

Chapitre 5

Contraintes phonologiques et sélection lexicale

Ce chapitre a fait l'objet d'une communications orale dans un congrès avec comité de lecture assorti d'une publication dans des actes :

XXII^{èmes} Journées d'Etude sur la Parole, 15-19 Juin 1998,
Martigny, Suisse.

CONTRAINTES PHONOLOGIQUES ET SELECTION LEXICALE

Dans cette première série d'expériences, nous confrontons les données expérimentales interprétées comme la preuve d'un recours à des connaissances sur la structure phonologique de la langue (McQueen, 1998; Vroomen & de Gelder, 1999) à la première interprétation alternative que nous avons proposée. Nous étudions par conséquent le rôle possible des phénomènes de sélection lexicale dans un modèle purement séquentiel de l'accès au lexique : le modèle COHORT (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987). Avant de passer plus précisément à l'analyse des rôles dissociés de la fréquence et des régularités phonologiques, il est en effet important de déterminer dans quelle mesure l'un des modèles qui a exercé le plus d'influence au début des années 80 peut rendre compte des résultats obtenus. Nous avons montré dans le chapitre précédent que fréquence et tautosyllabité sont liées. Ceci implique qu'un choix aléatoire des séquences de phonèmes supposées permettre la comparaison de séquences phonotactiquement légales ou illégales risque fort de conduire à un déséquilibre dans la structure des stimuli utilisés. Le phonème donnant lieu à une séquence légale risque en effet de donner lieu à une séquence constituant le début d'un nombre de mots plus élevé que celui qui donne lieu à une séquence illégale. Lorsque ce phonème est en cours de traitement, le nombre de candidats lexicaux qui sont maintenus dans la cohorte peut différer en fonction du type de groupe de

phonèmes comparé. Il est donc nécessaire de contrôler cette variable afin de restreindre le champ des interprétations possibles.

1. Expérience 1 : Contraintes phonologiques et sélection lexicale

Dans cette première expérience, nous étudions l'effet de la structure phonologique des groupes de consonnes en contrôlant l'influence éventuelle de phénomènes de sélection lexicale sérielle en nous référant au modèle COHORT (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987). Pour cela, nous comparons l'effet de trois catégories de groupes de consonnes (occlusive-liquide correspondant à des groupes tautosyllabiques / fréquents, occlusive-fricative et occlusive-occlusive constituant des groupes hétérosyllabiques / rares) en imposant des contraintes lexicales au matériel choisi. Ainsi, aucune des séquences mot + C utilisées dans l'expérience ne peut correspondre au début d'un mot dans la langue (en tout cas pas dans la base de données Brulex, Content et al., 1990). Si l'effet de la structure phonologique des séquences de phonèmes est répliqué, il sera alors possible d'affirmer que les effets mis en évidence (McQueen, 1998; Vroomen & de Gelder, 1999) ne constituent pas des artefacts de phénomènes de sélection lexicale sérielle.

1.1. Méthode

1.1.1. Sujets

Trente et un étudiants en second cycle de psychologie à l'Université Paris 5 - René Descartes ont participé à cette expérience soit volontairement, soit en échange de crédits d'enseignement. Ils étaient tous de langue maternelle française et n'avaient jamais souffert de troubles auditifs.

1.1.2. Matériel

Les stimuli sont des bisyllabes de structure C(C)VCCVC sans signification. La moitié d'entre eux commence par un mot d'une syllabe correspondant à la séquence de phonèmes initiale C(C)VC (par exemple /brytrin/, 'brute' + /rin/). Trente-quatre mots monosyllabiques ont été sélectionnés, dans la base de données BRULEX (Content et al., 1990), afin de respecter un certain nombre de critères structuraux et lexicaux. A l'exception de l'un d'entre eux, aucun ne contient une voyelle qui permettrait d'influencer le rattachement syllabique de la consonne

suyvante (Goslin, Content, & Frauenfelder, 1999). Nous avons donc évité d'utiliser des mots contenant les phonèmes /o/ ou /e/ qui sont ouverts dans une syllabe lourde (comme /o/ dans [rɔk], 'roc' et /e/ dans [sɛr], 'serres') mais pas dans une syllabe légère (comme /o/ dans [ro # kaj], 'rocaille' et /e/ dans [se # re], 'serrer'). Le seul mot du matériel contenant une voyelle de ce type (/o/ dans 'faute') n'a finalement pas été pris en compte dans les analyses qui sont présentées ici³⁵.

Les mots monosyllabiques sélectionnés se terminent tous par une consonne occlusive. Les non-mots auxquels ils sont intégrés présentent une structure C(C)VCCVC pour laquelle la suite CC médiane peut correspondre à une suite occlusive-liquide (tautosyllabique / fréquente, par exemple /gl/), occlusive-fricative (hétérosyllabique / rare, par exemple /gʒ/) ou occlusive-occlusive (hétérosyllabique / rare, par exemple /gd/, cf. Tableau 7). La fréquence des suites de consonnes médianes n'a pas été contrôlée ; les suites tautosyllabiques sont en moyenne plus fréquentes que les suites hétérosyllabiques (occl.-liq. vs. occl.-fric., $t(39) = 10.763$, $p < .001$; occl.-liq. vs. occl.-occl., $t(39) = 13.352$, $p < .001$). Les séquences occlusive-fricative et occlusive-occlusive ne présentent quant à elles pas de différence significative de leurs probabilités d'occurrence ($t(66) = 1.723$, $p > .1$). Les suites occlusive-liquide donnent lieu à un non-alignement entre frontières lexicale et phonologique alors que les deux autres catégories de groupes produisent un alignement de ces deux types de frontières.

Tableau 7 : Conditions expérimentales de l'Expérience 1.

	occlusive-liquide	occlusive-fricative	occlusive-occlusive
Exemple	/brytrin/	/brytsin/	/brytkin/
Classification du groupe de consonnes	tautosyllabique	hétérosyllabique	
Statut de la consonne finale du mot	attaque syllabique	coda syllabique	
Probabilité d'occurrence moyenne des groupes médians	88 (6.5)	44 (21.9)	38 (22.5)

Aucune des séquences C(C)VC + C initiales ne constitue un début de mot possible dans la base de données BRULEX (Content et al., 1990). Par exemple, le mot 'bague' peut être intégré dans les séquences /baglym/ (occlusive-liquide,), /bagʒym/ (occlusive-fricative) et /bagbym/

³⁵ Son insertion dans les analyses de variance ne modifie cependant pas les résultats obtenus.

(occlusive-occlusive). La séquence /bag/ constitue le début d'un certain nombre de mots dans la langue (bagarre, baguette, bagatelle, ...). Par contre, aucune des séquences /bag/ + C ne correspond à un début de mot dans les trois conditions expérimentales comparées. Ni /bagl/, ni /bagʒ/, ni /bagb/ ne constituent le début d'un mot dans la langue³⁶. Afin de nous assurer que l'ensemble des stimuli sélectionnés correspondait à cette contrainte, nous avons utilisé un script AWK similaire à celui qui avait été mis en œuvre pour étudier la fréquence des groupes de consonnes en position initiale de mot. Le script lisait un fichier contenant la liste des suites C(C)VC + C et recherchait la présence éventuelle de chaque suite en début de mot. A la fin du traitement, un fichier était créé qui contenait la liste des séquences données en entrée et le nombre de fois où chaque séquence avait été rencontrée en début de mot dans la base de données. Pour chaque séquence, on vérifiait que le nombre était bien égal à 0. Dans le cas contraire, un phonème différent était choisi et le script relancé sur les suites modifiées jusqu'à ce que l'ensemble des suites de la liste corresponde à des séquences ne constituant le début d'aucun mot dans la base de données. Trente-quatre autres stimuli ont été choisis dans lesquels aucun mot de la langue n'est enchâssé (par exemple /ʒalkãb/ ou /ʃylit/).

Les stimuli sélectionnés ont été enregistrés sur DAT par un locuteur masculin. Celui-ci était installé dans une cabine insonorisée et prononçait les stimuli à raison d'un non-mot de deux syllabes toutes les trois secondes. Chaque stimulus était prononcé trois fois de suite avec une intonation la plus neutre possible. Les stimuli ont ensuite été digitalisés sur 16 bits (échantillonnage à 16 kHz) par l'intermédiaire de l'entrée audio d'une carte son et stockés sur le disque dur d'un ordinateur compatible PC. Le fichier correspondant a alors été édité afin d'extraire le signal correspondant à chaque stimulus et de le sauvegarder dans un fichier individuel. Un exemplaire de chaque item a été sélectionné parmi les trois répétitions. Les stimuli ont ensuite été édités afin de mesurer la durée entre le début du stimulus et celui de l'occlusion de la consonne finale du mot.

1.1.3. Procédure

Nous avons utilisé la tâche de *Word-spotting* (Cutler & Norris, 1988; McQueen, 1996) dans laquelle les participants doivent repérer des mots monosyllabiques prononcés dans un non-

³⁶ Par attesté dans la langue, nous entendons désormais attesté dans la base de données Brulex (Content, et al., 1990). Il est évident que cette base de données ne recense pas l'ensemble des mots de la langue française. Elle fournit cependant un outil inestimable pour conduire nos recherches. Il est certain que de nombreux mots de la langue n'apparaissent pas dans Brulex mais l'on peut raisonnablement considérer que l'ensemble des mots recensés



mot de deux syllabes (par exemple ‘vague’ dans /vagyʒ/). Les mots étaient toujours prononcés à l’initiale du non-mot. Les participants étaient assis sur une chaise, face à un bouton réponse connecté au port parallèle d’un ordinateur. Les participants étaient prévenus qu’ils allaient entendre des stimuli sans signification dont certains commenceraient par un mot de la langue (la consigne est retranscrite en Annexe 12, p.XXIII). Ils n’étaient pas informés des mots qui allaient apparaître dans le matériel. Lorsqu’ils identifiaient un mot en position initiale de non-mot, ils devaient appuyer le plus rapidement possible sur le bouton réponse. Dans le cas contraire, ils devaient attendre l’essai suivant qui commençait 3 secondes après la présentation du stimulus. Chaque fois qu’ils fournissaient une réponse manuelle indiquant la détection d’un mot en position initiale de non-mot, ils devaient dire à voix haute quel mot ils avaient identifié puis se préparer à l’essai suivant. L’expérimentateur vérifiait alors que le mot identifié était bien celui qui était attendu. Dans le cas contraire, le fichier de résultats était édité à la fin de la passation expérimentale pour y noter que la réponse était incorrecte.

Les stimuli étaient présentés à un niveau d’écoute confortable (environ 65dB). Les temps de réaction étaient mesurés depuis le début du stimulus et corrigés automatiquement par le logiciel de passation expérimentale en soustrayant de la latence de chaque temps de réaction la durée entre le début du stimulus et le début de l’occlusion correspondant à la consonne finale du mot (les résultats de ces mesures sont fournis en Annexe 13, p.XXIII). L’expérience commençait par une phase de familiarisation avec la tâche. Chaque sujet était soumis aux trois conditions expérimentales de contexte phonologique mais n’entendait chaque mot qu’une seule fois. Trois listes expérimentales ont donc été créées, chaque sujet se voyant attribuer aléatoirement l’une d’entre elles.

1.2. Résultats

Pour l’ensemble des expériences dans lesquelles nous avons eu recours à la tâche de *word-spotting*, deux critères d’exclusion des données expérimentales ont été choisis. Le premier porte sur la sélection des participants. Les données d’un sujet ne sont prises en considération dans l’analyse que s’il fournit une réponse manuelle *doublée* d’une identification correcte du mot attendu dans plus de la moitié des essais (sans préjuger du temps de réaction associé). Le second critère est appliqué après celui de sélection des participants et permet de sélectionner les items pertinents pour l’analyse. On considère qu’un item ne peut être intégré dans l’analyse que si plus

constitue le vocabulaire de base d’un locuteur adulte et que les mots supplémentaires sont pour la plupart



de la moitié des sujets ont fourni pour cet item une réponse manuelle *doublée* d'une identification correcte du mot attendu. Pour chaque analyse de variance, deux analyses ont été effectuées, l'une sur les latences de détection (en ms) et l'autre sur les taux d'erreurs (en %). Dans l'ANOVA portant sur les latences de détection, les temps de réaction intégrés dans l'analyse ne devaient pas excéder la valeur du temps de réaction moyen \pm deux écarts-type (SD pour *Standard Deviation*). Dans le cas contraire, ils n'étaient pas pris en compte dans l'analyse. Aucun des temps enregistrés ne passait en fait en-dessous de cette limite. Le critère consiste donc, dans l'ensemble des analyses qui ont été effectuées dans cette thèse, à supprimer les temps de réaction qui dépassent la valeur de la moyenne *plus* deux écarts-type. Dans les analyses qui portent sur les erreurs, on a considéré comme erreur toute absence de réponse mais aussi (1) toute réponse ayant été suivie d'une prononciation du mot ne correspondant pas à celui qui était attendu et (2) tout temps de réaction supérieur à la limite déterminée dans l'analyse des temps de réaction. Nous présentons pour chaque ANOVA les résultats obtenus dans l'analyse par *Sujet* (F_1) et dans l'analyse par *Item* (F_2).

1.2.1. Analyse globale

Les données de l'un des participants n'ont pas pu être intégrées dans l'analyse en raison d'un taux d'erreurs ou d'omissions supérieur à 50%. Deux items ('blague' et 'type') ont été supprimés de l'analyse en raison d'un taux trop important d'erreurs d'identification ou d'absence de détection de la part des sujets. Un autre item a été retiré de l'analyse en raison de la présence d'une voyelle qui, selon les régularités phonologiques du français, est apte à fournir des indices sur la structure syllabique de la séquence (/o/ dans 'faute'). Les temps de réaction supérieurs à 1430 ms n'ont pas été introduits dans l'analyse. Au total, 15 % des données n'ont pas été intégrées dans l'analyse des temps de réaction.

1.2.1.1. Temps de réaction

La variable *Liste* n'interagit pas avec la variable *Contexte*. Elle n'est donc pas introduite dans les analyses de variance suivantes. Celles-ci sont conduites avec le *Contexte* (C_3) comme variable intra-sujet ($S * C_3$) et intra-item ($I * C_3$). La latence de détection moyenne est de 702 ms (SD = 181 ms). On observe des temps de réaction en moyenne plus longs pour les groupes occlusive-liquide (722 ms) que pour les deux autres catégories de groupes (cf. Figure 28). Les latences obtenues pour les groupes occlusive-fricative sont quant à elles plus élevées (703 ms) que celles de la condition occlusive-occlusive (680 ms). L'effet global du contexte phonologique

relativement rares.

est significatif dans l'analyse par sujet ($F_1(2,58) = 4.085, p < .05$) mais pas dans l'analyse par item ($F_2(2,60) = 1.452, p > .1$). Du fait du statut similaire des groupes occlusive-fricative et occlusive-occlusive en ce qui concerne leur syllabation, nous avons approfondi l'analyse en comparant les latences de détection observées dans ces deux conditions. Celles-ci ne diffèrent pas significativement ($F_1(1,29) = 3.486, p = .072$; $F_2 < 1$). Les temps de détection enregistrés dans la condition occlusive-liquide (tautosyllabique) sont quant à eux significativement plus élevés ($F_1(1,29) = 4.762, p < .05$) que pour les deux autres types de groupes de consonnes (hétérosyllabiques). Cette différence n'est cependant pas significative dans l'analyse par item ($F_2(1,30) = 2.205, p > .1$).

Cette expérience ne confirme qu'en partie les résultats obtenus avec la même tâche par McQueen (1998) dans l'analyse des erreurs ou les données de temps de réaction observées par Vroomen & De Gelder (1999) dans une tâche de détection de phonèmes. Seuls les effets étudiés dans l'analyse par sujet sont en effet significatifs. Aucun des effets de l'analyse par item n'atteint le seuil de significativité statistique.

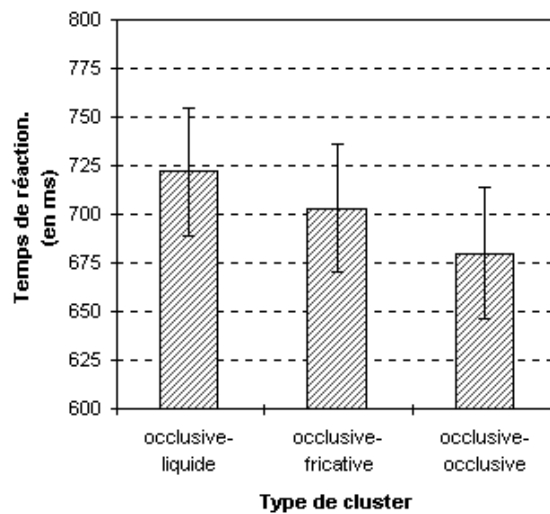


Figure 28 : Temps de réaction moyens (en ms) observés dans les trois conditions expérimentales de contexte phonologique de l'Expérience 1. Les barres verticales représentent l'erreur-standard³⁷ (SE).

1.2.1.2. Taux d'erreurs

Les mêmes analyses ont été conduites sur les taux d'erreurs (15 % de l'ensemble des données³⁸). Le facteur *Liste* n'interagit pas avec le facteur *Contexte*. On observe un pourcentage d'erreurs légèrement plus élevé dans la condition occlusive-liquide (17.8 %) que dans les

³⁷ Erreur-standard (*Standard-Error*) : Ecart-type (*Standard Deviation*) divisé par le nombre d'observations (i.e. de sujets).

conditions occlusive-fricative (13.7 %) et occlusive-occlusive (14.4 %). L'effet du facteur *Contexte* n'est cependant significatif dans aucune des analyses ($F_1(2,58) = 1.380, p > .1$; $F_2 < 1$).

1.2.1.3. Discussion

Les données de temps de réaction répliquent donc en partie les effets observés par McQueen (1998) et par Vroomen & De Gelder (1999). Les locuteurs français ont plus de difficultés à détecter un mot en position initiale de non-mot lorsque la frontière entre le mot et la suite du non-mot est alignée avec une frontière pouvant s'assimiler à une frontière phonologique (syllabique ou phonotactique). Cet effet se maintient même si la suite mot + C ne constitue à aucun moment un début de mot dans la langue. Les effets observés précédemment ne sont donc pas uniquement interprétables à partir d'un modèle de sélection lexicale sérielle comme COHORT (Marslen-Wilson & Welsh, 1978 ; Marslen-Wilson, 1987). Il reste cependant que les effets observés dans cette expérience ne sont significatifs que dans les analyses conduites avec le facteur *Sujet* comme variable aléatoire. Toutes les analyses par item fournissent des résultats non-significatifs. Il semblerait donc que les latences de détection moyennées sur les items (les mots) lorsqu'on les regroupe par catégorie de contexte phonologique ne soient pas suffisamment stables pour générer des valeurs moyennes significativement différentes. Il est probable que ce phénomène puisse s'expliquer par des différences liées à des caractéristiques des mots choisis dans le matériel (par exemple leur fréquence d'usage). En effet, la tâche de *word-spotting* est particulièrement difficile à accomplir. Cette difficulté se manifeste aussi bien dans les jugements des participants que dans les taux d'erreurs et les latences de détection moyenne observés dans toute expérience faisant appel à ce paradigme (Cutler & Norris, 1988; McQueen, 1996). Du fait de cette difficulté, on peut imaginer que la présence de mots particulièrement difficiles à isoler sur la base de leurs caractéristiques intrinsèques pourrait donner lieu à des effets moins nets du contexte phonologique ; des mots relativement faciles à isoler donnant au contraire lieu à des effets plus clairs du contexte. La présence de mots pour lesquels le contexte n'a qu'un effet limité pourrait alors empêcher le test statistique de rejeter l'hypothèse nulle pour l'autre ensemble de mots. Nous avons donc choisi de scinder le matériel expérimental à partir de critères propres aux mots de notre échantillon. Si c'est effectivement ce paramètre qui conduit à une absence de significativité dans l'analyse par item, ceci devrait nous permettre de stabiliser la variabilité propre aux items et de faire émerger des comparaisons significatives dans cette partie de l'analyse.

³⁸ Les temps de réaction supérieurs à la moyenne + deux écart-types sont considérés comme des erreurs.

1.2.2. Analyses complémentaires

1.2.2.1. Contexte phonologique et fréquence d'usage

Nous avons en premier lieu cherché à évaluer l'influence possible de la fréquence d'usage des mots de l'échantillon dans l'émergence de l'effet du contexte phonologique. Il serait tout à fait possible que ce paramètre puisse influencer non seulement les temps de réaction observés en *word-spotting* mais aussi l'effet des divers facteurs introduits dans l'analyse. C'est par ailleurs un paramètre immédiatement disponible dans la base de données BRULEX (Content et al., 1990). L'objet n'est pas ici de tester l'effet de la fréquence des mots sur les temps de réaction mais plutôt l'influence que peut avoir eu la *variabilité attachée à ce paramètre* sur la capacité du test statistique à faire émerger des effets de contexte phonologique significatifs. Notre but est donc de limiter la variabilité des latences de détection liée à la fréquence des mots afin de faciliter l'émergence d'effets significatifs dans les analyses par item. Nous avons estimé la médiane³⁹ de la distribution des fréquences d'usage⁴⁰ des mots de notre échantillon afin de scinder celui-ci en deux groupes de taille approximativement égale sans nous préoccuper de la signification de la valeur de fréquence d'usage en termes de fréquence absolue d'utilisation dans la langue. La médiane est à 19. Les mêmes analyses ont alors été conduites sur chacun des échantillons de mots, l'un contenant les mots les plus rares de l'échantillon ($FU < 19$, $N = 15$, moyenne = 11), l'autre les plus fréquents ($FU \geq 19$, $N = 16$, moyenne = 132). Puisqu'aucun effet, même marginal, n'émergeait dans l'analyse des erreurs, nous n'avons répliqué que les analyses de temps de réaction.

1.2.2.1.1. Résultats

Le facteur *Liste* n'interagit pas avec le facteur *Contexte* ; les analyses présentées ne le prennent donc pas en compte. Les latences de détection obtenues pour les 15 mots les plus rares de l'échantillon sont en moyenne de 713 ms ($SD = 201$ ms) et les moyennes observées ne correspondent pas à ce qui était obtenu dans l'analyse globale. Les groupes occlusive-fricative semblent donner lieu à des temps de réaction plus longs (741 ms) que les occlusive-liquide (708 ms) et les occlusive-occlusive (691 ms). On n'observe cependant aucun effet du contexte phonologique pour ce groupe de mots ($F_1(2,58) = 1.983$, $p > .1$; $F_2 < 1$). L'analyse conduite sur

³⁹ Médiane : cet indice statistique est la valeur qui sépare un échantillon en deux groupes de taille égale. La moitié des mots de l'échantillon a donc une fréquence d'usage supérieure à la médiane. Cet indice est ici plus intéressant que la moyenne puisque celle-ci peut séparer deux groupes de taille tout à fait inégale en raison du poids des valeurs extrêmes dans le calcul de la moyenne.

⁴⁰ Notée FU (Fréquence d'Usage sur 1 million). Les données de fréquence présentes dans BRULEX (Content et al., 1990) doivent être divisées par 100 pour aboutir à une fréquence d'usage sur 1 million.

les 16 mots les plus fréquents de l'échantillon (RT = 673 ms, SD = 198 ms) fournit des données sensiblement plus proches de ce qui était observé dans l'analyse globale. Les groupes occlusive-liquide donnent lieu à des latences de détection plus longues (720 ms) que les deux autres catégories de groupes (respectivement 662 ms et 637 ms pour les occlusive-fricative et les occlusive-occlusive). L'effet global du contexte n'est à nouveau significatif que dans l'analyse par sujet ($F_1(2,58) = 3.462, p < .05$) et pas dans l'analyse par item ($F_2(2,30) = 1.599, p > .1$). Nous avons à nouveau regroupé les groupes occlusive-fricative et occlusive-occlusive (qui ne diffèrent pas entre eux : $F_1 < 1$; $F_2 < 1$) et les avons comparés aux groupes occlusive-liquide. On observe de même que dans l'analyse globale un effet significatif par sujet ($F_1(1,29) = 7.838, p < .01$) mais celui-ci n'est que marginal dans l'analyse par item ($F_2(1,15) = 4.301, p = .056^\circ$).

1.2.2.1.2. Discussion

La restriction à une partie du matériel linguistique utilisée dans l'Expérience 1 semble avoir permis au test statistique de rejeter plus facilement l'hypothèse nulle dans l'analyse par item, ce qui conduit à observer un seuil statistique marginal alors qu'il était non-significatif dans l'analyse globale. Cet effet n'est cependant que marginal. Nous nous sommes alors posé la question de l'existence d'une autre variable lexicale qui aurait pu introduire une variabilité dans les temps de réaction observés et serait plus appropriée que la fréquence d'usage des mots. Nous pensons que le nombre de compétiteurs lexicaux pourrait constituer une information influençant les traitements effectués au cours de la tâche de *word-spotting*. Nous en sommes arrivé à choisir cette variable en raison du lien qu'elle peut avoir avec la fréquence d'usage. En effet, (Luce, 1990) montre que les mots fréquents ont en général un nombre de voisins lexicaux plus important que les mots rares et que les effets de fréquence peuvent se confondre avec les effets de voisinage lexical. Fréquence et nombre de compétiteurs sont donc liés, ce qui nous conduit à envisager que si la conduite de l'analyse sur les mots fréquents contribue à observer un seuil de probabilité plus bas (bien que toujours supérieur au seuil de significativité), il pourrait en être de même si l'on restreint l'analyse aux mots ayant beaucoup de compétiteurs. Par ailleurs, cette variable nous semblait mieux à même d'interagir réellement avec le contexte phonologique. En effet, Norris, McQueen, & Cutler (1995) citent un travail non-publié dans lequel seuls des effets de fréquence relativement importants auraient été observés dans une tâche de *word-spotting*. Dans ce même travail, des différences de fréquence limitées ne permettent pas de faire émerger des effets significatifs. Or la scission que nous avons effectuée dans notre matériel ne nous permet pas de dissocier groupes très rares et groupes très fréquents. En réalité, la plupart des mots que nous avons utilisés appartiennent à la classe des mots de fréquence d'usage moyenne (la moitié des mots utilisés se situent entre 11 et 52 sur 1 million sur l'échelle de fréquence

d'usage de BRULEX). La comparaison de mots présentant des différences relativement peu importantes en termes de fréquence ne devrait donc pas, dans une tâche de *word-spotting*, donner lieu à des différences importantes en termes de temps de réaction. Dans cette expérience nous avons pu observer, au cours de l'analyse ayant consisté à restreindre l'ANOVA aux mots fréquents d'une part et rares de l'autre, des temps de réaction moyens de 673 et 713 ms respectivement. Bien que cette différence soit significative ($t(29) = 2.099$, $p < .05$), il n'est pas certain qu'elle soit réellement liée à une différence de fréquence des items comparés. D'autres paramètres peuvent très bien être corrélés et déterminer une diminution des temps de réaction pour les mots fréquents sans que la différence soit effectivement liée à cette variable. On peut par exemple envisager que Si la fréquence d'usage n'est pas réellement en mesure d'influencer les traitements effectués au cours de la tâche, il est probable que la dichotomie réalisée ne soit pas à même de contribuer à faire émerger des effets significatifs pour une partie des items. Il nous semble plus probable que le nombre de compétiteurs d'un mot puisse influencer les processus qui sont à l'œuvre dans une tâche de *word-spotting*. En effet, la fréquence d'un mot n'a pas nécessairement de représentation dans le lexique. Son intervention pourrait d'ailleurs être envisagée en termes décisionnels ou intégratifs, c'est à dire qu'elle influencerait l'étape de prise de décision mais pas celle d'accès au lexique proprement dite. Le nombre de compétiteurs d'un mot constitue par contre une variable essentielle des processus de reconnaissance de la parole. Ainsi, dans le cadre d'un modèle comme COHORT (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987), ce paramètre est partie intégrante du traitement effectué sur le signal de parole. Le processus de reconnaissance des mots repose en effet essentiellement sur la sélection d'un mot parmi l'ensemble des candidats lexicaux activés dans la cohorte. Des modèles comme TRACE (McClelland & Elman, 1986) ou SHORTLIST (Norris, 1994) reprennent ce principe de sélection parmi un ensemble de candidats lexicaux en implémentant des processus de compétitions lexicales. Or ce type de processus fait implicitement intervenir le nombre de candidats lexicaux dans le déroulement des processus de traitement. Ce paramètre est donc en mesure d'influencer la durée nécessaire pour isoler un mot parmi l'ensemble des mots possibles et est devenu une caractéristique essentielle de tout modèle de l'accès au lexique. Il semble donc possible de choisir un critère alternatif à la fréquence pour tenter de réduire la variance liée aux items et de faire émerger un effet d'alignement significatif dans une analyse par item.

1.2.2.2. Contexte phonologique et nombre de compétiteurs

Afin d'affiner notre analyse du rôle des contraintes phonologiques dans les processus de segmentation de la parole en mots, nous avons donc choisi de conduire une analyse similaire à la

précédente en choisissant un critère plus adéquat pour restreindre l'étude à une partie du matériel : le nombre de compétiteurs lexicaux des mots utilisés dans l'expérience.

1.2.2.2.1. *Méthode*

L'estimation du nombre de compétiteurs des items expérimentaux a été réalisée avec le même script AWK que celui qui nous avait permis de vérifier qu'il n'existait pas de séquence 'mot' + C constituant le début d'un mot lors de la constitution du matériel de l'Expérience 1. Le script prenait en entrée la liste des mots utilisés dans cette expérience et recherchait les occurrences de la séquence de phonèmes correspondante au début des mots de la base de données BRULEX (Content et al., 1990). Le fichier de sortie contenait la liste des mots et le nombre de mots commençant par cette séquence dans la base de données. Cette opération nous a permis de calculer la médiane de cet indice et de séparer les items en 2 groupes. Les analyses de variance conduites sur l'intégralité du matériel de l'Expérience 1 ont été reconduites sur chacun de ces groupes de mots en se restreignant à l'analyse des temps de réaction.

La médiane de la distribution des items en terme de nombre de compétiteurs est à 6. Or trois des mots de l'échantillon ont exactement 6 compétiteurs dans la base de données BRULEX (Content et al., 1990). Cette valeur pose un problème. Soit l'on décide de dissocier les résultats obtenus pour les mots ayant plus de 6 compétiteurs et les autres (On compare alors les résultats obtenus dans des groupes de 16 et 15 items respectivement pour les mots ayant peu et beaucoup de compétiteurs), soit l'on choisit d'intégrer dans le groupe des mots ayant beaucoup de compétiteurs ceux qui ont plus de 5 compétiteurs (donc 6 ou plus de 6) et de restreindre l'autre groupe aux items restants, ce qui conduit à comparer des échantillons de 13 et 18 items respectivement pour les groupes de mots ayant peu et beaucoup de compétiteurs. Les mots que nous avons qualifiés de fréquents ont en moyenne un nombre de compétiteurs plus élevé (10.06) que les mots rares (6.44). Cette différence n'est pas significative ($t(28) = 1.500, p > .1$) mais elle nous incite à privilégier l'une des deux possibilités qui s'offrent à nous pour la dichotomie que nous souhaitons effectuer. C'est en effet pour les mots fréquents que l'on a observé la possibilité d'une différence statistiquement significative entre les conditions de contexte phonologique dans la section précédente. Du fait de la quantité relativement restreinte de mots dans notre matériel, nous avons choisi de sélectionner une méthode de dichotomie qui nous permettrait de conserver au test statistique une puissance raisonnable lui permettant de faire émerger des effets significatifs. Cette analyse restreinte en fonction du nombre de compétiteurs des mots de l'échantillon a donc été conduite d'une part sur les mots qui ont 6 compétiteurs ou plus (18 items), et d'autre part sur ceux qui en ont moins de 6 (les 13 autres items). En raison du nombre

de compétiteurs sensiblement plus élevé pour les mots fréquents que pour les mots rares, il est probable qu'un éventuel effet d'alignement devrait se manifester pour les mots ayant beaucoup de compétiteurs.

1.2.2.2.2. Résultats

Aucune des analyses effectuées ne fait ressortir un effet d'interaction entre les variables *Contexte* et *Liste* ; cette dernière n'est donc pas intégrée dans les analyses de variance présentées. De même que pour la restriction aux mots les plus rares, l'analyse des 13 mots pour lesquels le nombre de compétiteurs est le moins élevé ne permet pas de mettre en évidence des différences de temps de réaction entre les diverses conditions expérimentales. Les conditions occlusive-liquide, occlusive-fricative et occlusive-occlusive donnent lieu à des latences de détection qui sont respectivement de 690, 726 et 721 ms (tous les F sont inférieurs à 1, la latence moyenne de détection est de 712 ms, $SD = 197$).

L'analyse conduite sur les 18 mots de l'échantillon présentant la quantité la plus élevée de compétiteurs aboutit au contraire à des résultats très similaires à ceux que l'on a observé pour les mots les plus fréquents. Le temps de réaction moyen est de 698 ms ($SD = 201$ ms). Les groupes occlusive-liquide donnent lieu à des latences de détection plus longues (750 ms) que les groupes occlusive-fricative (691 ms) et occlusive-occlusive (655 ms). La Figure 29 fournit une représentation graphique de ces temps de réaction. L'effet global du contexte est significatif par sujet ($F_1(2,58) = 11.693$, $p < .01$) mais pas par item ($F_2(2,34) = 2.076$, $p > .1$). La comparaison des conditions occlusive-fricative et occlusive-occlusive montre que les temps de réaction

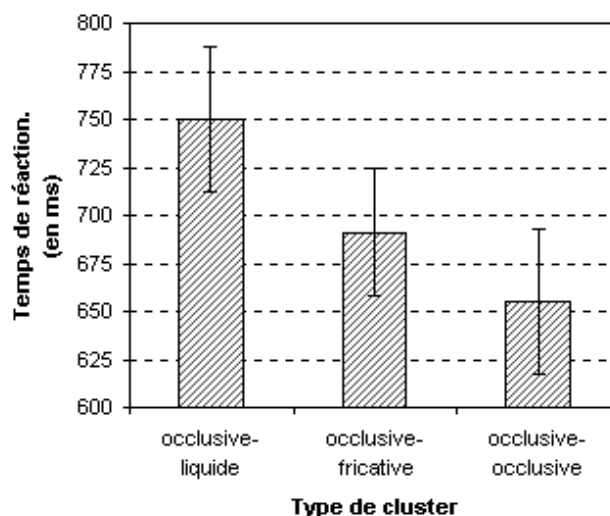


Figure 29 : Temps de réaction moyens (en ms) observés pour les 18 mots de l'échantillon ayant le nombre de compétiteurs le plus élevé. Les barres verticales représentent l'erreur-standard (SE).

associés ne diffèrent que marginalement dans l'analyse par sujet ($F_1(1,29) = 3.950$, $p = .056$) et pas dans l'analyse par item ($F_2 < 1$). Nous avons donc regroupé ces deux conditions et conduit une analyse permettant de comparer les groupes occlusive-liquide à ces deux autres catégories. Cette comparaison met en évidence des différences significatives aussi bien dans l'analyse par sujet ($F_1(1,29) = 19.614$, $p < .01$) que dans l'analyse par item ($F_2(1,17) = 4.643$, $p < .05$). Les mots sont détectés plus rapidement lorsque la consonne finale occlusive est suivie d'une liquide donnant lieu à un groupe de consonnes tautosyllabique / fréquent que lorsque cette même consonne est suivie d'une fricative ou d'une occlusive qui, quant à elles, génèrent un groupes de consonnes hétérosyllabique / rare.

1.3. Discussion

L'objet de cette expérience était de tester le rôle d'éventuelles connaissances sur les contraintes phonologiques de la langue avec un matériel permettant de neutraliser la contribution des processus de sélection sérielle dans l'émergence des effets d'alignement phonologique. Pour cela, nous avons utilisé des stimuli pour lesquels l'occurrence de la consonne contextuelle, celle qui suit le mot dans le matériel de *word-spotting*, permettait de mettre fin à toute compétition entre des candidats qui resteraient maintenus dans la cohorte lexicale initiale du modèle COHORT (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987). Pour l'ensemble des stimuli, la séquence 'mot' + C ne pouvait ainsi en aucun cas constituer le début d'un mot dans la base de données lexicale BRULEX (Content et al., 1990). Le choix de ce matériel rend donc impossible l'existence d'un déséquilibre du nombre de candidats lexicaux maintenus dans la cohorte à l'instant auquel est traitée la consonne contextuelle dans les conditions d'alignement et de non-alignement des frontières syllabique et lexicale.

Une première analyse a permis de répliquer l'effet du contexte phonologique, c'est à dire de l'alignement entre frontières syllabique et lexicale, malgré ce contrôle des processus de sélection lexicale sérielle. Cet effet n'a cependant pu être mis en évidence que dans l'analyse par sujet, celui-ci n'atteignant pas le seuil de significativité statistique dans l'analyse par item. La restriction de l'analyse de variance à une partie du matériel lexical nous a cependant permis de mettre en évidence un effet de la structure phonologique des groupes de consonnes sur les latences de détection enregistrées dans la tâche de *word-spotting*. Deux questions importantes se posent néanmoins à partir de ces résultats. Nous avons vu que l'effet du contexte phonologique n'est en réalité significatif que pour une partie des items expérimentaux : les mots ayant beaucoup de compétiteurs. Cette restriction de l'effet à une sélection de mots pose d'autant plus

le problème de l'interprétation qu'il convient de donner de ces résultats. Nous avons vu, lors de la description du travail de McQueen (1998), que les effets contextuels observés dans la tâche de *word-spotting* pourraient très bien s'expliquer par des caractéristiques intrinsèques aux mots que doivent détecter les participants. Ce problème se pose lors de toute étude de phénomènes contextuels impliquant la présentation d'exemplaires physiques différents en fonction des conditions de contexte. La limitation de l'effet à une partie des mots du matériel accentue la possibilité que l'effet de contexte observé soit en réalité lié à certaines caractéristiques de la portion acoustique qui correspond aux mots utilisés. Il est donc essentiel de déterminer la contribution de caractéristiques intrinsèques aux mots dans l'émergence même de l'effet avant de réfléchir aux raisons pour lesquelles un éventuel effet des contraintes phonologiques pourrait se restreindre à une partie des mots de l'échantillon.

2. Expérience 2 : Contrôle en décision lexicale

L'objet de la seconde expérience est d'évaluer la contribution de paramètres acoustiques propres aux exemplaires des mots utilisés dans la tâche de *word-spotting* à l'émergence de l'effet d'alignement observé dans l'Expérience 1. A cet effet, nous avons mis en œuvre une tâche de décision lexicale sur les fragments acoustiques correspondant à la portion C(C)VC initiale des stimuli utilisés dans la précédente expérience. Si l'effet contextuel observé précédemment peut s'expliquer par des caractéristiques acoustiques propres aux exemplaires utilisés dans la première expérience et n'a rien à voir avec le statut de la consonne contextuelle, les effets observés dans l'Expérience 1 devraient également émerger dans une tâche de décision lexicale dans laquelle les conditions contextuelles n'interviennent pas en tant que tel. Ainsi, les mots extraits des stimuli de la condition occlusive-liquide devraient alors donner lieu à des latences de décision lexicale plus longues que les mots extraits des conditions occlusive-fricative ou occlusive-occlusive. Pour cela, nous comparons les réponses fournies pour chaque portion C(C)VC initiale en fonction du contexte dont elle est extraite. On observe par exemple les réponses induites par le mot 'brute' en fonction du stimulus dans lequel il était produit dans l'Expérience 1 (occlusive-liquide, occlusive-fricative, occlusive-occlusive) alors que seule la portion lexicale est disponible. Si la prononciation de certaines consonnes contextuelles a influencé la prononciation du mot initial au point de faire émerger les effets que nous avons observés dans la tâche de *word-spotting*, ces effets devraient se répliquer malgré l'absence de contexte. Au contraire, si l'effet n'est pas reproduit, il sera possible d'affirmer que les différences observées dans l'expérience précédente n'étaient pas liées à des différences acoustiques des stimuli qui seraient déterminées par les

caractéristiques de la consonne contextuelle. Si l'effet observé dans l'expérience précédente est effectivement lié à la mise en œuvre de processus de segmentation lexicale reposant sur le statut de la consonne contextuelle, lequel influencerait l'alignement entre frontières phonologique-fréquentielle et lexicale, nous nous attendons à ne pas observer d'effet de l'origine des stimuli dans cette expérience.

2.1. Méthode

2.1.1. *Sujets*

Trente étudiants en second cycle de psychologie à l'Université Paris 5 - René Descartes ont participé à cette expérience, soit volontairement, soit en échange de crédits d'enseignement. Ceux-ci n'avaient pas pris part à la première expérience. Ils étaient tous de langue maternelle française et n'avaient jamais souffert de troubles auditifs.

2.1.2. *Matériel*

Les stimuli de l'Expérience 1 ont été édités avec un logiciel de traitement du signal. La fin acoustique de la portion correspondant à la séquence C(C)VC initiale a été estimée à partir du spectrogramme. Cette décision a ensuite été confirmée auditivement afin de s'assurer que la consonne contextuelle n'était effectivement pas audible jusqu'à ce point. La section comprise entre le début du stimulus et la fin acoustique de la portion C(C)VC a finalement été extraite et stockée sous le même format que les stimuli de l'Expérience 1 (digitalisation sur 16 bits, échantillonnage à 16 kHz). Cette opération a été effectuée avec les stimuli composés d'un mot en position initiale aussi bien qu'avec les stimuli de remplissage. Pour les stimuli-test, nous disposons alors de 3 exemplaires de chaque mot en fonction du contexte duquel il avait été extrait. Par exemple, la portion correspondant à la séquence /bag/ ('bague') a été extraite des stimuli /baglym/ (occlusive-liquide), /bagzym/ (occlusive-fricative) et /bagbym/ (occlusive-occlusive). Les trois stimuli obtenus contiennent uniquement la séquence de phonèmes /bag/, mais celle-ci provient dans chacun d'entre eux de 3 exemplaires acoustiques différents. La même opération a été réalisée avec les stimuli de remplissage.

2.1.3. *Procédure*

La procédure de passation expérimentale est identique à celle de l'Expérience 1. Seule la tâche diffère. Les participants devaient effectuer une tâche de décision lexicale sur des stimuli mono-syllabiques. Les stimuli obtenus à partir de la portion initiale des items de remplissage de

l'Expérience 1 correspondent à des non-mots. Ceux qui proviennent des stimuli-test forment des mots. Nous avons utilisé une procédure *go - no go* qui s'apparente à ce qui était demandé dans la tâche de word-spotting. Lorsque le sujet identifiait le stimulus comme étant un mot de la langue, il devait appuyer le plus rapidement possible sur le bouton réponse puis prononcer ce mot à voix haute. Dans le cas contraire, il n'appuyait pas et devait attendre l'essai suivant. L'intervalle inter-stimuli était le même que dans l'Expérience 1 (3 s). Les temps de réaction étaient mesurés depuis le début du stimulus et corrigés automatiquement par le logiciel de passation expérimentale en soustrayant de la latence de chaque temps de réaction la durée entre le début du stimulus et le début de l'occlusion correspondant à la consonne finale du mot. Cette durée est la même que celle qui était utilisée pour corriger les temps de réaction de l'Expérience 1. L'expérience commençait par une phase de familiarisation avec la tâche. Chaque sujet était soumis aux trois conditions d'origine du stimulus mais n'entendait chaque mot qu'une seule fois. Trois listes expérimentales ont été créées, chaque sujet se voyant attribuer aléatoirement l'une d'entre elles.

2.2. Résultats

Nous avons utilisé les mêmes critères de sélection des participants et des items que dans l'Expérience 1. Du fait de ces critères, les données de l'ensemble des participants ont été intégrées dans l'analyse. Cinq items ont cependant été supprimés de l'analyse en raison d'un taux trop important d'erreurs d'identification ou d'absence de détection de la part des sujets. Les temps de réaction supérieurs à 932 ms n'ont pas été introduits dans l'analyse. Les données non-intégrées dans l'analyse des latences de détection constituent 13 % de l'ensemble des observations.

2.2.1. Temps de réaction

Les temps de réaction et les taux d'erreur observés sont présentés dans le Tableau 8. La variable *Liste* n'interagit pas avec la variable *Origine*. Elle n'est donc pas introduite dans les analyses de variance présentées. Celles-ci sont conduites avec l'*Origine* (O_3) comme variable intra-sujet ($S * O_3$) et intra-item ($I * O_3$). La latence de détection moyenne est de 469 ms ($SD = 88$ ms). On observe des temps de réaction qui sont en moyenne plus *courts* pour les mots provenant de la condition occlusive-liquide (455 ms) que pour les deux autres catégories de groupes. Les latences obtenues pour les mots issus des conditions de contexte occlusive-fricative et occlusive-occlusive sont respectivement de 481 ms et 472 ms.

Le fait d'observer, pour les stimuli prononcés en condition occlusive-liquide, des temps de réaction en moyenne plus rapides que pour les stimuli prononcés dans les conditions occlusive-

fricative et occlusive-occlusive est déjà un point important pour affirmer que l’allongement des latences de détection observé dans la tâche de *word-spotting* ne peut s’expliquer par des paramètres acoustiques (liés à des phénomènes articulatoires) qui ralentiraient l’identification des mots lorsque ceux-ci sont suivis d’une consonne liquide par rapport à une situation dans laquelle la consonne suivante est une fricative ou une occlusive. Il est cependant intéressant d’analyser plus en détails les mesures obtenues dans cette tâche. L’effet global de l’origine des stimuli est marginal dans l’analyse par sujet ($F_1(2,58) = 3.118$, $p = .052$) mais n’est pas significatif dans l’analyse par item ($F_2(2,56) = 1.722$, $p > .1$). Du fait du statut similaire des groupes occlusive-fricative et occlusive-occlusive pour ce qui a trait à leur syllabation, nous avons approfondi l’analyse en regroupant ces deux modalités en une seule. On observe à nouveau une absence de différence significative entre ces deux conditions (tous les F sont inférieurs à 1). Par contre, les temps de réaction observés pour les stimuli qui étaient présentés à l’origine en condition occlusive-liquide sont significativement plus courts que ceux des deux autres modalités ($F_1(1,29) = 4.708$, $p < .05$). Cette différence n’est que marginale dans l’analyse par item ($F_2(1,28) = 3.543$, $p = .070$). On observe par conséquent un effet non-négligeable de l’origine des stimuli. Il est probable que cette différence dans les temps nécessaires à fournir une décision lexicale sur le statut des mots en fonction du contexte duquel ils ont été extraits est liée à des paramètres acoustiques qui seraient déterminés par des contraintes articulatoires ayant influencé la prononciation de l’occlusive finale du mot.

Tableau 8 : Temps de réaction (en ms, erreur-standard entre parenthèses) et taux d’erreurs (en %) observés dans l’Expérience 2.

	occlusive-liquide	occlusive-fricative	occlusive-occlusive
Temps de réaction	455 (75)	481 (99)	472 (89)
Taux d’erreur	8.1	17.6	13.3

Cet effet est cependant l’inverse de celui qui a été obtenu dans l’Expérience 1. L’analyse des temps de réaction mesurés dans la tâche de décision lexicale ne permet donc pas de rendre compte de l’effet d’alignement entre frontières syllabique et lexicale. Il est cependant courant d’observer des phénomènes de compensation entre latences de décision et taux d’erreurs. Dans ces situations, on peut par exemple observer une correspondance entre rétrécissement des temps de réaction et accroissement des taux d’erreur. Ce phénomène illustre généralement une attitude des sujets qui consiste à tenter de répondre plus rapidement dans certaines situations, ces réponses plus rapides induisant des erreurs plus nombreuses. Ce phénomène est donc plutôt lié à

des stratégies de réponse qui n'illustrent pas réellement un effet au niveau des processus de traitement mais peuvent s'expliquer en termes stratégiques. On peut néanmoins envisager que, dans la tâche de décision lexicale, certaines caractéristiques acoustiques des stimuli produits en condition occlusive-liquide aient incité les participants à répondre plus rapidement (par exemple, moins d'indices acoustiques sur l'identité du phonème contextuel). Les réponses plus rapides aux mots extraits de la condition occlusive-liquide pourraient donc induire des taux d'erreurs plus importants dans cette condition. Il est par conséquent nécessaire d'analyser d'éventuels effets dans les taux d'erreurs afin de vérifier que les différences de temps de réaction observées dans la première expérience ne se répliquent pas ici avec une autre variable dépendante.

2.2.2. Taux d'erreurs

Les mêmes analyses ont donc été conduites sur les taux d'erreurs (13 % de l'ensemble des données). Le facteur *Liste* n'interagit pas avec le facteur *Origine*. On observe un pourcentage d'erreurs nettement plus bas dans la condition occlusive-liquide (8.1 %) que dans les conditions occlusive-fricative (17.6 %) et occlusive-occlusive (13.3 %). Cette observation met en évidence une similarité avec l'effet qui est observé dans les temps de réaction mesurés dans la même tâche. L'effet du facteur *Origine* est statistiquement significatif dans l'analyse par sujet ($F_1(2,58) = 5.711, p < .01$) mais pas dans l'analyse par item ($F_2(2,56) = 2.322, p > .1$). Les taux d'erreur observés dans les conditions occlusive-fricative et occlusive-occlusive ne diffèrent cependant pas significativement ($F_1(1,29) = 1.771, p > .1$; $F_2 < 1$). La comparaison des taux d'erreur observés dans ces deux conditions et de ceux de la condition occlusive-liquide montre que les erreurs sont significativement plus nombreuses lorsque les mots sont extraits d'un contexte hétérosyllabique que lorsqu'ils ont été prononcés dans un contexte tautosyllabique ($F_1(1,29) = 13.865, p < .01$). Cet effet n'est à nouveau que marginal dans l'analyse par item ($F_2(1,28) = 3.679, p = .065$). Les participants ont donc éprouvé moins de difficultés à fournir une décision sur le statut lexical des mots présentés lorsque ceux-ci étaient extraits des stimuli de la condition occlusive-liquide utilisés dans l'Expérience 1 que lorsque ceux-ci étaient extraits des stimuli des conditions occlusive-fricative et occlusive-occlusive. Cet effet s'exprime aussi bien dans les latences de détection que dans les taux d'erreur. Le sens de l'effet est cependant l'inverse de celui qui est observé dans l'Expérience 1 avec la tâche de word-spotting.

2.3. Discussion

Dans l'Expérience 1, l'effet d'alignement entre frontières syllabique et lexicale se manifeste malgré les contraintes lexicales que nous nous sommes imposées pour le choix du

matériel expérimental. En effet, les stimuli que nous avons utilisés ne permettaient pas l'émergence d'un effet qui serait uniquement déterminé par un déséquilibre dans les quantités de candidats lexicaux maintenus dans la cohorte selon que le groupe formé par la consonne finale du mot et par la consonne qui lui succède est tautosyllabique / fréquent ou hétérosyllabique / rare. Ainsi, l'allongement des latences de détection observé lorsque frontières phonologique et lexicale sont discordantes ne peut s'expliquer intégralement à partir de processus qui seraient déterminés par un déséquilibre des phénomènes de sélection lexicale sérielle dans les conditions expérimentales comparées. Par ailleurs, cet effet n'est pas reproduit tel quel dans une expérience contrôle de décision lexicale dans laquelle les participants doivent donner une réponse concernant le statut lexical de la portion C(C)VC initiale extraite des stimuli utilisés dans la tâche de *word-spotting*. Au contraire, les mots extraits des stimuli de la condition occlusive-liquide présentent une tendance à être plus facilement identifiés que ceux extraits des deux autres conditions expérimentales.

Cette observation fait cependant émerger un certain nombre de questions concernant les données observées dans la tâche contrôle (8.1 %, 17.6 % et 13.3 % d'erreurs pour les groupes occl.-liq., occl.-fric. et occl.-occl. et des latences de détection de 455 ms, 481 ms, 472 ms) et la signification de celles qui sont obtenues dans l'Expérience 1 (pas d'effet sur les taux d'erreur ; un effet sur les temps de réaction : 750 ms, 691 ms, 655 ms respectivement pour les groupes de consonnes occl.-liq., occl.-fric. et occl.-occl.). On peut envisager que l'accroissement des taux d'erreur et des latences de décision lexicale observé dans la seconde expérience soit lié à une altération acoustique des stimuli qui se serait produite lors de l'extraction des portions initiales effectuée pour la constitution du matériel. On pourrait alors envisager que cette altération ait pu masquer un effet acoustique qui aurait le même sens que l'effet d'alignement mis en évidence dans la tâche de *word-spotting*. Il serait alors impossible d'affirmer que l'effet d'alignement de l'Expérience 1 est réel ; il pourrait tout simplement constituer un effet acoustique lui-même masqué par un autre effet se manifestant dans le sens inverse. Nous ne pensons cependant pas que l'effet de l'expérience contrôle puisse s'expliquer par une altération des portions acoustiques correspondant au mot dans les stimuli des conditions occlusive-fricative et occlusive-occlusive, celle-ci ayant eu lieu lorsque nous procédions à l'extraction de la séquence C(C)VC initiale. Un certain nombre de travaux nous incitent en effet à penser que ces deux catégories de groupes devraient au contraire être plus faciles à segmenter à partir d'une représentation acoustique du signal. Lors de la constitution des stimuli de l'Expérience 2, il nous a semblé plus difficile de séparer à l'aide du logiciel de traitement du signal la portion acoustique correspondant au mot lorsque celui-ci était suivi d'une liquide que lorsqu'il était suivi d'une fricative ou d'une

occlusive. On trouve dans la littérature un certain nombre de travaux réalisés dans la mouvance du courant de la *phonologie articulatoire* (Browman & Goldstein, 1992) qui montrent que les groupes de consonnes tautosyllabiques sont réalisés par le biais de processus de contrôle articulatoire tout à fait différents de ceux qui sont à l'œuvre pour la production des groupes hétérosyllabiques. Browman & Goldstein (1988) ont étudié divers indices dérivés de données obtenues en cinéradiographie et ont comparé des groupes de consonnes identiques mais dont la structure syllabique variait. Ils ont par exemple observé des données articulatoires pour un groupe comme /sp/ selon que celui-ci était prononcé à l'intérieur d'une syllabe ou dans deux syllabes différentes. Ils ont concentré leur étude sur des séquences $VC_1C_2VC_3$ intégrées dans des suites plus longues et ont mesuré, à partir de l'enregistrement des mouvements des articulateurs,

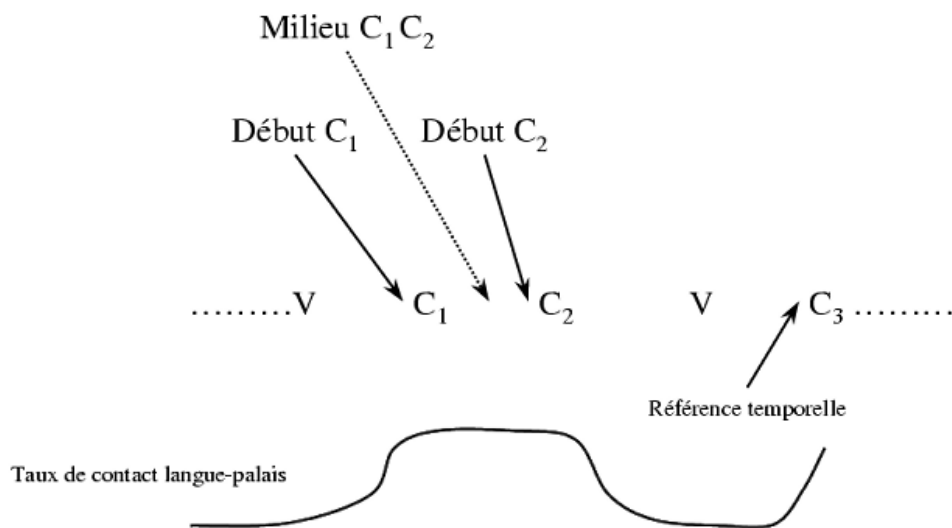


Figure 30 : Schéma représentant les résultats obtenus par Browman & Goldstein (1988) dans leur étude sur l'organisation articulatoire des séquences de consonnes en fonction de leur structure syllabique. La courbe du bas représente une mesure articulatoire des taux de contact entre langue et palais.

les durées qui séparent la consonne finale (C_3 , la référence) des phonèmes qui constituent le groupe C_1C_2 (cf. Figure 30). A partir de ces données de durée, ils ont mesuré la variabilité des données en fonction du point de mesure (début de C_1 , début de C_2 , milieu de C_1C_2).

Ils observent que les durées séparant les groupes de consonnes hétérosyllabiques de la consonne suivante présentent une variance minimale lorsque l'intervalle temporel est mesuré entre le début de C_1 et la référence. Au contraire, l'intervalle observé pour les groupes tautosyllabiques présente une variance minimale lorsque la mesure est effectuée entre le milieu de la séquence C_1C_2 et la consonne de référence. Cette observation est interprétée par les auteurs comme un signe de la cohésion articulatoire élevée des groupes tautosyllabiques dans le cadre des processus de production de la parole. Cette observation incite à penser que la production d'un groupe de consonnes en attaque de syllabe passe par une représentation unitaire de la paire

de consonnes alors que celle des groupes de consonnes hétérosyllabiques correspondrait à la juxtaposition de deux processus relativement indépendants pour chaque élément du groupe.

Les travaux de Byrd (1996) conduisent à des conclusions similaires à partir de l'analyse des recouvrements de gestes articulatoires en fonction du statut syllabique des séquences de consonnes. Les gestes correspondant à la production de chacun des phonèmes se recouvriraient beaucoup plus largement dans les groupes tautosyllabiques que dans le cas de la production de groupes hétérosyllabiques. Si ces phénomènes articulatoires ont aussi une manifestation acoustique spécifique, il serait certainement plus difficile de dissocier les composantes acoustiques correspondant aux segments de la séquence tautosyllabique que celles correspondant aux segments hétérosyllabiques. Ceci pourrait alors expliquer la difficulté que nous avons éprouvée à localiser la frontière entre les deux consonnes pour les stimuli de la condition occlusive-liquide. Cette difficulté a certainement influencé la possibilité d'extraire la portion correspondant au mot en évitant au maximum de supprimer des composantes acoustiques intervenant dans l'identification de celui-ci. Si c'est effectivement le cas, c'est à dire si les mots extraits de la condition occlusive-liquide ont été altérés par la procédure de découpage nécessaire à la mise en place de l'expérience contrôle et que l'effet observé dans la tâche de décision lexicale devait trouver sa source dans cette altération, nous aurions dû observer l'effet inverse : les stimuli extraits des séquences de la condition occlusive-liquide donnant lieu à des taux d'erreur plus importants et à des latences de détection plus longues. Il est donc peu probable que l'effet observé dans l'Expérience 1 puisse s'expliquer par des paramètres uniquement acoustiques qui auraient été masqués par une altération des stimuli extraits de la condition occlusive-liquide dans la tâche contrôle de décision lexicale.

Il est cependant possible de proposer une piste pour interpréter l'effet observé dans la tâche de décision lexicale. Il nous semble que les occlusives prononcées devant une liquide ont moins tendance à subir un phénomène qui pourrait s'apparenter à une réduction acoustique de la consonne. Ces phénomènes de réduction acoustique sont couramment observés en production spontanée. Ils consistent notamment, dans le cas des consonnes aussi bien que des voyelles, à modifier l'intensité et la composition spectrale de la portion acoustique subséquente à l'occlusion (Van Son & Pols, 1999). Ce phénomène se manifeste notamment lorsque les locuteurs sont dans une situation dans laquelle les processus de production de la parole sont moins contrôlés. Les caractéristiques d'une consonne réduite varient en fonction du type de consonne. Les fricatives et les occlusives réduites se caractériseraient par une diminution du rapport d'intensité consonne / voyelle. Cette réduction correspondrait également à une diminution du centre de gravité spectral pour les fricatives ainsi que pour les nasales (Van Son &

Pols, 1999). Cette réduction consonantique rend plus difficile l'identification de la consonne dans une tâche d'identification phonémique (Van Son & Pols, 1997). Lorsque nous avons procédé au découpage des stimuli afin d'en extraire les séquences C(C)VC initiales, il nous est apparu que l'occlusive finale des mots, qui est à l'initiale des groupes testés, était parfois difficilement visible sur la représentation spectrographique. Il est possible que la proportion de stimuli dans lesquels la consonne finale presque absente du signal soit plus importante parmi ceux des conditions occlusive-fricative et occlusive-occlusive que pour les items extraits de la condition occlusive-liquide. Nous n'avons pas conduit d'analyses acoustiques consistant à extraire l'enveloppe spectrale ou l'intensité moyenne de la section correspondant aux occlusives extraites des trois types d'environnements. En effet, quelle que soit l'interprétation à donner de l'effet observé en décision lexicale, son inversion par rapport à ce qui est observé dans l'Expérience 1 suffit à valider l'interprétation contextuelle proposée. Les travaux présentés par Browman & Goldstein (1988) et par Byrd (1996; Byrd & Tan, 1996) contribuent à penser qu'une altération liée à la méthode de découpage des stimuli aurait dû produire les effets inverses. S'il y a effectivement réduction consonantique plus marquée dans les stimuli occlusive-fricative et occlusive-occlusive, cette réduction est également présente dans les stimuli qui ont été utilisés dans l'Expérience 1 et n'est pas induite par la procédure d'extraction de la portion initiale des non-mots. L'éventuelle réduction consonantique aurait donc dû influencer également les réponses mesurées dans la tâche de *word-spotting*. Tout nous incite donc à penser que l'effet d'alignement observé dans la première expérience trouve bien sa source dans un effet de contexte qui est déterminé par le type de consonne qui suit la portion initiale du stimulus. Cet effet pourrait éventuellement -mais pas seulement- se décrire en termes d'alignement ou de non-alignement des frontières syllabique et lexicale. Nous sommes cependant confrontés à un effet qui se limite à une partie du matériel, ce qui nous incite à examiner plus en détails les relations qui peuvent exister entre des influences de type prélexical (structure phonologique, fréquence) et lexical (nombre de compétiteurs).

3. Discussion générale

Si l'effet d'alignement observé dans la première expérience n'est probablement pas lié à des caractéristiques acoustiques des stimuli, il reste que la restriction de l'effet à une partie du matériel doit être analysée plus en détails. Contrairement à ce que nous avons supposé pour justifier la séparation en deux groupes de mots, une différence dans la variabilité des résultats en fonction du type de mots ne semble pas justifier la restriction de l'effet à un sous-ensemble des

items. En effet, on observe entre les mots qui ont beaucoup ou peu de compétiteurs des écart-types très similaires (respectivement 184 et 179 ms) alors que cette dichotomie contribue à faire émerger un effet significatif pour les mots qui ont beaucoup de compétiteurs. Il est donc probable que l'explication de cette restriction de l'effet ne soit pas statistique mais trouve sa source dans une relation particulière entre deux catégories d'informations qui sont supposées correspondre à deux niveaux de traitement des processus de segmentation de la parole en mots : un niveau prélexical à l'intérieur duquel interviendraient les connaissances sur les régularités phonologiques ou probabilistes de la langue et un niveau lexical dans lequel s'opèrent des phénomènes de compétition entre candidats lexicaux. Il serait alors délicat d'affirmer que l'effet observé découle de processus de traitement prélexicaux ayant pour objet de procéder à une segmentation du signal de parole préalable à la reconnaissance des mots. En effet, si la segmentation phonologique a pour fonction de faciliter la localisation des frontières de mots, elle doit intervenir avant tout recours à des processus de traitement lexicaux (compétition entre candidats notamment). Si un paramètre lexical (le nombre de compétiteurs) peut influencer le recours à un paramètre infralexical (la structure phonologique de la séquence), on est en mesure de proposer deux classes de modèles pour expliquer ce lien. La première possibilité est que la présentation d'une séquence de phonèmes correspondant à un mot qui a peu de compétiteurs dans le lexique va faciliter sa reconnaissance. Cette facilitation rendrait inutile le recours à des indices prélexicaux de segmentation. Une discussion plus approfondie de cette proposition et de ses implications sera présentée dans le Chapitre 7 en guise de réflexion sur les niveaux et les modes d'implémentation de processus reposant sur le recours aux contraintes phonologiques de la langue ; l'objet du Chapitre 6 est d'évaluer la validité d'une seconde proposition. Ce phénomène d'interaction entre variables lexicale et prélexicale pourrait en effet être conçu comme la mise en évidence du caractère lexical de ce que nous avons envisagé comme un effet prélexical. Il est en effet beaucoup plus probable d'observer un effet d'interaction entre des variables intervenant à un même niveau de traitement qu'à des niveaux différents.

Rappelons que le lien entre fréquence des groupes de consonnes et structure syllabique nous a conduits à proposer plusieurs interprétations de l'effet d'alignement. L'une d'entre elles permet d'adopter une conception purement lexicale de cet effet (cf. section 4.2.3 du Chapitre 4) selon laquelle un groupe tautosyllabique / fréquent déclencherait l'activation d'une quantité importante de candidats lexicaux. C'est cet accroissement du nombre de candidats lexicaux qui induirait un allongement des latences d'identification dans la tâche de *word-spotting*. Lorsque nous avons dichotomisé l'échantillon des mots utilisés dans l'Expérience 1, nous avons choisi une méthode de calcul du nombre de compétiteurs d'un mot qui relève d'un modèle de

reconnaissance pour lequel les débuts de mots remplissent un rôle essentiel. En effet, nous avons procédé à ce calcul en recensant pour chaque mot les éléments du lexique commençant par cette séquence de phonèmes. La valeur que nous avons obtenue est donc un comptage du nombre de candidats qui sont maintenus dans cohorte initiale du modèle COHORT (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987) lorsque le phonème final du mot est en cours de traitement. Nous avons par conséquent choisi de manière arbitraire un mode de calcul qui accentue l'importance des débuts de mot. Lorsque nous avons envisagé que des groupes tautosyllabiques pourraient générer l'activation d'une quantité importante de candidats lexicaux en raison de leur fréquence d'occurrence élevée, nous avons fait référence à une conception différente de cette variable 'nombre de candidats lexicaux'. Cette dernière a alors été envisagée en rapport avec le modèle TRACE (McClelland & Elman, 1986), modèle pour lequel les phénomènes d'activation lexicale suivent une procédure d'alignement exhaustif sans que les débuts de mots se voient attribuer une importance particulière. Chaque phonème de la chaîne parlée pourrait alors générer l'activation d'un ensemble de candidats lexicaux quelle que soit la position du phonème dans les mots activés. Cette variable peut approximativement s'apparenter au 'nombre de voisins lexicaux' du *Neighborhood Activation Model* (NAM, Luce & Pisoni, 1998) et correspondrait au nombre d'unités lexicales qui subissent un accroissement de leur niveau d'activation dans le modèle TRACE (McClelland & Elman, 1986). Indépendamment du rôle essentiel de l'ordre des phonèmes, cet indice correspond au nombre de mots qui partagent 1 à N phonèmes avec le stimulus. Ainsi, dans le cadre d'un modèle à alignement initial, la séquence /vagrɜ/ donne lieu à l'activation de tous les mots commençant par /v/ ; ce qui correspond à 1040 mots dans la base de données BRULEX (Content et al., 1990). Le mot pertinent est alors progressivement sélectionné parmi les éléments de cette cohorte initiale. Cette sélection s'opère à partir d'un flux d'informations acoustico-phonétiques ascendant. Lorsque le phonème final /g/ est traité, il reste 15 éléments dans la cohorte de candidats. Dans le cadre d'un modèle à alignement exhaustif, le processus d'activation est très différent. De manière simplifiée, chaque phonème donne lieu à l'activation d'un ensemble indépendant de mots. Ces mots peuvent contenir le phonème en quelque position que ce soit. Ainsi, le /v/ de /vag/ activera tous les mots qui contiennent /v/ dans le lexique (3688 mots), puis tous ceux qui contiennent /a/ (14446 mots) et enfin les candidats qui intègrent le phonème /g/ (2940 mots). Les processus de compétition entre les divers candidats activés permettent d'aboutir à la reconnaissance du mot 'vague'⁴¹. Cette reconnaissance repose

⁴¹ En réalité, la taille du lexique utilisé dans les simulations du modèle TRACE (McClelland & Elman, 1986 ; voir aussi Frauenfelder & Peeters, 1990) est beaucoup plus modeste, ceci afin de faciliter la procédure de simulation.



-de même que dans COHORT- sur un flux ascendant d'informations acoustico-phonétiques mais aussi sur des phénomènes d'inhibition entre candidats lexicaux.

Même si cette procédure rend mieux compte de la reconnaissance des mots en parole continue en raison de l'existence de mots ayant un Point d'Unicité tardif (cf. Chapitre 2), le décours temporel du signal de parole incite cependant à ne pas abandonner totalement une composante séquentielle du traitement pour laquelle les activations lexicales seraient au moins partiellement alignées sur les débuts de mots. (cf. Mattys, 1997 pour une revue). Ces deux catégories d'alignement des activations pourraient donc constituer deux facettes du système de reconnaissance des mots en modalité auditive. On pourrait ainsi envisager un modèle dans lequel les procédures d'activation lexicale suivraient un alignement exhaustif mais où les candidats activés à partir d'un phonème présentant une forte probabilité de constituer un début de mot subiraient un accroissement de leur activation. Ceci accentuerait le poids du traitement séquentiel effectué sur le signal de parole tout en laissant la possibilité de faire intervenir des processus de reconnaissance fondés sur un alignement exhaustif afin de maintenir la possibilité de retarder la reconnaissance d'un mot s'il a par exemple un Point d'Unicité tardif.

On peut alors imaginer que les processus influencés par le nombre de compétiteurs du mot à partir de son début vont être sensibles à un accroissement du nombre de candidats activés en d'autres endroits du signal. Si l'on peut concevoir un modèle de la reconnaissance des mots dans lequel ces deux variables seraient relativement orthogonales, il devient possible d'envisager que la restriction de l'effet du type de groupe consonantique observée dans l'Expérience 1 trouve sa source dans une interaction entre deux variables lexicales plutôt qu'entre une variable lexicale et une variable infralexicale. L'effet d'alignement observé dans l'Expérience 1, de même que ceux mis en évidence par McQueen (1998) et par Vroomen & De Gelder (1999), pourrait alors se réduire à un effet du nombre de candidats lexicaux activés par le groupe de phonèmes constitué du phonème final du mot et du phonème contextuel. Cette interprétation remettrait en cause un modèle selon lequel les processus de segmentation lexicale peuvent avoir recours à des connaissances sur les régularités phonologiques de la langue.

Mais on peut considérer qu'un système de reconnaissance des mots reposant sur ces principes activerait approximativement cette quantité de mots.

Résumé

Les expériences présentées dans ce chapitre avaient pour objet de confronter l'effet d'alignement entre frontières phonologique et lexicale observé par McQueen (1998) et par Vroomen & De Gelder (1999) à une interprétation lexicale selon laquelle les deux types de conditions d'alignement correspondraient en réalité à un déséquilibre du nombre de compétiteurs maintenus dans la cohorte lorsque le phonème contextuel est en cours de traitement. L'Expérience 1 a été conduite avec un paradigme de *word-spotting* dans lequel l'occurrence de la consonne contextuelle donnait lieu à une rupture nette dans le processus de dé-activation des candidats de la cohorte. La séquence 'mot' + consonne contextuelle ne correspondait en effet au début d'aucun mot dans la base de données BRULEX (Content et al., 1990). L'effet d'alignement est répliqué pour une partie seulement des mots de l'échantillon. Une tâche contrôle de décision lexicale permet d'affirmer que l'effet observé n'est pas déterminé par des propriétés acoustiques des stimuli utilisés dans la tâche de *word-spotting*. La restriction de l'effet aux mots ayant beaucoup de compétiteurs incite cependant à envisager que cet effet soit lié à une variable lexicale plutôt qu'infralexicale.