Lexis

Journal in English Lexicology

18 | 2021

Lexical learning and teaching

Inférence et apprentissage du vocabulaire : effet de la lexicalisation en première langue et de la tâche

Enhao Léger-Zheng et Olga Théophanous

Édition électronique

URL : <https://journals.openedition.org/lexis/5937>

DOI : [10.4000/lexis.5937](https://doi.org/10.4000/lexis.5937)

ISSN : 1951-6215

Éditeur

Université Jean Moulin - Lyon 3

Référence électronique

Enhao Léger-Zheng et Olga Théophanous, « Inférence et apprentissage du vocabulaire : effet de la lexicalisation en première langue et de la tâche », *Lexis* [En ligne], 18 | 2021, mis en ligne le 18 décembre 2021, consulté le 21 décembre 2021. URL : <http://journals.openedition.org/lexis/5937> ; DOI : <https://doi.org/10.4000/lexis.5937>

Ce document a été généré automatiquement le 21 décembre 2021.

Lexis is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Inférence et apprentissage du vocabulaire : effet de la lexicalisation en première langue et de la tâche

Enhao Léger-Zheng et Olga Théophanous

1. Inférence lexicale et compréhension en lecture

- 1 L'inférence lexicale, c'est-à-dire la capacité de déduire le sens des mots inconnus en contexte est une stratégie cognitive fondamentale de la compréhension en lecture (Cornaire [1999] ; Fraser [1999] ; Paribakht & Wesche [1999]). Étant donné qu'il est impossible de connaître tous les mots d'une langue donnée, l'inférence lexicale fait en effet partie intégrante du processus de la lecture. Grâce à elle, le lecteur peut pallier ses lacunes lexicales et poursuivre sa lecture. La mise en œuvre de l'inférence soutient aussi psychologiquement le lecteur qui ne laisse pas tomber sa lecture au premier mot inconnu rencontré. Lire peut s'avérer anxiogène et ce, encore plus dans une langue seconde ou étrangère (L2) (Cornaire [1999]). Il est alors important de comprendre le fonctionnement de cette stratégie et les facteurs qui concourent à sa réussite. D'autant plus que si elle paraît assez « naturelle » et « intuitive » chez les natifs, elle est plus compliquée à mener en L2.
- 2 De nombreuses recherches examinent la façon dont les apprenants d'une L2 tentent de déduire le sens des mots en contexte (Haastrup [1991] ; Fraser [1999] ; Nassaji [2003], [2004] ; Wesche & Paribakht [2009]). Ces études révèlent que les apprenants utilisent diverses sources de connaissance pendant le processus de l'inférence : les connaissances intra et interlinguistiques (comme la morphologie lexicale, l'association de mots, l'étymologie, les emprunts, la ponctuation, les articulateurs logiques, la syntaxe, etc.), ainsi que les connaissances générales de toute sorte (Haastrup [1991] ; Paribakht & Wesche [1999] ; Bengelil & Paribakht [2004] ; Paribakht [2005] ; Wesche &

Paribakht [2009]). La recherche pointe aussi diverses techniques et procédures auxquelles les lecteurs recourent afin d'inférer le sens des mots inconnus comme la décomposition du mot en ses parties constituantes, la répétition à voix haute du mot inconnu, la régression pendant la lecture, l'introspection, l'analogie, la comparaison, le contrôle, l'évaluation (Walker [1983] ; Nassaji [2003], [2004] ; Théophanous & Dreyfus [2004]). La recherche révèle également un certain nombre de facteurs qui participent à la réussite de l'inférence. Parmi ces facteurs, mentionnons l'importance du mot cible dans la compréhension du texte (Brown [1993]), la capacité de l'apprenant de faire une utilisation effective des indices extralinguistiques (Haastrup [1991] ; De Bot, Paribakht & Wesche [1997]), l'idée que l'apprenant se fait du sens possible du mot inconnu (Frantzen [2003]), ses compétences en lecture (Haastrup [1991] ; Bengelil & Paribakht [2004]), l'étendue de son vocabulaire réceptif (Coady et al. [1993] ; Laufer [1997]), la qualité de ses connaissances lexicales (Nassaji [2004]) et sa familiarité avec le thème traité dans le texte (Pulido [2007]).

- 3 Nous savons peu de choses sur l'influence qu'exerce précisément la première langue (L1) dans le processus de l'inférence. Ainsi, bien qu'il soit généralement accepté que les connaissances acquises en L1 restent actives pendant le traitement d'informations lexicales en L2 (Jiang [2002] ; Kroll et al. [2010]), il n'est pas clair de quelle façon la L1 du lecteur contribue à l'utilisation de divers indices lors de l'inférence lexicale en L2 et à la réussite de celle-ci dans l'attribution du sens du mot inconnu. Certaines études sur les influences translinguistiques ont examiné l'effet que les différences orthographiques L1-L2 peuvent avoir sur le traitement lexical et les stratégies de lecture (Koda [1989], [2005] ; Suarez & Meara [1989] ; Ryan & Meara [1991] ; Jarvis [2000]). D'autres ont mis en avant l'effet de la présence et de la reconnaissance des mots transparents ou congénères interlinguaux (Tréville [1996] ; Théophanous [2001a] ; Hilton [2003]) et l'influence de la syntaxe de la L1 sur l'inférence en L2 (Nagy et al. [1997]).
- 4 Ceci dit, la recherche sur l'effet de la lexicalisation en L1 dans le processus de la lecture et dans l'inférence lexicale en L2 est rare. Par lexicalisation on entend le fait qu'un mot de la L2 ait un équivalent attesté (une unité lexicale simple ou une unité polylexicale préconstruite, cf. Paribakht [2010 : 62]) dans la L1 de l'apprenant. Par opposition, la non-lexicalisation renvoie à l'absence d'une unité lexicale équivalente attestée d'un mot de la L2 dans la L1 de l'apprenant.
- 5 Quand un mot est lexicalisé en L1, le concept correspondant à ce mot préexistant dans le lexique mental de l'apprenant facilite le lien forme-sens pendant la compréhension. En revanche, pour un mot non lexicalisé en L1, l'absence de concept disponible peut compliquer la mise en place du lien forme-sens de ce mot. L'inférence serait alors plus difficile à faire à cause de ce vide lexical en L1, ce qui peut créer aussi une charge cognitive supplémentaire chez le lecteur. Comme le dit Paribakht [2005 : 729], « (t)andis que l'inférence d'un mot lexicalisé en L1 peut être caractérisée comme un processus impliquant la reconnaissance d'un lemme, celui d'un mot non lexicalisé dans la langue maternelle implique un processus de la construction d'un lemme¹ ».
- 6 Dans une étude sur l'influence du farsi, la L1 d'un groupe d'apprenants iraniens, lors de l'inférence des mots anglais L2 qui leur étaient inconnus, Paribakht [2005] observe que les participants réussissent mieux l'inférence des mots lexicalisés pendant la lecture. Les mots non lexicalisés représentent ainsi un défi pour la compréhension et pourraient être une source de difficulté dans la compréhension en lecture et dans le développement subséquent du vocabulaire en L2. Dans une autre étude, Paribakht &

Tréville [2007], comparant les inférences de deux groupes linguistiques d'apprenants, francophones et persanophones, en anglais L2, observent que si les deux groupes linguistiques ont mieux réussi l'inférence des mots lexicalisés, le taux de réussite est plus élevé chez les francophones. La proximité linguistique L1-L2 (français-anglais) avait ainsi un effet positif sur la réussite des inférences lexicales (cf. aussi Wesche & Paribakht [2009]).

2. Inférence lexicale et apprentissage du vocabulaire

- 7 Inférer le sens des mots en contexte peut être le premier pas vers l'apprentissage d'un nouveau mot. Les natifs apprennent beaucoup de mots dans leur langue de cette façon, implicite et incidente, c'est-à-dire via la lecture et les interactions quotidiennes où l'objectif premier n'est pas l'apprentissage du vocabulaire. Cela ne se passe pas de la même façon pour les apprenants en L2 puisque ces derniers bénéficient rarement du luxe linguistique des rencontres multiples, variées et fréquentes avec les mots de la L2. En outre, en L2 la mise en œuvre à profit du contexte et de l'inférence avec les bénéfices qui peuvent en découler en termes d'apprentissage du vocabulaire n'est toujours pas claire et ce, malgré les nombreuses études sur ce sujet.
- 8 De nombreuses recherches ont examiné les limites du contexte et de l'inférence en contexte des mots inconnus dans l'apprentissage du vocabulaire en L2 (Hulstijn [1992] ; Hulstijn et al. [1996] ; Prince [1996] ; Laufer & Shmueli [1997] ; Nassaji [2003]). Tout d'abord, inférer en contexte n'est pas toujours possible et ce, même lorsque l'apprenant utilise tous les indices disponibles dans le texte (Nassaji [2003]). De plus, les apprenants ne prêtent pas nécessairement attention aux mots inconnus en contexte (Laufer [2003]). Par ailleurs, pour pouvoir tirer profit des indices contextuels et inférer correctement, le lecteur doit connaître la quasi-totalité des autres mots du texte, à savoir au moins 95% et idéalement entre 97%-98% des mots présents. Autrement dit, le lecteur peut se permettre de ne pas connaître dans le texte entre un mot sur vingt (au pire) et un mot sur cinquante (au mieux) (Hirsh & Nation [1992] ; Hu & Nation [2000]). La stratégie de l'inférence lexicale repose ainsi énormément sur le niveau des connaissances lexicales du lecteur et donc en bénéficient surtout les apprenants avancés connaissant déjà beaucoup de mots.
- 9 Une autre limite réside dans le fait que l'inférence peut aboutir à un sens erroné (cf. Kelly [1990] ; Hulstijn [1992]). Ainsi, les lecteurs font des inférences erronées sur la base de la morphologie lexicale, des faux-amis, des rapprochements entre formes lexicales similaires grapho-phonétiquement (Nassaji [2003]). Ils choisissent aussi le sens erroné des mots polysémiques ou échouent à identifier correctement la fonction grammaticale des mots. Les apprenants ont également tendance à se concentrer sur les caractéristiques du mot sans prêter attention au contexte (Haastrup [1989], [1991] ; Haynes [1993]) au point de déformer même sémantiquement celui-ci pour justifier le sens inféré (Bensoussan & Laufer [1984]).
- 10 Une autre limite tout aussi sérieuse de l'inférence en contexte réside dans la capacité de retenir en mémoire le vocabulaire inféré. Des recherches montrent que retenir les mots à l'issue de l'inférence n'est pas nécessairement automatique (Mondria & Wit-De Boer [1991] ; Parry [1993]). L'inférence, comblant un besoin ponctuel pendant la compréhension, ne semble pas fixer le mot dans la mémoire. Ainsi la recherche empirique en L2 montre que seulement un petit nombre (il peut varier entre 5-14%) de

mots L2 peut être appris d'une simple exposition au texte, sans qu'une pratique ou qu'une rencontre subséquente n'ait lieu (Hulstijn [1992] ; Hulstijn et al. [1996] ; Horst et al. [1998] ; Zahar et al. [2001]). Les gains lexicaux subséquents à l'inférence lexicale dépendent de nombreux paramètres, tels que la richesse informative du contexte (Li [1988]), la fréquence d'exposition au mot (Hulstijn et al. [1996] ; Waring & Takaki [2003]), le type de tâche à réaliser (Paribakht & Wesche [1997] ; Fraser [1999]) ou les caractéristiques du mot (Paribakht & Wesche [1997]). Prince [1996] trouve pour sa part que chez les apprenants, surtout chez les moins avancés, les nouveaux mots sont mieux retenus lorsqu'ils sont accompagnés de leur traduction en L1, que lorsqu'ils sont présentés en contexte L2. Après d'apprenants de l'espagnol L2, Pulido [2007] examine l'influence de la familiarité du thème du texte sur la rétention des mots cibles. Ses résultats montrent que, deux jours après la lecture, les apprenants retiennent mieux les mots cibles qui étaient présents dans un texte à thème familier plutôt que lorsque le thème était moins connu. Les résultats de Pulido pointent ainsi l'effet négatif que peut avoir la difficulté à conceptualiser le mot inféré sur la rétention de celui-ci.

- 11 Le portrait plutôt pessimiste de l'inférence lexicale sur l'apprentissage du vocabulaire de la L2 que nous venons de brosser ne devrait pourtant pas éclipser le bénéfice potentiel de cette stratégie. En effet, des études qui montrent un effet positif sur le développement du vocabulaire grâce à l'inférence existent aussi (cf. Qian [2003] ; Pigada & Schmitt [2006]). Par ailleurs, les apprenants eux-mêmes citent l'inférence lexicale comme une des cinq stratégies les plus utilisées et préférées lorsqu'ils lisent (cf. Schmitt [1997])². Une des clés de la réussite de l'inférence sur le long terme et en faveur de l'apprentissage du vocabulaire serait très probablement à chercher dans l'utilisation effective de cette stratégie conjointement avec d'autres.

3. L'étude

- 12 Après avoir constaté le manque d'études sur l'effet de la lexicalisation en L1 et son rôle dans l'inférence pendant la lecture et la rétention subséquente du vocabulaire inféré, nous avons entrepris une étude visant à comparer les démarches d'inférence lexicale adoptées par un groupe d'apprenants francophones du chinois L2 pendant qu'ils lisent des textes écrits en chinois et qu'ils sont confrontés à des mots inconnus. Notre étude vise ainsi un double objectif. D'une part, nous cherchons à cerner les difficultés spécifiques que pourraient avoir les apprenants francophones d'une L2 non alphabétique afin de déduire le sens correct des mots inconnus pendant la lecture et d'autre part, nous voulons examiner le développement du vocabulaire inféré à court et à long terme.
- 13 La stratégie d'inférence lexicale nous paraît d'autant plus importante en chinois L2 car non seulement le nombre des caractères à maîtriser est très élevé avant de pouvoir se lancer dans une quelconque lecture, mais aussi parce que les multiples distinctions graphiques différenciant les caractères entre eux demandent une très grande attention chez le lecteur L2. Ces distinctions graphiques en chinois dépassent le strict cadre de la différenciation entre formes lexicales similaires (*dépasser/dépenser*, *chemin/cheminée*) des langues alphabétiques (cf. synforms, Laufer [1991] ; confusibles lexicaux, Théophanous [2001b]). Une autre particularité de la langue chinoise à l'écrit consiste à aligner les caractères dans un espacement toujours égal entre eux. Cela ne permet pas d'identifier spatialement les mots qui sont formés d'un à plusieurs caractères et peut

littéralement provoquer un choc visuel chez le lecteur non expert en rendant ainsi difficile la mise en place de la stratégie de l'inférence lexicale. Enfin, l'acquisition de tel ou tel caractère en chinois L2 est le premier obstacle pour les apprenants de langues alphabétiques, car alors que la mémorisation d'un mot dans une langue alphabétique se fait en général par le biais de la phonétique, en chinois elle passe surtout par la voie visuelle. Enfin et surtout dans une L2 typologiquement lointaine de la L1 de l'apprenant, les probabilités pour que ce dernier rencontre des mots sans équivalent lexicalisé attesté dans sa L1 sont plus élevées.

- 14 À partir de notre question générale, à savoir comment un groupe de francophones étudiant le chinois L2 procède afin de déduire le sens des mots chinois qu'ils ne connaissent pas, quand ceux-ci sont lexicalisés dans leur L1 et quand ils ne le sont pas, nous formulons les trois questions suivantes :
- a) Lors de la lecture en chinois L2, les sources de connaissance utilisées pour inférer le sens des mots inconnus diffèrent-elles selon que ces mots sont lexicalisés ou non dans la L1 ?
 - b) Dans quelle mesure les apprenants parviennent-ils à déduire le sens des mots inconnus (lexicalisés ou non) présents dans les textes écrits en chinois ?
 - c) Quel est l'apport de l'inférence lexicale sur l'apprentissage subséquent du vocabulaire inféré ? Et est-ce qu'un investissement cognitif plus important de la part de l'apprenant garantit mieux cet apprentissage ?

4. Méthodologie

- 15 33 apprenants francophones ont participé à notre étude dont 25 femmes et 8 hommes. Au moment de l'expérience, les participants étaient âgés de 21,6 ans en moyenne. Parmi eux 32 étaient au sixième semestre de la Licence à l'Université de Toulouse II – Jean Jaurès (désormais UT2J) dont 27 en Langues Étrangères Appliquées (option anglais-chinois) et 5 en Langues, Littératures, Civilisations Étrangères en chinois. Un participant suivait un cours de chinois hebdomadaire dans une association depuis trois ans, ayant suivi préalablement des cours pendant trois ans au lycée et obtenu 20 sur 20 en chinois (Langue Vivante 3) au Baccalauréat. Les participants connaissaient entre 600 et 1200 mots en chinois au moment de l'expérience, ce qui correspond au niveau B1 du *Cadre Européen de Référence pour les Langues*, selon le barème du Bureau national pour l'enseignement du chinois langue étrangère³.
- 16 La démarche introspective via le protocole de la pensée à voix haute (*think-aloud*) a été utilisée pour la collecte des données. Ainsi chaque participant a été convoqué à une session individuelle où il s'est trouvé en tête-à-tête avec l'expérimentateur, natif du chinois parlant couramment le français. Chaque session individuelle, d'une durée approximative de deux heures, a consisté à observer les inférences lexicales du participant en lui demandant de décrire à haute voix ses raisonnements. Chaque participant a été prié de dire tout ce qu'il faisait ou pensait lorsqu'il essayait de déduire le sens des mots. Cette réflexion oralisée a eu lieu en français. L'expérimentateur incitait les participants à s'exprimer en s'abstenant toutefois de commenter leurs réponses. Toutes les sessions individuelles ont été enregistrées et transcrites ultérieurement à des fins d'analyse.
- 17 Afin de sélectionner les mots cibles, nous avons cherché dans la presse et dans le *Guide hiérarchique des caractères et des mots chinois* (汉语水平词汇与汉字等级大纲, 1992) qui

comprend 8828 mots chinois les plus utilisés dans les communications quotidiennes. Pour minimiser le risque de sélectionner des mots déjà connus par les participants, nous avons cherché parmi les mots les moins fréquents du Guide (Classe D comportant 3569 mots). Nous avons vérifié le sens des mots présélectionnés dans le Dictionnaire chinois-français (汉法词典, 2001) et puis vérifié auprès des enseignants du chinois à l'UT2J que ces mots n'avaient pas été enseignés ou n'allaient pas être enseignés pendant notre expérience.

- 18 Vingt mots chinois ont été testés, dix lexicalisés (L) en français et dix non lexicalisés (NL) en français. Parmi les mots lexicalisés, l'équivalent en français de quatre mots chinois : MC06 (雀跃, sauter de joie), MC08 (在乎, prendre quelque chose à cœur, faire attention à), MC14 (彩虹, arc-en-ciel) et MC16 (风筝, cerf-volant) est une séquence polylexicale figée, un mot composé ou une expression préfabriquée. Chacune de ces séquences polylexicales fait partie du lexique de la langue française. Deux enseignants francophones du chinois L2 ont aussi été consultés afin de valider notre sélection de mots cibles pour les catégories L et NL. Parmi eux, dix-sept sont constitués de deux caractères, un mot comporte un seul caractère, un mot trois caractères et un dernier quatre caractères. Les vingt mots cibles sont présentés dans le Tableau 1 dans l'ordre d'apparition dans les textes expérimentaux.

Tableau 1. Mots cibles

MC01	节俭	L	économe, frugal
MC02	窗花	NL	papier découpé pour la décoration d'une fenêtre
MC03	送灶	NL	saluer avec des offrandes le départ du Génie du foyer pour le Ciel, le 23 de la 12 ^e lune
MC04	腊月	NL	le douzième mois du calendrier lunaire
MC05	汤圆	NL	boulette de riz glutineux farcie de viande (fourrée à la confiture)
MC06	雀跃	L	sauter de joie
MC07	还礼	NL	faire un présent en retour, offrir un cadeau à son tour
MC08	在乎	L	prendre quelque chose à cœur, faire attention à
MC09	乏味	L	fade, insipide
MC10	盼望	L	espérer, désirer
MC11	安适	NL	tranquille et confortable
MC12	五颜六色	L	de diverses couleurs, multicolore
MC13	明丽	NL	clair et beau
MC14	彩虹	L	arc-en-ciel

MC15	踏青	NL	faire une excursion (se promener) à la campagne au printemps
MC16	风筝	L	cerf-volant
MC17	蝉	L	cigale
MC18	闲聊	L	bavarder
MC19	飘零	NL	tomber en voltigeant
MC20	白茫茫	NL	d'une blancheur éclatante et immaculée, une immensité argentée

- 19 Les mots cibles ont été regroupés par thème et nous avons ensuite rédigé deux textes, A et B. Chaque texte comporte cinq mots lexicalisés et cinq mots non lexicalisés. Chaque mot cible n'apparaît qu'une seule fois dans les textes. Le texte A relate des souvenirs d'enfance lors de la fête du printemps en Chine et le texte B donne une description d'un village pendant les quatre saisons. Les mots cibles représentent environ 5% de la totalité des mots présents dans les textes. Les autres mots présents dans les textes étaient considérés connus par les participants d'après leurs manuels, à l'exception d'un seul mot sans importance pour la compréhension globale du texte. Nous avons consulté l'avis des enseignants de chinois à l'UT2J et sollicité des apprenants de chinois du même niveau pour effectuer des pré-tests. Ces derniers nous ont confirmé que les deux textes étaient globalement compréhensibles pour un apprenant du niveau B1 en chinois L2. Chaque mot cible est intégré dans un contexte qui fournit des indices sur le sens général auquel renvoie ce mot. Par exemple, la phrase comportant MC14 (彩虹) « Parfois, à l'arrêt d'une pluie, on voit dans le ciel un(e) joli(e) MC14 » peut aider un apprenant à comprendre que MC14 signifie « arc-en-ciel ». Les deux textes contiennent 639 caractères (cf. le Tableau 2).

Tableau 2a. Textes utilisés (texte A)

我小时候，过年真的很热闹。那时候的经济不太好，大家平时都非常节俭，只有过年时才会花很多钱。新年前一个多月，大人们就开始忙着为过年买各种东西。很多人会用红纸做一些窗花，让家里更漂亮。新年前几天，每家都会做一些好吃的来送灶。孩子们从腊月第一天开始就每天想着新年快点儿到，因为过年时会有新衣服穿，还可以吃汤圆和很多别的好吃的。当然，过年时最让孩子们雀跃的事儿还是拿红包。很多孩子的红包最后不是自己的：父母没什么钱，别人给了孩子红包，孩子交给父母，父母再用这些钱去给其他人还礼。不过，孩子们拿到红包时都高兴极了，也不在乎钱最后是不是自己的。

现在日子好了，商店里的东西也多了，想什么时候买都可以，连发红包也能用手机。但大家都说过年越来越乏味了。以前盼望着过年穿新衣服、吃好吃的、拿红包，那是一种快乐，现在都没有了。

Tableau 2b. Textes utilisés (texte B)

我的家乡是一个人口不多的小城市，交通很方便，一年四季都非常美，是一个很安适的地方。家乡的春天很温暖，城里的树和草很早就换上了绿色的新衣，路边都开着五颜六色的小花儿。春天也是一个多雨的季节。每场春雨过后，小城都会变得更明丽。有时下完雨后，天上还会有漂亮的彩虹。每到天气好的时候，人们会到城市以外的地方去踏青。要是风比较大的话，还可以去小城旁边的山上放风筝。夏天很热，树上的蝉从早叫到晚。到了晚上，吃完了晚饭的人们都走出家门，大人们坐在一起闲聊，孩子们就在一边玩儿着，那是一天中最热闹的时候。一到秋天，天气就开始转凉，鸟儿们都飞去温暖的南方过冬。草和树叶慢慢地变黄，秋风刮过，树叶飘零。小城安静下来，等着冬天的到来。家乡的冬天常常下雪。大雪过后，城市白茫茫的，城里的草地上、树上、房子上和城外的山上都是雪，真是漂亮极了。

- 20 À la suite de la tâche de l'inférence, nous avons demandé aux participants de compléter une série de vingt phrases lacunaires. Pour cette tâche, nous avons divisé le groupe des participants de façon aléatoire en deux : le groupe I (17 participants) et le groupe II (16 participants). Les participants du groupe I devaient accomplir la tâche lacunaire à l'aide d'une liste des mots cibles fournie, alors que les participants du groupe II devaient chercher les sens des mots cibles dans le *Dictionnaire chinois-français* (汉法词典, 2001) mis à leur disposition avant d'accomplir la tâche lacunaire.
- 21 Le jour de la session individuelle, chaque participant s'est exercé à la méthode de la pensée à voix haute, avant la tâche de l'inférence à proprement parler. Nous lui avons ensuite demandé de lire le texte A (dans lequel les mots cibles n'étaient pas soulignés) silencieusement en le prévenant que quelques questions lui seraient posées ensuite. Après la lecture, l'expérimentateur demandait à chaque participant de déduire, l'un après l'autre, le sens des mots cibles qui cette fois-ci étaient soulignés dans le texte. La même procédure a été répétée pour le texte B. Par la suite, chaque participant a complété le test lacunaire.
- 22 La rétention du vocabulaire inféré a été testée à deux reprises, à court et à long terme, via l'échelle des connaissances lexicales (*Vocabulary Knowledge Scale VKS*, cf. encadré). Cet outil, développé par Paribakht & Wesche [1997], se basant sur les différentes étapes de l'acquisition du vocabulaire, demande à l'apprenant de dire à quel niveau se situe sa connaissance pour tel ou tel mot testé. Paribakht & Wesche déclinent ainsi la qualité des connaissances lexicales à des niveaux successifs, allant de la connaissance zéro jusqu'à l'emploi du mot testé dans une phrase. Cette mesure d'évaluation est souvent utilisée pour mesurer la « profondeur » des connaissances lexicales. Si composer une phrase contenant le mot testé (Niveau 6) dépasse le strict cadre de l'évaluation de la connaissance sémantique, les cinq premiers niveaux permettent d'obtenir une bonne idée de cette connaissance, étant donné que les apprenants montrent de quelle façon exactement ils comprennent (ou pas) le sens de chaque mot testé⁴. Dans notre recherche, cet instrument a été utilisé aux deux post tests (VKS2 et VKS3), soit deux et quatre semaines après la session d'inférence. Il a aussi été utilisé avant la session de l'inférence (VKS1) pour bien vérifier la non-connaissance des mots cibles.

- Niveau 1 : Je n'ai jamais vu ce mot.
- Niveau 2 : Je pense avoir déjà vu ce mot, mais je ne sais pas ce qu'il veut dire.
- Niveau 3 : J'ai déjà vu ce mot, mais je ne sais pas ce qu'il veut dire.
- Niveau 4 : J'ai déjà vu ce mot. Il veut peut-être dire ..., mais je n'en suis pas sûr(e).
- Niveau 5 : Je connais ce mot et je sais qu'il veut dire...

- Niveau 6 : Je connais ce mot et je sais qu'il veut dire... Je sais écrire une phrase en chinois avec ce mot.

5. Présentation des résultats

5.1. Inférence lexicale et sources de connaissance

- 23 Nous avons relevé au total 654 inférences lexicales⁵ dont 324 ont porté sur les mots lexicalisés et 330 sont sur les mots non lexicalisés. Lors du premier test des connaissances lexicales (VKS1), 62,7% des mots cibles ont été jugés jamais vus et 98,5% (ce qui correspond au nombre total des réponses aux niveaux 1, 2 et 3) sémantiquement inconnus (cf. le Tableau 3).

Tableau 3. Nombre et répartition des réponses de différents niveaux de connaissance au VKS1

Niveau de connaissance	1	2	3	4	5	6	Total
Nb de réponses	410	113	121	10	0	0	654
Répartition (%)	62,7%	17,3%	18,5%	1,5%	0%	0%	100%

- 24 Les sources de connaissance sur lesquelles se sont basés les apprenants pour déduire le sens des mots inconnus relèvent essentiellement de la L2, c'est-à-dire le chinois, et ce, à différents niveaux linguistiques : mot, phrase, discours. Plus rarement, les apprenants ont utilisé leurs connaissances générales du monde. Jamais ils n'ont fait de transferts depuis leur L1, le français, ou d'une autre langue connue par eux. Nous avons classé les sources de connaissance utilisées en dix catégories, selon la classification de Paribakht & Tréville [2007], dans laquelle nous avons inclus des sources qui concernent certaines particularités de la langue chinoise. Ainsi quatre sources de connaissance y ont été ajoutées, à savoir le sens d'un caractère constitutif du mot (SC07), la forme d'un caractère constitutif du mot (SC08), le son d'un caractère constitutif du mot (SC09) et le mode de constitution du mot (SC10). À titre d'illustration nous présentons la définition de chaque catégorie accompagnée d'un exemple extrait des entretiens d'inférence lexicale dans le Tableau 4.

Tableau 4. Typologie des sources de connaissance utilisées pour inférer les mots chinois

Sources de nature non linguistique (SNL)

SC01 Connaissances générales : Le lecteur puise dans ses connaissances générales non linguistiques y compris celles du thème traité dans le texte.

Exemple : (MC14 « 彩虹 ») : Ça doit sûrement être un « arc-en-ciel », parce qu'en général au printemps quand il pleut et parfois il fait soleil... (il y a un arc-en-ciel.)

Sources propres à la L2 - Niveau du discours (ND)

SC02 Contexte lointain : Le lecteur s'appuie sur le sens général du texte, du paragraphe ou des phrases qui se trouvent avant ou après la phrase contenant le mot inconnu.

Ex. (MC09 « 乏味 ») : *On parle de la situation générale de la Chine. ... [A]u début du texte ... le narrateur ... compare les faits actuels, aujourd'hui à son enfance.*

Sources propres à la L2 - Niveau de la phrase contenant le mot (NP)

SC03 Ponctuation : Le lecteur s'appuie sur ses connaissances des règles de ponctuation et de leur signification.

Ex. (MC07 « 还礼 ») : *Déjà, on a une explication juste après, puisque on a (un) deux-points. Donc on doit dire le contexte je pense.*

SC04 Contexte proche : Le lecteur s'appuie sur le sens général de la phrase contenant le mot.

Ex. (MC19 « 飘零 ») : *[S]i je relis la phrase, ça parle encore de l'herbe et des feuilles ou des arbres. ... Oui. Là, c'est l'« automne ». Euh... donc on parle de l'automne et des feuilles.*

SC05 Grammaire : Le lecteur s'appuie sur les propriétés syntaxiques du mot, sur l'ordre des mots ou sur les catégories grammaticales.

Ex. (MC19 « 飘零 ») : *[J]e dirais plutôt un verbe avec le sujet, donc les feuilles des arbres. Les feuilles des arbres, peut-être qu'ils tournent ou avec du vent qu'ils tournent...*

Sources propres à la L2 - Niveau du mot (NM)

SC06 Association/Collocation : Le lecteur fait l'association entre le mot et un autre mot qui lui est familier ou une expression dont il croit que le mot fait partie.

Ex. (MC05 « 汤圆 ») : *Hm, ah, peut-être c'est quelque chose à boire. « tang », c'est la « soupe ». Mais c'est pas ça. Euh, ah oui, en même temps c'est pas <((Le participant dit un mot qui signifie « boire »))> « he ». Donc c'est plutôt quelque chose à manger.*

SC07 Sens d'un caractère constitutif du mot : Le lecteur s'appuie sur le sens d'un caractère contenu dans un mot pour deviner son sens.

Ex. (MC02 « 窗花 ») : *Euh je pense que c'est quelque chose qui s'accroche à la fenêtre, puisqu'il me semble que le premier caractère, c'est pour « fenêtre ».*

SC08 Forme d'un caractère constitutif du mot : Le lecteur cherche à identifier un caractère contenu dans un mot à partir d'un bloc composant du caractère, ou puise dans la signification d'un bloc composant du caractère pour deviner le sens du caractère et du mot.

Ex. (MC03 « 送灶 ») : *Le deuxième caractère, c'est le « feu » avec la « terre » ... Des plats à emporter peut-être. Le « feu », peut-être le feu pour cuisiner, pour réchauffer les plats.*

SC09 Son d'un caractère constitutif du mot : Le lecteur cherche à identifier la prononciation d'un caractère contenu dans un mot par le biais d'un bloc composant du caractère.

Ex. (MC03 « 送灶 ») : *[L]a clé du feu. ... (Le deuxième caractère,) pourquoi pas quelque chose qui peut se prononcer en « du » ou en « tu » ?*

SC10 Mode de constitution du mot : Le lecteur puise dans la façon dont est constitué un mot ou dans la relation entre les caractères constitutifs.

Ex. (MC02 « 窗花 ») : *[L]e deuxième (caractère) c'est le nom et le premier c'est un qualifiant par rapport au deuxième caractère qui est pour moi « fleur ».*

- 25 Nos participants ont utilisé plusieurs sources de connaissance différentes dans 537 inférences lexicales, soit 82,1% de la totalité des 654 inférences observées. Pour donner un exemple, MC02 « 窗花 » :

Ah le « hua », c'est « fleur ». ... Oui ça peut être ça, parce que beaucoup de gens vont utiliser ... des cartes rouges ? ... Je pense à des couronnes de fleurs ou quelque chose comme ça. En tout cas, c'est un nom parce qu'il y a « yi xie » devant, « plusieurs », « plusieurs » quelque chose.

- 26 Dans cet exemple, le participant a) a reconnu le sens du deuxième caractère de MC002, « 花 (*huā*, fleur) », b) a compris de la phrase comportant MC02 que les gens utilisent du papier rouge et c) a repéré l'adjectif « 一些 (*yìxiē*, plusieurs) » devant MC02. Autrement dit, pour déduire le sens de MC02, le participant a utilisé trois sources de connaissance, le sens d'un caractère constitutif (SC07), le contexte proche (SC04) et la grammaire (SC05). Nos participants ont utilisé en moyenne 2,4 sources de connaissance différentes pour inférer un mot chinois inconnu.
- 27 Le Tableau 5 résume les différentes sources de connaissance utilisées par les participants pour inférer (à haute voix) le sens des mots inconnus. La source la plus utilisée est le contexte proche (SC04) et ce, indépendamment de la nature lexicalisée ou non des mots cibles. Le pourcentage s'élève à 37,8% pour les mots lexicalisés et à 36% pour les mots non lexicalisés. D'autres sources fréquemment utilisées sont la grammaire (SC05), le sens d'un caractère constitutif (SC07) et la forme d'un caractère constitutif (SC08). Cependant, le poids relatif de SC05 et de SC07 n'est pas indépendant de la nature lexicalisée ou non des mots cibles. On observe ainsi des différences quant à l'ordre des sources utilisées selon que le mot cible est lexicalisé ou non. Par exemple, alors que pour les mots lexicalisés les apprenants se fient à la grammaire (21,5%), pour les mots non lexicalisés ils se fient au sens d'un caractère constitutif du mot inconnu (23,8%) (Tableau 5). Une analyse univariée par modèle linéaire général⁶ montre que cette différence entre mots lexicalisés et mots non lexicalisés est significative au niveau de la fréquence d'utilisation de la grammaire (SC05) ($D(1, 652) = 17,042, p = 0,000$) et du sens d'un caractère constitutif du mot (SC07) ($D(1, 652) = 38,703, p = 0,000$).

Tableau 5. Fréquence d'utilisation et poids relatif des sources de connaissance utilisées pour inférer les mots chinois lexicalisés (L) et non lexicalisés (NL)

	Mots L*		Mots NL**	
	Fréquence d'utilisation	Poids relatif	Fréquence d'utilisation	Poids relatif
SC01 (Connaissances générales)	34	4,5%	41	5,1%
SC02 (Contexte lointain)	32	4,2%	17	2,1 %
SC03 (Ponctuation)	2	0,3%	2	0,3%
SC04 (Contexte proche)	285	37,8%	290	36,0%
SC05 (Grammaire)	162	21,5%	113	14,0%
SC06 (Association/collocation)	8	1,0%	9	1,1%

SC07 (Sens du caractère)	112	14,9 %	192	23,8%
SC08 (Forme du caractère)	98	13,0%	121	15,0%
SC09 (Son du caractère)	9	1,2%	6	0,7%
SC10 (Mode de constitution)	12	1,6%	15	1,9%
Total	754	100%	806	100%
* 324 inférences lexicales observées				
** 330 inférences lexicales observées				

28 La différence dans l'ordre de la fréquence des quatre sources principales utilisées lors de l'inférence des mots lexicalisés (L) et non lexicalisés (NL) se résume de cette façon :

- **Mots L** : Contexte proche > Grammaire > Sens d'un caractère constitutif du mot > Forme d'un caractère constitutif du mot
- **Mots NL** : Contexte proche > Sens d'un caractère constitutif du mot > Forme d'un caractère constitutif du mot > Grammaire

5.2. Degré de réussite des inférences lexicales

29 Les tentatives d'inférence des participants ont été notées selon une échelle à cinq niveaux comme ceci :

- Une inférence abandonnée ou une réponse complètement erronée est notée 0 ;
- Une réponse complètement correcte est notée 1 ;
- Une réponse partiellement correcte est notée 0,25, 0,5 ou 0,75, en fonction de son rapprochement au sens correct.

30 Pour donner un exemple, pour le mot MC04 « 腊月, le douzième mois du calendrier lunaire », la réponse « une certaine période » est notée 0,25 ; la réponse « un nom d'un mois particulier » est notée 0,5 et la réponse « un mois lié au Nouvel an chinois » est notée 0,75. Nous avons demandé à deux Chinois natifs bilingues en français d'effectuer une double correction de la notation. En cas de désaccord sur la notation de certains mots, nous avons opté pour le score le plus fréquent.

31 Le Tableau 6 présente le score total pour les 654 inférences lexicales observées (141,25, soit en moyenne de 0,22). Le score moyen pour les inférences des mots lexicalisés est de 0,19, alors que celui des mots non lexicalisés est de 0,25. Selon les résultats d'analyse par Test-T apparié⁷, cette différence est significative ($t(32) = -2,61, p = 0,014$).

Tableau 6. Score d'inférence lexicale pour les mots cibles lexicalisés (L) et non lexicalisés (NL)

	Mots cibles	Mots cibles L	Mots cibles NL
Nb d'observation	654	324	330
Score total	141,25	60,25	81

Score moyen*	0,22	0,19	0,25
Écart-type	0,32	0,36	0,28
* Score maximal = 1			

- 32 Enfin, si les apprenants ont produit proportionnellement plus d'inférences partiellement correctes pour les mots non lexicalisés (47,9%) que pour les mots lexicalisés (10,2%), ils ont produit plus d'inférences entièrement correctes pour les mots lexicalisés (13,9%) que pour les mots non lexicalisés (3,6%). (Tableau 7)

Tableau 7. Nombre et répartition des réponses de différents scores d'inférence pour les mots cibles lexicalisés (L) et non lexicalisés (NL)

	Nb. d'observ.	Score 0		Score 0~1		Score 1	
		Nb.	%	Nb.	%	Nb.	%
Mots L	324	246	75,9%	33	10,2%	45	13,9%
Mots NL	330	160	48,5%	158	47,9%	12	3,6%
Total	654	406	62,1%	191	29,2%	57	8,7%

- 33 Parmi les quatre principales sources de connaissance utilisées, nous avons constaté un effet significatif de l'utilisation du sens d'un caractère constitutif (SC07) sur le score d'inférence à l'aide d'une analyse univariée par modèle linéaire général ($D(1, 638) = 23,962, p = 0,000$). Plus précisément, le score en moyenne est de 0,33 pour les inférences pendant lesquelles les participants ont utilisé cette source de connaissance, mais seulement de 0,07 pour les inférences n'ayant pas engagé cette source de connaissance.

5.3. Rétention des mots inférés

- 34 L'un des objectifs de cette étude visait la rétention du vocabulaire à l'issue de l'inférence lexicale. Plus précisément, nous voulions savoir de quelle manière une condition d'apprentissage cognitivement plus chargée (c'est-à-dire la réalisation de deux tâches après l'inférence lexicale) pouvait favoriser la rétention du mot inféré dans le lexique mental de nos participants. Concrètement, nous nous demandions si la tâche supplémentaire (la consultation du dictionnaire bilingue), réservée seulement aux participants du groupe II, pouvait contribuer à une meilleure rétention des mots inférés. Enfin, nous voulions aussi savoir quel type de mots serait le mieux retenu, les mots lexicalisés ou les mots non lexicalisés.

5.3.1. Évolution de la connaissance lexicale multidimensionnelle pour les mots cibles

- 35 La Figure 1 montre les nombres et les répartitions des réponses portant sur six niveaux de connaissance (voir encadré dans la méthodologie) obtenues aux trois tests VKS. À

partir de ces réponses, nous avons examiné l'évolution de la connaissance lexicale sous trois aspects (la forme, le sens et l'emploi) des mots cibles chez nos participants.

Figure 1. Évolution de la répartition des réponses de différents niveaux de connaissance des mots cibles, aux trois tests de connaissance lexicale

	Nb d'observ.		Niveau 1		Niveau 2		Niveau 3		Niveau 4		Niveau 5		Niveau 6	
VKS1	654	100%	410	62,7%	113	17,3%	121	18,5%	10	1,5%	0	0%	0	0%
VKS2	654	100%	110	16,8%	104	15,9%	272	41,6%	65	9,9%	71	10,9%	32	4,9%
VKS3	654	100%	78	11,9%	93	14,2%	295	45,1%	71	10,9%	72	11,0%	45	6,9%

- 36 Selon les analyses statistiques par modèle mixte à lien cumulatif⁸, la connaissance lexicale des mots cibles chez nos participants est significativement différente entre VKS1 et VKS3 (LR χ^2 (2) = 832,48, $p < 0,000$).
- 37 Cette évolution s'avère positive : la proportion des réponses du niveau 1, correspondant à la connaissance « jamais-vu » pour un mot donné, a considérablement diminué de 62,7% au VKS1 à 16,8% au VKS2 et à 11,9% au VKS3 ; par contre, la connaissance sur le plan de la forme lexicale, à savoir la reconnaissance partielle ou complète de la forme qui correspond respectivement aux niveaux 2 et 3 de l'échelle, s'est améliorée. En effet, de 35,8% au VKS1, la proportion des réponses du niveau 2 et du niveau 3 est passée à 57,5% au VKS2 et à 59,3% au VKS3 ; quant à la connaissance du sens des mots cibles qui correspond aux niveaux 4, 5 et 6, par rapport au VKS1 (1,5%), la proportion des réponses de ces trois niveaux a connu une augmentation considérable au VKS2 (25,7%) et au VKS3 (28,8%) ; enfin, concernant les connaissances productives des mots, à savoir la capacité à produire une phrase correcte avec tel ou tel mot cible qui correspond au niveau 6 de l'échelle, les participants ont réussi, lors du VKS2 et du VKS3, à produire une phrase correcte respectivement pour 4,9% et 6,9% des mots cibles.

5.3.2. Évolution de la connaissance sémantique pour les mots cibles

- 38 En dehors des analyses sur l'évolution de la connaissance lexicale sous divers aspects à l'aide de six niveaux de connaissance du VKS, nous avons également examiné l'évolution de la connaissance seulement sur le plan sémantique des mots cibles chez nos participants. Rappelons ici que les niveaux 4, 5 et 6 du VKS portent sur le sens du mot testé, et que les participants doivent fournir une réponse sémantique à ces niveaux. Les réponses sur les sens des mots cibles obtenues aux trois tests VKS, ainsi que les réponses d'inférence lexicale, ont été notées sur le même barème⁹, c'est-à-dire sur l'échelle de 0, 0,25, 0,5, 0,75 et 1 selon le rapprochement de la réponse au sens correct du mot. Les réponses aux niveaux 1, 2 et 3 du test VKS qui correspondent à une non-connaissance du sens du mot testé sont notées 0.

- 39 Le Tableau 8 montre les scores de la connaissance sémantique, totaux et moyens, des mots cibles, au VKS1, à l'inférence lexicale et aux VKS2 et VKS3. Selon les résultats d'analyse par modèle mixte linéaire¹⁰, la connaissance sur le sens des mots cibles chez nos participants a eu une évolution positive significative entre VKS1 et VKS3 ($F(3, 217) = 53,890, p = 0,000$).

Tableau 8. Score de la connaissance sémantique des mots cibles obtenus aux trois tests de connaissance lexicale et à l'inférence lexicale

	VKS1 (pré-inférence)	Inférence (J0)	VKS2 (J14)	VKS3 (J28)
Nb. d'observation	654	654	654	654
Score total	4,25	141,25	134,75	151,5
Score moyen*	0,01	0,22	0,21	0,23
Écart-type	0,05	0,32	0,38	0,39
* Score maximal = 1				

5.3.3. Différence d'évolution entre les deux groupes

- 40 Nous avons eu au total 338 observations pour 17 participants dans le groupe I et 316 observations pour 16 participants dans le groupe II.
- 41 La comparaison de l'évolution de la connaissance lexicale sous divers aspects (forme, sens et emploi) entre les deux groupes, analysée par modèle de régression ordinale, montre que la différence d'évolution entre deux groupes est significative ($LR \chi^2(2) = 15,94, p = 0,000$). La différence principale entre les deux groupes se situe dans l'évolution des proportions des réponses aux niveaux 1, 2 et 3 : les participants du groupe II présentent, par rapport au groupe I, une diminution plus importante des réponses du niveau 1 et du niveau 2 mais une augmentation plus grande des réponses du niveau 3. Les participants du groupe II semblent donc avoir mieux retenu le vocabulaire inféré (Figure 2). Cela signifie que la consultation du dictionnaire bilingue, la tâche supplémentaire réservée aux participants du groupe II, a été bénéfique en termes de gains lexicaux pour les mots cibles.

Figure 2. Évolution de la répartition des réponses de différents niveaux de connaissance (groupe I et groupe II), aux trois tests de connaissance lexicale

	Nb d'observ.		Niveau 1		Niveau 2		Niveau 3		Niveau 4		Niveau 5		Niveau 6	
VKS1	338	100%	207	61,2%	66	19,5%	62	18,4%	3	0,9%	0	0,0%	0	0,0%
VKS2	338	100%	69	20,4%	72	21,3%	128	37,9%	25	7,4%	32	9,5%	12	3,5%
VKS3	338	100%	43	12,7%	63	18,6%	153	45,3%	27	8,0%	31	9,2%	21	6,2%

Groupe II

	Nb. d'observ.		Niveau 1		Niveau 2		Niveau 3		Niveau 4		Niveau 5		Niveau 6	
VK	316	100%	203	64,2%	47	14,9%	59	18,7%	7	2,2%	0	0,0%	0	0,0%
VK	316	100%	41	13,0%	32	10,1%	144	45,6%	40	12,7%	39	12,3%	20	6,3%
VK	316	100%	35	11,1%	30	9,5%	142	44,9%	44	13,9%	41	13,0%	24	7,6%

- 42 Quant à l'évolution de la connaissance sémantique des mots cibles chez les participants des deux groupes, les scores moyens sur le sens des mots aux quatre phases sont représentés dans la Figure 3. Selon les résultats d'analyse par modèle mixte linéaire, la différence d'évolution entre deux groupes n'est pas significative ($F(3, 217) = 1,679, p = 0,172$). Cela signifie que la tâche supplémentaire du groupe II n'a pas eu d'effet sur le gain sémantique pour les mots cibles.

Figure 3. Évolution du score de la connaissance sémantique en moyenne des mots cibles du groupe I et du groupe II

5.3.4. Différence d'évolution entre les mots lexicalisés et non lexicalisés

- 43 L'évolution de la connaissance sémantique chez les participants pour les mots lexicalisés et non lexicalisés est représentée dans la Figure 4, qui montre les scores moyens sur le sens des mots cibles à différentes phases. L'analyse statistique par modèle mixte linéaire montre que l'appropriation sémantique des mots inconnus chez les participants est différente selon que les mots sont lexicalisés ou non lexicalisés. L'évolution de la connaissance sémantique pour les mots lexicalisés est significativement différente de celle pour les mots non lexicalisés ($F(3, 217) = 3,598, p = 0,014$).

Figure 4. Évolution du score de la connaissance sémantique en moyenne des mots cibles lexicalisés (L) et non lexicalisés (NL)

6. Discussion

6.1. Choix des sources de connaissance pendant l'inférence

- 44 Les apprenants, afin de comprendre le sens des mots chinois inconnus, se fient principalement au contexte proche et ce, indépendamment de la nature lexicalisée ou non des mots cibles. Autrement dit, ils cherchent à comprendre le sens du mot inconnu à travers le sens général de la phrase contenant ce mot. Paribakht & Tréville [2007] observent un même phénomène chez des apprenants francophones et persanophones de l'anglais L2 et Fan & Zhang [2014] chez des apprenants sinophones de l'anglais L2. Cette stratégie est courante en L1 aussi (Cornaire [1999]).
- 45 Par ailleurs, nous avons vu que nos participants, en dehors du contexte immédiat, s'appuyaient souvent sur les sens des caractères constitutifs des mots inconnus, surtout lorsqu'il s'agissait des mots non lexicalisés. Cela semble être lié au fait qu'en chinois, le sens d'un caractère constitutif d'un mot est souvent en rapport étroit avec le sens du mot dans son ensemble. Par exemple, 书 (livre) et 包 (sac) signifient 书包 (cartable). Ce fait, conjointement à celui selon lequel les mots non lexicalisés dans notre étude contenaient plus de caractères courants connus par nos participants, expliquerait, selon nous, que les apprenants aient pu tirer profit de ces connaissances afin de comprendre le sens du mot dans son ensemble. Autrement dit, à la rencontre d'un mot non lexicalisé inconnu, si le sens d'un caractère constitutif de ce nouveau mot est connu, l'apprenant francophone préfère utiliser cette stratégie plutôt que de s'appuyer sur les propriétés syntaxiques du mot, sur l'ordre des mots ou sur sa catégorie grammaticale. Les apprenants du chinois L2 au niveau intermédiaire semblent ainsi conscients de l'apport sémantique d'un caractère constitutif au sens global d'un mot chinois.

6.2. Lexicalisation et degré de réussite de l'inférence lexicale

- 46 Dans l'ensemble nos participants ont éprouvé beaucoup de difficultés à déduire correctement les sens des mots cibles inconnus, lexicalisés ou non, puisqu'en moyenne ils ont réussi leurs interprétations à hauteur seulement de 0,22.

- 47 La comparaison des réponses obtenues montre que nos participants ont mieux réussi la déduction du sens des mots non lexicalisés (0,25) que celle des mots lexicalisés (0,19). Ce résultat inattendu contredit celui que des études réalisées en anglais L2 ont obtenu pour des mots anglais de types variés (simples, composés et dérivés) (Paribakht [2005] ; Paribakht & Tréville [2007]). Dans notre étude nous avons vu que le sens des caractères constitutifs des mots chinois semble avoir aidé les participants dans leurs inférences.
- 48 Ceci dit, l'effet de la lexicalisation dans la L1 de l'apprenant n'est pas sans importance, car même si les participants ont produit proportionnellement plus d'inférences partiellement correctes pour les mots non lexicalisés (47,9%) que pour les mots lexicalisés (10,2%), les inférences entièrement correctes pour les mots lexicalisés (13,9%) dépassent largement celles pour les mots non lexicalisés (3,6%). Autrement dit, si les mots non lexicalisés ont donné lieu à davantage d'inférences partiellement correctes, ce sont les mots lexicalisés, et de loin, qui ont comptabilisé le plus d'inférences entièrement correctes. Ce grand écart en termes d'inférence entièrement correcte entre les mots lexicalisés et non lexicalisés a déjà été constaté dans les études en anglais L2 de Paribakht [2005] auprès d'apprenants persanophones (17% pour les mots lexicalisés vs 6% pour les mots non lexicalisés) et de Fan & Zhang [2014] auprès d'apprenants sinophones (48% pour les mots lexicalisés vs 18% pour les mots non lexicalisés). Ces résultats valideraient l'idée selon laquelle un concept existant dans la L1 de l'apprenant favorise la compréhension d'un mot L2 inconnu (Wesche & Paribakht [2009]).

6.3. Rétention du vocabulaire

- 49 Les post tests (VKS2 et VKS3) ont montré que les connaissances lexicales pour les mots cibles ont évolué à court et à long terme. Si l'apport de l'inférence n'est pas à exclure de l'apprentissage incident du vocabulaire, la focalisation sur les mots cibles par le biais des tâches post-inférence a donné l'occasion aux apprenants de porter leur attention spécifiquement sur ces mots une nouvelle fois avec comme corrélat une meilleure rétention subséquente. La tâche supplémentaire réservée aux participants du groupe II, la consultation du dictionnaire bilingue, a favorisé globalement les gains lexicaux pour les mots cibles, à court et à long terme. Cependant l'effet bénéfique de la tâche de consultation du dictionnaire bilingue sur les gains lexicaux pour les mots cibles ne se situe pas au niveau du sens retenu (écart non significatif entre les deux groupes) mais au niveau de la forme des mots. Ce résultat contredit Zhu [2004] et Laufer & Rozovski-Roitblat [2015] probablement pour les raisons suivantes :
- 1) Dans notre étude, la consultation du dictionnaire bilingue n'a pas servi à la compréhension d'un mot présent dans le texte mais à la complétion de l'exercice lacunaire. Il se pourrait ainsi que le besoin de vérifier le sens des mots cibles et *a fortiori* d'apprendre ces mots, ait été relativement faible chez nos participants (cf. Laufer & Hulstijn [2001]).
 - 2) Une seule consultation du dictionnaire pour compléter les phrases lacunaires semble insuffisante pour que nos apprenants acquièrent un gain sémantique pérenne pour ce mot. Si cela n'a rien d'extraordinaire étant donné qu'un nouveau mot doit être vu entre 10 et 15 fois avant qu'il puisse être connu par l'apprenant (Uchihara et al. [2019]), il semble encore plus vrai pour les mots qui renvoient à des concepts inexistant dans le lexique de la L1 de l'apprenant.

- 50 Quant à l'effet de la lexicalisation sur la rétention des mots cibles, les participants ont eu un gain sémantique plus important pour les mots lexicalisés que pour les mots non lexicalisés à court et à long terme. Si le score d'inférence moyen des mots non lexicalisés est supérieur à celui des mots lexicalisés (0,25 vs 0,19), lors du VKS2, le score sur le sens des mots non lexicalisés est inférieur à celui des mots lexicalisés (0,19 vs 0,22). Par ailleurs, lors du VKS3, l'écart entre les scores de la connaissance sémantique des mots lexicalisés et des mots non lexicalisés a continué à croître (0,19 vs 0,27). Nous schématisons ces comparaisons du score entre les mots lexicalisés (L) et non lexicalisés (NL), sur le court et long terme, comme suit :

L < NL	>>	L > NL	>>	L > NL
(Inférence)		(VKS2)		(VKS3)

- 51 Tenant compte de ces effets opposés concernant la rétention sémantique des mots lexicalisés et non lexicalisés, il est raisonnable de conclure que s'il a été plus facile pour les participants de deviner le sens des mots non lexicalisés grâce à leur connaissance préalable des caractères constitutifs des mots, il a été plus difficile pour eux de retenir le sens de ces mots, à court et à long terme (par rapport aux mots lexicalisés). Il paraît que les apprenants ont tendance à mieux retenir en mémoire les mots qui ont un équivalent direct en L1, par rapport aux mots qui renvoient à des référents nouveaux, qui doivent aussi être appris. Parmi les mots cibles, ceux qui ont obtenu le meilleur score de connaissance sémantique aux post tests sont des mots qui évoquent clairement quelque chose chez des apprenants francophones du chinois (par exemple, MC12 (五颜六色, *de diverses couleurs, multicolore*) et MC16 (风筝, *cerf-volant*)).
- 52 Afin de comprendre un mot L2 n'ayant pas de concept correspondant dans son lexique mental, l'apprenant doit mobiliser plusieurs concepts connexes. Par exemple, la compréhension du MC04 (腊月, *le douzième mois du calendrier lunaire*) implique les concepts « mois », « dernier » ou « douzième », « calendrier chinois », etc. Cela signifie qu'en vue de la rétention du mot, l'apprenant doit créer un nouveau lemme à partir de la forme 腊月, qui est lui-même relié à un ensemble de concepts connexes (un lien forme-concepts). Il est raisonnable de penser qu'un lien forme-concepts soit plus complexe à établir qu'un lien forme-concept unitaire, comme pour un mot déjà lexicalisé en L1. La complexité caractérise également la reconnaissance d'un mot : l'apprenant active un concept (existant) si le mot est lexicalisé, mais plusieurs concepts connexes si le mot est non lexicalisé. Et plus les concepts concernés à activer sont complexes, plus l'apprenant risque de ne récupérer que partiellement les informations conceptuelles pertinentes. Nos résultats confirment qu'au VKS3, par rapport au VKS2, les réponses entièrement correctes pour les mots non lexicalisés retombent, tandis que celles pour les mots lexicalisés augmentent.
- 53 Parmi les mots non lexicalisés, MC07 (还礼, *faire un présent en retour, offrir un cadeau à son tour*) est le seul d'avoir reçu une augmentation continue de score sur la connaissance sémantique aux phases postérieures à l'inférence. Nous expliquons cela par le fait que le sens de ce mot est étroitement lié aux sens de ses caractères constitutifs, 还 (*rendre, retourner*) et 礼 (*cadeau*). Le mot cible 07 était donc un mot sémantiquement transparent pour nos participants, puisqu'ils connaissaient le sens de

ses deux caractères constitutifs. Les concepts connexes de ce mot étant déjà disponibles dans la mémoire des apprenants, ils ont pu élaborer le sens global de MC07 qui renvoie pourtant à un concept qui n'existe pas dans leur L1. En revanche, le mot MC03 (送灶, *saluer avec des offrandes le départ du Génie du foyer pour le Ciel, le 23 de la 12^e lune*) a reçu le plus mauvais score sémantique lors des post tests. C'est également le mot le moins réussi à l'inférence lexicale. La compréhension de ce mot nécessite des connaissances socio-culturelles sur les mythes et les coutumes chinois. La complexité des connaissances sémantiques liés à ce mot non lexicalisé rend plus difficiles l'inférence de son sens, ainsi que la rétention ultérieure du mot.

Conclusion

- 54 Cette étude analyse l'effet de la lexicalisation en français sur les démarches d'inférence lexicale adoptées par un groupe d'apprenants francophones du chinois L2 et sur la rétention du vocabulaire inféré à court et à long terme. De façon prévisible, la source principale sur laquelle se basent les apprenants afin de déduire les mots inconnus est la phrase contenant le mot – qu'il soit lexicalisé ou non. Une autre source importante est le sens des caractères constitutifs. Nos apprenants de niveau intermédiaire en chinois L2 se sont ainsi montrés conscients du fait que les sens individuels des caractères constitutifs d'un mot chinois sont souvent en rapport étroit avec le sens global du mot.
- 55 Nos participants sont parvenus à mieux comprendre les mots non lexicalisés que les mots lexicalisés. La lexicalisation en L1 n'est donc pas nécessairement l'élément le plus important dans la compréhension d'un mot chinois inconnu par des apprenants francophones. Si le mot inconnu est sémantiquement transparent et les apprenants connaissent les sens de ses caractères constitutifs, il est possible pour eux de comprendre, ne serait-ce que partiellement, le sens de ce mot, que ce dernier soit lexicalisé ou non dans la L1 des apprenants. La lexicalisation en français L1 n'a pas eu un rôle concluant dans la compréhension d'un mot chinois inconnu, mais elle a eu un effet significatif sur la rétention de celui-ci. Les apprenants francophones ont eu plus de difficultés à retenir le sens d'un mot chinois non lexicalisé en français L1, à court et à long terme, qu'un mot chinois lexicalisé. L'existence d'un concept correspondant dans la L1 de l'apprenant favorise donc la rétention du sens du mot.
- 56 L'autre intérêt de cette étude est son illustration de l'effet de la charge cognitive sur les gains lexicaux des mots inconnus rencontrés en lecture. Si un investissement cognitif plus important favorise l'apprentissage des mots chinois inconnus à court et à long terme, le gain lexical se situe surtout au niveau de la forme de ces mots, plutôt qu'au niveau de leur sens. Il est possible que la tâche supplémentaire dans notre étude (la consultation d'un dictionnaire bilingue pour compléter les phrases lacunaires), n'ait pas permis aux apprenants un investissement cognitif suffisant pour obtenir un gain sémantique important. De futures études sont alors à mener au sujet de l'investissement cognitif de l'apprenant sur le plan de l'appropriation sémantique des mots en chinois L2.
- 57 Enfin les apprenants francophones retiennent mieux, à court et à long terme, les mots chinois sémantiquement transparents, constitués de caractères qu'ils connaissent déjà. Autrement dit, les concepts connexes disponibles dans la mémoire des apprenants francophones favorisent l'apprentissage d'un nouveau mot chinois, lexicalisé ou non. En définitive, l'apport de l'inférence lexicale peut être renforcé si cette stratégie de

compréhension en lecture s'accompagne de tâches complémentaires focalisant l'attention des apprenants à nouveau sur les mots inférés.

BIBLIOGRAPHIE

- BENGLEIL N. F. & PARIBAKHT T. S., 2004, "L2 reading proficiency and lexical inferencing by university EFL learners", *The Canadian Modern Language Review*, 61(2), 225-249.
- BENSOUSSAN M. & LAUFER B., 1984, "Lexical guessing in context in EFL reading comprehension", *Journal of Research in Reading*, 7(1), 15-32.
- BROWN C., 1993, "Factors affecting the acquisition of vocabulary: frequency and saliency of words", in HUCKIN T., HAYNES M. & COADY J. (Eds.), *Second Language Reading and Vocabulary Learning*, NJ: Ablex Pub. Corp., 263-286.
- COADY J. *et al.*, 1993, "High frequency vocabulary and reading proficiency in ESL readers", in HUCKIN T., HAYNES M. & COADY J. (Eds.), *Second Language Reading and Vocabulary Learning*, NJ: Ablex Pub. Corp., 217-226.
- CORNAIRE C., 1999, *Le Point sur la lecture*, CLE International.
- DE BOT K., PARIBAKHT T. S. & WESCHE M. B., 1997, "Toward a lexical processing model for the study of second language vocabulary acquisition: evidence from ESL reading", *Studies in Second Language Acquisition*, 19(3), 309-329.
- Dictionnaire chinois-français 汉法词典*, 2001, 商务印书馆 The Commercial Press.
- FAN L. & ZHANG J., 2014, "汉语词汇化对英语词汇推理加工过程影响的研究 A study on the influence of Chinese lexicalization on lexical inferencing in English", *外语与外语教学 Foreign Languages and Their Teaching*, (6), 50-55.
- FRANTZEN D., 2003, "Factors Affecting How Second Language Spanish Students Derive Meaning from Context", *The Modern Language Journal*, 87(2), 168-199.
- FRASER C., 1999, "Lexical processing strategy use and vocabulary learning through reading", *Studies in Second Language Acquisition*, 21(2), 225-241.
- HAASTRUP K., 1989, "The learner as word processor", in NATION I. P. & CARTER R. (Eds.) *Vocabulary Acquisition*, 34-46.
- HAASTRUP K., 1991, *Lexical Inferencing Procedures, or, Talking about Words: Receptive Procedures in Foreign Language Learning with Special Reference to English*, Gunter Narr Verlag.
- HAYNES M., 1993, "Patterns and perils of guessing in second language reading", in HUCKIN T., HAYNES M. & COADY J. (Eds.), *Second Language Reading and Vocabulary Learning*, NJ: Ablex Pub. Corp., 46-64.
- HILTON H. E., 2003, « L'accès au lexique mental dans une langue étrangère : le cas des francophones apprenant l'anglais », *CORELA*.
- HIRSH D. & NATION I. S. P., 1992, "What vocabulary size is needed to read unsimplified texts for pleasure?", *Reading in a Foreign Language*, 8(2), 689-696.

- HORST M., COBB T. & MEARA P., 1998, "Beyond a clockwork orange: Acquiring second language vocabulary through reading", *Reading in a Foreign Language*, 11(2), 207–223.
- HU H.-C. M. & NATION I. P., 2000, "Unknown vocabulary density and reading comprehension", *Reading in a Foreign Language*, 13(1), 403–430.
- HULSTIJN J., 1992, "Retention of inferred and given word meanings: Experiments in incidental vocabulary learning", in ARNAUD P. J. L. & BÉJOINT H. (Eds.), *Vocabulary and Applied Linguistics*, Houndmills, Basingstoke, Hampshire: Macmillan, 113–125.
- HULSTIJN J., HOLLANDER M. & GREIDANUS T., 1996, "Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words", *The Modern Language Journal*, 80(3), 327–339.
- JARVIS S., 2000, "Methodological rigor in the study of transfer: identifying L1 influence in them interlanguage lexicon", *Language Learning*, 50(2), 245–309.
- JIANG N., 2002, "Semantic transfer and its implications for vocabulary teaching in a second language", *The Modern Language Journal*, 88, 416–432.
- KELLY P., 1990, "Guessing: No substitute for systematic learning of lexis", *System*, 18(2), 199–207.
- KODA K., 1989, "The effects of transferred vocabulary knowledge on the development of L2 reading proficiency", *Foreign Language Annals*, 22(6), 529–540.
- KODA K., 2005, *Insights into Second Language Reading*, Cambridge University Press.
- KROLL J. F. et al., 2010, "The revised hierarchical model: A critical review and assessment", *Bilingualism (Cambridge, England)*, 13(3), 373–381.
- LAUFER B., 1991, "How much lexis is necessary for reading comprehension?", in ARNAUD P. J. L. & BÉJOINT H. (eds.), *Vocabulary and Applied Linguistics*, Houndmills, Basingstoke, Hampshire: Macmillan, 126–132.
- LAUFER B., 1997, "The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess", in COADY J. & HUCKIN T. (Eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*, New York: Cambridge University Press, 20–34.
- LAUFER B., 2003, "Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? Some empirical evidence", *The Canadian Modern Language Review*, 59(4), 567–587.
- LAUFER B. & HULSTIJN J., 2001, "Incidental vocabulary acquisition in a second language: the construct of task-induced involvement", *Applied Linguistics*, 22(1), 1–26.
- LAUFER B. & ROZOVSKI-ROITBLAT B., 2015, "Retention of new words: quantity of encounters, quality of task, and degree of knowledge", *Language Teaching Research*, 19, 687–711.
- LAUFER B. & SHMUELI K., 1997, "Memorizing new words: Does teaching have anything to do with it?", *RELC Journal*, 28(1), 89–108.
- LI X. L., 1988, "Effects of contextual cues on inferring and remembering meanings of new words", *Applied Linguistics*, 9(4), 402–413.
- MONDRIA J.-A. & WIT-DE BOER M., 1991, "The effects of contextual richness on the guessability and the retention of words in a foreign language1", *Applied Linguistics*, 12(3), 249–267.
- NAGY W. E., MCCLURE E. F. & MONTSERRAT M., 1997, "Linguistic transfer and the use of context by Spanish-English bilinguals", *Applied Psycholinguistics*, 18(4), 431–452.

- NASSAJI H., 2003, "L2 vocabulary learning from context: strategies, knowledge sources, and their relationship with success in L2 lexical inferencing", *TESOL Quarterly*, 37(4), 645–670.
- NASSAJI H., 2004, "The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success", *The Canadian Modern Language Review*, 61(1), 107–135.
- PARIBAKHT T. S., 2005, "The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language", *Language Learning*, 55(4), 701–748.
- PARIBAKHT T. S., 2010, "The effect of lecalization in the native language on second language lexical inferencing: A cross-language study", in CHACÓN-BELTRÁN D. R., ABELLO-CONTESSE C. & TORREBLANCA-López M. del M. (Eds.) *Further Insights into Non-Native Vocabulary Teaching and Learning*, Clevedon: Multilingual Matters.
- PARIBAKHT T. S. & TRÉVILLE M.-C., 2007, « L'inférence lexicale chez des locuteurs de français et des locuteurs de persan lors de la lecture de textes anglais : effet de la lexicalisation en première langue », *The Canadian Modern Language Review*, 63(3), 399–428.
- PARIBAKHT T. S. & WESCHE M., 1997, "Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition", in COADY J. & HUCKIN T. (Eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*, New York: Cambridge University Press, 174–200.
- PARIBAKHT T. S. & WESCHE M., 1999, "Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing", *Studies in Second Language Acquisition*, 21(2), 195–224.
- PARRY K., 1993, "Too many words: Learning the vocabulary of an academic subject", in HUCKIN T., HAYNES M. & COADY J. (Eds.), *Second Language Reading and Vocabulary Learning*, NJ: Ablex Pub. Corp., 109–129.
- PIGADA M. & SCHMITT N., 2006, "Vocabulary acquisition from extensive reading: A case study", *Reading in a Foreign Language*, 18(1), 1–28.
- PRINCE P., 1996, "Second language vocabulary learning: The role of context versus translations as a function of proficiency", *The Modern Language Journal*, 80(4), 478–493.
- PULIDO D., 2007, "The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading", *Applied Linguistics*, 28(1), 66–86.
- QIAN X. J., 2003, "汉语阅读中的伴随性词汇学习研究 On incidental vocabulary learning in Chinese L2 reading", *北京大学学报 (哲学社会科学版) Journal of Peking University (Philosophy and Social Sciences)*, 40(4), 135–142.
- RYAN A. & MEARA P., 1991, "The case of the invisible vowels: Arabic speakers reading English words", *Reading in a Foreign Language*, 7(2), 531–540.
- SCHMITT N., 1997, "Vocabulary learning strategies", in SCHMITT Norbert & MCCARTHY M. (Eds.), *Vocabulary: Description, Acquisition and Pedagogy*, 1st edition, Cambridge: Cambridge University Press, 199–227.
- SUAREZ A. & MEARA P., 1989, "The effects of irregular orthography on the processing of words in a foreign language", *Reading in a Foreign Language*, University of Hawaii National Foreign Language Resource Center, 6(1), 349–356.
- THÉOPHANOUS O., 2001a, « Déviations lexicales de forme et de sens chez les apprenants de français langue étrangère », *Revue française de linguistique appliquée*, 6(1), 107–120.

THÉOPHANOUS O., 2001b, « Identification des congénères interlinguaux : Quels facteurs sont en jeu ? Un état de la question », *ITL Review of Applied Linguistics*, 131-132, 85-105.

THÉOPHANOUS O. & DREYFUS M., 2004, « Attitudes d'inférence lexicale lors d'une tâche de lecture en FLS », *Les cahiers de l'association de didactique du français langue étrangère*, 15, 173-192.

TRÉVILLE M.-C., 1996, "Lexical learning and reading in L2 at the beginner level: The advantage of cognates", *The Canadian Modern Language Review*, 53(1), 173-190.

UCHIHARA T., WEBB S. & YANAGISAWA A., 2019, "The effects of repetition on incidental vocabulary learning: a meta-analysis of correlational studies", *Language Learning*, 69(3), 559-599.

WALKER L. J., 1983, "Word identification strategies in reading a foreign language", *Foreign Language Annals*, 16(4), 293-299.

WARING R. & TAKAKI M., 2003, "At what rate do learners learn and retain new vocabulary from reading a graded reader?", *Reading in a Foreign Language*, 15(2), 130-163.

WESCHE M. & PARIBAKHT T. S., 2009, *Lexical Inferencing in a First and Second Language: Cross-Linguistic Dimensions*, Bristol; Buffalo: Multilingual Matters.

ZAHAR R., COBB T. & SPADA N., 2001, "Acquiring vocabulary through reading: Effects of frequency and contextual richness", *The Canadian Modern Language Review*, 57(5), 41-72.

ZHU Y., 2004, "边注和查词典等输入调整对留学生伴随性词汇学习的作用 The effect of modified input methods of marginal glosses and dictionary-consulting on the incidental vocabulary learning by foreign students", *世界汉语教学 Chinese Teaching in the World*, (04), 74-84.

NOTES

1. "Whereas inferring the meaning of a word lexicalized in L1 can be characterized as a process involving lemma recognition, that of a word not lexicalized in the native language involves a process of lemma construction."
2. Les quatre autres stratégies par ordre décroissant sont le dictionnaire bilingue, la répétition orale, la répétition écrite et l'étude de l'orthographe.
3. <http://www.chinesetest.cn/gosign.do?id=1&lid=0#> (consulté le 25 mai 2021)
4. Nous avons ajouté un niveau (Niveau 2) à l'échelle d'origine afin d'avoir une vue plus complète de la reconnaissance de la forme d'un mot chez nos participants.
5. Selon les réponses des participants obtenues au VKS1, trois mots lexicalisés se sont avérés connus par certains participants (MC08 在乎 et MC12 五颜六色 par un participant, MC14 彩虹 par quatre participants). Ces six cas ont donc été exclus de nos analyses. Certains participants ont donné des réponses correctes sur les sens de certains mots cibles, mais leurs rapports verbaux lors de l'entretien d'inférence lexicale ne peuvent pas justifier une connaissance complète pour ces mots. Il est possible que les participants aient déjà rencontré les mots sans bien les retenir ou simplement confondu le mot concerné avec un autre mot formellement similaire. Dans ce cas-là, nous avons inclus ces mots dans nos observations en ramenant le niveau de connaissance au niveau 4 qui correspond à une connaissance incertaine du sens du mot.
6. Le modèle linéaire général (logiciel SPSS) fournit une analyse de variance pour des variables dépendantes par un facteur, en l'occurrence, pour l'utilisation des sources de connaissance par la nature de lexicalisation des mots cibles.
7. Test-T apparié (bilatéral) (logiciel SPSS) permet d'analyser la différence entre les moyennes du score d'inférence des mots cibles lexicalisés et des mots non lexicalisés.

8. L'échelle à six niveaux de connaissance du VKS est une variable ordinale. Nous avons donc opté pour le modèle mixte à lien cumulatif (*Cumulative Link Mixed Model*) du logiciel R qui analyse par régression les effets des facteurs en mesure répétée (dans notre cas, le temps, le groupe et la lexicalisation) sur une variable ordinale (le niveau de connaissance).

9. Nous rappelons que, lors du VKS1, certains participants ont donné une réponse correcte pour certains mots cibles pour lesquels ils n'avaient en réalité pas une connaissance stabilisée. Ces réponses ont été rectifiées à 0,5.

10. Un modèle mixte linéaire (logiciel SPSS) permet d'analyser par régression l'effet des facteurs en mesure répétée (dans notre cas, le temps, le groupe et la lexicalisation) sur une variable dépendante continue (le score du sens).

RÉSUMÉS

L'inférence lexicale est une importante stratégie de lecture et d'apprentissage du vocabulaire, en L2 comme en L1. Nous nous intéressons à l'utilisation de cette stratégie en L2 afin de déduire le sens d'un mot inconnu pendant la lecture et l'apprentissage subséquent du mot inféré, en nous penchant sur l'effet de la lexicalisation (c'est-à-dire l'existence ou pas d'une unité lexicale équivalente attestée dans la L1 de l'apprenant pour un mot donné de la L2) et l'effet de l'investissement cognitif de l'apprenant sur le mot inconnu. Nous avons observé 33 apprenants francophones du chinois L2 lorsqu'ils infèrent 20 mots chinois inconnus intégrés dans deux textes et avons examiné l'évolution des connaissances lexicales des participants pour les mots cibles. Nos résultats montrent que la nature de la lexicalisation d'un mot chinois inconnu n'est pas le facteur concluant dans la compréhension du mot chez les apprenants francophones. Ces derniers ont pourtant plus de difficultés à retenir le sens d'un mot chinois non lexicalisé qu'un mot lexicalisé, à court et à long terme. Nos résultats montrent également qu'un investissement cognitif plus important des apprenants sur un mot inconnu peut favoriser l'apprentissage du mot sur le plan formel mais non sur le plan sémantique.

Lexical inferencing is an important strategy in reading comprehension and vocabulary learning, both in L2 and in L1. Focusing both on the effect of L1 lexicalization (i.e., the existence of lexical equivalents in learner's L1 for a L2 word) and on that of learner's cognitive investment, we are interested in the process of inferring an unknown L2 word encountered in reading and the subsequent learning of the inferred word. We observed 33 French-speaking learners of Chinese while they inferred 20 unknown Chinese words contained in two passages and examined the evolution of their lexical knowledge of these words. The results show that the L1 lexicalization is not a critical factor for French-speaking learners to understand an unknown Chinese word. However, it is more difficult for them to retain the meaning of a non-lexicalized Chinese word than a lexicalized one, in the short and long term. Moreover, a more important cognitive involvement may help a learner obtain a better grasp of an unknown word, but only of its form, not its meaning.

INDEX

Keywords : lexical inferencing, lexicalization, lexical processing, vocabulary acquisition, retention, strategy, reading, reading comprehension, Chinese L2, input

Mots-clés : inférence lexicale, lexicalisation, traitement lexical, acquisition du vocabulaire, rétention, stratégie, lecture, compréhension écrite, chinois L2, input

AUTEURS

ENHAO LÉGER-ZHENG

Université de Toulouse Jean Jaurès
enhao.leger-zheng@univ-tlse2.fr

OLGA THÉOPHANOUS

Université de Toulouse Jean Jaurès
olga.theophanous@univ-tlse2.fr