

Chapitre 3

Les verbes (ir)réguliers français

L'élève : Les racines des mots sont-elles carrées ?

Le professeur : Carrées ou cubiques. C'est selon.

Eugène Ionesco (La leçon ; p. 125.)

3.1 Introduction

Dans la série d'expériences présentée dans ce chapitre nous testons la nature des représentations stockées dans le lexique mental des formes verbales en français. Une particularité du système français concernant les formes verbales est que celles-ci sont toutes décomposables quelle que soit la catégorie de verbes considérée, régulière ou irrégulière. En effet, ces formes verbales sont composées d'une racine et d'un suffixe. Néanmoins, ce qui distingue les formes régulières des formes irrégulières est que la racine de ces dernières peut connaître une variation orthographique idiosyncrasique (par ex. le verbe *boire* donne la forme fléchie *buvait*). De ce fait, le système flexionnel français comporte une spécificité par rapport à la langue anglaise pour laquelle toutes les formes verbales ne sont pas décomposables -les formes verbales régulières au passé le sont, mais pas les formes verbales irrégulières. Étant donné que de nombreuses études en psycholinguistique ont été réalisées en langue anglaise, il est intéressant d'exploiter le système flexionnel français afin d'élargir

nos connaissances sur certaines questions. Par exemple, du point de vue du traitement lexical engagé sur les formes verbales, la question que pose les formes composées d'une racine irrégulière est de savoir si une procédure décompositionnelle peut permettre à un locuteur donné d'accéder à la signification de la forme verbale. En effet, la décomposition des formes verbales régulières comme *manges* peut permettre l'activation de la racine *mang-* et du suffixe *-es*, et de là, retrouver le sens relatif au verbe *manger* (*mang-er*, avec *-er* qui indique l'infinitif) et les indications de temps, de mode et de personne du suffixe. Par contre cette procédure semble moins évidente dans le cas des formes verbales irrégulières. En effet, pour celles-ci, la racine isolée par la procédure de décomposition est de nature idiosyncrasique et donc très différente dans sa forme de celle du verbe dont elle est issue. Par exemple, la racine idiosyncrasique *bu-* dans *buvais* est une racine du verbe *boire*, tout comme la racine *ir-* dans *irons*¹⁶ est une racine idiosyncrasique du verbe *aller*. Par ailleurs, ces verbes irréguliers ont parfois des formes régulières : *allons* est composé de la racine régulière *all-* du verbe *aller*. De ce fait, il est légitime de se demander si les racines irrégulières ont une représentation propre au sein du lexique mental permettant d'activer le sens relatif au verbe dont elles sont issues, ou, si les formes irrégulières qui sont ponctuelles, sont traitées via leur forme globale. Nous allons également étudier la nature des représentations des formes verbales morphophonologiques (ou semi-régulières). Celles-ci sont composées d'une racine dont la forme phonologique connaît une légère transformation. Cette transformation suit une règle morphophonologique. Par exemple, *sèmeront* contient la racine allomorphique *sèm-* du verbe *semer*.

Pour tester si ces racines sont représentées au sein du lexique mental, nous avons utilisé une tâche de décision lexicale simple (sans amorce). Les formes verbales testées (régulières à irrégulières) peuvent être de deux types en fonction de la fréquence de surface de la forme verbale et de la fréquence cumulée de la racine (haute versus faible).

¹⁶En fait le *-r* dans *ir-* ou *-rons* peut être considéré comme la marque du futur, auquel cas la racine du verbe *aller* minimaliste est *i*.

3.2 La morphologie flexionnelle et les formes verbales en anglais et en français

Comme nous l'avons noté dans le Chapitre 1 (p.21), les études sur le traitement lexical des formes fléchies en linguistique et psycholinguistique n'ont d'abord suscité que peu d'intérêt car elles étaient considérées comme des cas particuliers reliés au domaine de la syntaxe et non à celui du lexique. Cependant, les cas d'irrégularités et les questions qu'ils soulèvent quant à leur traitement cognitif connaissent aujourd'hui un regain d'intérêt.

3.2.1 Les formes fléchies en anglais et en français

Les formes verbales fléchies au passé en anglais se divisent en deux catégories, avec d'une part les formes verbales régulières, composées du verbe à l'infinitif et de la terminaison *-ed* (par ex. *to talk* 'parler' donne *talked* au passé), et d'autre part les formes verbales irrégulières idiosyncrasiques (par ex. *to teach* 'enseigner' donne *taught* au passé). Ainsi, en anglais certains verbes se conjuguent au passé par l'application d'une règle (ajout de *-ed*), pour d'autres, la forme au passé doit être encodée en mémoire. A ce niveau-ci, nous devons noter que, d'un point de vue purement linguistique, les verbes irréguliers anglais ne sont pas toujours considérés comme entièrement idiosyncrasiques. Halle et Marantz (1993) soulignent que le choix du suffixe est soit $-\emptyset$ (par ex. *hit-hit* 'frapper' en français) soit *-t* (par ex. *sleep-slept* 'dormir' en français) adjoint à la forme phonologique alternative de la racine, elle-même déterminée par l'identité du verbe et spécifiée dans une liste stockée en mémoire. La construction des formes fléchies en français est différente de l'anglais. Premièrement, le français est une langue que l'on peut qualifier de très flexionnelle dans le sens où il existe plusieurs suffixes pour indiquer à la fois le temps auquel est conjugué un verbe, mais également le mode et la personne. Deuxièmement, nous notons deux principales différences avec l'anglais concernant leur forme de surface. La première est que toutes les formes verbales ont un suffixe et sont donc toutes décomposables quelle que soit la catégorie de verbes considérée. La seconde, qui découle de la première, est que

les verbes irréguliers sont composés d'un suffixe régulier mais d'une racine idiosyncrasique dans certaines formes fléchies (par ex. le verbe *boire* donne *buw-ait* à la troisième personne de l'imparfait de indicatif). Par ailleurs, nous pouvons observer le cas des verbes semi-réguliers, appelés aussi morphophonologiques, dont les racines subissent une variation allomorphique pour certaines formes fléchies (par ex. le verbe *semer* comporte deux racines possibles, *sem-* [səm] dans *semait* [səme] et *sèm-* [sɛm] dans *sèment* [sɛm]). La règle qui sous-tend ce changement allomorphique de la racine est la suivante : le [ɛ] dans la racine devient un [ə] quand la racine est suivie d'une syllable contenant un schwa (Deyhime, 1967).

3.2.2 Les études sur les formes verbales régulières et irrégulières.

Le cas de l'anglais

Un grand nombre d'études a été mené afin de tester les deux hypothèses alternatives concernant le traitement lexical des formes verbales régulières et irrégulières en anglais. D'un côté les partisans de la double voie ou double procédure (par ex. Baayen, Dijkstra & Schreuder, 1997 ; Clashen, 1999 ; Marslen-Wilson & Tyler, 1997, 1998, 2003 ; Pinker & Prince, 1988 ; Pinker & Ullman, 2002 ; Ullman & al., 1997), prônent l'application d'une procédure analytique pour les formes fléchies régulières et une procédure par accès direct pour les formes fléchies irrégulières non décomposables. De l'autre côté, les tenants des modèles connexionnistes qui proposent que toutes les formes fléchies sont traitées par une procédure unique, l'accès direct (McClelland & Patterson, 2002 ; Plunckett & Marchman, 1993 ; Seidenberg & Gonnerman, 2000 ; voir le Chapitre 1 pour plus de détails sur ces deux hypothèses). Plusieurs études ont essayé de tester ces hypothèses en utilisant le paradigme d'amorçage en intra ou inter modalité ¹⁷ avec en amorce une forme fléchie régulière ou irrégulière (*talked* ou *drove*) et en cible la racine de ces formes (*talk* ou *drive*). Il ressort

¹⁷Voir le Chapitre 2 pour plus de détails sur ce paradigme.

de ces études que les formes fléchies régulières amorcent leur racine (par ex. *talked* amorce sa racine *talk*) alors que ce n'est pas observé dans le cas où l'amorce est une forme fléchie irrégulière (amorçage très diminué ou absence d'amorçage entre *drove* et *drive*, voir par exemple Kempley & Morton, 1982 ; Napps, 1989). Ce résultat a été interprété dans le cadre des modèles à double voie : celle par décomposition pour traiter les formes régulières (*talk-ed*) et celle par accès direct pour les formes irrégulières. La première voie conduit à l'isolation de la racine qui reçoit alors de l'activation (*talk*) ce qui permet quand elle est présentée en tant que cible d'être reconnue plus rapidement. Nous rappelons ici que les verbes irréguliers anglais au passé ne sont pas complètement idiosyncrasiques selon une étude de Hallé et Marantz (1993). Néanmoins, la décomposition de *slept* en *slep-* et *-t* (puisque selon ces auteurs *slep* est une forme allomorphique de *sleep* et *-t* est dans ce cas le suffixe) conduit à l'isolation d'une racine qui n'en est pas une en anglais. Il n'existe pas d'autre forme polymorphémique construite sur la base de *slep-*. De ce fait, il semble difficile à ce niveau de réflexion de caractériser ces formes verbales comme décomposables dans le cadre des procédures cognitives mises en œuvre au cours de leur accès lexical. Nous reviendrons sur cette question à la fin de ce chapitre. Bien que les résultats présentés ci-dessus aient été abondamment cités dans le cadre des études défendant l'hypothèse d'une double voie de traitement pour accéder aux sens des formes polymorphémiques régulières et irrégulières, une interprétation dans le cadre des modèles connexionnistes a également été proposée. Dans ce cadre-ci, les effets d'amorçage sont des effets émergeant de l'organisation même du lexique mental. Le lexique est composé de différentes unités pour chaque mot d'une langue donnée, avec des connexions lexicales entre les unités qui partagent des caractéristiques orthographiques, phonologiques et sémantiques¹⁸. Quand la représentation d'un mot polymorphémique est activée (*talked*), cela propage de l'activation à toutes les autres représentations lexicales qui lui sont interconnectées. Dans le cas des formes polymorphémiques composées d'une même racine (*talk*, *talking*, *talks*),

¹⁸Nous renvoyons le lecteur au Chapitre 1 pour l'interprétation de ce résultat selon l'approche connexionniste.

cela conduit, par un phénomène d'émergence, à l'activation de la racine commune (*talk* dans l'exemple cité). Par ailleurs, des études réalisées sur des langues romanes (par ex. l'italien ou le français) montrent des effets d'amorçage entre une amorce irrégulière et sa racine (Orsolini & Marslen-Wilson, 1997; Meunier & Marslen-Wilson, 2004), suggérant un processus unique d'accès lexical pour les formes verbales régulières et irrégulières.

Le cas du français

Meunier et Marslen-Wilson (2004) ont observé dans deux études utilisant le protocole d'amorçage, une en intra et l'autre en inter modalité, un effet d'amorçage similaire entre une forme fléchie et sa racine, que la forme fléchie soit régulière, morphophonologique ou irrégulière. En d'autres termes, l'effet d'amorçage qu'ils observent entre les paires *mangeons-manger*, ou *boirons-boire* et *semons-semer* n'est pas différent de celui observé pour les paires *buvais-boire* et *sèment-semer*¹⁹. Si nous reprenons le principe de l'amorçage, la facilitation observée pour traiter un mot cible lorsque celui-ci est précédé d'une amorce polymorphémique dont la base est la cible, s'explique par le fait que cette dernière a été pré-activée au cours du traitement de l'amorce. Ainsi, dans le cas des formes verbales régulières comme *walked* en anglais, l'effet d'amorçage sur la reconnaissance de *walk* indique que sa représentation a été pré-activée, vraisemblablement au cours du traitement de l'amorce *walked* décomposée en *walk-* et *-ed*. En ce qui concerne les effets d'amorçage obtenus en français, la même explication peut être retenue concernant les formes régulières : *mangeons* est décomposé en *mang-eons* ce qui propage de l'activation à la représentation de la racine *mang-* présente dans la cible *manger* (*-er* indique la forme infinitive des verbes du premier groupe). Concernant les formes morphophonologiques, l'effet d'amorçage peut s'expliquer à partir de l'hypothèse proposée par Lahiri et Marslen-Wilson (1991) puis reprise par Marslen-Wilson et Zhou en 1999 dans une étude qui porte

¹⁹Bien que l'effet d'amorçage soit numériquement différent entre la condition "verbe régulier" et "verbe irrégulier", celui-ci n'est jamais significatif dans les analyses statistiques (voir Meunier & Marslen-Wilson, 2004 pour plus de détails).

sur les formes dérivées dont les racines sont allomorphiques. Selon cette hypothèse, une forme allomorphique partage une représentation lexicale phonologiquement abstraite avec la forme canonique. Cette représentation aurait des traits phonétiques sous-spécifiés. En d'autres termes la différence dans la forme phonologique des deux racines ne conduit pas à l'activation de deux représentations lexicales distinctes. En effet, ces auteurs ont observé que la reconnaissance du mot cible *sane* [sem] en anglais, est à la fois amorcée par le mot dérivé *sanely* [semɛli], et par *sanity* [saenɪti] dans une tâche de décision lexicale avec un protocole d'amorçage en inter modalité (l'amorce est présentée auditivement et la cible en modalité visuelle). Ainsi, lorsque que les amorces *sanely* ou *sanity* sont présentées, une procédure de décomposition isole la racine *sane* prononcée différemment dans les deux amorces sans qu'une différence d'amorçage sur la reconnaissance du mot cible *sane* ne soit observée. Cela suggère que les formes allomorphiques et canoniques partagent une représentation commune et phonologiquement abstraite au sein du lexique mental. Cet effet d'amorçage a également été obtenu entre des mots dérivés dont l'un est composé d'une racine allomorphique et l'autre de la racine canonique (*sanely* amorce *sanity*). De la même manière, en français, la forme verbale *sèment* peut être décomposée en *sèm-* et *-ent*, et si les racines *sèm-* et *sem-* partagent une représentation commune, alors il n'est pas surprenant d'observer un effet d'amorçage comme observé dans l'étude de Meunier et Marslen-Wilson (2004). A partir du résultat observé en anglais pour *sanely* et *sanity*, qui amorcent tous deux leur racine *sane*, Marslen-Wilson et Zhou (1999) ont aussi proposé l'hypothèse d'une organisation du lexique mental à un seul niveau de représentation (appelé *lemma*). Selon ce modèle, les informations sémantiques, syntaxiques et celles sur la forme phonologique (abstraite) d'un mot, sont disponibles à partir d'une seule représentation amodale (voir Figure 3.1).

Concernant maintenant l'effet d'amorçage observé en français entre les formes verbales irrégulières et leur verbe correspondant, l'explication qui consiste à dire que la racine-cible est pré-activée au cours du traitement de l'amorce (ce qui lui permet d'être reconnue plus rapidement quand elle est présentée de manière isolée une seconde fois pour la tâche

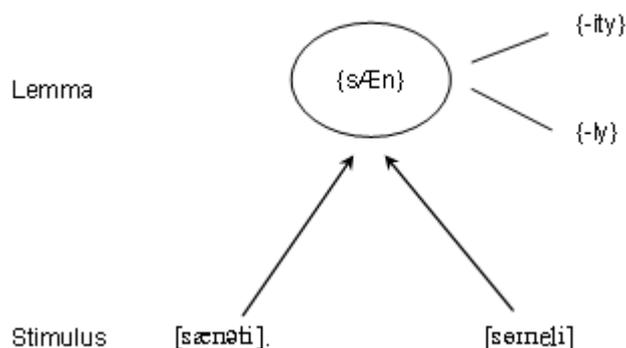


FIG. 3.1 – Représentation schématisée du modèle à un niveau de représentation (lemma) d’après Marslen-Wilson et Zhou (1999).

de décision lexicale) ne semble plus être valable. En effet, la racine d’une forme fléchie irrégulière en français, comme *buvais* [byvɛ] n’a que très peu de traits orthographiques communs avec la cible *boire* [buar]. Il semble donc difficile de concevoir que *buvais* puisse pré-activer la représentation de *boire*. Une explication de cet effet d’amorçage pourtant observé dans l’étude de Meunier et Marslen-Wilson (2004) peut cependant être avancée avec le modèle à deux niveaux de représentation de Allen et Badecker (2002) que nous allons à présent détailler.

Un ou des niveau(x) de représentation au sein du lexique mental

Lorsque nous reconnaissons un mot, nous accédons à différentes informations le concernant comme sa forme phonologique ou graphique, les affixes qui peuvent s’y adjoindre (*mangeable, mangeoire, *mangisme*) ainsi que les informations syntaxiques (*ils mangent, *ils manges*) et sémantiques correspondant. Toutes ces propriétés doivent être encodées au sein du lexique mental et activées au cours de l’accès lexical. Cela conduit à s’interroger quant au(x) niveau(x) de représentation qui est/sont alors engagé(s). Cette question peut être légitimement envisagée dans le cadre des études sur l’accès lexical des mots

polymorphémiques. Etant donné que le système lexical doit permettre de retrouver la représentation sémantique exacte depuis n'importe quelle forme polymorphémique (régulière ou irrégulière), Allen et Badecker (2002) envisagent, par opposition au modèle d'organisation de Marslen-Wilson et Zhou (1999), une organisation lexicale qui comprend au moins deux niveaux de représentation : un niveau qui encode les informations lexicales de nature abstraite (sémantique, syntaxe) et un niveau encodant les détails orthographiques ou phonologiques des mots (voir Figure 3.2).

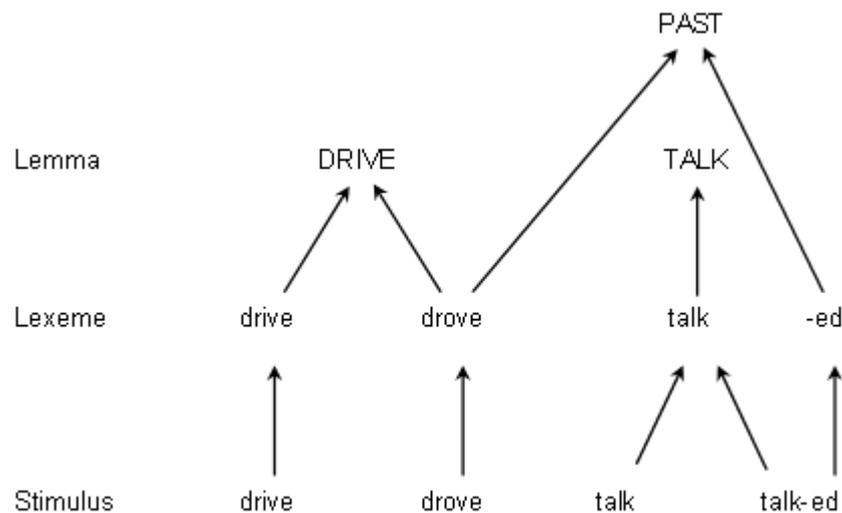


FIG. 3.2 – Représentation schématisée du modèle à deux niveaux (lemma-lexeme) d'après Allen et Badecker, (2002) et appliqué aux verbes *to drive* et *to talk*

Ces deux niveaux peuvent être nommés *lemma* et *lexeme* en reprenant la nomenclature utilisée dans les études en production de parole (par ex. Levelt, Roelofs & Meyers, 1999). Cette conception quant à l'organisation du lexique mental se base sur une étude d'amorçage en inter modalité proposée par les auteurs (Allen & Badecker, 2002) en anglais. Ils ont observé un effet d'amorçage dans le cas d'un faible recouvrement orthographique entre une forme verbale irrégulière et son infinitif, (par ex. *bought-bring* du verbe *to bring* 'apporter'). En revanche, cet effet n'est pas observé lorsque le recouvrement orthographique

est important (par ex. *drive-drove* du verbe *to drive* ‘conduire’). Les auteurs concluent de cette manière : les représentations au niveau du *lexeme* entrent en compétition quand elles partagent trop de caractéristiques formelles (inhibition orthographique). Si elles en partagent peu, un effet facilitateur peut être observé par l’activation de la représentation commune au niveau du *lemma*. (La Figure 3.2 représente le modèle à deux niveaux de représentation et la Figure 3.3 représente un modèle à un seul niveau de représentation).

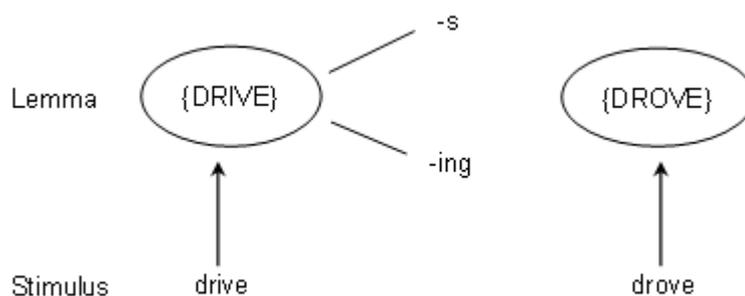


FIG. 3.3 – Représentation schématique du modèle à un niveau de représentation de Marslen-Wilson et Zhou, (1999) appliqué au verbe irrégulier *to drive*.

Grâce aux expériences proposées dans ce chapitre, nous testons la nature des représentations des racines dans les trois classes verbales du français : régulière, morphophonologique et irrégulière. Plus précisément, nous testons l’existence d’une représentation commune, phonologiquement abstraite, pour les racines de formes verbales régulières à irrégulières. Afin de tester cette hypothèse nous utilisons une tâche de décision lexicale simple en manipulant l’effet de fréquence de surface et de fréquence cumulée des verbes utilisés.

3.3 Manipulation de la fréquence de surface et de la fréquence cumulée des formes verbales

Burani, Salmaso et Caramazza (1984) ont testé en italien l'influence de la fréquence de surface et de la fréquence cumulée de formes verbales régulières dans une tâche de décision lexicale simple (sans la présentation d'une amorce). Trois conditions étaient testées. Dans la première, les formes verbales ont une fréquence de surface et une fréquence cumulée élevées. Dans la seconde, les formes verbales ont une fréquence de surface élevée et une fréquence cumulée faible. Enfin dans la troisième condition, la fréquence de surface et la fréquence cumulée sont faibles. Nous rappelons que la fréquence de surface correspond à la fréquence d'occurrence du mot polymorphémique considéré comme un item lexical isolé (par ex. la fréquence d'occurrence de *chantait*) et que la fréquence cumulée correspond à la somme des fréquences de tous les mots qui partagent une même racine (par ex. la fréquence des mots comme *chant*, *chantait*, *chantons*...). Ces notions impliquent que lors du traitement d'un mot polymorphémique, sa représentation globale est activée dans le cas de la fréquence de surface et la représentation de sa racine est activée si on considère la fréquence cumulée. Burani et al., (1984) observent que la fréquence de surface et la fréquence cumulée influencent les temps de reconnaissance de ces formes. Ils obtiennent des temps de réponse plus rapides lorsque les formes verbales ont une fréquence de surface élevée, et également si elles ont une fréquence cumulée élevée. Ces résultats montrent que le traitement lexical des formes fléchies engage l'activation de la représentation de leur racine ainsi que leur représentation holistique.

Concernant notre étude sur la reconnaissance de formes verbales régulières à irrégulières en français, nous nous attendons à observer un effet de la fréquence de surface avec des temps de reconnaissance plus rapides pour les formes fléchies de fréquence élevée. Pour les formes régulières, si nous observons un effet de la fréquence cumulée, cela indiquera que deux racines distinctes sont activées au cours du traitement de deux formes fléchies de deux verbes différents. Ensuite, tout en utilisant la même procédure, nous testons ces

effets sur des verbes morphophonologiques et irréguliers avec toujours le même raisonnement concernant l'effet de la fréquence cumulée : la présence d'un tel effet signifiera que deux racines distinctes ont été activées bien que dans le cas de ces verbes, l'opposition des fréquences cumulées hautes *vs.* basses se fasse à partir d'un même verbe. Un effet de la fréquence de surface montrerait l'activation d'une représentation globale au cours de l'accès lexical aux formes fléchies.

3.4 Expérience 1 : Verbes réguliers

3.4.1 Méthode

Participants

Quarante huit étudiants de l'Institut d'Etudes Politiques de l'Université Lumière Lyon 2 ont participé à l'expérience. Ils étaient tous de langue maternelle française, avaient une vision normale ou corrigée et aucun ne présentait de trouble du langage. Leur âge était compris entre 20 et 30 ans. Ils ont été indemnisés pour leur participation. Les mêmes 48 étudiants ont participé aux Expériences 2 et 3 pour lesquelles les formes verbales utilisées étaient différentes. L'ordre de passation des Expériences 1, 2 et 3 était contre-balancé.

Stimuli

Dans cette expérience, les verbes testés sont tous réguliers. Les formes fléchies dans ce groupe de verbes sont construites par la concaténation d'une racine et d'un suffixe différent selon le temps, la personne et le mode auquel est conjugué le verbe. Par exemple, pour le verbe *manger*, on peut trouver les formes suivantes : *mangeais* au passé du mode indicatif première personne du singulier ou *mangeons* au présent du mode indicatif, première personne du pluriel. Quinze paires de verbes ont ainsi été sélectionnées. Pour chacune de ces paires, l'un des verbes possède une fréquence cumulée élevée (par ex. *manger*, dont la fréquence cumulée est de 158.6 occurrences par million) et l'autre verbe possède une

fréquence cumulée faible (par ex. *grimper*, qui possède une fréquence cumulée de 27.3 occurrences par million). Un exemple est donné dans le Tableau 3.1.

Verbes réguliers				
	Fréquence cumulée haute		Fréquence cumulée basse	
	Fréquence de surface		Fréquence de surface	
	haute	basse	haute	basse
Exemple	mang-ent	mang-eas	grimp-e	grimp-era

TAB. 3.1 – Exemples de formes verbales utilisées dans l’Expérience 1.

Pour chacun des verbes réguliers, nous avons sélectionné une forme fléchie avec une fréquence de surface élevée (par ex. *mangent* et *grimpe* qui ont respectivement des fréquences de 5.23 et 4.71 occurrences par million) et une forme fléchie avec une fréquence de surface basse (par ex. *mangeas* et *grimpera* possédant toutes les deux une fréquence de 0.06 occurrences par million). Nous avons choisi notre matériel verbal de sorte que la différence entre le logarithme de la fréquence des formes verbales contrastées (formes verbales avec une fréquence cumulée élevée *vs.* basse ou une fréquence de surface élevée *vs.* basse) soit supérieur à 1. En tout, nous avons sélectionné 60 formes verbales qui constituent nos stimuli expérimentaux. Les fréquences de surface moyennes, les fréquences cumulées et la moyenne de la longueur des formes verbales sélectionnées sont reportées dans les annexes. De plus, 60 non-mots ont été construits pour équilibrer les réponses des sujets à la tâche de décision lexicale qui leur était demandée. Ces non-mots ont été construits en associant une pseudo-racine et un vrai suffixe choisi parmi ceux utilisés dans la liste de mots expérimentaux. Ce sont donc des pseudo formes verbales. Dix essais ont également été proposés aux participants en entraînement avant de passer à la phase de test.

Quatre listes expérimentales ont été construites. Les formes verbales étaient présentées

dans un ordre semi aléatoire, de façon à ce que : (a) un même type de stimulus, entraînant une même réponse, “oui” ou “non”, n’apparaisse pas plus de trois fois à la suite ; (b) il n’y ait pas plus de deux items à la suite avec une fréquence similaire ; (c) au moins 10 stimuli soient présentés entre deux formes verbales différentes mais appartenant à un même verbe ; (d) il n’y ait pas plus de deux stimuli successifs partageant leurs lettres initiales ou leur suffixe. Par ailleurs, quatre autres listes expérimentales sont construites pour lesquelles l’ordre de présentation des stimuli est l’inverse des quatre premières listes. Chaque participant est testé une seule fois sur une seule des huit listes expérimentales.

Procédure

L’expérience est pilotée par un ordinateur PC grâce au logiciel DMDX (Forster & Forster, 2003). Chaque essai commence par la présentation d’un point de fixation au centre de l’écran pendant 500 ms, suivi d’un stimulus verbal écrit en majuscules avec accentuation. Le stimulus reste à l’écran jusqu’à ce que le participant donne une réponse. La tâche cognitive demandée est une tâche de décision lexicale. Après 2000 ms, si le participant n’a pas répondu, l’essai suivant est automatiquement proposé. Les participants ont pour consigne de répondre le plus vite et le plus correctement possible “oui” si l’item présenté à l’écran est un mot de la langue française et “non” si cela n’est pas le cas. Les réponses sont données grâce à une manette de jeu WingMan Logitech. La réponse “oui” est donnée avec la main dominante du sujet. Avant l’expérience à proprement parlée les participants font une série de 10 essais pour s’entraîner.

3.4.2 Résultats

Les temps de réponse inférieurs à 300 ms et supérieurs à 1200 ms sont exclus des analyses car considérés comme déviants au vu de la tâche demandée (cela exclut 1.5 % des données). Par ailleurs les 1.9 % d’erreurs (correspondant à une réponse différente de la réponse attendue) ont été exclus des analyses. Etant donné ce faible taux d’erreurs, nous avons uniquement analysé les temps de réponse. Deux analyses de variance (ANOVA)

ont été conduites sur les temps de réponses : une analyse prenant le facteur sujet comme variable aléatoire ($F1$) et l'autre prenant le facteur item ($F2$). Un résumé des résultats est présenté dans le Tableau 3.2 et la Figure 3.4.

Temps de réponse des sujets pour les formes verbales régulières, en fonction de leur fréquence cumulée et de leur fréquence de surface.

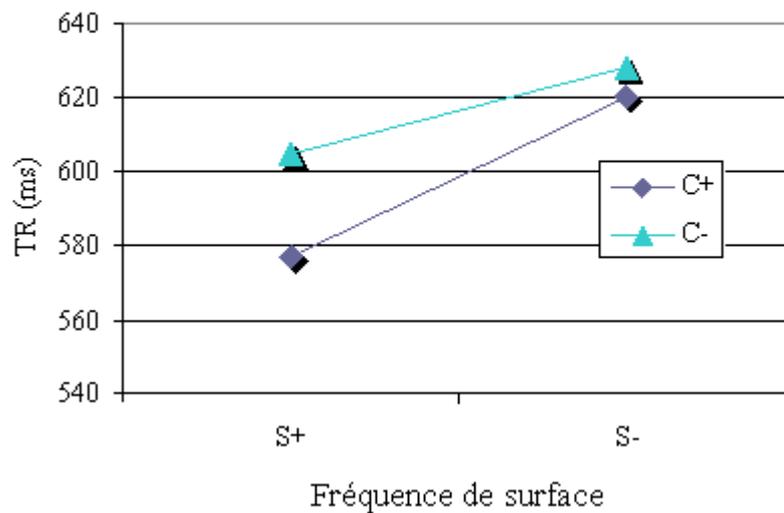


FIG. 3.4 – Graphique représentant les temps de réponse obtenus dans l'Expérience 1.

Les analyses montrent un effet significatif de la fréquence de surface [$F1(1,47) = 47.35$, $p < .001$, $F2(1,28) = 20.26$, $p < .001$]. Les formes verbales de basse fréquence sont reconnues plus lentement que celles de fréquence élevée. L'effet de la fréquence cumulée est présent uniquement dans l'analyse par sujet, alors qu'une tendance est observée dans l'analyse par item [$F1(1,47) = 14.75$, $p < .001$; $F2(1,28) = 2.93$, $p = .097$] : les formes verbales de fréquence cumulée élevée sont reconnues plus rapidement que celles de fréquence cumulée faible. L'interaction entre les deux facteurs fréquence est marginalement significative dans l'analyse par sujet et non significative dans l'analyse par item [$F1(1,47) = 3.21$, $p = .08$; $F2(1,28) = 1.58$, n.s.]. L'effet de fréquence est plus important dans la condition où les formes verbales ont une fréquence de surface élevée (29 ms) que dans la condition où les formes verbales ont une fréquence de surface basse (8 ms). Une comparaison *post-*

hoc confirme cette tendance, l'effet de la fréquence cumulée pour les formes verbales de fréquence de surface haute est significative par sujet et présente une forte tendance par item : [t_1 (47) = 3.37, $p < .001$, t_2 (11) = 1.57, $p = .07$]. Pour les formes verbales de fréquence de surface faible nous n'observons pas d'effet significatif [t 's <1]. Nous discutons de ce résultat dans la discussion générale, après avoir analysé les résultats obtenus dans les Expériences 2 et 3 présentées ci-dessous.

Verbes réguliers				
	<u>Fréquence cumulée haute</u>		<u>Fréquence cumulée basse</u>	
	Fréquence de surface		Fréquence de surface	
	haute	basse	haute	basse
Exemple	mang-ent	mang-eas	grimp-e	grimp-era
TR	576	620	605	628

TAB. 3.2 – Résumé des temps de réponse obtenus dans l'Expérience 1.

3.4.3 Discussion

Les résultats de cette première expérience montrent un effet de la fréquence cumulée et de la fréquence de surface dans la reconnaissance des formes verbales régulières. L'effet de fréquence de surface est une observation classique dans ce type d'étude et indique l'implication de la forme globale dans le traitement lexical des formes verbales (Scarborough, Cortese & Scarborough, 1977). Taft (2004) de son côté a proposé que cet effet puisse refléter une procédure de recombinaison entre la représentation de la racine et celle du suffixe. Nos résultats ne permettent pas de conclure en faveur de l'une ou l'autre interprétation. Ce qui retient notre attention en revanche, est l'effet de fréquence cumulée. Cet effet peut être interprété comme une indication que la représentation de la racine est activée au cours de la reconnaissance des formes verbales régulières. Nous allons à

présent étudier le cas des verbes dit morphophonologiques. Pour ces verbes il existe deux racines possibles, par exemple *semais* [səm] vs. *sèment* [sɛm] pour le verbe *semer* [səmɔ]. Ces deux racines sont orthographiquement et phonologiquement proches, et alternent en fonction d'une règle phonologique. Comme nous l'avons évoqué précédemment, une étude réalisée par Marslen-Wilson et Zhou (1999) soutient que la forme allomorphique et la forme canonique d'une racine partagent une représentation phonologiquement abstraite dans le lexique mental. Les travaux en français de Meunier et Marslen-Wilson (2004) montrent un effet d'amorçage entre une forme verbale à racine allomorphique (*sèment*) et sa forme à l'infinitif (*semer*). Dans l'Expérience 2 présentée ci-dessous, nous testons l'hypothèse d'une représentation commune pour les deux types de racines avec un paradigme permettant de tester cette hypothèse de manière plus directe.

3.5 Expérience 2 : Verbes morphophonologiques

Dans cette deuxième expérience, nous testons la reconnaissance de formes verbales appartenant au groupe des verbes morphophonologiques. Nous voulons établir si leur reconnaissance implique l'activation de représentations spécifiques en fonction de la nature d'une racine. Dans ce but, nous avons calculé séparément la fréquence cumulée des deux racines possibles (i.e. la racine canonique et allomorphique) pour les verbes de cette classe. Nous avons testé si cette variable a un effet sur les temps de reconnaissance de ces formes verbales. Comme dans l'expérience précédente, nous avons testé les temps de réponse dans la condition où un verbe de cette classe (par exemple le verbe *semer*) est présenté sous une forme fléchie de fréquence cumulée faible (*sèm-*) et lorsqu'il est présenté sous une forme fléchie de fréquence cumulée élevée (*sem-*), mais aussi quand la forme fléchie de fréquence de surface basse (*semais - sèmerai*) ou élevée (*semée - sème*). Contrairement à l'Expérience 1, nous avons sélectionné ces quadruplets à partir d'un seul verbe (dans l'exemple *semer*). En tout, quinze quadruplets ont été sélectionnés.

3.5.1 Méthode

Les participants et la procédure l'Expérience 2 sont identiques à ceux de l'Expérience 1.

Stimuli

Dans cette expérience, les verbes de la classe morphophonologique sont testés. Leurs formes conjuguées ont deux racines possibles (par ex. *semais* [səmə] vs. *sément* [sɛm] pour le verbe *semer* [səmɛ]), qui alternent en fonction d'une règle phonologique. Quinze verbes de ce type ont été sélectionnés pour lesquels nous pouvions comparer 4 formes verbales :

1. une de fréquence cumulée et de fréquence de surface élevée (par ex. *semée* avec une fréquence cumulée de 12.24 occurrences par million et une fréquence de surface de 2.45 occurrences) ;
2. une forme verbale de fréquence cumulée élevée et de fréquence de surface faible (par ex. *semais* avec une fréquence cumulée de 12.24 occurrences par million et une fréquence de surface de 0.03 occurrences) ;
3. une forme verbale de fréquence cumulée et de fréquence de surface faible (par ex. *sèmerais*, avec une fréquence cumulée de 2.5 occurrences par million et une fréquence de surface de 0.03 occurrences) ;
4. une quatrième forme de fréquence cumulée faible et de fréquence de surface élevée (par ex. *sème* avec une fréquence cumulée de 2.5 occurrences par million et une fréquence de surface de 0.03 occurrences).

Voir le Tableau 3.3 pour un récapitulatif des 4 formes verbales testées.

La différence de logarithme de la fréquence des formes verbales qui sont comparées (formes verbales de fréquence cumulée élevée vs. basse ou de fréquence de surface élevée vs. basse) est supérieure à 1.

Verbes morphophonologiques				
	Fréquence cumulée haute		Fréquence cumulée basse	
	Fréquence de surface		Fréquence de surface	
	haute	basse	haute	basse
Exemple	sem-ée	sem-ais	sèm-e	sèm-erai

TAB. 3.3 – Exemples de stimuli expérimentaux utilisés dans l’Expérience 2.

3.5.2 Résultats

Par erreur, un quadruplet de formes conjuguées a été présenté deux fois à chaque participant (quadruplet du verbe *protéger*). Par conséquent nous avons exclu de nos analyses les temps de réponse des 4 formes verbales de ce quadruplet. Les temps de réponses déviants (TR compris entre 300 et 1200 ms) ont été exclus des analyses (3.6 % des données). Le taux d’erreur est de 1.25 %. Etant donnée ce très faible taux, aucune analyse statistique n’a été réalisée sur les erreurs. Un résumé des temps de réponse est présenté le Tableau 3.4 et la Figure 3.5.

Nous observons un effet significatif de la fréquence de surface sur les temps de réponse dans l’analyse par sujet et par item [$F1(1, 47) = 29.18, p < .001, F2(1, 22) = 11.80, p = .002$]. Les temps de réponse sont plus courts pour les formes verbales de fréquence de surface élevée. Aucun effet de la fréquence cumulée n’est observé [$F_s < 1$] et nous n’observons pas d’interaction entre ces deux facteurs [$F_s < 1$].

3.5.3 Discussion

L’analyse statistique des résultats montre que la fréquence de surface influence les temps de réponse sur la reconnaissance des formes verbales morphophonologiques. Comme nous l’avons signalé dans la discussion précédente, cet effet indique l’activation de la forme

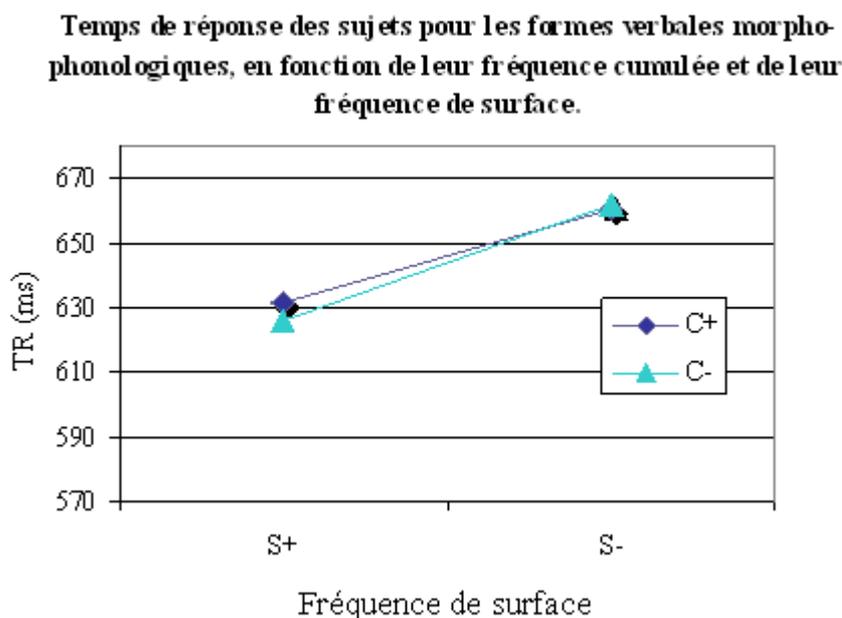


FIG. 3.5 – Graphique représentant les temps de réponse observés dans l’Expérience 2.

globale dans le traitement lexical des formes verbales, ou tout du moins, une recombinaison entre la racine et le suffixe. Cependant, nous n’observons pas d’effet de la fréquence cumulée. Cela pourrait indiquer que ces formes verbales ne sont pas décomposées au cours de leur traitement lexical. Une telle interprétation semble quelque peu hâtive. En effet, il serait surprenant que les formes verbales composées d’un suffixe régulier et d’une racine régulière ou allomorphique (orthographiquement proche de la première) ne soient pas décomposées comme les formes verbales régulières. Une interprétation possible serait que la décomposition de ces formes verbales morphophonologiques active une représentation commune pour les deux type de racines (comme proposé par Lahiri & Marslen-Wilson, 1991, et Marslen-Wilson & Zhou, 1999). Si tel est le cas, les deux formes de surface (*sema**is*, *sème**ra**is*) activeraient la représentation d’une racine de fréquence cumulée unique. En d’autres termes, les formes verbales de notre catégorie “fréquence cumulée faible” auraient la même représentation et la même fréquence cumulée que nos formes verbales de “fréquence cumulée haute”. Si cette interprétation est correcte, alors nous de-

Verbes morphophonologique				
	Fréquence cumulée haute		Fréquence cumulée basse	
	Fréquence de surface		Fréquence de surface	
	haute	basse	haute	basse
Exemple	sem-ée	sem-ais	sèm-e	sèm-erai
TR	631	660	626	662

TAB. 3.4 – Résumé des temps de réponses observés dans l’Expérience 2.

vrions obtenir un effet de la fréquence cumulée quand celle-ci est calculée sur la base des fréquences cumulées des deux racines possibles. Pour tester cette idée, nous avons calculé la fréquence cumulée totale (somme des fréquences cumulées de chacune des deux racines possibles) pour chacun des verbes composant notre matériel. Puis nous avons comparé les TR des verbes de fréquence cumulée totale élevée *vs.* basse²⁰. Analysés ainsi, les temps de réponse sur les formes verbales morphophonologiques indiquent un effet significatif de la *fréquence cumulée totale* par participant et par item [$F1(1,47) = 22.33, p < .001$; $F2(1,12) = 4.53, p = .05$] ainsi que de la fréquence de surface [$F1(1,47) = 32.39, p < .001$; $F2(1,12) = 8.52, p = .01$]. L’interaction n’est pas significative [$F1(1,47) = 2.85, p = .097$; $F2 < 1$]. Cependant, nous avons tout de même analysé l’effet d’amorçage en fonction de la fréquence de surface pour voir si nous observerions la même tendance que celle observée avec les verbes réguliers. Nous observons un effet significatif de la fréquence cumulée par sujet et par item pour les formes verbales de fréquence de surface élevée, avec une différence de 43 ms entre les TR pour les formes de fréquence cumulée totale élevée et les formes de fréquence cumulée totale basse : [$F1(1,47) = 21.21, p < .001$; $F2(1,12) = 8.72, p = .012$]. En revanche, pour les formes verbales de fréquence de surface basse

²⁰Cette méthode de calcul de la fréquence cumulée revient à celle utilisée dans l’expérience avec les verbes réguliers.

l'effet de la fréquence cumulée totale est significatif seulement par participant [$F1(1,47) = 4.89, p = .031; F2(1,12) = 1.15, n.s.$]. La différence des TR entre les formes de haute et de basse fréquence cumulée totale est de 21 ms. Ainsi, les temps de reconnaissance des formes verbales morphophonologiques montrent un effet de la fréquence de surface et un effet de la fréquence cumulée totale. Selon les hypothèses que nous suivons, l'effet de la fréquence cumulée totale montre qu'une racine phonologiquement abstraite est activée au cours du traitement lexical des verbes morphophonologiques, et qu'il y a donc eu décomposition. Dans la dernière expérience, nous nous penchons sur le cas des verbes irréguliers. Ces verbes, dans leurs formes conjuguées, ont au-moins deux racines possibles qui alternent de manière idiosyncratique. Dans une expérience d'amorçage visuel masqué en inter modalité (l'amorce est présentée auditivement et la cible visuellement) réalisée en français par Meunier & Marslen-Wilson (2004), les formes fléchies irrégulières amorcent leur forme infinitive (*buvais* amorce *boire*), tout comme les formes régulières du même verbe (*boirons* amorce *boire*). Dans l'expérience qui suit nous testons l'existence d'une représentation commune et phonologiquement abstraite pour ce type de verbes.

3.6 Expérience 3 : Verbes irréguliers

Dans cette troisième expérience nous testons la reconnaissance des formes verbales irrégulières afin de connaître la nature de la représentation des racines de cette classe verbale mais aussi la nature des processus qui leur sont appliqués. De la même manière que dans l'expérience précédente, nous avons calculé la fréquence cumulée de chacune des deux racines possibles d'un verbe irrégulier. Nous comparons de nouveau les formes verbales de fréquence cumulée élevée *vs.* basse (par ex. *buv-* qui a une fréquence cumulée de 28.1 occurrences par million et *boiv-* une fréquence cumulée de 4.6 occurrences par million) ainsi que les formes de fréquence de surface, soit élevée (*buvai-* - *boivent* qui ont des fréquences de surface de 3.48 et 3.61 occurrences par million respectivement) soit basse (*buviez* - *boives* qui ont des fréquences de surface de 0.13 et 0.1 occurrences par

million respectivement). Les formes verbales de quinze verbes différents sont sélectionnées en quadruplet en fonction de ces facteurs.

3.6.1 Méthode

Les participants et la procédure sont identiques à ceux des expériences précédentes.

Stimuli

Dans cette expérience, 15 verbes ont été utilisés pour lesquels nous avons sélectionné 4 formes verbales :

1. une forme verbale de fréquence cumulée et de fréquence de surface élevée
2. une forme verbale de fréquence cumulée élevée et de fréquence de surface basse
3. une forme verbale de fréquence cumulée basse et de fréquence de surface élevée
4. une forme verbale de fréquence cumulée et de fréquence de surface basse.

Voir le Tableau 3.5 pour des exemples.

Verbes irréguliers				
Fréquence cumulée haute			Fréquence cumulée basse	
Fréquence de surface			Fréquence de surface	
haute basse			haute basse	
Exemple	buv-aient	buv-iez	boiv-ent	boiv-es

TAB. 3.5 – Exemple de stimuli expérimentaux utilisés dans l'Expérience 3.

La différence de logarithme de la fréquence des formes verbales qui sont comparées (formes verbales de fréquence cumulée élevée *vs.* basse ou de fréquence de surface élevée *vs.* basse) est supérieure à 1.

3.6.2 Résultats

Les temps de réponses déviants (inférieurs à 300 ms et supérieurs à 1200 ms) sont exclus des analyses statistiques. Ainsi nous avons retiré 2.6 % des temps de réponse. Par ailleurs, 2.3 % d'erreurs ont été relevés. Etant donné ce taux peu élevé nous ne réalisons pas d'analyse statique sur les taux d'erreurs. Un résumé des résultats est présenté dans le Tableau 3.6 et la Figure 3.6.

Temps de réponse des sujets pour les formes verbales irrégulières, en fonction de leur fréquence cumulée et de leur fréquence de surface.

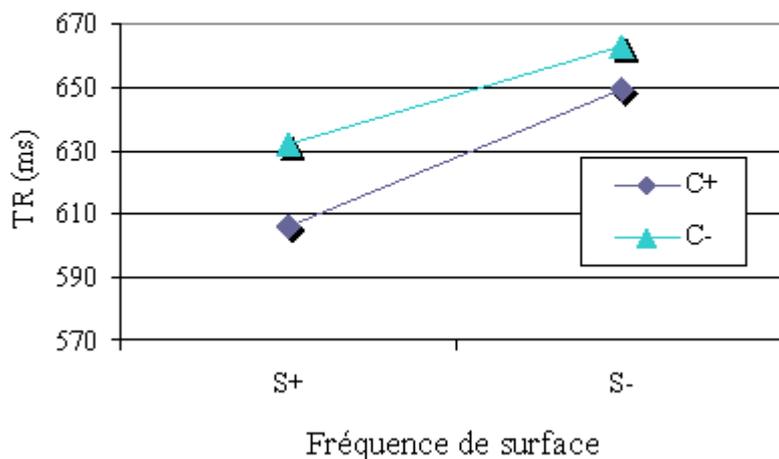


FIG. 3.6 – Graphique représentant les temps de réponse observés dans l'Expérience 3.

Les résultats montrent un effet de la fréquence de surface sur les temps de réponse [$F1(1,47) = 54.41, p < .001, F2(1,28) = 12.22, p = .001$]. Les formes verbales de fréquence de surface élevée sont reconnues plus rapidement que celles de fréquence de surface basse. Nous observons aussi un effet de la fréquence cumulée, significatif par participant seulement [$F1(1,47) = 14.51, p < .001; F2(1,28) = 1.47, n.s.$]. Les formes de fréquence cumulée élevée sont reconnues plus rapidement que celles de fréquence cumulée faible. Il n'y a pas d'interaction entre les deux facteurs [$F's < 1$].

Verbes irréguliers				
	<u>Fréquence cumulée haute</u>		<u>Fréquence cumulée basse</u>	
	Fréquence de surface		Fréquence de surface	
	haute	basse	haute	basse
Exemple	buv-aient	buv-iez	boiv-ent	boiv-es
TR	600	643	623	657

TAB. 3.6 – Résumé des temps de réponse observés dans l’Expérience 3.

3.6.3 Discussion

Les analyses statistiques montrent que la fréquence de surface et la fréquence cumulée influencent les temps de reconnaissance des formes verbales irrégulières en français. L’effet de la fréquence cumulée est significatif par participant. Ce résultat est important, car il suggère que les deux représentations distinctes de la racine (canonique et idiosyncratique) sont activées au cours du traitement lexical des verbes irréguliers. Il est intéressant de noter qu’à nouveau l’effet de la fréquence cumulée est numériquement plus faible pour les fréquences de surface basses (14 ms) que pour les fréquences de surface élevées (23 ms ; voir la Figure 3). Les résultats obtenus dans cette expérience sont très proches de ceux observés dans l’Expérience 1 avec les verbes réguliers et de ceux de l’Expérience 2 avec les verbes morphophonologiques (dans l’analyse des temps de réponse en fonction de la fréquence cumulée totale). Afin de comparer plus précisément ces trois expériences, nous avons réalisé une méta-analyse incluant les données des trois expériences (les stimuli de l’Expérience 2 restant organisés en fonction de la fréquence cumulée totale). Nous testons donc trois facteurs, le facteur expérience avec trois niveaux (verbes réguliers, morphophonologiques et irréguliers), le facteur fréquence de surface à deux variables (élevée *vs.* basse) et le facteur fréquence cumulée à deux variables (élevée *vs.* basse). Globalement, l’analyse ANOVA montre un effet du type d’expérience significatif par item [$F2(2,40) = 7.64, p = .001$] et une tendance par participant [$F1(2, 141) = 2.64, p = .07$], un effet

de la fréquence de surface [$F1(1,141) = 125.86, p < .001$; $F2(1,40) = 36.01, p < .001$], et un effet de la fréquence cumulée [$F1(1,141) = 51.55, p < .001$; $F2(1,40) = 9.81, p = .003$]. L'interaction entre la fréquence de surface et la fréquence cumulée est significative par participant uniquement [$F1(1,141) = 6.85, p < .01$; $F2(1,40) = 1.20, n.s.$]. La triple interaction n'est pas significative ($F_s < 1$). Une analyse spécifique sur l'effet de la fréquence cumulée montre que celui-ci est très significatif pour les formes verbales de fréquence de surface élevée [$F1(1,141) = 48.59, p < .001$; $F2(1,40) = 16.36, p < .001$], avec en moyenne un avantage de 40 ms pour les formes de fréquence cumulée élevée. Pour les formes verbales de fréquence de surface faible, nous observons un effet de la fréquence cumulée significatif par participant et une tendance par item [$F1(1,141) = 8.84, p < .01$; $F2(1,40) = 2.93, p = .09$]. En moyenne on observe une différence de 13 ms entre les temps de réponses pour les formes de fréquence cumulée élevée et basse. Cette méta-analyse indique que l'effet de la fréquence cumulée est plus robuste pour les formes verbales de fréquence de surface élevée que celles de fréquence de surface basse. Nous discuterons ce point dans la discussion générale.

3.7 Discussion générale

En français, toutes les formes fléchies sont composées d'une racine et d'un suffixe. Pour les verbes réguliers, il existe une seule racine possible pour toutes les formes fléchies. Pour les verbes morphophonologiques, deux racines coexistent et alternent dans leur forme fléchie selon une règle morphophonologique. Concernant les verbes irréguliers, au-moins deux racines coexistent et alternent de manière idiosyncrasique. Pour ces trois classes de verbes, le suffixe est toujours régulier. Dans notre étude, nous avons testé les temps de reconnaissance des formes verbales pour ces trois classes de verbe. Nous avons manipulé la fréquence de surface et la fréquence cumulée de ces formes verbales afin de déterminer par quel processus nous accédons à la représentation de ces mots polymorphémiques et si cela dépend de la nature régulière ou irrégulière de la racine. Les résultats indiquent :

- Un effet de la fréquence de surface sur la reconnaissance de formes verbales pour les trois classes verbales.

- Un effet de la fréquence cumulée sur la reconnaissance des formes verbales régulières et irrégulières.

Concernant la reconnaissance des formes verbales morphophonologiques, les résultats sont plus spécifiques. Nous n'avons pas observé d'effet de la fréquence cumulée quand celle-ci est calculée sur la base des deux racines possibles. Cependant, en calculant la fréquence cumulée totale de chaque verbe, (fréquences cumulées de la forme canonique et de la forme allomorphique) et que nous séparons ensuite les verbes en deux catégories - les formes verbales avec une fréquence cumulée totale haute d'un côté et celles avec une fréquence cumulée totale basse de l'autre- alors nous observons un effet de fréquence cumulée sur les temps de réponse.

Ces résultats nous apportent des indications quant au traitement des formes fléchies et sur la représentation des racines régulières et irrégulières dans le lexique mental.

L'effet de la fréquence cumulée

Nous devons noter que l'effet de fréquence cumulée observé dans ces expériences est robuste dans l'analyse par sujet alors qu'il reste faible dans l'analyse par item. Cette situation pourrait être due à un manque de puissance dans les analyses statistiques par item. En effet, les contraintes imposées par la manipulation de la fréquence de surface et de la fréquence cumulée pour un verbe donné nous obligent à tester uniquement 15 formes verbales dans chacune des conditions. Néanmoins, la méta-analyse nous a permis d'analyser un plus grand nombre de données et montre un effet de la fréquence cumulée très significatif dans l'analyse par sujet et par item. Par ailleurs, nous devons noter que l'effet de la fréquence cumulée observé est toujours plus faible pour les formes fléchies de

fréquence de surface basse. Nous décidons sur cet état de fait de mettre ces résultats en rapport avec la discussion engagée dans l'article de Taft publié en 2004. Dans cet article, il mentionne que l'effet de la fréquence cumulée peut être masqué avec une condition distractive. Pour cette condition, les non-mots sont composés d'une racine et d'un suffixe dont la combinaison est illégale (par ex. *sportation*). Face à cette condition distractive, la procédure de recombinaison en racine plus suffixe (postulée dans le modèle de Taft) est obligatoire pour que les sujets fassent la tâche de décision lexicale. Pour décider si un item est un vrai mot ou non, les participants sont obligés de traiter les suites de lettres présentées dans leur globalité, ce qui ajoute plus d'importance à la procédure de recombinaison (et donc l'effet de fréquence de surface) qu'à la procédure de décomposition. En effet, si plus d'importance est attribuée à la procédure de recombinaison, alors cela désavantage l'observation d'un effet de la décomposition (révélé par l'effet de la fréquence cumulée). Par contre, dans le cas où les non-mots sont tous composés d'une racine n'existant pas dans la langue étudiée et d'un suffixe existant (par ex. *plaurifier*), les participants peuvent suivre la stratégie de ne repérer que la racine pour réaliser la tâche de décision lexicale, en répondant systématiquement "mot", quand la racine existe et "non mot" quand la racine n'existe pas. Dans notre étude, la condition distractive était constituée de non-mots composés d'une racine non existante adjointe à un vrai suffixe du français. Ainsi, nous nous mettions dans les meilleures conditions pour observer un effet de la fréquence cumulée, chose que l'on n'observe pourtant pas de manière tranchée pour les formes verbales de fréquence de surface basse. Cependant, si nous reprenons la logique proposée par Taft (2004), il est tout à fait possible que l'effet de fréquence cumulée soit moins important pour les formes verbales de basse fréquence, car elles correspondent généralement à des formes verbales peu familières. Ainsi, plus de temps serait nécessaire à la procédure de recombinaison permettant de déterminer si la forme globale existe en français ou non. Si plus d'importance est attribuée à la procédure de recombinaison, cela favorise l'observation d'un effet de la fréquence de surface et minimise les chances d'observer un effet de la fréquence cumulée.

L'effet de la fréquence cumulée et les verbes morphophonologiques

Concernant les verbes morphophonologiques, nous n'avons pas observé d'effet de la fréquence cumulée en comparant les temps de reconnaissance entre deux formes verbales composées chacune d'une racine allomorphique différente. Cependant, un effet de la fréquence cumulée totale (i.e., la somme des fréquences cumulées des deux formes allomorphiques pour un verbe) a été observé dans une analyse *post-hoc* sur les temps de reconnaissance des formes fléchies de ce groupe de verbes. Les formes verbales de fréquence cumulée totale élevée sont reconnues plus rapidement que celles de fréquence cumulée basse. L'interprétation de cet effet *post-hoc* va dans le sens de l'hypothèse selon laquelle une représentation abstraite et unique est activée au cours du traitement de différentes formes polymorphémiques composées d'une base allomorphique (racine ou radical). En d'autres termes la même représentation est activée indépendamment des changements allomorphiques que peut subir une forme verbale. Cette interprétation est en accord avec les résultats de Marslen-Wilson et Zhou (1999), qui ont utilisé une tâche de décision lexicale avec un paradigme d'amorçage inter-modal. Ils ont observé qu'une forme dérivée allomorphique comme *sanity* ([saenəti]) produit un effet d'amorçage sur sa racine *sane* (prononcé [sem]). Cet effet montre que malgré la différence de réalisation phonologique des mots dérivés, une même représentation de la racine est activée deux fois : une fois au cours du traitement de l'amorce, et une fois au cours du traitement de la racine cible. De même, Meunier et Marslen-Wilson (2004) ont observé un effet d'amorçage entre une forme verbale morphophonologique composée d'une racine allomorphique et sa racine canonique.

L'effet de la fréquence cumulée et les verbes irréguliers

L'effet de la fréquence cumulée observé pour la classe des verbes irréguliers montre que les racines irrégulières sont activées au cours de l'identification des formes verbales irrégulières. Nous montrons ainsi que les racines canoniques et les racines irrégulières ne partagent pas de représentation commune au sein du lexique mental. Ceci ne permet pas d'expliquer l'effet d'amorçage entre *buvais* et *boire* observé avec un paradigme d'amorçage

masqué et non masqué (Meunier et Marslen-Wilson, 2004). Il est important de noter que la différence d'amorçage non significative observée dans cette étude entre deux formes alternatives d'un verbe irrégulier (une forme canonique et une forme irrégulière) sur la reconnaissance du verbe à l'infinitif est de 20 ms. L'absence de significativité peut être due à une variabilité ou une erreur de type 1. Une autre possibilité pourrait être que les effets d'amorçage entre des formes verbales et leurs racines ne soient pas aussi importants en anglais qu'en français. En effet, l'observation répétée d'un effet d'amorçage entre une forme verbale régulière et sa racine et l'absence de cet effet dans le cas des verbes irréguliers a été reprise abondamment dans la littérature pour soutenir l'hypothèse de la double voie (par ex. Kempley & Morton, 1982 ; Napps, 1989 ; Stanners, Neiser, Herson, & Hall, 1979). Cependant, et comme nous l'avons déjà présenté dans l'introduction de ce chapitre, Allen et Badecker (2002) ont montré un effet d'amorçage entre une forme verbale irrégulière et sa racine lorsque celles-ci partagent peu de recouvrement orthographique (par exemple *taught* amorce *teach*) en comparaison d'une paire amorce-cible qui partage un fort recouvrement (*gave* n'amorce pas *give*). Les auteurs rendent compte de ces résultats dans un modèle à deux niveaux de représentation, dans lequel les verbes irréguliers ont des représentations distinctes au niveau des représentations formelles (niveau des *lexemes*), mais une représentation commune au niveau des représentations abstraites (niveau des *lemma*). Au vu de nos résultats, qui indiquent que les racines sont représentées au sein du lexique mental, l'effet d'amorçage entre une forme irrégulière et le verbe correspondant à l'infinitif pourrait s'expliquer par la théorie de Allen et Badecker (2002), c'est-à-dire comme l'amorçage de *teach* par *taught*). La forme verbale irrégulière active la représentation de sa racine au niveau du *lexeme*, qui elle-même active la représentation commune et abstraite au niveau du *lemma*. L'observation d'un effet d'amorçage entre la forme verbale irrégulière et sa forme infinitive rendrait compte de l'activation des représentations stockées au niveau du *lemma* par les représentations stockées au niveau du *lexeme* (voir Figure 3.7). Néanmoins, le fait que l'effet d'amorçage observé en français pour les verbes irréguliers soit observé dans un paradigme d'amorçage masqué (présentation de l'amorce pendant 47

ms) pose une limite à cette interprétation. Ce temps de présentation de l'amorce semble en effet insuffisant pour que la représentation au niveau du *lemma* ait pu être activée ²¹. Nous reviendrons sur ce modèle dans le Chapitre 6.

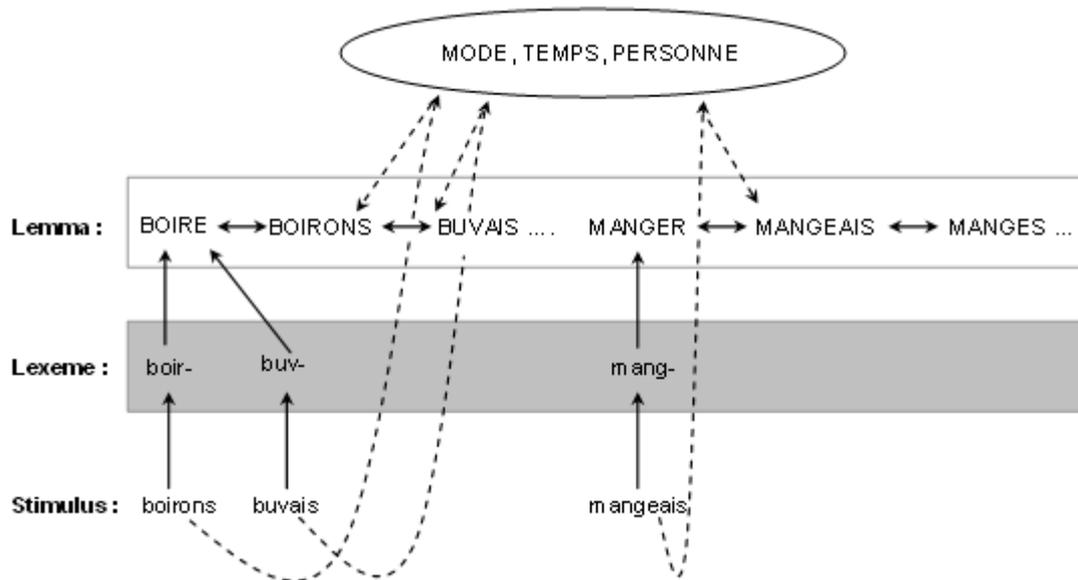


FIG. 3.7 – Schéma des représentations des verbes réguliers et irréguliers en français selon le modèle à deux niveaux d'Allen et Badecker (2002).

Une proposition récente envisagée pour rendre compte des effets d'amorçage entre les formes verbales irrégulières et leurs racines dans la langue anglaise, est proposée par Stockall et Marantz (2006). Ces auteurs postulent que l'accès lexical des formes verbales repose sur une procédure unique de décomposition et donc un accès par la racine. La distinction entre formes régulières et irrégulières n'intervient pas. Tous les mots morphologi-

²¹Néanmoins, il est probable que nous connaissons insuffisamment les temps de présentation nécessaires à la propagation d'activation. Si les études comportementales ne montrent pas d'effet sémantique pour des temps de présentation très courts de l'amorce (comme 47 ms), cela ne veut pas dire qu'un peu d'activation ne s'est pas propagé jusqu'à ces représentations (encodées au niveau du *lemma* dans les modèles à deux niveaux : *lexeme/lemma*) au sein du lexique mental. Peut-être que l'activation de ces représentations est réalisée, mais ne rend pas forcément disponibles de manière consciente les informations qui leur sont associées).

quement complexes, et pas seulement dans leur forme de surface (comme le verbe *taught* ‘enseigner’ au passé), activent leur racine indépendamment de la variation phonologique de l’allomorphe. Rappelons que selon ces auteurs, les formes verbales telles que *taught* ne sont pas idiosyncrasiques, mais composées du suffixe *-t* et de la variation allomorphique de *teach* en *taugh-*, comme les verbes *bring*, *buy*, *catch*... qui deviennent “ought” : *brought*, *bought*, *caught* au passé. D’autres groupes de verbes peuvent être construits en fonction de la variation que prend leur racine au passé. Par exemple *sleep*, *keep*, *sweep* possèdent une forme au passé “ept” : *slept*, *kept*, *swept*. Dans un tel modèle, lors du traitement d’une forme verbale comme *taught*, la représentation de sa racine *teach* est activée ainsi que l’information fonctionnelle “forme verbale au passé”. Pour que la suite de lettres *taught* active la représentation de *teach*, il faut que la forme sonore “ought” soit reconnue comme le résultat en sortie d’un mécanisme qui s’opère sur la séquence “each” (voir Figure 3.8).

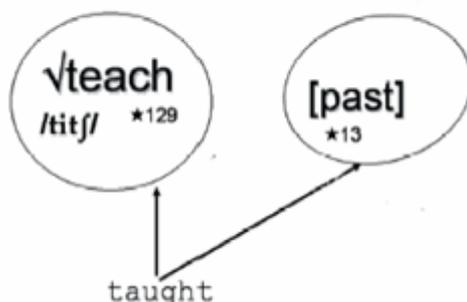


FIG. 3.8 – Schéma représentant le traitement de la forme passé *taught*, qui active la racine *teach* et le morphème fonctionnel [PAST]. * indique une règle morphophonologique spécifique. Par exemple, la règle 129 permet de générer la forme irrégulière *taught*.

Afin d’argumenter leur point de vue, les auteurs proposent une expérience de décision lexicale avec un paradigme d’amorçage non masqué couplée à une acquisition magnéto-encéphalographique. Les formes verbales utilisées en amorce sont irrégulières et possèdent soit un fort recouvrement orthographique avec leur racine-cible (par ex. *gave-give*) soit un faible recouvrement (*taught-teach*). Les temps de réponse pour la décision lexicale

sur le mot cible ainsi que les mesures encéphalographiques sont enregistrés. La magnétoencéphalographie (MEG) fonctionnelle permet d'obtenir des informations sur l'activité cérébrale. Elle mesure les champs magnétiques induits par une tâche cognitive donnée. En reconnaissance visuelle de mots, trois champs magnétiques sont généralement évoqués, notamment la M350 (qui apparaît 350 ms après la présentation de l'item lexical) (Embick, Hackl, Schaeffer, Kelepir & Marantz (2001); Pykkänen, Stringfellow, Flagg & Marantz (2000). Par ailleurs, deux études réalisées par Pykkänen, Stringfellow & Marantz (2002) et Stockall, Stringfellow & Marantz (2004), ont montré que des facteurs liés à des processus cognitifs précoces comme la probabilité d'apparition de bigrammes, affectent le décours temporel du champ magnétique M350, alors que ce n'est pas le cas pour des facteurs post-lexicaux comme la densité des voisins phonologiques (voir Stockall et al., 2004 pour plus de détails). Dans l'expérience de Stockall et Marantz (2006), les résultats montrent une dissociation pour la condition orthographique et la condition irrégulière concernant les mesures de la MEG et les temps de réponse. Dans la condition orthographique, les latences de la M350 ne sont pas significativement différentes lorsque l'amorce et la cible partagent un lien par rapport à la condition contrôle non liée, alors que les temps de réponse le sont (amorçage inhibiteur quand l'amorce et la cible sont liées). Dans la condition irrégulière on observe l'inverse. Les latences de la M350 sont significativement plus rapides dans la condition liée vis-à-vis de la condition non liée, alors que les temps de réponse ne le sont pas. Cela indique, d'une part que la M350 n'est pas sensible aux effets de compétition lexicale (effet post-lexical) entre items orthographiquement proches alors que les temps de réponse le sont. D'autre part, l'effet d'amorçage observé avec la M350 et non observé sur les temps de réponse pour la condition où l'amorce est une forme verbale irrégulière avec peu de recouvrement orthographique avec sa racine-cible, est interprété de cette manière : Les formes verbales irrégulières (*taught*) activent leur racine (*teach*) au cours de l'accès lexical (effet d'amorçage). Par la suite, la forme au passé du verbe traité (*teach*) est générée par l'application de règles morphologiques spécifiques. Cela conduit à un effet de compétition (amorçage inhibiteur). Une limite que l'on peut poser à cette proposition

tient au fait qu'elle est difficilement applicable aux formes verbales irrégulières en français. Comme le montre notre étude, les racines verbales du français sont représentées au sein du lexique mental. L'accès au sens d'une forme verbale irrégulière *buvait*, par exemple, ne semble pas activer la représentation de la racine *boir-* par une règle morphologique et l'information "verbe au passé", mais bien la représentation de *buw-* elle-même.

3.8 Conclusion

Nous avons proposé dans ce chapitre une étude sur les processus d'accès lexical aux formes verbales régulières, morphophonologiques et irrégulières. Nous avons observé que la fréquence de surface et cumulée de ces formes influencent leur temps de reconnaissance. Ce résultat montre que les racines des formes verbales en français ont une représentation au sein du lexique mental, et cela, quel que soit le degré de régularité de la racine. En confrontant nos données à des résultats d'amorçage en français, entre des formes verbales irrégulières et leurs formes infinitives respectives, on s'aperçoit qu'elles sont en faveur d'un modèle à deux niveaux de représentation. L'effet d'amorçage ne s'expliquerait donc pas (en tout cas pour le cas des verbes irréguliers) par l'activation double de la racine des formes verbales irrégulières (une première fois au cours du traitement de l'amorce et une deuxième fois au cours du traitement de la cible) car les formes irrégulières en français contiennent une racine orthographiquement dissimilaire de sa forme infinitive. L'effet d'amorçage s'expliquerait plutôt par l'activation de la représentation même de la racine irrégulière au niveau du *lexeme*, qui propagerait de l'activation aux représentations stockées au niveau du *lemma*, ces représentations étant communes pour la racine irrégulière et canonique issues du même verbe. Par ailleurs cette étude met en exergue l'importance des études translinguistiques, en montrant les limites de certains modèles d'accès lexical par rapport aux formes fléchies régulières et irrégulières lorsque l'on tente de les appliquer à des langues dont les caractéristiques sont différentes des langues pour lesquelles ces modèles ont été conçus. Les racines des formes verbales régulières et irrégulières sont

donc représentées au sein du lexique mental. Cela pourrait-il être lié à la co-occurrence de ces racines dans plusieurs mots polymorphémiques d'une même famille (*buvais*, *buvette*, *buvard*; *manger*, *mangeable*, *mangeoire*...)? Si on imagine que cela est le cas, alors peut-être qu'à la lecture de ces mots, ces unités infra-lexicales sont repérées et extraites du mot dans le but d'entamer une procédure de recherche lexicale précoce sur les représentations contenant cette racine et susceptibles de correspondre au mot. Cependant, nous devons garder en tête que les suffixes sont eux aussi des unités pouvant être retrouvées dans plusieurs mots polymorphémiques (*mangeable*, *buvable*), et que ces unités sont régulières et pourvues de sens. Ces suffixes eux aussi sont susceptibles d'être repérés au cours de la lecture de ces mots pour en être extraits et activer le sens qui va modifier celui de la racine à laquelle ils sont adjoints. Dans le chapitre qui suit nous abordons la question du déclenchement de la procédure de décomposition, par le biais d'une expérience originale qui manipule les caractéristiques linguistiques de certains mots de la langue française.

