliser le degré de voisement, sans pour autant manipuler délibérément le matériel expérimental.

Nous avons proposé que la présence des traces acoustiques de voisement dans les segments partiellement assimilés en français suffirait pour pouvoir accéder à la forme sous-jacente du mot, sans devoir recourir à d'autres mécanismes perceptifs supplémentaires. Quant aux segments complètement assimilés, nous avons suggéré que le contexte phonologique est employé par l'auditeur par inférence régressive de la forme sous-jacente du segment assimilé (cf. Gaskell & Marslen-Wilson, 1996, 1998). Si les deux types de mécanismes cognitifs ont fréquemment été présentés comme des hypothèses alternatives pour rendre compte des données empiriques du traitement perceptif de l'assimilation, nous proposons ici qu'il pourrait en fait s'agir de deux hypothèses complémentaires, dès lors que l'on prend en considération le degré d'assimilation.

Comme nous l'avons déjà mentionné préalablement, Gaskell (2003) a proposé un nouveau modèle connexionniste, inspiré largement de son ancien modèle (Gaskell, Hare, & Marslen-Wilson, 1995), mais limité aux niveaux des traits de façon à pouvoir se focaliser sur les effets d'assimilation de lieu d'articulation. Après entraînement, le modèle était capable d'interpréter une séquence telle que *leam* (*bacon*) comme une représentation de surface de la forme canonique *lean* (*bacon*). Ce qui distingue le modèle actuel de sa version précédente, c'est la prise en compte de la nature graduée de l'assimilation. C'est-à-dire qu'un continuum de 0-100% d'assimilation a été utilisé dans les simulations. Un modèle probabiliste tel que celui présenté par Gaskell rend donc bien compte des données expérimentales obtenues sur l'assimilation de lieu d'articulation en anglais. Nos données empiriques en français sont également en accord avec une telle approche probabiliste. Bien entendu, ceci ne peut être confirmé que par la modélisation de nos données acoustiques et perceptives.

Le fait que les variations systématiques observées dans une langue soient motivées par les règles phonologiques a conduit certains à proposer que le traitement perceptif des variations phonologiques serait en fait guidé par les mécanismes cognitifs spécifiques à la langue en

question. Nous avons déjà vu un tel mécanisme sous la forme de l'inférence régressive proposée par Gaskell et collègues (1995, 1996, 1998). Il est présupposé que l'auditeur de langue maternelle anglaise prend appui sur le contexte phonologique, afin d'inférer la forme sous-jacente du segment assimilé, car il a des connaissances implicites des règles phonologiques de sa langue stipulant qu'une consonne coronale peut prendre la forme d'une labiale dans le contexte labial (mais jamais l'inverse). En revanche, Gow (2002, 2003) propose que l'alignement et le regroupement perceptifs des traits feraient partie des mécanismes perceptifs plus généraux. En conséquence, leur mise en œuvre ne dépendrait pas des expériences linguistiques préalables de l'auditeur. Il est clair que seules les études intra-langue avec plusieurs populations linguistiques permettent de pouvoir trancher sur la question de savoir si le traitement perceptif est guidé par des mécanismes spécifiques à la langue en question ou des mécanismes perceptifs généraux (cf. Otake, Yoneyama, van der Lugt, & Cutler, 1996 ; Weber, 2001, 2002). Très récemment, Gow et Im (2004) ont comparé les performances entre des participants hongrois et des participants américains n'ayant aucune compétence en hongrois. Ces auteurs ont présenté des stimuli hongrois ayant subi une assimilation régressive de voisement. Des analyses acoustiques sur ces stimuli ont révélé que l'assimilation de voisement en hongrois est un processus gradué plutôt que dichotomique. Les deux groupes de participants ont détecté plus rapidement des cibles voisées quand celles-ci étaient suivies par des segments assimilés. Ces résultats suggèrent que les mécanismes sous-tendant le traitement des segments assimilés par voisement reflètent des propriétés générales dans le traitement perceptif de l'assimilation de voisement en hongrois, et opèrent indépendamment de l'expérience linguistique des auditeurs²⁸. Dans la même étude, ces auteurs ont également examiné le cas d'assimilation de place vélaire-à-labiale, un phénomène qui est régulièrement observé en langue coréenne mais jamais en anglais. Cette fois-ci, les analyses acoustiques ont montré que ce type d'assimilation implique un processus discret plutôt que gradué. Des participants américains n'ayant aucune connaissance de la langue

²⁸ Comme le notent Gow et Im (2004) à juste titre, le modèle de Gaskell (2003) est également capable de rendre compte de ces résultats. Des formes d'assimilation de voisement progressif existent dans cette langue, ce qui laisse suggérer que les auditeurs de langue maternelle anglaise, étant quelque peu familiarisés avec ce type de variation, avaient des performances comparables à celles des participants hongrois. Il est d'ailleurs à noter que certaines formes d'assimilation régressive ont été attestées en anglais américain (Lansing, communication personnelle), même s'il s'agit de formes vernaculaires, comme le changement de voisement /v/ → /f/ dans *we have to*, la création de formes plurielles écrites avec /s/ mais prononcées comme /z/ (e.g., *cabs*), ou encore les suffixes « -ed » dans les verbes au passé, qui peuvent être prononcés comme /t/ (e.g., *he backed the car*).

coréenne avaient pour tâche de détecter le phonème-cible dans des paires syllabiques. Il n'y avait pas d'effet de facilitation pour les phonèmes-cibles suivant les phonèmes assimilés, ce qui est compatible avec l'observation en anglais qu'il n'y a pas ou peu d'effet progressif quand il s'agit des assimilations complètes (cf. Gaskell & Marslen-Wilson, 1998). Qui plus est, il y avait également une absence d'effet progressif quand la même expérience était effectuée avec des participants coréens. Dans l'ensemble, ces résultats montrent donc que l'absence d'effet de contexte progressif ne pouvait pas être attribuée à des mécanismes perceptifs spécifiques à la langue coréenne, étant donné les résultats similaires obtenus avec les participants américains qui ne parlaient pas le coréen. Gow et Im (2004) ont alors proposé que ce sont des mécanismes perceptifs généraux qui guident la perception de la parole assimilée, bien que ces auteurs admettent également que les auditeurs s'engagent probablement aussi dans des processus de groupement de type descendant, sous forme de différents schémas linguistiques :

In addition to the more familiar bottom-up grouping principles that have been identified in the perception of auditory and visual stimuli, it is clear that listeners also engage in top-down schema driven grouping processes. [...] While the mechanism of schema-driven grouping may be universal, listeners with different linguistic backgrounds should engage different schemata. (Gow et Im, 2004 : 293).

3. LIGNES DIRECTIVES POUR DE FUTURES RECHERCHES

3.1. La modélisation des variations phonologiques

L'objectif de l'utilisation d'un modèle computationnel en psychologie cognitive est de simuler les processus cognitifs impliqués dans le traitement étudié. L'implémentation d'un modèle précis obéit à plusieurs raisons dont les suivantes (d'après Dijkstra & De Smedt 1996) : premièrement, une simulation sur ordinateur peut permettre de vérifier de façon exhaustive si les différentes parties d'un modèle théorique proposé sont cohérentes et si le modèle est complet. Deuxièmement, une simulation computationnelle peut révéler des aspects inconnus dans l'interprétation des données qui semblent, à première vue, contradictoires. Enfin, une simulation n'est pas seulement capable de prédire les résultats pour les conditions qui sont empiriquement connues et vérifiées, mais elle fait aussi des prédictions en ce qui concerne les conditions inconnues. (Par le passé, nous avons déjà eu l'occasion de compléter les données empiriques avec un modèle computationnel dans nos recherches sur l'organisation du lexique mental lors des premières étapes de l'apprentissage d'une langue étrangère, cf. Snoeren, Dijkstra, Grainger, & Van Heuven, soumis.)

3.1.1. *Modélisation de l'assimilation en anglais et en français*

Il serait intéressant d'étendre nos études sur l'assimilation de voisement en français par sa modélisation informatique. L'objectif principal serait ainsi de déterminer si le modèle de Gaskell (2003) est capable de simuler ces données du français. Pour cela ce modèle sera enrichi en incorporant des facteurs phonétiques (i.e. le voisement sous-jacent) ainsi que des facteurs lexicaux (i.e. la fréquence, l'ambiguïté lexicale et le voisinage phonologique).

Afin de pouvoir dégager les aspects du traitement qui sont spécifiques à la langue et ceux qui sont censés être universels, il serait alors très pertinent de comparer le traitement de l'assimilation entre deux langues (l'anglais et le français en l'occurrence). Nous pouvons envisager d'élaborer un modèle computationnel similaire à celui de Gaskell (2003) qui est entraîné à des mécanismes perceptifs plus généraux qui sont supposés se produire indépendamment de la langue, en prenant en compte des propriétés acoustico-phonétiques de la variation, et en particulier la nature graduée dans sa réalisation phonétique qui ne nécessite qu'un traitement ascendant. Un deuxième modèle pourrait être entraîné à repérer les régularités statistiques qui sont spécifiques à chaque langue, nécessitant un traitement descendant, en prenant en compte les transformations complètes des sons. Nous pourrions ainsi tester l'hypothèse selon laquelle la mise en œuvre des mécanismes cognitifs généraux ou spécifiques dans la résolution des segments de parole transformés dépend de l'ampleur de la variation. Il serait évidemment souhaitable de combiner les résultats de ces simulations avec les données des expériences comportementales, nécessitant des populations de participants avec pour langue maternelle le français ou l'anglais.

3.1.2. *Modélisation d'autres types de variations phonologiques*

L'assimilation de voisement en français est illustrative du problème de la *variabilité* de la parole. D'autres formes de variations phonologiques que l'on rencontre fréquemment en français comme celles de la liaison relèvent principalement de la problématique de la *segmentation*. Ainsi, les règles phonologiques qui motivent la liaison donnent lieu à des processus de re-syllabification (Gow & Gordon, 1995 ; Gaskell, Spinelli, & Meunier, 2002 ; Spinelli, McQueen, & Cutler, 2003). Cette problématique de la segmentation en mots du signal de parole a donné lieu à une multiplicité de travaux et reste centrale en psycholinguistique. Toutefois si, comme le suggèrent nos données acoustiques sur l'assimilation, les auditeurs sont capables d'exploiter la présence d'indices acoustico-phonétiques subtils présents dans le signal qui reflètent la réalisation d'une opération de liaison, alors la frontière entre les mots serait susceptible d'être prédite à partir de ces indices

acoustiques, tout comme le segment suivant (i.e. un liquide ou une voyelle). Cette situation est très similaire à celle de l'assimilation. Par conséquent, il devrait être possible de simuler les processus cognitifs mis en jeu lors de la segmentation des mots en français (cf. Gaskell, Spinelli, & Meunier, 2002). Néanmoins, le modèle actuel (Gaskell, 2003) comporte également un certain nombre de limitations. En particulier, le modèle ne semble pas être en mesure de rendre compte des variations qui impliquent la réduction des segments, comme c'est le cas pour l'effacement du schwa (e.g., Racine & Grosjean, soumis). Il semble important d'apporter des modifications au modèle afin de pouvoir étendre sa capacité à simuler d'autres types de variations que l'assimilation.

3.2. Assimilation et morphologie

Dans la présente recherche, nous avons examiné le cas de l'assimilation qui se produit entre mots. Cependant, nous avons mentionné dans l'introduction que ce genre de variation phonologique peut aussi se produire au sein même d'un mot, ce qui change sa structure interne (e.g., la consonne intervocalique voisée dans *médecin* peut être réalisée comme la consonne non voisée /t/ par la chute de schwa et par assimilation régressive provoquée par la fricative non voisée, cf. Rigault, 1967). Nous citons ici deux articles pour illustrer que les variations phonologiques peuvent avoir des conséquences sur le plan morphologique.

Les assimilations au sein des mots semblent particulièrement appropriées pour étudier les interactions entre les informations orthographiques et phonologiques d'une part, et la structure morphologique sous-jacente d'autre part, comme certains l'ont montré (cf. Hallé, Chéreau, et Segui, 2000 ; Gumnior, Zwitserlood, & Bólte, 2005). Hallé et al. (2000) ont présenté des portions des mots tels que *absurde* prononcé [apsyrd] où l'occlusive /b/ s'est dévoisée au profit de la fricative non voisée /s/, dans une tâche de dévoilage progressif. Il a été trouvé que les auditeurs entendent /b/ plus facilement que /p/ dans des mots comme *absurde* où il y a assimilation, alors que les résultats d'une expérience de détection de phonèmes ont révélé que les temps de détection du /p/ étaient plus rapides que ceux de /b/. Même si les auteurs ont favorisé le modèle de la cohorte pour expliquer leurs résultats, la décomposition morphologique peut également être à l'origine de ces résultats (i.e., l'activation dérivationnelle des préfixes comme {ab-} ou {sub-}).

Ernestus & Baayen (sous presse, voir aussi Ernestus & Baayen, 2003) ont examiné le rôle de la neutralisation incomplète dans le dévoisement final sur la création des pseudo-verbes au passé. En néerlandais, deux suffixes « -de » et « -te » sont employés pour indiquer la forme

au passé (première personne au singulier, e.g., *ik lach* vs. *ik lachte*, « je ris » vs. « je riais »). Ces auteurs ont présenté deux listes comportant des pseudo-verbos avec pour consonnes finales des obstruents ou fricatives. L'une des deux listes ne contenait que des obstruents non voisés, alors que l'autre liste contenait des obstruents voisés qui étaient censées être réalisées sans voisement (à cause de la règle du dévoisement final). Comme le dévoisement final en néerlandais est un processus de neutralisation incomplète (cf. Warner et al., 2004), il était présumé que les obstruents voisés en position finale contenaient tout de même des traces fines de voisement. Les analyses acoustiques ont effectivement montré que les réalisations des phonèmes /p/ et /t/ étaient significativement différentes par rapport aux phonèmes /b/ et /d/, malgré la règle qui stipule le dévoisement de ces phonèmes en position finale. Les deux listes de mots étaient ensuite présentées auditivement dans une expérience pendant laquelle les participants avaient pour tâche d'écrire le plus rapidement possible la forme au passé du non-mot qu'ils venaient entendre. Les résultats indiquaient que les participants attribuaient plus de préfixes voisés (« -de ») que non voisés (« -te ») aux mots contenant des obstruents ayant subi un dévoisement final, comportant toutefois des traces de voisement. Ceci indique que les auditeurs peuvent exploiter des indices acoustiques fins de voisement lors de la création des pseudo-verbos en néerlandais.

Relativement peu d'études ont été consacrées à des questions théoriques à mi-chemin entre les niveaux de traitement phonologique et morphologique. Nous pensons que les variations phonologiques telles que l'assimilation peuvent fournir de nouvelles pistes intéressantes pour l'étude morphologique.

3.3. Vers une perspective d'intégration de la production et perception de la parole

Tout au long de ce travail, nous n'avons cessé d'insister sur l'importance d'étudier les propriétés physiques des variations phonologiques pour mieux saisir les processus cognitifs qui sous-tendent leur traitement perceptif. Nous espérons avoir démontré qu'une telle approche peut s'avérer fructueuse pour l'étude psycholinguistique sur la parole. Qui plus est, nous sommes d'avis qu'une perspective d'intégration de la production et de perception de la parole va bien au-delà des implications pour l'étude psycholinguistique et qu'elle peut même s'inscrire dans un cadre phénoménologique.

Si de nombreux études en psycholinguistique ont été consacrées soit à des modèles de perception, soit à des modèles de production de la parole, les psycholinguistes admettent de plus en plus souvent la nécessité d'un modèle d'intégration de ces deux volets :

[...] We genuinely believe that the production-comprehension correspondence should be taken seriously if we want to understand how the mapping of speech input onto the lexicon is accomplished. After all, most listeners are also speakers. (Coenen et al., 2001 : 559).

Très récemment, Schiller et Meyer (2003) ont édité un ouvrage, dans lequel on a également souligné l'importance de prendre en compte les travaux en production et en perception pour la théorie psycholinguistique. Ils ont résumé ainsi le point de vue sur ce sujet partagé par de nombreux psycholinguistes :

First, concerning the representation of phonological and phonetic knowledge, there is wide agreement that language users either have one shared store of knowledge which they access in different ways in speech production and comprehension, or separate, but closely linked stores. Nobody assumes that there are entirely independent representations used exclusively in production and comprehension. (Schiller & Meyer, 2003 : 6).

Plusieurs cadres théoriques ont d'ores et déjà été proposés dans le but de réconcilier la perception et la production de la parole (cf. Lindblom, 1990), dont notamment la théorie motrice élaborée par Liberman et collègues (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Selon la théorie motrice de la parole, l'auditeur interprète le signal de parole en se basant sur les mouvements articulatoires dont celui-ci est le produit. Plus récemment, des propos similaires ont été formulés en phonologie articulatoire. Ainsi, Goldstein et Fowler (2003) partent de l'idée de l'inter-subjectivité qui est propre au langage :

[...] language use is almost an entirely *a between-person activity*, and it matters whether or not listeners perceive the language forms speakers intend to convey. [...] We start from the premise that languages evolved to be spoken and to be heard and, therefore, that language forms [...] are likely to be public events. (Goldstein & Fowler, 2003 : 196, nos italiques).

D'après ces auteurs, il faut qu'il y ait un système de formes linguistiques (*language forms*) communes à ce qui est produit par le locuteur d'une part, et à ce qui est perçu par l'auditeur d'autre part. Il est intéressant de noter que dans le cadre de la phénoménologie de la

perception, le philosophe Merleau-Ponty (1945, 1966)²⁹ a déjà avancé des idées similaires quand il dit qu' :

[...] il faut que la gesticulation phonétique utilise un alphabet de significations déjà acquises, le geste verbal s'exécute dans un *certain panorama commun aux interlocuteurs*, comme la compréhension des autres gestes suppose un monde perçu commun à tous. (Merleau-Ponty, 1945 : 226, nos italiques).

Si cet *alphabet de significations* correspond à des traits acoustiques, des gestes sous forme de variables de l'appareil vocal, des phonèmes ou des syllabes reste encore à savoir. Dans tous les cas, l'auditeur qui perçoit les segments de parole de sa langue, est également locuteur de ces mêmes segments de parole. Ceci a amené certains à suggérer que la production et la perception de la parole doivent être intimement liées. Nous sommes portés à croire que leur fonctionnement peut être réellement compris dès lors que la production et la perception de la parole seront considérées comme les deux versants du même système cognitif.

²⁹ Il est important de souligner que dans ses travaux sur la parole, Merleau-Ponty (1945, voir aussi *La prose du monde*, 1966) a rigoureusement rompu avec la tradition saussurienne classique de l'arbitraire du signe, qui veut que le lien unissant signifiant (i.e. le concept) et signifié (i.e. l'image acoustique) est strictement arbitraire. Dans la conception phénoménologique, la distinction entre signifiant et signifié ne s'impose plus (cf. Eluerd, 1977 ; Touchet, 2002).

CONCLUSION

Dans la présente recherche, nous avons abordé l'étude de la production et la perception de l'assimilation de voisement, une variation phonologique fréquemment observée en langue française. Les données empiriques antérieures (essentiellement conduites en anglais), ont suggéré que la mise en œuvre de deux mécanismes de traitement perceptif peut varier en fonction du degré d'assimilation, d'où la nécessité d'avoir une description acoustico-phonétique adéquate de la variation. Nos données empiriques en français permettent de confirmer que les deux mécanismes de traitement perceptif proposés dans la littérature portant sur l'assimilation de lieu d'articulation, jouent un rôle selon que les segments sont plus ou moins fortement affectés par l'assimilation. En cas d'assimilation complète, l'auditeur prendrait appui sur le contexte droit pour inférer l'identité sous-jacente du segment assimilé, alors qu'en cas d'assimilation partielle, la présence des traces acoustiques suffirait pour y accéder.

Le fait que l'assimilation de voisement en français soit un phénomène de nature graduée plutôt que dichotomique, est en accord avec de nombreuses données acoustiques recueillies dans d'autres langues. Les mesures acoustiques rapportées ont eu le mérite de démontrer que le signal de parole ayant subi toutes sortes de variabilités, contient apparemment des « traces » des segments d'origine. Qui plus est, l'auditeur semble particulièrement sensible à la structure détaillée du signal acoustique, et ces informations sont utilisées lors du traitement lexical. Vues sous cet angle, les variations systématiques ne peuvent plus vraiment être considérées comme du bruit qui gêne la reconnaissance de la parole, mais plutôt comme une source riche d'informations, mise au profit par l'auditeur dans ses recherches lexicales.

Nous tenons enfin à souligner que tout modèle théorique visant à décrire les processus impliqués dans la perception de la parole devrait prendre en compte les propriétés physiques du signal. Avec les études rapportées dans cette thèse, nous espérons avoir contribué à l'élaboration à la fois méthodologique et théorique d'un modèle psycholinguistique, selon lequel la perception et la production de la parole sont considérées comme deux versants du même système de traitement langagier.

REFERENCES

- Andruski, J. E., Blumstein, S. E., & Burton, M. W. (1994). The effect of subphonetic differences on lexical access. *Cognition*, *52*, 163-187.
- Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, *5*, 183-207.
- Baayen, R. H. (2004). Statistics in Psycholinguistics: A critique of some current gold standards. In *Mental Lexicon Working Papers I*, The University of Alberta : Edmonton, pp. 1-45.
- Baayen, R. H., Tweedie, F. J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family size effect in the mental lexicon. *Brain & Language*, *81*, 55-65.
- Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, *44*, 568-591.
- Banel, M.-H., & Bacri, N. (1997). Reconnaissance de la parole et indices de segmentation métriques et phonotactiques. *L'Année Psychologique*, *97*, 77-112.
- Barry, M.C. (1985). A palatographic study of connected speech processes. *Cambridge Papers in Phonetics & Experimental Linguistics*, *4*, 1-16.
- Bates, D. (2005). Fitting linear mixed models in R. *R News. The newsletter of the R project*, *5*(1), 27-30.
- Berg, R. J. H. van den (1986). The effect of varying voice and noise parameters on the perception of voicing in Dutch two-obstruent sequences: Covariation of voicing cues. *Speech communication*, *5*(3-4), 355-367.
- Berg, R. J. H. van den (1989). Perception of voicing in Dutch two-obstruent sequences: Covariation of voicing cues. *Speech Communication*, *8*(1), 17-25.
- Best, C.T., Morongillo, B., & Robson, R. (1981). Perceptual relevance of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, *29*, 191-211.
- Blumstein, S.E. (1986). On acoustic invariance in speech. In J.S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 178-201). Hillsdale, NJ: Erlbaum.
- Boersma, P., & Weenink. (1992-2004). *Praat : A system for doing phonetics by computer*. <http://www.praat.org>.
- Boland, J. E., & Cutler, A. (1996). Interaction with autonomy: Multiple output models and the inadequacy of the Great Divide. *Cognition*, *58*, 309-320.
- Booij, G. (1995). *The phonology of Dutch*. New York : Oxford University Press ; Oxford : Clarendon Press.
- Bowers, J. S., & Davis, C. J. (2004). Is speech perception modular or interactive? *Trends in Cognitive Sciences*, *8*(1), 3-5.
- Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge: MIT Press.
- Browman, C. P., & Goldstein, L.M. (1990a). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*. (pp. 341-376). Cambridge: Cambridge University Press.
- Browman, C. P., & Goldstein, L.M. (1990b). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, *18*, 299-320.
- Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 567-575.

- Burton, M.W., & Robblee, K.E. (1997). A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics*, 25, 97-114.
- Bybee, J. (2001). *Phonology and Language Use*. Cambridge : Cambridge University Press.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 261-290.
- Calliope (1989). *La parole et son traitement automatique*. Paris : Masson.
- Carton, F. (1974). *Introduction à la phonétique du français*. Paris : Bordas.
- Casagrande, J. (1984). *The sound system of French*. Washington DC: Georgetown University Press.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.
- Clements, G.N. (1985). The Geometry of Phonological Features. *Phonology Yearbook*, 2, 223-252.
- Clements, G.N. (1993). Lieu d'articulation des consonnes et des voyelles: une théorie unifiée. In B. Laks & A. Riolland (Eds.), *L'architecture des représentations phonologiques*. Paris : CNRS Editions.
- Clements, G.N. & Hume, E.V. (1995). The internal organization of speech sounds. In John A. Goldsmith (Eds.), *The Handbook of Phonological Theory*. Cambridge, Mass. & Oxford: Blackwell.
- Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, 16(5-6), 535-564.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- 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, 193-210.
- Content, A., Mousty, P., & Radeau, M. (1990). Brulex : une base de données informatisée pour le français. *L'Année Psychologique*, 90, 551-566.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language-specific comprehension strategy. *Nature*, 304(5922), 159-160.
- Delgutte, B. (1986). Analysis of French stop consonants using a model of the peripheral auditory system. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 163-177): Hillsdale, NJ: Erlbaum.
- Dell, F. (1995). Consonant clusters and phonological syllables in French. *Lingua*, 95, 5-26.
- Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and Phonology in Language Comprehension and Production* (pp. 9-37). Berlin: Mouton de Gruyter.
- Dijkstra, T., & De Smedt, K. (Eds.). (1996). *Computational psycholinguistics. AI and connectionist models of human language processing*. London: Taylor & Francis.
- Dommelen, W. A. van (1983). Parameter interaction in the perception of French plosives. *Phonetica*, 40(1), 32-62.
- Dommelen, W. A. van (1985). Further experiments in the perception of French plosives. *Phonetica*, 42(2-3), 100-115.
- Duez, D. (1995). On spontaneous French speech: aspects of the reduction and contextual assimilation of voiced stops. *Journal of Phonetics*, 23, 407-427.

- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27, 429-446.
- Dufour, S., Peereman, R., Pallier C., & Radeau, M. (2002). Vocolex : Une base de données lexicales sur les similarités phonologiques entre les mots français. *L'Année Psychologique*, 102, 725-746.
- Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (sous presse). WordGen: A tool for word selection and nonword generation in Dutch, German, English and French. *Behaviour Research Methods, Instruments & Computers*.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Elman, J. L. (sous presse). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*.
- Elman, J. L., & McClelland, J. L. (1986). Exploiting lawful variability in the speech wave. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 360-385). Hillsdale, NJ: Erlbaum.
- Eluerd, R. (1977). *Pour aborder la linguistique*. Paris : Les Editions ESF.
- Ernestus, M., (2000). *Voice assimilation and segment reduction in casual Dutch, a corpus-based study of the phonology-phonetics interface*. LOT: Utrecht.
- Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79(1), 5-38.
- Ernestus, M., & Baayen, R. H. (sous presse). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In Goldstein, L.M., Whalen, D.H., & Best, C.T. (Eds.), *Papers in Laboratory Phonology VIII*.
- Ferrand, L., & Alario, F.-X. (1998). Normes d'association verbales pour 366 noms d'objets concrets. *L'Année Psychologique*, 98, 659-709.
- Ferrand, L., & New, B. (2003). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 25-43). Hauppauge, NY: Nova Science Publisher.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, 35, 116-124.
- Fowler, C. A. (1987). Perceivers as realists, talkers too: Commentary on papers by Strange, Diehl et al., and Rakerd and Verbrugge. *Journal of Memory and Language*, 26, 874-587.
- Fowler, C. A., & Housum, J. (1987). Talkers' signaling of new and old words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489-504.
- Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory ? *Journal of Experimental Psychology: Human Perception & Performance*, 16(1), 77-91.
- Fujimura, O., Macchi, M., & Streeter, L. A. (1978). Perception of stop consonants with conflicting transitional cues: A cross-linguistic study. *Language & Speech*, 21, 337-346.
- Gafos, A. (sous presse). Dynamics : The non-derivational alternative to modeling phonetics phonology. In : Goldstein, L.M., Whalen, D. H., & Best, C.T. (Eds.) *Papers in Laboratory Phonology VIII*.
- Ganong, W.F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology : Human Perception & Performance*, 6(1), 110-125.
- Gaskell, M. G., Hare, M., & Marlsen-Wilson, W. D. (1995). A connectionist model of phonological representation in speech perception. *Cognitive Science*, 19, 407-439.

- Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception & Performance*, 22(1), 144-158.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception & Performance*, 24(2), 380-396.
- Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoken word recognition: bridging the gap. *Journal of Memory and Language*, 44, 325-349.
- Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, 45(2), 220-266.
- Gaskell, M. G., Spinelli, E., & Meunier, F. (2002). Perception of resyllabification in French. *Memory & Cognition*, 30(5), 798-810.
- Gaskell, M. G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics*, 31, 447-463.
- Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and Phonology in language comprehension and production* (pp. 159-196). Berlin: Mouton de Gruyter.
- Goldinger, S.D. (1998). Echoes of echoes ? An episodic theory of lexical access. *Psychological Review*, 105(2), 251-279.
- Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language*, 24, 631-645.
- Gow, D.W., and Gordon, P.C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 344-359.
- Gow, D.W., & Hussami, P. (1999). *Acoustical modification in English place assimilation*. Article présenté au *Meeting of the Acoustical society of America*, Columbus, OH.
- Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, 45, 133-159.
- Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance*, 28, 163-179.
- Gow, D. W., & Zoll, C. (2002). Phonological answers (and their corresponding questions). In A. Csirmaz, Z. Li, A. Nevins, O. Vaysman & M. Wagner (Eds.), *The MIT Working Papers in Linguistics* (Vol. 42, pp. 55-68). Cambridge: MIT Press.
- Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, 65(4), 575-590.
- Gow, D. W., & IM, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, 51, 279-296.
- 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.
- Grainger, J. & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and marked identification latencies. *Perception & Psychophysics*, 47, 191-198.
- Grainger, J., Van Kang, M. N., & Segui, J. (2001). Cross-modal repetition priming of heterographic homophones. *Memory & Cognition*, 29(1), 53-61.

- Grammont, M. (1933). *Traité de phonétique*. Paris : Delagrave.
- Grosjean, F. (1980). Spoken word-recognition processes and the gating paradigm. *Perception & Psychophysics*, 28, 267-283.
- Gunnior, H., Zwitserlood, P., & Bölte, J. (2005). Assimilation in existing and novel compounds. *Language and Cognitive Processes*, 20(3), 465-488.
- Gussenhoven, C. Jacobs, H. (1998). *Understanding Phonology*. New York: Arnold Publishers.
- Goslin, J. & Frauenfelder, U.H. (2000). A comparison of theoretical and human syllabification. *Language and Speech*, 44(4), 409-436.
- Hallé, P. A., Chéreau, C., & Segui, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in the French listeners' mind. *Journal of Memory and Language*, 43, 618-639.
- Hawkins, S. (2003). Roles and representations systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, 31, 373-405.
- Hay, J., Pierrehumbert J., & Beckman, M. (2003). Speech perception, wellformedness, and the statistics of the lexicon. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in Laboratory Phonology VI* (pp. 58-74). Cambridge: Cambridge University Press.
- Isel, F., & Bacri, N. (1999). Spoken-word recognition: The access to embedded words. *Brain & Language*, 68, 61-67.
- Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.
- Jacobs, H. (1993). La palatalisation gallo-romane et la représentation des traits distinctifs. In B. Laks & A. Riolland (Eds.), *Architecture des représentations phonologiques* (pp. 147-171). Paris : CNRS Editions.
- Jakobson, R., & Halle, M. (1956). *Fundamentals of language*. The Hague : Mouton.
- Jakobson, R., Fant, G., & Halle, M. (1963). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.
- Jessen, M. (2001a). *Phonetics and phonology of the tense and lax obstruents in German*. Thèse de doctorat, Cornell University, Ithaca, NY.
- Jessen, M. (2001b). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In T. A. Hall (Ed.), *Distinctive Feature Theory* (pp. 237-294). Berlin: Mouton de Gruyter.
- Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach*. Thèse de doctorat, UCLA, Los Angeles.
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W.D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee, & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229-254). Amsterdam : John Benjamins.
- Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonological Yearbook 2*, 85-138.
- Klatt, D. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279-312.
- Kohler, K.J. (1979). Dimensions in the perception of fortis and lenis plosives. *Phonetica*, 36, 332-343.
- Kohler, K. J. (1985). The perception of lenis and fortis plosives in French: A critical re-evaluation. *Phonetica*, 42, 116-123.

- Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanation. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modeling* (pp. 69-92). Dordrecht: Kluwer.
- Kuzla, C. (2003). Prosodically-conditioned variation in the realization of domain-final stops voicing assimilation of domain-initial fricatives in German. In M.J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2829-2832). Barcelona, Spain.
- Labov, W. (1994). *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.
- Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38, 245-294.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7* (pp. 637-675). Berlin: Mouton de Gruyter.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- Lindblom, B. (1989). Phonetic invariance and the adaptive nature of speech. In B. A. G. Elsendoorn & H. Bouma (Eds.), *Working models of human perception* (pp. 139-173). London: Jovanovich Publishers.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H Theory. In W. J. Hardcastle & A. Marchal (eds.), *Speech production and speech modelling* (pp. 403-439). Netherlands: Kluwer Academic Publishers.
- Lisker, L., & Abramson, A.S., (1964). A cross-language study of voicing in initial stops: Acoustic measurements. *Word*, 20, 324-422.
- Lisker, L. & Abramson, A.S. (1971). Distinctive features and laryngeal control. *Language*, 47, 767-785.
- Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics*, 39(3), 155-158.
- Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighbourhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech processing* (pp. 122-147). Cambridge, MA: MIT Press.
- Luce, P.A., & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.
- Lukatela, G., Eaton, T., Sabadini, L., & Turvey, M. T. (2004). Vowel duration affects visual word identification: Evidence that the mediating phonology is phonetically informed. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1), 151-162.
- Magnuson, J.S., McMurray, B., Tanenhaus, M.K., & Aslin, R.N. (2003). Lexical effects on compensation for coarticulation: The ghost of Chrismash past. *Cognitive Science*, 27, 285-298.
- Marslen-Wilson, W. D., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 1376-1392.
- Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, 101(4), 653-675.
- Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101, 3-33.
- Mattys, S. (2004). Stress versus coarticulation: Toward an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 30(2), 397-408.

- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.
- McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*(4), 539-553.
- McQueen, J.M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 621-638.
- McQueen, J.M. (1996). Phonetic categorisation. *Language and Cognitive Processes*, *11*(6), 655-664.
- McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and Phonology in Language Comprehension and Production* (pp. 39-78). Berlin: Mouton de Gruyter.
- McQueen, J.M. (2003). The ghost of Christmas future: Didn't Scrooge learn to be good? Commentary on Magnuson et al. (2003). *Cognitive Science*, *27*, 795-799.
- Mehler, J., Dommergues, J., Frauenfelder, U.H., & Segui, J. (1981). The syllable's role in speech segmentation. *Language and cognitive Processes*, *4*, 57-67.
- Meunier, F., & Segui, J. (1999). Frequency effects in auditory word recognition: The case of suffixed words. *Journal of Memory and Language*, *41*, 327-344.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Paris : Editions Gallimard.
- Merleau-Ponty, M. (1969). *La prose du monde*. Paris : Editions Gallimard.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227-234.
- Milberg, W., Blumstein, S. E., & Dworetzky, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, *26*, 305-308.
- Mirman, D., McClelland, J.L., & Holt, L.L. (sous presse). Computational and behavioral investigations of lexically induced delays in phoneme recognition. *Journal of Memory and Language*.
- Misiurski, C., Blumstein, S. E., Rissman, J., & Berman, D. (sous presse). The role of lexical competition and acoustic-phonetic structure in lexical processing: Evidence from normal subjects and aphasic patients. *Brain and Language*.
- Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics*, *65*(6), 956-969.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: Lexique. *L'Année Psychologique*, *101*, 447-462.
- Nguyen, N. (sous presse). Perception de la parole. In N. Nguyen, S. Wauquier-Graveline & J. Durand (Eds.), *Phonologie et phonétique: Forme et substance*. Hermès.
- Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In G. J. Docherty & D. R. Ladd (Eds.), *Laboratory phonology II: Gesture, segment, prosody* (pp. 261-280). Cambridge: Cambridge University Press.

- Nolan, F., Holst, T., & Kuhnert, B. (1996). Modeling [s] to [ʃ] accommodation in English. *Journal of Phonetics*, 24, 113-137.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioural Brain Sciences*, 23(3), 299-325; Discussion : 325-270.
- Norris, D., McQueen, J. M., & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory & Cognition*, 30(3), 399-411.
- Ohala, J.J. (1983). The origin of sound patterns in vocal tract constraints. In P.F. MacNeilage (Eds.), *The production of speech*, (pp. 189-216). Berlin: Springer-Verlag.
- Ohde, R.N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America*, 75, 224-230.
- Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, 100(6), 3831-3842.
- O'Shaughnessy, D. (1981). A study of French vowel and consonant durations. *Journal of Phonetics*, 9, 385-406.
- Peterson, G.I., & Lehiste, I. (1960). Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America*, 32, 693-703.
- Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 57-282.
- Plunkett, K., & Elman, J. L. (1997). *Exercices in rethinking innateness. A handbook for connectionist simulations*. Cambridge, Massachusetts: The MIT Press.
- Quené, H., van Rossum, M., & van Wijck, M. (1998, December). *Assimilation and anticipation in word perception*. Paper presented at the Proceedings of the Fifth International Conference on Spoken Language Processing, Sydney, Australia.
- R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>.
- Raphael, L.J. (1972). Preceding vowel duration as a cue to voicing characteristics of word-final consonants in English. *Journal of the Acoustical Society of America*, 51, 1296-1303.
- Racine, I. & Grosjean, F. (soumis). Le coût de l'effacement du schwa lors de la reconnaissance des mots en français. *Canadian Journal of Experimental Psychology*.
- Radeau, M., Morais, J., & Segui, J. (1994). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1297-1311.
- Repp, B. (1982). Phonetic tranding relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92, 81-110.
- Rigault, A. (1967). *L'assimilation consonantique de sonorité en français : Etude acoustique et perceptuelle*. Paper presented at the Proceedings of the 6th International Congress of Phonetic Sciences, Prague.
- Roelofs, A. (2003). Modeling the relation between the production and recognition of spoken words forms. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and Phonology in Language Comprehension and Production* (pp. 115-158). Berlin: Mouton de Gruyter.

- Roodenrys, S., Hulme, C., Lethbridge, A., Hinton, M., & Nimmo, L. M. (2002). Word-frequency and phonological neighborhood effects on verbal short-term memory. *Journal of Experimental Psychology, Learning, Memory, & Cognition*, 28(6), 1019-1034.
- Saerens, M., Serniclaes, W., & Beeckmans, R. (1989). Acoustic versus contextual factors in stop voicing perception in spontaneous French. *Language and Speech*, 32, 291-314.
- Samuel, A.G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, 110 ; 474-494.
- Samuel, A. G. (2001). Some empirical tests of Merge's architecture. *Language and Cognitive Processes*, 16, 709-714.
- Saussure, F. de (1916). *Cours de linguistique générale*. Paris : Grande Bibliothèque Payot.
- Shinohara, S. (A paraître). Perceptual effects in final cluster reduction patterns. *Lingua*.
- Snoeren, N. D., Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (Manuscrit soumis pour publication). Organization of the mental lexicon during the earliest stages of foreign language acquisition.
- Snoeren, N. D., Hallé, P. A., & Segui, J. (sous presse). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics*.
- Solé, M.-J. (2002). Assimilatory processes and aerodynamic factors. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7* (pp. 351-386). Berlin: Mouton de Gruyter.
- Spinelli, E., & Alario, F.-X. (2002). Gender context effects on homophone words. *Language and Cognitive Processes*, 17(5), 457-469.
- Spinelli, E., McQueen, J. M., & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, 48(412-432).
- Steriade, D. (2000). Directional asymmetries in place assimilation: A perceptual account. In E. Hume & K. Johnson (Eds.), *Perception in phonology* (pp.219-250). Academic Press.
- Stevens, K.N., Keyser, S.M., Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 432-469).
- Streeter, L.A., & Nigro, G. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America*, 65(6), 1533-1541.
- Sumner, M. & Samuel, A.G. (sous presse). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language*.
- Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, 27, 324-340.
- Thomas, J. M.-C., Bouquiaux, L., & Cloarec-Heiss, F. (1976). *Initiation à la phonétique*. Presses Universitaires de France.
- Touchet, P. (2002). Merleau-Ponty : Le langage et la parole, <http://www.philosophies.ac-versailles.fr>.
- Tuller, B. Case, P. Ding, M. & Scott Kelso, J.A. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception and Performance*, 20(1), 3-16.
- Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America*, 58, 434-455.

- Utman, J. A., Blumstein, S. E., & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics*, 62(6), 1297-1311.
- Utman, J. A., Blumstein, S. E., & Sullivan, K. (2001). Mapping from sound to meaning: Reduced lexical activation in Broca's aphasia. *Brain and Language*, 79, 444-472.
- Van Alphen, P. M. (2004). *Perceptual relevance of prevoicing in Dutch*. Thèse de doctorat, Max Planck Institut for Psycholinguistics & Radboud Universiteit Nijmegen, Pays-Bas.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 28(4), 735-747.
- Vitevitch, M. S. (2002). Influence of onset density on spoken-word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, 28(2), 270-278.
- Wajskop, M. (1979). Segmental durations of French intervocalic plosives. In B. Lindblom & S. Ohman (Eds.), *Frontiers of speech communication research* (pp. 109-123). New York: Academic Press.
- Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251-276.
- Warren, R.M. (1970). Restoration of missing speech sounds. *Science*, 167, 392-393.
- Warren, P., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*, 41 (3), 262-275.
- Warren, P., & Marslen-Wilson, W. D. (1988). Cues to lexical choice: Discriminating place and voice. *Perception & Psychophysics*, 43, 21-30.
- Weber, A. (2001). Help or hindrance: How violation of different assimilation rules affects spoken-language processing. *Language & Speech*, 44, 95-118.
- Weber, A. (2002). Assimilation violation and spoken-language processing: A supplementary report. *Language and Speech*, 45 (2), 37-46.
- Wetzels, W.L., & Mascaró, J. (2001). The typology of voicing and devoicing. *Language*, 77, 207-244.
- Whalen, D. H.(1991a). Infrequent words are longer in duration than frequent words [Abstract]. *Journal of the Acoustical Society of America*, 90, 2311.
- Whalen, D.H. (1991b). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, 50(4), 351-360.
- Whalen, D.H. (1992). Further results on the duration of infrequent and frequent words [Abstract]. *Journal of the Acoustical Society of America*, 91, 2339-2340.
- Wright, S., & Kerswill, P. (1989). Electropalatography in the analysis of connected speech. *Clinical linguistics & Phonetics*, 3, 49-57.

ANNEXE 1. MATERIEL EXPERIMENTAL DES EXPERIENCES 1-3

Tableau 1. Expériences 1a-1b : phrases test contenant les mots stimuli.

<i>Contexte non assimilatoire</i>	<i>Contexte assimilatoire</i>
<i>Phrases contenant les MAA</i>	
1. J'ai écrit <i>bled</i> vraiment bien	1. J'ai écrit <i>bled</i> sans erreur
2. Ces <i>jades</i> ne sont pas bon marché	2. Ces <i>jades</i> sont vraiment bon marché
3. Ces <i>rides</i> n'ont rien d'extraordinaire	3. En fait, ces <i>rides</i> sont bien naturelles
4. La <i>soude</i> ne sent pas si mauvais que ça	4. La <i>soude</i> pue vraiment fort
5. J'ai écrit <i>vide</i> vraiment bien	5. J'ai écrit <i>vide</i> sans erreur
6. J'ai écrit <i>blatte</i> sans erreur	6. J'ai écrit <i>blatte</i> vraiment bien
7. Ces <i>jattes</i> sont vraiment bon marché	7. Ces <i>jattes</i> ne sont pas bon marché
8. Ces <i>rites</i> sont bien naturels	8. Ces <i>rites</i> n'ont rien d'extraordinaire
9. La <i>soute</i> pue vraiment fort	9. La <i>soute</i> ne sent pas si mauvais
10. J'ai écrit <i>vite</i> sans erreur	10. J'ai écrit <i>vite</i> vraiment bien
<i>Phrases contenant les MNA</i>	
1. On a mangé des <i>frites</i> croustillantes	1. Les <i>frites</i> de Bruxelles sont les meilleures
2. Il a attrapé une <i>grippe</i> contagieuse	2. Il a attrapé la <i>grippe</i> de son frère
3. Il a pris des <i>gouttes</i> pour le nez	3. Il a vu une <i>goutte</i> de sang sur le lit
4. Les <i>jupes</i> courtes sont à la mode	4. Les <i>jupes</i> droites lui vont bien
5. J'ai trouvé la <i>note</i> plutôt salée	5. Chantez-moi toutes les <i>notes</i> de la gamme
6. C'est un <i>stock</i> très difficile à écouler	6. Il a épuisé son <i>stock</i> de cigarettes
7. Le comédien a un <i>trac</i> fou	7. Il a le <i>trac</i> des débutants
8. Lisez bien le <i>mode</i> d'emploi	8. Ce manteau est très à la <i>mode</i> ³⁰ cet hiver
9. Il a fait beau dans le <i>sud</i> de la France	9. Je n'ai jamais vu un vent du <i>sud</i> si fort
10. J'ai oublié mon <i>tube</i> de dentifrice	10. Ce <i>tube</i> passe à la radio sans arrêt

³⁰Le mot *mode* peut être prononcé comme le mot *motte* [mot] après assimilation, mais nous avons compté ce mot comme un mot non ambigu étant donné la différence de fréquence d'usage entre *mode* et *motte*, (65.81 et 2.74 respectivement, d'après la base de données *Lexique*).

Tableau 2. Expérience 2 : mots stimuli et contextes phrastiques

<i>Phonèmes suivis par des occlusives /ptkbdg/</i>		<i>Phonèmes suivis par des fricatives /fsvʒ/</i>	
[-voisé]	[+voisé]	[-voisé]	[+voisé]
1. Une <i>bllette</i> grillée	1. Un <i>bled</i> pourri	1. La <i>bllette</i> verte	1. Un <i>bled</i> souriant
2. Des <i>jattes</i> gauloises	2. Les <i>rides</i> profondes	2. Des <i>jattes</i> vernissées	2. Les <i>rides</i> saillantes
3. Des <i>rites</i> bizarres	3. Les <i>jades</i> précieux	3. Les <i>rites</i> japonais	3. Les <i>jades</i> superbes
4. la <i>soute</i> britannique	4. La <i>soude</i> puante	4. La <i>soute</i> vide	4. La <i>soude</i> frémissante
5. C'est <i>vite</i> balayé	5. Un <i>vide</i> critique	5. C'est <i>vite</i> jugé	5. Un <i>vide</i> surprenant
6. Les <i>frites</i> belges	6. La <i>mode</i> passagère	6. Des <i>frites</i> jaunes	6. La <i>mode</i> séduisante
7. Une <i>grippe</i> dangereuse	7. Le <i>sud</i> pauvre	7. Un <i>grippe</i> virale	7. Le <i>sud</i> frileux
8. La <i>Note</i> bleue	8. Un <i>tube</i> terrifiant	8. Une <i>note</i> juridique	8. Un <i>tube</i> silencieux
9. Des <i>jupes</i> droites	9. Un <i>crabe</i> croustillant	9. Des <i>jupes</i> variées	9. Un <i>crabe</i> farci
10. Un <i>trac</i> dingue	10. Une <i>drogue</i> populaire	10. Un <i>truc</i> gênant	10. Une <i>drogue</i> stimulante

Tableau 3a. Expérience 3 : mots stimuli ambigus par assimilation. Sont indiqués les jugements subjectifs, la fréquence d'usage et les champs informatifs sur le voisinage (*voty* = nombre total de voisins par type, *vosty* = nombre de voisins par substitution (par type), *vosty* 3/4 = nombre de voisins par substitution de la (troisième ou quatrième) consonne finale, cf. base de données *Vocolex*).

<i>Item</i>	<i>Jugement subj.</i>	<i>Fréquence</i>	<i>Voty</i>	<i>Vosty</i>	<i>Vosty 3/4</i>
Black	4.31	9.29	*	*	*
Blague	4.69	3.84	11	6	3
Trombe	2.44	17.52	7	3	3
Trompe	3.63	16.10	16	9	3
Grade	3.13	7.48	15	10	3
Gratte	3.44	4.35	18	8	7
Soude	2.31	0.77	19	14	7
Soute	2.88	3.13	29	24	7
Jade	2.31	1.29	17	16	8
Jatte	1.93	16.87	28	26	8
Bec	3.75	0.61	28	22	10
Bègue	2.19	2.45	18	16	10
Ride	4	10.68	30	21	10
Rite	3.13	0.39	34	25	10
Daube	3.56	1.06	29	22	11
Dope	2.75	9.55	26	21	11
Bac	4.38	9.10	38	31	14
Bague	4.13	1	28	23	14
Bock	1.25	0.19	36	29	14
Bogue	1.63	31.16	26	23	14
Code	4.44	31.16	31	23	14
Cote	3	4.45	38	34	14
Rade	1.88	5.29	40	29	15
Rate	2.81	3.03	43	32	15

* La base de données *Vocolex* ne contient pas le mot *black*.

Tableau 3b. Expérience 3 : Mots stimuli non ambigus

<i>Item</i>	<i>Fréquence</i>	<i>Voty</i>	<i>Vosty</i>	<i>Vosty 3/4</i>
Vogue	8.52	14	8	2
Fraude	4.42	11	6	3
Steppe	5.16	9	7	4
Floc	1.35	18	12	4
Glèbe	0.97	8	7	4
Jupe	18.13	13	10	5
Crotte	2.16	24	14	6
Bride	6.29	22	12	7
Lac	23.55	31	23	9
Coupe	46.23	27	18	10
Digue	5.1	23	20	10
Tube	23.52	19	15	10
Bouc	5.74	35	24	11
Fac	2.26	31	26	11
Mob	1.03	21	19	11
Soc	1.39	32	27	12
Pape	17.16	37	29	14
Botte	5.55	41	32	14
Coq	10.81	39	29	14
Gag	0.68	23	21	14
Bob	7.29	27	24	14
Rabe	0.16	31	23	15
Bide	4.58	29	24	16
Cab	0.16	31	25	17

ANNEXE 2. PRE-TEST JUGEMENTS SUBJECTIFS (EXPERIENCE 3)

Chers participants,

Vous trouvez ci-dessous une liste de mots français. Veuillez indiquer, selon vous, la fréquence du mot sur une échelle de 1 à 5 (1 = mot rare, 5 = mot très fréquent) **en encerclant un chiffre.**

Lemme	1	2	3	4	5
Match	1	2	3	4	5
Miche	1	2	3	4	5
Mouette	1	2	3	4	5
Pale	1	2	3	4	5
Phare	1	2	3	4	5
Poule	1	2	3	4	5
Seau	1	2	3	4	5
Soie	1	2	3	4	5
Tronc	1	2	3	4	5
Soude	1	2	3	4	5
Baume	1	2	3	4	5
Bogue	1	2	3	4	5

...

Merci pour votre participation

ANNEXE 3. CONSIGNE PRODUCTION DE LA PAROLE (EXPERIENCES 2-3A)

Chers participants,

Nous vous présentons un certain nombre d'énoncés en français sur plusieurs feuilles. Nous vous demandons de lire ces énoncés à haute voix. Il est très important de lire ces phrases de manière claire et fluide. Afin de vous familiariser avec cette procédure, vous allez d'abord lire 15 énoncés qui ne font pas partie de l'expérience proprement dite.

Merci pour votre participation.

ANNEXE 4. CONSIGNE POUR LA CATEGORISATION PHONETIQUE (EXPERIENCE 1B)

L'expérience est constituée de deux courtes phases d'entraînement et d'une phase de test proprement dit. Chaque phase consiste en une série d'essais. A chaque essai, vous entendrez un « mot parlé ». Dans la phrase « test », deux réponses vous sont demandées à chaque essai :

1. La consonne du français qui vous semble la plus proche de la consonne finale de l'item que vous venez d'entendre.
2. Une note subjective indiquant si la correspondance entre le son entendu et votre réponse est de très bonne (5) à très médiocre (1).

Au cours de chaque essai, il vous est possible de rejouer l'item qui vous est présenté.

Dans la première phase d'entraînement, on ne vous demande que la réponse 1, et on vous donne la « bonne » réponse. La seconde phase d'entraînement est simplement destinée à vous familiariser avec la façon de répondre dans la phase de test (il n'y a pas de « retour »).

Merci pour votre participation.

ANNEXE 5. CONSIGNE CATEGORISATION DES PHONEMES (EXPERIENCE 3B-4)

Cette expérience est constituée d'une phase d'entraînement et de trois phases test. Chaque phase consiste en une série d'essais et à chaque essai, vous entendrez un item. Deux réponses vous sont demandées à chaque essai :

1. La consonne du français qui vous semble la plus proche de **la consonne finale** de l'item que vous venez d'entendre.
2. Une **note subjective** indiquant la correspondance entre le son entendu et votre réponse. Cette note va de **5** (très bonne) à **1** (très médiocre).

Vous effectuez votre réponse en cliquant avec la souris sur (1) la consonne présentée sur l'écran et en cliquant sur (2) un des chiffres en dessous.

Merci pour votre participation.

ANNEXE 6. MATERIEL EXPERIMENTAL DES EXPERIENCES 4-6

Tableau 4a. Expériences 4-6 : mots stimuli (en italiques) et leurs contextes phrastiques.

<i>Canonique</i>	<i>Assimilé</i>	<i>Contrôle</i>
<i>Plosives finales : /p/</i>		
1. Une <i>coupe</i> transversale	Une <i>coupe</i> droite	Un <i>acte</i> final
2. Une <i>grippe</i> contagieuse	Une <i>grippe</i> durable	Un <i>orgue</i> portatif
3. Une <i>jupe</i> serrée	Une <i>jupe</i> grise	Un <i>blanc</i> laiteux
4. Une <i>nappe</i> tâchée	Une <i>nappe</i> déchirée	Une <i>phase</i> critique
5. Le <i>pape</i> triste	Le <i>pape</i> débonnaire	Une <i>aire</i> protégée
6. Une <i>soupe</i> corse	La <i>soupe</i> délicieuse	La <i>endre</i> volcanique
7. Une <i>troupe</i> comique	Une <i>troupe</i> gaie	Une <i>anse</i> métallique
8. Un <i>type</i> sensé	Un <i>type</i> galant	Une <i>boîte</i> noire
9. Une <i>lampe</i> cassée	Une <i>lampe</i> de cheveux	Une <i>cible</i> monumentale
10. Une <i>pompe</i> tordue	Une <i>pompe</i> grinçante	Une <i>firme</i> japonaise
11. Une <i>rampe</i> tympanique	Une <i>rampe</i> glissante	Un <i>jour</i> férié
12. Un <i>groupe</i> solidaire	Un <i>groupe</i> difficile	Un <i>moine</i> réfugié
<i>Plosives finales : /t/</i>		
1. Des <i>bottes</i> confortables	Des <i>bottes</i> brillantes	Un <i>nœud</i> double
2. Une <i>brute</i> sanguinaire	Une <i>brute</i> violente	Une <i>gamme</i> complète
3. Une <i>chute</i> chaotique	Une <i>chute</i> brutale	Un <i>œuf</i> cuit
4. Un <i>doute</i> persistant	Un <i>doute</i> grandissant	Une <i>Pierre</i> cassée
5. Une <i>faute</i> prévisible	Une <i>faute</i> grossière	Une <i>pluie</i> diluvienne
6. Une <i>note</i> salée	Une <i>note</i> grave	Le <i>ventre</i> plein
7. La <i>route</i> perdue	La <i>route</i> goudronnée	Un <i>angle</i> différent
8. Un <i>vote</i> secret	Un <i>vote</i> blanc	Une <i>bague</i> empruntée
9. La <i>datte</i> séchée	La <i>datte</i> garnie	Une <i>base</i> militaire
10. Une <i>grotte</i> préhistorique	Une <i>grotte</i> blanche	Le <i>beurre</i> naturel
11. La <i>lutte</i> continue	La <i>lutte</i> brutale	Une <i>blouse</i> blanche
12. Des <i>gouttes</i> scintillantes	Des <i>gouttes</i> brûlantes	Le <i>but</i> principal

Tableau 4b. Expériences 4-6 : suite matériel expérimental.

<i>Canonique</i>	<i>Assimilé</i>	<i>Contrôle</i>
<i>Plosives finales : /k/</i>		
1. La <i>banque</i> populaire	La <i>banque</i> d'Algérie	Un <i>couple</i> heureux
2. Un <i>bloc</i> plastifié	Un <i>bloc</i> défectueux	Un <i>drap</i> humide
3. Des <i>briques</i> posées	Des <i>briques</i> déstabilisées	Un <i>casque</i> protecteur
4. Un <i>choc</i> terrifiant	Un <i>choc</i> brutal	Une <i>gloire</i> éphémère
5. Un <i>grec</i> patriotique	Un <i>grec</i> drôle	Un <i>cuir</i> souple
6. Un <i>lac</i> pollué	Un <i>lac</i> desséché	La <i>clef</i> verte
7. La <i>nuque</i> tendue	La <i>nuque</i> dégagée	Un <i>cadre</i> familial
8. Une <i>plaque</i> tordue	Une <i>plaque</i> découpée	Un <i>centre</i> culturel
9. Un <i>sac</i> troué	Un <i>sac</i> démesuré	Une <i>canne</i> jaune
10. Un <i>truc</i> particulier	Un <i>truc</i> débile	Un <i>dieu</i> omniscient
11. Un <i>flic</i> pointilleux	Un <i>flic</i> décidé	Une <i>dose</i> forte
12. Des <i>claques</i> sonores	Des <i>claques</i> violentes	Un <i>masque</i> facial
<i>Plosives finales : /b/</i>		
1. Une <i>bombe</i> destructrice	Une <i>bombe</i> terrifiante	Un <i>loup</i> domestiqué
2. Un <i>club</i> gastronomique	Un <i>club</i> touristique	Un <i>vase</i> simple
3. Un <i>globe</i> doré	Un <i>globe</i> terrestre	La <i>dette</i> nationale
4. Une <i>jambe</i> galbée	Une <i>jambe</i> cassée	Un <i>titre</i> national
5. Une <i>robe</i> droite	Une <i>robe</i> serrée	Une <i>branche</i> professionnelle
6. Une <i>tombe</i> grandiose	Une <i>tombe</i> somptueuse	Une <i>corse</i> sociale
7. Le <i>tube</i> digestif	Le <i>tube</i> cathodique	La <i>ferme</i> conservatoire
8. L' <i>aube</i> glacée	L' <i>aube</i> colorée	Une <i>fiche</i> personnelle
9. Des <i>bribes</i> disséminées	Des <i>bribes</i> significantes	Un <i>poste</i> permanent
10. Un <i>crabe</i> délicieux	Un <i>crabe</i> farci	Une <i>poche</i> pleine
11. Un <i>cube</i> dense	Un <i>cube</i> saillant	Une <i>plaie</i> profonde
12. Un <i>snob</i> agréable	Un <i>snob</i> silencieux	Le <i>linge</i> sale

Tableau 4c. Expériences 4-6 : suite matériel expérimental.

<i>Canonique</i>	<i>Assimilé</i>	<i>Contrôle</i>
<i>Plosives finales : /d/</i>		
1. Une <i>aide</i> bancaire	Une <i>aide</i> sociale	Une <i>feuille</i> quadrillée
2. La <i>bande</i> magnétique	La <i>bande</i> passante	Le <i>prince</i> charmant
3. La <i>blonde</i> belge	La <i>blonde</i> suédoise	Un <i>arc</i> traditionnel
4. Le <i>coude</i> blessé	Le <i>coude</i> plié	Un <i>gosse</i> gâté
5. Le <i>guide</i> breton	Le <i>guide</i> prévoyant	La <i>cloche</i> royale
6. La <i>mode</i> britannique	La <i>mode</i> parisienne	Une <i>source</i> disparue
7. Le <i>stade</i> bruyant	Le <i>stade</i> complet	Une <i>panne</i> majeure
8. Une <i>viande</i> braisée	Une <i>viande</i> saignante	Une <i>poudre</i> suspecte
9. Une <i>bride</i> mauve	Une <i>bride</i> soudée	Une <i>gare</i> sympathique
10. La <i>dinde</i> gratinée	La <i>dinde</i> savoureuse	Une <i>zone</i> industrielle
11. Les <i>soldes</i> budgétaires	Les <i>soldes</i> précédentes	Un <i>pacte</i> secret
12. La <i>sonde</i> gastrique	La <i>sonde</i> perdue	Une <i>marge</i> supérieure
<i>Plosives finales /g/</i>		
1. La <i>langue</i> basque	La <i>langue</i> pendue	La <i>face</i> cachée
2. Les <i>seringues</i> doseuses	Les <i>seringues</i> trouées	Les <i>armes</i> chimiques
3. Une <i>figue</i> délicieuse	Une <i>figue</i> sucrée	La <i>reine</i> norvégienne
4. Un <i>gang</i> dangereux	Un <i>gang</i> terrifiant	Un <i>culte</i> privé
5. La <i>ligue</i> dissociée	La <i>ligue</i> portugaise	Le <i>sable</i> rouge
6. Les <i>digues</i> basses	Les <i>digues</i> submersibles	Les <i>cerises</i> mûres
7. Un <i>dingue</i> bruyant	Un <i>dingue</i> paumé	Une <i>bosse</i> douloureuse
8. La <i>drogue</i> brute	La <i>drogue</i> parfaite	Le <i>verbe</i> conjugué
9. La <i>vogue</i> branchée	La <i>vogue</i> française	Le <i>peuple</i> migrateur
10. Les <i>fringues</i> bizarres	Les <i>fringues</i> sportives	Une <i>marche</i> funèbre
11. Les <i>fugues</i> d'adolescents	Les <i>fugues</i> proposées	Le <i>miel</i> contaminé
12. La <i>fougue</i> disciplinée	La <i>fougue</i> passionnée	La <i>housse</i> moulante

ANNEXE 7. CONSIGNE DECISION LEXICALE AVEC AMORÇAGE DE REPETITION (EXPERIENCES 5 ET 6)

A chaque essai, vous allez d'abord entendre un item puis voir apparaître sur l'écran une suite de lettres. La phase d'entraînement qui consiste en 12 essais, n'est destinée qu'à vous familiariser avec la tâche et elle ne sera pas prise en compte. Après cette phase, l'expérience proprement dite va commencer et à partir de ce moment, l'ordinateur enregistrera vos réponses.

Votre tâche consiste à dire **le plus rapidement possible mais sans faire d'erreur**, si la suite de lettres qui apparaît à l'écran est un mot de la langue française ou un pseudo-mot (exemple : tabe). S'il s'agit d'un mot, vous répondez OUI en appuyant avec votre index droit sur le bouton droit. Si c'est un pseudo-mot, vous appuierez avec votre index gauche sur le bouton de gauche pour répondre NON.

Votre tâche consiste non seulement à répondre aux suites de lettres présentées à l'écran mais également à **écouter les items auditifs qui précèdent les items présentés sur l'écran** car une épreuve de reconnaissance de ces items vous sera présentée à la fin de l'expérience.

Questions ? Résumons :

- 1/ Ecoutez attentivement les items présentés auditivement.
- 2/ Fixez l'écran tout au long de l'expérience et laissez vos index sur les deux boutons de réponse, sans appuyer.
- 3/ Si la suite de lettres présentées à l'écran constitue un mot, appuyez le plus vite possible sur le bouton de droite.
- 4/ Si la suite de lettres présentées à l'écran n'est pas un mot, appuyez le plus vite possible sur le bouton de gauche.
- 5/ Enfin, si vous vous rendez compte que vous avez fait une erreur, ne vous en préoccupez pas, mais continuez à répondre le plus vite possible à l'item suivant.

ANNEXE 8 □ POST-TEST DE RECONNAISSANCE (EXPERIENCES 5-6)

EPREUVE DE RECONNAISSANCE

Voici une suite de mots dont certains vous ont été présentés auditivement au cours de l'expérience. Entourez O si vous reconnaissez le mot et N si vous pensez qu'il ne vous a jamais été présenté.

1. coupe	O	N
2. table	O	N
3. chat	O	N
4. nappe	O	N
5. chaise	O	N
6. soupe	O	N
7. écran	O	N
8. banque	O	N
9. perle	O	N
10. livre	O	N
11. tête	O	N
12. fugue	O	N
13. bouton	O	N
14. style	O	N
15. membre	O	N
16. montre	O	N
17. doigt	O	N
18. gouttes	O	N
19. lutte	O	N
20. bras	O	N
21. œil	O	N
22. gang	O	N
23. trou	O	N
24. feuille	O	N

25. carte	O	N
26. cube	O	N
27. crabe	O	N
28. truc	O	N
29. claques	O	N
30. box	O	N

ANNEXE 9. PRE-TEST D'ASSOCIATION LIBRE (EXPERIENCES 7 ET 8)

Ci-dessous vous trouvez une liste avec des mots français.

1/ Veuillez écrire en face de chaque mot le premier mot qui vous vient à l'esprit.

2/ Ne donnez qu'une seule réponse.

3/ Suivez l'ordre des mots.

4/ Ne sautez aucun mot.

5/ Allez le plus vite possible.

6/ Ecrivez lisiblement.

7/ Restez concentré et silencieux.

Exemple Chien : chat.

nappe :

cartable :

rose :

note :

bac :

faute :

...

Indiquez ici votre sexe :

Indiquez ici votre âge :

MERCI POUR VOTRE PARTICIPATION !

ANNEXE 10. MATERIEL EXPERIMENTAL (EXPERIENCES 7 ET 8)

Tableau 6. Expérience 7 : mots stimuli et leurs contextes phrastiques.

<i>Contexte non assimilatoire (formes canoniques)</i>	<i>Contexte assimilatoire (formes assimilées)</i>
<i>Phrases contenant les MAA</i>	
1. Un <i>black</i> policier	1. Un <i>black</i> bourgeois
2. Une <i>trompe</i> tendue	2. Une <i>trompe</i> d'éléphant
3. Une <i>gratte</i> cassée	3. Une <i>gratte</i> blues
4. Une <i>soute</i> pleine	4. Une <i>soute</i> bondée
5. Une <i>jatte</i> poreuse	5. Une <i>jatte</i> beurrée
6. Un <i>bec</i> tordu	6. Un <i>bec</i> droit
7. Un <i>rite</i> populaire	7. Un <i>rite</i> barbare
8. La <i>dope</i> coupée	8. la <i>dope</i> danoise
9. Le <i>bock</i> profond	9. Le <i>bock</i> bavarois
10. Une <i>cote</i> prometteuse	10. Une <i>cote</i> boursière
11. Un <i>bac</i> professionnel	11. Un <i>bac</i> difficile
12. Une <i>rate</i> cancéreuse	12. Une <i>rate</i> bileuse
<i>Phrases contenant les MNA</i>	
1. Une <i>coupe</i> teintée	1. Une <i>coupe</i> dorée
2. Une <i>jupe</i> courte	2. Une <i>jupe</i> grise
3. Une <i>nappe</i> carrée	3. Une <i>nappe</i> décolorée
4. Une <i>troupe</i> théâtrale	4. Une <i>troupe</i> grotesque
5. Une <i>grotte</i> profonde	5. Une <i>grotte</i> bretonne
6. Une <i>lutte</i> permanente	6. Une <i>lutte</i> guerrière
7. Une <i>note</i> pertinente	7. Une <i>note</i> grave
8. Une <i>faute</i> criminelle	8. Une <i>faute</i> grave
9. La <i>nuque</i> tondue	9. La <i>nuque</i> dégagée
10. Un <i>choc</i> thermique	10. Un <i>choc</i> brutal
11. Une <i>brique</i> tombée	11. Une <i>brique</i> brisée
12. La <i>plaque</i> tectonique	12. La <i>plaque</i> dentaire

Tableau 7. Expérience 7 : paires amorce-cible.

<i>Amorce</i>	<i>Cible associée</i>	<i>Jugement d'association (%)</i>
1. black	NOIR	40
2. trompe	ELEPHANT	69
3. gratte	GUITARE	40
4. bock	BIERE	28
5. jatte *	LAIT	
6. bec	OISEAU	64
7. soute	BAGAGE	47
8. bac	EXAMEN	36
9. rate	FOIE	19
10. rite	COUTUME	16
11. dope	DROGUE	69
12. cote *	CHIFFRE	
1. coupe	CHEVEUX	31
2. jupe	ROBE	36
3. nappe	TABLE	83
4. troupe	THEATRE	21
5. grotte	CAVERNE	24
6. lutte	COMBAT	28
7. brique	MAISON	31
8. plaque *	EGOUT	
9. faute	ERREUR	55
10. note	MUSIQUE	19
11. nuque	COU	59
12. choc	ACCIDENT	16

* Les jugements d'associations du pré-test étant trop dispersés pour ces mots-là, nous avons attribué les associations en demandant à dix étudiants, n'ayant pas participé au pré-test, de faire une association libre sur ces trois mots. Les réponses unanimes étaient retenues.

Tableau 8. Expérience 8 : équivalents canoniques des formes assimilées (cf. Exp. 7).

<i>Mots sources (cf. Exp. 7)</i>	<i>Contexte non assimilatoire (formes canoniques)</i>
1. black	1. Une <i>blague</i> belge
2. trompe	2. Une <i>trombe</i> d'eau
3. gratte	3. Un <i>grade</i> militaire
4. soute	4. Une <i>soude</i> brute
5. jatte	5. Une <i>jade</i> magique
6. bec	6. Un <i>bègue</i> gêné
7. rite	7. Une <i>ride</i> volcanique
8. dope	8. Une <i>daube</i> monumentale
9 bac	9. Une <i>bague</i> métallique
10. boc	10. Une <i>bogue</i> verte
11. cote	11. Un <i>code</i> valide
12. rate	12. Un <i>rade</i> branchouille

Tableau 9a. Expérience 8 : mots de remplissage, paires amorce cible reliées.

<i>Amorce</i>	<i>Cible associée</i>	<i>Jugement d'association (%)*</i>
1. beurre	TARTINE	20
2. chat	CHIEN	30
3. lettre	POSTE	19
4. nœud	PAPILLON	27
5. règle	MESURE	21
6. singe	GRIMACE	12
7. taupe	TERRE	23
8. mouche	INSECTE	12
9. évier	CUISINE	16
10. crabe	MER	25
11. glace	VANILLE	17
12. flèche	ARC	53
13. montre	HEURE	57
14. pince	LINGE	38
15. phare	PORT	17
16. luge	NEIGE	74
17. ours	PELUCHE	26
18. malle	VALISE	17
19. cloche	EGLISE	37
20. coffre	BIJOU	90
21. cintre	ARMOIRE	90
22. puits	EAU	53
23. carte	JEU	90
24. balai	BROSSE	27

* D'après les normes d'association de Ferrand et Alario (1998).

Tableau 9b. Expérience 8 : mots de remplissage, paires amorce-cible non reliées

Amorce	Cible
1. larve	LOISIR
2. mouette	ESSAI
3. gant	ECRAN
4. jouet	FRAICHEUR
5. craie	CANON
6. gain	STATUT
7. zinc	CLIMAT
8. globe	AMANT
9. pic	APPUI
10. croupe	DOSSIER
11. hôte	VIRUS
12. bond	ACCENT
13. dette	PHOTO
14. site	HEROS
15. bougre	PARFUM
16. orgue	CHEMIN
17. meute	ABRI
18. dogme	PAQUET
19. tigre	PAPIER
20. cuisse	FAMILLE
21. deuil	ORGANE
22. four	ESPRIT
23. ongle	CELLULE
24. fable	RIGUEUR

ANNEXE 11. CONSIGNE DECISION LEXICALE AVEC AMORÇAGE SEMANTIQUE (EXPERIENCES 7 ET 8)

A chaque essai, vous allez d'abord entendre un mot puis voir apparaître sur l'écran une suite de lettres. La phase d'entraînement qui consiste en 10 essais, n'est destinée qu'à vous familiariser avec la tâche et les résultats ne seront pas prise en compte. Après cette phase, l'expérience proprement dite va commencer et à partir de ce moment, l'ordinateur enregistrera vos réponses.

Votre tâche consiste à décider **le plus rapidement possible mais sans faire d'erreur**, si la suite de lettres qui apparaît à l'écran est un mot de la langue française ou un pseudo-mot (exemple : tabe). S'il s'agit d'un mot, vous répondez OUI en appuyant avec votre index droit sur le bouton droit. Si c'est un pseudo-mot, vous appuierez avec votre index gauche sur le bouton de gauche pour répondre NON.

Votre tâche consiste non seulement à répondre aux suites de lettres présentées à l'écran mais également à **écouter les items auditifs qui précèdent les items présentés sur l'écran.**

Résumons :

- 1/ Ecoutez attentivement les items présentés auditivement.
- 2/ Fixez l'écran tout au long de l'expérience et laissez vos index sur les deux boutons de réponse ;
- 3/ Si la suite de lettres présentées à l'écran constitue un mot, appuyez le plus vite possible sur le bouton de droite ;
- 4/ Si la suite de lettres présentées à l'écran n'est pas un mot, appuyez le plus vite possible sur le bouton de gauche.

ANNEXE 12. QUESTIONNAIRE SUR LES CONNAISSANCES LINGUISTIQUES

QUESTIONNAIRE

Date _____

Sujet Numéro _____

1. Age _____

2. Année universitaire : 1^{ère} 2^e 3^e 4^e +

3. Quel diplôme préparez-vous ? _____

4. Est-ce que vous entendez parfaitement des deux oreilles ? (O/N) _____

Si non, quelles difficultés avez-vous remarquées ? _____

5. Si vous ou un de vos proches a jamais eu des problèmes de langage (tendance à bégayer, difficulté à trouver des mots, etc.) ou de lecture (dyslexie, etc.), indiquez s'il vous plaît :

La personne affectée : _____

La nature du problème : _____

Cette personne a-t-elle reçu un traitement spécial ? (O/N) _____

6. Quelle est votre langue maternelle ? _____

Quelle est la langue maternelle de votre mère ? _____

Quelle est la langue maternelle de votre père ? _____

7. Pour toutes les langues que vous parlez en dehors du français, indiquez s'il vous plaît :

<i>Langue</i>	<i>L'âge auquel vous avez commencé à l'étudier</i>	<i>Indiquez le niveau de vos compétences en cerclant un numéro ci-dessous où 1 = un peu et 5 = très bien</i>			
		<i>Parlé</i>	<i>Compris</i>	<i>Lu</i>	<i>Écrit</i>
_____	_____	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
_____	_____	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
_____	_____	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
_____	_____	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

8. Dans quelle(s) région(s) avez-vous habité ? Pouvez-vous en indiquer une liste ?

Région

Le nombre de mois ou d'années que vous y avez passés

9. Dans quelle région votre mère a-t-elle grandi ? _____

Dans quelle région votre père a-t-il grandi ? _____

10. En grandissant, avez-vous eu beaucoup de contact avec quelqu'un qui parlait une

langue étrangère ? (O/N) _____

Si oui, quelle était cette langue ? _____

11. Etes-vous gaucher ou droitier ? (D/G) _____

Merci pour votre participation

ANNEXE 13. ANALYSE SUPPLEMENTAIRE DE REGRESSION MULTIPLE (LME)

L'objectif principal des expériences 5 et 6, rapportées dans le chapitre 3 était d'examiner le traitement perceptif des segments plus ou moins fortement assimilés. Nos études phonétiques préalables (cf. chapitre 2) avaient d'ores et déjà révélé que les segments non voisés (e.g., *jupe* → [Zyb]) sont plus complètement assimilés que les segments voisés ne se dévoient (e.g., *robe* → [rop]). Ainsi, nous avons opérationnalisé le degré d'assimilation en termes de voisement sous-jacent, d'où l'attribution du facteur Voisement V_2 jacent à deux modalités (V_1 = obstruente non voisée ; V_2 = obstruente voisée). Cependant, si l'asymétrie entre voisés et non voisés s'avère être un effet phonétique assez robuste (tout au moins en français), il n'en reste pas moins que cette factorisation, exigée dans la méthodologie des analyses de variance, est une simplification de la réalité. Notre étude sur la production des l'assimilation en est une illustration, car nous avons par ailleurs démontré que l'assimilation de voisement en français est souvent un processus phonétique gradué plutôt que dichotomique. Quand nous examinons en détail les analyses acoustiques de la parole assimilée, nous nous apercevons que certaines obstruents non voisées ne subissent pas toujours une assimilation complète, alors que certaines obstruents voisées se dévoient complètement. C'est pour cette raison que nous avons alors décidé d'effectuer une analyse statistique de régression multiple (Baayen, Tweedie, & Schreuder, 2002 ; Baayen, 2004) en collaboration avec le Prof. R. Harald Baayen (*Interfaculty Research Unit for Speech*, Radboud Universiteit Nijmegen, et *Max Planck Institute for Psycholinguistics*, Nijmegen, Pays-Bas). Cette analyse avait pour but majeur d'examiner le rôle précis du degré d'assimilation sur les temps de réponse, sans pour autant avoir recours à une estimation approximative de l'assimilation par la factorisation du voisement sous-jacent.

Nous avons effectué des analyses acoustiques (cf. chapitre 2 pour plus de détails) sur les items prononcés dans les contextes assimilatoires (e.g., *jupe grise*). Le continuum des degrés d'assimilation était pris en compte dans l'analyse de régression³¹.

³¹ En moyenne, le profil asymétrique entre non voisés et voisés était de nouveau observé : le degré moyen d'assimilation était de 95 % pour les items non voisés (/p,t,k/) alors qu'il était de 53 % pour les items voisés (/b,d,g/).

Les analyses rapportées ici sont effectuées avec les modèles Effets Linéaires Mixtes (désormais LME, d'après *Linear Mixed Effects Models*, Pinheiro & Bates, 2000 ; Bates, 2005)³². Traditionnellement, une analyse de variance a pour but d'évaluer l'influence des variables indépendantes sur le comportement du sujet, généralement mesuré en termes de temps de réponses, ou en taux d'erreurs. Ce genre d'analyses se fait toujours en deux parties ; une analyse par sujets (F_1) et par items (F_2). Cependant, comme l'ont noté Baayen, Tweedie et Schreuder (2002), le processus de moyennage dans ces analyses a pour inconvénient que cela peut résulter en une perte considérable de données potentiellement riches en information. En utilisant LME, il est possible d'étudier les effets des variables indépendantes sur le comportement général d'une population, tout examinant la variabilité de chaque participant, sans pour autant recourir à des processus coûteux de moyennage sur les temps de réponse des participants. De plus, LME donne plus d'informations concernant les valeurs des écart-types des effets aléatoires (i.e. bruit), fournit d'emblée des précisions sur les participants « lents » ou « rapides », et des interactions possibles entre les prédicteurs manipulés (Baayen, Tweedie, & Schreuder, 2002 ; Baayen, communication personnelle).

Le fichier de données qui a été soumis à l'analyse LME comportait les temps de réponse de 116 participants (i.e. deux groupes de 58 participants des expériences 5 et 6) pour 72 mots stimuli. De plus, les prédicteurs suivants étaient pris en compte dans le modèle : la fréquence d'usage (en valeur log, « Freq »), le degré d'assimilation (en %, « Assim ») ; le nombre de voisins plus fréquents que la cible (« NVD ») ; la présence ou absence du Contexte (« Cont »), et type de consonne finale (/p,t,k/ vs. /b,d,g/, ConsF)³³.

Les résultats ont montré que « degré d'assimilation » en tant que tel n'a pas influencé les TRs, $F(1, 2916) = 0.29, p = 0.63$, alors qu'une interaction significative était observée entre « degré d'assimilation » et « contexte », $F(1, 2916) = 10.51, p < .01$. Ce résultat confirme donc bien

³² L'analyse requiert l'installation de l'application « lme4 » (*Linear Mixed Effects Models using S4 classes*), sous le logiciel R (*La Fondation R pour le Calcul Statistique*, 2005, Version 1.11 pour Mac OS Aqua GUI, voir les pages web de R pour plus de détails, <http://www.R-project.org>).

³³ Dans des analyses préliminaires, nous avons pris en compte d'autres prédicteurs acoustiques tels que la durée de la voyelle précédant le segment assimilé, l'intensité de la voyelle, et le nombre de voisins phonologiques. Ces prédicteurs n'ont pas influencé les temps de réponses

nos résultats préalables, dans le sens où la sensibilité contextuelle (i.e. la présence ou absence du contexte droit) varie en fonction du degré d'assimilation.

Tableau 10. Tableau ANOVA séquentiel du LME.

	<i>Df</i>	<i>Sum Sq</i>	<i>Mean Sc</i>	<i>Denom</i>	<i>F-value</i>	<i>Pr (>F)</i>
Freq	1	1.27	1.27	2619	36.53	1.72e-09***
Assim	1	0.01	0.01	2619	0.23	0.63
NVD	1	0.06	0.06	2619	1.70	0.19
ConsF	1	0.20	0.20	2619	5.65	0.017*
Cont	1	0.10	0.10	2619	2.96	0.086
Assim:ConsF	1	0.12	0.12	2619	3.47	0.063
NVD:ConsF	1	0.17	0.17	2619	4.94	0.026*
Assim:Cont	1	0.36	0.36	2619	10.51	0.0012 **

Tableau 11. Ecart-types des effets aléatoires (nbr.d'observations : 2628 , 116 sujets, 72 items).

<i>Group</i>	<i>Variance</i>	<i>Std. Deviation</i>
Subjets	0.0166425	0.129006
Items	0.0030887	0.055576
Residual	0.0347114	0.186310

Tableau 13. Les effets fixes [].

	<i>Estimate</i>	<i>Std. Error</i>	<i>DF</i>	<i>t value</i>	<i>Pr (> t)</i>
Intercept	6.32 e+00	3.01 e-02	2619	210.17	2.2 e-16
Freq	-2.87 e-02	6.29 e-03	2619	-4.56	5.28 e-06
Assim	6.56 e-02	3.64 e-02	2619	1.80	0.0715
NVD	1.70 e-02	6.40 e-03	2619	2.65	0.008
ConsF	-1.08 e-01	7.99 e-02	2619	-1.35	0.1777
Cont	5.49 e-03	2.74 e-02	2619	0.20	0.8413
Assim : ConsF	1.00 e-01	8.43 e-02	2619	1.19	0.2344
NVD : ConsF	-2.65 e-02	1.19 e-02	2619	-2.22	0.0265
Assim : Cont	-6.76 e-02	2.085e-02	2619	-3.24	0.0012