

DICTIONNAIRE DES MOTS COMPOSÉS¹

Daniel Vojtek

Faculté des lettres, Université de Prešov

Abstract: The aim of the article is to present a summary of the activities undertaken within the project APVV-0342-11. The goals of the project and of the individual researchers, as well as the methodology used are exposed. The activities are hierarchically presented and they are stratified according to their relevance.

Keywords: multi-word unit, multilingual dictionary, lexicology, lexicography, notion

Abstrakt: Cieľom článku je podať prehľad činností realizovaných v rámci projektu APVV-0342-11. Článok predstavuje ciele projektu a jednotlivých výskumných pracovníkov, rovnako ako aj použité metódy. Tieto činnosti sú prezentované hierarchicky a sú rozložené podľa ich dôležitosti.

Kľúčové slová: viacсловné pomenovanie, viacjazyčný slovník, lexikológia, lexikografia, pojem

En 2012, une équipe de chercheurs slovaques a été formée sous la direction de M. Ološtiak (maître de conférences à la Faculté de lettres de l'Université de Prešov en Slovaquie) pour entamer un projet qui, à première vue, pourrait paraître titanesque et dont les buts s'avèrent être irréels. L'Agence pour le soutien de la science et de la recherche (APVV) l'a subventionné et les travaux ont pu commencer en juillet 2012.

L'ambition du présent article est simple, à savoir informer sur l'équipe, l'objet de recherche, les objectifs, la méthodologie, l'originalité, les étapes et les résultats visés du projet qui s'intitule *Dictionnaire des mots composés (recherches lexicographiques, lexicologiques et comparatives / Slovník viacсловných pomenovaní (lexikografický, lexikologický a komparatívny výskum)*.

¹ This contribution is the result of the project implementation APVV-0342-11 *Slovník viacсловných pomenovaní (lexikografický, lexikologický a komparatívny výskum)*.

ÉQUIPE

Du fait de la tradition assez importante de la slavistique au sein de la Faculté des lettres à Prešov, cette initiative des chercheurs slavissants n'est pas étonnante. Par conséquent, ces derniers représentent la plupart des membres du projet. Ont participé trois linguistes spécialisés en philologies slovaque et/ou slave (M. Gavurová, M. Ivanová, L. Ološtiaková), trois linguistes lexicographes (M. Ološtiak, M. Kysel'ová, D. Viňarská), un spécialiste en linguistique historique et dialectologie (P. Karpinský), cinq linguistes traductologues (M. Blaho, M. Bilá, M. Kášová, A. Koželová, D. Vojtek) et trois linguistes informaticiens de l'Université technique de Košice (B. Czéreová, J. Genči, J. Staš).

OBJET DE RECHERCHE

L'objet de cette recherche est clair, néanmoins pour la partie qui concerne la traduction, il peut poser problème. L'intérêt se concentre sur les mots composés, y compris ce que l'on appelle les faux composés en linguistique française. En slovaque, les *viacslovné pomenovania*² (mots composés) sont toutes les expressions multiverbales contenant au moins deux composants auto-syntagmatiques dépourvus de toute caractéristique phraséologique ou terminologique. On peut trouver également ainsi dénommés les mots composés qui ont perdu leur valeur de terme et qui sont aujourd'hui usités dans la communication courante.

Le choix de cette partie du lexique peut se justifier par le fait que, de nos jours, ce genre (ou bien ce type) de lexèmes n'a pas subi de traitement lexicographique complexe, ni dans les philologies slaves ni dans les autres philologies.

Dans les premières étapes, la tâche des linguistes traductologues est la recherche et l'analyse de mots composés slovaques et de leurs équivalents anglais, français, allemands, espagnols et russes, ainsi que la contribution à l'élaboration d'un ensemble d'unités lexicales slovaques.

OBJECTIFS

Comme nous l'avons déjà esquissé, il s'agira de profiter du potentiel créatif des départements philologiques et techniques des deux universités concer-

² *Viacslovné pomenovania* veut dire, littéralement, dénominations (noms) contenant plusieurs mots. On trouve également les équivalents français *lexies multiverbales*, *dénominations multiverbales*, *expressions multiverbales*, *expressions nominales*, *expressions composées*.

nées. L'équipe constituée devrait être capable d'élaborer, d'une part, une œuvre de qualité et contribuant au domaine des recherches linguistiques de base, et, d'autre part, une œuvre utile dans la pratique.

Toutes les activités effectuées dans le cadre du projet devraient déboucher sur l'élaboration d'un dictionnaire, qui serait un dictionnaire de langue et en même temps un dictionnaire multilingue de mots composés. La version finale contiendrait 17 000 unités lexicales.

D'autres buts partiels sont visés avant de parvenir au dictionnaire lui-même. Il faudra rassembler le savoir concernant les mots composés avec l'accent mis notamment sur les notions suivantes :

1. Statut et délimitation des mots composés par rapport aux
 - dénominations syntagmatiques libres (*starý otec / veselý otec*)
 - termes d'un domaine de spécialité (recherche de la fréquence d'un terme dans les textes autres que les textes de spécialité, surtout dans la communication des médias)
 - phrasèmes (niveau d'expressivité et de métaphore d'une telle unité lexicale)
2. Typologie des mots composés (termes de spécialité, phrasèmes, noms propres, catégorie de mots invariables, etc.)
3. Caractéristique structurale (types de structure, par exemple les composés substantifs avec un élément adjectival antéposé ou postposé – *hladká múka / kult osobnosti*, etc.)
4. Caractéristique sémantique (sémantique d'un élément et du mot composé dont il fait partie)
5. Rapport entre les mots composés et les unités lexicales univerbales et les processus dynamiques qui y sont liés.

MÉTHODOLOGIE

L'organisation des travaux s'appuiera sur deux principes de base. Le premier est l'acquisition des données des différentes sources et le second est le principe de la paramétrisation et de l'unification.

1. Acquisition des données

Pour cette partie, *Krátky slovník slovenského jazyka* (Petit dictionnaire de langue slovaque) édité en 2003 représente le point de départ pour l'élaboration du dictionnaire, surtout parce qu'il s'agit d'un dictionnaire de langue slovaque contenant le lexique le plus général et le plus fréquemment usité, ce qui répond exactement aux exigences et à l'objectif du traitement des mots composés. La

base de données ainsi acquises sera ensuite complétée par les données provenant d'autres sources comme :

- Dictionnaires de langue et dictionnaires bilingues
- Dictionnaires terminologiques
- Encyclopédies
- Livres et manuels scolaires des écoles primaires et secondaires (surtout parce qu'ils représentent un cadre de savoir essentiel et général)
- Base de données des mots composés motivés temporellement et territorialement, dont notamment les archaïsmes extraits des sept volumes de *Historický slovník slovenského jazyka* (Dictionnaire historique de langue slovaque) et les lexèmes dialectaux extraits des deux volumes de *Slovník slovenských nárečí* (Dictionnaire des dialectes slovaques).

2. Principes de paramétrisation et d'unification

Tous les paramètres des mots composés seront traités dans Microsoft Excel selon les principes de Ďurčo (1991) et Mlacek (1993, Paramétrisation phraséologique) et selon ceux de Furdík (1990, Paramétrisation en dérivation). L'ensemble des données ainsi traitées posera les bases de l'élaboration du dictionnaire et d'une monographie collective (deux publications majeures du projet).

Le point de départ méthodologique et théorique sera la théorie de la motivation lexicale de Juraj Furdík (2008). Parmi les dix-sept types de motivation lexicale imposés par cet auteur (mais également par son successeur M. Ološtiak qui y a consacré son ouvrage *Aspekty teórie lexikálnej motivácie* en 2011) c'est la motivation syntaxique qui se s'avère primordiale par rapport aux objectifs du projet. En effet, toutes les unités lexicales contenant au moins deux composants auto-syntaxmatiques sont syntaxiquement motivées.

ORIGINALITÉ

Étant donné l'absence des travaux lexicographiques consacrés aux mots composés, il s'agira d'un nouveau type de dictionnaire *sui generis* basé sur des principes spécifiques. La structure de chaque entrée (unité fondamentale) du dictionnaire en dépendra directement.

Chaque entrée du dictionnaire comprendra le composant de l'entrée, le mot composé qui contient ce composant, la caractéristique grammaticale, la caractéristique communicative et pragmatique, la sémantisation, l'exemplification et les équivalents en cinq langues. Dans ce sens, le dictionnaire pourrait représenter la continuation de *Slovník české frazeologie a idiomatiky* en quatre volumes (2009) dans lesquels on trouve les équivalents des phrasèmes tchèques

en quatre langues (anglais, allemand, français, russe). Vu la nature du dictionnaire et des recherches qui y sont liées, on suppose pouvoir élargir et enrichir la théorie et la pratique de la lexicographie bilingue.

ÉTAPES

L'hypothèse de départ est celle qu'un mot composé est un type d'unité lexicale spécifique, caractérisé par la configuration des traits suivants :

- composition de deux ou plusieurs éléments
- caractère de notion
- caractère non-terminologique

Cette spécificité exige une approche appropriée du traitement des mots composés dans deux étapes, heuristique et analytique-synthétique.

1. Étape heuristique

La recherche des mots composés dans tout corpus nécessite une méthodologie appropriée, basée sur la combinaison de méthodes computationnelles, statistiques et linguistiques. Une fois les mots composés recherchés, il faut les soumettre à l'analyse linguistique. Dans les cas particuliers ou contradictoires, on peut appliquer de nouveau les informations statistiques, fréquentielles ou autres, selon le cas précis.

2. Étape analytique et synthétique

L'analyse et la synthèse des données acquises seront réalisées en mettant l'accent sur deux publications majeures, le dictionnaire et la monographie collective.

Pour **le dictionnaire**, l'ensemble des paramètres (P) suivants sera analysé:

- P d'orthoépée
- P morphologiques (le dictionnaire comprendra les données morphologiques fondamentales du composant supérieur du mot composé)
- P communicatifs et pragmatiques (l'appartenance de l'unité lexicale dans un groupe thématique, l'expressivité, la langue parlée)
- P sémantiques (explication du sens)
- P d'exemplification et de collocation (l'exemplification est une partie intégrale d'un terme de dictionnaire et sert d'élément support pour la sémantisation)
- P structuraux et P de nomination (les mots composés participent considérablement aux processus dynamiques)
- P traductologiques

Pour **la monographie collective**, l'ensemble des paramètres (P) suivants sera analysé:

- P structuraux (nombre total des composants, nombre de composants auto-sémantiques, structure syntaxique du mot composé, etc.)
- P supra-sémantiques (informations sur la temporalité, la territorialité, l'expressivité, le registre de langue)
- P interlinguistiques (types de symétrie et d'asymétrie entre les mots composés slovaques et leurs équivalents en cinq langues étrangères)

RÉSULTATS

Les résultats que ce projet promet d'obtenir reflètent la totalité des activités prévues. Tout d'abord, le dictionnaire fournira aux usagers du slovaque une revue complexe et originale d'un type d'unités lexicales précis et il illustrera la spécificité des processus de lexicalisation sous forme de glissements sémantiques différents. Il présentera les résultats des processus dynamiques dans la langue. Il servira d'outil à la pratique de la traduction car il offrira les équivalents des mots composés slovaques en cinq langues. Du point de vue lexicographique, il contribuera à la théorie et à la pratique des recherches lexicographiques en introduisant des remarques théoriques, méthodologiques et terminologiques.

La monographie collective présentera de nouveaux points de départ théoriques et méthodologiques en développant la théorie de la motivation lexicale. Elle traitera la typologie stratifiée des mots composés et analysera les moyens de leur naissance (lexicalisation d'une construction syntaxique libre, réduction, etc.) La monographie analysera systématiquement les liens et les limites entre les types de mots composés, et les mots composés et les unités lexicales à un seul mot (univerbales). Elle offrira une analyse comparative des mots composés en slovaque et en cinq langues étrangères et contribuera ainsi aux connaissances sur la typologie des langues.

Bibliographie:

- ČERMÁK, F. et al.: *Slovník české frazeologie a idiomatiky 1. – 4.*, 2. prepracované a doplnené vydanie, Praha : Leda, 2009.
- DESSAINTEs, M.: *Recherche linguistique et enseignement*, Gembloux : Duculot, 1971.
- ĐURČO, P.: « Modelovanie vo frazeológii. Poznámky k moskovskej frazeologickej škole », *Slovo a slovesnosť*, 1991, Vol. 52, n° 2, 138 – 144.

- FURDÍK, J.: *Teória motivácie v lexikálnej zásobe*, M. Ološtiak (ed.). Košice : Vydavateľstvo LG, 2008.
- MLACEK, J.: « K metodologickým zmenám v slovenskej frazeológii », In *Frazeológia vo vzdelávaní, vede a kultúre. Phraseology in Education, Science and Culture*. Nitra : Vysoká škola pedagogická, Fakulta humanitných vied, 1993, s. 282 – 287.
- OLOŠTIAK, M.: *Aspekty teórie lexikálnej motivácie*, Prešov : FF PU v Prešove, 2011.
- POLGUÈRE, A.: *Lexicologie et sémantique lexicale: notions fondamentales*, Montréal : Les presses de l'Université de Montréal, 2003.
- POLGUÈRE, A., CLAS, A.: *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve : Duculot, 1995.

Daniel Vojtek
Institut d'études romanes
Faculté des lettres
Université de Prešov
1, rue 17. Novembre, 080 01 Prešov
Slovaquie
woyteg2@yahoo.ca
daniel.vojtek@unipo.sk