

## *Pieux mensonge* ou *mensonge blanc*? Les collocations dans les dictionnaires bilingues

### 1. Introduction

« Comment dit-on *hvít lygi* en français ? ». Voici une question posée par un étudiant de français à l'Université d'Islande lors d'un cours consacré à l'expression écrite. Après avoir consulté des dictionnaires existants entre les deux langues, l'islandais et le français, les étudiants ont trouvé, grâce à un dictionnaire anglais-français en ligne, qu'on dit *pieux mensonge* et non *\*mensonge blanc* comme le propose Google Traduction en traduisant chaque composant du syntagme islandais littéralement.

Ce petit exemple illustre bien le fait que les utilisateurs des dictionnaires bilingues n'ont pas seulement besoin d'informations sur les unités lexicales simples mais également sur les exemples de constructions ainsi que sur les unités polylexicales comme les collocations<sup>1</sup> qui constituent une unité significative et qui apparaissent généralement comme une paire au sein d'une phrase. En effet, la langue n'est pas seulement constituée de mots et de leurs combinaisons libres mais également de combinaisons plus ou moins figées. Pour bien maîtriser une langue donnée, il est fondamental de pouvoir employer correctement des syntagmes non libres, ou comme le formule Forkl (2005, 197) : « La vraie maîtrise d'une langue étrangère repose en grande partie sur la possession des milliers – ou même

---

1 En islandais, le terme *ordastæða* est utilisé pour collocation (cf. Jónsson 2005, XX).

millions ! – de ses collocations, qui paraissent si banales au locuteur natif. »

Ainsi, un bon dictionnaire bilingue fournit non seulement des équivalents pour chaque mot-entrée, mais également de nombreux exemples d'utilisation qui illustrent l'emploi du mot-entrée, des collocations (comme *hvít lygi* mentionnée ci-dessus) et des locutions figées au sens figuré.

Dans cet article nous allons nous concentrer sur la collocation en tant qu'entité à prendre en compte dans les dictionnaires bilingues. Plus précisément, nous examinerons le traitement des collocations dans un nouveau dictionnaire islandais-français en ligne, LEXIA, qui est en cours d'élaboration. Avant d'étudier la place des collocations dans la lexicographie bilingue, nous ferons un point sur les deux principales approches en ce qui concerne la définition du concept de collocation.

## II. Les collocations

Le terme *collocation* a été introduit dans les années trente par Firth dans le cadre de l'école contextualiste britannique<sup>2</sup> pour décrire « certains phénomènes linguistiques de cooccurrence qui relèvent essentiellement de la compétence linguistique des locuteurs natifs » (Williams 2003, 33).<sup>3</sup> Le terme lui-même est récent en français, comme le constate Tutin, alors que la notion est plus ancienne car on la trouve déjà chez Bally (1951 [1909]) « sous le terme de série phraséologique » (Tutin 2013, 47–48). On peut également parler d'*affinités syntagmatiques* car il s'agit d'une affinité entre, le plus souvent, deux composants d'un syntagme. Cop (1991) emploie, en effet, l'adjectif anglais *affinitive* dans sa définition de collocation et constate que les collocations se situent dans un continuum entre séquences libres et séquences complètement figées et/ou figurées : « Collocations are affinitive, bipartite lexical combinations which, in terms of the attractive force between their two component parts, can be situated between free combinations and idioms [. . .] ». Hausmann

2 Le contextualisme, développé par Firth et ses successeurs, « prône l'étude de la langue en contexte, et ce pour pouvoir mieux la comprendre, l'enseigner et la traduire » (Williams 2003, 34).

3 Il convient de préciser que la définition du concept de la collocation est plus large selon l'école contextualiste britannique, que d'après l'approche qualitative (cf. II.1) qui définit la collocation comme une « cooccurrence lexicale restreinte » (Hausmann et Blumenthal 2006, 3).

et Blumenthal (2006, 4) constatent également que les collocations comme *un célibataire endurci* et *passer un examen* se situent entre la combinaison libre et la locution idiomatique ou figurée :

Étant une unité polylexicale codée en langue, la collocation se distingue premièrement de la **combinaison libre**, non combinée en langue, qu'elle soit banale (*oublier un examen, un célibataire rancunier*) ou originale (*lait noir, la route se rabougrit, le jour se fissure*). Mais elle se distingue aussi d'autres unités polylexicales codées en langue, notamment de la **locution idiomatique** ou **figurée** (*prendre la mouche, casser les pieds à qqn*) qui est codée en langue mais n'a pas de base (ni, par conséquent, de collocatif), puisqu'elle signifie et est sélectionnée par le locuteur en bloc (le locuteur ne parlant pas de mouche ou de pieds).

C'est donc la cooccurrence restreinte qui distingue les collocations des séquences libres tandis que c'est leur transparence sémantique qui les distingue des locutions figurées comme le constate Laufer (2011, 30) :

Collocations are defined as habitually occurring lexical phrases that are characterized by relative transparency in meaning and form-restricted cooccurrence of elements. Restricted cooccurrence distinguishes collocations from free combinations in which the individual words are easily replaceable following the rules of grammar. Relative semantic transparency of collocations, on the other hand, distinguishes them from idioms whose meaning is opaque since it cannot be understood from the words that compose it.

En effet, Murano (2010, 79) a constaté que « le figement le plus faible se trouve dans les collocations, qui respectent la compositionnalité du sens, mais limitent la liberté de cooccurrence ».<sup>4</sup> Les collocations sont donc des unités syntagmatiques qui, d'une manière générale, peuvent être définies comme étant des locutions figées ne comportant pas (ou très rarement) de sens figuré,<sup>5</sup> c'est-à-dire que les composants d'une

4 La polylexicalité mise à part, le figement et la noncompositionnalité sémantique (ou l'idiomaticité) sont les deux principaux critères employés pour définir les phrasèmes (ou locutions figées).

5 Pour une discussion plus détaillée sur la typologie de différents phrasèmes et sur l'établissement d'une terminologie pour désigner ces différents phrasèmes cf. Burger (2010) et Sverrisdóttir (2009).

collocation donnée ont la même signification que lorsqu'ils sont employés séparément en dehors de cette collocation (Sverrisdóttir 2009, 163–164).

Généralement, on considère qu'il y a deux grandes approches phraséologiques, l'approche dite qualitative et l'approche quantitative. Selon l'approche qualitative, les unités phraséologiques définies par des critères linguistiques se situent sur un continuum, allant des phrasèmes les plus opaques et figés jusqu'aux phrasèmes qui sont plus transparents et variables. Quant à l'approche quantitative ou distributionnelle, elle identifie des cooccurrences lexicales à partir de leur fréquence dans des corpus (Granger et Paquot 2008, 28).

Ces deux approches se manifestent notamment dans deux conceptions de la collocation. Selon l'approche essentiellement qualitative ou sémantique (chez des auteurs comme Mel'čuk 1998, González Rey 2002, Hausmann 2003, Grossmann et Tutin 2003), on suppose qu'il y a une relation hiérarchique entre les constituants d'une collocation qui ont un statut sémantique différent (Siepmann 2005, 410) et le terme *collocation* est défini comme « une cooccurrence lexicale restreinte » (Hausmann et Blumenthal 2006, 3).

Selon l'approche orientée sur la fréquence, par exemple chez Sinclair (1991), ce sont les cooccurrences statistiquement significatives de deux ou plusieurs mots qui sont étudiées (Siepmann 2005, 410). Le terme *collocation* est alors employé pour « l'apparition, soit fréquente, soit statistiquement significative (compte tenu des fréquences pondérées) de deux unités lexicales données dans un contexte plus ou moins étroit » (Hausmann et Blumenthal 2006, 3).

Afin d'avoir un aperçu plus détaillé sur la définition d'une collocation, nous présenterons ces deux conceptions de la collocation. Il convient de préciser, conformément à ce que constate Williams (2003, 44), que les deux approches « représentent en réalité deux façons différentes d'appréhender les mêmes phénomènes de cooccurrences textuelles ».

## II.1 L'approche qualitative ou sémantique

L'acception essentiellement qualitative de la collocation s'inscrit dans la lignée de Bally qui distinguait deux types de « locutions

phraséologiques » : les « unités phraséologiques » et les « séries phraséologiques » (ou « groupements usuels ») (Tutin 2013, 48). Ce dernier groupe correspond à la notion de collocation d'aujourd'hui. Bally (1951 [1909], 70) en donne de nombreux exemples, qui sont régulièrement cités dans la bibliographie sur les collocations, comme *grièvement blessé*, *gravement malade*, *désirer ardemment*, *aimer éperdument*, *intimement lié*, *refuser catégoriquement*, *chaleur tropicale* (cf. également Hausmann et Blumenthal 2006, 8). Selon cette conception étroite, la collocation est une construction binaire constituée d'une « base » superordonnée au niveau cognitif (par exemple *blessé* dans *grièvement blessé*) et d'un « collocatif » qui lui est cognitivement subordonné (*grièvement*).

La collocation est considérée comme une unité binaire selon l'acception qualitative. D'après Siepmann (2005, 415), cela est vrai pour la plupart des collocations. En revanche, il constate qu'on peut argumenter que plusieurs collocations à trois constituants peuvent être réduites à une structure binaire. Ainsi, la collocation *prendre une bouffée d'air* est constituée de (*air* + *bouffée*) et du verbe *prendre*, et *ulcère gastrique bénin* se décompose en (*ulcère* + *gastrique*) et *bénin*. Si l'on accepte l'idée des collocations à trois constituants, on constate aisément, selon Siepmann (2006, 416), que plusieurs combinaisons binaires que l'on considère traditionnellement comme relevant de la combinaison libre (par exemple *accepter des pièces*) sont en réalité emboîtées dans une structure d'une nature collocationnelle plus grande. Il cite l'exemple *le parcomètre n'accepte que des pièces de 20 centimes*, pour illustrer une collocation à trois constituants avec un complément d'objet direct non animé. Siepmann (2005, 416) constate, au sujet des collocations comme *tomber à gros flocons* (sujet : *neige*) et *emporter la conviction* (sujet : *argument*), qu'il est difficile d'identifier une relation systématique entre le verbe et le substantif puisque toute la collocation dépend sémantiquement d'un sujet spécifique (par exemple *neige* dans le cas de *tomber à gros flocons*).

## II.2 L'approche quantitative ou distributionnelle

L'approche orientée sur la fréquence ne se limite pas à la définition d'une collocation comme une unité binaire dotée d'une structure

hiérarchique mais considère l'importance statistique des cooccurrences de deux ou plusieurs mots. John Sinclair (1991, 170), qui est souvent considéré comme étant le précurseur de cette école, définit la collocation de la manière suivante : « Collocation is the occurrence of two or more words within a short space of each other in a text. The usual measure of proximity is a maximum of four words intervening. »

À la différence de ceux qui adhèrent à l'approche traditionnelle de la phraséologie, Sinclair et ses partisans s'occupent moins de la distinction de différentes catégories linguistiques et de sous-catégories des unités polylexicales (angl. *word combinations*) (Granger et Paquot 2008, 29). L'approche distributionnelle ne s'appuie pas sur des critères sémantiques pour identifier des unités polylexicales et ainsi tout type de combinaisons de mots peut faire l'objet d'étude dans le cadre de cette approche de la phraséologie. Toujours d'après les mêmes auteures, la signification n'est pas sans importance, mais on doit porter sur elle un point de vue différent, celui du contextualisme (Granger et Paquot 2008, 31).

Selon Granger et Paquot (2008, 29), l'approche orientée sur la fréquence produit une large gamme de combinaisons de mots qui ne rentrent pas toutes dans des catégories linguistiques prédéfinies. Ainsi, cette approche prend également en compte des séquences nommées en anglais « frames », « collocational frameworks »,<sup>6</sup> et « colligations »<sup>7</sup> (*ibid.*). Ces séquences illustrent, selon Granger et Paquot, le principe d'idiome (angl. *Idiom principle*)<sup>8</sup> développé par Sinclair (1991) qui considère le langage comme étant essentiellement constitué de formules ou chaînes de mots choisis ensemble comme un seul mot (*ibid.*). Williams (2003, 40) explique que, d'après Sinclair, le langage est régi par deux principes, les choix ouverts et le principe

6 Williams (2003, 40) les définit comme des « formules discontinues qui entourent d'autres mots ».

7 Le terme anglais *colligation*, inventé par Firth, désigne un mot et son environnement grammatical (la structure grammaticale qu'un mot préfère ou évite). Hoey (1998) entend par *colligation* : « the grammatical company a word keeps (or avoids keeping) either within its own group or at a higher rank », « the grammatical functions that the word's group prefers (or avoids) » et « the place in a sequence that a word prefers (or avoids) » (Hoey, cité par Siepmann 2005, 439).

8 Sinclair (1991, 173) fait le constat suivant au sujet de ce principe d'idiome : « One of the main principles of the organization of language is that the choice of one word affects the choice of others in its vicinity. Collocation [...] is one of the patterns of mutual choice, and idiom [...] is another. The name given to this principle of organization is the idiom principle. The other main principle of organization which contrasts with the idiom principle is the openchoice principle [...] ».

d'idiome. Ainsi, nous pouvons formuler un énoncé librement « en choisissant des mots et des formes grammaticales, mais, en même temps, nos choix sont limités par des conventions et des phrases préfabriquées qui font partie de la compétence linguistique d'un locuteur natif ». D'après Granger et Paquot (2008, 29), l'approche distributionnelle a montré que plusieurs unités qui n'ont pas été considérées comme appartenant à la phraséologie sont en fait répandues et centrales dans la langue tandis que des séquences très restreintes telles que les locutions idiomatiques et les proverbes sont rares.

L'approche distributionnelle est donc moins étroite que celle dite traditionnelle. En revanche, comme le remarque Siepmann (2008, 186), les résultats obtenus doivent toujours être analysés pour ne pas inclure des cooccurrences entièrement dues au hasard comme « hotel at », « either hotel » ou « nature because ».

Afin d'éviter un flou terminologique, les auteures Granger et Paquot (2008, 42) proposent d'employer le terme *collocation* dans le sens traditionnel pour des préférences dans les relations syntagmatiques entre deux lexèmes (cette préférence étant déterminée par l'usage), et de parler simplement de « cooccurrent » pour faire référence aux cooccurrences statistiques (résultat d'extraction automatique d'un corpus) (Granger et Paquot 2008, 41).

Williams (2003, 33), qui constate que « le concept de collocation est difficile à formaliser et [qu']aucune définition ne fait l'unanimité », précise également que si les grandes maisons d'édition ont adopté une définition contextualiste du sens, ce qui a encouragé l'inclusion des collocations dans les dictionnaires, c'est cette inclusion dans les dictionnaires qui a « obligé les chercheurs à un traitement plus formel ». Il constate également que l'approche traditionnelle « a cependant permis de jeter les bases d'une meilleure prise en compte des collocations dans les dictionnaires » (Williams 2003, 41).

En effet, même si l'approche dite sémantique se concentre sur la hiérarchie sémantique dans la définition de ce qu'est une collocation, la fréquence est quand même un facteur en ce qu'il s'agit d'une cooccurrence habituelle ou usuelle dans la langue.

Conformément à la conception traditionnelle de la collocation, nous considérons qu'il y a une relation hiérarchique entre les constituants qui ont un statut sémantique différent. Ainsi, nous

employons, comme le préconisent Granger et Paquot, le terme *collocation* pour désigner des préférences dans les relations syntagmatiques entre deux lexèmes (cette préférence étant déterminée par l'usage). Nous considérons les collocations comme étant des syntagmes ou phrasèmes sémantiquement compositionnels mais dont la sélection des constituants est restreinte. Il faut néanmoins garder à l'esprit que la collocation peut être constituée de plus que deux constituants, même si souvent ces constituants peuvent être réduits à une structure binaire.

### II.3 L'importance des collocations dans les dictionnaires bilingues

Le fait de pouvoir utiliser correctement des collocations est important dans l'acquisition d'une langue étrangère, d'autant plus qu'elles expriment souvent des activités de la vie quotidienne comme *bursta tennurnar*<sup>9</sup> (« se brosser les dents ») (Sverrisdóttir 2009, 164, Hausmann et Blumenthal 2006, 4–5). Les collocations méritent une attention particulière dans un dictionnaire bilingue parce qu'un choix erroné du collocatif, par exemple *blanc* au lieu de *pieux* pour formuler la collocation *pieux mensonge*, peut donner un effet comique au discours du locuteur étranger. L'emploi correct des collocations, en revanche, donne l'impression que le locuteur étranger parle la langue couramment (Cop 1991, 2775–2776). Nous allons examiner quelques exemples de traitement de collocations et les défis que leur traduction peut poser dans un nouveau dictionnaire islandaisfrançais en ligne, LEXIA, sur lequel nous travaillons.

LEXIA est élaboré à partir d'une nouvelle base de données lexicographique pour la langue islandaise. Cette base de données comporte environ 50 000 entrées et a été créée à l'Institut Árni Magnússon d'études islandaises parallèlement à l'élaboration d'un dictionnaire en ligne multilingue, ISLEX (ouvert au public depuis novembre 2011), avec l'islandais comme langue source et le danois, le suédois, le norvégien (bokmål et nynorsk) et, plus récemment, le finnois et le feringien comme langues cibles. Il s'agit de la première

9 L'équivalent allemand est également une collocation (*sich die Zähne putzen*) (Sverrisdóttir 2009, 164).

base de données en ligne pour la langue islandaise spécialement conçue pour l'élaboration des dictionnaires bilingues.<sup>10</sup> Comme ISLEX, le dictionnaire LEXIA comporte de nombreux exemples (environ 33 000). Ces exemples sont pour la plupart inventés mais inspirés d'exemples réels tirés d'un corpus réuni par l'Institut Árni Magnússon d'études islandaises.

Dans la base de données lexicographique utilisée pour le dictionnaire LEXIA, la conception de ce qu'est une locution figée est assez large : elle est définie comme deux ou plusieurs mots en relation contrainte ou figée. Une distinction est faite entre deux types de locutions : les locutions au sens figuré et les collocations. Les locutions au sens figuré, c'est-à-dire les locutions dont le sens n'est pas compositionnel, sont au nombre de 4 400 tandis que les collocations sont environ 5 500. La plupart (90 %) des collocations et des locutions figées au sens figuré proviennent de la base de données du dictionnaire *Stóra orðabókin um íslenska málnotkun* (Jónsson 2005), le reste venant des collections privées des éditeurs (Úlfarsdóttir 2013, 55, Sigurðardóttir *et al.* 2008, 783).

Le dictionnaire islandais-français LEXIA est destiné à deux groupes d'utilisateurs : d'une part aux locuteurs islandais et d'autre part aux locuteurs francophones. Ainsi LEXIA doit aider les utilisateurs islandais à s'exprimer en français et les utilisateurs francophones à comprendre un texte islandais. De plus, grâce à la fonctionnalité de recherche inversée, l'utilisateur francophone peut taper un mot français dans le moteur de recherche et accéder à tous les articles du dictionnaire où ce mot est proposé comme équivalent ou apparaît dans une traduction d'un exemple d'utilisation. Il convient néanmoins de préciser que le dictionnaire n'est pas bidirectionnel, c'est-à-dire qu'il est uniquement composé de la seule partie islandais-français. Enfin, le dictionnaire sera sans doute utile à tous ceux qui sont amenés à traduire entre les deux langues-objets du dictionnaire.

Comme nous l'avons vu (cf. Davíðsdóttir, 2013), les besoins des utilisateurs d'un dictionnaire bilingue varient selon qu'ils l'utilisent

---

10 LEXIA est accessible depuis un URL différent que celui d'ISLEX, <https://lexia.hi.is/is/>. À l'avenir il est prévu d'ajouter d'autres langues cibles, premièrement l'allemand et par la suite l'anglais et l'espagnol.

pour comprendre un texte dans une langue étrangère ou pour s'exprimer dans une langue étrangère. Quand la langue source est une langue étrangère pour l'utilisateur et la langue cible sa langue maternelle, on parle de dictionnaire de décodage (L2->L1) et sa fonction première est d'aider l'utilisateur à comprendre le sens des mots entrées et de fournir des informations sur l'emploi de ce mot-entrée.

Quand la langue source du dictionnaire est la langue maternelle de l'utilisateur et la langue cible une langue étrangère pour lui, on parle d'encodage (L1->L2). Grâce à un dictionnaire d'encodage, l'utilisateur peut avoir un accès direct aux équivalents qu'il ignore dans la langue étrangère et cela vaut également pour des équivalents de tournures et d'unités lexicales comme les collocations.

Puisque les collocations sont sémantiquement compositionnelles,<sup>11</sup> elles ne posent généralement pas de problèmes de compréhension (de décodage) aux locuteurs étrangers. Selon Farina (2006, 150), le principal problème concernant les collocations (ou « syntagmes de grande fréquence » comme elle les nomme) est de reconnaître les syntagmes comme tels, c'est-à-dire qu'un locuteur étranger risque de traduire les composants d'une collocation par leur équivalent sémantique « comme s'il s'agissait d'une association libre », ce qui pourrait produire une tournure maladroite.

Forkl (2005, 197–198) constate qu'il est important d'inclure des collocations dans un dictionnaire puisque le locuteur étranger qui apprend relativement vite « les équivalents de nombreux mots-bases de sa langue maternelle en fera usage comme points de départ pour développer ses idées et formuler ses phrases dans la langue apprise ». Or, comme le fait remarquer Forkl (2005, 198), c'est la traduction des collocatifs qui est difficile car leur traduction peut uniquement être faite « en fonction de la base » et c'est là où le locuteur étranger a besoin d'aide.

Même les apprenants avancés peuvent ainsi rencontrer des difficultés quand il s'agit de s'exprimer dans une langue étrangère.

11 Mel'čuk (2008, 5) signale néanmoins une complication concernant certaines collocations du type *café noir* « où le collocatif n'a son sens ('sans ajout de produit laitier') qu'au sein de la collocation », mais selon lui cela n'affecte pas la compositionnalité sémantique car « CAFÉ exprime le sens 'café', et NOIR exprime le sens 'sans ajout de produit laitier' » et ainsi « leur combinaison se [fait] par les règles générales ».

D'après Laufer (2011, 29), les apprenants d'une langue étrangère ne perçoivent pas les collocations comme des entités (angl. *chunks*) et les produisent souvent en combinant des mots séparés qui ne vont pas ensemble dans la langue étrangère en question.

D'après Hausmann et Blumenthal (2006, 3), c'est la hiérarchie base-collocatif<sup>12</sup> qui fait qu'en s'exprimant le locuteur progresse en partant de la base pour arriver au collocatif et non inversement. D'après eux une collocation (p. ex. *passer un examen, une bouffée de colère*) est « la combinaison phraséologique codée en langue d'une base (*examen, colère*) et d'un collocatif (*passer, bouffée*) » (*ibid.*). Toujours selon les mêmes auteurs, il s'agit d'une « réalité psychologique observable en discours, notamment de la part d'un locuteur étranger qui perçoit l'apprentissage des bases comme celui d'équations simples (angl. *exam* = fr. *examen*), alors que l'équation du collocatif varie en fonction de la base (pour dire *passer un examen*, l'anglais dit *take* ('prendre') et le collocatif anglais *pass* veut dire en français 'réussir'). » (Hausmann et Blumenthal 2006, 4) et ne convient donc pas pour traduire la collocation française *passer un examen* en anglais. Les mêmes auteurs (*ibid.*) constatent qu'un locuteur étranger, même très avancé, peut éprouver des difficultés à trouver le bon collocatif à partir de la base qu'il connaît dans la langue étrangère.

Un exemple de choix erroné de collocatif que nous avons pu constater chez les apprenants islandais est l'utilisation du verbe *demander* avec le substantif *question* (\**demander une question*), qui est la traduction mot à mot de la collocation *spyrja spurningar* en islandais, au lieu de choisir le collocatif à partir de la base, c'est à dire le verbe *poser* : *poser une question*. Cette erreur dans le choix du collocatif a attiré notre attention sur le fait que cette collocation ne figure dans LEXIA que dans un exemple d'utilisation sous le verbe *spyrja* (« demander ») (cf. image 1).

12 Cop (1991, 2776) a également relevé la hiérarchie basecollocatif : « Collocations consist of a base and a collocator which are not on the same hierarchical level, and so they fulfil differing functions in dictionaries ».

## spyrja so ⓘ

🔊 **framburður**

➔ BEYGING

FALLSTJÓRN: POLFALL (+ EIGNARFALL)

🗉 demander

**hún spurði mig bara einnar spurningar**

🗉 elle m'a uniquement posé une question

**ég spurði hann fréttu af fjölskyldunni**

🗉 je lui ai demandé des nouvelles de la famille

**hann leit á hana en spurði einskis**

🗉 il lui a jeté un regard mais il n'a rien demandé

Image 1.

Dans l'objectif d'enrichir davantage la description lexicographique dans LEXIA, nous suggérons d'ajouter la collocation *spyrja* <einnar> *spurningar* (« poser <une> question ») à l'article *spurning* (« question ») (cf. image 2).

## spurning no kvk

🔊 **framburður**

➔ BEYGING

spur-ning

1

(það sem spurt er um)

🗉 question

**bera fram spurningu**

🗉 poser une question

**leggja fyrir <hana> spurningu**

🗉 &lt;lui&gt; poser une question (poser une question à &lt;quelqu'un&gt;)

**spyrja <einnar> spurningar**

🗉 poser &lt;une&gt; question

2

(vafamál)

🗉 question, doute

**það er engin spurning**

🗉 cela ne fait aucun doute

**það er spurning <hvort þetta sé heppilegt>**

🗉 on peut se demander &lt;si c'est une bonne idée&gt;

Image 2.

Les dictionnaires présentent souvent de nombreuses collocations sans les distinguer en tant que telles, c'est-à-dire qu'elles figurent parmi des exemples de constructions libres (Hausmann et Blumenthal 2006, 7). Même si l'utilisateur n'a pas nécessairement besoin de savoir que *spyrja* <einnar> *spurningar* (« poser <une> question ») est

une collocation, il est important qu'il voie qu'il ne s'agit pas d'un simple exemple relevant d'une combinaison libre. C'est pourquoi les lexicographes devraient donner de l'importance et faciliter l'accès aux collocations, comme le constate Laufer (2011, 45) :

Both lexicographers and teachers should draw learner's attention to the existence of a restricted cooccurrence of lexical items which may be very different from the 'equivalent' combination in their L1. Lexicographers may have to do more than inserting a collocation in an example which illustrates the meaning of the headword. The collocations should be given both prominence and easy access.

Dans LEXIA (comme dans ISLEX), les collocations sont distinguées des exemples, leur forme canonique est donnée et elles apparaissent avec un style typographique différent des autres informations données dans l'entrée du dictionnaire. Ainsi, les collocations sont en caractères gras tandis que les exemples d'utilisation sont en italiques.

Selon Laufer (2011, 45), les collocations peuvent être mises en relief, entre autres, en présentant des exemples de phrases courtes pour chaque collocation dans un encadré clairement mis en valeur.<sup>13</sup> Quand il s'agit de collocations dont la base est un verbe, les collocations sont parfois illustrées dans LEXIA avec un exemple, comme nous montre la capture d'écran ci-dessous (image 3) de la collocation *breiða úr sér* (s.v. *breiða* « étendre »).

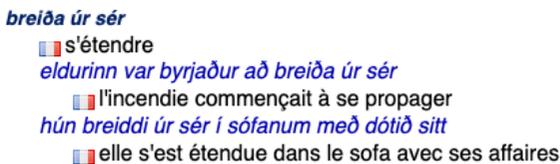
***breiða úr sér***  

 The image shows a dictionary entry for the Icelandic verb 'breiða'. The collocation 'breiða úr sér' is highlighted in bold blue text. Below it, there is an example sentence in italics: 'eldurinn var byrjaður að breiða úr sér'. The English translation of the example is 'l'incendie commençait à se propager'. The full example sentence in Icelandic is 'hún breiddi úr sér í sófanum með dótið sitt', with the English translation 'elle s'est étendue dans le sofa avec ses affaires'.

Image 3.

Il peut parfois être difficile de distinguer une collocation d'une construction libre. Un syntagme comme *bursta temurnar* ou *se brosser les dents* en français peut être considéré comme une collocation ou simplement comme un exemple de l'emploi typique du verbe *bursta*

13 Il faut préciser que Laufer parle des dictionnaires monolingues d'apprentissage (angl. *learners' dictionaries*).

(« se brosser ») avec un des compléments d'objet direct qu'il peut régir (ces compléments étant limités, ainsi en islandais on peut se brosser les dents et les cheveux et on peut également employer ce verbe pour parler du fait de cirer des chaussures ou de brosser des vêtements). Il convient de préciser que la ligne de séparation entre une combinaison libre et une collocation n'est pas toujours nette, les affinités syntagmatiques relevant de différences graduelles et non discrètes. En considérant *bursta tennur(nar)* (« se brosser les dents ») dans le dictionnaire LEXIA on voit que cette affinité syntagmatique est considérée comme une collocation sous l'article *tönn* (« dent ») (cf. image 4) mais qu'elle est seulement insérée dans un exemple qui doit illustrer l'emploi du mot entrée dans l'article *bursta* (« brosser ») (cf. image 5).

## tönn no kvk

 **framburður**

→ BEYGING

1

(bein í munni)

 **dent**

**bursta tennur(nar)**

 se brosser les dents

**falskar tennur**

 dents artificielles, dentier

**gnísta tönnum**

 grincer des dents

**taka tennur**

 faire ses dents, percer ses dents

Image 4.

## bursta so ⓘ

 **framburður**

→ BEYGING

1

FALLSTJÓRN: POLFALL

 **brosser**

*ég ætla að bursta á mér hárið*

 je vais me brosser les cheveux

*hann burstar tennurnar á kvöldin*

 il se brosse les dents le soir

*hún burstaði skóna sína*

 elle a brossé ses chaussures

Image 5.

La plupart des phrasèmes de la base de données LEXIA ont été choisis à partir du dictionnaire monolingue des locutions et collocations *Stóra orðabókin um íslenska málnotkun* (Jónsson 2005). Selon l'auteur de cet ouvrage (Jónsson 2005, XX), il est naturel que les informations sur les collocatifs soient accessibles sous l'entrée de la base puisque la base est plus proche du concept et de la signification de la collocation que le collocatif. C'est ainsi généralement le cas dans LEXIA.<sup>14</sup>

Aujourd'hui, les dictionnaires en ligne offrent la possibilité de chercher par unité polylexicale dans le champ de recherche et ainsi d'accéder directement au traitement d'un syntagme ou d'une locution figée dans le dictionnaire, et l'utilisateur n'a donc plus besoin de savoir sous quelle entrée un syntagme ou une locution est traitée. Il faut préciser que pour obtenir dans LEXIA des résultats en cherchant par une unité polylexicale, l'utilisateur doit d'abord choisir l'option de recherche avancée intitulée « recherche dans le texte intégral ».

Dans LEXIA (comme dans ISLEX), les représentants des composants modifiables ou substituables d'une locution figée ou d'une collocation, par exemple le sujet ou les compléments, sont entourés de chevrons (également nommés crochets obliques) : <...>.<sup>15</sup> Ainsi, la présentation suivante : *spyrja <hana> að nafni* (en français : <lui> demander son nom), indique que le complément d'objet direct *hana* (« lui ») peut être modifié (cf. image 6).

**spyrja <hana> að nafni**  
 <lui> demander son nom (demander <à quelqu'un> son nom)  
**spyrja <hann> spjörunum úr**  
 <le> presser de questions (presser <quelqu'un> de questions)  
 <'>interroger (interroger <quelqu'un>)

Image 6.

Cette présentation permet de rendre compte, au moins partiellement, du degré de figement car l'utilisateur voit les composants qui

14 Parfois les collocations dans LEXIA figurent sous la base ainsi que sous le collocatif comme par exemple *falskar tenur* (« dents artificielles ») que l'on peut trouver sous le substantif *iönn* (« dent ») et également sous l'adjectif *falskur* (« faux »).

15 Cette manière de présenter les locutions figées et les collocations est basée sur le modèle du dictionnaire de Jónsson (2005). Comme le précise Úlfarsdóttir (2013, 55-56), les composants présentés entre crochets ont été simplifiés par rapport au dictionnaire de Jónsson (2005) pour qu'il soit facile de les traduire dans la langue cible.

peuvent être modifiés ou substitués et également quelle est la partie complètement figée d'une locution ou d'une collocation.

En revanche, cette façon de présenter les composants variables d'une collocation est parfois problématique puisque cela peut donner à l'utilisateur une fausse impression par rapport au choix de l'équivalent. Ainsi, la collocation *bíta í* <*brauðið*> qui figure sous l'entrée *bíta* (« mordre ») est traduite par *mordre dans <le pain>* en français, et l'utilisateur islandais pourrait par conséquent en déduire qu'on emploie aussi le verbe *mordre* pour tout autre complément d'objet direct, comme par exemple *bíta í* <*eplið*> (fr. « mordre dans <la pomme> »). Or, en français, il convient d'utiliser le verbe *croquer* avec ce complément d'objet direct : *croquer dans un/une pomme*. Afin de mieux illustrer les traductions possibles en français de la collocation islandaise, il serait envisageable d'ajouter d'autres compléments d'objets indirects typiques régis par la construction verbale islandaise *bíta í*.

Au début du présent article, nous avons cité la question posée par un étudiant de comment dire *hvít lygi* (« pieux mensonge ») en français. Cette question a attiré notre attention sur ce que la collocation islandaise ne se trouvait pas dans LEXIA et nous avons proposé de l'ajouter sous l'article *lygi* (cf. image 7). Cet exemple ainsi que les autres que nous avons mentionnés proviennent de notre expérience de travail lexicographique sur LEXIA et confirment le constat de Williams (2003, 37) que « l'intuition du linguiste a toujours un rôle » dans la lexicographie.

## lygi no kvk

 **framburður**

→ BEYGING

 mensonge

*bessi saga er örugglega lygi*

 cette histoire est sûrement un mensonge

*lygarnar komust upp að lokum*

 finalement, le mensonge a été avéré

**fara með lygar**

 dire des mensonges

**hvít lygi**

 pieux mensonge

---

**þetta er lyginni líkast**

 c'est à ne pas y croire

Image 7.

En revanche, comme le précise Williams (2003, 37), les idées ou « hypothèses doivent désormais être formulées à partir des données réelles ». Ainsi, les deux approches de conception de la collocation se croisent. Dorénavant, nous avons à notre disposition un corpus très large pour la langue islandaise, *Risamálbeild* (Steingrímsson *et al.*, 2018) qui permet d'étudier plus précisément les cooccurrences d'un mot donné ainsi que les collocations. Ce corpus a déjà servi pour trouver des unités lexicales qui, y étant fréquemment attestées, méritent d'être ajoutées à la base de données lexicographique et au dictionnaire ISLEX en tant que motentrées. Il serait intéressant d'explorer ce corpus afin de trouver davantage de collocations ou affinités syntagmatiques à ajouter dans un dictionnaire bilingue dans lequel l'islandais est la langue source. En effet, un des moyens de décourager l'emploi fréquent par les apprenants d'outils médiocres comme Google Traduction est de fournir des dictionnaires en ligne qui proposent une microstructure riche d'exemples avec davantage d'informations syntagmatiques.

## Conclusion

Vu leur importance dans la langue, nous avons voulu porter une attention particulière aux collocations. Nous avons étudié leurs caractéristiques et leur traitement dans les dictionnaires bilingues en nous appuyant sur des ouvrages et des articles dans les domaines de la lexicographie et de la phraséologie. Plusieurs linguistes (cf. Hausmann et Blumenthal 2006, Cop 1991) ont constaté que les collocations se situent sur un continuum entre combinaisons libres et séquences complètement figées. Or, la ligne de séparation entre une combinaison libre, une collocation et une locution figée n'est pas nette et les dictionnaires bilingues doivent prendre toutes sortes de syntagmes en considération, et non pas uniquement les locutions strictement figées et figurées.

Nous avons pu constater que les collocations sont particulièrement importantes dans l'acquisition d'une langue étrangère et, par conséquent, dans un dictionnaire bilingue. En général, l'utilisateur étranger comprend les collocations puisque leur sens est compositionnel. En revanche, leur production dans une langue étrangère est plus

difficile car l'utilisateur ne peut pas savoir si une collocation peut se traduire littéralement dans la langue étrangère. Ainsi, un dictionnaire bilingue devrait non seulement fournir des équivalents pour les entrées de la langue source mais également donner plusieurs exemples d'utilisation ainsi que des collocations dont le sens est compositionnel mais qui sont restreintes dans le choix du collocatif à partir de la base.

Les dictionnaires bilingues ne peuvent pas répondre à tous les besoins des utilisateurs ; un apprenant avancé d'une langue étrangère sera toujours amené à un moment donné à consulter des dictionnaires monolingues pour avoir des informations plus précises sur la langue qu'il étudie. Néanmoins, nous estimons qu'un dictionnaire bilingue peut et devrait être utile pour l'utilisateur avancé d'une langue étrangère, et pas uniquement pour les débutants. Les lexicographes peuvent atteindre cet objectif en développant la partie phraséologique d'un dictionnaire bilingue, intégrant plus d'exemples, ainsi que des collocations présentées séparément de ces exemples.

## ÚTDRÁTTUR

## Orðastæður í tvímála orðabókum

Í þessari grein er sjónum beint að orðastæðum í tvímála orðabókum, einkum í nýrri íslensk-franskri veforðabók, LEXÍU. Orðastæður eru laustengd orðasambönd sem standa sem ein eining innan setningar og mynda merkingarlega heild, t.d. *bursta tennur(nar)* og *hvít lygi*. Merking orðastæðna er yfirleitt ekki yfirfærð, það er að segja að hægt er að ráða merkingu þeirra af einstökum hlutum þeirra. Hins vegar getur reynst málnotendum erfitt að finna jafnheiti orðastæðna í erlendu máli og hætt við að þeir reyni að þýða þær orð fyrir orð. Því er mikilvægt að gera orðastæðum góð skil í tvímála orðabókum. Fjallað er um tvær mismunandi nálganir í skilgreiningum á orðastæðum. Því næst er framsetning orðastæðna í LEXÍU skoðuð og dæmi tekin um viðbætur sem gerðar hafa verið með það að markmiði að bæta orðabókarlýsinguna.

*Lykilorð:* orðastæður, orðasambönd, tvímála orðabækur, íslensk-frönsk veforðabók

## ABSTRACT

**Collocations in bilingual dictionaries**

In this paper the focus is on collocations in bilingual dictionaries, in particular in a new Icelandic-French online dictionary, LEXIA. Collocations are common lexical combinations which can be situated between free combinations and idioms, e.g. *hvít lygi* ('white lie'). The meaning of collocations is usually transparent since it can be deduced from their components. However, it can be a challenge for foreign language users to find equivalents in the foreign language since the co-occurrence of the components of a collocation is restricted. Therefore, collocations are important in bilingual dictionaries, especially for the users whose mother tongue is the source language of the dictionary. It is also important that collocations are easily accessible in online dictionaries. After a discussion of the two main approaches to defining the concept of collocation, the challenges related to the presentation of collocations in LEXIA are examined and illustrated with examples of collocations added to the database.

*Keywords:* collocations, phraseology, bilingual dictionaries, Icelandic-French Online Dictionary.

## BIBLIOGRAPHIE

- Atkins B. T., Sue. 2002 [1996]. « Bilingual Dictionaries. Past, Present and Future ». Dans *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*, édité par M.H. Corréard, 1–29. United Kingdom : EURALEX 2002.
- Bally, Charles. 1951 [1909]. *Traité de stylistique française*. 3<sup>e</sup> édition. Volumes 1–2. Genève : Librairie Georg ; Paris : Librairie C. Klincksieck.
- Burger, Harald. 2010. *Phraseologie. Eine Einführung am Beispiel des Deutschen*. 4<sup>e</sup> édition. Berlin : Erich Schmidt Verlag.
- Cop, Margaret. 1991. « Collocations in the Bilingual Dictionary ». Dans *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, vol. 3, édité par F. J. Hausmann *et al.*, 2775–2778.
- Cowie, Anthony Paul. 1981. « The Treatment of Collocations and Idioms in Learners' Dictionaries », *Applied Linguistics* 2(3), 223–235.
- Cowie, Anthony Paul, éd. 1998. *Phraseology. Theory, Analysis and Applications*. Oxford : Clarendon Press.
- Davíðsdóttir, Rósa Elín. 2013. « Hlutverk tvímála orðabóka. Ólíkar notendaþarfir í íslensk-frönsku ljósi », *Orð og tunga* 15, 73–93.
- Farina, Annick. 2006. « Traduction de syntagmes : une utilisation dynamique des ressources lexicales sur support électronique ». Dans *Lessicografia bilingue e traduzione : metodi, strumenti, approcci attuali*, édité par F. San Vicente, 147–164. Monza : Polimetrica.
- . 2009. « Problèmes de traitement des 'pragmatèmes' dans les dictionnaires bilingues ». Dans *Le dictionnaire maître de langue : lexicographie et didactique. Actes des « Deuxièmes Journées Allemandes des Dictionnaires » à la Mémoire de Josette Rey-Debove*, édité par M. Heinz, 245–263, Berlin : Franck & Timme.
- Forkl, Yves. 2005. « L'accès aux collocations dans le dictionnaire électronique : le cas du *Trésor de la langue française informatisé* (TLFi) ». Dans *L'exemple lexicographique dans les dictionnaires français contemporains. Actes des « Premières Journées allemandes des dictionnaires » (Klingenberg am Main, 25–27 juin 2004)*, édité par M. Heinz, 197–214. Tübingen : Max Niemeyer.
- González Rey, Isabel. 2002. *La Phraséologie du français*. Toulouse : Presses Universitaires du Mirail.
- Granger, Sylviane et MarieAude Lefer. 2012. « Towards more and better phrasal entries in bilingual dictionaries ». Dans *Proceedings of the 15th EURALEX International Congress*, édité par R. Vatvedt Fjeld et J. M. Torjusen, 682–692. Disponible à l'adresse : [http://www.euralex.org/proceedings-toc/euralex\\_2012/](http://www.euralex.org/proceedings-toc/euralex_2012/) [Consulté le 30/04/2021].
- Granger, Sylviane et Magali Paquot. 2008. « Disentangling the phraseological web ». Dans *Phraseology in Foreign Language Learning and Teaching*, édité par S. Granger et F. Meunier, 27–49. Amsterdam, Philadelphia : John Benjamins.
- Grossmann, Francis et Agnès Tutin, édcs. 2003. *Les collocations. Analyse et traitement*. Amsterdam : De Werelt.

- Hannesdóttir, Anna Helga et Jón Hilmar Jónsson. 2001. « Að hafa í sig og á » Isländsk fraseologi í ett isländsktsvenskt perspektiv ». *LexicoNordica* 8, 67–91.
- Hausmann, Franz Josef et Peter Blumenthal, éds. 2006. « Présentation : collocations, corpus, dictionnaires ». Dans *Langue française* 150. *Collocations, corpus, dictionnaires*, édité par F. J. Hausmann et P. Blumenthal, 3–13. Paris : Larousse ; Armand Collin.
- Hoey, Michael (1998). « 'Introducing Applied Linguistics' : 25 years on ». *'Language and Literacies' : Selected papers from the Annual Meeting of the British Association of Applied Linguistics held at the University of Manchester, September 1998*, cité dans Siepmann (2005).
- ISLEX. Þórdís Úlfarsdóttir (rédactrice en chef). Reykjavík : L'Institut Árni Magnússon d'études islandaises. Disponible à l'adresse : [www.islex.is](http://www.islex.is) [Consulté le 30/04/2021].
- Jónsson, Jón Hilmar. 2005. *Stóra orðabókin um íslenska málnotkun*. Reykjavík : JPV.
- Kromann, HansPeder. 1990. « Selection and presentation of translational equivalents in monofunctional and bifunctional dictionaries ». *Cahiers de Lexicologie* 56–57, 17–26.
- Laufer, Batia. 2011. « The contribution of dictionary use to the production and retention of collocations in a second language ». *International Journal of Lexicography* 24(1), 29–49.
- LEXIA. Þórdís Úlfarsdóttir (rédactrice en chef) et Rósa Elín Davíðsdóttir (rédactrice en chef de la partie française). L'Institut Árni Magnússon d'études islandaises. Disponible à l'adresse <https://lexia.hi.is/is/> [Consulté le 30/06/2021].
- Mel'čuk, Igor A. 1998. « Collocations and Lexical Functions ». Dans *Phraseology. Theory, Analysis, and Applications*, édité par A. P. Cowie, 23–53. Oxford : Oxford Clarendon Press.
- . 2004. « Collocations dans le dictionnaire ». Dans *Les écarts culturels dans les dictionnaires bilingues*, dirigé par T. Szende, 19–64. Paris : Honoré Champion.
- . 2008. « Phraséologie dans la langue et dans le dictionnaire ». Dans *Repères & Applications (VI), 2008, XXIV Journées Pédagogiques sur l'Enseignement du Français en Espagne (Barcelone 3–5 septembre 2007)*. Disponible à l'adresse : <http://olst.ling.umontreal.ca/pdf/MelcukPhraseme2008.pdf> [Consulté le 30.04.2021].
- Murano, Michela. 2010. *Le traitement des Séquences Figées dans les dictionnaires bilingues françaisitalien, italienfrançais*. Milano : Polimetrica.
- Siepmann, Dirk. 2005. « Collocation, colligation and encoding dictionaries. Part I Lexicological Aspects ». *International Journal of Lexicography* 18(4), 409–443.
- . 2006 « Collocations et dictionnaires d'apprentissage onomasiologiques bilingues : questions aux théoriciens et pistes pour l'avenir ». Dans *Collocations, corpus, dictionnaires. Langue française* 150, édité par F. J. Hausmann et P. Blumenthal, 99–117. Paris : Larousse.
- . 2008. « Phraseology in learners' dictionaries. What, where and how? ». Dans *Phraseology in Foreign Language Learning and Teaching*, édité par F.

- Meunier et S. Granger, 185–202. Amsterdam : John Benjamins.
- Sigurðardóttir, Aldís *et al.* 2008. « ISLEX - An Icelandic-Scandinavian Multilingual Online Dictionary ». Dans *Proceedings of the XIII Euralex International Congress*, édité par E. Bernal et J. DeCesaris, 779–789. Disponible à l'adresse : <https://euralex.org/category/publications/euralex-2008/> [Consulté le 30/04/2021].
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.
- Steingrímsson, Steinþór *et al.* 2018. Risamálheild: A Very Large Icelandic Text Corpus. *Proceedings of LREC 2018*, 4361–4366. Myazaki, Japan.
- Svensén, Bo. 2009. *A Handbook of Lexicography. The Theory and Practice of Dictionary Making*. Cambridge : Cambridge University Press.
- Sverrisdóttir, Oddný. 2009. « Orð til taks. Af eiginleikum og flokkum fastra orðasambanda. », *Milli mála. Ársrit Stofnunar Vigdísar Finnbogadóttur í erlendum tungumálum*. Reykjavík : Háskólaútgáfan, 149–171.
- Tutin, Agnès. 2013. « Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicatarargument », *Langages* 2013/1 (189), 47–63.
- Úlfarsdóttir, Þórdís. 2013. « ISLEX - norræn margmála orðabók ». *Orð og tunga* 15, 41–71.
- Varantola, Krista. 2002. « Use and Usability of Dictionaries: Common Sense and Context Sensibility? ». Dans *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*, édité par M.H. Corréard, 30–44. United Kingdom : EURALEX 2002.
- Williams, Geoffrey. 2003. « Les collocations et l'école contextualiste britannique ». Dans *Les collocations : analyse et traitement*, édité par F. Grossmann et A. Tutin, 33–44. Amsterdam : De Werelt.