Une approche psycholinguistique des phénomènes phraséologiques : le cas des expressions conventionnelles

Amanda Edmonds

▶ **To cite this version:**

Amanda Edmonds. Une approche psycholinguistique des phénomènes phraséologiques : le cas des expressions conventionnelles. Langages, 2013, 189, pp.121-138. hal-02879443

HAL Id: hal-02879443
https://hal.science/hal-02879443

Submitted on 9 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

PRE-PRINT VERSION
Edmonds, A. (2013). Une approche psycholinguistique des phénomènes phraséologiques: le cas des expressions conventionnelles. *Langages, 189,* 121-138.

Une approche psycholinguistique des phénomènes phraséologiques : le cas des expressions conventionnelles

Amanda EDMONDS
CRPHL (Centre de recherche en poétique et histoire littéraire)
Université de Pau et des Pays de l'Adour

Résumé de l'article
L'omniprésence d'unités phraséologiques dans toute langue découle, selon de nombreux chercheurs, de leur statut psycholinguistique tout à fait particulier ; de telles unités seraient stockées et extraites en bloc du lexique mental, ce qui rend leur traitement plus « facile » et/ou plus « rapide ». Ce présupposé psycholinguistique est généralement accepté, en dépit d'un manque général de preuves psychométriques. Le présent article tente de déterminer si un sous-ensemble d'unités phraséologiques – *les expressions conventionnelles* – jouit d'avantages psycholinguistiques pour un groupe de locuteurs natifs et deux groupes de non natifs du français. Pour ce faire, nous avons mis au point une expérience chronométrée, et l'analyse des temps de réaction suggère que les expressions conventionnelles ont en effet un corrélat mental pour les natifs comme pour les non natifs. Deux hypothèses (l'hypothèse lexicaliste et l'hypothèse de la compétence pragmatique) seront confrontées afin d'expliquer ces résultats.
Mots-clés : phraséologique, psycholinguistique, psychométrique, expression conventionnelle, français langue seconde

Abstract
The pervasiveness of phraseological units has been argued to follow from the fact that such sequences are stored holistically in the lexicon, implying that they are "easier" and/or "faster" to process. This psycholinguistic assumption is widespread in the literature, despite the fact that psychometric-based evidence for it is rare. This article attempts to address the question of whether the processing of a subset of phraseological units – *conventional expressions* – is associated with processing advantages for native and nonnative speakers of French. To this end, an online moving window task was created and administered to 20 natives and 40 nonnatives. The reaction time results suggested that conventional expressions do have a mental correlate for both groups, and two hypotheses (the lexicalist hypothesis and the pragmatic competence hypothesis) are compared in order to explain these findings.
Key words: phraseology, psycholinguistics, psychometrics, conventional expression, French as a second language

1. Introduction

L'intérêt porté aux phénomènes phraséologiques n'est pas nouveau ; ainsi « dès le quatrième siècle, dans son commentaire sur le Psaume 77 : 69, saint Augustin déclarait que des expressions telles que *in saeculum* pouvaient être traitées comme des unités » (Roberts, 1993, p. 36, citant Kelly, 1979). En linguistique, c'est surtout depuis une trentaine d'années que l'unité phraséologique attire de plus en plus l'attention des chercheurs. Cette importance croissante des unités phraséologiques – en linguistique générale aussi bien qu'en recherche sur l'acquisition des langues – est due, d'une part au développement de la linguistique de corpus, qui a permis d'analyser des masses de données et de constater de nombreuses régularités d'emploi (Kucera & Francis 1967; Sinclair 1991) et, d'autre part aux approches théoriques qui tendent à privilégier le lexique, parfois même au détriment de la grammaire (Fillmore, 1988 ; Jackendoff, 2002 ; Sinclair, 1991). Pour Legallois et Gréa (2006), nous pouvons caractériser cette période de « tournant phraséologique de la linguistique », dans la mesure où la phraséologie « propose un renouvellement de l'analyse syntaxique, sémantique, textuelle et

psycholinguistique sur la base d'observations plus empiriques, grâce, notamment, au recours aux corpus de grandes dimensions » (p. 5).

En recherche sur l'acquisition des langues, ce sont Pawley et Syder qui, dans leur article phare publié en 1983, furent parmi les premiers à souligner l'importance des unités phraséologiques pour les apprenants d'une deuxième langue. Ces auteurs ont identifié deux caractéristiques d'un parler natif, qui constituent un défi majeur pour tout apprenant : *nativelike selection* (le natif est capable de choisir une formulation à la fois grammaticale, appropriée et naturelle face à de nombreuses autres tournures possibles) et *nativelike fluency* (le natif est capable de produire un discours relativement long sans disfluences). Ces auteurs, comme beaucoup d'autres, expliquent cette compétence native en faisant appel à ce qu'ils supposent être la nature psycholinguistique bien spécifique des unités phraséologiques. Comme le disent Pawley et Syder, « la majorité, et de loin, du lexique d'un locuteur de l'anglais consiste en unités lexicales complexes, qui comprennent plusieurs centaines de milliers d'amorces de phrase lexicalisées » (p. 215)[1]. En d'autres termes, toute unité phraséologique aurait sa place dans le lexique mental comme toute unité lexicale simple (i.e., le mot), et ce serait donc le stockage et l'extraction en bloc d'unités plus grandes que le mot qui expliqueraient *nativelike selection* et *nativelike fluency* chez le natif. Moreau (1986) décrit cette spécificité psycholinguistique ainsi : « On aurait donc affaire à des combinaisons dotées d'un statut régulier du point de vue linguistique, mais d'un statut particulier du point de vue psycholinguistique, puisqu'elles ne semblent pas résulter d'une sélection réelle et d'un assemblage effectif d'unités » (p. 139).

Cette position, selon laquelle l'unité phraséologique aurait une inscription mémorielle, est répandue en linguistique générale comme en recherche sur l'acquisition des langues, et cela en dépit du fait que les expériences psycholinguistiques qui pourraient étayer cette position sont relativement rares – constat qui vaut pour les études sur les locuteurs natifs comme pour celles sur les locuteurs non natifs. Dans le présent article, nous présentons les résultats d'une étude psycholinguistique qui met à l'épreuve l'idée d'une inscription mémorielle pour une sous-partie du paysage phraséologique qui, jusqu'à présent, n'a pas été l'objet d'études psychométriques : les expressions conventionnelles. Plus concrètement, notre étude examine le traitement de telles expressions par un échantillon de locuteurs natifs et non natifs du français, et avance des conclusions concernant les éventuels avantages psycholinguistiques associés à ces expressions pour ces deux groupes de locuteurs.

2. Les unités phraséologiques

2.1. Définir et identifier les unités phraséologiques

Pendant longtemps, l'unité phraséologique a été surtout conçue comme une unité lexicologique, et les combinaisons étaient généralement étudiées pour être répertoriées dans les dictionnaires (ex. Heinz, 1997, 2003 ; Lamiroy, 2003). De nos jours, la phraséologie – et, par conséquent, l'unité phraséologique – ne se restreint plus à cette tradition lexicologique, mais elle s'est également étendue à la linguistique générale et à la recherche sur l'acquisition des langues. Depuis cette extension, la terminologie, les définitions et les critères d'identification qui lui sont propres n'ont cessé de se multiplier et de se complexifier (voir la discussion de la *polyfactorialité* dans la contribution de Schmale dans le présent numéro). Dans une tentative récente de démêler ce qu'elles ont baptisé « l'écheveau phraséologique »[2], Granger et Paquot (2008) ont souligné la prédominance de deux approches dans la recherche en phraséologie : les approches phraséologiques (ou traditionnelles) et les approches distributionnelles. Selon ces auteurs, une approche traditionnelle définit une unité phraséologique grâce à des critères

[1] "by far the largest part of the English speaker's lexicon consists of complex lexical items including several hundred thousand lexicalized sentence stems." (la traduction est la nôtre)

[2] Le titre du chapitre est *Disentangling the phraseological web*.

syntaxiques, sémantiques et/ou pragmatiques (voir, par exemples, les études de Tutin et de Dalmas & Gautier dans ce numéro), alors qu'une approche distributionnelle se fonde surtout sur des informations statistiques (ex. la fréquence d'une séquence ou des mesures de la force de la cohésion entre une suite de mots[3]), qui peuvent être obtenues grâce à l'utilisation de corpus en linguistique (l'étude de Longrée & Mellet, ce numéro, en constitue un exemple). Il est pourtant intéressant de noter que, indépendamment de l'approche adoptée, la nature psycholinguistique (à savoir, la représentation mentale) attribuée à une unité phraséologique est devenue une préoccupation centrale en phraséologie, jusqu'à constituer le critère clé de la définition d'*unité phraséologique* pour nombre de spécialistes. À titre d'exemple, la définition de *formulaic sequence* proposée par Wray (2002, p. 9), définition dominante dans la recherche phraséologique anglophone, définit une *séquence formulaïque* comme « une séquence, continue ou discontinue, de mots ou d'autres éléments, qui est ou a l'air d'être préfabriquée ; c'est-à-dire, stockée et extraite en bloc de la mémoire lors de son emploi plutôt que d'être fabriquée ou analysée par la grammaire »[4].

Selon une telle définition, l'objectif n'est plus de découper du langage en *expressions idiomatiques*, *collocations* et *combinaisons libres*, comme c'était traditionnellement le cas dans les études lexicologiques, mais plutôt de distinguer les suites de mots qui seraient stockées dans la mémoire en tant que telles et les suites de mots « fabriquée[s] ou analysée[s] par la grammaire ». Or, s'il est évident que le repérage de cette distinction psycholinguistique – qui n'est pas directement observable – pose d'importants problèmes méthodologiques à tout linguiste qui s'y prête~~. Pourtant~~, dans la pratique, ces problèmes ne sont que rarement évoquées.~~, puisque~~ En général, deux caractéristiques psycholinguistiques sont communément acceptées comme preuves (quoique indirectes) d'une inscription mémorielle, à savoir la rapidité et la facilité de traitement (*processing*). En d'autres termes, une séquence formulaïque serait appréhendée et produite plus rapidement qu'une séquence semblable mais sans inscription mémorielle, et le coût cognitif imposé par la compréhension ou l'emploi de cette séquence serait moindre. Ces supposés avantages psycholinguistiques sont évoqués régulièrement dans la littérature, ce qui est illustré dans les citations suivantes pour le domaine de la recherche sur l'acquisition des langues :

(1) On compte sur de tels blocs afin d'atténuer les problèmes de traitement. On les utilise pour 'gagner' du temps pendant que d'autres processus de traitement se déroulent, ce qui permet de planifier le contenu de ce que l'on va dire, ainsi que sa forme linguistique. (Skehan, 1998, p. 40)[5]

> Dans le stock de collocations familières, on trouve des expressions pour un large éventail de concepts familiers et d'actes de langage, et le locuteur est capable de les extraire de la mémoire à long terme sous forme de blocs ou de chaines automatiques; il minimise ainsi la charge de travail au niveau de l'encodage à l'intérieur de chaque proposition et se rend disponible pour effectuer d'autres tâches dans l'échange, y compris planifier des unités du discours plus grandes. (Pawley & Syder, 1983, p. 192)[6]

[3] Deux mesures sont utilisées fréquemment : les scores MI (*mutual information*) et les *t*-scores (Church & Hanks, 1990 ; Gries, 2010).

[4] "sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at time of use, rather than being subject to generation or analysis by the language grammar." (la traduction est la nôtre)

[5] "We rely on such chunks to ease processing problems, using them to 'buy' processing time while other computation proceeds, enabling us to plan ahead for the content of what we are going to say, as well as the linguistic form." (la traduction est la nôtre)

[6] "In the store of familiar collocations there are expressions for a wide range of familiar concepts and speech acts, and the speaker is able to retrieve these as wholes or as automatic chains from the long term memory; by doing

> Une grande majorité des concepts et des actes de langage les plus familiers peut être exprimée d'une manière formulaïque, et si un locuteur peut extraire ces séquences en bloc facilement de la mémoire, la fluence en est améliorée. (Wood, 2002, p. 7)[7]

Les deux caractéristiques de rapidité et de facilité (qui découleraient d'un stockage dans le lexique mental) ont conduit au postulat que certaines conséquences concrètes et surtout quantifiables de ces caractéristiques devraient se manifester, et celles-ci sont alors devenues des critères permettant d'identifier les séquences formulaïques. Ainsi, la rapidité et la facilité du traitement de ces séquences amèneraient un locuteur à les utiliser plus fréquemment (critère 2a), et leur utilisation accorderait plus de temps au locuteur pour la planification de son discours (critère 2b). De surcroît, plusieurs auteurs ont émis l'hypothèse selon laquelle le stockage d'une séquence en tant qu'item lexical influe sur sa forme et sur sa prononciation, dans la mesure où cette suite de mots a une forme invariable (critère 2c) et est prononcée sans disfluences (critère 2d). Ainsi, on obtient les quatre critères d'identification suivants :

(2) Critères d'identification
- a. Fréquence élevée
- b. Temps accru pour la planification du discours
- c. Forme invariable
- d. Prononciation sans disfluences

Si ces critères sont en effet fréquemment évoqués, il n'existe toutefois pas de consensus sur leur opérationalisation. En effet, c'est notamment l'interprétation de ce qui constitue une fréquence *élevée* ou une forme *invariable* qui varie d'un auteur à l'autre.

2.2. L'unité phraséologique : un phénomène psycholinguistique à part ?

Si de nombreux chercheurs acceptent qu'une séquence fréquente, dotée d'une forme invariable et prononcée sans disfluences soit stockée et extraite en bloc de la mémoire, le développement assez récent de plusieurs outils psychométriques a permis de mettre à l'épreuve les suppositions ~~fondamentales~~ concernant la nature psycholinguistique des séquences formulaïques, notamment par rapport à la rapidité de traitement accrue qui accompagnerait toute représentation holistique. Un grand nombre d'expériences a essayé de démontrer qu'une séquence dite formulaïque (et, donc, stockée et extraite en bloc) est en effet appréhendée plus rapidement qu'une séquence similaire mais non phraséologique. Ces études ont, pour la plupart, ciblé deux sous-ensembles parmi les nombreux types de séquences formulaïques, à savoir les expressions idiomatiques et les séquences qui, dans un corpus, ont une fréquence élevée et/ou une forte cohésion interne.

Le premier sous-ensemble – les expressions idiomatiques – correspond à toute séquence d'au moins deux mots qui se reconnaît soit par une syntaxe irrégulière et non compositionnelle *(sans mot dire)*, soit par une sémantique opaque *(lui donner le bon Dieu sans confession)*, soit par les deux *(sans coup férir)*. Dans le paysage phraséologique, ce sont les seules suites qui sont unanimement considérées comme étant stockées en bloc, leur non compositionnalité caractéristique fournissant une preuve particulièrement convaincante de leur mémorisation holistique. Les nombreuses expériences faites avec des locuteurs natifs (surtout de l'anglais) ont en général étayé cette conviction, révélant une rapidité de réaction significative face aux expressions idiomatiques par comparaison avec des expressions semblables mais non idiomatiques (Cacciari & Tabossi, 1993 ; Cronk & Schweigert, 1992 ; Nenonen, Niemi, & Laine, 2002 ; Qualls, Treaster, Blood, & Hammer, 2003 ; Swinney & Cutler, 1979, etc.). En

---

this he minimizes the amount of clause-internal encoding work to be done and frees himself to attend to other tasks in talk-exchange, including the planning of larger units of discourse." (la traduction est la nôtre)

[7] "A great proportion of the most familiar concepts and speech acts can be expressed formulaically, and if a speaker can pull these readily from memory as wholes, fluency is enhanced." (la traduction est la nôtre)

revanche, les résultats pour les apprenants d'une langue seconde sont moins concluants. Ainsi, deux études ont identifié une rapidité significative dans les réactions des non natifs aux expressions idiomatiques (Conklin & Schmitt, 2008 ; Underwood, Schmitt, & Galpin, 2004), alors que Schmitt et Underwood (2004) et Siyanova-Chanturia, Conklin, et Schmitt (2011) n'ont trouvé aucune asymétrie significative dans les temps de réaction de leurs locuteurs non natifs.

Le deuxième sous-ensemble objet des études psychométriques correspond aux séquences identifiées à l'aide des méthodes employées par les approches distributionnelles de la phraséologie (voir Granger & Paquot, 2008). Dans ces études, il est question de relever les suites de mots les plus fréquentes et/ou ayant une cohésion interne particulièrement forte. Étant donné que les défenseurs de la grammaire émergente (comme Bybee & Hopper, 2001 ou bien Tomasello, 2000) suggèrent que la fréquence élevée de certaines séquences peut donner lieu à un stockage et une extraction en bloc, il n'est pas surprenant que plusieurs auteurs tentent de déterminer si la rapidité des réactions varie en fonction du profil distributionnel d'une séquence. Considérons tout d'abord les séquences ayant une fréquence élevée. Il a été démontré que les locuteurs natifs mais aussi les non natifs répondent à de telles suites de mots plus rapidement qu'aux séquences moins fréquentes (natifs : Siyanova & Schmitt, 2008 ; Siyanova-Chanturia, Conklin, & van Heuven, 2011 ; non natifs : Ellis, Simpson-Vlach, & Maynard, 2008 ; Jiang & Nekrasova, 2007). En revanche, les natifs et les non natifs ne semblent pas se comporter de façon identique face aux séquences identifiées à l'aide des mesures qui identifient la force de cohésion d'une suite de mots (scores MI et *t*-scores). Ellis et al. (2008) ont trouvé que la force de cohésion interne d'une séquence était une variable significative dans leur modèle de régression qui explique la rapidité des réponses des natifs, alors que cette même variable n'était pas significative dans l'analyse des temps de réaction des non natifs. Ces résultats, considérés conjointement, semblent indiquer que les locuteurs natifs et les locuteurs non natifs sont tous deux sensibles au profil distributionnel des suites de mots : les temps de réaction des natifs révèlent qu'ils sont sensibles et à la fréquence et à la force de la cohésion d'une séquence, alors que les non natifs ne sont sensibles qu'à la fréquence.

Les études passées en revue, représentatives de la recherche à la fois psychométrique et phraséologique, ont choisi d'examiner le traitement d'expressions idiomatiques ou de séquences définies en fonction d'un profil distributionnel particulier. Toutefois, les phénomènes phraséologiques ne se limitent pas à ces seules séquences, et l'application de la définition de Wray (2002) et des quatre critères d'identification mentionnés précédemment (2a-d) conduit à la prise en compte d'une grande variété de séquences. Il paraît donc naturel et souhaitable d'élargir l'objet de telles études, ce qui permettra de déterminer si l'ensemble des séquences phraséologiques bénéficie des mêmes avantages psycholinguistiques que les expressions idiomatiques et les séquences à fréquence élevée ou à forte cohésion, et ainsi de juger de la véracité de la définition purement psycholinguistique de Wray pour l'ensemble des unités phraséologiques. Notre projet vise à contribuer à cette discussion, avec une expérience psychométrique qui nous a permis d'évaluer si la rapidité que l'on associe aux phénomènes phraséologiques est à l'œuvre dans le traitement d'un échantillon d'expressions phraséologiques tout à fait compositionnelles, une expérience que nous avons proposée à des locuteurs natifs et non natifs du français.

3. Les expressions conventionnelles jouissent-elles d'avantages psycholinguistiques ?

3.1. Les expressions conventionnelles : les séquences ciblées

Pour ce projet, nous avons choisi d'examiner un sous-ensemble du paysage phraséologique, à savoir les *expressions conventionnelles* (Bardovi-Harlig, 2009). Appelées *énoncés liés* par Fónagy (1998) et *routines linguistiques* par Coulmas (1979), de telles suites se distinguent des autres unités phraséologiques de par leur association à une situation sociale ou à un contexte

communicationnel précis[8]. En d'autres termes, nous avons affaire ici à un phénomène qui est à la fois phraséologique et pragmatique. Dans le cadre de ce projet, nous avons eu recours à un questionnaire afin d'identifier des expressions susceptibles d'être conventionnelles dans la communauté où s'est déroulé le projet (dans le Sud-Ouest de la France). Un questionnaire de trente-cinq contextes a été proposé aux participants. Pour chacun de ces contextes, le participant devait se mettre dans la situation décrite et, à la fin, une phrase inachevée qui finissait par un deux-points (ex. « Tu lui dis : ») invitait le participant à répondre comme il l'aurait fait dans une situation de communication réelle. Les consignes encourageaient les participants à donner jusqu'à quatre réponses pour chaque contexte, et les 86 locuteurs natifs qui ont rempli le questionnaire ont fourni au total entre 116 et 179 réponses par contexte. Nous avons tenté, grâce à l'analyse de ces réponses, de dégager les expressions conventionnelles à l'aide de quatre critères d'identification. Pour être considérée comme une expression conventionnelle, une séquence devait comporter plus d'un mot, avoir une forme invariable[9], jouir d'une fréquence élevée parmi toutes les réponses au contexte en question et les éléments de la séquence devaient entretenir une relation syntaxique entre eux (ce qui exclut des suites telles que *et le*). Cette analyse a identifié 31 expressions conventionnelles parmi les réponses des natifs, dont 13 ont été retenues pour l'expérience psychométrique.

3.2. La création d'une expérience psychométrique

Une expérience chronométrée a ensuite été créée, qui avait pour objectif de répondre à la question de recherche suivante :

(3) Les expressions conventionnelles jouissent-elles d'avantages psycholinguistiques aussi bien chez les locuteurs natifs que chez les non natifs ?

Dans le cadre de ce projet, le terme d'*avantages psycholinguistiques* renvoie à la rapidité de traitement, ce qui a été examiné à l'aide d'une expérience psychométrique bâtie autour de deux comparaisons de temps de réaction. Afin de rendre les deux comparaisons possibles, nous avons manipulé chacune des 13 expressions conventionnelles de deux manières différentes : d'abord, nous avons procédé à une manipulation de *mot* et, ensuite, à une manipulation de *cadre*. Pour la manipulation de mot, un mot de chacune des 13 expressions conventionnelles (*mot original*) a été remplacé par un synonyme proche (*substitut*), ce qui nous a permis de comparer les temps de réaction pour le mot original et son substitut dans une expression conventionnelle. Si une telle expression est stockée et extraite en bloc, il est logique que les mots qui la composent soient lus plus rapidement que les substituts qui remplacent ces mots. Pour la manipulation de cadre, le mot original de chaque expression conventionnelle a été inséré dans un cadre alterné, qui était à la fois grammatical et sémantiquement approprié sans que l'expression qui en résulte ait été identifiée comme étant conventionnelle dans la première phase de ce projet (le questionnaire). De cette manière, nous avons pu comparer le temps de réaction au mot original dans un cadre conventionnel à celui relevé dans un cadre alterné. Si une expression conventionnelle est bel et bien stockée et extraite en bloc, on peut s'attendre à ce que, lorsqu'un mot appartient à une expression conventionnelle, le temps de réaction à ce mot soit plus rapide que lorsque ce même mot se trouve dans un cadre alterné (et non conventionnel). Les manipulations de mot et de cadre, appliquées à chacune des 13 expressions conventionnelles, ont généré 52 items expérimentaux ; nous y avons ajouté 28 distracteurs, pour un total de 80 items. Ces manipulations de mot et de cadre sont illustrées dans le tableau 1 à partir d'une des 13 expressions conventionnelles testées (*c'est gentil*).

[8] Cette description se distingue de celle de certains auteurs, tels que Schmale (ce numéro), qui considèrent que toute unité phraséologique a un ancrage pragmatique.

[9] Dans la pratique, une certaine souplesse dans l'opérationalisation de *forme variable* est souvent acceptée, ce qui a été également le cas dans ce projet. Par exemple, nous avons considéré que les suites *je sais pas* et *je ne sais pas* avaient la même forme.

Tableau 1. *Les manipulations de mot et de cadre*

	Cadre	
Mot	Conventionnel	Alterné
Original	Merci, / [c'est / **gentil**] / de / votre part	C'est / bien / **gentil** / à vous / de me le / proposer, / merci.
Substitut	Merci, / c'est / **aimable** / de / votre part	C'est / bien / **aimable** / à vous / de me le / proposer, / merci.

Note. Les [] délimitent l'expression conventionnelle ; le mot original et le substitut sont en gras.

Il s'agissait ensuite de découper chaque item en segments, car il était essentiel, pour l'expérience chronométrée, que les mots originaux et les substituts soient présentés séparément pour que la longueur du temps de réaction à ces segments puisse être mesurée (en millisecondes). Dans le tableau 1, les barres obliques indiquent le découpage des items. À l'aide du logiciel *Linger*[10], une expérience psychométrique a été conçue et présentée sur ordinateur. Il s'est agi de reprendre les contextes utilisés dans la première phase du projet (le questionnaire) et de proposer comme réponse pour chacun d'entre eux un des 80 items cités plus haut. Après la présentation du contexte, le premier segment de la réponse appariée à ce contexte apparaissait. Le participant contrôlait l'apparition des segments suivants en appuyant sur la barre d'espacement : à chaque appui, le segment présent à l'écran disparaissait et le segment suivant se présentait. Après avoir lu la réponse en entier, le participant devait décider si oui ou non la réponse était naturelle dans le contexte. Le temps écoulé entre chaque appui de la barre d'espacement était enregistré par *Linger*, et ce sont les données ainsi collectées qui ont servi de variable dépendante dans notre analyse statistique.

3.3. Les participants

Au total, 60 participants ont accepté de prendre part à cette expérience psycholinguistique, dont 20 locuteurs natifs du français et 40 locuteurs anglophones vivant en France. Tous les participants vivaient dans le Sud-Ouest de la France au moment de prendre part à ce projet ; aucun n'avait participé à la première phase de l'étude (le questionnaire). Les non natifs comportaient deux groupes de 20 participants chacun : un groupe de court séjour dont les membres étaient dans le Sud-Ouest depuis entre 4 et 6 mois (M = 5 mois) et un groupe de long séjour dont les locuteurs y vivaient depuis entre 1 an et 33 ans (M = 10,5 ans). Tous les participants anglophones avaient appris le français à l'âge adulte et ne parlaient que l'anglais pendant leur enfance. Les non natifs de court séjour avaient, en moyenne, 26,5 ans, alors que les non natifs de long séjour étaient dans l'ensemble plus âgés (M = 41 ans). Le groupe le plus jeune était celui des locuteurs natifs, qui étaient, en moyenne, âgés de 18,4 ans.

3.4. L'analyse

L'objectif de ce projet de recherche était de déterminer si les expressions conventionnelles étudiées jouissent d'avantages psycholinguistiques chez des locuteurs natifs ainsi que chez des locuteurs non natifs du français. Grâce à la manipulation de deux variables, trois cas de figure peuvent constituer des preuves d'avantages psycholinguistiques (sous forme d'une asymétrie significative dans la rapidité de traitement) dans notre expérience, et chacun d'entre eux était examiné dans l'analyse statistique effectuée. Les trois scénarii sont présentés en (4a-c) :

(4a) Premier scénario : manipulation de mot

Un temps de réaction plus rapide au mot original qu'au substitut dans le cadre conventionnel (ex. *c'est **gentil*** < *c'est **aimable***) est interprété comme étant la preuve d'une différence significative de rapidité en faveur des expressions conventionnelles, à condition seulement que cette même asymétrie ne soit pas avérée dans un cadre alterné (ex. *c'est bien **gentil** à vous* ≥ *c'est bien **aimable** à vous*). En effet, ces deux conditions

[10] Le logiciel *Linger* est disponible à l'URL suivant : <*http://tedlab.mit.edu/~dr/Linger/*>

sont nécessaires afin de s'assurer que l'asymétrie ne soit pas due au fait que le temps de réaction au mot original est toujours plus rapide que celui enregistré pour le substitut, et ce quel que soit le cadre.

(4b) Deuxième scénario : manipulation de cadre
Les expressions testées bénéficient d'avantages psycholinguistiques si deux paramètres sont réunis : 1) lorsque le mot original appartient à un cadre conventionnel, le temps de réaction à ce mot est plus rapide que lorsqu'il est situé dans le cadre alterné (ex. *c'est* ***gentil*** < *c'est bien* ***gentil*** *à vous*) ; 2) la même asymétrie ne caractérise pas les temps de réaction au substitut (ex. *c'est* ***aimable*** ≥ *c'est bien* ***aimable*** *à vous*). Le respect des deux conditions garantit que l'asymétrie ne découle pas d'un cadre qui confère des avantages psycholinguistiques à tout item lexical (mot original ou substitut) qui s'y trouve.

(4c) Troisième scénario : inhibition
Le troisième cas de figure concerne cette fois-ci les temps de réaction à un substitut placé dans un cadre conventionnel. Plus précisément, la présence d'un substitut dans un cadre conventionnel peut surprendre un locuteur, dans la mesure où on pourrait logiquement s'attendre à trouver l'expression conventionnelle (et, donc, le mot original). Sur le plan psycholinguistique, cet « effet de surprise » conduirait à un temps de réaction plus élevé au substitut dans un cadre conventionnel, par comparaison avec le temps de réaction au mot original dans ce même cadre (ex. *c'est* ***gentil*** < *c'est* ***aimable***) et au temps de réaction au substitut dans un cadre alterné (ex. *c'est bien* ***aimable*** *à vous* < *c'est* ***aimable***).

Ces différents cas sont explicités dans le tableau 2.

Tableau 2. *Rapidité significative : trois cas de figure*

Cas de figure	Résultats psychométriques		
Manipulation de mot	Mot original (cadre conventionnel)	<	Substitut (cadre conventionnel)
		&	
	Mot original (cadre alterné)	≥	Substitut (cadre alterné)
Manipulation de cadre	Mot original (cadre conventionnel)	<	Mot original (cadre alterné)
		&	
	Substitut (cadre conventionnel)	≥	Substitut (cadre alterné)
Inhibition	Mot original (cadre conventionnel)	<	Substitut (cadre conventionnel)
		&	
	Substitut (cadre alterné)	<	Substitut (cadre conventionnel)

Ces différents scénarios ont été examinés à l'aide de plusieurs analyses de variance à mesures répétées, d'abord sur l'ensemble des temps de réaction et ensuite sur les données pour chacune des 13 expressions conventionnelles. Toute analyse significative a été suivie par des tests post-hoc.

3.5. Les résultats

3.5.1. Les résultats globaux

L'analyse de variance de l'ensemble des données a révélé que les locuteurs répondaient plus rapidement aux mots originaux qu'aux substituts ($F[1, 57] = 42.232$, $p < .001$), qu'il n'y avait pas de différence globale dans les temps de réaction en fonction du cadre ($F[1, 57] = 1.833$, $p = .181$), mais que l'interaction entre les variables de *mot* et de *cadre* était significative ($F[2, 57] = 17.27$, $p < .001$). Cette première analyse a également montré que le profil de traitement des natifs différait de celui des deux groupes de non natifs, tandis que la comparaison des non

natifs de court séjour et des non natifs de long séjour a révélé que ces deux groupes se comportaient de manière semblable. À la lumière de ce résultat, nous avons fusionné les deux groupes de non natifs dans les analyses subséquentes. Les moyennes des temps de réaction sont présentées dans le tableau 3[11].

Table 3. *Les moyennes des temps de réaction en millisecondes*

	Cadre conventionnel		Cadre alterné	
Groupe	Mot original	Substitut	Mot original	Substitut
Non natifs – court séjour	756.78 (298)	1085.41 (466)	790.07 (284)	981.73 (382)
Non natifs – long séjour	713.49 (219)	913.46 (289)	656.44 (171)	806.86 (274)
Natifs	479.9 (129)	518.53 (150)	502.82 (103)	518.32 (116)

Note. Les écarts types sont entre parenthèses

L'interaction significative entre la variable de *mot* et la variable de *cadre* pourrait être due à un ou plusieurs des trois cas de figure exposés précédemment ; ceci sera examiné dans les sections qui suivent à l'aide des tests post-hoc.

3.5.2. Premier cas de figure : manipulation de mot

En ce qui concerne le premier cas de figure, ni les temps de réaction des locuteurs natifs, ni ceux des deux groupes de non natifs n'ont conforté l'hypothèse selon laquelle les expressions conventionnelles jouiraient d'avantages psycholinguistiques. Chez les natifs, il n'y avait pas de différence significative, dans les cadres conventionnels, entre les temps de réaction pour les mots originaux et pour les substituts (t[19] = .02, p = .984) ; au contraire, les non natifs étaient plus rapides pour les mots originaux que pour les substituts dans ces mêmes cadres (t[39] = 7.973, p < .001). Toutefois, cette asymétrie chez les non natifs ne semble pas être due à un traitement plus rapide des expressions conventionnelles, puisque ces locuteurs ont tout simplement *toujours* réagi aux mots originaux plus rapidement, que le cadre soit conventionnel ou alterné (t[39] = 4.164, p < .001).

3.5.3. Deuxième cas de figure : manipulation de cadre

Les tests post-hoc ont démontré que les natifs répondaient de manière plus rapide quand ils lisaient un mot original dans son cadre conventionnel que lorsqu'ils lisaient ce même mot dans un cadre alterné (t[19] = 3.307, p < .01) ; cependant, les temps de réaction au substitut ne variaient pas en fonction du cadre (t[19] = .158, p = .876). Ce résultat suggère ainsi une rapidité plus importante dans le traitement des expressions conventionnelles chez les natifs. Chez les non natifs, aucune asymétrie significative n'a été trouvée en lien avec la manipulation de cadre.

3.5.4. Troisième cas de figure : inhibition

Aucune trace d'une inhibition pour les substituts dans les cadres conventionnels n'a été identifiée dans les réactions des natifs. En revanche, les non natifs, eux, réagissaient plus lentement au substitut dans son cadre conventionnel qu'au substitut dans un cadre alterné (t[39] = 3.803, p < .001). Ils étaient également plus lents dans leurs réactions au substitut en comparaison au mot original dans son cadre conventionnel (t[39] = 7.973, p < .001). Ces résultats, considérés conjointement, indiquent une inhibition claire dans les temps de réaction des non natifs au substitut inséré dans une expression conventionnelle.

3.5.5. Résumé des résultats

Les résultats psychométriques de cette expérience ont en effet mis en évidence des avantages psycholinguistiques pour l'échantillon d'expressions conventionnelles testées par comparaison avec les expressions non phraséologiques, résultats qui étaient valables pour les natifs tout autant que pour les non natifs. Nous pouvons donc répondre à notre question initiale par

[11] Pour toute analyse statistique, nous avons utilisé les résiduels et non pas les temps de réaction en millisecondes comme variable dépendante. Ce choix nous a permis de contrôler la différence de longueur entre les mots originaux et les substituts.

l'affirmative : l'échantillon d'expressions conventionnelles testées dans cette expérience jouit d'avantages psycholinguistiques chez les participants natifs et non natifs. Toutefois, ces avantages psycholinguistiques se manifestent différemment chez les locuteurs natifs et les locuteurs non natifs. Alors que les francophones réagissent au même mot plus rapidement quand celui-ci se trouve dans un cadre conventionnel (par opposition à un cadre alterné), les temps de réaction enregistrés pour les non natifs indiquent un phénomène d'inhibition dans le traitement d'un substitut placé dans un cadre conventionnel.

4. Discussion

Si les avantages psycholinguistiques identifiés dans l'expérience décrite sont certes compatibles avec le stockage et l'extraction en bloc du lexique mental, qui forment la base de la définition de *séquence formulaïque* de Wray (2002), cette compatibilité n'équivaut pourtant pas à une preuve formelle d'une inscription mémorielle. Dans cette dernière section, nous nous interrogerons sur la nature du corrélat mental associé aux expressions conventionnelles en comparant deux possibles interprétations des avantages psycholinguistiques découverts : l'hypothèse lexicaliste et l'hypothèse de la compétence pragmatique.

L'hypothèse lexicaliste correspond à l'explication traditionnelle, selon laquelle les unités phraséologiques sont des phénomènes lexicaux stockés dans le lexique mental. Si cette position ne rencontre guère de détracteurs quand il s'agit des expressions idiomatiques, dont la non compositionnalité plaide en faveur de la mémorisation, l'extension de cette hypothèse à des séquences tout à fait compositionnelles est plus récente et va de pair avec l'importance croissante du lexique mental par rapport à la grammaire, importance soulignée par Bolinger (1976, p. 2) quand il écrit : « les locuteurs dépendent au moins autant de leur mémoire que de leur capacité d'assemblage »[12]. Au vu des résultats de notre étude, l'hypothèse lexicaliste nous semble problématique pour deux raisons, l'une générale, l'autre spécifique aux unités phraséologiques dont il est question dans ce projet. La première remarque, d'ordre général, concerne les suppositions qui sous-tendent cette hypothèse et son application. En effet, les chercheurs qui font appel à l'hypothèse lexicaliste ont tendance à considérer les avantages psycholinguistiques (qui correspondent, le plus souvent, à une rapidité relative) comme étant une indication claire et incontestable d'une représentation mentale holistique (ex. Jiang & Nekrasova, 2007, p. 442 ; Qualls et al., 2003, p. 255 ; Underwood et al., 2004, p. 167). Suivant cette même logique, nous pourrions être tentée de déclarer que les résultats présentés dans cet article constituent une preuve formelle d'une représentation mentale holistique pour les expressions conventionnelles testées, interprétation qui serait, à nos yeux, quelque peu précipitée ; en effet, si l'association entre mémorisation et rapidité est logique, elle demeure tout de même supposée. Cette association supposée n'est d'ailleurs pas exclusive dans le sens où il peut y avoir d'autres explications pour une rapidité relative dans les temps de réaction (ex. une plus grande automatisation dans l'appréhension des séquences fréquentes – phénomène qui peut être tout à fait indépendant d'une inscription mémorielle).

Notre seconde remarque met en doute la capacité de l'hypothèse lexicaliste à tenir compte de la particularité des expressions conventionnelles, à savoir leur association à une situation sociale ou à un contexte communicationnel précis. De Cock (1998), dans sa discussion de la suite *you see* en anglais, souligne bien ce problème d'une seule et même séquence qui serait phraséologique dans certains contextes (*well, **you see**, I knew this guy*), mais qui dans d'autres, n'a rien d'une unité phraséologique (*do **you see** the dog?*). Sur le plan psycholinguistique, on s'attendrait à ce que *you see* jouisse d'avantages psycholinguistiques associés à l'unité phraséologique dans le premier exemple mais non dans le deuxième. Pourtant, l'hypothèse lexicaliste, qui place les expressions conventionnelles dans le lexique

[12] "speakers do at least as much remembering as they do putting together."

mental, aurait des difficultés à expliquer pourquoi, dans les contextes où la séquence n'est pas phraséologique, un locuteur fabriquerait systématiquement une suite de mots identique à une séquence qu'il a déjà stockée en bloc dans son lexique mental, démarche qui prendrait plus de temps et entraînerait un coût cognitif plus important. Ceci semble particulièrement peu probable quand on considère ce que l'on sait sur le traitement des expressions idiomatiques auxquelles l'on peut attribuer deux interprétations, l'une figurée (et, donc, non compositionnelle et holistique) et l'autre littérale. Les études de Colombo (1993) et de Peterson, Burgess, Dell et Eberhard (2001) ont montré que l'interprétation figurée (et holistique) d'une expression idiomatique est *toujours* activée, qu'il s'agisse d'un contexte privilégiant une interprétation figurée ou d'un contexte favorisant une interprétation littérale, résultat qui suggère que, une fois stockée dans le lexique mental, la représentation holistique d'une unité phraséologique sera mobilisée à chaque fois qu'il est question de la suite qui la compose. L'application de ce résultat aux expressions conventionnelles laisse supposer que, si celles-ci sont en effet stockées dans le lexique mental, leur représentation holistique sera mobilisée à chaque fois que la suite en question sera appréhendée ou produite, peu importe le contexte. Pour revenir à l'exemple de De Cock, *you see* bénéficierait donc d'avantages psycholinguistiques aussi bien lorsqu'il s'agit d'une unité phraséologique – *well, you see, I knew this guy* – que lorsque cette suite n'a rien de phraséologique – *do you see the dog ?*

Ces réflexions nous ont amenée à éclairer nos résultats d'expérience à la lumière de ce que nous appelons l'hypothèse de la compétence pragmatique. Au lieu de placer les expressions conventionnelles directement dans le lexique mental, nous nous demandons s'il ne serait pas plus cohérent d'expliquer les avantages psycholinguistiques décelés dans cette étude pour de telles séquences en tenant compte de la compétence pragmatique de chaque locuteur Notre compétence pragmatique nous permet de reconnaître, d'interpréter et d'encoder de nombreux actes de langage. Pourquoi cette compétence ne régirait-elle pas également les associations entre un acte de langage et des expressions conventionnelles ? Une fois un acte de langage reconnu (en production ou en compréhension), toute expression conventionnelle pouvant réaliser l'acte en question serait activée, ce qui expliquerait les avantages psycholinguistiques associés à de telles séquences, avantages qui se limiteraient aux contextes déclencheurs.

Départager l'hypothèse lexicaliste et l'hypothèse de la compétence pragmatique nécessiterait une nouvelle expérience psychométrique qui viserait à évaluer si le contexte est réellement déterminant dans le traitement d'expressions conventionnelles. Si des avantages psycholinguistiques sont détectés pour les expressions conventionnelles quel que soit le contexte dans lequel elles se trouvent, on aurait une preuve plus concluante du stockage et de l'extraction en bloc du lexique mental pour les expressions conventionnelles. Par contre, si les avantages psycholinguistiques démontrés dans la présente étude sont bel et bien limités à des contextes déclencheurs, et ne caractérisent pas les expressions conventionnelles en toute circonstance, nous pourrions conclure en faveur de l'hypothèse de la compétence pragmatique.

5. Conclusion

L'unité phraséologique a-t-elle une validité psycholinguistique ? Alors que de nombreuses définitions, de multiples termes et maints critères d'indentification reposent sur une réponse affirmative à cette question, rares sont les études qui ont tenté de vérifier ce qui peut sembler si logique. Dans le cadre de la présente étude, nous nous sommes efforcée de déterminer si un sous-ensemble d'unités phraséologiques – un sous-ensemble jusqu'à présent négligé dans les études psychométriques – pourrait jouir d'avantages psycholinguistiques. À la différence de la plupart des études psychométriques en phraséologie, nous avons choisi d'étudier le traitement d'un échantillon d'*expressions conventionnelles*, expressions qui sont en général parfaitement compositionnelles et qui sont associées à une situation ou à un contexte précis. Les résultats suggèrent que de telles expressions ont effectivement un corrélat mental, dans le sens où les

profils de traitement associés à ces expressions se distinguent de manière significative de ceux associés à des séquences non phraséologiques, pour les participants natifs comme pour les participants non natifs. Ce corrélat mental se manifeste par une rapidité plus importante pour les expressions conventionnelles chez les natifs et par une inhibition pour les substituts chez les non natifs. Deux interprétations de ces résultats ont été présentées et analysées, et nous avons exprimé nos réserves sur la position traditionnelle – l'hypothèse lexicaliste – à deux niveaux, l'un général, l'autre spécifique au traitement d'expressions conventionnelles. Nous avons par la suite proposé qu'une deuxième hypothèse – l'hypothèse de la compétence pragmatique – serait mieux à même d'expliquer les résultats, ce qui nous permettrait de plaider en faveur d'une nouvelle extension du domaine de la phraséologie au-delà du lexique et de remettre en question la définition purement psycholinguistique proposée par Wray (2002) pour l'ensemble des *formulaic sequences*. Toutefois, les données récoltées dans cette étude ne nous permettent pas de trancher de manière définitive entre une représentation mentale holistique des expressions conventionnelles et une activation liée à la compétence pragmatique de chaque locuteur afin d'expliquer les avantages psycholinguistiques observés. Une réponse définitive à cette question devra donc attendre un autre jour et surtout une nouvelle étude.

Références

BARDOVI-HARLIG K. (2009), "Conventional expressions as a pragmalinguistic resource: Recognition and production of conventional expressions in L2 pragmatics", *Language Learning* 59, 755-795.

BOLINGER D. (1976), "Meaning and memory", *Forum Linguisticum* 1, 1-14.

BYBEE J. & HOPPER P. (2001), *Frequency and the Emergence of Linguistic Structure*, Amsterdam: John Benjamins.

CACCIARI, C., & TABOSSI, P. (eds.). (1993), *Idioms: Processing, structure, and interpretation*, Mahwah, NJ: Erlbaum.

CHURCH, K., & HANKS, P. (1980), "Word association norms, mutual information and lexicography", *Computational Linguistics* 16, 22-29.

COLOMBO L. (1993), "The comprehension of ambiguous idioms in context", *in* C. Cacciari & P. Tabossi (eds.), *Idioms: Processing, Structure, and Interpretation*, Mahwah, NJ: Erlbaum, 163-200.

CONKLIN K. & SCHMITT N. (2008), "Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers?", *Applied Linguistics* 29, 72-89.

COULMAS F. (1979), "On the sociolinguistic relevance of routine formulae", *Journal of Pragmatics* 3, 239-266.

CRONK C. B. & SCHWEIGERT W. A. (1992), "The comprehension of idioms: The effects of familiarity, literalness, and usage", *Applied Psycholinguistics* 13, 131-146.

DE COCK S. (1998), "A recurrent word combination approach to the study of formulae in the speech of native and nonnative speakers of English", *International Journal of Corpus Linguistics* 3, 59-80.

ELLIS N. C., SIMPSON-VLACH R. & MAYNARD, C. (2008), "Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL", *TESOL Quarterly* 42, 375-396.

FILLMORE C. (1988), "The mechanisms of construction grammar", *Berkeley Linguistics Society* 14. 35–55.

FONAGY I. (1998), « Figement et changement sémantique », *in* M. Martins-Baltar (éd), *La locution: entre langue et usages*, Paris: ENS Editions Fontenay/Saint-Cloud, 131-164.

GRANGER S. & PAQUOT M. (2008), "Disentangling the phraseological web", *in* S. Granger & F. Meunier (eds), *Phraseology: An Interdisciplinary Perspective*, Amsterdam: Benjamins, 27–50.

GRIES S. Th. (2010), "Useful statistics for corpus linguistics", *in* S. Aquilino & A. Moises (eds.), *A Mosaic of Corpus Linguistics: Selected Approaches,* Frankfurt am Main: Peter Lang, 269–291.

HEINZ M. (1997), « L'À-peu-près dans les locutions et son traitement lexicographique », *in* P. Fiala, P. Lafon & M.-F. Piguet (éds.), *La locution : entre lexique, syntaxe et pragmatique. Identification en corpus, traitement, apprentissage,* Paris : Klincksieck, 213-229.

HEINZ M. (2003), « Les locutions figurées dans le *Nouveau Petit Robert :* évolution de quelques traitements entre 1993 et 2003 », *in* M. C. Cormier, J.-C. Boulanger & A. Francoeur (éds.), *Les dictionnaires Le Robert : genèse et évolution*, Montréal : Les presses universitaires de Montréal, 227-245.

JACKENDOFF R. (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford: Oxford University Press.

JIANG N. & NEKRASOVA T. M. (2007), "The processing of formulaic sequences by second language speakers", *The Modern Language Journal* 91, 433-445.

KUCERA H. & FRANCIS W. N. (1967), *Computational Analysis of Present-Day American English*, Providence, RI: Brown University Press.

LAMIROY B. (2003), « Les notions linguistiques de figement et de contrainte », *Lingvisticae Investigationes* 26, 1-14.

LEGALLOIS D. & GREA P. (2006). « La grammaire de construction », *Cahier du CRISCO* 21, 5-27.

MOREAU M.-L. (1986), « Les séquences préformées: entre les combinaisons libres et les idiomatismes; le cas de la négation avec ou sans ne », *Le français moderne* 54, 137-160.

NENONEN M., NIEMI J. & LAINE M. (2002), "Representation and processing of idioms: Evidence from aphasia", *Journal of Neurolinguistics* 15, 43-58.

PAWLEY A. K. & SYDER F. H. (1983), "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency", *in* J. C. Richards & R. W. Schmidt (eds.), *Language and communication*. London: Longman, 191–225.

PETERSON R. R., BURGESS C., DELL G. S. & EBERHARD K. M. (2001), "Dissociation between syntactic and semantic processing during idiom comprehension", *Journal of Experimental Psychology* 27, 1223-1237.

QUALLS C. D., TREASTER B., BLOOD G. W. & HAMMER C. S. (2003), "Lexicalization of idioms in urban fifth graders: A reaction time study", *Journal of Communication Disorders* 36, 245-261.

SCHMITT N. & UNDERWOOD G. (2004), "Exploring the processing of formulaic sequences through a self-paced reading task", *in* N. Schmitt (ed.), *Formulaic sequences: Acquisition, Processing and Use*. Amsterdam: Benjamins, 173-190.

SINCLAIR J. (1991), *Corpus, Concordance, and Collocation*, Oxford: Oxford University Press.

SIYANOVA A. & SCHMITT N. (2008), "L2 learner production and processing of collocation: A multi-study perspective", *The Canadian Modern Language Review* 64, 429-458.

SIYANOVA-CHANTURIA A., CONKLIN K. & SCHMITT N. (2011), "Adding fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers", *Second Language Research* 27, 251-272.

SIYANOVA-CHANTURIA A., CONKLIN K. & VAN HEUVEN W. J. B. (2011), "Seeing a phrase « time and again » matters: The role of phrasal frequency in the processing of

multiword sequences", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37, 776-784

SKEHAN P. (1998), *A Cognitive Approach to Language Learning*, Oxford: Oxford University Press.

SWINNEY D. & CUTLER A. (1979), "The access and processing of idiomatic expressions", *Journal of Verbal Learning and Verbal Behavior* 18, 523-534.

TOMASELLO M. (2000), "First steps towards a usage-based theory of language acquisition", *Cognitive Linguistics* 11, 61-82.

UNDERWOOD G., SCHMITT N. & GALPIN A. (2004), "The eyes have it: An eye-movement study into the processing of formulaic sequences", *in* N. Schmitt (ed.), *Formulaic sequences: Acquisition, processing and use*, Amsterdam: Benjamins, 153-172.

WOOD D. (2002), "Formulaic language in acquisition and production: Implications for teaching", *TESL Canada Journal* 20, 1-15.

WRAY A. (2002), *Formulaic Language and the Lexicon*, New York: Cambridge University Press.