



Mémoire de recherche

L'accès au lexique chez les bilingues
Français-Anglais :
étude de l'influence du voisinage inter-
langue et de la typicalité orthographique

 **Manon Wattiau**

Master 1 Psychologie clinique du développement :
Evolution, involution et handicap
Université de Strasbourg

Remerciements

Mes remerciements s'adressent d'emblée à Mme Eva Commissaire qui, par son investissement et sa patience, a réussi à me transmettre le goût et l'intérêt pour la recherche. Elle a su répondre à mes questionnements et aiguiller ma pensée tout au long de ce travail. Ses nombreuses relectures m'ont permis de mieux diriger ma réflexion et de remettre en question mon raisonnement afin de répondre au mieux aux exigences de ce travail.

Je tiens également à remercier tous les participants de cette étude, et notamment mes collègues de Master 1, qui n'ont pas hésité à donner de leur temps, en démontrant un grand intérêt pour cette recherche.

Je terminerai par remercier mon compagnon Matthieu ainsi que Mme Isabelle Fornasieri, dont les judicieux conseils m'ont aidé à persévérer dans ce travail et à définir mes objectifs.

Table des matières

I.Introduction.....	4
II.Partie Théorique.....	5
II.1.Comment reconnaît-on les mots ?.....	5
<i>II.1.1 La lettre comme unité de base de la lecture.....</i>	<i>5</i>
<i>II.1.2 Le modèle à Activation Interactive de McClelland & Rumelhart.....</i>	<i>5</i>
<i>II.1.3 La tâche de décision lexicale, l'amorçage et le masquage.....</i>	<i>7</i>
II.2.Le fonctionnement des bilingues : l'accès au lexique.....	8
<i>II.2.1 Le voisinage orthographique, ses paramètres et conséquences : du monolingue au bilingue.....</i>	<i>11</i>
<i>II.2.1.1 La densité ou taille du voisinage orthographique.....</i>	<i>11</i>
<i>II.2.1.2 La fréquence du voisinage orthographique.....</i>	<i>13</i>
<i>II.2.2 Les effets du voisinage orthographique inter-langue.....</i>	<i>14</i>
<i>II.2.3 Le Bilingual Interactive Activation Model, Grainger & Dijkstra (1992) ; Dijkstra & Van Heuven (1998).....</i>	<i>17</i>
<i>II.2.4 Les effets de la typicalité.....</i>	<i>18</i>
II.3.Notre étude.....	20
III.Méthode.....	21
III.1.Participants.....	21
III.2.Matériel et stimuli.....	21
III.3.Procédure.....	24
IV.Résultats.....	25
V.Discussion.....	27
VI.Bibliographie.....	27
VII.Annexes.....	33
<i>Annexe 1 : Liste des items à traduire en anglais présentée aux participants.....</i>	<i>33</i>
<i>Annexe 2 : Questionnaire proposé aux participants.....</i>	<i>34</i>

<i>Annexe 3 : Tâche de lecture chronométrée en français présentée aux sujets :.....</i>	<i>35</i>
<i>lecture de texte et fiche d'évaluation.....</i>	<i>35</i>
<i>Annexe 4 : Liste des stimuli utilisés.....</i>	<i>39</i>
<i>Annexe 5 : Tableau des propriétés linguistiques des cibles pseudo-mots.....</i>	<i>41</i>

I. Introduction

*« Le mot est aussi important pour la psycholinguistique
que ne l'est la cellule pour la biologie. »*

Balota (1994, p.303)

Comment un lecteur reconnaît-il un mot écrit ? La réponse à cette question anime de nombreux chercheurs depuis plus d'un siècle, James McKeen Cattell étant un des pionniers du sujet avec son article « The time taken up by cerebral operation » publié en 1886. En effet, la lecture est un talent cognitif étonnant, dont les processus sont automatisés, irrépessibles, non-conscients et surtout très rapides. L'introspection ne nous permet pas, en tant que lecteur, de prendre conscience des mécanismes à l'oeuvre dans l'identification des mots. C'est pour cela que les chercheurs établissent des techniques expérimentales spécifiques afin de mettre en lumière ces processus si complexes. Si nous avons aujourd'hui assez d'éléments pour comprendre comment un monolingue reconnaît un mot de sa langue, nous pouvons nous demander comment un bilingue fait-il pour distinguer les mots de deux langues lors de la lecture ? Chaque langue est gouvernée par des contraintes qui lui sont propres, relatives à la structure **orthographique**, **morphologique** et **phonologique** des mots. Par exemple, la combinaison « WH » n'existe pas en français, alors qu'elle est très commune en anglais (white, who, where, wish, etc.). Nous verrons que ces contraintes ont une influence sur les mécanismes d'identification des mots, en plus d'autres éléments tel que le voisinage orthographique que nous définirons plus loin. Nous pouvons à présent nous demander comment les représentations internes des mots de différentes langues sont-elles organisées dans notre cerveau ? Cette question nous amène à définir la notion de **lexique mental** désignant l'ensemble des représentations internes des mots (il comporte toute l'information phonologique, orthographique, morphologique, syntaxique et sémantique concernant les mots connus de la langue). Nous étudierons donc comment ce lexique est organisé et comment le sujet bilingue y accède.

II. Partie Théorique

II.1. Comment reconnaît-on les mots ?

II.1.1 La lettre comme unité de base de la lecture

Actuellement la position consensuelle concernant la reconnaissance des mots écrits est celle qui propose que l'unité de base de la lecture est la **lettre** en tant qu'identité abstraite¹ (Grainger & Jacobs, 1999 ; Jacobs & Grainger, 1994). En effet, reconnaître un mot c'est reconnaître les lettres qui le compose afin d'avoir accès au lexique mental. La reconnaissance des mots dépend donc de l'activation de l'identité de lettres abstraites. Donc pour reconnaître un mot, nous codons ses lettres constitutives ainsi que leurs positions, c'est ce qui nous permet de distinguer *LION* de *LOIN* par exemple. Nous pouvons à présent nous demander comment ce codage est réalisé, et quelles sont les étapes qui succèdent au codage. C'est ce que le modèle à Activation Interactive de McClelland et Rumelhart (1981) propose d'expliquer.

II.1.2 Le modèle à Activation Interactive de McClelland & Rumelhart

Ce modèle hiérarchique à activation parallèle et interactif de la reconnaissance des mots, que l'on nommera modèle IA (McClelland & Rumelhart, 1981) dans le présent travail (voir Figure 1), fait partie des premiers modèles computationnels « connexionnistes » des années 1980, utilisé chez les monolingues. Il utilise des représentations locales : chaque trait visuel, chaque lettre de l'alphabet et chaque mot est représenté par une unité distincte. Modèle à activation interactive signifie que la perception d'un mot est influencée par des informations de bas niveau (le niveau des lettres) mais aussi par des informations de haut niveau (niveau des mots). Il possède donc 3 niveaux : le niveau des traits visuels, des lettres et des mots (niveau des représentations orthographiques lexicales). Le premier niveau concerne la détection des traits visuels distinctifs (au nombre total de 14) et plus précisément les lignes horizontales, verticales et diagonales constituant les lettres de l'alphabet. Le second niveau code l'identification des lettres (26 unités lettres) associées à ces traits visuels. Enfin, le troisième niveau correspond à l'identification des mots par le lexique mental dans lequel il existe une unité par mot. Au sein de chaque niveau, les unités ne correspondant pas aux entrées activées seront inhibées.

¹ Cette identité est indépendante de la typographie, de la taille et de la forme (Evet & Humphreys, 1981)

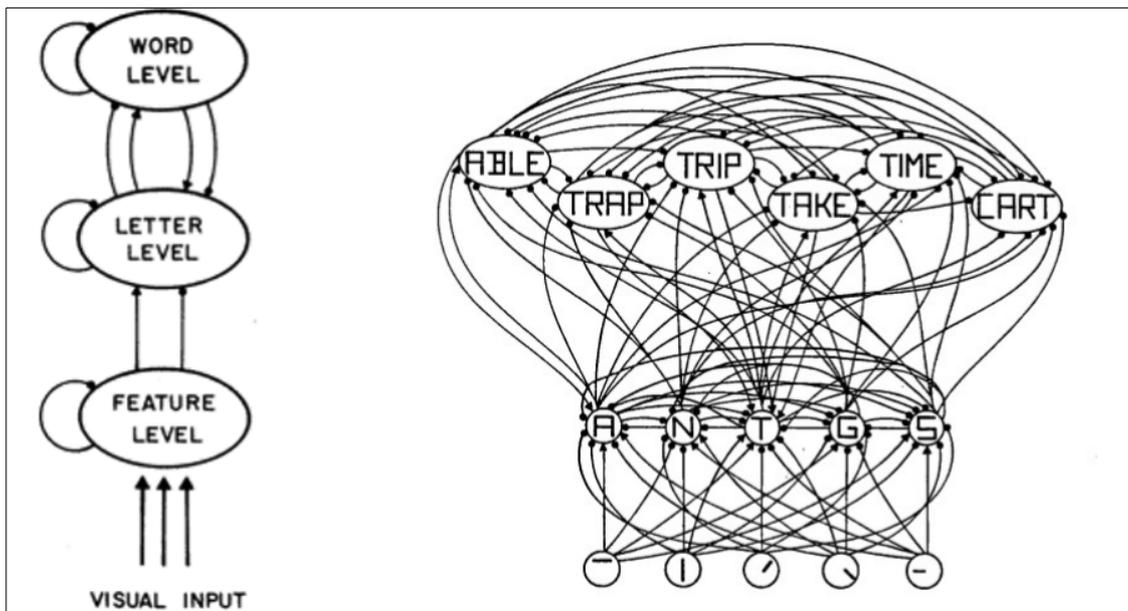


Figure 1. The Interactive Activation Model

En effet, il existe des connexions bi-directionnelles excitatrices et inhibitrices dans ce modèle. Par exemple, lorsque notre œil perçoit une lettre, les traits visuels de cette dernière vont activer les unités-trait correspondantes (niveau 1), lesquelles vont ensuite venir activer l'unité lettre compatible (niveau 2) et pour finir activer le mot correspondant (niveau 3), via des connexions excitatrices. De plus, la perception des lettres individuelles dans un mot est améliorée grâce à la rétroaction d'informations du niveau 3 des mots vers le niveau 2 des lettres. Il y a donc un transfert descendant de facilitation excitatrice vers le niveau des lettres, augmentant la probabilité que ces lettres soient choisies. Globalement, les connexions sont excitatrices entre les différents niveaux et inhibitrices à l'intérieur des niveaux. Ce mécanisme d'inhibition intra-niveau va particulièrement nous intéresser dans cette étude. En effet, au niveau des mots (niveau 3 lexical), il existe un mécanisme d'inhibition lexicale entraînant une inhibition mutuelle de tous les mots candidats (inhibition des mots entre eux), afin d'accélérer le processus de reconnaissance du mot cible. Ce phénomène vient introduire la notion de **compétition lexicale** entre les mots. Plusieurs candidats seraient activés et nécessiteraient d'être inhibés afin qu'un seul stimulus soit sélectionné. Il est important de préciser que dans ce modèle, chaque unité lexicale a un niveau d'activation « au repos » qui est toujours plus élevé pour les mots de haute fréquence d'apparition dans le langage que pour les mots de basse fréquence. C'est le principe qui sous-tend l'**effet de fréquence**, qui affecte le processus de reconnaissance, et au sein duquel les mots fréquents sont reconnus plus rapidement que les mots dits rares.

C'est à partir de ce modèle que Grainger et Dijkstra (1992) puis Dijkstra et Van Heuven (1998) élaborent le *Bilingual Interactive Activation Model* (appelé BIA) que nous étudierons plus loin, afin de décrire les étapes et mécanismes en jeu dans la reconnaissance des mots écrits chez les bilingues. Les études ayant mené à l'élaboration de ce modèle ont utilisé une procédure expérimentale bien spécifique de la psycholinguistique, appelée la tâche de décision lexicale, que nous proposons d'explicitier dès à présent et qui sera utilisée dans la présente étude.

II.1.3 La tâche de décision lexicale, l'amorçage et le masquage

Cette tâche de lecture est très fréquemment utilisée par les psycholinguistes car très facile à mettre en place. Sa première utilisation par Rubenstein, Garfield et Milikan, date de 1970. Cette tâche consiste à décider le plus rapidement possible et le plus correctement possible si une suite de lettres présentée sur un écran est un mot ou non de la langue. Dans la plupart des expériences, les chercheurs étudient deux variables : le **temps de réaction** (correspondant au temps entre la présentation du stimulus sur l'écran et le début de la réponse du sujet) et le **taux d'erreurs** en pourcentage. Malgré la consigne, il se peut que certains sujets fassent tout de même des erreurs de décision, d'autant plus qu'ils peuvent manipuler la vitesse de leur réponse. En effet, s'ils répondent trop rapidement, ils vont commettre plus d'erreurs que s'ils répondent plus lentement. Posner (1986) a théorisé cette relation entre rapidité et précision par la **fonction rapidité/précision** représentée par une courbe. Il montre alors un biais personnel du sujet : si le sujet interprète la consigne avec une emphase sur la rapidité, alors il risque de commettre beaucoup d'erreurs et inversement, s'il interprète la consigne avec une emphase sur la précision, il commettra moins d'erreurs mais ses temps de réaction seront plus longs. C'est pour cela qu'il est important pour l'expérimentateur d'énoncer clairement la consigne de cette tâche, en mettant autant d'emphase sur la rapidité que sur la précision, et de vérifier que le sujet respecte bien les consignes en s'assurant qu'il n'existe pas de conflit entre rapidité et précision. Nous retiendrons deux facteurs principaux qui influencent le temps de réaction et le taux d'erreurs : le premier dépend de la consigne (et donc de la stratégie utilisée par le sujet) et le deuxième dépend des propriétés des mots et non-mots présentés.

Une technique très efficace qui peut être couplée à la décision lexicale est la **technique d'amorçage**. Elle consiste à présenter au lecteur un premier stimulus (appelé « amorce ») suivi d'un deuxième qu'il devra traiter (appelé « cible »). Concernant les études

citées dans le présent travail, l'amorce est présentée très rapidement, de manière subliminale (de 17 à 67 millisecondes), afin qu'elle ne puisse pas être traitée consciemment par le sujet. L'amorce laisse alors une trace résiduelle inconsciente qui influencera ou non la lecture de la cible. Cette technique permet d'étudier l'influence (facilitation ou inhibition) d'une propriété donnée de l'amorce sur la cible, comparativement à une amorce « contrôle » neutre. Par exemple, est-ce que l'amorce « FRAISE » va faciliter ou inhiber la lecture de la cible « BANANE » comparativement à une amorce neutre « VIOLON » ? Plusieurs relations entre l'amorce et la cible peuvent être manipulées : relation visuelle (EPJO-FBIQ), orthographique (LION-LOIN), phonologique (RANG-REND), syntaxique (UNE-CHAISE) ou encore sémantique (comme dans notre exemple). Mais il est aussi possible de présenter une amorce pendant des temps plus longs (supérieurs à 150 millisecondes), cependant les sujets peuvent alors avoir la possibilité d'utiliser des stratégies de prédiction, car ils remarquent la nature de la relation entre l'amorce et la cible. Si l'objectif est d'étudier des traitements inconscients et automatisés (comme dans notre étude), il est plus adapté d'utiliser des temps de présentation de l'amorce très courts (jusqu'à 67 millisecondes) ainsi qu'un autre procédé appelé le **masquage** (Evelt & Humphreys, 1981 ; Forster & Davis, 1984). Cela consiste à limiter la visibilité de l'amorce en ajoutant un masque (une série de dièses ##### la plupart du temps) précédant ou suivant le stimulus amorce. Les effets observés dans cette condition d'amorçage avec masquage sont interprétés par le fait que le traitement de l'amorce laisse un effet résiduel qui va influencer ou non le traitement de la cible qui suit. Etant donné la haute proximité temporelle entre l'amorce et la cible, ces effets sont transitoires. Grâce à ce procédé, nous pouvons donc étudier des processus de traitement automatisés, non-conscients et irrépressibles sans que le sujet puisse développer des stratégies de réponse (ce que nous voulons à tout prix éviter).

Etudions à présent comment, grâce à l'utilisation de la tâche de décision lexicale avec amorçage, les chercheurs ont réussi à mettre en évidence les mécanismes d'accès au lexique qui permettent la reconnaissance d'un mot, chez les individus bilingues.

II.2. Le fonctionnement des bilingues : l'accès au lexique

Il existe deux types d'hypothèses concernant l'accès au lexique chez les bilingues : une première plaidant l'existence d'un lexique intégré et commun aux deux langues et une deuxième défendant l'indépendance des lexiques (respectifs à chaque langue). Passons à présent en revue les quatre principaux modèles proposés pour expliquer la reconnaissance des

mots chez les bilingues (voir Figure 2, d'après Van Heuven, Dijkstra & Grainger, 1998).

Le modèle A étudié par Macnamara et Kushnir (1971) ; Scarborough, Gerard, et Cortese (1984) et enfin Soares et Grosjean (1984) expose l'hypothèse d'un accès sélectif au langage avec un lexique indépendant. Chez le bilingue, le lexique de la langue lue serait directement activé grâce à un mécanisme de sélection appelé «input switch». C'est un système très sélectif qui met en relation l'input langagier au niveau phonologique et orthographique avec son lexique associé, dans une seule langue. Si aucune correspondance n'est trouvée dans le lexique activé, le système va rechercher dans l'autre lexique correspondant à celui de la seconde langue.

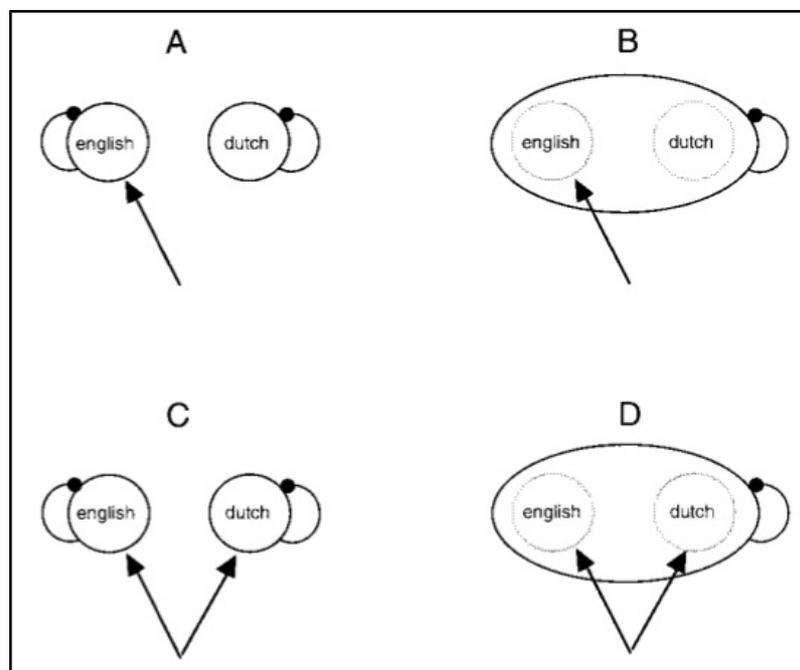


Figure 2. Quatre points de vue théoriques sur la reconnaissance des mots chez les bilingues : (A) accès sélectif au langage, lexiques indépendants ; (B) accès sélectif, lexique intégré ; (C) accès non-sélectif au langage, lexiques indépendants ; (D) accès non-sélectif au langage, lexique intégré. Les courbes se terminant par un point noir indiquent des connexions inhibitrices.

L'hypothèse B propose également un accès sélectif mais cette fois-ci associé à un lexique intégré. Elle est fonctionnellement similaire à l'hypothèse A. Les mots de la langue qui n'a pas été sélectionnée n'auront donc aucune influence sur ceux de la langue activée selon ces deux modèles. De nombreux auteurs dans leurs études initiales ont tenté de prouver l'hypothèse d'un accès sélectif (hypothèses A et B) dont Soares et Grosjean (1984). Ils ont utilisé une tâche utilisant des « code-switched words » (c'est-à-dire insérer des mots d'une autre langue dans une phrase), afin de répliquer les résultats précédents de Macnamara et Kushnir (1971). Avec des monolingues Anglais et des bilingues Portugais-Anglais, ils testent

trois conditions : (1) English : *After lunch, the children asked for a piece of cake for dessert.* (2) Portuguese : *Depois do almoco os miudos pediram uma fatia de bolo para sobremesa.* (3) Code-switched : *Depois do LUNCH os miudos pediram uma fatia de CAKE para DESSERT.* Ils trouvent que les bilingues répondent plus lentement dans la troisième condition où les deux langues sont mélangées au sein d'une même phrase, comparativement aux deux premières conditions dites monolingues. De plus, lorsqu'ils proposent aux participants d'identifier des non-mots dans cette même tâche, ils observent que les bilingues sont plus lents que les monolingues dans les trois conditions. Ils interprètent ces résultats dans le sens de l'hypothèse d'un **accès sélectif au langage** : un délai est observé car les sujets doivent changer de lexique grâce au commutateur « input switch ». Macnamara et Kushnir (1971) avaient déjà trouvé que les bilingues mettaient plus de temps pour passer d'une langue à l'autre pendant la lecture silencieuse.

Les modèles C et D proposent un **accès au lexique non sélectif à la langue**. Ce qui veut dire que pour reconnaître un mot, tous les candidats potentiellement compatibles avec le stimulus sont activés en parallèle, quelle que soit leur langue d'appartenance. Le modèle C (Forster, 1976) présente un accès non sélectif où chaque lexique est indépendant. Les mots d'une première langue sont analysés avant ceux d'une deuxième langue, la recherche est donc organisée en fonction de la langue. Dans ce modèle, il n'y a pas de connexions inhibitrices entre les mots des différentes langues (donc pas d'inhibition lexicale). Enfin, le dernier modèle D, dont s'inspire la présente recherche, propose l'hypothèse d'un **accès non sélectif avec un lexique intégré**. Selon ce modèle il existerait des connexions inhibitrices entre les mots de différentes langues (mécanisme de compétition lexicale induisant de l'inhibition lexicale). Ce mécanisme de compétition lexicale viendrait par contre allonger le temps de reconnaissance du mot chez les bilingues, car les mots des différents lexiques langagiers entreraient en ligne de compte. Des preuves ont été apportées en faveur d'un accès non sélectif avec notamment l'utilisation de paradigmes d'interférence comme la tâche de Stroop mot-couleur (Dyer, 1971 pour l'étude princeps et Coderre, Van Heuven & Conklin, 2012) ou encore la tâche d'interférence image-mot (Ehri & Bouchard-Ryan, 1980), pour n'en citer que quelques-uns. Dyer (1971) a mis en évidence, grâce à une tâche de Stroop (interférence : couleur/image - mot) dans laquelle les participants doivent nommer le plus rapidement possible la couleur d'écriture des mots, un **effet d'interférence significatif** chez les **bilingues** par rapport aux monolingues. L'accès non sélectif a aussi été mis en évidence par des tâches de décision lexicale (Altenberg & Cairns, 1983 ; Nas, 1983). Altenberg et Cairns (1983) ont

montré que les représentations de la langue non ciblée sont activées durant la reconnaissance des mots chez les bilingues.

L'accès non sélectif au lexique a aussi été mis en évidence grâce aux effets du **voisinage orthographique**, que nous proposons de présenter dans la section suivante. Nous reviendrons sur les modèles C et surtout D après cette section sur le voisinage, afin d'apporter des arguments qui permettront de les départager.

II.2.1 Le voisinage orthographique, ses paramètres et conséquences : du monolingue au bilingue

Le voisinage orthographique est un élément essentiel influant la reconnaissance des mots. Nous distinguerons les effets de la taille ou densité du voisinage orthographique des effets de fréquence du voisinage orthographique. Ces facteurs influencent tous deux la reconnaissance des mots de manière significative. Commençons par définir le voisinage orthographique : « *c'est l'ensemble des mots de même longueur partageant toutes les lettres sauf une à la même position avec un mot cible* » (Coltheart, Davelaar, Jonasson & Besner, 1977). Par exemple le mot PAGE possède trois catégories de voisins (en fonction de la position de la lettre intervertie) : NAGE, RAGE et CAGE pour la position 1 – PIGE pour la position 2 et PAIE, PALE, PAPE pour la position 3. A l'inverse, un « ermite » est un mot sans aucun voisin orthographique (ex : DRAP).

En ce qui concerne les bilingues, le **voisinage orthographique inter-langue** (qui sera étudié dans notre étude) est relatif aux voisins orthographiques des deux langues. Par exemple le mot français RIRE possède 3 voisins anglais ; FIRE, HIRE et RIDE. Nous distinguerons du voisinage inter-langue le **voisinage partagé**, qui signifie l'existence d'un voisin commun entre l'amorce et la cible (Van Heuven, Dijkstra, Grainger & Schriefers, 2001). Par exemple NARE et GARE ont comme voisin partagé MARE. Il existe également des voisins partagés entre deux langues, que l'on nommera le voisinage partagé inter-langue. Par exemple le mot anglais BEAN et son voisin français BEAU ont comme voisin partagé inter-langue le mot anglais BEAR.

II.2.1.1 La densité ou taille du voisinage orthographique

Coltheart et al. (1997) dans leur expérience princeps chez les monolingues, démontrent que les non-mots ayant beaucoup de voisins orthographiques (entre 6 et 20), entraînent des temps de réaction à une tâche de décision lexicale significativement plus longs

que des non-mots ayant peu de voisins (entre 0 et 4). Chez les bilingues, Van Heuven, Dijkstra et Grainger (1998) montrent dans une tâche de décision lexicale avec démasquage progressif que plus un mot a de voisins orthographiques² (et ce dans les 2 langues), plus le sujet met de temps à l'identifier. Dans leur étude de 2001, Van Heuven, Dijkstra, Grainger et Schriefers manipulent la taille du voisinage orthographique partagé entre les amorces et les cibles, au sein de la même langue (le néerlandais). Ils utilisent aussi une tâche de décision lexicale avec amorçage masqué. Les amorces étaient orthographiquement reliées aux cibles, c'est-à-dire qu'elles différaient uniquement d'une lettre (ex : HOLK-DOLK), par contre une catégorie d'amorces partageait au moins un voisin orthographique avec la cible (holk-DOLK partageant les voisins KOLK-VOLK-WOLK-TOLK) alors qu'une autre n'en partageait aucun (dols-DOLK ne partagent aucun voisin). Ils ont aussi observé que les sujets mettaient significativement moins de temps à répondre lorsque les amorces ne partageaient aucun voisin orthographique avec les cibles. Ils ont ici mis en évidence un **effet inhibiteur** de la taille du voisinage orthographique partagé entre l'amorce et la cible : plus il y a de voisins partagés, plus les sujets répondent lentement.

La nature facilitatrice ou inhibitrice des effets du voisinage orthographique a des implications théoriques importantes concernant les modélisations des mécanismes traitement impliqués dans la lecture : un **effet inhibiteur** soutiendrait l'hypothèse d'un processus de sélection avec **compétition lexicale** (cf. Modèle IA de McClelland & Rumelhart, 1981 et Modèle BIA, Grainger & Dijkstra, 1992 ; Dijkstra & Van Heuven, 1998) et donc l'hypothèse d'un accès non sélectif avec lexicale intégré, alors qu'un effet de facilitation supporterait l'existence d'un mécanisme sans compétition lexicale pour la sélection. En effet, selon le Modèle BIA que nous verrons plus loin, les voisins orthographiques d'un mot entreraient en compétition pour leur sélection, au niveau du lexique mental, entraînant un mécanisme **d'inhibition lexicale** intra-niveau.

De nombreux chercheurs ont alors investigué ce paramètre et trouvent des résultats **contradictoires** en ce qui concerne les effets du voisinage orthographique dans la tâche de décision lexicale : certains trouvent des effets de facilitation de la taille du voisinage orthographique pour les mots (Andrews, 1989, 1992 ; Forster & Shen, 1996 ; Sears, Hino & Lupker, 1995), pour n'en citer que quelques-uns, d'autres ne trouvent pas d'effet facilitateur

2 Nous savons que plus un mot est long, moins il possède de voisins orthographiques. La plupart des recherches que nous étudieront utilisent des mots relativement courts de 4 ou 5 lettres afin de pouvoir contrôler leur voisinage.

(Carreiras, Perea & Grainger, 1997 dans leur expérience 2; Coltheart et al., 1997) et d'autres encore trouvent même un effet inhibiteur dû à la fréquence du voisinage orthographique et non pas à la taille de ce dernier. Pour reprendre un des résultats montrant un effet facilitateur de la taille du voisinage orthographique, Andrews dans son expérience de 1989, avait manipulé le nombre de voisins orthographiques, appelé N, pour des mots anglais de **basse fréquence** uniquement. L'effet facilitateur qu'elle avait trouvé n'était donc valable que pour les mots de basse fréquence. D'autres auteurs ont alors commencé à étudier le paramètre de **fréquence** du voisinage orthographique, qui apparaît alors comme un paramètre important à prendre en considération dans la manipulation du voisinage orthographique.

II.2.1.2 La fréquence du voisinage orthographique

Il s'agit de la fréquence d'apparition des voisins orthographiques dans le lexique, mesurée en occurrences par million (o.p.m.). Andrews (1989) ainsi que Grainger, O'regan, Jacobs, et Segui (1989), ont étudié simultanément la taille du voisinage orthographique et sa **fréquence** en français. Dans une tâche monolingue de décision lexicale sans amorçage chez des participants français, ils ont montré que quelle que soit la taille du voisinage (N), les temps de décision lexicale étaient significativement plus longs pour les mots ayant au moins un voisin orthographique plus fréquent comparé aux mots n'ayant pas de voisins orthographiques plus fréquents (par exemple le mot NERF ayant 3 voisins orthographiques dont 1 de haute fréquence, entraînait des temps de réaction plus longs que le mot JUPE ayant aussi 3 voisins orthographiques dont aucun de haute fréquence). Ils concluent donc que la présence dans le voisinage orthographique d'au moins un voisin de plus haute fréquence que le mot stimulus induit une interférence dans le processus de traitement du mot stimulus. Cet effet est aussi expliqué par le phénomène de **compétition lexicale** (Grainger & Jacobs, 1996 ; McClelland & Rumelhart, 1981).

Ces effets d'inhibition de la fréquence du voisinage initialement obtenus par Grainger et al., (1989) avec une tâche de décision lexicale ont été répliqués dans d'autres langues (hollandais, français, espagnol), et les chercheurs ont obtenu des résultats **contradictaires en anglais**. Certains auteurs (Huntsman & Lima, 1996 ; Perea & Pollatsek, 1998 ; Ziegler & Perry, 1998) trouvent des effets d'inhibition en anglais, tout comme Grainger et al. (1989) alors que d'autres observent un effet de facilitation (Huntsman & Lima, 2002 ; Sears et al., 1995). Certains ne trouvent aucun effet (Forster & Shen, 1996 ; Sears et al., 1995, 2006). Dans sa revue de la littérature, Andrews (1997) relate que l'effet d'inhibition de la fréquence

du voisinage orthographique a été observé essentiellement en français, hollandais et espagnol, mais pas en anglais. A l'inverse, l'effet de facilitation de la taille du voisinage orthographique a été obtenu essentiellement en anglais mais pas en français, espagnol ou hollandais. Selon cette même auteure, le mécanisme d'inhibition lexicale serait différent en fonction des langues : plus important pour le français, le hollandais ou encore l'espagnol et plus faible pour l'anglais. En effet, ces effets de facilitation tiendraient leur origine de l'influence facilitatrice de la rime orthographique, dont l'influence joue un rôle important en anglais (Ziegler & Perry, 1998) par rapport au français, à l'espagnol ou au hollandais. Rappelons que la rime orthographique est une rime consistant en des mots orthographiquement similaires mais qui ne riment pas quand ils sont prononcés (par exemple en anglais SAVE et HAVE) ce qui est le cas dans les langues dites opaques ou irrégulières. Mais Grainger & Jacobs (1996) n'adhèrent pas à ce point de vue et proposent alors un autre modèle général expliquant ces différents effets (de facilitation et d'inhibition de la taille et de la fréquence du voisinage orthographique) associés au type de la tâche, de la consigne et de la nature des non-mots dans l'expérience. Ils appellent ce modèle le MROM (Multiple Read-Out Model). Notons, afin de faire un résumé, que les effets du voisinage orthographique semblent varier en fonction (1) de la tâche utilisée ; (2) selon que l'on manipule la taille ou la fréquence du voisinage ; (3) de la langue considérée.

Nous avons donc vu que la densité ainsi que la fréquence du voisinage orthographique influait de manière importante la reconnaissance des mots. Cet effet inhibiteur va en faveur de l'hypothèse non-sélective de l'accès au langage. Qu'en est-il des effets de voisinage orthographique inter-langue, lorsque amorces et cibles sont de deux langues différentes ? C'est ce que nous proposons d'examiner à présent, et que nous allons étudier dans cette étude.

II.2.2 Les effets du voisinage orthographique inter-langue

Chez les bilingues, les effets du voisinage inter-langue sont des paramètres particulièrement pertinents pour mettre en lumière les principes d'organisation et d'accès au lexique. Rappelons qu'il existe à ce niveau 2 types d'hypothèses : celles concernant le type d'accès au lexique : sélectif versus non-sélectif à la langue et celles concernant l'indépendance de ce lexique : intégré versus indépendant. Selon l'hypothèse d'un accès sélectif au langage (modèles A & B cités plus haut) où seul le lexique du mot cible est activé, la reconnaissance du mot cible sera déterminée par les caractéristiques des voisins de cette langue donnée uniquement. Par contre, selon le modèle de l'accès non-sélectif avec lexique

intégré (modèle D), la reconnaissance d'un mot cible va venir activer les voisins orthographiques de plusieurs langues.

Afin de tester les hypothèses C et D à savoir, pour un accès non-sélectif, la présence d'un lexique indépendant (C) ou intégré (D), la manipulation du voisinage orthographique inter-langue est tout à fait pertinente. En effet, selon l'hypothèse (D) d'un lexique intégré, la reconnaissance d'un mot cible dans une langue sera affectée par les voisins orthographiques de la langue ciblée mais aussi de la langue non ciblée. Le mot-stimulus présenté va activer des mots de plusieurs langues en même temps donc les voisins orthographiques respectifs de chaque langue vont entrer en compétition. En revanche, selon l'hypothèse d'un lexique indépendant (mais toujours dans le cadre de l'accès non-sélectif, modèle C), la reconnaissance du mot cible ne devrait pas être affectée par le voisinage orthographique de la langue non ciblée car il n'y a pas d'interaction directe entre les deux lexiques. On voit donc que la manipulation du voisinage orthographique est un élément majeur qui nous permettra de départager entre ces deux hypothèses : si le lexique est indépendant ou partagé. Van Heuven, Dijkstra et Grainger (1998) ont étudié à travers quatre expériences la nature du lexique (indépendant ou partagé). Dans leurs quatre expériences, ils ont contrôlé la taille respective du voisinage Anglais et Néerlandais des mots de manière à ce que chaque participant soit confronté aléatoirement à quatre conditions : la taille du voisinage Anglais et Néerlandais est grande – celle du Néerlandais est petite alors que celle de l'Anglais est grande – celle du Néerlandais est grande alors que celle de l'Anglais est petite – les deux sont petites.

Dans une première expérience, ils font appel à des participants bilingues Néerlandais-anglais pour une tâche de démasquage progressif. Les participants voient une série de mots cibles anglais puis une autre de mots néerlandais (ou inversement). Ils mettent en évidence un effet inhibiteur, pour la reconnaissance des mots anglais (L2) donc la langue ciblée, de la densité du voisinage orthographique néerlandais : plus il y a de voisins orthographiques néerlandais associés au mot cible anglais, plus les participants mettent du temps à répondre. La reconnaissance des mots cibles est donc très influencée par le nombre de voisins orthographiques de la langue non ciblée. Ces résultats sont en accord avec ceux précédemment trouvés par Beauvillain (1992) ainsi que Grainger et Dijkstra (1992) et mettent en évidence l'existence d'un lexique intégré au sein duquel les unités des deux langues sont inter-connectées via des liens inhibiteurs. Cet effet inhibiteur était également observé lorsque les mots des deux langues étaient mélangés (Expérience 2). Ils trouvent le même résultat dans

leurs expériences 3 et 4 avec une tâche de décision lexicale : un effet inhibiteur de la densité du voisinage néerlandais sur l'identification de mots cibles anglais.

L'hypothèse d'un accès non-sélectif à un lexique intégré est donc vérifiée, étant donné qu'on peut observer à maintes reprises chez les bilingues, que l'identification d'un mot est fortement influencée par la densité du voisinage orthographique de la langue non ciblée (à cause du phénomène d'inhibition latérale). Toutes ces données viennent prouver que la perception d'un mot active automatiquement les voisins orthographiques des deux langues du bilingue. Mais la reconnaissance d'un mot chez l'individu bilingue est aussi influencée par la fréquence du voisinage orthographique (e.g., Andrews, 1989, 1992; Carreiras, Perea, & Grainger, 1997; Grainger et al., 1989; Grainger & Jacobs, 1996; Grainger & Segui, 1990; Snodgrass & Mintzer, 1993). Bijeljac-Babic, Biardeau et Grainger (1997) ont alors apporté des preuves supplémentaires concernant **l'inhibition orthographique latérale inter-langue** chez les bilingues, avec des amorces de **haute fréquence**. Plus précisément, ils ont démontré que des amorces françaises reliées orthographiquement aux cibles anglaises (donc des voisins orthographiques) engendraient des temps de réaction supérieurs, comparativement à des amorces et cibles non reliées. Par exemple les sujets mettaient plus de temps à répondre lorsque l'amorce française « BOUT » (de haute fréquence) précédait la cible « BOAT » (condition reliée) comparativement à l'amorce non reliée « AGIR » (de haute fréquence aussi). Ils ont également mis en évidence que la performance des bilingues dans leur seconde langue (L2) influait significativement sur le phénomène d'inhibition orthographique : lorsque les mots cible et amorce étaient de deux langues différentes, le phénomène d'inhibition s'accroissait en fonction de la performance du sujet dans la langue de l'amorce (L2). Plus les sujets étaient performants dans leur seconde langue, plus cet effet d'inhibition était prégnant³.

Nous avons donc vu que le voisinage orthographique inter-langue est un paramètre influent fortement la reconnaissance des mots chez les bilingues. Le modèle BIA (*Bilingual Interaction Activation Model*) est le seul modèle à prédire une influence du voisinage orthographique inter-langue avec un accès non sélectif au langage et un lexique intégré. D'après ce modèle, c'est le nombre relatif de voisins orthographiques de la langue ciblée et non ciblée qui détermine la rapidité avec laquelle l'activation du noeud langagier ciblé va dominer sur celui de la langue non ciblée. C'est à juste titre que nous utiliserons ce modèle en

3 Rappelons qu'un débutant dans l'apprentissage d'une seconde langue est beaucoup moins exposé aux mots de cette nouvelle langue, donc les mots utilisés en amorce auront toujours une fréquence basse pour ces sujets, malgré le fait que leur fréquence intrinsèque d'apparition dans la langue soit élevée.

tant que modèle théorique de base pour la présente étude.

II.2.3 Le Bilingual Interactive Activation Model, Grainger & Dijkstra (1992) ; Dijkstra & Van Heuven (1998)

Ce modèle partage les caractéristiques structurelles et fonctionnelles du modèle à activation interactive (McClelland & Rumelhart, 1981), voir Figure 3 ci-dessous.

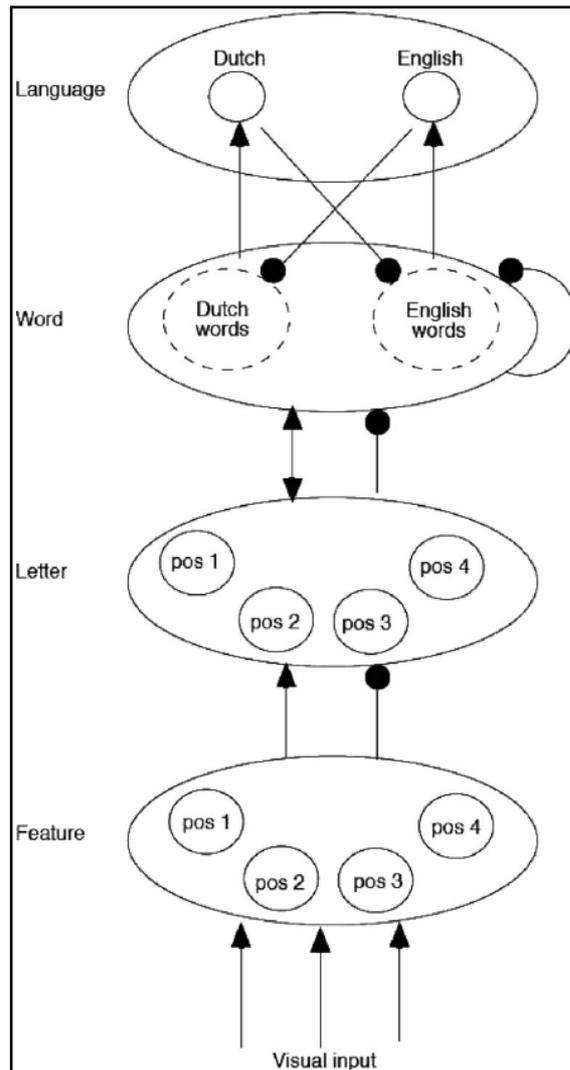


Figure 3. The Bilingual Interactive Activation Model, Grainger et Dijkstra (1992) ; Dijkstra et Van Heuven (1998)

Tous deux possèdent trois niveaux communs, mais ce modèle ajoute un quatrième et dernier niveau, celui des représentations langagières dans lequel on retrouve les nœuds langagiers « language nodes » qui ne sont plus de simples filtres langagiers mais des étiquettes ou représentations d'appartenances langagières (un nœud pour chaque langue). Ces nœuds

permettent la sélection d'un mot du langage ciblé du niveau 3 via une connexion descendante inhibitrice (inhibition du lexique non ciblé). Au sein du niveau des mots (le niveau 3), tous les mots des deux langues respectives peuvent s'inhiber entre eux (inhibition intra-niveau), indépendamment de la langue d'appartenance : c'est l'hypothèse de l'**accès au lexique non sélectif** à la langue. Les mots activés au niveau du lexique envoient alors leur activation au dernier niveau, celui des nœuds langagiers (niveau 4) de la langue correspondante. Ce nœud activé peut alors envoyer une inhibition descendante vers les mots de la langue qui n'a pas été sélectionnée. Le nœud langagier activé collecte les activations venant du niveau inférieur des mots de la langue correspondante et inhibe les mots de l'autre langue. On peut alors conclure que le degré d'activation des nœuds langagiers reflète la quantité d'activité de chaque lexique. Pour finir, ce modèle présume qu'un nœud langagier peut faciliter la sélection d'un mot du langage ciblé après, et seulement après, un accès non sélectif. Ce nœud ne peut donc pas faciliter la sélection depuis l'étape initiale de reconnaissance du mot, cela vient préciser que les mots de la langue non ciblée pourront toujours être activés. Ce modèle avance donc l'hypothèse d'un **accès non sélectif à la langue**, avec un **stockage lexical indépendant** pour chaque langue sous la forme d'un **lexique intégré** pour les mots de différentes langues.

Depuis, Dijkstra et Van Heuven (2002) ont étendu ce modèle en y intégrant des représentations phonologiques et sémantiques, qu'ils nomment le BIA+ afin de répondre aux limitations du modèle BIA, qui ne considérait que des représentations orthographiques. Le modèle BIA + démontre, tout comme le BIA, que lorsqu'un mot écrit est présenté à un bilingue, les deux lexiques sont activés. D'autres auteurs encore proposent une extension de ce modèle (Kesteren, Dijkstra & Smedt, 2012), et suggèrent que lorsque des indices orthographiques concernant la typicalité des mots sont présents, les lecteurs se serviraient de ces indices afin d'accéder au lexique de manière plus sélective. Ils ajoutent un niveau de traitement sublexical contenant les bigrammes entre autres.

II.2.4 Les effets de la typicalité

Le Français et l'Anglais partagent le même système orthographique et sont donc de ce point de vue très similaires (comparativement à l'Hébreu et à l'Anglais par exemple). Étudier deux langues partageant la même écriture permet donc de comprendre quels sont les indices ou les caractéristiques propres aux mots que les bilingues analysent afin de distinguer les deux langues. Chaque langue a ses propres règles ortho-tactiques et il se pourrait que les bilingues s'appuient sur ces règles pour reconnaître la langue des mots qu'ils lisent. Vaid et Frenck-

Mestre (2002) avaient montré que les bilingues Français-Anglais identifiaient plus rapidement la langue d'un mot quand ce dernier avait une orthographe spécifique à la langue. Westbury et Buchanan (2002) utilisent une nouvelle mesure ; la **fréquence minimale du bigramme**, à savoir la fréquence du bigramme le moins fréquent d'un mot, que nous utiliserons dans la présente étude. Ils montrent, avec une tâche de décision lexicale, que les mots dont la fréquence minimale du bigramme est basse étaient traités plus rapidement que ceux dont la fréquence minimale du bigramme est élevée, indépendamment de la taille du voisinage. Contrairement aux effets du voisinage orthographique qui n'apparaissent qu'avec des mots de basse fréquence (Andrews, 1992, 1999), l'effet de typicalité est apparu pour des mots de haute fréquence (et non pour des mots de basse fréquence). Ils ont montré que la mesure de la fréquence minimale du bigramme était très adéquate pour mesurer la typicalité orthographique, c'est pour cette raison que nous utiliserons cette mesure dans notre étude. Plus récemment, Casaponsa, Carreiras et Duñabeitia (2014) ont aussi retrouvé cet effet chez des bilingues Basque-Espagnol (la langue d'appartenance de mots typiques était reconnue plus rapidement). Ces résultats montrent donc que les bilingues utilisent des indices orthographiques sublexicaux (relatifs à la typicalité) afin d'accéder au lexique de la langue. Le lexique des bilingues serait alors organisé différemment en fonction du degré de typicalité (ou spécificité) des mots. Lorsque de tels indices sur la typicalité sont absents, les bilingues auraient un accès non-sélectif au langage, augmentant ainsi le degré d'interférence inter-langue. Il serait alors plus long d'accéder au lexique, ce processus demandant plus d'efforts à cause du mécanisme de compétition plus important. En revanche, lorsque des indices orthographiques sur la typicalité sont présents (donc des séquences de lettres spécifiques à la langue), les bilingues auraient un accès plus spécifique à la langue (plus sélectif), réduisant le degré d'interférence entre les langues et donc la compétition inter-langue. Dans leur étude de 2015, Casaponsa et Duñabeitia démontrent que l'accès au lexique chez les bilingues est fortement déterminé par les régularités orthographiques des deux langues du bilingue, et plus particulièrement que l'absence d'indices orthographiques concernant la typicalité d'une langue favorise un accès au lexique non spécifique au langage alors que la présence de ces indices réduit l'activation parallèle des deux langues (et donc le degré de compétition lexicale). Dans notre étude, nous étudierons ce paramètre, couplé à celui de la relation orthographique entre amorce et cible, ces deux paramètres n'ayant encore jamais été étudiés conjointement. L'effet de la typicalité inter-langue n'a pas non plus été étudié dans une tâche de décision lexicale avec amorçage masqué de la L1 (français) vers la L2 (anglais).

II.3. Notre étude

La reconnaissance et l'identification des mots chez les personnes bilingues serait sous-tendue par un processus d'accès au lexique non-sélectif à la langue dans lequel les deux lexiques du bilingue seraient intégrés au sein d'une unité commune. Cependant l'accès au lexique pourrait être modulé par des facteurs relatifs à la typicalité de séquences de lettres spécifiques à une langue. Partant de ce modèle théorique, notre objectif sera d'étudier l'accès au lexique chez des bilingues français-anglais avec une tâche de décision lexicale en anglais (afin d'étudier les effets de la L1 à la L2), à l'aide de deux variables d'intérêt : le voisinage orthographique inter-langue et la typicalité orthographique. Nous avons sélectionné des adultes Français ayant un niveau de compétence en anglais considérée comme faible (notons les effets d'inhibition dépendent également du niveau de performance dans la langue 2, Bijeljac-Babic, 1997). Plus particulièrement il sera question d'étudier comment la reconnaissance d'une cible anglaise chez les bilingues Français-Anglais peut être influencée par une amorce française reliée orthographiquement (un voisin orthographique inter-langue) ou non à cette cible, et comment un bigramme d'un mot (une séquence de deux lettres) typique de l'anglais mais rare du français, tel que « WR » peut influencer la reconnaissance des mots anglais dans cette tâche de décision lexicale. Nous nous attendons à ce que les amorces françaises reliées orthographiquement aux cibles anglaises inhibent leur reconnaissance (comme l'avaient démontré Dijkstra & Van Heuven Grainger, 1998), à cause du phénomène de compétition lexicale inter-langue. Les sujets devraient alors répondre plus rapidement, avec moins d'erreurs dans la condition non-reliée orthographiquement. Nous nous attendons également à ce que les mots anglais de la condition « spécifique » (avec des bigrammes fréquents en anglais mais pas en français) soient reconnu plus facilement par rapport à la condition « partagée ».

III. Méthode

III.1. Participants

Pour cette expérience, nous avons sélectionné 20 participants adultes (âge moyen : 28.75 (8.3)) dont la langue maternelle est le français, possédant des connaissances en anglais, mais qui ne se considèrent pas comme bilingues français-anglais. Nous qualifierons ces participants d'un niveau de type « low proficient » ou de faible compétence en anglais. Ces participants ont appris l'anglais en tant que seconde langue à l'école à partir de 11 ans et pratiquent toujours régulièrement l'anglais autant à l'écrit qu'à l'oral. Ils sont donc considérés comme des bilingues « tardifs ». Le niveau de chacune des langues (français et anglais) a été évalué, et ce de deux manières. Afin de contrôler leur niveau d'anglais, deux éléments ont été examinés : nous nous sommes assurées qu'aucun n'avait vécu dans un pays anglophone plus de deux mois, et leur niveau d'anglais a été contrôlé via un exercice de traduction français-anglais proposé à la fin de chaque passation (75 items à traduire, présentés en Annexe 1). Nous nous sommes également assurées que leur pratique de l'anglais était régulière, avec un questionnaire (présenté en Annexe 2). En moyenne pour la pratique orale des langues, les participants ont estimé qu'ils passaient 15.5 % de leur temps quotidien à parler ou entendre l'anglais, 76 % du temps à parler ou entendre le français et 8.4 % du temps à parler ou entendre une autre langue. En ce qui concerne le versant écrit, ils ont estimé qu'ils passaient 23.8 % du temps de leur temps quotidien à lire ou écrire l'anglais, 69.4 % à lire ou écrire le français et 6.8 % à lire ou écrire une autre langue. Afin de contrôler leur niveau de français, une tâche de lecture chronométrée d'un texte à voix haute leur a été proposée : ce texte provient du cahier de passation de l'Evaluation des Compétences de Lecture chez l'Adulte de plus de 16 ans (ECLA 16+), sur le thème de la pollution et d'un volume de 296 mots (présenté également en Annexes 3). Le nombre d'erreurs (nombre moyen d'erreurs : 1.45 (1.1)), le nombre de mots lus (nombre moyen de mots lus : 295.4 (0.6)) et le temps de lecture (temps moyen : 1.59 min (.28)) ont été comptabilisés durant la tâche. Enfin, nous préciserons que tous les participants ont appris d'autres langues dans leurs parcours (allemand, espagnol, italien, japonais, turque, arabe, portugais, luxembourgeois, etc.).

III.2. Matériel et stimuli

Le matériel de cette expérience est constitué de 80 cibles anglaises (40 cibles mots et 40 cibles pseudo-mots), ainsi que de 80 amorces françaises (40 amorces reliées et 40 non-

reliées). Au total il y a donc 160 paires d'amorces-cibles (la liste complète des stimuli est présentée en Annexe 4). Les moyennes et écarts-types pour chaque condition sont présentés dans le Tableau 1. Pour cette tâche de décision lexicale, 40 mots cibles anglais monosyllabiques d'une longueur variant entre 4 et 5 lettres (longueur moyenne = 4.3 lettres) ont été sélectionnés selon deux conditions : une première condition « spécifique » et une seconde « partagée », qui représentent différents degrés de typicalité orthographique avec le français. Ces mots cible sont issus de la base de données MCWords (Medler, D.A., & Binder, J.R., 2005).

La condition « spécifique » est constituée de mots anglais et pseudo-mots qui contiennent des séquences de lettres (bigrammes) peu typiques du français (par exemple le mot « WRAP » et le pseudo-mot « WROP », dont le bigramme « WR » est très peu typique du français). La condition « partagée » est constituée de mots anglais et de pseudo-mots qui contiennent des séquences de lettres plutôt typiques du français (par exemple le mot « TRAP » et le pseudo-mot « TRAN » qui contiennent des bigrammes typiques du français). Dans les deux cas les séquences de lettres sont typiques pour un Anglais. Afin de construire ces deux conditions, nous avons calculé la fréquence moyenne du bigramme le moins fréquent en français¹ : nous avons regroupé dans la condition « spécifique » les mots dont la moyenne des fréquences du bigramme le moins fréquent en français était la plus faible, comparativement aux mots de la condition « partagée » qui ont une moyenne des fréquences du bigramme le moins fréquent significativement plus élevée. En d'autres termes, ce qui différencie les deux conditions est uniquement la fréquence moyenne du bigramme le moins fréquent en français ($t(38) = 4.5, p < .001$). En ce qui concerne les autres variables appariées entre les deux conditions « partagée » et « spécifique », nous ne noterons aucune différence statistiquement significative concernant la longueur des mots (longueur moyenne condition « spécifique » : 4.25 (.44) ; longueur moyenne condition « partagée » : 4.3 (.47)), la fréquence lexicale ($t < 1$, n.s.), la taille du voisinage orthographique anglais de la cible ($t(38) = 1.08, p = .287$, n.s.) et sa fréquence ($t < 1$, n.s.), la taille du voisinage français de la cible ($t < 1$, n.s.) et sa fréquence cumulée ($t < 1$, n.s.) et enfin la fréquence lexicale moyenne des bigrammes en anglais ($t(38) = 1.129, p = .266$, n.s.). Tous ces paramètres sont donc comparables d'une condition à l'autre.

Les cibles étaient précédées d'amorces mots français. Les 80 amorces françaises ont été sélectionnées selon deux conditions : une première condition reliée orthographiquement (par exemple pour la cible anglaise « BEAN » l'amorce française « beau », ces deux mots

sont donc des voisins orthographiques) et une deuxième condition non-reliée (toujours pour la cible « BEAN », l’amorce « ciel »). Chaque amorce, qu’elle soit reliée ou non-reliée, a la même longueur que sa cible. Les deux types d’amorces sont appariées selon leur fréquence lexicale (pour les amorces précédant les cibles mots, $t(39) = 1.608, p = .116$, n.s. et pour les amorces précédant les pseudo-mots, $t(39) = .891, p = .378$, n.s.). Pour les cibles mot, la fréquence moyenne des amorces reliées est de 148.7 (203.2), et pour les amorces non-relières la fréquence moyenne est de 139.6 (180.6). Pour les cibles pseudo-mots, la fréquence moyenne des amorces reliées est de 105.4 (173.4), et pour les amorces non-relières, la fréquence moyenne est de 103.1 (167.3).

Tableau 1 : propriétés linguistiques (moyennes et écarts-types) des cibles mots

Cibles MOTS		
	Condition « specific »	Condition « shared »
Longueur	4.2 (.4)	4.3 (.47)
Fréquence	102.8 (99.5)	100.5 (157.7)
Taille du voisinage (N)	9 (4.5)	7.3 (5.4)
Fréquence du voisinage (N)	54.3 (75.1)	54.6 (65)
Fréquence bigrammes	2942 (997.3)	3399 (1508.3)
Taille du N français	1.7 (.9)	2 (1.1)
Fréquence cumulée N français	155.2 (195.3)	170.8 (201.3)
Fréquence amorce reliée	134.5 (182.6)	162.8 (225.8)
Fréquence Amorce non-reliée	133.8 (180.4)	145.4 (185.2)
Fréquence bigramme min. français	321 (435)	2670 (2287.6)

Ensuite 40 pseudo-mots cibles ont été construits en changeant une lettre d’un mot cible anglais précédemment sélectionné soit au début soit à la fin du mot (ex : BOAT – FOAT ou FLAT – FLAD). Ces 40 pseudo-mots suivent exactement les mêmes conditions que les cibles mots (les propriétés linguistiques des pseudo-mots sont présentées en Annexe 5 dans le Tableau 2).

Les cibles pouvant être précédées de deux amorces différentes, nous avons constitué deux listes (1 et 2) de paires amorces-cibles dont l’ordre a été contrebalancé à chaque passation (chaque participant voit les deux listes donc les 80 cibles), afin que chaque participant puisse, pour chaque cible, être confronté aux deux types d’amorces (reliée et non-

reliée). Les mêmes cibles sont utilisées dans les deux listes, seule la condition d'amorçage diffère : par exemple pour la cible « BEAN » l'amorce de la première liste sera « beau », alors que l'amorce de la seconde liste sera « ciel ». Chaque sujet verra alors autant de cibles mots que de cibles pseudo-mots.

III.3. Procédure

Chaque session commence avec l'apparition sur fond blanc d'un point de fixation (+) pendant 1000 ms. Puis un masque constitué de 5 dièses (#####) est présenté pendant 500 ms, après lequel apparaît l'amorce écrite en lettres minuscules pendant 50 ms. La cible est enfin présentée en lettres majuscules jusqu'à la réponse du sujet (voir figure 4).

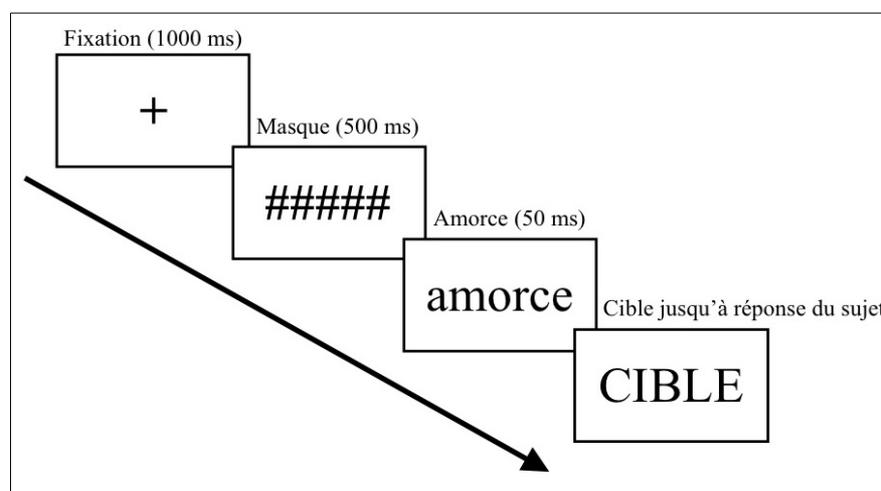


Figure 4. Représentation schématique de la procédure expérimentale utilisée

Il s'agit d'une tâche de décision lexicale, c'est-à-dire que les participants doivent décider si la cible présentée est un mot anglais ou pas. S'ils estiment que la cible présentée est un mot anglais existant, ils appuieront sur une touche positive (identifiée sur le clavier avec une gommette de couleur) avec la main dominante, et s'ils estiment que la cible présentée n'est pas un mot anglais, ils appuieront sur la touche négative avec la main non-dominante. La latéralité de chaque sujet est demandée afin d'adapter le système de réponse clavier du logiciel. C'est une tâche de décision lexicale avec amorçage masqué, les sujets ne sont donc pas informés de l'apparition des amorces. Nous leur avons demandé de répondre aussi vite que possible tout en commettant le moins d'erreurs possible et nous avons pris soin d'énoncer cette consigne de la même manière pour tous les sujets (afin de ne pas induire de biais concernant l'utilisation préférentielle d'une stratégie liée à la vitesse ou à la précision). Des pauses sont proposées tous les 20 items. Après la présentation de l'exercice chaque participant

a effectué un entraînement à la tâche de décision lexicale, suivi de la première session de décision lexicale, puis une tâche de lecture orale avec un questionnaire, ensuite la deuxième session de décision lexicale et pour finir un exercice de traduction. Au total l'expérience dure vingt minutes.

IV. Résultats

Dans la présente étude nous avons conduit une ANOVA dans laquelle étaient manipulés les facteurs de typicalité orthographique des cibles avec 2 conditions « spécifique » et « partagée », et la relation orthographique entre amorce et cible avec 2 conditions ; une première reliée orthographiquement et une deuxième non-reliée. Toutes ces variables sont à mesures répétées. Dans la présente analyse nous avons mesuré les temps de réaction des sujets en millisecondes à la tâche de décision lexicale proposée, ainsi que le taux d'erreur. Les moyennes et écarts-types seront présentés dans le Tableau 3. Nous avons également analysé les résultats par session (1 et 2) afin de mettre en évidence un éventuel effet de cette dernière.

Afin de procéder au nettoyage des données, nous avons calculé la moyenne des temps de réaction des sujets pour les cibles mots sans les erreurs (676.6 ms), et déterminé l'écart-type (183.8) afin de supprimer toutes les données au-delà de 2.5 écart-types au-dessus de la moyenne. Ainsi, toutes les données supérieures à la valeur seuil de 1136 ms ont été enlevées, ce qui représente 39 données sur 1287 soit 3 % du nombre total de données (< 5%). Nous avons également supprimé les données de 2 participants dans l'analyse des temps de réaction et des erreurs, dont les temps de réaction anormalement élevés témoignaient d'une mauvaise compréhension des consignes.

L'analyse des temps de réaction a montré un effet significatif de la condition session, $F(1,17) = 18.861, p < .001, \eta^2p = .53$. Cet effet reflète que les sujets ont en moyenne répondu plus rapidement en session 2 (633.33 ms) par rapport à la session 1 (683.58 ms), amorces et cibles confondues. L'effet de la condition amorce était également significatif, $F(1,17) = 26.669, p < .001, \eta^2p = .61$. Les participants étaient en moyenne plus rapides dans la condition non-reliée (645.46 ms) par rapport à la condition reliée (671.46 ms). Par contre les effets de la condition cible, de l'interaction entre la session et la cible, entre la cible et l'amorce ainsi qu'entre la session, la cible et l'amorce étaient tous non significatifs, $F < 1, n.s.$, tout comme l'effet d'interaction entre la session et l'amorce, $F(1,17) = 1.027, p = .3, n.s.$

L'analyse du taux d'erreur a mis en évidence un effet significatif de la condition cible, $F(1,17) = 36.599, p < .001, \eta^2p = .68$. Les sujets ont eu plus de bonnes réponses dans la condition des cibles « spécifiques » (.929 ou 7.1 % d'erreurs) que dans celle des cibles « partagées » (.857 ou 14.3 % d'erreurs). L'effet de la session était non significatif, $F(1,17) = 1.863, p = .19, n.s.$ L'effet d'interaction entre la condition cible et la condition amorce était

non significatif, $F(1,17) = .789, p = .2, n.s.$ L'effet de la condition d'amorçage, de l'interaction entre la session et la condition cible et entre la session et la condition amorcée étaient également non significatifs, $F < 1, n.s.$ L'effet d'interaction d'ordre 3 entre la session, la condition cible et la condition amorcée était aussi non significatif, $F < 1, n.s.$

Tableau 3 : Moyennes et écart-types des temps de réaction et du taux d'erreurs pour chaque condition et pour les 2 sessions

Session 1	Cible Partagée		Cible Spécifique	
	Amorce reliée	Amorce non reliée	Amorce reliée	Amorce non reliée
Temps de réaction	675.2 (90.7)	703 (88.5)	659.2 (78.7)	696.9 (71.5)
% d'Erreurs	16 (0.2)	12 (0.1)	6.5 (0.1)	7.5 (0.1)
Session 2	Cible Partagée		Cible Spécifique	
	Amorce reliée	Amorce non reliée	Amorce reliée	Amorce non reliée
Temps de réaction	619.8 (53.3)	644 (67)	627.7 (61)	641.8 (60)
% d'Erreurs	15 (0.1)	16.5 (0.1)	7 (0.1)	8 (0.1)

V. Discussion

L'objectif de cette recherche était d'étudier les effets du voisinage orthographique inter-langue couplés à ceux de la typicalité orthographique, de la L1 à la L2, chez des bilingues Français-Anglais moyennement compétents en anglais (L2). Dans cette expérience nous avons testé l'influence de deux facteurs sur la rapidité et l'exactitude de l'identification de mots et pseudo-mots dans une tâche de décision lexicale en anglais. La première variable concernait la relation orthographique inter-langue entre la cible et l'amorce avec deux modalités ; une condition reliée (beau-BEAN) et une condition non-reliée (ciel-BEAN). La deuxième variable concernait la typicalité orthographique des cibles anglaises ; la première condition « spécifique » était constituée de cibles anglaises qui contenaient des bigrammes peu typiques du français (par exemple le mot WRAP), et la deuxième condition « partagée » comprenait des cibles anglaises qui contenaient des séquences de lettres typiques du français (comme TRAP ou BIND). Dans cette dernière condition les cibles françaises et amorces anglaises partagent une orthographe qui pourrait être retrouvée autant en anglais qu'en français (règles ortho-tactiques similaires), rendant les mots de cette condition à priori plus difficile à discerner (d'après les travaux de Vaid & Frenck-Mestre, 2002 ; Casaponsa, Carreiras & Duñabeitia, 2014 ; Casaponsa & Duñabeitia, 2015). Nous nous attendions donc à ce que les mots anglais de la condition « spécifique » (avec des bigrammes fréquents en anglais mais pas en français) soient reconnu plus facilement par rapport à ceux de la condition « partagée ». Nous nous attendions également à ce que les amorces françaises reliées orthographiquement aux cibles anglaises inhibent leur reconnaissance (comme l'avaient entre autres démontré Bijeljac-Babic et al., 1997), à cause du phénomène de compétition lexicale.

Les résultats ne sont pas tout à fait allés dans le sens de ces hypothèses. En ce qui concerne l'analyse des temps de réaction, seuls deux résultats étaient significatifs ; il y a eu un effet de la session ainsi qu'un effet de l'amorce (cibles confondues). C'est-à-dire que les sujets ont été plus rapides pour répondre lors de la session 2, et plus rapides pour répondre dans la condition où l'amorce était reliée à la cible, démontrant ici un effet de facilitation de l'amorçage. Cet effet de facilitation n'était pas attendu, et va à l'encontre de nos hypothèses de départ. De nombreux auteurs dont Bijeljac-Babic et al. (1997) avaient montré un effet inhibiteur des amorces françaises reliées orthographiquement aux cibles anglaises mais avec des amorces de haute fréquence et des cibles de basse fréquence chez des bilingues performants. Dans notre étude, la fréquence des amorces et des cibles était plus modérée (la

fréquence moyenne des mots cibles était de 101.7 o.p.m., et la fréquence moyenne des amorces était de 144.1 o.p.m.). Par contre la fréquence des amorces françaises était relativement hétérogène (de 13.79 à 790 o.p.m.). Il est important de noter que les fréquences des cibles anglaises sont les fréquences d'apparition des mots dans la langue anglaise ; un Français qui apprend l'anglais n'aura pas le même contact avec ces fréquences. Il serait intéressant pour notre étude, d'utiliser une base de données de la fréquence des mots anglais auxquels sont exposés les Français qui apprennent la langue, afin de pouvoir manipuler plus justement la fréquence des mots anglais. Malheureusement cette base de données n'existe pas et constituerait un travail considérable. Ce pourrait cependant être l'objet d'un prochain travail.

Comme nous l'avons vu précédemment, l'effet de l'inhibition de l'amorce est un indice du mécanisme de compétition lexicale. Notre résultat illustrant un effet de facilitation ne met pas en évidence le mécanisme de compétition lexicale proposé dans le modèle BIA (et BIA+, Dijkstra & Van Heuven, 2002). Commissaire (2012) avait déjà trouvé un effet de facilitation des amorces française (L1) reliées de basse fréquence sur les cibles anglaises (L2) de haute fréquence, chez des participants peu compétents en anglais. Souvent ce mécanisme d'inhibition a été trouvé dans le sens de la L2 vers la L1 mais pas dans le sens de la langue 1 vers la langue 2 que nous avons testé ici. Il serait donc intéressant de répliquer cette étude dans le sens opposé, de la langue 2 vers la langue 1, car les liens ne sont pas de la même intensité dans les deux sens. Il semblerait qu'un mot de la langue 2 inhibe plus la reconnaissance d'un mot de la langue 1. Dans notre étude les cibles et les amorces étaient d'une fréquence comparable et plutôt modérée (entre 100 et 200 o.p.m.). Un effet de facilitation a été trouvé (dans l'étude de Commissaire, 2012, et celle-ci) pour les cibles anglaises de haute et de relativement haute fréquence, précédées par des amorces françaises reliées de basse et relativement haute fréquence. Afin d'avoir un effet d'inhibition (ou au moins une absence de facilitation) dans notre étude, peut-être aurions-nous pu utiliser des amorces françaises de haute fréquence, avec les mêmes cibles. Nous savons que plus la fréquence d'un mot est basse, plus son seuil d'identification sera haut et plus la fréquence d'une amorce est haute, plus elle peut envoyer de l'inhibition à ses voisins orthographiques. Utiliser des amorces de haute fréquence mettrait peut-être en lumière d'autres effets (un effet d'inhibition notamment).

Il me semble que de prendre en compte les aspects phonologiques de la langue

anglaise pourrait être pertinent afin d'expliquer ce résultat. Ziegler et Perry (1998) avaient démontré l'importance de la rime orthographique en anglais (SHAVE et HAVE sont bien des voisins orthographiques, mais se prononcent très différemment), par rapport à d'autres langues dont le français. Ils ont montré qu'un mot anglais qui avait beaucoup de voisins de sonorité (appelés « body neighbours » comme le mot cheat qui a comme voisins de sonorité wheat, meat et heat) étaient reconnus plus facilement. Ils distinguent alors les voisins orthographiques des voisins de sonorité (body neighbours). Si nous répliquions cette étude, il serait intéressant de contrôler ce voisinage, afin de réduire éventuellement l'effet de facilitation. Mais encore, peut-être que beau a réussi à faciliter la reconnaissance de BEAN car ces deux mots sont prononcés très différemment bien qu'ils soient voisins ([bin] vs. [bo]). Cette différence phonologique n'amènerait donc pas à de la compétition lexicale entre ces deux mots. Seule la phonétique de la première lettre « B » est commune et suffirait peut-être à faciliter la reconnaissance de la cible. Peut-être que les sujets peu performants en L2 accèdent en premier à des niveaux de traitement phonologique avant d'atteindre le niveau lexical.

Mais si « beau » a pu faciliter la reconnaissance de « BEAN » chez ces bilingues peu performants, peut-être que les sujets sont restés à des niveaux de traitement plus superficiels. Peut-être qu'avec des mots de la L2 moyennement fréquents (et, je le rappelle, fréquents pour des Anglais mais pas pour des Français), les bilingues peu compétents n'accèdent pas au lexique de leur seconde langue et restent à des niveaux de traitement plus superficiels. Il serait donc intéressant de refaire la même expérience avec des bilingues fortement compétents en anglais. A ce sujet, Grainger, Midgley et Holcomb (2010) ont proposé une extension du modèle BIA, le BIA-d présenté en figure 5.

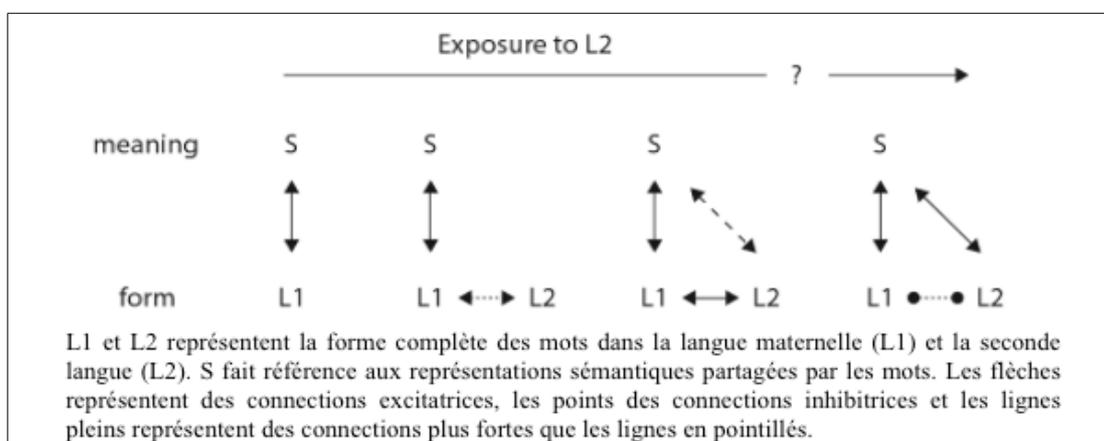


Figure 5. Le BIA-d, de Grainger, Midgley et Holcomb (2010)

Ces auteurs ont étudié l'acquisition du vocabulaire chez les bilingues ayant appris une L2 tardivement dans un contexte scolaire, et proposent que la compétition lexicale entre les mots de différentes langues n'émerge qu'après un certain temps d'exposition à la seconde langue (L2). Le lexique de la L2 aurait d'abord besoin de s'étendre et de devenir plus autonome afin de faire des relations directes avec le niveau sémantique (S). Il serait donc possible que nos participants n'aient pas été exposés suffisamment à l'anglais pour que l'on puisse voir émerger un phénomène de compétition lexicale (via l'inhibition inter-langue). Si l'objectif est toujours d'étudier des bilingues Français ayant un faible niveau de performance en anglais (ayant appris dans un contexte scolaire) mais ayant été exposés de manière plus intensive à l'anglais, il serait intéressant de sélectionner des participants Québécois. En effet, ils apprennent l'anglais dans un contexte scolaire mais sont beaucoup plus exposés à la langue anglaise que les Français de France (via des films, la culture Américaine très proche géographiquement, etc.). Il pourrait également être intéressant d'énoncer les consignes de l'exercice en anglais, afin de mettre le participant dans un contexte anglophone avant qu'il commence la tâche (de manière à pré-activer son lexique anglais).

La taille ainsi que la fréquence du voisinage anglais et français des cibles ont été contrôlées et sont statistiquement comparables pour toutes les conditions, donc ces deux paramètres n'ont pas pu influencer les résultats. Cependant, nous savons que la présence d'un voisin partagé entre l'amorce et la cible peut influencer la reconnaissance de cette dernière (Van Heuven, Dijkstra, Grainger & Schriefers, 2001, chez les monolingues). En effet ces chercheurs avaient montré, en Néerlandais, que les amorces reliées qui avaient au moins un voisin partagé avec les cibles inhibaient la reconnaissance des cibles. Il serait intéressant d'investiguer ce paramètre entre deux langues (le voisinage partagé inter-langue) dans notre étude afin de contrôler ce paramètre pour avoir le même nombre de voisins partagés entre l'amorce et la cible, dans toutes les conditions.

L'effet d'amorçage a également interagi avec le nombre de présentations de la cible. Les participants ont vu deux fois la même cible, une fois précédée par son amorce reliée et une fois par son amorce non-reliée. L'analyse des temps de réaction a montré un effet de facilitation pour la session 2 : les sujets ont répondu plus rapidement en session 2. Ce résultat peut tout simplement être expliqué par le fait que les sujets ont vu deux fois la même cible (une fois dans la session 1 et une deuxième dans la session 2) donc ils sont plus rapides pour reconnaître cette cible, car ils l'ont déjà identifiée une fois. Dijkstra et al. (2010) avaient

trouvé un effet d'inhibition uniquement en session 1 pour cette même raison, cependant dans notre étude, même en session 1 nous n'observons aucun effet d'inhibition.

Nous pouvons également investiguer les paramètres propres aux sujets afin d'expliquer ces résultats. En effet, les adultes que nous avons sélectionnés étaient assez hétérogènes en termes d'âge (de 23 à 53 ans, avec un âge moyen de 28.7 ans (8.3)) ; certains ont appris l'anglais avant les années 2000 avec des méthodes différentes de celles d'aujourd'hui. Il est important de préciser que la catégorie socio-professionnelle n'a pas été contrôlée, la majorité des participants étaient étudiants, mais d'autres étaient des employés. Ces participants ont eu un contact avec la langue anglaise assez différent les uns des autres, ce qui pourrait influencer d'une certaine manière les résultats. Etant donné que ces participants n'ont pas la même expérience avec l'anglais (certains ont une utilisation plutôt universitaire de l'anglais alors que d'autres ont une utilisation à des fins relationnelles) il se pourrait que des participants aient été peu exposés à l'anglais écrit par rapport à d'autres.

Pour les effets de la typicalité, nous nous attendions à ce que les mots de la catégorie « spécifique » soient reconnus plus facilement que ceux de la catégorie partagée. L'analyse des erreurs va dans le sens de notre hypothèse initiale en tant que les sujets ont fait moins d'erreurs pour reconnaître les mots anglais « spécifiques ». Ce résultat est en accord avec les résultats précédents de Casaponsa et al. (2014) et démontre que les sujets ont eu recours à une analyse des indices sub-lexicaux. Les mots ayant une orthographe non partagée avec le français en terme de fréquence de bigrammes minimale ont été identifiés avec moins d'erreurs, montrant ici que les bilingues peu performants semblent identifier ces mots avec plus d'exactitude. Peut-être que lorsque nous débutons dans l'apprentissage d'une langue, il nous est plus facile d'apprendre des mots spécifiques, qui partagent peu de caractéristiques orthographiques avec notre première langue (il y a alors moins de compétition lexicale). Par contre les sujets n'ont pas été plus rapides pour identifier les cibles « spécifiques » par rapport aux cibles « partagées ». Vu que les sujets ont en moyenne fait moins d'erreurs pour identifier les mots « spécifiques », peut-être ont-ils privilégié une stratégie mettant l'accent sur la justesse plutôt que sur la rapidité (comme l'avaient déjà étudié Posner, 1986). Etant donné qu'ils sont plutôt débutants en anglais, peut-être que les participants ont eu besoin de plus de temps d'analyse afin de répondre le plus correctement possible. Peut-être que la présentation de mots « spécifiques » a activé rapidement des indices sublexicaux poussant les sujets à répondre, mais que par doute à cause de leur faible niveau, ils ont pris plus de temps pour

vérifier leur réponse. Casaponsa et Duñabeitia (2015) reprennent l'idée que le lexique chez les bilingues serait organisé différemment, en fonction du degré de spécificité des mots par rapport au langage. Lorsque les mots ne seraient pas spécifiques d'une langue (comme dans notre catégorie de cibles « partagée »), les bilingues auraient un accès non sélectif au lexique alors que lorsque les mots seraient spécifiques d'une langue, ils pourraient avoir un accès plus sélectif à la langue, réduisant ainsi le degré d'interférence inter-langagière. Si on se réfère toujours aux principes du modèle BIA-d, étant donné que les liens entre les deux langues ne seraient pas tout à fait construits chez ces participants, et que leur lexique anglais est en construction lui aussi, l'effet de facilitation normalement attendu pour reconnaître les mots spécifiques n'advierait pas car le sujet ne pourrait pas accéder à son lexique anglais de manière indépendante du lexique français (même si les liens entre ces deux lexiques sont faibles). De plus, à ce stade d'apprentissage de la langue, toujours selon le BIA-d, il y aurait des connexions excitatrices entre les mots des deux langues (L1 et L2 sur la figure 5), ce qui ne favorise pas un accès plus indépendant au lexique, et pas non plus une certaine rapidité d'accès à celui-ci.

Ces résultats ne suffisent pas, à mes yeux, à remettre en cause l'accès non-sélectif au langage proposé par Grainger et Dijkstra (1992) ; Dijkstra et Van Heuven (1998). Comme nous l'avons vu plus haut, cet accès non-sélectif serait complètement mis en place seulement après un certain temps d'exposition à la langue, que les sujets de cette étude n'ont pas eu. Il serait donc intéressant de répliquer cette étude avec des participants fortement compétents en anglais, afin d'avoir des résultats plus clairs. Il serait également intéressant de contrôler d'autres facteurs, tel que la fréquence des cibles avec une nouvelle base données des mots anglais auxquels sont exposés les Français, la présence de voisins partagés inter-langue, ou encore les voisins de sonorité (body neighbours) de la cible. Refaire cette même expérience de la L2 vers la L1 pourrait également mettre en lumière d'autres résultats. Nous avons donc pu remarquer qu'étudier l'accès au lexique chez des bilingues peu compétents et tardifs mettait en évidence d'autres effets que ceux trouvés pour les bilingues natifs ou fortement compétents. Les mécanismes impliqués dans l'accès au lexique seraient donc différents en fonction du degré de compétence, et peut-être à envisager comme un continuum.

VI. Bibliographie

- Altenberg, E. P., & Cairns, H. S. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 22(2), 174–188.
- Andrews, S. (1989). Frequency and neighborhood effects on lexical access : Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 802–814.
- Andrews, S. (1992). Frequency and neighborhood effects on lexical access : Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2), 234–254.
- Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval : Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, 4(4), 439–461.
- Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology*, 4(2), 207–228.
- Beauvillain, C. (1992). Orthographic and Lexical Constraints in Bilingual Word Recognition. In R. J. Harris (Ed.), *Advances in Psychology* (Vol. 83, pp. 221–235). North-Holland.
- Bijeljac-babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition*, 25(4), 447–457.
- Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, 8(3), 531–544.
- Carr, T. H., Davidson, B. J., & Hawkins, H. L. (1978). Perceptual flexibility in word recognition : Strategies affect orthographic computation but not lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 4(4), 674–690.
- Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition : cross-task comparisons. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 23(4), 857–871.
- Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Frontiers in Psychology*, 5.

- Casaponsa, A., & Duñabeitia, J. A. (2016). Lexical organization of language-ambiguous and language-specific words in bilinguals. *The Quarterly Journal of Experimental Psychology*, *69*(3), 589–604.
- Cattell, J. M. (1886). The Time Taken Up by Cerebral Operations. *Mind*, *42*, 220–242.
- Coderre, E. L., Van Heuven, W. J. B., & Conklin, K. (2013). The timing and magnitude of Stroop interference and facilitation in monolinguals and bilinguals. *Bilingualism : Language and Cognition*, *16*(2), 420–441.
- Coltheart, M., & Coltheart, V. (1997). Reading Comprehension Is Not Exclusively Reliant upon Phonological Representation. *Cognitive Neuropsychology*, *14*(1), 167–175.
- Coltheart, M., Davelaar, E., Jonasson, T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI*. Lawrence Erlbaum Associates.
- Commissaire, E., Orthographic and phonological coding during L2 visual word recognition in L2 learners : lexical and sublexical mechanisms. Psychology. Université Charles de Gaulle - Lille III; University of Dundee - Ecosse -UK, 2012. English.
- Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301.
- Dijkstra, T., & Van Heuven, W. J. B., (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connexionist Approaches to Human Cognition* (pp. 189–225). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dijkstra, T., Van Heuven, W. J. B., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, *38*(3–4), 177–196.
- Dijkstra, T., & Heuven, W. J. B. van. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197.
- Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *10*(3), 297–302.
- Ehri, L. C., & Ryan, E. B. (1980). Performance of bilinguals in a picture-word interference task. *Journal of Psycholinguistic Research*, *9*(3), 285–302.

- Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *The Quarterly Journal of Experimental Psychology Section A*, 33(4), 325–350.
- Ferrand, L. (2007). *Psychologie Cognitive de la Lecture : Reconnaissance des mots écrits chez l'adulte*. Bruxelles : De Boeck.
- Forster, K. I. (1976). Accessing the mental lexicon. *New approaches to language mechanisms*, 30, 231-256.
- Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 680–698.
- Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood : Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology-Learning Memory and Cognition*, 22(3), 696-713.
- Grainger, J., & B, J. (2004). Modeling Letter Position Coding in Printed Word Perception. In *Mental lexicon: "Some words to talk about words"* (pp. 1–23). Hauppauge, NY, US: Nova Science Publishers.
- Grainger, J., & Dijkstra, T. (1992). On the Representation and Use of Language Information in Bilinguals. In R. J. Harris (Ed.), *Advances in Psychology* (Vol. 83, pp. 207–220). North-Holland.
- Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, 20(6), 1158–1176.
- Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103(3), 518–565.
- Grainger, J., & Jacobs, A. M. (1999). Temporal Integration of Information in Orthographic Priming. *Visual Cognition*, 6(3–4), 461–492.
- Grainger, J., & Jacobs, A. M. (2013). *Localist connexionist Approaches To Human Cognition*. Psychology Press.

- Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail & M. Hickmann (Eds.), *Language Acquisition across Linguistic and Cognitive Systems*. New York: John Benjamins (pp. 267-284).
- 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 masked identification latencies. *Perception & Psychophysics*, *47*(2), 191–198.
- Haber, L. R., Haber, R. N., & Furlin, K. R. (1983). Word Length and Word Shape as Sources of Information in Reading. *Reading Research Quarterly*, *18*(2), 165–189.
- Humphreys, G. W., Evett, L. J., & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology*, *22*(4), 517–560.
- Huntsman, L. A., & Lima, S. D. (2002). Orthographic Neighbors and Visual Word Recognition. *Journal of Psycholinguistic Research*, *31*(3), 289–306.
- Huntsman, L. A., & Lima, S. D. (1996). Orthographic neighborhood structure and lexical access. *Journal of Psycholinguistic Research*, *25*(3), 417–429.
- Kesteren, R. van, Dijkstra, T., & Smedt, K. de. (2012). Markedness effects in Norwegian–English bilinguals: Task-dependent use of language-specific letters and bigrams. *The Quarterly Journal of Experimental Psychology*, *65*(11), 2129–2154.
- Macnamara, J. (1967). The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *6*(5), 729–736.
- Macnamara, J., & Kushnir, S. L. (1971). Linguistic independence of bilinguals : The input switch. *Journal of Verbal Learning and Verbal Behavior*, *10*(5), 480–487.
- McClelland, J. L., & Johnston, J. C. (1977). The role of familiar units in perception of words and nonwords. *Perception & Psychophysics*, *22*(3), 249–261.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception : Part I. An account of basic findings. *Psychological Review*, *88*(5), 375–407.

- Medler, D.A., & Binder, J.R. (2005). MCWord : An On-Line Orthographic Database of the English Language. <http://www.neuro.mcw.edu/mcword/>
- Monk, A. F., & Hulme, C. (1983). Errors in proofreading : Evidence for the use of word shape in word recognition. *Memory & Cognition*, *11*(1), 16–23.
- Nakayama, M., Sears, C. R., & Lupker, S. J. (2008). Masked priming with orthographic neighbors: A test of the lexical competition assumption. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1236–1260.
- Nas, G. (1983). Visual word recognition in bilinguals : evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior*, *22*(5), 526–534.
- Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767–779.
- Posner, M.I. (1986). *Chronometric explorations of the mind*. New York : Oxford University Press.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275–280.
- Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 487–494.
- Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 84–99.
- Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception : Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333–367.
- Sears, C. R., Campbell, C. R., & Lupker, S. J. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1040–1062.
- Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology : Human Perception and Performance*, *21*(4), 876–900.

- Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors : Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 65–76.
- Snodgrass, J. G., & Mintzer, M. (1993). Neighborhood effects in visual word recognition : Facilitatory or inhibitory? *Memory & Cognition*, *21*(2), 247–266.
- Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, *12*(4), 380–386.
- Spoehr, K. T., & Smith, E. E. (1975). The role of orthographic and phonotactic rules in perceiving letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(1), 21–34.
- Traxler, M., & Gernsbacher, M. A. (2011). *Handbook of Psycholinguistics*. Academic Press.
- Vaid, J., & Frenck-Mestre, C. (2002). Do orthographic cues aid language recognition? A laterality study with French–English bilinguals. *Brain and Language*, *82*(1), 47–53.
- Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483.
- Van Heuven, W. J., Dijkstra, T., Grainger, J., & Schriefers, H. (2001). Shared neighborhood effects in masked orthographic priming. *Psychonomic Bulletin & Review*, *8*(1), 96–101.
- Westbury, C., & Buchanan, L. (2002). The Probability of the Least Likely Non-Length-Controlled Bigram Affects Lexical Decision Reaction Times. *Brain and Language*, *81*(1–3), 66–78.
- Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, *1*(1), 59–85.
- Woodworth, R. S. (1938). *Experimental Psychology*. H. Holt.
- Ziegler, J.C., & Perry, C. (1998). No more problems in Coltheart’s neighborhood : resolving neighborhood conflicts in the lexical decision task. *Cognition*, *68*(2), B53–B62.

VII. Annexes

Annexe 1 : Liste des items à traduire en anglais présentée aux participants

NOM
PRENOM
DATE

Écrire la traduction en anglais

(Commencer par la colonne 1 puis la colonne 2 et pour finir la colonne 3)

Colonne 1		Colonne 2		Colonne 3	
Dormir		La colline		Brancher	
Vraiment		Le papillon		Le pont-levis	
La maison		Partager		La prune	
La tante		La larme		Creuser	
La fille		Le naufrage		Tondre	
Aimer		Suivre		Cambrioler	
Tard		Le barrage		Ronfler	
Vendre		Gagner		Siffler	
Le doigt		Renverser		La méduse	
Parfois		L'ongle		L'échafaudage	
L'église		Le dessin		Le cerf-volant	
Le sable		Crier		Poignarder	
Fort		Le rideau		Semer	
L'année		Rêver		Couler	
Lire		Le poing		Espiègle	
Dire		Traverser		La jetée	
La voiture		La maladie		La brouette	
Envoyer		Le toit		Entêté	
La sœur		Se raser		Les fléchettes	
Croire		La craie		Ombragé	
Attendre		L'aiguille		Hisser	
Jouer		La chèvre		Insonorisé	
Noir		Le menton		La boussole	
Petit		Le ruisseau		Élire	
Le pays		L'éventail		Le poignard	

Annexe 2 : Questionnaire proposé aux participants

Prénom: Numéro: Date :
Date de naissance : Latéralité : Niveau :

Langue maternelle ?

Quand as-tu commencé à apprendre l'anglais ? Dans quel contexte ?

Apprends- tu l'anglais encore aujourd'hui ?

As- tu appris une autre langue que celle-ci ?

Pratique orale de l'anglais actuellement :

En %, quelle proportion de temps parles- tu/entends-tu l'anglais au quotidien ?

En %, quelle proportion de temps parles- tu/entends-tu le français au quotidien ?

En %, quelle proportion de temps parles- tu/entends-tu une autre langue au quotidien ?

TOTAL: 100%

Pratique écrite de l'anglais actuellement :

En %, quelle proportion de temps lis-tu/écris-tu l'anglais au quotidien ?

En %, quelle proportion de temps lis-tu/écris-tu le français au quotidien ?

En %, quelle proportion de temps lis-tu/écris-tu une autre langue au quotidien ?

TOTAL: 100%

Comment situes- tu ton niveau d'anglais ? 1 – 2 – 3 – 4

1 : Je peux dire et comprendre quelques mots et expressions.

2 : Je peux comprendre et parler dans une conversation simple concernant le quotidien.

3 : Je suis à l'aise dans la langue pour les conversations du quotidien. Dans les conversations plus complexes, je saisis les idées principales et peux répondre de façon approximative.

4 : Je suis très à l'aise pour communiquer dans la langue y compris pour les notions complexes ou abstraites. Je peux exprimer tout ce que je souhaite et de la façon que je le souhaite.

Annexe 3 : Tâche de lecture chronométrée en français présentée aux sujets : lecture de texte et fiche d'évaluation

Lecture de texte⁵²

Le pollueur peut payer ses dégâts sur le net.

En France, le phénomène a seulement un an et demi. Trois sites Internet proposent aux particuliers de calculer leurs émissions de dioxyde de carbone (un gaz à effet de serre) lorsqu'ils utilisent leur voiture, chauffent leur maison ou prennent l'avion. Les internautes peuvent ensuite payer une somme correspondant à leur niveau de pollution pour financer des projets contre le réchauffement climatique. Exemple : un vol aller-retour Paris-Rome rejette mille trois cents kilos de dioxyde de carbone dans l'atmosphère par personne. Il coûte vingt-six euros sur le site Climat Mundi. Ce système est appelé « mécanisme de compensation ». Un peu comme le principe du pollueur/payeur mis en œuvre dans le cadre du protocole de Kyoto. « Les adhérents sont déjà des écolos convaincus. Ils veulent compenser les rejets de dioxyde de carbone qu'ils ne peuvent pas éviter. On repousse les limites de l'écologie », explique Eric Parent, cofondateur de Climat Mundi.

Une initiative récente qui prend de l'ampleur.

Sur ce site Internet accessible depuis juin, l'argent sert à financer trois projets : des fours moins gourmands en énergie en Erythrée, des éoliennes en Turquie et une usine de production propre en Australie. Un autre site, Action Carbone, finance la reforestation en Colombie. Selon Climat Mundi, les sommes récoltées sont encore faibles mais le phénomène prend de l'ampleur. « Environ cinq personnes compensent chaque jour sur notre site leurs rejets de dioxyde de carbone », affirme Eric Parent. Ce nouveau moyen de protéger la planète n'a pas que des adeptes. Jean-Marc Janovici, ingénieur en énergie, est dubitatif : « le meilleur moyen de réduire nos émissions de dioxyde de carbone est encore de...les réduire. Même en payant, les gaz à effet de serre sont toujours dans l'atmosphère. Compenser c'est aussi penser qu'on peut ne rien changer. »

⁵² L'actu, le quotidien d'actualité dès 14 ans, www.playbac.fr

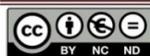


9. Lecture de texte : Le pollueur

Elise Clérin dans L'actu –mercredi 8 novembre 2006⁵¹

Le pollueur peut payer ses dégâts sur le net. En France, le phénomène a seulement un	16
an et demi. Trois sites Internet proposent aux particuliers de calculer leurs émissions	29
de dioxyde de carbone (un gaz à effet de serre) lorsqu'ils utilisent leur voiture,	43
chauffent leur maison ou prennent l'avion. Les internautes peuvent ensuite payer une	55
somme correspondant à leur niveau de pollution pour financer des projets contre le	68
réchauffement climatique. Exemple : un vol aller-retour Paris-Rome rejette mille	77
trois cents kilos de dioxyde de carbone dans l'atmosphère par personne. Il coûte	90
vingt-six euros sur le site Climat Mundi. Ce système est appelé « mécanisme de	103
compensation ». Un peu comme le principe du pollueur/payeur mis en œuvre dans le	117
cadre du protocole de Kyoto. « Les adhérents sont déjà des écolos convaincus. Ils	130
veulent compenser les rejets de dioxyde de carbone qu'ils ne peuvent pas éviter. On	144
repousse les limites de l'écologie », explique Eric Parent, cofondateur de Climat	155
Mundi. Une initiative récente qui prend de l'ampleur. Sur ce site Internet accessible	168
depuis juin, l'argent sert à financer trois projets : des fours moins gourmands en	181
énergie en Erythrée, des éoliennes en Turquie et une usine de production propre en	195
Australie. Un autre site, Action Carbone, finance la reforestation en Colombie. Selon	207
Climat Mundi, les sommes récoltées sont encore faibles mais le phénomène prend de	220
l'ampleur. « Environ cinq personnes compensent chaque jour sur notre site leurs	231
rejets de dioxyde de carbone », affirme Eric Parent. Ce nouveau moyen de protéger la	245
planète n'a pas que des adeptes. Jean-Marc Janovici, ingénieur en énergie, est	257
dubitatif : « le meilleur moyen de réduire nos émissions de dioxyde de carbone est	270
encore de...les réduire. Même en payant, les gaz à effet de serre sont toujours dans	286
l'atmosphère. Compenser c'est aussi penser qu'on peut ne rien changer. »	296
Nombre de mots lus : __ __ __ Nombre d'erreurs : __ __	
MCLM : __ __ __	

⁵¹ L'actu, le quotidien d'actualité dès 14 ans, www.playbac.fr



Annexe 4 : Liste des stimuli utilisés

Amorces Françaises		Cibles MOTS Anglais	
Reliées	Non-reliées	« specific »	« shared »
beau	ciel	BEAN	
blond	reine	BLOOD	
bloc	vase	BLOW	
bout	gens	BOAT	
brave	guise	BRAKE	
chat	neuf	COAT	
fond	vite	FOOD	
goût	lire	GOAT	
grec	pipe	GREY	
mont	chic	MONK	
peur	bois	PEAR	
rend	vain	READ	
rond	juin	ROAD	
sent	cuir	SEAT	
phare	brise	SHARE	
sein	clef	SKIN	
soin	rage	SOON	
chose	femme	WHOSE	
torse	paume	WORSE	
drap	flic	WRAP	
bond	cerf		BIND
perd	lame		HERD
bain	poil		BARN
grand	heure		BRAND
soir	bras		STIR
juive	sabre		JUICE
trop	chez		TRAP
huit	paix		HUNT
fleur	grise		FLOUR
blond	proie		BLIND
bord	chef		BIRD
raide	texte		RAISE
fasse	mince		FALSE
huit	dame		HURT
faim	sept		FARM
plat	nord		FLAT
mode	banc		MOVE
part	mois		PAST
mine	quai		MIND
lait	port		LAST

Amorces Françaises		Cibles PSEUDO-MOTS Anglais	
Reliées	Non-reliées	« specific »	« shared »
peau	cour	PEAN	
froid	pense	FROOD	
clou	fade	CLOW	
fort	lieu	FOAT	
grave	butte	GRAKE	
cour	vide	COAR	
pond	muse	POOD	
doit	cinq	DOAT	
bref	faim	BREY	
rond	quai	RONK	
veau	code	VEAR	
pend	froc	PEAD	
roux	vase	ROAX	
seau	brin	SEAF	
phare	fente	THARE	
sien	choc	SKEN	
foin	cire	FOON	
phase	beige	WHASE	
forme	genre	WORME	
trop	main	WROP	
gond	surf		GIND
vert	pain		HERT
brun	flot		BAUN
blanc	ombre		BRANC
soie	banc		STIE
tuile	panse		TUICE
trac	bise		TRAN
nuit	voix		NUNT
pleur	jauge		PLOUR
blond	proie		BLUND
fard	tact		FIRD
taise	tinte		NAISE
valse	verbe		NALSE
fort	pied		FURT
daim	pope		DARM
flan	snob		FLAD
mode	bain		MOGE
sait	pied		SAST
vend	maux		VIND
part	vite		LART

Annexe 5 : Tableau des propriétés linguistiques des cibles pseudo-mots

Tableau 2 : propriétés linguistiques (moyennes et écarts-types) des cibles pseudo-mots

Cibles PSEUDO - MOTS		
	Condition « specific »	Condition « shared »
Longueur	4.2 (.4)	4.3 (.5)
Taille du voisinage (N)	8.2 (3.9)	6.7 (4.1)
Fréquence du voisinage (N)	97.6 (240.6)	73.3 (82.1)
Fréquence bigrammes	2570.6 (815.1)	3123.7 (1523.1)
Taille du N français >1	1.8 (1.1)	2 (1.1)
Fréquence cumulée N français	189.4 (305.9)	120.4 (173.9)
Fréquence amorce	109.9 (180.4)	101 (170.7)
Fréquence bigramme min. français	374.7 (492.8)	2781.9 (2032.1)
Fréquence Amorce non-reliée	108.8 (178.4)	97.5 (159.7)