

La reconnaissance des mots cognats anglais-français chez les apprenants finnophones

Le rôle de la similarité phonologique

Anni Hintikka

Mémoire de master

Programme de master : Enseignement et apprentissage des langues,
Département de français

Institut de langues et de traduction

Faculté des Lettres

Université de Turku

Mai 2021

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

UNIVERSITÉ DE TURKU

Institut de langues et de traduction / Faculté des Lettres

HINTIKKA ANNI : La reconnaissance des mots cognats anglais-français chez les apprenants finnophones. Le rôle de la similarité phonologique

Mémoire de master, 48 p. (15 p. d'annexes)

Département de français

Programme de master : Enseignement et apprentissage des langues, Département de français

Mai 2021

Dans ce mémoire de master, nous étudions la reconnaissance des mots chez les apprenants finnophones de français L3, qui maîtrisent déjà une autre langue étrangère, l'anglais (L2). Nous nous intéressons aux mots cognats. Les cognats sont des équivalents de traduction qui ont une forme phonologique et orthographique commune dans deux ou plusieurs langues (comme le mot *rose* en anglais et en français). Nous visons à savoir, dans un premier temps, si ces mots sont reconnus différemment des mots non-cognats. Dans un deuxième temps, puisqu'il existe une lacune de recherche concernant le rôle de la phonologie dans la reconnaissance des mots cognats, nous tentons à clarifier si le degré de similarité phonologique des mots cognats joue un rôle dans la reconnaissance des mots.

Afin de répondre à ces questions, une expérience psycholinguistique était menée avec l'outil en ligne PsyToolkit. Nous avons utilisé la tâche de décision lexicale visuelle et collecté les temps de réaction et les taux de réussite. L'expérience consistait de 168 mots : 28 mots cognats, 56 mots non-cognats et 84 non-mots. Les mots cognats avaient une forme orthographique identique en français et en anglais mais le degré de leur similarité phonologique variait. Les données collectées de 20 participants ont été analysées avec les modèles linéaires (généralisés) à effets mixtes avec le logiciel R.

Les résultats suggèrent que la L2 non-cible, l'anglais, était activée chez les participants lors du traitement lexical en français L3. Cela indique qu'il existerait une différence représentationnelle dans le lexique mental entre les mots cognats et les non-cognats. De plus, le degré de similarité entre les formes phonologiques de l'anglais et du français a influencé la reconnaissance des mots. Conséquemment, les résultats de cette étude mettent en évidence qu'il est important de prendre en compte le rôle de la similarité phonologique des mots cognats dans les études futures.

Mots-clés : psycholinguistique, phonologie, cognat, reconnaissance des mots

Table des matières

1. Introduction	1
2. Cadre théorique	3
2.1. La reconnaissance des mots	3
2.2. Le lexique en cours d'acquisition – particularités des apprenants de langues étrangères	6
2.3. Les mots cognats	8
2.3.1. L'effet de cognat	10
2.3.2. Comment les mots cognats sont représentés dans le lexique mental ?	12
2.3.3. L'influence de la similarité phonologique	14
2.4. La présente étude.....	19
3. Méthode	20
3.1. Les participants	21
3.2. Les stimuli	21
3.3. La procédure.....	24
4. Résultats	26
4.1. L'effet de cognat	26
4.1.1. L'analyse des taux d'erreur.....	27
4.1.2. L'analyse des temps de réaction	28
4.2. L'effet du degré de similarité phonologique des cognats	31
4.2.1. L'analyse des taux d'erreur.....	31
4.2.2. L'analyse des temps de réaction	33
5. Discussion	35
6. Conclusion	39
Bibliographie	42
Annexes	i
Annexe 1. Les mots non-cognats	i
Annexe 2. Les mots cognats P-	iii
Annexe 3. Les mots cognats P+	iv
Annexe 4. Le questionnaire en finnois et en français	v
Annexe 5. Résumé en finnois – Suomenkielinen tiivistelmä.....	viii

Liste des figures

Figure 1. Distribution des temps de réaction avant et après la transformation inverse ..	29
Figure 2. Inspection visuelle de l'hypothèse d'homoscédasticité des résidus, diagramme de dispersion (modèle lmer : statut de cognat).....	30
Figure 3. Inspection visuelle de l'hypothèse de normalité des résidus, courbe Q-Q (modèle lmer : statut de cognat).....	30
Figure 4. Inspection visuelle de l'hypothèse d'homoscédasticité des résidus, diagramme de dispersion (modèle lmer : statut de cognat compte tenu de la phonologie)	34
Figure 5. Inspection visuelle de l'hypothèse de normalité des résidus, courbe Q-Q (modèle lmer : statut de cognat compte tenu de la phonologie)	34

Liste des tableaux

Tableau 1. Les caractéristiques des stimuli.....	22
Tableau 2. Résultats de l'analyse effectuée avec un modèle linéaire généralisé à effets mixtes (Taux de réussite, effet de cognat)	27
Tableau 3. Résultats de l'ANOVA (Comparaison des modèles effectués avec glmer, effet de cognat).....	28
Tableau 4. Résultats de l'analyse effectuée avec modèle linéaire à effets mixtes (Temps de réaction, effet de cognat)	29
Tableau 5. Résultats de l'ANOVA (Comparaison des modèles effectués avec lmer, effet de cognat).....	30
Tableau 6. Résultats de l'analyse effectuée avec un modèle linéaire généralisé à effets mixtes (Taux d'erreur, effet du degré de similarité phonologique).....	32
Tableau 7. Résultats de l'ANOVA (Comparaison des modèles effectués avec glmer, effet du degré de similarité phonologique).....	32
Tableau 8. Résultats de l'analyse effectuée avec un modèle linéaire à effets mixtes (Temps de réaction, effet du degré de similarité phonologique)	33
Tableau 9. Résultats de l'ANOVA (Comparaison des modèles effectués avec lmer, effet du degré de similarité phonologique).....	34

1. Introduction

Dans ce mémoire de master, nous étudions le processus de reconnaissance des mots. Nous nous concentrerons sur les mots cognats qui sont des équivalents de traduction ayant une forme similaire d'une langue à l'autre, comme le mot *science* en anglais et en français. L'objectif de ce travail est double. D'une part, dans une perspective plus large, nous chercherons à savoir quel type d'effet les cognats produisent dans la reconnaissance. D'autre part, nous nous intéresserons plus précisément au rôle que la similarité phonologique des mots cognats entre l'anglais et le français joue dans la reconnaissance des mots.

Pour ce faire, nous effectuerons une expérience psycholinguistique avec 23 participants. Nous utiliserons la tâche de décision lexicale en français. Les temps de réaction et des taux d'erreur recueillis seront traités avec des méthodes statistiques. La population étudiée sera formée d'apprenants finnophones qui ont des connaissances en anglais et en français. Ainsi, la langue maternelle ou première (L1) sera le finnois et les langues étrangères (L2 et L3), les langues au centre de cette étude, seront l'anglais et le français (voir la définition des termes de L1, de L2 et de L3 plus bas).

L'importance du lexique pour l'expression et la compréhension d'une langue est évidente : il est l'élément central dans l'apprentissage des langues (Lindqvist, 2016 : 152). Dans les études sur le lexique mental non monolingue, les participants sont très souvent des bilingues, et les langues étudiées sont la L1 et la L2. Il existe beaucoup moins d'études sur les personnes parlant trois ou plusieurs langues (Kroll et al., 2013 : 103).

Une motivation pour notre recherche est son actualité dans le domaine. Il existe des études qui ont montré l'effet du degré de similarité phonologique sur la reconnaissance des mots cognats, mais elles ont été menées dans des langues qui ne partagent pas le même alphabet (par exemple Nakayama et al., 2014). Relativement peu d'études ont manipulé systématiquement le degré de similarité orthographique *et* phonologique des mots cognats. Comme l'ont souligné Comesaña et al. (2015), avant la parution de leur article, seulement quatre études avaient traité de manière systématique la similarité phonologique conjointement avec la similarité orthographique. Plus récemment, l'étude de Carrasco-

Ortiz et al. (2019) s'est également intéressée à cet aspect. Il convient de noter que dans ce travail, comme les langues visées (l'anglais et le français, mais aussi le finnois) ont le même alphabet, nos propos seront valables pour les langues ayant le même alphabet.

Nous étudierons donc le rôle de la phonologie dans la reconnaissance des mots cognats, un aspect peu étudié, le rapport entre les langues L2 et L3, un rapport peu étudié, avec des participants multilingues, une population peu étudiée.

Cette étude apportera des implications théoriques : les résultats peuvent aider à décrire l'architecture représentationnelle des mots cognats dans le lexique mental, et à évaluer si la similarité phonologique est un facteur qui doit être pris en compte dans les études futures sur les mots cognats.

Les questions de recherche sont les suivantes :

Q1 : Est-ce que la reconnaissance des mots cognats anglais-français diffère de celle des non-cognats chez les apprenants finnophones ? Si oui, quelle est la nature de cette différence ?

Q2 : Est-ce que le degré de similarité phonologique entre l'anglais et le français a un effet sur la reconnaissance des mots cognats ? Si oui, quelle est la nature de cet effet ?

Nous commencerons le travail par la partie théorique qui se divise en quatre sous-parties. Nous étudierons d'abord la reconnaissance des mots et la nature non sélective de ce processus (2.1). Par la suite, nous aborderons quelques particularités des apprenants de langues (2.2). Après cela, nous nous concentrerons sur les mots cognats (2.3) : l'effet de cognat (2.3.1), les représentations des cognats (2.3.2) et le rôle de la similarité phonologique dans la reconnaissance des mots cognats (2.3.3). À la fin du chapitre 2, nous présenterons les hypothèses (2.4).

Dans le chapitre 3, nous expliquerons comment nous avons recueilli notre matériau de recherche. Nous continuerons par présenter nos résultats (l'analyse statistique des données, chapitre 4) et leur interprétation ainsi que les limites éventuelles de notre recherche (5). Nous finirons avec les conclusions et offrirons quelques perspectives futures (6).

Il convient de noter ici que pour distinguer les différentes langues des personnes bi- ou multilingues, nous utilisons les termes de L1, de L2 et de L3. La définition de ces termes

n'est pas univoque. Si « L » dans cette abréviation signifie toujours la « langue », la numérotation peut relever de deux perspectives différentes. D'un côté, le numéro peut faire référence à l'ordre d'apprentissage (De Groot, 2011 : 343). Par exemple, la L1 signifie la première, la L2 la deuxième et la L3 la troisième langue apprise. Ainsi, on peut également utiliser les termes de L4, de L5, etc. D'un autre côté, la L2 peut désigner, en plus de la deuxième langue apprise, n'importe quelle langue apprise après la première langue (Hammarberg, 2001 : 22). C'est ainsi que, par exemple une quatrième langue apprise par quelqu'un peut être appelée tout simplement la L2. Dans son étude, Hammarberg distingue la L3 de la L2 pour faire référence à la langue en cours d'apprentissage et fait la remarque que dans ce cas, la L3 n'est pas forcément la troisième langue dans l'ordre d'apprentissage pour l'apprenant en question (ibid.). Nous optons pour ce point de vue et appelons l'anglais la L2 et le français la L3 pour faire la distinction entre ces langues étrangères chez les apprenants.

2. Cadre théorique

Le cadre théorique de ce mémoire est constitué de quatre sous-parties. Dans la première sous-partie, nous nous concentrerons sur le traitement lexical en général et aborderons la notion d'accès non sélectif au lexique. Dans la deuxième sous-partie, nous nous intéresserons aux particularités des apprenants de langues étrangères. Dans la troisième sous-partie, nous étudierons plus précisément les mots cognats et surtout le rôle du degré de similarité phonologique dans la reconnaissance des mots cognats. Pour finir, dans la quatrième sous-partie, nous présenterons nos hypothèses.

2.1. La reconnaissance des mots

Un mot est une suite sonore constituée de phonèmes ou une suite de lettres (graphèmes) associée à une ou plusieurs significations. Cette association doit être acquise pour pouvoir produire ou comprendre un mot (Spinelli & Ferrand, 2005 : 21, 23-24). Les mots sont considérés comme des unités linguistiques de base (id. : 20), stockés dans la mémoire à long-terme (De Groot, 2011 : 5). Une représentation lexicale est constituée de différentes parties composantes (id. : 5-6). Plus précisément, dans le lexique d'un adulte lettré, les informations suivantes de différents niveaux sont associées à un mot : orthographiques, phonologiques, morphologiques, sémantiques, et syntaxiques (Spinelli & Ferrand, 2005 :

22). On utilise la métaphore du **lexique mental** pour faire référence aux informations lexicales dans la mémoire (Treisman, 1960, cité par Spinelli & Ferrand, 2005 : 21-22).

La reconnaissance des mots est le processus le plus important pour la compréhension d'une langue (De Groot, 2011 : 155). Il s'agit d'un processus très rapide : chez un locuteur compétent, il est exécuté dans un quart de seconde. La reconnaissance des mots peut être comprise de deux façons, dont la première est constituée seulement d'une phase. Au sens strict, la reconnaissance des mots fait référence au processus, où une forme physique du mot écrit (une suite de lettres) ou parlé (un signal sonore) est associé à la représentation correspondante, c'est-à-dire, l'information stockée dans la mémoire au niveau de la forme (représentation orthographique ou représentation phonologique du mot). La deuxième définition, la définition au sens large, inclut non seulement cette correspondance mais aussi l'étape qui suit : l'activation des autres informations associées à un mot. Les autres informations sont, comme mentionné, par exemple les informations morphologiques et syntaxiques, mais la plus importante est la représentation sémantique qui est également activée. Ainsi, au sens large, la reconnaissance des mots commence par la perception de la forme physique, acoustique ou visuelle, et est complète quand toute information stockée avec la représentation lexicale est disponible. La reconnaissance des mots décrit donc toute activité mentale qui prend place pendant ce processus. Le terme d'**accès lexical** est utilisé pour désigner ce même processus, ce qui peut prêter à confusion. L'accès lexical est également le terme utilisé pour la production - dans ce cas, il réfère à tout processus entre l'intention de produire un mot et la sélection accomplie d'un élément lexical. (De Groot, 2011 : 155-156).

Il est important de comprendre que le processus de reconnaissance des mots peut être influencé par plusieurs facteurs. La fréquence (combien un mot est utilisé dans une langue) est un facteur qui produit un effet très robuste dans le traitement lexical (Spinelli & Ferrand, 2005 : 134). Les mots plus fréquents sont reconnus plus rapidement et avec moins d'erreurs que les mots peu fréquents (ibid.). La fréquence est la variable qui présente la plus forte corrélation avec les temps de réaction. Typiquement entre 30 et 40 % de la variance des temps de réaction est expliquée par la fréquence (selon une revue de méga-études sur les bases de données lexicales par Brysbaert et al., 2016 : 442). L'effet de fréquence se manifeste également lors de la reconnaissance des mots dans une L2 (Cop et al., 2015). Il est à noter que la reconnaissance des mots aussi bien dans la L1 que dans la L2 prend place dans un certain contexte (Duyck et al., 2007 : 663). Les phrases qui

entourent les mots et les relations sémantiques des mots sont des facteurs importants dans la reconnaissance des mots (Schwanenflugel & LaCount 1988 : 353). Le contexte exerce une influence dans l'accès lexical : il est plus facile de reconnaître un mot dans une phrase s'il est prévisible dans le contexte (Schwanenflugel & LaCount, 1988). De plus, le contexte peut interagir avec d'autres variables lexicales (Duyck et al., 2007 : 676). Encore, un facteur qui est important dans la reconnaissance des mots est le voisinage orthographique. Selon la définition classique de Coltheart et al. (1977), les voisins orthographiques sont des mots de même longueur qui diffèrent par une lettre. Le voisinage orthographique est une des sources importantes de la variabilité des temps de réaction (Brysbaert et al., 2016 : 443). Un grand nombre de voisins orthographiques peut rendre la reconnaissance plus rapide (ibid.) ou bien ralentir le processus d'accès lexical (van Heuven et al., 1998 : 464).

De très nombreuses études ont fourni des preuves en faveur de l'**accès lexical non sélectif** chez les bi- ou multilingues (par exemple Schwartz et al., 2007 ; Dijkstra et al., 1999 ; Lemhöfer et al., 2004 ; Duyck et al., 2007 ; Cop et al., 2017). La non-sélectivité veut dire que lors de la reconnaissance ou de la production qui se déroule dans une des langues de la personne bi- ou multilingue, les autres langues sont également activées (Kroll et al., 2013 : 104). Si l'accès au lexique était sélectif, la langue utilisée dans une certaine situation serait la seule langue activée (id. : 107). Autrement dit, dans ce cas, les autres langues de la personne bilingue ne pourraient pas être activées et ainsi n'auraient pas d'influence sur le processus d'accès lexical (id. : 107). Le terme de **l'activation parallèle** est également utilisé pour décrire l'activation des éléments lexicaux de toutes les langues (id. : 106). Toutes les langues d'une personne bi- ou multilingue sont sujettes à l'activation. Par exemple, il a été montré que toutes les trois langues d'une personne trilingue sont activées pendant la reconnaissance des mots dans la L3 (Lemhöfer et al., 2004). La L1, étant la langue la plus dominante, exerce son influence facilement sur une L2 (Kroll et al., 2013 :108). Tout de même, il est important de comprendre que même le processus d'accès lexical dans une L1 dominante peut être modifié par l'influence d'une L2, qui est une langue plus faible des deux (Schwartz et al., 2007 : 120).

L'activation parallèle de toutes les langues prend place aussi bien en ce qui concerne la compréhension que la production, comme l'ont montré les études récentes (Kroll et al., 2013 : 102). La plupart des études sur l'activation non sélective sont menées avec des bilingues et/ou concernant la L2, le nombre d'études sur les L3, L4, etc. restant plus

restreint (id. : 103). Elles utilisent très souvent des mots **cognats**, grâce à leurs propriétés particulières en comparaison avec d'autres mots (Lijewska, 2020 : 71), c'est-à-dire leur similitude en ce qui concerne le sens et la forme entre deux (ou plusieurs) langues différentes. Nous consacrons la troisième sous-partie théorique (2.3) à l'étude plus détaillée des mots cognats. Avant cela, nous aborderons quelques particularités des apprenants de langues étrangères (2.2), car dans ce mémoire, nous nous intéressons aux connaissances linguistiques d'une langue étrangère, qui exercent une influence sur une autre langue étrangère d'un apprenant.

2.2. Le lexique en cours d'acquisition – particularités des apprenants de langues étrangères

Pour un apprenant d'une langue étrangère, il est essentiel d'apprendre des mots de cette langue (De Groot, 2011 : 83 ; Lindqvist, 2016 : 152). Comme nous venons de le constater dans 2.1, pour apprendre le vocabulaire d'une langue (maternelle ou étrangère), il faut acquérir les associations entre les formes et les sens des mots. Effectivement, faire les associations efficaces entre la forme et la sémantique est essentiel pour utiliser une L2 avec aisance (Segalowitz, 2010 : 75).

En ce qui concerne l'apprenant d'une langue étrangère, la tâche d'acquisition du vocabulaire est différente que pour les enfants apprenant leur L1. Une raison pour cette différence est qu'un système phonologique a déjà été établi dans la langue maternelle. Ainsi, identifier les éléments phonémiques qui ne font pas partie du système phonologique de la L1 représentent un défi pour les apprenants (cf. par exemple Broersma, 2005). Ce n'est cependant pas la seule source des difficultés quant à l'accès lexical dans L2 concernant la phonologie (Cook et al., 2016 : 3).

En effet, dans le lexique mental d'un locuteur non natif, la représentation phonolexicale (le terme utilisé par Cook et al., 2016, pour faire référence à la représentation phonologique du mot) peut être imprécise ou incertaine. Cela veut dire que la représentation phonolexicale n'est pas constituée d'une séquence phonologique figée : elle manque de détails. Certains phonèmes de la représentation phonologique du mot peuvent être mal définis ou bien se présenter comme facultatifs. L'apprenant peut par exemple penser qu'un (ou plusieurs) des phonèmes dans la représentation ne fait pas une partie obligatoire du mot. Ce type de représentations phonolexicales qui manquent de spécificité (« floues ») peuvent mener à des associations forme-sens incorrectement

établies. Par exemple, pour l'apprenant de l'anglais L2, la représentation phonolexicale du mot *parent* peut se présenter comme [ˈpɛrə(n)t]. Le phonème [n] est pensé être optionnel. Par conséquent, l'apprenant associera deux différents mots, *parent* [ˈpɛrɪənt] et *parrot*, [ˈpɛrɪət] à la même représentation floue, ce qui peut avoir des effets néfastes pour l'accès lexical des mots présentés oralement. (Cook et al., 2016 : 3-4).

De fait, il a été indiqué que les représentations phonolexicales des apprenants sont significativement plus faibles que les représentations ortholexicales dans leur lexique. Cela paraît être le cas surtout pour les apprenants de L2 dans un environnement guidé. L'imprécision des représentations phonolexicales dépend des compétences linguistiques : les apprenants plus avancés sont susceptibles d'avoir des représentations phonolexicales plus robustes. (Veivo, 2017 : 59-60).

Remarquons aussi, que l'apprentissage d'une L3 diffère encore de celui de L2 sur certains points. En plus du système langagier de L1, l'apprenant a des connaissances préalables dans la L2, qui est potentiellement une grande source d'influence chez l'apprenant d'une L3 (Hammarberg, 2001 : 37). En ce qui concerne une L2, la source d'une éventuelle influence interlinguistique (l'influence d'une langue non-cible sur la langue cible) est naturellement la L1, la seule langue acquise. Mais, quand il s'agit d'une L3, la source du transfert peut être la L1, la L2 ou les deux (De Groot, 2011 : 343). Au cours de l'apprentissage de la L3, il semble qu'une autre langue étrangère apprise (la L2) devienne la source de l'influence interlinguistique au lieu de la L1. Ce facteur affectant l'apprentissage des langues est appelé **le facteur du statut de L2** selon Hammarberg (2001 : 23). Deux raisons, suggérées par Williams & Hammarberg (1998), sont à la base de ce facteur. Selon la première, il existe un mécanisme d'acquisition différent pour l'apprentissage de la L1 et de la L2. Ensuite, lors de l'apprentissage de la L3, le mécanisme d'apprentissage de la L2 est à nouveau activé. Selon la deuxième raison, les apprenants ne peuvent pas faire confiance à la L1 dans l'apprentissage de la L3, car la L1 n'est pas une langue étrangère. (Williams & Hammarberg, 1998 : 323). Autres éléments qui jouent également un rôle sont la compétence, la typologie et la récence (ibid.). Si le locuteur a un niveau de compétence élevé dans la L2, cette langue devient la source de l'influence interlinguistique plus facilement. Cela est également le cas si les L2 et L3 sont les langues typologiquement proches. De plus, la L2 est une source d'influence plus accessible si l'apprenant l'a utilisée récemment et a ainsi maintenu son activation. (Hammarberg, 2001 : 22-23). En outre, la maîtrise relative des langues influe sur le choix

de la langue source : une langue non-cible bien maîtrisée est plus susceptible d'avoir une influence sur la langue cible qu'une langue non-cible faiblement maîtrisée (De Groot, 2011 : 346).

2.3. Les mots cognats

Dans cette section, après la définition du terme **cognat**, nous nous pencherons sur l'effet de cognat (angl. *cognate effect*) (2.3.1) afin d'expliquer quel est le rôle de cet effet dans le traitement lexical et quels sont les facteurs principaux qui peuvent l'influencer. Nous traiterons ensuite les différentes visions sur les représentations des mots cognats dans le lexique mental (2.3.2) et finalement nous nous intéresserons à la contribution de la similarité phonologique dans le traitement lexical des mots cognats en abordant les études pertinentes pour ce point de vue (2.3.3).

Que faut-il entendre par le terme cognat ? Nous parlons de mots cognats lorsque les mots dans deux langues ont des formes similaires et ils sont des équivalents de traduction. Par exemple, le mot *hotel* en anglais et en espagnol ont le même sens et la même forme (exemple repris de De Groot, 2011 : 121). Selon la définition stricte, les mots cognats proviennent d'un ancêtre étymologique commun (Myers-Scotton, 2005 : 302 ; De Groot, 2011 : 122) et, par conséquent, les mots empruntés ne seraient pas des mots cognats. Néanmoins, comme le remarquent Lijewska (2020 : 71) et Otwinowska (2015 : 44), les psycholinguistes tendent à utiliser une définition plus large qui ignore ce critère étymologique. Nous adoptons également ce point de vue. Ainsi, pour nous, les cognats sont des équivalents de traduction qui ont une forme phonologique – et orthographique, selon les systèmes d'écriture utilisés dans les langues en question – similaire ou identique dans deux ou plusieurs langues.

Pour récapituler, la similarité des mots cognats peut être évaluée sur trois axes : l'orthographe, la phonologie et la sémantique.

- Similarité orthographique : dans les langues avec le même alphabet, les mots cognats ont différents degrés de similarité orthographique. Les paires cognats qui ont exactement la même forme orthographique entre deux ou plusieurs langues (comme le mot *rose* en anglais et en français) s'appellent des cognats identiques (Lijewska, 2020 : 77). Les cognats non-identiques sont des cognats dont les

formes orthographiques se chevauchent partiellement mais ne sont pas identiques (ibid.), comme les mots *musique* et *music* en français et en anglais.

- Similarité phonologique : la variation entre les paires cognats peut aussi se situer au niveau de la similarité phonologique, la forme phonologique pouvant être identique ou différer par un ou plusieurs phonèmes d'une langue à l'autre. Par exemple le mot *film* diffère, quant à sa phonologie, seulement légèrement entre le français (/film/) et l'anglais (/film/). En revanche, le mot *cousin*, identique quant à l'orthographe, est prononcé de manière très différente : /kʌz.n/ en anglais, mais /ku.zɛ̃/ en français.
- Similarité sémantique : la variation concerne également la similarité sémantique. Par exemple, le mot *expérience* en français peut être traduit par *experience* ou *experiment* en anglais (De Groot 2011 : 121). De Groot (ibid.) fait la remarque qu'il est courant que les chercheurs ne prennent pas en compte le degré spécifique de similarité sémantique, car le sens n'est que très rarement complètement identique dans les deux langues.

Pour identifier des mots avec les sens et les formes similaires et/ou identiques afin de les utiliser dans les expériences psycholinguistiques, on peut faire appel aux bases de données des textes traduits (comme dans Schepens et al., 2013) et, peut-être encore plus souvent, aux évaluations sur la similarité par des personnes bilingues (cf. par exemple Schwartz et al., 2007).

Quant à l'apprentissage du vocabulaire d'une langue étrangère, les cognats semblent être appris plus rapidement que les non-cognats, car en grande partie, la forme est déjà acquise dans une autre langue de l'apprenant. Ainsi, la forme d'un mot cognat est acquise avec un nombre inférieur d'essais d'acquisition que la forme d'un mot non-cognat. Toutefois, une forme similaire dans deux langues n'est pas toujours une garantie d'une similarité sémantique. C'est le cas des mots qui sont appelés **faux-amis** : la forme est la même dans les deux langues, mais le sens est différent. Par exemple, l'équivalent de traduction du mot français *actuellement* n'est pas *actually* en anglais (notre exemple). Puisqu'il existe des éléments linguistiques de ce type, qui peuvent prêter à confusion, l'apprenant ne sait pas s'il peut faire confiance en toute occasion à ses connaissances préalables dans une autre langue. (De Groot, 2011 : 119, 121-122).

2.3.1. L'effet de cognat

Les mots cognats sont souvent utilisés dans les tâches expérimentales pour étudier l'organisation du lexique bilingue. Les recherches ont trouvé que le fait qu'un mot est un cognat est un facteur qui influence le processus de reconnaissance des mots et qui produit un effet facilitateur dans le traitement des mots. Cet effet, appelé « effet de cognat »¹, consiste essentiellement en une différence entre les temps de réaction et les taux de réussite dans le cas du traitement des mots cognats et le traitement des mots non-cognats (De Groot, 2011 : 200) chez les bi- ou multilingues. Cela implique que le traitement lexical est différent pour des mots cognats et des mots non-cognats. Souvent, les mots cognats sont reconnus plus rapidement et avec un taux d'erreur moins élevé que les mots non-cognats. Toutefois, la recherche antérieure a également trouvé des effets d'inhibition (les mots cognats sont reconnus moins rapidement et avec plus d'erreurs que les non-cognats), ou pas d'effets (cf. les études de Dijkstra et al. 2010, expérience 2 et de Schwartz et al., 2007)

Il s'agit en général d'un effet robuste (Lijewska, 2020 : 76) qui prend place aussi bien dans la compréhension (cf. Cop et al., 2017 ; Dijkstra et al., 2010 ; Lemhöfer et al., 2004) que dans la production (cf. Costa et al., 2000). L'effet peut être observé avec une variété de tâches comportementales (par exemple la décision lexicale, cf. Lemhöfer et al., 2004) ainsi qu'avec l'imagerie cérébrale (par exemple l'étude des potentiels évoqués, cf. Comesaña et al., 2012). De plus, cet effet est observable dans les études utilisant aussi bien des mots isolés (cf. Comesaña et al., 2015 ; Lemhöfer et al., 2004 ; Szubko-Sitarek, 2011 ; Dijkstra et al., 2010) que des mots en contexte (cf. Duyck et al., 2007 pour les phrases, et Cop et al., 2017 pour les textes plus larges). Des facteurs différents peuvent changer la direction et/ou la taille de l'effet (Lijewska, 2020 : 77). Dans ce qui suit, nous présenterons les facteurs les plus importants pour notre étude.

Un facteur qui influence le traitement lexical est l'élément définissant des mots cognats (comme le dit Lijewska, 2020 : 77), c'est-à-dire **le degré de similarité** orthographique et phonologique. Effectivement, un aspect important concernant la similarité est la distinction entre les cognats identiques (par exemple *rose* en anglais et en français) et

¹ Le terme n'est pas figé en français. Par exemple, Manolescu & Jarema (2018) parlent de « l'effet de cognat », Voga (2020) de « l'effet cognat », et Laxén et al. (2011) plus généralement des effets facilitateurs ou des effets de facilitation pour les mots cognats. En anglais, l'effet est appelé « cognate effect » (De Groot, 2011 : 200) ou « cognate facilitation effect » (voir Lijewska, 2020 : 73 ; Peeters et al., 2013 : 316).

non-identiques (par exemple *paradis-paradise* en français et en anglais) – qui pourtant n’a pas été souvent faite dans les études du domaine, d’après Comesaña et al. (2015 : 615). Cette distinction est primordiale car les mots cognats identiques et non-identiques sont traités différemment (cf. les études de Dijkstra et al., 2010 et de Duyck et al., 2007) : l’effet de cognat est plus robuste pour les cognats identiques.

Un autre facteur qui peut influencer la taille et la direction de l’effet est **la tâche utilisée** dans l’expérience (Dijkstra et al., 2010 : 298, 300). Un des objectifs de l’étude de Dijkstra et al. (2010) était de savoir comment la tâche modifie la reconnaissance des mots cognats (l’autre objectif était le rôle du degré de similarité, voir sous-section 2.3.3.). Ils ont mené trois expériences avec les participants qui avaient le néerlandais comme la L1 et l’anglais comme la L2. Les participants avaient au moins huit ans d’expérience en anglais (id. : 288). Dans la première expérience, la tâche expérimentale était la décision lexicale (id. : 289). Dans une tâche de décision lexicale (visuelle), les participants voient des mots et des non-mots un par un et ils doivent décider s’il s’agit d’un mot réel ou non (d’où le nom décision lexicale). Concrètement, ils appuient sur un bouton « oui » quand il s’agit d’un mot réel et sur un bouton « non » quand il s’agit d’un non-mot. Le temps de réaction est mesuré à partir du moment où le stimulus apparaît jusqu’à ce que le participant réponde. (De Groot, 2011 : 157). Selon les résultats de Dijkstra et al. (2010 : 300), dans cette tâche les mots cognats ont été reconnus plus rapidement que les mots non-cognats. L’effet de cognat observé était donc facilitateur. Cependant, les effets étaient inverses dans une tâche où les participants ont dû décider à laquelle des langues utilisées (anglais ou néerlandais) un mot appartenait (expérience 2). Dans ce cas, l’effet était inhibiteur et les mots cognats ont été reconnus moins rapidement que les mots non-cognats (ibid.). Selon Poort, la différence dans la taille de l’effet selon la tâche s’explique par le niveau de représentations qu’une tâche expérimentale touche (2018 : 177).

Considérant les langues au centre d’intérêt de notre étude (la L2 et la L3 des apprenants), il est important de souligner que l’effet de cognat ne se produit pas seulement entre une L1 et une L2. Il se manifeste également entre les langues étrangères apprises plus tard dans la vie (Lemhöfer et al., 2004 : 603-604) : un effet de cognat a été observé dans la L3 des personnes trilingues, quand les mots étaient des cognats entre la L2 et la L3 (Zhu & Mok, 2020). De plus, l’effet est présent dans une L3 quand les mots sont cognats dans toutes les trois langues d’une personne trilingue (la L1, la L2 et la L3) (Szubko-Sitarek, 2011 ; Lemhöfer et al., 2004). En effet, quand un mot est cognat **en trois langues**, l’effet

est encore plus robuste qu'au cas de cognats en deux langues (Szubko-Sitarek, 2011 ; Lemhöfer et al., 2004). Outre cela, l'effet de cognat est plus robuste dans **la langue moins utilisée ou moins bien maîtrisée** de la personne bilingue (Comesaña et al., 2015 : 618). Par exemple Costa et al. (2000 : 1290) ont montré que l'effet est plus robuste dans la langue non-dominante des bilingues. Toutes les langues (par exemple la L1, la L2 et la L3) peuvent donc avoir une influence l'une sur l'autre, mais, la langue ou les langues mieux maîtrisées exercent une plus grande influence sur les langues moins bien maîtrisées.

2.3.2. Comment les mots cognats sont représentés dans le lexique mental ?

Nous venons de présenter les facteurs qui peuvent influencer la taille et/ou la direction de l'effet, mais quelle est la raison derrière l'effet dans la reconnaissance des mots ? Selon De Groot (2011 : 218), la source de l'effet de cognat dans la reconnaissance des mots est ambiguë. Elle constate que l'effet de cognat résulte de l'activation non sélective des langues (id. : 201). Cela veut dire que les équivalents de traduction de la langue non-cible sont activés : l'effet dans le traitement s'explique ainsi par l'excitation des représentations des formes similaires à un stimulus (id. : 199-200, 218). En ce cas, la co-activation des nœuds de mémoire séparés entre langues prend place (id. : 205). Selon cette approche, il existe alors des représentations différentes, séparées pour chaque langue, qui sont activées lors de la présentation d'un stimulus similaire. Cependant, toujours selon De Groot (id. : 203), il existe d'autres théories selon lesquelles les représentations des mots cognats diffèrent de celles des mots non-cognats d'une manière qualitative. Ainsi, il est en effet possible que les représentations séparées, spécifiques à une seule langue n'existent pas (et en conséquence l'effet ne résulte pas de la co-activation de ces éléments) (id. : 205). Selon cette approche, l'effet peut être considéré comme le résultat de l'activation d'une seule représentation qui est commune pour les deux langues. De Groot propose encore que, même si ces représentations étaient séparées, une possibilité serait que l'effet ne résulte pas (complètement) d'elles (ibid.). Cette question sur les représentations, et conséquemment, le traitement des mots cognats, est encore débattue – en effet, Comesaña et al. (2015 : 615) qualifient ce débat d'intense.

Comme nous l'avons indiqué ci-dessus (2.3), les mots cognats peuvent être identiques ou non-identiques quant à leur forme orthographique. Cet aspect peut être pris en compte en développant les théories sur les représentations aussi. Lemhöfer et al. (2004), qui ont

étudié la reconnaissance des cognats trilingues, offrent une explication éventuelle : les représentations seraient différentes pour les cognats identiques et pour les cognats non-identiques. Cette différence représentationnelle serait reflétée dans la taille de l'effet. Par conséquent, les cognats non-identiques seraient dotés d'un mécanisme de traitement moins efficace, ce qui résulte en un effet moins grand en comparaison avec l'effet pour les cognats identiques. (Lemhöfer et al., 2004 : 602-603). Ce mécanisme moins efficace pourrait simplement signifier que pour un cognat non-identique, il existe deux représentations orthographiques, une pour chaque langue (Peeters et al., 2013 : 316), tandis que pour un cognat identique, il en existe seulement une qui commune pour les deux langues (Dijkstra et al., 2010 : 285). Nous présenterons par la suite quelques positions théoriques différentes qui sont plus spécifiques.

Premièrement, selon Sanchez-Casas & García-Albea (2005), il existe une différence représentationnelle entre les mots cognats et les non-cognats au niveau de la morphologie. Même s'ils n'appartiennent pas à la même langue, les cognats ont une même racine morphologique (Kirsner et al., 1993 : 236). Cette position théorique est appelée « le point de vue des morphèmes partagés »² par Peeters et al. (2013 : 316, notre traduction). Selon cette position, il n'existe pas de différence entre les représentations des cognats identiques et non-identiques ; elle est donc valable pour les deux (voir Dijkstra et al., 2010 et Peeters et al., 2013). Deuxièmement, la différence qualitative entre les représentations des mots cognats et des mots non-cognats peut se situer au niveau conceptuel. Ainsi, De Groot & Nas (1991 : 117) proposent que les mots cognats partagent une même représentation sémantique et que, de plus, les équivalents de traduction, cognat ou non, sont connectés au niveau lexical. Cette position peut être appelée celle des représentations de forme liées d'une manière associative et est valable au moins pour les cognats non-identiques (cf. Dijkstra et al., 2010 : 285). Reprenons, troisièmement, une position localiste et connexionniste, selon laquelle l'effet facilitateur pour des mots cognats résulte d'un amorçage orthographique-sémantique : les représentations orthographiques qui se chevauchent dans les deux langues sont activées quand l'une ou l'autre des formes orthographiques est présentée (Dijkstra et al., 2010 : 286). Par conséquent, une représentation sémantique, qui est commune, est également activée. En effet, plus le degré de similarité orthographique est élevé, plus l'activation de la représentation sémantique est robuste (ibid.). Selon cette position, il existe, en plus de deux représentations

² Shared-morpheme view

orthographiques et phonologiques et une représentation sémantique, deux nœuds séparés qui indiquent à quelle langue le mot appartient (id. : 286 (figure 1), 300). Cette position décrit les cognats non-identiques. La quatrième position fait, à son tour, référence aux représentations des cognats identiques. Il s'agit de la position de « deux morphèmes »³ suggérée par Peeters et al. (2013 : 317, notre traduction). Selon eux (id. : 317, 327), la différence représentationnelle prend place au niveau morphologique, comme dans la proposition de Sanchez-Casas & García-Albea et Kirsner et al., mais au lieu d'une représentation morphologique commune, il existerait deux représentations morphologiques différentes à cause des propriétés différentes des mots dans différentes langues (par exemple le genre ou le nombre grammatical) et, chez les apprenants, des contextes d'apprentissage différents pour la L1 et la L2. Comme la troisième position mentionnée ci-dessus (celle des représentations de forme liées d'une manière associative), cette position présuppose que l'effet résulte d'une résonance entre les représentations (id. : 317). Toutes ces propositions se basent sur le principe du lexique non sélectif : l'activation des éléments similaires ou identiques ne peut pas être sélective quant à la langue.

D'ailleurs, en ce qui concerne les représentations des mots cognats chez les apprenants d'une langue étrangère, nous pouvons hypothétiser que les représentations au niveau de la forme (et surtout orthographiques) des mots cognats sont plus robustes que celles des non-cognats, car elles profitent de l'activation au niveau supérieur en comparant avec des non-cognats à cause de l'architecture représentationnelle. Au fait, dans l'expérience de Hintikka (2019), les participants ont réussi à attribuer un sens correct plus fréquemment pour les mots cognats que pour les non-cognats. De plus, les mots cognats écrits ont été reconnus avec un taux d'erreur moins élevé que les mots cognats présentés oralement. Le taux d'erreur pour les mots cognats présentés oralement était moins élevé que celui des mots non-cognats présentés oralement. Ainsi, les résultats ont montré l'asymétrie entre les représentations phonolexicales et ortholexicales (cf. les résultats de Veivo, 2017), aussi dans le cas des mots cognats. (Hintikka, 2019).

2.3.3. L'influence de la similarité phonologique

Dans cette étude, nous nous intéressons, en plus de l'effet de cognat, à la façon dont la forme phonologique des mots cognats affecte le traitement visuel des mots. Van Leerdam

³ Two-morpheme view

et al. (2009) et Carrasco-Ortiz et al. (2012) ont montré que quand un mot est traité visuellement, c'est-à-dire lu dans une L2, l'encodage phonologique prend place (Van Leerdam, 2009 : 328 ; Carrasco-Ortiz et al., 2012 : 538). Cet encodage phonologique est non sélectif quant aux langues (Van Leerdam, 2009 : 330 ; Carrasco-Ortiz et al. 2012 : 538). Par encodage phonologique, nous comprenons ici l'activation des représentations phonologiques en parallèle dans deux langues. Cela veut dire que les deux langues d'une personne bilingue font l'objet de l'activation du code phonologique. Les participants de l'étude de Van Leerdam et al. étaient des bilingues néerlandais-anglais, dont la L1 était le néerlandais et l'anglais était la langue étrangère la plus dominante (la L2). Carrasco-Ortiz et al. étudiaient les bilingues français-anglais, dont le français était la L1 et l'anglais la L2. Les stimuli des deux études ne contenaient pas de cognats.

Déjà en 1999, Dijkstra et al. ont signalé que la phonologie est un élément négligé dans les études sur la reconnaissance des mots chez les bilingues (Dijkstra et al., 1999). Ils ont étudié la reconnaissance des mots chez les bilingues dont la L1 était le néerlandais et la L2 l'anglais. Les mots variaient selon leur similarité orthographique, phonologique et sémantique entre les deux langues. Trois expériences ont été menées. Dans la première expérience, les participants ont effectué une tâche de démasquage progressif en anglais. Dans cette tâche, un mot cible et un masque qui cache le mot alternent sur l'écran. Progressivement, la durée de la présentation du mot est prolongée tandis que la durée de la présentation du masque est réduite. Le participant appuie sur une touche quand il reconnaît le mot. Dans la deuxième expérience, les participants choisis de la même population ont effectué une tâche de décision lexicale en anglais. La troisième expérience était la même que la deuxième, mais les participants étaient des monolingues américains d'anglais. L'étude a montré que l'orthographe, la phonologie et la sémantique jouent chacune un rôle distinct dans la reconnaissance des mots chez les bilingues. (Dijkstra et al., 1999 : 501-508, 511).

Carrasco-Ortiz et al. proposent également, à partir de leurs résultats, que la phonologie et l'orthographe contribuent différemment à la reconnaissance des mots chez les bilingues (2019 : 22). Les participants de leur étude étaient divisés en deux groupes. Les participants du premier groupe étaient des bilingues anglais-espagnol, qui avaient parlés espagnol à la maison et acquis l'anglais pendant l'enfance. Les participants du deuxième groupe étaient des apprenants en espagnol L2 et avaient l'anglais comme la L1. Les deux groupes ont effectué une tâche de décision lexicale visuelle en anglais et en espagnol, qui

contenait des mots non-cognats et des cognats. Le degré de similarité orthographique et phonologique des cognats était contrôlé. Les participants ont reconnu les mots cognats plus rapidement quand les deux codes (orthographique et phonologique) étaient différents entre eux. Cela veut dire que les participants ont reconnu les mots plus rapidement quand la forme était à la fois phonologiquement dissimilaire et orthographiquement similaire entre les deux langues, ou bien phonologiquement similaire et orthographiquement dissimilaire entre les deux langues. (2019 : 6-8, 10-11, 16). Selon Dijkstra et al. (1999 : 500) les niveaux orthographique et sémantique sont généralement au centre d'intérêt dans les études du domaine. Ainsi, inclure la contribution exacte de la phonologie dans les études de la reconnaissance des mots chez les bilingues doit être un but important (ibid.).

Il existe beaucoup d'études qui ont montré l'effet de cognat sur le traitement lexical chez les bilingues dont les langues ont des alphabets différents. Dans l'étude de Nakayama et al. (2014), le degré de similarité phonologique des mots cognats japonais-anglais était manipulé. La technique de l'amorçage masqué était utilisée dans l'étude. Les participants ont effectué une tâche de décision lexicale en anglais. Avant de voir un mot anglais (mot cible), un masque visuel (#####) était présente aux participants, suivi par la présentation très rapide (50 ms) d'un mot japonais (amorçage) (id. : 718). Les mots cognats japonais-anglais (L1-L2) étaient reconnus plus rapidement s'ils étaient phonologiquement similaires que s'ils étaient phonologiquement dissimilaires entre les deux langues (id. : 719-720). L'effet d'amorçage était plus grand si l'amorçage et le mot cible étaient phonologiquement similaires que si l'amorçage et le mot cible étaient phonologiquement dissimilaires (ibid.). Étant donné que l'orthographe similaire est complètement absente, l'effet de cognat dans les études utilisant des langues avec des alphabets différents se situe forcément au niveau des représentations phonologiques.

En ce qui concerne les langues ayant le même alphabet, selon Comesaña et al. (2015 : 615), les études réalisées jusque-là utilisant des mots cognats ont généralement pris en compte la variation de la similarité orthographique. Cependant, le rôle de la similarité phonologique est souvent laissé de côté. Peeters et al. (2013 : 327) le remarquent également : la différence phonologique dans le cas des mots cognats orthographiquement identiques n'est pas beaucoup prise en compte dans les études antérieures. Comme le soulignent Comesaña et al., avant la parution de leur article, seulement quatre études ont traité de manière systématique la similarité phonologique avec la similarité orthographique. Quant aux études réalisées jusque-là qui ont pris en compte la variation

de degré de similarité orthographique et phonologique, elles ont montré qu'elle influence le processus de reconnaissance des mots écrits. Cependant, les résultats de ces études sont inconsistants. (Comesaña et al., 2015 : 616).

Il se peut que le degré de similarité phonologique n'ait pas d'effet. Dans ce cas, un mot cognat avec un degré haut de similarité entre l'anglais et le français comme *cause* et un mot avec un degré bas de similarité comme *cousin* sont traités de la même manière, et il n'existe pas de différence dans les temps de réaction et dans les taux de réussite. Il est également possible qu'il existe une différence dans les temps de réaction et dans les taux de réussite : si les mots cognats avec un degré haut de similarité (*cause*) sont reconnus plus rapidement et avec un taux d'erreur plus bas, l'influence de la phonologie est facilitatrice. Si les mots cognats avec un degré bas de similarité (*cousin*) sont reconnus plus rapidement et avec un taux d'erreur plus bas, l'influence de la phonologie est inhibitrice.

L'objectif de l'étude de Comesaña et al. (2015) était d'étudier les représentations précises des mots cognats identiques et non-identiques ainsi que leur traitement. Leur étude était la première à manipuler le degré de similarité orthographique et phonologique ainsi que la composition de la liste des stimuli. Ils ont mené deux expériences. Les participants étaient des bilingues catalan-espagnol, qui avaient acquis les deux langues dans la petite enfance. Dans les deux expériences, les participants ont effectué une tâche de décision lexicale visuelle en espagnol. Dans la première, les mots cognats étaient orthographiquement identiques et non-identiques. Le degré de similarité des mots cognats était varié quant à leur orthographe (identique ou dissimilaire) et à leur phonologie (similaire ou dissimilaire). La première expérience était répliquée avec un groupe de contrôle des natifs monolingues pour vérifier que les autres caractéristiques des stimuli n'ont pas un effet perturbateur. Dans la deuxième expérience, tous les mots cognats étaient orthographiquement non-identiques, variant toujours quant à leur orthographe (similaire ou dissimilaire) et à leur phonologie (similaire ou dissimilaire). Selon les résultats de l'étude, un effet de cognat facilitateur (temps de réaction plus rapides) se présentait seulement quand les listes de stimuli contenaient des mots cognats orthographiquement identiques. Quant à l'effet du degré de similarité phonologique dans la première expérience, il était inhibiteur aussi bien pour les mots cognats orthographiquement identiques que pour ceux non-identiques. Les temps de réaction

étaient donc plus bas quand la phonologie était similaire. (Comesaña et al., 2015 : 617-620)

Comme mentionné dans la sous-section 2.3.1, l'étude de Dijkstra et al. (2010) visait à savoir comment la tâche et le degré de similarité affectent la reconnaissance des mots cognats. Dans la tâche de décision lexicale de l'étude, le degré de similarité phonologique avait un grand effet facilitateur qui se manifestait seulement pour les mots cognats orthographiquement identiques (2010 : 292). Dans l'étude de Schwartz et al. (2007), l'effet du degré de la similarité phonologique était également facilitateur. Les participants étaient des bilingues anglais-espagnol, qui avaient appris l'espagnol tôt dans la vie (id. : 112). Dans l'expérience, les participants ont effectué une tâche de dénomination des mots en anglais et en espagnol. Cela veut dire qu'ils ont lu à haut voix les mots présentés visuellement dans les deux langues (id. : 113-114). Selon les résultats, les mots cognats de haut degré de similarité orthographique ont été traités différemment selon leur degré de similarité phonologique. Si le degré de similarité phonologique était bas, la dénomination des mots était moins rapide et les taux de réussite étaient moins élevés (id. : 120). Selon Schwartz, cela indique que l'activation se propage de l'orthographe vers la phonologie entre les langues (*angl. feed-forward activation*) (id. : 116). Rencontrer un stimulus orthographique active le code orthographique, qui à son tour mène à l'activation des représentations phonologiques. Si ces représentations phonologiques sont dissimilaires entre eux, elles causent un effet inhibiteur qui rend le processus de reconnaissance des mots plus lent (id. : 112).

Ce que nous trouvons intéressant est la suggestion de Schwartz et al. (2007 : 120) selon laquelle les effets de la phonologie de la L2 sur le traitement lexical de la L1 seraient plus saillants chez les bilingues dont le niveau des compétences linguistiques est plus élevé. Selon ces auteurs, une meilleure maîtrise de la L2 pourrait aider à activer plus rapidement des codes lexicaux de leur L2. Peeters et al. (2013), qui pourtant n'ont pas étudié l'activation phonologique en premier lieu, font une remarque similaire en interprétant les résultats de leur expérience. Ils n'ont pas trouvé d'effet de chevauchement phonologique dans le traitement des mots cognats français-anglais (L1-L2) dans une tâche de décision lexicale en anglais. Ils proposent que si les participants avaient un niveau de compétence plus élevé de la L2, il serait plus probable d'observer l'effet (id. : 327).

Outre cela, Peeters et al. expliquent le manque d'effet par la paire des langues : ils comparent leurs résultats avec ceux de Dijkstra et al. (2010 ; expérience 1 avec décision lexicale), qui ont utilisé le néerlandais et l'anglais, et constatent que « le degré de chevauchement phonologique est beaucoup plus grand pour néerlandais-anglais » (notre traduction) que pour les cognats français-anglais utilisés dans leur expérience (2013 : 327). Effectivement, le français et l'anglais ont beaucoup de cognats orthographiques (Schepens et al. 2013 : 6), tandis que le nombre des cognats phonétiques/phonologiques est inférieur (id. : 9). Selon Schepens et al. (id. : 6-7), le français et l'anglais forment une paire de langues qui se caractérise par une relation complexe entre les systèmes de parole et les systèmes d'écriture. Il convient de noter que dans la L1 de nos participants, le finnois, les correspondances phonème-graphème sont très transparentes : en règle générale, une lettre correspond à un phonème.

Comme nous l'avons déjà constaté en 2.3.1., la tâche expérimentale peut modifier l'effet de cognat. En ce qui concerne l'influence de la similarité phonologique, même si dans une tâche de décision lexicale l'effet de la phonologie est facilitateur, dans une tâche où les participants doivent décider à laquelle des deux langues le mot appartient, le degré de similarité phonologique n'a aucun effet. Cela est également le cas pour la tâche de démasquage progressif dans l'étude de Dijkstra et al. (2010). (Dijkstra et al., 2010 : 294, 296).

2.4. La présente étude

Il convient de rappeler que nous cherchons à savoir, en premier lieu, si la reconnaissance des mots cognats anglais-français diffère de celle des non-cognats chez les apprenants finnophones. Si c'est le cas, nous visons à clarifier la nature de cette différence. En deuxième lieu, notre question est « Est-ce que le degré de similarité phonologique entre l'anglais et le français a un effet sur la reconnaissance des mots cognats ? Si oui, quelle est la nature de cet effet ? ».

En se concentrant sur ces objectifs, récapitulons ici brièvement les points principaux de notre cadre théorique (chapitre 2). Selon les études antérieures, le traitement des mots cognats est différent de celui des mots non-cognats. Les cognats sont, de règle générale, reconnus plus rapidement et avec un taux d'erreur moins élevé que les non-cognats. Cette différence dans le traitement peut s'expliquer par les différences au niveau représentationnel. De plus, il a été hypothétisé que les représentations des mots cognats

sont plus robustes que celles des mots non-cognats chez les apprenants. Puis, sachant que les mots cognats varient en ce qui concerne leur phonologie, leur orthographe (dans les langues ayant le même alphabet) et leur sémantique, il a été abordé que le degré de similarité orthographique des mots cognats a reçu plus d'attention dans les études du domaine, tandis que le rôle de la phonologie a été négligé. Les rares études qui ont systématiquement pris en compte le degré de similarité phonologique ont donné des résultats inconsistants. De plus, il est à noter que dans ces études, les langues au centre d'intérêt sont la L1 et la L2. Il existe alors une lacune de recherche concernant le rôle de la phonologie dans le traitement des mots cognats L2-L3 chez les apprenants.

En se basant sur ces points théoriques, nous formulons les hypothèses suivantes :

H1 : Les mots cognats sont traités différemment des mots non-cognats. Les mots cognats sont reconnus plus rapidement et avec un taux d'erreur moins élevé.

H2 : Les mots cognats phonologiquement similaires entre les langues sont reconnus plus rapidement et avec un taux d'erreur moins élevé que les mots cognats phonologiquement dissimilaires.

Dans les chapitres suivants, nous décrirons comment nous avons recueilli nos données (3), nous présenterons l'analyse statistique et les résultats (4) et discuterons comment ces résultats peuvent être interprétés compte tenu des théories sur lesquelles nous nous appuyons (5).

3. Méthode

Afin de savoir si un effet de cognat se manifeste et afin de prendre en compte le rôle négligé de la phonologie, nous avons décidé d'utiliser la tâche de décision lexicale chez les apprenants de français L3 et de varier le degré de similarité phonologique des mots cognats. Dans ce chapitre, nous décrirons nos participants (3.1) et les stimuli (3.2). Nous présenterons également la procédure (3.3) en introduisant l'outil pour collecter les données et la tâche de décision lexicale. Dans la section 3.3, nous expliquerons également le déroulement de notre expérience en détail.

3.1. Les participants

Les participants étaient des étudiants du département de français à l'université de Turku. Au total 23 personnes ont participé à l'étude : 20 femmes et 3 hommes, avec une moyenne d'âge de 26,5 ans (voir le questionnaire sur les données de base en Annexe 4). Ils avaient commencé à étudier le français à l'âge de 11,8 ans et l'anglais à l'âge de 8,6 ans en moyenne. Neuf participants avaient séjourné dans un pays francophone durant au moins quatre mois et six participants durant moins de trois mois. Huit participants n'avaient jamais séjourné dans un pays francophone. Tous les participants ont confirmé de parler finnois comme la langue maternelle. Tous les participants avaient une vision normale ou corrigée à la normale. Aucun participant n'avait de troubles d'apprentissage (comme la dyslexie) ni de troubles du langage et de la parole. Le score moyen de l'auto-évaluation sur trois aspects (compétence à l'oral, en compréhension et en lecture) sur l'échelle de 1 (aucune compétence) à 10 (compétence parfaite) chez les participants était 7,9 en français et 9,2 en anglais (les questions sur la compétence étaient basées sur celles utilisées dans le LEAP-Q ; Marian et al., 2007). Les participants avaient des profils linguistiques variés selon les langues qu'ils ont déclaré parler. Tous les participants ont déclaré parler le finnois, l'anglais, le français et le suédois. Parmi d'autres langues parlées étaient l'espagnol, le chinois, l'allemand, l'italien, le russe, le coréen, le polonais et le hongrois.

3.2. Les stimuli

Les stimuli utilisés dans l'expérience étaient des mots isolés (voir Annexes 1-3). L'expérience contenait au total 168 mots, dont la moitié étaient des mots réels et l'autre moitié étaient des non-mots. Toutes les caractéristiques des stimuli sont regroupées dans le Tableau 1.

Parmi les 84 mots réels, 28 étaient des mots cognats anglais-français choisis à partir du matériel fourni dans Schepens et al. (2013). Schepens et al. (2013) ont identifié automatiquement un grand nombre de mots cognats dans différentes langues en utilisant les bases de données lexicales et de traduction. Conséquemment, le critère étymologique, lié à la définition stricte de mots cognats (cf. section 2.3) n'est pas respecté quant à nos stimuli.

Tableau 1. Les caractéristiques des stimuli

	Mots cognats		Mots non-cognats	Non-mots
	P-	P+		
Nombre	14	14	56	84
Nombre de mots monosyllabiques	8	7	30	45
Nombre de mots bisyllabiques	6	7	26	39
Distance orthographique	1.000	1.000		
Distance phonologique (de min à max)	de 0.515 à 0.794	de 0.840 à 0.933		
Fréquence moyenne	57.164	59.168	55.857	
Longueur moyenne	5.64	5.71	5.68	5.68
Nombre des voisins orthographiques en moyenne	2.6	4.1	3.4	3.4

Dans le matériel de Schepens et al. (2013), la distance entre deux formes, orthographique ou phonologique, varie entre 0 et 1 de manière que les formes identiques ont la valeur la plus grande, 1. Il est donc à remarquer qu'une grande distance orthographique ou phonologique entre deux mots correspond à de plus petites valeurs, tandis qu'une petite distance orthographique ou phonologique correspond à de plus grandes valeurs. L'unité de mesure pour calculer la distance entre les mots est appelée la distance de Levenshtein. La distance de Levenshtein se base sur la théorie de l'information qui prend en compte le nombre minimal d'insertions, de suppressions et de remplacements pour modifier une chaîne de caractères à une autre (Schepens et al., 2013 : 2). Le nombre de ces modifications constitue la distance. Comme la longueur des mots varie, la distance est normalisée (id. : 5). Nous avons choisi d'inclure des mots dont l'orthographe était identique (c'est-à-dire dont la distance orthographique a une valeur de 1.000) mais dont la similarité phonologique a varié.

Parmi les 28 mots cognats, 14 étaient des mots pour lesquels la distance phonologique entre l'anglais et le français était grande. Cela veut dire que la prononciation d'un même mot était dissimilaire d'une langue à l'autre. La distance phonologique de ces mots variait entre 0.515 et 0.794. Nous appelons ces 14 mots "Groupe P-" (Annexe 2). Un exemple du mot de Groupe P- est *science*, prononcé /sjãs/ en français et /'saɪəns/ en anglais. Parmi

des 28 mots cognats, les 14 restants étaient des mots dont la distance phonologique était petite (les valeurs entre 0.840 et 0.933). Cela veut dire que la prononciation d'un même mot était similaire d'une langue à l'autre. Nous appelons ces 14 mots restants "Groupe P+" (Annexe 3). Un exemple du mot de Groupe P+ est *cause*, prononcé /koz/ en français et /kɔːz/ en anglais.

Comme nous l'avons vu dans 2.1, la reconnaissance des mots est un processus qui est influencé par une multitude de facteurs. Pour isoler l'effet de la phonologie, les autres facteurs les plus importants étaient contrôlés pour tous les mots réels (cognat ou non-cognat). Ces facteurs étaient la fréquence des mots, la longueur des mots, le nombre des syllabes et le voisinage orthographique. Ces caractéristiques étaient équilibrées entre les mots cognats phonologiquement similaires et phonologiquement dissimilaires et entre les mots cognats et non-cognats.

Nous avons rejeté les mots qui pourraient être traduits vers le finnois par un mot cognat (par exemple *tunnel*, *tunneli* en finnois). Nous avons procédé ainsi même lorsque l'orthographe en finnois était très dissimilaire (par exemple *temple*, *temppeli* en finnois). Également les mots homographes entre le finnois et le français étaient rejetés (par exemple *villa*, *vain*).

La fréquence moyenne⁴ des mots dans le Groupe P- était de 57.164 et des mots dans le Groupe P+ de 59.168. Nous avons choisi des mots d'une longueur de 4 à 7 lettres. Les mots du Groupe P- avaient une moyenne de 5.64 lettres. Les mots du Groupe P+ avaient une moyenne de 5.71 lettres. Parmi les mots du Groupe P-, huit étaient des mots monosyllabiques et six étaient des mots bisyllabiques. Quant au Groupe P+, sept mots étaient monosyllabiques et sept mots étaient bisyllabiques. En moyenne, les mots dans P- avaient 2.6 voisins orthographiques et les mots dans P+ en avaient 4.1.

Pour les mots contrôles (non-cognats), nous avons choisi 56 mots réels en français depuis la base de données lexique.org (New & Pallier, 2020). Les caractéristiques de ces mots étaient équilibrées : nous avons choisi item par item deux mots contrôles dont les caractéristiques (fréquence, longueur en lettres, nombre de syllabes et voisins orthographiques) correspondaient à celles du mot cognat. Pour vérifier qu'aucun mot

⁴ La mesure de fréquence que nous utilisons est « la fréquence du lemme selon le corpus de sous-titres par million d'occurrences » (New & Pallier, s.d. : 9) dans la base de données Lexique 3.83 (New & Pallier, 2020)

contrôle n'était cognat entre l'anglais et le français, nous avons vérifié qu'ils n'étaient pas inclus dans le matériel de Schepens et al. (2013).

Dans la tâche de décision lexicale, traditionnellement, la moitié des mots sont des mots réels et la moitié des mots sont des non-mots (cf. par exemple Comesaña et al. 2015, Peeters et al 2013, Lemhöfer et al. 2004, Szubko-Sitarek 2011). Ainsi, nous avons créé 84 non-mots à l'aide d'un générateur de non-mots WordGen (Duyck et al., 2004). Les non-mots doivent avoir des caractéristiques similaires aux mots réels quant à leur longueur et quant au nombre des syllabes, comme dans l'étude de Comesaña et al. (2015 : 617). Les non-mots doivent correspondre aux règles phonotactiques du français (De Groot, 2011 : 157) et ils ne peuvent pas être des mots réels dans une autre langue du participant (cf. Lemhöfer, Dijkstra & Michel, 2004 : 539). La longueur, le nombre des syllabes et le nombre des voisins orthographiques des non-mots que nous avons choisis correspondaient à la longueur, au nombre des syllabes et au nombre des voisins orthographiques des mots réels (cognats et non-cognats). Les non-mots qui avaient au moins un voisin orthographique étaient générés à partir d'un mot réel en français en remplaçant une des lettres par une autre, par exemple *fapine* (à partir de *famine*). Pour les mots ayant zéro voisin orthographique, cette fonctionnalité n'a pas été utilisée afin de pouvoir équilibrer le nombre de voisins orthographiques entre les différents types de mots (cognat, non-cognat, non-mot).

3.3. La procédure

L'expérience a été créée et effectuée avec PsyToolkit, une boîte à outils pour programmer et mener les expériences et les questionnaires de psychologie cognitive en ligne (Stoet, 2010 et 2017 ; <https://www.psychtoolkit.org/>, consulté le 5 octobre 2020). Le questionnaire sur le profil des participants a été créé par REDCap (Research Electronic Data Capture) hébergé à l'université de Turku (Harris et al. 2009, 2019).

Nous avons effectué une expérience psycholinguistique, basée sur la tâche de décision lexicale. Nous avons choisi d'utiliser cette tâche car elle est largement utilisée dans les études du domaine (De Groot, 2011 : 157 ; Hilton, 2003 : 3). Cependant, nous sommes consciente des inconvénients éventuels. En effet, comme le constate De Groot (2011 : 157), il s'agit d'une tâche artificielle. De plus, le temps de réaction ne reflète pas purement le processus lexical, mais contient également un stade post-lexical (le temps utilisé pour donner la réponse) (ibid.).

Dans ce qui suit, nous décrivons le déroulement de l'expérience. Avant de commencer l'expérience, les participants ont reçu une déclaration de confidentialité et des informations sur l'étude. Ils ont ensuite donné leur consentement explicite pour leur participation et pour l'utilisation des données recueillies à des fins de recherche. Les participants ont reçu des instructions écrites en finnois sur le déroulement de l'expérience et se sont engagés à porter des lunettes ou des lentilles de contact si nécessaire. On leur a ensuite expliqué d'appuyer sur la touche « a » avec l'index gauche et « l » avec l'index droit. Pour les droitiers la touche « l » était utilisée pour indiquer un mot réel et « a » pour un non-mot ; pour les gauchers le cas était inverse. Les participants ont ainsi appuyé sur la touche pour indiquer un mot réel avec leur main dominante (cf. Comesaña et al., 2015 : 618).

Avant l'expérience, le participant a eu dix essais pour s'entraîner. Les instructions brèves (quelles touches utiliser, être le plus rapide et faire le moins d'erreurs possible, l'information sur les pauses) ont été répétées avant et après le bloc d'entraînement.

L'expérience était divisée en trois blocs expérimentaux, chacun de 56 essais. Entre les blocs, il y avait une pause d'une minute (un rectangle jaune au milieu de l'écran). Un essai était construit comme il suit. Un point de fixation, qui apparaît au centre de l'écran pendant 500 ms, était suivi par un délai de 500 ms (un écran noir). Après cela, un mot apparaissait au centre de l'écran jusqu'à ce que le participant réponde. Le temps pour répondre était limité à 5000 ms. Après la réponse ou l'intervalle de 5000 ms, il y avait un délai de 500 ms (écran noir), et puis un nouvel essai commençait.

L'ordre des mots dans les blocs expérimentaux était aléatoire, mais de telle manière qu'on ne pouvait pas avoir de suite plus de quatre mots réels ou non-mots. Les participants étaient divisés dans trois groupes qui ont tous les trois fait le test avec tous les 168 mots, mais avec un ordre de mots différent.

Les temps de réaction (le temps entre l'apparition du mot et la touche du clavier) ainsi que les taux de réussite (réponse correcte/incorrecte) ont été recueillis pendant le test. Après avoir effectué la tâche de décision lexicale, les participants ont rempli un questionnaire concernant leur profil (Annexe 4).

4. Résultats

Nos données sont constituées de temps de réaction et de taux de réussite dans une tâche de décision lexicale, qui reflètent le processus d'accès lexical de la forme du mot (Hilton, 2003 : 3). Nous nous sommes intéressée aux différences entre les types des mots, non au temps de réaction absolu pour chaque mot (cf. Harley, 2008 : 171). L'analyse se base donc sur des données numériques qui ont été traitées avec des méthodes statistiques. Pour l'analyse statistique des données nous avons utilisé le logiciel R et le package lme4 (version 1.1-26 ; Bates et al., 2015). Nous présenterons d'abord les résultats des analyses concernant l'effet de cognat (4.1) et continuerons par les résultats des analyses sur le degré de similarité phonologique des mots cognats (4.2).

Le taux d'erreur pour les 84 mots réels était de 5,9 %. Les taux d'erreurs élevés sont typiques pour les tâches faites en L2 ou L3 (Lemhöfer et al., 2004 : 595). Trois participants étaient exclus de l'analyse ayant un taux d'erreur supérieur de 15 % (cf. Dijkstra et al., 2010 : 290 ; Comesaña et al., 2015 : 618). Nous avons ainsi analysé les données de 20 participants. Seulement le traitement des mots réels a été analysé.

Dans nos analyses statistiques, nous avons utilisé les modèles linéaires (généralisés) à effets mixtes qui permettent d'étudier les facteurs fixes au centre d'intérêt tout en tenant compte de la variabilité causée par les facteurs aléatoires comme les différences inhérentes entre les participants. Pour les analyses des temps de réaction, nous avons utilisé les modèles linéaires à effets mixtes (la fonction `lme4::lmer` du logiciel R), car la variable dépendante est continue. Pour les analyses des taux d'erreur, nous avons utilisé les modèles linéaires généralisés à effets mixtes (la fonction `lme4::glmer` du logiciel R), car la variable dépendante est binaire. Nous avons utilisé les méthodes d'estimation par défaut. Pour la fonction `lme4::lmer`, il s'agit de REML (Restricted Maximum Likelihood) et pour la fonction `lme4::glmer`, de l'approximation de Laplace.

4.1. L'effet de cognat

Dans cette section, nous présenterons les résultats de l'analyse des taux d'erreur et de l'analyse des temps de réaction pour répondre à notre première question de recherche (Q1) : est-ce que la reconnaissance des mots cognats anglais-français diffère de celle des non-cognats chez les apprenants finnophones ? Si oui, quelle est la nature de cette différence ? Nous avons formulé l'hypothèse suivante (H1) : Les mots cognats sont

traités différemment. Ils sont reconnus plus rapidement et avec un taux d'erreur moins élevé.

4.1.1. L'analyse des taux d'erreur

Afin d'effectuer l'analyse, nous avons exclu six observations à cause d'une réponse manquante : dans ce cas, les participants n'ont pas répondu dans le délai imparti de 5000 millisecondes. Conséquemment, 1674 observations ont été analysés avec un modèle linéaire généralisé à effets mixtes (GLMM ou general linear mixed model). Dans un GLMM, la variable dépendante est binaire : ici, c'était l'exactitude des réponses (réponse correcte/réponse incorrecte). Le facteur fixe du modèle était 'le statut de cognat' (cognat vs. non-cognat). Les facteurs aléatoires étaient le participant et l'item.

En ce qui concerne les mots cognats, le taux erreur était de 7.5 %. Quant aux mots non-cognat, le taux d'erreur était de 5.1 %. Les résultats de l'analyse effectuée avec un GLMM sont présentés dans le tableau 2.

Tableau 2. Résultats de l'analyse effectuée avec un modèle linéaire généralisé à effets mixtes (Taux de réussite, effet de cognat)

Effets fixes	Estim.	Erreur stand.	Valeur z	Pr(> z)
(Intercept)	-4.7496	0.5356	-8.868	<2 ^e -16 ***
Cognat	1.1238	0.6081	1.848	0.0646 .

Signification : 0 **** 0.001 *** 0.01 ** 0.05 ' . 0.1 ' ' 1

Les résultats de l'analyse montrent que les participants ont donné moins de réponses erronées pour les mots non-cognats que pour les mots cognats ($z = 1.848$). Cette différence est marginalement significative au niveau alpha de 0.05 ($p = 0.0646$). Ces résultats suggèrent que la reconnaissance des mots cognats diffère de celle des mots non-cognats. Une différence représentationnelle présupposée entre les mots cognats et non-cognats a donc influencé le traitement lexical. L'effet de cognat observé, qui est marginalement significatif, est de nature inhibitrice.

Nous avons ajouté l'auto-évaluation des compétences langagières en français au modèle afin d'examiner si cette variable apporte une amélioration à l'analyse. Les résultats d'ANOVA sont illustrés dans le tableau 3. Ils montrent que l'ajout de l'auto-évaluation des compétences langagières affecte le modèle d'une manière marginalement significative ($Pr(>Chisq) = 0.088$). Les valeurs AIC et BIC donnent des résultats différents : selon BIC, le modèle original est meilleur, tandis que selon AIC, l'ajout de

cette variable améliore le modèle. Même s'il était très intéressant de voir quel rôle, le cas échéant, la compétence joue dans l'effet, cela dépasse le cadre de ce travail et nous ne présenterons donc pas ici d'autres résultats de cette analyse.

Tableau 3. Résultats de l'ANOVA (Comparaison des modèles effectués avec glmer, effet de cognat)

Modèle	Valeur AIC	Valeur BIC	Pr (>Chisq)
glmer : statut de cognat	636.27	657.96	
glmer statut de cognat * auto-évaluation	635.40	667.93	0.08756

4.1.2. L'analyse des temps de réaction

Pour l'analyse des temps de réaction, nous avons exclu cinq mots pour lesquels 30 % des participants ont donné une réponse erronée : baleine, drap, humble, missile et nain (100 observations au total). Six observations supplémentaires ont été exclues parce que leur temps de réaction était supérieur à 5000 millisecondes qui était le temps limite pour donner une réponse. Par la suite, nous avons exclu de l'analyse les observations avec un temps de réaction éloigné de plus de deux écarts-types de la moyenne par participant (79 observations). Finalement, pour analyser les temps de réaction, nous avons inclus seulement les réponses correctes, ce qui nous a laissé avec 1452 observations.

Avant d'appliquer un modèle statistique, nous avons inspecté la distribution des temps de réaction pour savoir si l'échantillon suit une distribution normale. Généralement, les temps de réaction ont une tendance d'être positivement asymétriques. L'inspection visuelle ainsi que le coefficient d'asymétrie (3.43) ont montré une asymétrie fortement positive. Effectivement, selon un test de normalité, l'échantillon ne suit pas une distribution normale. C'est pourquoi nous avons opté pour une transformation des données. Nous avons comparé visuellement la distribution des TR après les transformations log, racine carrée et inverse. La transformation inverse (1/TR) donne une distribution la plus proche à la loi normale (le coefficient d'asymétrie est de -0.33). Les distributions de la variable temps de réaction avant et après sont représentées dans la figure 1.

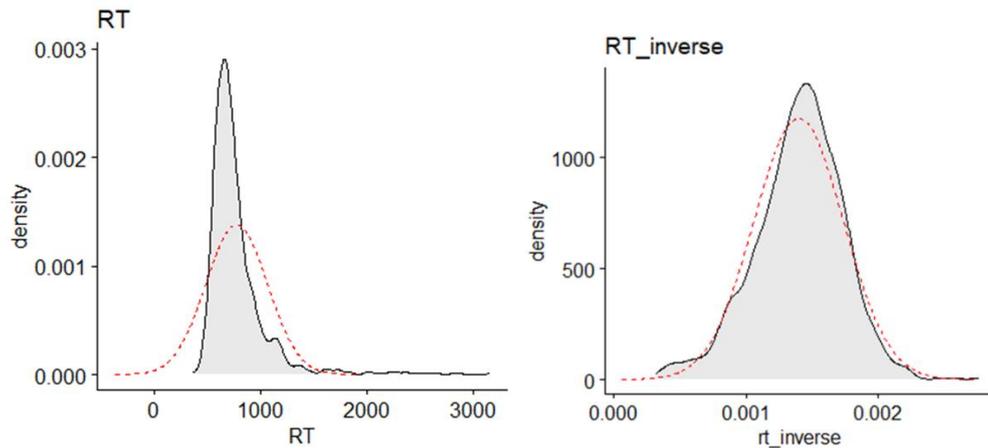


Figure 1. Distribution des temps de réaction avant et après la transformation inverse

Nous cherchions à savoir si les temps de réaction des mots cognats diffèrent des mots non-cognats. Pour ce faire, nous avons utilisé un modèle linéaire à effets mixtes. La variable dépendante était le temps de réaction (TR). Le facteur fixe dans notre modèle était le statut de cognat (est-ce que le mot est cognat ou non). Pour prendre en compte la variabilité liée à chaque item et à chaque participant, nous avons inclus ‘participant’ et ‘item’ dans le modèle comme facteurs aléatoires.

Le temps de réaction moyenne pour les mots cognats était 795 millisecondes et 764 millisecondes pour les non-cognats. Les résultats de l’analyse statistique sont illustrés dans le tableau 4.

Tableau 4. Résultats de l’analyse effectuée avec modèle linéaire à effets mixtes (Temps de réaction, effet de cognat)

Effets fixes	Estim.	Erreur stand.	Valeur <i>t</i>
(Intercept)	0.001398	0.00004302	32.488
Cognat	-0.00001636	0.00002713	-0.603

Les résultats nous montrent que contrairement à notre hypothèse, les mots cognats ont été reconnus plus lentement que les mots non-cognats ($t = -0.603$). La direction de l’effet de cognat aurait donc été inhibitrice. Cependant, la différence entre les temps de réaction des mots cognats et des mots non-cognats est petite, et nous pouvons constater qu’elle n’est pas statistiquement significative. Le fait qu’un mot a le même sens et la même forme dans les deux langues n’a pas eu un effet dans le traitement lexical. Cela veut dire que le traitement lexical n’est pas influencé par l’activation présumée qui prend place au niveau des représentations.

Les hypothèses d'homoscédasticité et de normalité des résidus du modèle ont été vérifiées visuellement. Cette inspection visuelle n'a révélé aucune tendance évidente (voir figures 2 et 3). Cela justifie également l'utilisation d'une transformation (inverse) des données.

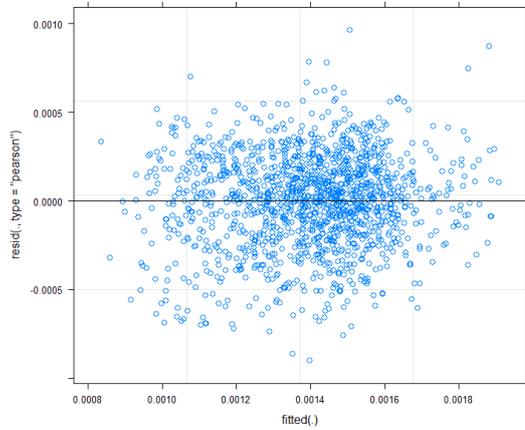


Figure 2. Inspection visuelle de l'hypothèse d'homoscédasticité des résidus, diagramme de dispersion (modèle lmer : statut de cognat)

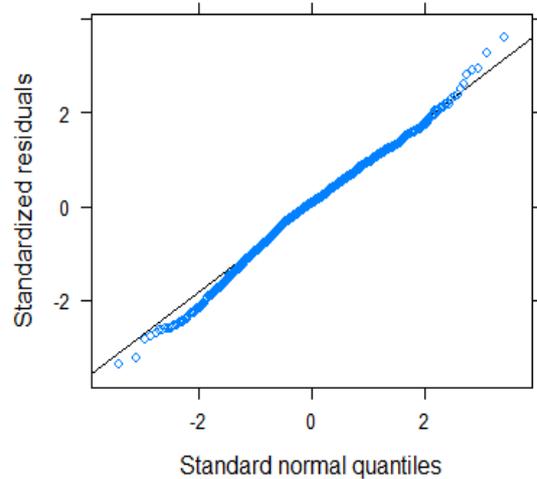


Figure 3. Inspection visuelle de l'hypothèse de normalité des résidus, courbe Q-Q (modèle lmer : statut de cognat)

Comme dans l'analyse des taux d'erreur, nous avons ajouté l'auto-évaluation des compétences du français à notre modèle. Les résultats de la comparaison ANOVA des modèles, illustrés dans le tableau 5, indiquent que le modèle original est légèrement meilleur (les valeurs AIC et BIC sont plus petites pour le modèle original). La différence n'est pas statistiquement significative au niveau alpha de 0.05 et ainsi nous pouvons constater que l'ajustement du modèle n'est pas significativement amélioré par l'ajout de cette variable.

Tableau 5. Résultats de l'ANOVA (Comparaison des modèles effectués avec lmer, effet de cognat)

Modèle	Valeur AIC	Valeur BIC	Pr (>Chisq)
lmer : statut de cognat	-19604	-19578	
lmer : statut de cognat * auto-évaluation	-19601	-19564	0.752

L'analyse des taux d'erreur et l'analyse des temps de réaction ont donné des résultats différents. Nous fournirons des explications pour ces résultats dans la discussion (5.). Avant cela, nous examinerons l'effet du degré de similarité phonologique.

4.2. L'effet du degré de similarité phonologique des cognats

Comme nous venons de le constater dans la sous-section 4.1.1., selon l'analyse des taux d'erreur, l'effet de cognat marginalement significatif était observé. La direction de l'effet était inhibitrice. Cependant, selon l'analyse des temps de réaction (4.1.2), la différence entre les TR pour les mots cognats et les mots non-cognats n'était pas statistiquement significatif et l'effet de cognat n'était pas être observé chez nos participants.

Dans cette partie, nous nous intéressons plus spécifiquement à la différence entre les cognats phonologiquement similaires (P+) et dissimilaires (P-). Dans ce qui suit, nous présenterons des résultats de l'analyse des taux d'erreur (4.2.1) et de l'analyse des temps de réaction (4.2.2) pour répondre à notre deuxième question de recherche (Q2) : Est-ce que le degré de similarité phonologique entre l'anglais et le français a un effet sur la reconnaissance des mots cognats ? Si oui, quelle est la nature de cet effet ? En autres termes, nous tenterons à clarifier si le traitement lexical des mots est différent pour les cognats phonologiquement similaires (P+) et dissimilaires (P-), alors si l'effet de cognat apparaît différemment en fonction de la similarité phonologique. Nous avons formulé l'hypothèse suivante (H2) : les mots cognats phonologiquement similaires sont reconnus plus rapidement et avec un taux d'erreur moins élevé que les mots cognats phonologiquement dissimilaires.

4.2.1. L'analyse des taux d'erreur

Cette analyse a été menée avec les mêmes données que dans 4.1.1. Comme pour l'analyse des taux d'erreur pour observer un effet de cognat, nous avons utilisé un modèle linéaire généralisé à effets mixtes (GLMM). La variable dépendante était l'exactitude des réponses (réponse incorrecte/réponse correcte). Le facteur fixe dans le modèle était le statut de cognat tenant compte de la similarité phonologique : le mot est non-cognat, cognat phonologiquement similaire (P+) ou cognat phonologiquement dissimilaire (P-). Les facteurs aléatoires étaient de nouveau le participant et l'item.

6,1 % de réponses données étaient incorrectes en ce qui concerne les mots cognats phonologiquement similaires (P+). Quant aux mots cognats phonologiquement dissimilaires (P-), le taux d'erreur était de 9,0 %. Le taux d'erreur des mots non-cognats était de 5,1 %. Les résultats de l'analyse réalisée avec un GLMM sont illustrés dans le tableau 6.

Tableau 6. Résultats de l'analyse effectuée avec un modèle linéaire généralisé à effets mixtes (Taux d'erreur, effet du degré de similarité phonologique)

Effets fixes	Estim.	Erreur stand.	Valeur z	Pr (> z)
(Intercept)	-4.7349	0.5324	-8.893	<2 ^e -16***
P+	0.7909	0.7677	1.030	0.3029
P-	1.4131	0.7463	1.893	0.0583 .

Signification : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Le taux d'erreur était plus élevé pour les P+ ($z = 1.030$) en comparant avec les mots non-cognats. Mais, cette différence était petite et elle n'est pas statistiquement significative au niveau alpha de 0.05 ($p = 0.3029$). Ainsi, il nous paraît que la reconnaissance des mots cognats phonologiquement similaires n'a pas différée de celle de non-cognats : aucun effet de cognat n'était observé. Cependant, les mots cognats P- étaient également reconnus avec un taux d'erreur plus élevé que les mots non-cognats ($z = 1.893$), mais ce résultat était marginalement significatif au niveau alpha de 0.05 ($p = 0.0583$). Les participants ont donc arrivé à attribuer une réponse correcte plus souvent pour les non-cognats que pour les cognats phonologiquement dissimilaires entre l'anglais et le français. Les mots cognats phonologiquement dissimilaires ont donc été reconnues différemment des mots non-cognats. Cet effet de cognat observé pour les mots phonologiquement dissimilaires était de nature inhibitrice.

De nouveau, nous avons ajouté l'auto-évaluation des compétences du français au modèle pour voir si l'ajout de cette variable apporte une amélioration à l'ajustement du modèle. Comme représenté dans le tableau 7, le modèle original est meilleur avec les valeurs AIC et BIC légèrement plus petites. La différence entre les modèles n'est pas statistiquement significative au niveau alpha de 0.05. Ajouter la variable n'a pas amélioré l'ajustement du modèle.

Tableau 7. Résultats de l'ANOVA (Comparaison des modèles effectués avec glmer, effet du degré de similarité phonologique)

Modèle	Valeur AIC	Valeur BIC	Pr (>Chisq)
glmer : statut de cognat compte tenu de la phonologie	637.80	664.91	
glmer : statut de cognat compte tenu de la phonologie * auto-évaluation	638.17	681.55	0.131

4.2.2. L'analyse des temps de réaction

Pour cette analyse, nous avons utilisé les mêmes données que précédemment dans la sous-section 4.1.2 (1452 observations, les temps de réaction transformés avec la transformation inverse $1/TR$).

Cette fois-ci, le facteur fixe était le statut de cognat tenant compte de la similarité phonologique (c'est-à-dire si les mots sont cognats phonologiquement similaires (P+), cognats phonologiquement dissimilaires (P-) ou non-cognats). De nouveau, nous nous sommes intéressée aux temps de réaction (variable dépendante) et nous avons les mêmes facteurs aléatoires que dans le modèle linéaire à effets mixtes précédent, participant et item.

La moyenne des temps de réaction était 835 millisecondes pour les mots cognats P+, 754 ms pour les cognats P- et 764 ms pour les non-cognats. Les résultats de cette analyse sont illustrés dans le tableau 8.

Tableau 8. Résultats de l'analyse effectuée avec un modèle linéaire à effets mixtes (Temps de réaction, effet du degré de similarité phonologique)

Effets fixes	Estim.	Erreur stand.	Valeur t
(Intercept)	0.001398	0.00004291	32.572
P+	-0.0000599	0.00003435	-1.744
P-	0.00002742	0.00003449	0.795

Comme catégorie de référence, nous avons les mots non-cognats. Selon les résultats, les mots cognats phonologiquement similaires ont été reconnus plus lentement ($t = -1.744$) que les mots non-cognats. Cet effet est marginalement significatif. Il est intéressant de noter que la direction a changé quant aux mots cognats dissimilaires : ils ont été reconnus plus rapidement ($t = 0.795$) que les mots non-cognats. Cependant, cette différence n'est pas statistiquement significative. Nous pouvons voir clairement une différence entre le traitement des mots cognats P+ et P-. Ces résultats impliquent qu'un effet de cognat est modifié par le degré de similarité phonologique. Le fait que les formes phonologiques d'un mot cognat étaient similaires entre eux avait un effet inhibiteur dans la reconnaissance des mots.

Nous avons vérifié visuellement les hypothèses d'homoscédasticité et de normalité des résidus du modèle. Cette inspection visuelle n'a révélé aucune tendance évidente (voir figures 4 et 5). Cela justifie, de nouveau, l'utilisation d'une transformation des données.

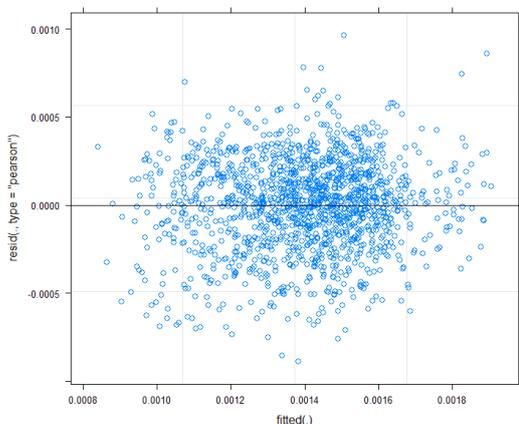


Figure 4. Inspection visuelle de l'hypothèse d'homoscédasticité des résidus, diagramme de dispersion (modèle lmer : statut de cognat compte tenu de la phonologie)

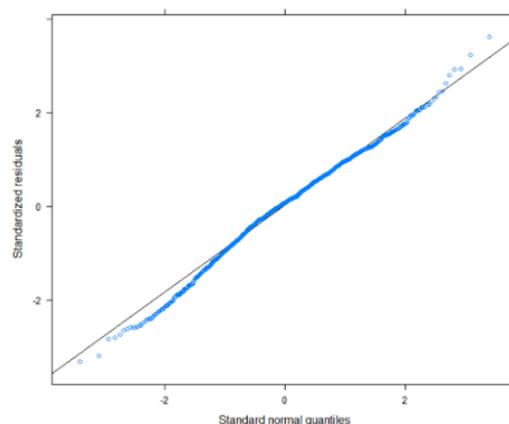


Figure 5. Inspection visuelle de l'hypothèse de normalité des résidus, courbe Q-Q (modèle lmer : statut de cognat compte tenu de la phonologie)

Nous avons ajouté le score d'auto-évaluation également à ce modèle. Les résultats de la comparaison ANOVA entre le modèle original et le modèle avec le score d'auto-évaluation (illustrés dans le tableau 9) indiquent que le modèle original est meilleur (les valeurs AIC et BIC plus petites). La différence entre les deux modèles n'est pas statistiquement significative au niveau alpha de 0.05 et ainsi nous pouvons constater que l'ajout de la variable n'a pas amélioré l'ajustement du modèle.

Tableau 9. Résultats de l'ANOVA (Comparaison des modèles effectués avec lmer, effet du degré de similarité phonologique)

Modèle	Valeur AIC	Valeur BIC	Pr (>Chisq)
lmer : statut de cognat compte tenu de la phonologie	-19606	-19574	
lmer : statut de cognat compte tenu de la phonologie * auto-évaluation	-19601	-19554	0.7073

Dans le chapitre suivant, nous discuterons de l'interprétation des résultats présentés dans ce chapitre (4).

5. Discussion

Dans ce chapitre, après un bref résumé des résultats, nous discuterons comment les résultats peuvent être interprétés compte tenu des recherches antérieures présentées dans le chapitre 2. Nous aborderons également les limitations éventuelles de notre étude.

Nous avons trouvé un effet de cognat inhibiteur statistiquement marginal dans notre analyse des taux d'erreur mais aucun effet de cognat dans notre analyse des temps de réaction. Lorsque nous avons examiné l'effet du chevauchement phonologique, dans notre analyse des taux d'erreur, nous n'avons observé aucun effet quant aux cognats phonétiquement similaires (P+). En revanche, nous avons observé un effet inhibiteur statistiquement marginal pour les cognats phonétiquement dissimilaires (P-). Dans l'analyse des temps de réaction, le résultat est inverse. Les cognats dissimilaires (P-) ne différaient pas des non-cognats, mais les cognats similaires (P+) produisaient un effet inhibiteur statistiquement marginal dans la reconnaissance des mots.

Tout d'abord, nous pouvons conclure que les apprenants ont activé une langue non-cible pendant la reconnaissance des mots. Ainsi, nos résultats sont une preuve en faveur de l'accès lexical non sélectif. De plus, notre étude montre que les codes phonologiques d'une autre langue étrangère sont activés lors de la lecture de formes orthographiques d'une langue étrangère. En outre, notre étude confirme que la L2 peut affecter la reconnaissance des mots dans la L3. Compte tenu de la robustesse de l'hypothèse de l'accès non sélectif, ce résultat n'est peut-être pas surprenant, mais comme la majorité des études ont utilisé la L1 et la L2, il convient de le souligner.

Considérons premièrement plus en détail l'interprétation des résultats concernant l'effet de cognat (4.1). Nous avons observé un effet de cognat inhibiteur marginalement significatif dans notre analyse des taux de réussite. Ce résultat n'est pas en accord avec les études antérieures réalisées avec une décision lexicale où seulement une langue est présente pendant la tâche (le français dans la nôtre ; voir par exemple Dijkstra, 2010, expérience 1 ; Poort, 2018, expérience 1 ; Lemhöfer et al. 2004). Nous pouvons toutefois offrir certaines explications qui pourraient rendre compte de ce résultat.

Rappelons que les effets inhibiteurs peuvent être dus à une tâche différente, comme celle utilisée dans l'expérience 2 de Dijkstra et al. (2010). Il s'agissait d'une tâche où il faut décider à laquelle des deux langues un mot appartient (id. : 293). Notre résultat d'un effet

inhibiteur des mots cognats dans une tâche de décision lexicale unilingue est très similaire à celui de Zhu & Mok (2020, expérience 2). Dans cette expérience, les participants ont effectué une tâche de décision lexicale en leur L3 (l'allemand). Des mots cognats L2-L3 (anglais-allemand) étaient inclus dans l'expérience (id. : 452). Dans l'analyse des temps de réaction, un effet de cognat n'a pas été observé. Cependant, les taux d'erreur indiquaient un effet inhibiteur mais seulement pour les mots cognats orthographiquement identiques. Les auteurs suggèrent qu'en voyant les mots, les participants les ont reconnus en tant que mots en anglais, la langue non-cible. Il a fallu, par la suite, supprimer l'activation de l'anglais, ce qui devient une tâche exigeante sur le plan cognitif. (id. : 455). Les auteurs proposent que l'effet soit dû à un niveau de compétence limité (id. : 464). Nos participants sont plus compétents que ceux dans l'étude de Zhu & Mok, mais cette explication nous paraît quand même logique, car les cognats dans notre étude étaient tous identiques et avaient donc entièrement la même forme que les mots en anglais. Pensant qu'il faut éviter l'anglais dans la tâche, nos participants ont pu changer leur manière de répondre. Ainsi, la tâche de décision lexicale peut commencer à être semblable à la tâche où il faut décider à quelle des deux langues un mot appartient (Dijkstra et al., 2010, expérience 2 ; voir aussi Zhu & Mok, 2020 : 464). En adoptant éventuellement une stratégie de réponse différent, d'une part, nos participants n'auraient pas profité de l'activation supérieure au niveau conceptuel propre aux cognats, parce qu'ils se seraient penchés davantage sur la forme (cf. Zhu & Mok, 2020 : 464). D'autre part, si nous présumons qu'il existe des nœuds séparés liés aux représentations lexicales qui indiquent à quelle langue un mot appartient (selon la position localiste et connexionniste, voir la sous-section 2.3.2), une autre possibilité est qu'une activation supérieure des représentations des cognats par rapport aux non-cognats mène à l'activation supérieure du nœud de la langue non-cible. Ainsi, c'est le temps utilisé pour supprimer cette activation du nœud de la langue non-cible qui rend la direction de l'effet inhibitrice (cf. Dijkstra et al., 2010 : 299).

Toutefois, il existe une autre explication pour ce résultat atypique. Comme l'ont indiqué Comesaña et al. (2015 : 623), l'effet que les différents niveaux de la similarité phonologique des mots cognats ont sur le traitement lexical peut être un facteur à l'origine des résultats contradictoires dans la recherche qui utilise les mots cognats, car ce facteur n'a pas été souvent pris en compte dans les études du domaine. Dans notre première analyse des taux d'erreur, les cognats phonologiquement similaires et phonologiquement

dissimilaires n'ont pas été analysés séparément. Pourtant, il est possible que les différents degrés de similarité affectent le traitement lexical différemment et influencent les résultats. C'est pourquoi il faudrait être prudent lors de l'interprétation des résultats de cette première analyse.

Avec notre deuxième question de recherche, nous visions à savoir si la reconnaissance des mots était modifiée par l'effet de similarité phonologique. Passons maintenant à l'interprétation des résultats de ces analyses (présentés dans la section 4.2). Selon les résultats de l'analyse des taux d'erreur, les mots cognats qui ont des formes phonologiques plus distinctes entre elles dans les deux langues (P-) avaient un effet inhibiteur sur la reconnaissance et donc ils ont causé plus d'erreurs chez les participants. Ces résultats sont partiellement en accord avec ceux de Dijkstra et al. (2010, expérience 1) et Schwartz et al. (2007). Dans ces expériences, les participants ont reconnu les mots cognats (orthographiquement identiques dans l'étude de Dijkstra et al., 2010 et orthographiquement très similaires dans l'étude de Schwartz et al., 2007) plus rapidement lorsqu'ils étaient phonologiquement similaires que lorsqu'ils étaient phonologiquement plus distincts (Dijkstra et al., 2010 : 292 ; Schwartz et al., 2007 : 116). Deux formes phonologiques plus distinctes entre les deux langues ont donc ralenti les temps de réaction. Ainsi, la direction de l'effet que nous avons obtenu pour les mots phonologiquement plus dissimilaires (P-) est conforme à ce type de résultat. Il est quand même à noter que les formes phonologiques similaires entre elles (P+) n'ont pas produit un effet facilitateur, comme dans les études de Dijkstra et al. (2010) et de Schwartz et al. (2007). Comme mentionné dans la sous-section 2.3.3, Schwartz et al. (2007) parlent de l'activation qui se propage de l'orthographe vers la phonologie (*feed-forward activation*). L'activation du code orthographique suscite alors l'activation des représentations phonologiques. Si ces représentations phonologiques sont dissimilaires entre elles, elles causent un effet inhibiteur. Nous pouvons hypothétiser qu'un processus de ce type prend place dans le lexique mental chez nos participants : deux représentations phonologiques s'activent et le processus est ralenti par la dissimilarité des représentations.

Selon les résultats de l'analyse des temps de réaction, les cognats phonologiquement similaires entre les deux langues (P+) étaient reconnus plus lentement (effet marginalement significatif). Ce résultat est inattendu, car il est opposé par rapport aux résultats de l'analyse des taux d'erreur. Un effet inhibiteur de ce type, cette-fois ci produit par les cognats phonologiquement similaires, est quand même conforme à une étude

antérieure (l'expérience 1 de Comesaña et al., 2015). Rappelons que dans leur étude, les formes phonologiques similaires entre elles ont produit des effets inhibiteurs aussi bien pour les cognats identiques que non-identiques. Même si ce type de résultat n'est pas conforme aux résultats de Schwartz et al. (2007) et de Dijkstra et al. (2010), selon Comesaña et al. (2015 : 619) l'idée de deux représentations phonologiques qui s'activent et entrent en compétition n'est pas exclue.

Les résultats des analyses de temps de réaction et les taux d'erreur sont contradictoires : en ce qui concerne l'effet de cognat (sans prendre compte le degré de similarité phonologique), il était observé seulement avec l'analyse des taux d'erreur. Quant à la similarité phonologique, les analyses des temps de réaction et des taux d'erreur ont donné des résultats différents. Trouver des interprétations qui pourraient expliquer en même temps tous ces résultats contradictoires, qui sont quand même obtenus dans la même tâche, s'avère une tâche compliquée. Nous nous demandons si ces résultats contradictoires sont dus à la plateforme utilisée. Bien que PsyToolkit (Stoet, 2010 et 2017) ait été prouvée fiable dans une étude utilisant une tâche complexe (Kim et al., 2019), nous sommes consciente qu'elle peut influencer la fiabilité de nos résultats. Il est à noter que l'analyse des taux d'erreur est une mesure plus robuste que l'analyse des temps de réaction. Les différents appareils, navigateurs Internet ou autres questions liées au matériel ou au logiciel n'affectent pas la fiabilité des taux d'erreur, mais ce sont des facteurs susceptibles à influencer les temps de réaction. Par conséquent, nous nous appuyons davantage sur nos données sur les taux d'erreur que sur nos données sur les temps de réaction que nous devons interpréter avec beaucoup de prudence.

En plus les défis que la collecte des données sur Internet pose pour notre étude, dans ce qui suit, nous aborderons d'autres limites éventuelles de cette étude. Un défi évident dans notre étude est que les profils linguistiques de nos participants étaient très variés : beaucoup d'entre eux parlent plusieurs langues étrangères qui peuvent être activées lors l'accès lexical et ainsi influencer le traitement lexical. Tenir compte de toutes les langues différentes en choisissant les items est une tâche très compliquée et cela pose des défis pour la recherche sur les personnes multilingues (cf. Kroll, 2013 : 109-110). La majorité de nos participants ont déclaré l'anglais en tant que leur langue étrangère la plus dominante et elle était également la première langue étrangère dans l'ordre d'apprentissage chez la majorité de nos participants. Nous pouvons donc nous attendre à ce que cette langue ait une influence la plus robuste parmi les différentes langues

étrangères. Il serait utile de reproduire ce type d'étude avec des participants « purement » trilingues.

Une limite potentielle est le nombre de participants. Encore, un autre aspect que nous devons mentionner est l'influence de la L1 des participants. Selon l'activation du code phonologique parallèle dans toutes les langues de la personne bilingue (voir sous-section 2.3.3, Carrasco-Ortiz et al., 2012 ; Van Leerdam, 2009), le code phonologique du finnois serait activé aussi. Comme nous l'avons constaté dans 2.3.3, l'anglais et le français forment une paire de langues dont la relation entre les systèmes orthographique et phonologique est complexe. Cependant, le finnois a des correspondances graphème-phonèmes très consistantes. Les systèmes orthographique et phonologique du finnois peuvent donc avoir une influence sur la manière de traiter les stimuli visuels chez les participants. Dans le cadre de cette étude, cet aspect n'a pas été pris en compte, mais dans les études futures il conviendrait de le faire.

De toute façon, comme l'analyse des taux d'erreur et l'analyse des temps de réaction ont montré, il semble exister une différence entre les mots cognats phonologiquement similaires et dissimilaires dans nos données. Nous présenterons nos conclusions dans le chapitre suivant (6.).

6. Conclusion

Dans ce mémoire de master, nous cherchions à savoir si un effet de cognat se manifeste dans la reconnaissance des mots en français L3 chez les apprenants finnophones ayant l'anglais comme la L2 et le français comme la L3. De plus, nous visions à clarifier si la similarité phonologique des mots cognats jouait un rôle dans la reconnaissance. Pour répondre à ces questions, nous avons effectué une expérience psycholinguistique : 20 personnes ont effectué une tâche de décision lexicale. Les temps de réaction et les taux d'erreur ont été analysés avec des modèles linéaires (généralisés) à effets mixtes en utilisant le logiciel R.

Premièrement, un effet de cognat inhibiteur marginalement significatif était observé dans l'analyse des taux d'erreur. Cela montre que le processus de reconnaissance en L3 est affecté quand un mot a un sens et une forme similaire dans une autre langue étrangère maîtrisée par les participants. Ce résultat (effet inhibiteur) obtenu avec une tâche de décision lexicale n'est pas conforme aux études antérieures. Il peut s'expliquer par une

stratégie de réponse spécifique. Une autre explication est que les cognats phonologiquement similaires et dissimilaires n'étaient pas examinés séparément dans cette partie d'analyse, ce qui est le cas dans la plupart des études antérieures du domaine.

C'est pourquoi, deuxièmement, nous avons étudié le degré de similarité phonologique des cognats, qui effectivement a influencé la reconnaissance des mots. Dans l'analyse des taux d'erreur, un effet inhibiteur marginalement significatif se manifestait quand les mots avaient des formes phonologiques plus distincts entre les deux langues. Cela est en accord avec les études antérieures. Cependant, les temps de réaction des cognats phonologiquement similaires présentaient un effet inhibiteur marginalement significatif, un résultat inverse.

Dans la discussion, nous avons proposé différentes interprétations pour nos résultats. Nous avons également abordé les limites éventuelles de notre étude, dont une est la fiabilité des temps de réaction. D'autres facteurs mentionnés sont les profils linguistiques variés des participants, l'influence du finnois et le nombre restreint des participants, qui sont tous les facteurs susceptibles à influencer les résultats.

Nous avons traité un sujet peu étudié : l'influence de la L2 sur la L3 dans la reconnaissance des mots cognats et le rôle de la similarité phonologique. Notre étude apporte des implications théoriques : nos résultats montrent qu'un effet de cognat est présent dans la L3 des apprenants, quand il s'agit des mots cognats entre la L2 et la L3. Cela indique qu'il existerait une différence représentationnelle entre les mots cognats et les mots non-cognats également quand il s'agit d'une L2 et d'une L3. De plus, cette étude montre que la phonologie d'une langue étrangère non-cible est activée lors la reconnaissance des mots dans la L3 et le degré de similarité phonologique entre les L2 et L3 des mots cognats joue un rôle dans le traitement des mots cognats. Ainsi, nos résultats signalent que dans la recherche future, il serait nécessaire de prendre en compte le degré de similarité phonologique les mots cognats. Il conviendrait de mettre en évidence l'importance de la phonologie également dans le développement des théories qui traiteront les mots cognats et l'accès non sélectif au lexique mental.

Nous voulons également exprimer les implications pédagogiques éventuelles, qui certes seront plus indirectes. Nous nous demandons s'il serait utile de sensibiliser les apprenants aux similitudes dans l'orthographe et la prononciation entre les différentes langues qu'ils sont en train d'apprendre et favoriser la comparaison des langues en classe des langues.

Finalement, nous considérons quelques perspectives futures. Pour avoir plus d'informations de l'interaction entre la phonologie et l'orthographe, il serait intéressant de varier également le degré de similarité orthographique. Les mots cognats trilingues finnois-anglais-français (comme *tunnel*, *tunneli*) pourraient également être un objet d'étude intéressant. En outre, les différents facteurs au niveau des participants pourraient faire l'objet d'étude. Surtout le niveau de compétences langagiers s'avère être un facteur qui joue un rôle important dans la reconnaissance des mots chez les apprenants. Afin de l'inclure dans une étude, il conviendrait de comparer des apprenants de différents niveaux et d'utiliser les méthodes plus objectifs et précis pour évaluer la compétence qu'une simple auto-évaluation. Pour compléter et/ou comparer les résultats de cette étude, d'autres méthodes comme la dénomination de mots et la décision lexicale auditive fourniraient de l'information supplémentaire. De plus, avec des tests de vocabulaire et de fréquence subjective, nous pourrions recueillir de l'information importante sur la nature des représentations chez les apprenants.

Cette étude a fourni de l'information importante quant à l'influence d'une langue étrangère (L2) sur une autre langue étrangère (L3). Elle souligne l'importance du degré de la similarité phonologique des mots cognats dans les études sur le lexique non sélectif. En fin de compte, des recherches ultérieures sont nécessaires pour étudier le rôle de la phonologie dans l'effet de cognat plus en détail.

Bibliographie

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *The Journal of the Acoustical Society of America*, 117(6), 3890-3901. <https://doi.org/10.1121/1.1906060>
- Brysbaert, M., Stevens, M., Mander, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, 42(3), 441–458. DOI: 10.1037/xhp0000159
- Carrasco-Ortiz, H., Amengual, M., & Gries, S. T. (2019). Cross-language effects of phonological and orthographic similarity in cognate word recognition. *Linguistic Approaches to Bilingualism*, 1-29. doi:10.1075/lab.18095.car
- Carrasco-Ortiz, H., Midgley, K. J., & Frenck-Mestre, C. (2012). Are phonological representations in bilinguals language specific? An ERP study on interlingual homophones. *Psychophysiology*, 49(4), 531-543. doi:10.1111/j.1469-8986.2011.01333.x
- Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. Dans S. Dornic (dir.), *Attention and performance VI* (p. 535-555). Hillsdale, NJ: Erlbaum.
- Comesaña, M., Ferré, P., Romero, J., Guasch, M., Soares, A. P., & García-Chico, T. (2015). Facilitative effect of cognate words vanishes when reducing the orthographic overlap: The role of stimuli list composition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 614-635. doi:10.1037/xlm0000065
- Comesaña, M., Sánchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., & Fraga, I. (2012). The interplay of phonology and orthography in visual cognate word recognition: An ERP study. *Neuroscience Letters*, 529(1), 75-79. doi:10.1016/j.neulet.2012.09.010

- Cook, S. V., Pandža, N. B., Lancaster, A. K., & Gor, K. (2016). Fuzzy nonnative phonolexical representations lead to fuzzy form-to-meaning mappings. *Frontiers in Psychology, 7*, 1-17. doi:10.3389/fpsyg.2016.01345
- Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2017). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition, 20*(4), 747-769. doi:10.1017/S1366728916000213
- Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic bulletin & review, 22*(5), 1216-1234. <https://doi.org/10.3758/s13423-015-0819-2>
- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283-1296. <https://doi.org/10.1037/0278-7393.26.5.1283>
- De Groot, A. M. B., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*(1), 90-123. [https://doi.org/10.1016/0749-596X\(91\)90012-9](https://doi.org/10.1016/0749-596X(91)90012-9)
- De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. New York, NY: Psychology Press.
- Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518. doi:10.1006/jmla.1999.2654
- 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. doi:10.1016/j.jml.2009.12.003
- Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A Tool for Word Selection and Non-Word Generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments & Computers, 36*(3), 488-499. <https://doi.org/10.3758/BF03195595>

- Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 663-679. doi:10.1037/0278-7393.33.4.663
- Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. Dans J. Cenoz, B. Hufeisen, & U. Jessner (dir.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (p. 21-41). Clevedon : Multilingual Matters.
- Harley, T. A. (2008). *The psychology of language: From data to theory*. Troisième édition. Hove: Psychology Press.
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S.N. & REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of biomedical informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics*, 42(2), 377-381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Hilton, H. (2003). L'accès au lexique mental dans une langue étrangère : Le cas des francophones apprenant l'anglais. *Corela*, 1(2). doi:10.4000/corela.676
- Hintikka, A. (2019). *La reconnaissance des mots cognats anglais-français dans deux modalités chez les apprenants finnophones*. (Mémoire de licence inédit). Université de Turku.
- Kim, J., Gabriel, U., & Gygax, P. (2019). Testing the effectiveness of the Internet-based instrument PsyToolkit: A comparison between web-based (PsyToolkit) and lab-based (E-Prime 3.0) measurements of response choice and response time in a complex psycholinguistic task. *PloS one*, 14(9). doi : <https://doi.org/10.1371/journal.pone.0221802>

- Kirsner, K., Lalor, E., & Hird, K. (1993). The bilingual lexicon: Exercise, meaning and morphology. Dans R. Schreuder, & B. Weltens (dir.), *The bilingual lexicon* (p. 215-248). Amsterdam: John Benjamins.
- Kroll, J. F., Gullifer, J. W., & Rossi, E. (2013). The multilingual lexicon: The cognitive and neural basis of lexical comprehension and production in two or more languages. *Annual Review of Applied Linguistics*, 33, 102-127. doi:10.1017/S0267190513000111
- Laxén, J., Lavour, J. M., & Aparicio, X. (2011). Reconnaissance en traduction et homographie interlangue. *Psychologie française*, 56(3), 161-172. <https://doi.org/10.1016/j.psfr.2011.07.001>
- Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19(5), 585-611. doi:10.1080/01690960444000007
- Lijewska, A. (2020). Cognate processing effects in bilingual lexical access. Dans R. Heredia, & A. Cieśliska (dir.), *Bilingual lexical ambiguity resolution* (p. 71-95). Cambridge: Cambridge University Press. doi:10.1017/9781316535967.005
- Lindqvist, C. (2016). Le développement de la taille du vocabulaire en français L2 chez les élèves suédophones. *Synergies Pays Scandinaves*, (11/12), 151-161.
- Manolescu, A., & Jarema, G. (2018). Influence du genre grammatical et du statut de cognat sur la production des bilingues de haut niveau. *SHS Web of Conferences*, 46(10002). doi : <https://doi.org/10.1051/shsconf/20184610002>
- Marian, V., Blumenfeld, H.K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research*, 50(4), 940-967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067)). Version finnoise traduite par Ari Huta and Mika Läheemäki, Université de Jyväskylä, Finlande. Version française adaptée par Marilyn Hall et traduite par Bhatara, Michaud, et Gervain (2011).
- Myers-Scotton, C. (2006). *Multiple voices : An introduction to bilingualism*. Malden (MA): Blackwell.

- Nakayama, M., Verdonschot, R. G., Sears, C. R., & Lupker, S. J. (2014). The masked cognate translation priming effect for different-script bilinguals is modulated by the phonological similarity of cognate words: Further support for the phonological account. *Journal of Cognitive Psychology*, 26(7), 714-724. doi:10.1080/20445911.2014.953167
- New, B. & Pallier, C. (s.d.). *Manuel de Lexique 3. Document version 3.2*. Disponible sur http://lexique.org/_documentation/Manuel_Lexique.3.2.pdf
- New, B., & Pallier, C. (2020). Lexique 3.83. Disponible sur <http://www.lexique.org>
- Otwinowska, A. (2015). *Cognate vocabulary in language acquisition and use : Attitudes, awareness, activation*. Blue Ridge Summit, PA: Multilingual Matters.
- Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, 68(4), 315-332. doi:10.1016/j.jml.2012.12.003
- Poort, E. D. (2018). *The representation of cognates and interlingual homographs in the bilingual lexicon*. (Thèse de doctorat). University College London. Disponible sur <https://discovery.ucl.ac.uk/id/eprint/10064100>
- Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory. Can cognate status be characterized as a special kind of morphological relation? Dans J. F. Kroll, & A. M. B. De Groot (dir.), *Handbook of bilingualism: Psycholinguistic approaches* (p. 226-250). Oxford: Oxford University Press.
- Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, Walter J. B. (2013). Cross-language distributions of high frequency and phonetically similar cognates. *PLoS One*, 8(5). doi:10.1371/journal.pone.0063006
- Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 344-354. <https://doi.org/10.1037/0278-7393.14.2.344>

- Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22(1), 106-129. doi:10.1080/01690960500463920
- Segalowitz, Norman. 2010. *Cognitive Bases of Second Language Fluency*, New York : Routledge.
- Spinelli, E., & Ferrand, L. (2005). *Psychologie du langage. L'écrit et le parlé, du signal à la signification*. Paris: Armand Colin.
- Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104. <https://doi.org/10.3758/BRM.42.4.1096>
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31. <https://doi.org/10.1177/0098628316677643>
- Szubko-Sitarek, W. (2011). Cognate facilitation effects in trilingual word recognition. *Studies in Second Language Learning and Teaching*, 1(2), 189-208.
- 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. <https://doi.org/10.1006/jmla.1998.2584>
- Van Leerdam, M., Bosman, A. M. T., & De Groot, A. M. B. (2009). When MOOD rhymes with ROAD: Dynamics of phonological coding in bilingual visual word perception. *The Mental Lexicon*, 4(3), 303-335. doi:10.1075/ml.4.3.01van
- Veivo, O. (2017). *Orthographe et reconnaissance des mots parlés chez les apprenants tardifs de L2*. (Thèse de doctorat.) Turku: Université de Turku. Disponible sur <http://urn.fi/URN:ISBN:978-951-29-6869-5>
- Voga, M. (2020). Traitement bilingue inter-alphabet et morphologie : Le rôle des variables paradigmatiques dans l'amorçage masqué grec-français. *SHS Web of Conferences*, 78(08001). doi : <https://doi.org/10.1051/shsconf/20207808001>

- Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied linguistics*, 19(3), 295-333. <https://doi.org/10.1093/applin/19.3.295>
- Zhu, Y., & Mok, P. P. K. (2020). Visual recognition of cognates and interlingual homographs in two non-native languages: Evidence from Asian adult trilinguals. *Linguistic Approaches to Bilingualism*, 10(4), 441-470. <https://doi.org/10.1075/lab.17049.zhu>

Annexes

Annexe 1. Les mots non-cognats

Mot	Fréquence	Partie du discours	Longueur	Nombre de syllabes	Nombre de voisins orthographiques
drap	19,220	NOM	4,00	1	3
chou	29,860	NOM	4,00	1	2
chien	223,530	NOM	5,00	1	3
sœur	184,990	NOM	5,00	1	3
souhait	10,410	NOM	7,00	1	2
chiffre	27,280	NOM,VER	7,00	1	1
chèvre	14,080	NOM	6,00	1	0
douane	6,120	NOM	6,00	1	0
maigre	12,730	ADJ (NOM)	6,00	1	1
meuble	19,480	NOM,ADJ,VER	6,00	1	3
caisse	38,030	NOM	6,00	1	6
douche	36,060	NOM,VER	6,00	1	8
crêpe	8,440	NOM,VER	5,00	1	2
ongle	20,370	NOM	5,00	1	2
œuf	39,400	NOM	4,00	1	2
goût	57,940	NOM	4,00	1	3
savon	16,680	NOM	5,00	2	3
orange	17,140	NOM	5,00	2	2
valise	50,990	NOM	6,00	2	3
poulet	41,250	NOM	6,00	2	5
équipe	129,970	NOM,VER	6,00	2	3
épaule	32,410	NOM,VER	6,00	2	2
visage	141,230	NOM	6,00	2	4
cadeau	125,790	NOM	6,00	2	3
époque	68,440	NOM	6,00	2	2
oiseau	77,730	NOM	6,00	2	2
morceau	64,780	NOM	7,00	2	2
oreille	73,540	NOM	7,00	2	2
fuite	29,090	NOM	5,00	1	6
sucré	32,010	NOM (VER)	5,00	1	5
honte	103,400	NOM	5,00	1	7
libre	134,420	ADJ	5,00	1	5
baie	7,090	ADJ,NOM,VER	4,00	1	14
sain	26,630	ADJ	4,00	1	14
verre	176,570	NOM	5,00	1	10
reine	59,050	NOM	5,00	1	7
foie	23,480	NOM	4,00	1	9

nain	13,900	NOM, ADJ	4,00	1	10
trou	90,720	NOM	4,00	1	5
faim	128,040	NOM	4,00	1	4
poids	34,420	NOM	5,00	1	2
neige	39,340	NOM (VER)	5,00	1	2
boisson	18,550	NOM	7,00	2	3
hauteur	19,990	NOM	7,00	2	3
lumière	134,830	NOM	7,00	2	1
poisson	81,510	NOM	7,00	2	3
niveau	50,700	NOM	6,00	2	0
espoir	64,220	NOM	6,00	2	0
couteau	58,150	NOM	7,00	2	0
sommeil	44,530	NOM	7,00	2	0
colline	26,390	NOM	7,00	2	1
chapeau	54,910	NOM	7,00	2	1
épreuve	23,430	NOM	7,00	2	1
soutien	18,870	NOM	7,00	2	1
couloir	29,170	NOM	7,00	2	2
baleine	16,680	NOM	7,00	2	2

Annexe 2. Les mots cognats P-

Mot	Forme phonol. français	Forme phonol. anglais	Distance orthog.	Distance phonol.	Fréq.	Partie du discours	Longueur	Nb. de syllabes	Nb. de voisins orthog.
lion	ljo	l2@n	1,000	0,565	20,860	NOM	4,00	1	3
point	pwe	p4nt	1,000	0,550	225,320	NOM	5,00	1	3
science	sjAs	s2@ns	1,000	0,640	32,970	NOM	7,00	1	0
humble	^bl	hVmbP	1,000	0,700	12,430	ADJ	6,00	1	0
angle	Agl	{NgP	1,000	0,794	13,290	NOM	5,00	1	3
jungle	Z^gl	_VNpP	1,000	0,760	16,470	NOM	6,00	1	2
centre	sAtr	sEnt@R	1,000	0,770	57,000	NOM	6,00	1	8
fruit	fr-i	frut	1,000	0,700	39,450	NOM	4,00	1	2
royal	rwAjAl	r4@l	1,000	0,731	22,650	ADJ	5,00	2	2
danger	dAZe	d1n_@R	1,000	0,669	80,610	NOM	6,00	2	4
urgent	yrZA	3_@nt	1,000	0,515	31,560	ADJ	6,00	2	1
client	klijA	kl2@nt	1,000	0,647	112,120	NOM	6,00	2	3
cousin	kuze	kVzH	1,000	0,625	73,870	NOM	6,00	2	2
patient	pAsjA	p1SHt	1,000	0,544	61,690	NOM	7,00	2	3

Annexe 3. Les mots cognats P+

Mot	Forme phonol. français	Forme phonol. anglais	Distance orthog.	Distance phonol.	Fréq.	Partie du discours	Longueur	Nb. de syllab es	Nb. de voisins orthog.
sauce	sos	s\$S	1,000	0,933	19,240	NOM	5,00	1	5
cause	koz	k\$Z	1,000	0,933	218,600	NOM	5,00	1	6
vote	vOt	v5t	1,000	0,892	15,040	NOM	4,00	1	14
force	fOrs	f\$S	1,000	0,850	151,960	NOM	5,00	1	8
dose	doz	d5s	1,000	0,900	18,440	NOM	4,00	1	10
zone	zon	z5n	1,000	0,900	53,940	NOM	4,00	1	4
sport	spOr	sp\$t	1,000	0,925	30,040	NOM	5,00	1	2
dispute	dispyt	dIspjut	1,000	0,871	16,540	NOM	7,00	2	3
bizarre	bizAr	bIz#R	1,000	0,920	131,410	ADJ	7,00	2	1
effort	efOr	Ef@t	1,000	0,840	42,700	NOM	6,00	2	0
miracle	mirAkl	mIr@kP	1,000	0,897	56,890	NOM	7,00	2	0
fortune	fOrtyn	f\$Jun	1,000	0,850	33,620	NOM	7,00	2	1
divorce	divOrs	dIv\$S	1,000	0,867	23,230	NOM	7,00	2	1
missile	misil	mIs2l	1,000	0,900	16,700	NOM	7,00	2	2

Annexe 4. Le questionnaire en finnois et en français

Page 1

Taustakysely / Questionnaire

Vastaa allaoleviin kysymyksiin.

Veillez répondre aux questions ci-dessous.

Etu- ja sukunimi
Prénom et nom

Ikä
Age

(Merkitse vuosina / En années)

Sukupuoli
Sexe

- Mies / Homme
 Nainen / Femme
 Muu/En halua kertoa / Autre/ Je ne veux pas dire

Äidinkieli/Äidinkielet
La/les langue(s) maternelle(s)

Minkä ikäisenä olet aloittanut ranskan opiskelun?
A quel âge avez-vous commencé à étudier le français?

(Merkitse vuosina / En années)

Mitä kaikkia muita kieliä puhut ja minkä ikäisenä olet aloittanut näiden muiden kielten opiskelun?
- Merkitse esim.
englanti, 8-vuotiaana
ruotsi, 13-vuotiaana

Quels sont les autres langues que vous parlez? À quel âge avez-vous commencé à les étudier?
- Répondez par exemple ainsi :
anglais à l'âge de 8 ans,
suédois à l'âge de 13 ans

(Merkitse jokainen opiskelemasi vieras kieli (muu kuin ranska) ja minkä ikäisenä aloit opiskelemaan kieltä / Pour chaque langue étrangère (autre que le français) que vous avez étudiée, indiquez l'âge auquel vous avez commencé à étudier la langue.)

Luettele paremmuusjärjestyksessä kaikki kielet joita osaat, myös äidinkieli/äidinkielet. Aloita kielestä jota osaat parhaiten.

Listez toutes les langues que vous parlez, y compris la langue maternelle, par ordre de compétence. Commencez par la langue que vous maîtrisez le mieux.

Oletko oleskellut jossakin ranskankielisessä maassa?

Jos olet oleskellut useamman eri kerran, valitse vaihtoehto, joka kuvaa oleskelujen kestoa yhteensä.

Avez-vous séjourné dans pays francophone?

Si vous avez séjourné à plusieurs reprises, choisissez l'option qui reflète le temps séjourné total.

- En koskaan / Jamais
 Kyllä, alle kuukauden / Oui, moins d'un mois
 Kyllä, 1-3 kuukautta / Oui, entre 1 et 3 mois
 Kyllä, 4-6 kuukautta / Oui, entre 4 et 6 mois
 Kyllä, 7-12 kuukautta / Oui, entre 7 et 12 mois
 Kyllä, yli vuoden / Oui, plus d'un an

Oletko oikea- vai vasenkätinen?

Êtes-vous droitier ou gaucher?

- Oikeakätinen / Droitier
 Vasenkätinen / Gaucher

Arvioi asteikolla nollasta kymmeneen, kuinka hyvin ymmärrät puhetta ranskaksi.

Évaluez sur une échelle de 0 à 10 vos compétences en compréhension de la langue orale en français.

- 0 - Ei osaamista / Aucun
 1 - Erittäin vähäinen / Très faible
 2 - Vähäinen / Faible
 3 - Välttävä / Correct
 4 - Vähän huonompi kuin kohtuullinen / Un peu moins qu'adéquat
 5 - Kohtuullinen / Adéquat
 6 - Vähän parempi kuin kohtuullinen / Un peu plus qu'adéquat
 7 - Hyvä / Bon
 8 - Erittäin hyvä / Très bon
 9 - Erinomainen / Excellent
 10 - Täydellinen / Parfait

Arvioi osaamistasi asteikolla nollasta kymmeneen: lukeminen ranskaksi.

Évaluez sur une échelle de 0 à 10 votre niveau de compétence concernant la lecture en français.

- 0 - Ei osaamista / Aucun
 1 - Erittäin vähäinen / Très faible
 2 - Vähäinen / Faible
 3 - Välttävä / Correct
 4 - Vähän huonompi kuin kohtuullinen / Un peu moins qu'adéquat
 5 - Kohtuullinen / Adéquat
 6 - Vähän parempi kuin kohtuullinen / Un peu plus qu'adéquat
 7 - Hyvä / Bon
 8 - Erittäin hyvä / Très bon
 9 - Erinomainen / Excellent
 10 - Täydellinen / Parfait

Näetkö tietokoneen ruudulle ilman silmälaseja?

Pouvez vous voir l'écran de votre ordinateur sans lunettes?

- Kyllä / Oui
 En / Non

Onko sinulla todettu jotakin kielen tai puheen häiriötä tai oppimisvaikeutta (esim. lukivaikeus)?

A-t-on diagnostiqué chez vous des troubles du langage ou de la parole ou des difficultés d'apprentissage (par exemple, la dyslexie) ?

- Kyllä / Oui
 Ei / Non

Arvioi asteikolla nollasta kymmeneen, kuinka hyvin osaat puhua ranskaa.

Évaluez sur une échelle de 0 à 10 vos compétences orales en français.

- 0 - Ei osaamista / Aucun
 1 - Erittäin vähäinen / Très faible
 2 - Vähäinen / Faible
 3 - Välttävä / Correct
 4 - Vähän huonompi kuin kohtuullinen / Un peu moins qu'adéquat
 5 - Kohtuullinen / Adéquat
 6 - Vähän parempi kuin kohtuullinen / Un peu plus qu'adéquat
 7 - Hyvä / Bon
 8 - Erittäin hyvä / Très bon
 9 - Erinomainen / Excellent
 10 - Täydellinen / Parfait

Arvioi asteikolla nollasta kymmeneen, kuinka hyvin osaat puhua englantia.

Évaluez sur une échelle de 0 à 10 vos compétences orales en anglais.

- 0 - Ei osaamista / Aucun
 1 - Erittäin vähäinen / Très faible
 2 - Vähäinen / Faible
 3 - Välttävä / Correct
 4 - Vähän huonompi kuin kohtuullinen / Un peu moins qu'adéquat
 5 - Kohtuullinen / Adéquat
 6 - Vähän parempi kuin kohtuullinen / Un peu plus qu'adéquat
 7 - Hyvä / Bon
 8 - Erittäin hyvä / Très bon
 9 - Erinomainen / Excellent
 10 - Täydellinen / Parfait

Arvioi asteikolla nollasta kymmeneen, kuinka hyvin ymmärrät puhetta englanniksi

Évaluez sur une échelle de 0 à 10 vos compétences en compréhension de la langue orale en anglais.

- 0 - Ei osaamista / Aucun
- 1 - Erittäin vähäinen / Très faible
- 2 - Vähäinen / Faible
- 3 - Välttävä / Correct
- 4 - Vähän huonompi kuin kohtuullinen / Un peu moins qu'adéquat
- 5 - Kohtuullinen / Adéquat
- 6 - Vähän parempi kuin kohtuullinen / Un peu plus qu'adéquat
- 7 - Hyvä / Bon
- 8 - Erittäin hyvä / Très bon
- 9 - Erinomainen / Excellent
- 10 - Täydellinen / Parfait

Arvioi osaamistasi asteikolla nollasta kymmeneen: lukeminen englanniksi.

Évaluez sur une échelle de 0 à 10 votre niveau de compétence concernant la lecture en anglais.

- 0 - Ei osaamista / Aucun
- 1 - Erittäin vähäinen / Très faible
- 2 - Vähäinen / Faible
- 3 - Välttävä / Correct
- 4 - Vähän huonompi kuin kohtuullinen / Un peu moins qu'adéquat
- 5 - Kohtuullinen / Adéquat
- 6 - Vähän parempi kuin kohtuullinen / Un peu plus qu'adéquat
- 7 - Hyvä / Bon
- 8 - Erittäin hyvä / Très bon
- 9 - Erinomainen / Excellent
- 10 - Täydellinen / Parfait

Sähköpostiosoitteesi

- Syötä tähän sähköpostiosoitteesi, jos olet halukas osallistumaan mahdolliseen lisätutkimukseen. Sähköpostin ilmoittaminen ei sido sinua osallistumaan lisätutkimukseen. Sähköpostiosoitettasi käytetään vain tutkimuskutsun lähettämiseen.

Adresse e-mail

- Fournissez votre adresse e-mail, si vous souhaitez participer à une autre expérience. Le fait de fournir un adresse e-mail ne vous oblige pas à participer à une autre expérience. Votre adresse électronique ne sera utilisée que pour vous envoyer une invitation à l'expérience.

Annexe 5. Résumé en finnois – Suomenkielinen tiivistelmä

1. Johdanto

Tässä pro gradu -tutkielmassa tarkastellaan vieraan kielen sanantunnistusta. Kiinnostuksen kohteena ovat kognaatit eli sanat, joiden kirjoitus- ja äänneasu sekä merkitys ovat samankaltaiset kahdessa tai useammassa kielessä. Esimerkki tällaisesta sanasta on *science*, joka merkitsee tiedettä sekä ranskaksi että englanniksi. Tutkimuksella on kaksi tavoitetta. Ensiksi tavoitteena on selvittää, onko kognaattien ja ei-kognaattien prosessoinnissa eroa. Toiseksi tarkastellaan, vaikuttaako kognaattien fonologisen samankaltaisuuden aste sanantunnistukseen.

Tutkimuksen kohteena ovat ranskan kielen oppijat, jotka osaavat englantia ja joiden äidinkieli on suomi. Tässä tutkimuksessa käytetään eri kielistä lyhenteitä L1, L2 ja L3. Lyhenne L1 viittaa lähes aina ensikieleen, mutta lyhennettä L2 käytetään monella tavalla. Tässä tutkielmassa käytetään lyhennettä L2 kuvaamaan osallistujien aiemmin oppimaa kieltä (englanti) ja lyhennettä L3 kuvaamaan kieltä, jota he ovat oppimassa (ranska), jotta kyseiset vieraat kielet voidaan erottaa toisistaan. Hammarbergia (2001: 22) mukailleen, L3 ei siis välttämättä ole oppimisjärjestyksessä kolmas kieli.

Sanantunnistukseen liittyvien prosessien tutkiminen on tärkeää, sillä sanasto on hyvin tärkeä osa-alue vieraan kielen oppimisessa (Lindqvist, 2016: 152). Tämän lisäksi on perusteltua tutkia monikielisiä oppijoita, sillä heitä on tutkittu vähemmän verrattuna kaksikielisiin (Kroll et al., 2013: 103). Kognaattien prosessointia on tutkittu melko paljon, mutta tähän mennessä vain harvat tutkimukset ovat ottaneet huomioon fonologian roolin. Kognaattitutkimuksessa on keskitytty enimmäkseen L1:n ja L2:n välisiin kognaatteihin. Sen sijaan kahden vieraan kielen (L2 ja L3) väliset kognaatit ovat jääneet vähemmälle huomiolle. Tässä tutkimuksessa tutkitaan monikielisten oppijoiden L2:n ja L3:n välisten kognaattien prosessointia ottaen huomioon kognaattien fonologisen samankaltaisuuden aste.

Tutkimuskysymykset ovat:

- 1) Eroaako ranskan ja englannin välisten kognaattien tunnistus ei-kognaattien tunnistuksesta suomenkielisillä oppijoilla? Jos eroaa, millainen tämä ero on?

2) Onko fonologisen samankaltaisuuden asteella vaikutusta ranskan ja englannin välisten kognaattien tunnistukseen? Jos on, millainen tämä vaikutus on?

Näihin kysymyksiin vastaamiseksi 23 koehenkilöä osallistui psykolingvistiseen kokeeseen, jossa suoritettiin leksikaalisen päätöksenteon tehtävä. Kokeessa mitattiin reaktioaikaa ja virheiden määrää. Aineisto analysoitiin tilastollisilla menetelmillä.

2. Teoriatausta

Sana koostuu kirjaimista (sanan kirjoitettu muoto) tai foneemeista (sanan puhuttu muoto), jotka yhdistetään merkitykseen. Oppiakseen sanastoa oppijan täytyy oppia yhdistämään sanan muoto ja merkitys (Spinelli & Ferrand, 2005: 21, 23-24). Aikuisella lukutaitoisella henkilöllä on pitkäkestoiseen muistiin varastoitua tietoa sanan ortografisesta ja fonologisesta muodosta, semantiikasta, morfologiasta sekä syntaksista (Spinelli & Ferrand, 2005: 22; de Groot, 2011: 5). Kun viitataan näihin tietoihin, puhutaan mentaalileksikosta (Treisman 1960, viitannut Spinelli & Ferrand, 2005: 21-22). Sanantunnistus kuvaa prosessia, joka tapahtuu sanan havaitsemishetkestä aina siihen pisteeseen asti, että kaikki sanaan liittyvä tieto on aktivoitunut mentaalileksikossa (De Groot, 2011: 155-156). On tärkeää ymmärtää, että sanantunnistusprosessiin vaikuttaa monia tekijöitä, kuten sanan yleisyys eli frekvenssi (Spinelli & Ferrand, 2005: 134), konteksti (esim. Duyck et al., 2007) sekä sanan ortografinen naapurusto (esim. van Heuven et al., 1998). Useat tutkimukset ovat osoittaneet, että monikielisten henkilöiden sanantunnistusprosessi on luonteeltaan sellainen, että henkilön kaikki kielet ovat aktivoituneita, kun kieltä tunnistetaan tai tuotetaan (esim. Schwartz et al., 2007; Lemhöfer et al., 2004).

Oppijoiden sanastonoppimisessa saattaa esiintyä vaikeuksia esimerkiksi sen takia, että oppijoiden fonoleksikaalinen representaatio eli yksittäiseen kokonaiseen sanaan liittyvä fonologinen muistiedustuma voi olla epätarkka (Cook et al., 2016). L3:n oppiminen eroaa L2:n oppimisesta siten, että kieltenvälisen vaikutuksen lähteenä on oppijalla usein pikemminkin L2 kuin L1 (Hammarberg, 2011: 37).

Kognaatti voidaan kirjoittaa hieman eri tavalla eri kielissä, jolloin kyse on epäidenttisistä kognaateista. Kun kirjoitusasu on täysin sama eri kielten välillä, puhutaan identtisistä kognaateista. Lisäksi kognaattien fonologisen muodon ja merkityksen samankaltaisuuksien aste vaihtelee. Kognaatit tunnistetaan usein helpommin kuin sanat,

jotka eivät ole kognaatteja. Niiden kohdalla tehdään siis vähemmän virheitä ja niiden tunnistus on yleensä nopeampaa (esim. Lemhöfer et al., 2004). Tämä niin sanottu kognaattiefekti on yleensä hyvin erottuva efekti, mutta se on kuitenkin altis erilaisille tekijöille. Esimerkiksi identtiset kognaatit tunnistetaan usein epäidenttisiä paremmin (katso esim. Dijkstra et al., 2010). Myös tutkimuksissa käytettyjen tehtävien tyyppi voi vaikuttaa efektiin (Dijkstra et al., 2010). Tämän pro gradu -tutkielman näkökulmasta on tärkeää huomioida, että kognaattien aiheuttama efekti prosessoinnissa näkyy myös L2:n ja L3:n välillä (esim. Lemhöfer et al., 2004; Zhu & Mok, 2020).

Kognaattien nopeampaa prosessointia selitetään eroilla sanojen representaatioissa. Kognaattien ajatellaan siis representoituneen eri tavalla mentaalileksikkoon kuin sanojen, joilla ei ole tällaista samankaltaista muotoa sekä merkitystä eri kielten välillä. Kognaattien representaatioista on esitetty useita eri teorioita. Niiden representaatiot saattavat erota muista sanoista esimerkiksi semanttisella tasolla (esim. De Groot & Nas, 1991) tai morfologisella tasolla (esim. Peeters et al., 2013).

Kun monikielinen henkilö lukee tekstiä jollakin kielellä, myös muiden kielten kuin lukemisen kohteena olevan kielen fonologia aktivoituu (Van Leerdam et al., 2009). Aiemmissä tutkimuksissa on osoitettu, että kognaattien fonologisen samankaltaisuuden aste eri kielten välillä vaikuttaa sanantunnistukseen (esim. Carrasco-Ortiz et al., 2019). Tutkimusten mukaan keskenään fonologisesti samankaltaiset kognaatit tunnistetaan nopeammin kuin keskenään fonologisesti eriävät (esim. Schwartz et al., 2007). Toisaalta samankaltaisten fonologisten muotojen aiheuttama efekti voi olla myös inhibitorinen eli hidastaa tunnistusta (Comesaña et al., 2015). On huomattava, että vain muutama tutkimus on systemaattisesti ottanut huomioon sekä ortografisen että fonologisen samankaltaisuuden asteen, kun on kyse kognaateista.

Teoriapohjaan tukeutuen tälle tutkimukselle muodostettiin kaksi hypoteesia:

- 1) Kognaattien sanantunnistus eroaa muiden sanojen tunnistamisesta. Kognaatit tunnistetaan muita sanoja nopeammin ja virheettömämmin.
- 2) Fonologisesti keskenään samankaltaiset kognaatit tunnistetaan fonologisesti keskenään erilaisia kognaatteja nopeammin ja virheettömämmin.

3. Aineistonkeruumenetelmä

Koehenkilöiden suorittamassa kokeessa oli yhteensä 168 sanaa, joista puolet oli oikeita sanoja ja puolet epäsanoina. Oikeista sanoista 24 oli kognaatteja. Kognaatit valittiin Schepensin et al. (2013) lisämateriaaleista siten, että ne olivat kirjoitusasultaan identtisiä, mutta niiden ranskan ja englannin välinen fonologinen etäisyys vaihteli. Kognaatit olivat joko fonologisesti samankaltaisia keskenään (P+) tai fonologisesti erilaisempia keskenään (P-). Mukaan ei otettu sanoja, jotka ovat kognaatteja myös suomessa (esimerkiksi sana *tunnel*, *tunneli*). Mukaan ei myöskään otettu sanoja, jotka kirjoitetaan samalla tavalla kuin suomessa, mutta joiden merkitys eroaa suomenkielisen sanan merkityksestä. Sanat, jotka eivät olleet kognaatteja, valittiin lexique.org -tietokannasta (New&Pallier, 2020). Huomioon otettiin kaikkien oikeiden sanojen frekvenssi, pituus, tavujen määrä sekä ortografinen naapurusto. Epäsanat luotiin WordGen-epäsanageneraattorilla (Duyck et al. 2004). Epäsanat noudattivat ranskan fonotaktisia sääntöjä (De Groot, 2011: 157), ja niiden pituutta, tavumäärää ja ortografista naapurustoa kontrolloitiin.

Itse koe ohjelmoitiin ja suoritettiin PsyToolkit-nimisellä työkalulla verkossa (Stoet, 2010 & 2017). Taustatietokysely luotiin REDCap-työkalulla (Harris et al. 2009 & 2019). Kokeessa suoritettiin leksikaalisen päätöksenteon tehtävä, koska sitä käytetään laajasti alan tutkimuksissa (esim. Hilton, 2003: 3).

Kokeeseen osallistui 23 henkilöä. Koehenkilöt olivat ranskan oppiaineen yliopistoprosentteja, jotka ilmoittivat puhuvansa suomea äidinkielen tasoisesti. Koehenkilöillä ei ollut luki- tai oppimisvaikeuksia.

Ennen kokeeseen osallistumista koehenkilöt lukivat tietosuojailmoituksen sekä antoivat suostumuksensa tutkimukseen osallistumisesta. Koehenkilöt saivat kirjalliset ohjeet kokeesta ja heitä ohjeistettiin käyttämään silmälasia tai piilolinssejä tarvittaessa kokeen ajan. Kokeessa ohjeistettiin painamaan näppäimistön painiketta "a" vasemmalla etusormella ja näppäimistön painiketta "l" oikealla etusormella. Oikeakätisille osallistujille "a" merkitsi epäsanaa ja "l" oikeaa sanaa. Vasenkätisille ohjeistus oli päinvastainen: "a" merkitsi oikeaa sanaa ja "l" epäsanaa.

Ennen varsinaista koetta koehenkilöt näkivät kymmenen harjoituskierrosta. Itse koe oli jaettu kolmeen osioon. Jokaisessa osiossa oli 56 sanaa. Osioden välillä oli minuutin mittainen tauko. Yksi kierros oli seuraavanlainen. Ensin näytöllä näkyi fiksaatiopiste 500 millisekunnin ajan. Tämän jälkeen oli 500 ms viive, jonka aikana näkyi musta näyttö.

Tämän jälkeen sana ilmestyi ruudulle, ja pysyi siinä, kunnes vastaaja vastasi tai 5000 ms oli kulunut. Vastauksen tai 5000 ms jälkeen oli jälleen 500 ms viive (tyhjä näyttö). Sen jälkeen alkoi uusi kierros.

Sanojen järjestys oli satunnaistettu, kuitenkin niin, että yli neljä sanaa tai epäsanaa ei voinut esiintyä peräjälkeen. Osallistujat jaettiin kolmeen ryhmään, joissa sanojen järjestys oli eri. Kokeessa kerättiin reaktioaika sekä tieto siitä, vastasiko koehenkilö oikein vai väärin. Leksikaalisen päätöksenteon tehtävän jälkeen koehenkilöt vastasivat taustatietokyselyyn.

4. Tulokset

Kolme koehenkilöä jätettiin pois analyysistä, koska he tekivät yli 15 % virheitä (vrt. Lemhöfer et al. 2004: 595). Täten 20 koehenkilön reaktioajat ja virheiden osuus vastauksista analysoitiin. Epäsanoja ei huomioitu analyysissä.

Analyysiin käytettiin R-tilasto-ohjelmaa. Analyysit tehtiin käyttäen lineaarista sekamallia sekä yleistettyä lineaarista sekamallia (lme4-paketti, Bates et al., 2015). Niiden avulla saadaan otettua huomioon kiinnostuksen kohteena olevien vaikutusten eli kiinteiden vaikutusten lisäksi satunnaisvaikutukset, jotka liittyvät esimerkiksi luonnollisiin eroihin eri koehenkilöiden välillä. Reaktioaika-analyysiin käytettiin lineaarista sekamallia (lmer-paketin funktio lme4), koska muuttuja on jatkuva. Virheanalyysiin käytettiin yleistettyä lineaarista sekamallia (lmer-paketin funktio glmer), koska muuttuja on binaarinen. Käytetyt estimointimenetelmät ovat mallien oletusmenetelmiä.

Ensin tarkasteltiin kognaattien ja muiden sanojen välisiä prosessointieroja (tutkimuskysymys 1).

Virheanalyysiä varten aineistosta poistettiin ne havainnot, joissa ei ollut kyllä/ei -vastausta, eli 5000 ms:n aikaraja oli tullut täyteen. Näitä vastauksia oli kuusi kappaletta. Täten analysoitiin 1674 havaintoa yleistetyllä lineaarisella sekamallilla. Riippuva muuttuja oli vastausten tarkkuus (kuinka paljon virheitä osallistujat tekivät). Kiinteä tekijä mallissa oli kognaattius, satunnaiset tekijät koehenkilö ja sana.

Kognaattien virheprosentti oli 7,5 % ja ei-kognaattien 5,1 %. Analyysin tulokset osoittivat, että vastaajat tekivät tilastollisesti marginaalisesti merkitsevästi enemmän virheitä, kun kyse on kognateista ($z=1.848$, $p = 0.0646$, merkitsevyystaso 0.05).

Reaktioaika-analyysiä varten aineistosta poistettiin 5 sanaa eli 100 havaintoa, joiden kohdalla yli 30 % vastauksista oli väärin (baleine, drap, humble, missile, nain). Lisäksi poistettiin kuusi havaintoa, joiden vastausaika oli yli 5000 ms aikarajan. Tämän jälkeen poistettiin vielä 79 havaintoa, joiden reaktioajat olivat yli kahden keskihajonnan päästä osallistujakohtaisesta keskiarvosta. Vain ne havainnot, joissa vastaus oli oikein, huomioitiin analyysissä. Loppujen lopuksi analysoitiin siis 1452 havaintoa.

Ennen tilastollisen mallin sovittamista reaktioajan jakaumaa tarkasteltiin. Visuaalisen tarkastelun sekä vinousluvun mukaan jakauma oli vahvasti oikealle eli positiivisesti vino. Myös normalisuustestien mukaan jakauma ei noudattanut normaalijakaumaa. Siksi käytettiin muunnosta. Logaritmi-, neliöjuuri- ja käänteistä muunnosta kokeiltiin ja näistä käänteinen muunnos antoi parhaimman tuloksen.

Reaktioajan analyysiin käytettiin lineaarista sekamallia. Riippuva muuttuja oli reaktioaika. Kiinteä tekijä mallissa oli kognaattius, satunnaiset tekijät osallistuja ja sana. Kognaattien keskimääräinen reaktioaika oli 795 millisekuntia ja ei-kognaattien 764 millisekuntia. Kognaatit tunnistettiin ei-kognaatteja hitaammin, mutta ero ei ollut tilastollisesti merkitsevä ($t=0.603$).

Seuraavaksi tarkasteltiin, vaikuttaako kognaattien fonologisen samankaltaisuuden aste sanantunnistukseen (tutkimuskysymys 2).

Tässä virheanalyysissä käytettiin samaa dataa kuin ensimmäisessä virheanalyysissä. Analyysiin käytettiin yleistettyä lineaarista sekamallia. Riippuva muuttuja oli vastausten tarkkuus (oikein/väärin). Kiinteä tekijä mallissa oli sanan kognaattius fonologisen samankaltaisuuden aste huomioon ottaen. Satunnaiset tekijät olivat jälleen koehenkilö ja sana.

Koehenkilöt tekivät enemmän virheitä fonologisesti keskenään erilaisten kognaattien kohdalla ($z = 1.893$) kuin niiden sanojen kohdalla, jotka eivät olleet kognaatteja. Tämä tulos osoittautui marginaalisesti merkitseväksi tasolla 0.05 ($p= 0.0583$). Myös fonologisesti keskenään erilaisten kognaattien kohdalla tehtiin enemmän virheitä kuin niiden sanojen kohdalla, jotka eivät olleet kognaatteja, mutta tämä ero ei ollut tilastollisesti merkitsevä ($p= 0.3029$).

Reaktioaika-analyysi tehtiin samalla datalla kuin edeltävä reaktioaika-analyysi. Jälleen käytettiin lineaarista sekamallia, ja kiinteä tekijä oli sanan kognaattius huomioiden

fonologisen samankaltaisuuden aste. Satunnaiset tekijät olivat koehenkilö ja sana. Riippuva muuttuja oli reaktioaika.

Fonologisesti samankaltaisten kognaattien reaktioaikojen keskiarvo oli 835 millisekuntia, fonologisesti erilaisten kognaattien 754 ms ja muiden kuin kognaattien 764 ms. Fonologisesti keskenään erilaiset kognaatit tunnistettiin nopeammin ($t= 0.795$), kuin ne sanat, jotka eivät olleet kognaatteja, mutta tämä ero ei ollut tilastollisesti merkitsevä. Sen sijaan fonologisesti keskenään samanlaiset kognaatit tunnistettiin tilastollisesti marginaalisesti merkitsevästi hitaammin ($t=-1.744$) verrattuna sanoihin, jotka eivät olleet kognaatteja.

5. Tulosten pohdinta ja päätelmät

Tuloksista voidaan päätellä, että oppijoiden L2 aktivoitui sanantunnistuksen aikana, sillä kognaattien prosessointi oli erilaista kuin niiden sanojen, jotka eivät olleet kognaatteja. Voidaan siis todeta, että L2:n ja L3:n väliset kognaatit vaikuttavat L3:n prosessointiin. Kun otetaan myös fonologia huomioon, voidaan todeta, että L2:n fonologia aktivoitui, ja vaikutti sanantunnistukseen. Tämä osoittaa, että L2:n fonologia aktivoituu, kun luetaan L3:lla, ja että kognaattien fonologisen samankaltaisuuden aste vaikuttaa niiden tunnistukseen.

Kun tarkasteltiin, millaisia eroja kognaattien ja ei-kognaattien tunnistamisessa oli, virheanalyysin mukaan kognaatit tuottivat inhibitorisen efektin: niiden kohdalla tehtiin enemmän virheitä. Tämä efekti oli päinvastainen kuin useissa aiemmissa tutkimuksissa, joissa kognaattien tunnistus osoittautui helpommaksi (esim. Lemhöfer et al., 2004). Tämä voi selittyä sillä, että vastaajat muuttivat vastausstrategiaansa, koska he näkivätkin kokeessa sanoja, jotka ovat englantia (seuraten Zhun ja Mokin, 2020, päätelmiä). Toisaalta on huomioitava, että tässä analyysissä fonologisen samankaltaisuuden vaikutusta ei otettu huomioon. Se on saattanut vaikuttaa tuloksiin (vrt. Comesaña et al., 2015: 623), joten niitä kannattaa tulkita varoen.

Seuraavaksi tarkasteltiin, miten fonologisen samankaltaisuuden aste vaikutti kognaattien tunnistukseen. Virheanalyysin tulokset vahvistivat osittain aiempien tutkimusten havaintoja, sillä kahdessa aiemmassa tutkimuksessa fonologisesti keskenään erilaiset kognaatit aiheuttivat inhibitorisen vaikutuksen, eli tunnistus hidastui verrattuna fonologisesti samankaltaisiin kognaatteihin (Dijkstra et al., 2010; Schwartz et al., 2007).

Tämän voidaan ajatella tukevan ajatusta siitä, että kognaatin ortografian aktivoituminen leviää edelleen myös sen fonologisiin representaatioihin (Schwartz et al., 2007). Jos nämä fonologiset representaatiot ovat keskenään erilaisia, aiheutuu inhibitorinen efekti, joka hidastaa tunnistusta. Reaktioaika-analyysissä saatiin päinvastaisia tuloksia, eli fonologinen samankaltaisuus hidasti tunnistamista, kuten eräässä aiemmassa tutkimuksessakin oli havaittu (Comesaña et al., 2015).

Keskenään ristiriitaiset tulokset virheanalyysissä ja reaktioaika-analyysissä voivat selittyä aineistonkeruuseen käytetyllä työkalulla. Internet-pohjainen reaktioajan mittaus saattaa olla altis esimerkiksi eri selainten ja laitteistojen välisille eroille. Virheanalyysin tuloksiin nämä eivät vaikuta, joten niitä voidaan pitää luotettavampina.

Muita tutkimuksen rajoitteita on esimerkiksi koehenkilöiden kielellisten taustojen moninaisuus: taustaolettamuksena on, että oppijan kaikki kielet aktivoituvat, kun yhtä kieltä käytetään. Näin myös muut osatut kielet ovat voineet vaikuttaa sanantunnistukseen. Suurin osa vastaajista kuitenkin piti englantia vahvimpana vieraana kielenään ja se oli suurimman osan ensimmäinen vieras kieli oppimisjärjestyksessä. Lisäksi on huomioitava, että suomi ja sen konsistentit kirjain-äänne -vastaavuudet ovat saattaneet vaikuttaa kielelliseen prosessointiin tehtävän aikana.

Tulevaisuudessa aihetta voisi tutkia lisää esimerkiksi eri aineistonkeruumenetelmillä. Olisi myös kiinnostavaa nähdä, millaisen vaikutuksen kolmikieliset suomi-englanti-ranska -kognaatit (kuten *tunneli*, *tunnel*) aiheuttaisivat sanantunnistuksessa. Lisäksi oppijoiden kielitaidon tason vaikutusta voisi analysoida tarkemmin, sillä se saattaa osaltaan vaikuttaa kognaattien prosessointiin.

Tämän tutkimuksen tärkeä viesti on, että kognaattien fonologisen samankaltaisuuden aste tulisi ottaa huomioon, kun tutkitaan kognaatteja ja muodostetaan teorioita monikielisestä sanantunnistuksesta. On myös aihetta pohtia, voisiko kieltenopetuksessa kiinnittää huomiota eri kielten sekä ortografian ja fonologian vertailuun.