

Agnieszka DRYJAŃSKA

Uniwersytet Warszawski
a.dryjanska@uw.edu.pl
<https://orcid.org/0000-0003-1649-8408>

Vitalija KAZLAUSKIENĖ

Universit t de Vilnius
Vitalija.kazlauskiene@ff.vu.lt
<https://orcid.org/0000-0003-0505-7770>

JOUR, JOURN E AU SENS DE ‘F TE’ ET LEURS  QUIVALENTS COLLOCATIONNELS EN POLONAIS ET LITUANIEN : D FIS POUR LE FLE

1. INTRODUCTION

Cette  tude s’inscrit dans la continuit  d’une recherche collaborative ant rieure (Kazlauskien , Dryjańska 2022) qui visait   analyser le concept de *f te*   travers sa repr sentation linguistique¹. L’une des conclusions formul es indiquait la synonymie entre *f te* et *journ e*, source de difficult s potentielles pour les apprenants, traitant le sujet des festivit s, en tant que contamination des combinaisons de mots, tant au niveau lexical que syntaxique. Cette observation a  t  consid r e comme point de d part de la pr sente  tude, dont l’objectif  tait d’identifier les sp cificit s des RLR des mots *jour* et *journ e* au sens de ‘f te’ en fran ais et de les comparer  

¹ La repr sentation linguistique de la r alit  (RLR) peut  tre d finie, conform ment   la conception des cognitivistes am ricains, comme une « conceptualisation du monde pr sente dans la langue » (Cholewa 2008).

leurs équivalents en polonais et en lituanien. Notre recherche adopte une approche pluridisciplinaire, avec un objectif principalement didactique. Il est pourtant à noter qu'elle ne vise pas à proposer une tâche ou une méthodologie directement applicable en FLE. Nous nous proposons d'identifier des pistes conceptuelles qui pourront ensuite servir de base didactique. Les réflexions et la méthodologie utilisées se situent à l'intersection de la linguistique de corpus et de la linguistique culturelle, en prenant également en compte les apports de la lexicologie et de la sémantique. Afin d'accéder au sens des mots, nous recourons à la sémantique distributionnelle, favorisant une approche quantitative et statistique au sens (Firth et al. 1962 [1957]). Ainsi, tout d'abord, nous abordons les objectifs didactiques liés à l'intégration des corpus de textes dans la didactique des langues. Nous examinons ensuite la définition de la notion clé de l'analyse du corpus – collocation. Par la suite, nous identifions les besoins didactiques propres aux collocations lexicales, en nous appuyant sur les données des corpus nationaux polonais (NKJP²) et lituanien (DLKT³), des corpus français tels que Frantext et la Leipzig Corpora Collection (LCC), ainsi que des données de trois langues fournies par Sketch Engine.

CADRE THÉORIQUE

2. ANALYSE DE CORPUS EN DIDACTIQUE DES LANGUES ÉTRANGÈRES

L'importance des usages stéréotypés de la langue est reconnue par la plupart des études didactiques sur le lexique (Boulton 2010 ; Boulton, Tyne 2014 ; Granger, Paquot 2008; Debono 2018, André 2020 ; Lafontaine 2021 ; Szita 2023 ; Ruggia, Gaillat 2023 et autres). Ce type de relation lexicale constitue un domaine du lexique difficile à maîtriser par les apprenants non natifs. Ainsi se pose la question des aides et des moyens possibles pour atteindre ce but. Les dictionnaires proposant certainement des solutions, ils ont pourtant tendance à s'orienter plutôt vers des mots isolés. « On (...) retrouve la plupart du temps [les cooccurrences lexicales, les collocations] dans la section réservée aux exemples, mais toutes les collocations pos-

² NKJP : Narodowy Korpus Języka Polskiego.

³ DLKT : Corpus du lituanien contemporain.

siblement contrôlées par un mot n'y figurent pas forcément » (Tremblay 2014 : 75). Avec le développement de la linguistique de corpus et d'outils permettant de les explorer sous différentes formes, nous envisageons leur intégration dans la didactique des langues en tant que *data-driven-learning* (DDL) (Johns 1991 ; Boulton 2010) – apprentissage sur corpus (Boulton, Tyne 2014) dans la version *hands-on* qui implique une approche empirico-inductive au sens. Elle permet aux apprenants d'explorer eux-mêmes les données de corpus afin d'en tirer des conclusions sémantico-syntaxiques. Ainsi, le sens devient plus aisément « traitable » et « traçable ». Cette approche est fondée sur le langage authentique avec l'apprenant jouant le rôle de *language detective* de Johns (Chirobocea 2017), qui perfectionne son autonomie (Boulton, Tyne 2014) et ses aptitudes heuristiques.

Nonobstant les acquis de la linguistique de corpus et leur potentiel pour la didactique des langues et cultures (DLC), à notre connaissance, la pratique de classe en Pologne et en Lituanie ne les exploite pas suffisamment, notamment dans le cadre interculturel. Cela explique notre intérêt pour cette problématique.

3. COLLOCATION (COOCCURRENCE, VOISINAGE) EN LINGUISTIQUE DE CORPUS ET EN FLE

L'origine de l'intérêt pour la collocation remonte au moins aux travaux de A. Secheyay et de Ch. Bally, comme le remarquent D. Śliwa (2011) et S. Mejri (2018). A. Secheyay et Ch. Bally, linguistes du début du XX^e siècle, ont ouvert le chemin au développement de la phraséologie, dont le sens est expliqué par ce passage souvent cité : « Notre parole est en bonne partie faite de redites, de membres de phrases, de phrases entières qui se déclenchent mécaniquement dans certaines circonstances et dont nous ne retenons que l'intention générale » (Secheyay 1929 : 98 cité par Mejri 2018).

Il en ressort une conséquence didactique : l'identification et l'intégration des structures fréquemment utilisées par les locuteurs natifs dans des contextes bien précis dans la DLC. Les « fragments redits » de Secheyay (1921 cité par Mejri 2018) ou les « séries » de Bally (1921 cité par Mejri 2018), ne sont « rien d'autre que ce qu'on désigne actuellement par le terme de collocations » (Mejri 2018). L'importance de la combinatoire lexicale est particulièrement accentuée par le terme *la troisième articulation du langage* forgé par Mejri (Ibidem). Dans la lignée de Bally, dans le cadre de l'approche

phraséologique, on trouve, notamment, Hausmann (1989) (cité par Tutin 2005), Meličuk (1998), Tutin, Grossmann (2002), Tutin (2005, 2008). Cette dernière auteure définit les collocations comme « lexicalised phrases where two recurrent lexical elements have a syntactic relationship »⁴ (Tutin 2008). La collocation se compose d'une base et d'un collocatif. Dans la caractéristique de la collocation, on accentue son caractère semi-figé, ce qui a trouvé son expression dans l'appellation de *semi-phasème* proposée par Meličuk. Tutin et Grossmann (2002) sont d'accord sur le fait que « *semi* suggère en effet que le figement n'est pas complet » (Tutin, Grossmann 2002), mais ils optent pour le terme *collocation*.

Le degré de figement des collocations peut être discuté en termes de linguistique, il se traduit pourtant difficilement en langage de formules mathématiques qui est employé dans la linguistique de corpus. C'est elle qui a fait émerger une définition plus large des collocations : ce sont « des mots qui tendent à apparaître ensemble » (Tutin, Grossmann 2002). Cette définition s'applique aux corpus de texte pour lesquels on trouve des outils permettant de mesurer cette « tendance des certains mots à apparaître ensemble ».

Dans Frantexte et le Corpus National de la Langue Polonaise (NKJP) ainsi que dans le Corpus du Lituanien Contemporain (DLKT), on utilise le terme *collocation*. Les corpus dans la LCC sont dotés de fonctionnalités pour identifier des *cooccurrences*, mots qui apparaissent souvent ensemble dans une phrase. Il y a également le terme voisinage (gauche, droit) désignant les mots qui se trouvent directement à gauche ou à droite d'un lexème analysé. Dans la LCC, nous pouvons observer une partie commune entre les cooccurrences significatives et les voisinages (gauche ou droit).

L'extraction automatique des collocations permet de relever tous les mots qui apparaissent fréquemment ensemble, mais elle doit être souvent suivie d'un traitement « manuel », indispensable pour révéler des nuances sémantiques ou pour identifier des classements fautifs. À titre d'exemple, nous pouvons évoquer le classement incorrect du lexème *jour* comme le COD du verbe *pouvoir* dans la phrase suivante : *il puisse un jour tomber malade*. Les outils automatiques ne font pas non plus de distinction, importante pour notre étude, entre les collocations et les locutions verbales telles que *voir le jour*. Une analyse manuelle approfondie est également nécessaire

⁴ 'Les phrases lexicalisées où deux éléments lexicaux récurrents ont une relation syntaxique'.

afin d'identifier des collocations prépositionnelles correctes, à savoir, dans la séquence de mots *les jours contre la maladie*, l'occurrence du mot *jours* est accidentelle, sa signification n'étant pas liée au sens de 'fête', ce qui devient clair quand nous analysons un contexte plus large : *personnes qui luttent tous les jours contre la maladie*. Dans le cadre didactique, ces types d'observations doivent être explicités avant l'intégration des outils d'analyse automatique en classe afin de sensibiliser les apprenants à leurs limites.

Malgré tous les méandres terminologiques dans les approches phraséologique et formelle, il est crucial, à notre sens, de rappeler le lien entre la distribution d'un mot, son sens et sa caractéristique syntaxique, qui se traduit par la constatation de Firth « You shall know a word by the company it keeps » (1962 [1957]). Dans le cadre de la DLC, les apprenants peuvent inférer le sens d'un mot en analysant ses collocations. De plus, afin d'identifier les caractéristiques sémantiques et syntaxiques les plus importantes du point de vue didactique, les apprenants peuvent se focaliser sur la fréquence des mots étant donné que « les lexèmes les plus fréquents désignent les réalités les plus saillantes » (Koselak 2003). Le recours aux corpus de textes constitue donc une voie ouverte pour les apprenants d'une langue étrangère vers les *conceptosphères* de lexèmes.

PARTIE EMPIRIQUE

4. ANALYSE ÉTYMOLOGIQUE ET LEXICOGRAPHIQUE

L'étymologie est une composante énumérée comme première dans le processus de la reconstruction de la RLR (Cholewa 2008). Ainsi, les mots *jour* et *journée* ayant une origine latine (du latin *diurnum* 'jour' (Antidote, TLFi)) signifient un espace de temps. Le mot *journée* désigne un « espace de temps compris entre le lever et le coucher du soleil », tandis que le *jour* se marque par la durée plus longue, il s'agit de « l'intervalle de temps écoulé pendant une rotation complète de la terre servant à situer un événement, une action, un fait » (TLFi).

En ce qui concerne les langues polonaise et lituanienne, l'espace de temps exprimé par deux mots *jour* et *journée* est désigné respectivement par un équivalent – *dzień* et *diena*.

Le mot polonais *dzień* 'jour' vient du proto-slave *dьnъ*, de racine *diw* 'briller', 'luire' (Brückner 1927), *jasna pora doby* 'période de clarté' (Maziarz

2015), *okres od wschodu do zachodu słońca, kiedy jest jasno (w opozycji do nocy)* 'la période du lever au coucher du soleil, lorsqu'il fait clair (par opposition à la nuit)' (WSJP).

Le mot lituanien *dienà* 'jour' provient de *dei-n-ā* iš *dei-* baltes, ce qui signifie 'clarifier, éclairer' (Smoczyński 1982 : 216). La parenté avec *dīnam*, [lie. *dienà*, la. *diena*, sl. *dnov*], a été aussi établie il y a longtemps (Çabej 1961a : 61). La relation avec le sens de 'quotidiennement' (Endzelīns DI IV (2) : 196). Les Prussiens nommaient le matin *deinō*. Dans les dialectes lituaniens, on trouve la signification de *diena* en tant que 'matin', le début de la journée, ex. *prieš pat dienėlę užmigau* 'Je me suis endormi juste avant l'aube' (Mažiulis 1966a : 105).

Ainsi, l'étymologie et la définition de *dzień* et *diena* se ressemblent par leur signification 'briller', 'éclairer', 'luire', cependant la signification de l'intervalle de temps qui sépare le lever du coucher du soleil révèle certaines différences. Ainsi, en lituanien, on trouve encore le sens de matin et de quotidienneté. Il est à noter que le sens de 'fête' des mots *dzień*, *diena*, *jour* et *journée* ne figure pas dans leurs définitions lexicographiques, cependant, il émerge immédiatement d'une analyse de corpus.

5. PRÉSENTATION DE LA RECHERCHE

5.1. La méthodologie et les corpus utilisés dans la recherche

Notre recherche applique la méthodologie *corpus-driven*, selon laquelle le linguiste ne cherche pas à imposer des modèles préexistants ou à faire correspondre ses intuitions aux données extraites lors de l'analyse du corpus. Au contraire, il s'agit de partir de l'observation des données elles-mêmes. Cette démarche adopte une approche inductive et a été documentée par Tognini-Bonelli (2001). Cette méthodologie se traduit en DLC par l'approche *hands-on*. Notre exploration n'a donc pas uniquement pour but de fournir un certain savoir linguistique dont les éléments pourraient être intégrés en FLE, mais également d'identifier des éléments méthodologiques qui se prêteront à l'emploi *hands-on* en classe.

Notre recherche effectuée pour les lexèmes *fête*, *święto* et *šventė* a permis d'extraire les motifs collocationnels N(de)(Dét)N, NA, NV, qui seront examinés cette fois-ci pour *jour*, *journée*, *dzień* et *diena*. Ensuite, l'analyse collocationnelle de ces quatre lexèmes se focalisera sur de nombreux

patrons prépositionnels. ces caractéristiques sont importantes en DLC, car elles permettent de travailler le lexique dans sa dimension sémantique s'associant à la construction syntaxique d'une collocation et à ses variétés.

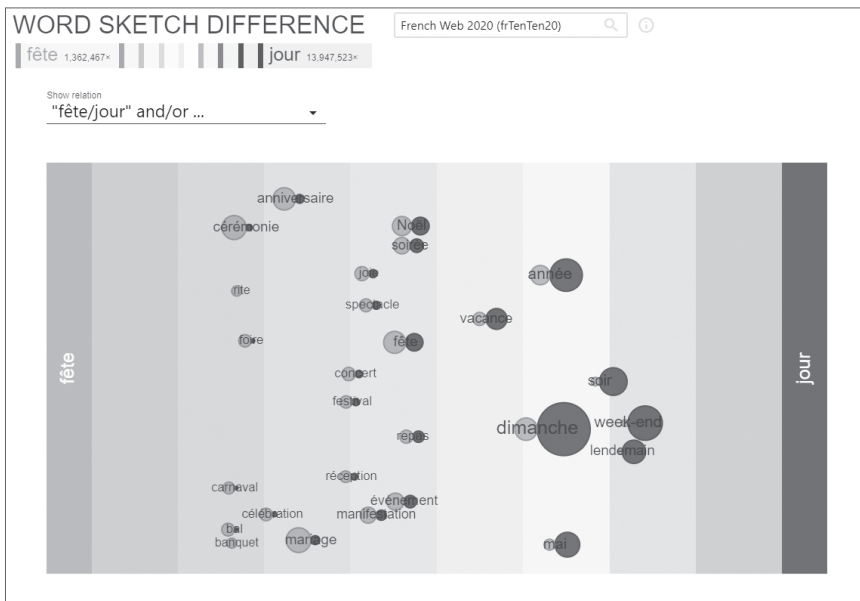
Dans la présente recherche nous avons recouru aux corpus de textes suivants :

1. le Corpus National de la Langue Polonaise (NKJP) équilibré, de 240 192 461 de mots, comprenant des ressources littéraires, journalistiques, spécialisées et d'Internet ;
2. le Corpus du lituanien contemporain (DLKT) équilibré, 140 921 288 de mots, comprenant des ressources littéraires, journalistiques, spécialisées ;
3. Frantext – un corpus français principalement littéraire comprenant 266 millions de mots en décembre 2022 ;
4. la Leipzig Corpora Collection (LCC) – un site contenant de nombreux corpus dans de nombreuses langues basés sur des données d'Internet ;
 - les corpus français utilisés – fra_mixed_2012 et fra_news_2020 comptant respectivement 1 468 766 604 et 149 802 727 mots ;
 - les corpus polonais utilisés – pol_newscrawl_2011 et pol_news_2020 comprenant respectivement 96 476 260 et 28 297 398 mots ;
 - les corpus lituaniens utilisés – lit_newscrawl_2011, lit_news_2020, lit_wikipedia_2021 comptant respectivement 66 635 052 ; 15 943 045 et 7 768 582 mots.
5. Sketch Engine
 - le corpus français utilisé – French Web 2017 comprenant 5 752 261 039 mots ;
 - le corpus polonais utilisé – Polish Web 2019 comprenant 4 253 636 443 mots ;
 - le corpus lituanien utilisé – ELEXIS Lithuanian Web 2021 comprenant 846 563 251 mots.

5.2 Concept de célébration au carrefour des mots *fête*, *jour* et *journée*.

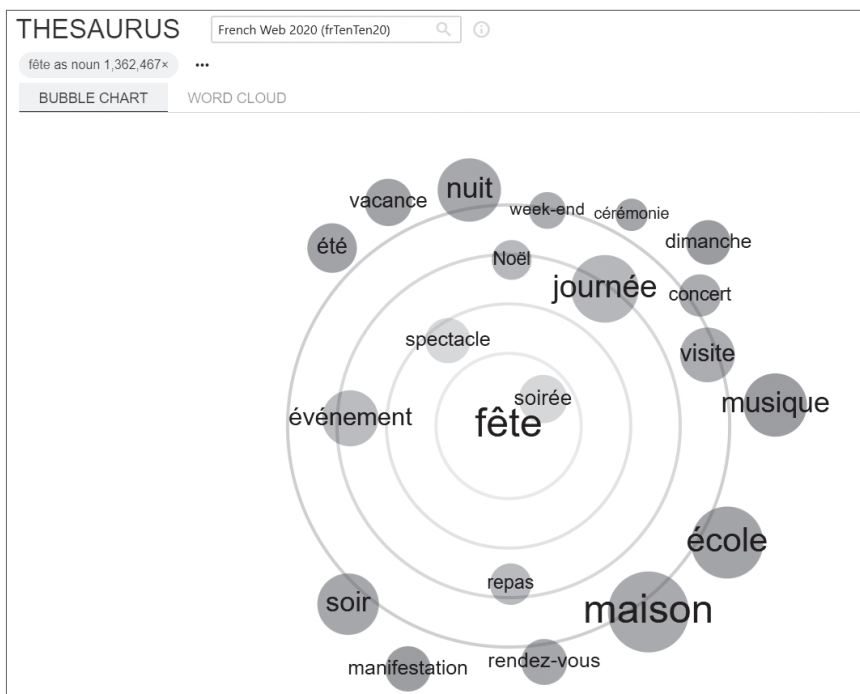
Cette partie de l'étude est focalisée sur les lexèmes *jour* et *journée* employés dans des contextes de festivités de façon synonymique avec le lexème *fête*, ce qui est présenté, à titre d'exemple, par les figures ci-dessous générées par deux fonctionnalités de Sketch Engine – Word Sketch Difference et Thesaurus :

Figure 1. Représentation graphique contrastive de Word Sketch.
Différence des deux lexèmes : *fête* et *jour* (FR)



Source : frTenTen20 Sketch Engine.

Figure 2. Représentation graphique des résultats de la fonctionnalité
Thésaurus dans Sketch Engine du lexème *fête* (FR)



Source : frTenTen20 Sketch Engine.

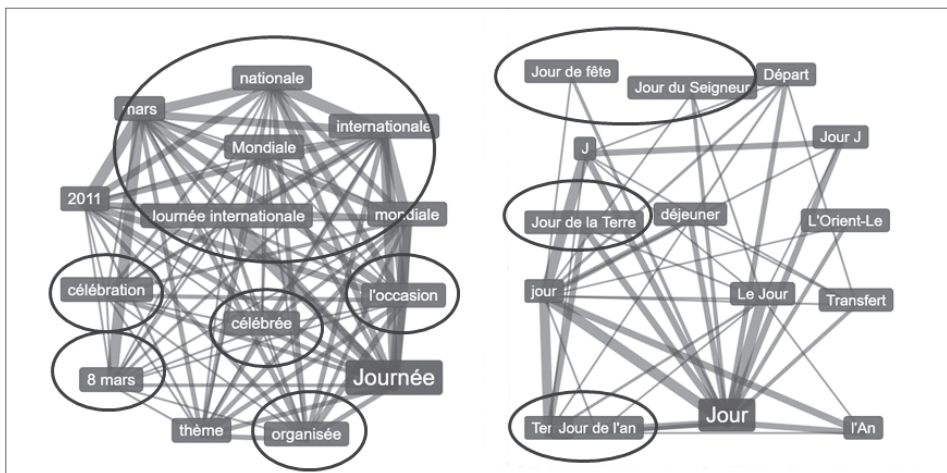
Ces graphiques visualisent l'existence d'une partie commune de l'emploi des lexèmes *fête*, *jour*, *journée*, ce qui confirme la nécessité de l'analyse de ces deux derniers lexèmes et de leurs équivalents polonais et lituanien afin d'appréhender le concept 'fête' en français et dans la perspective contrastive franco-polono-lituanienne.

5.3. Analyse collocationnelle de *jour*, *journée*, *dzień* et *diena* fondée sur la LCC

5.3.1. FR : *jour*, *journée*

L'analyse lexicographique et étymologique a fait émerger une certaine difficulté définitoire concernant les deux lexèmes *jour* et *journée* qui se traduit ensuite par des défis didactiques. L'étape initiale de l'analyse, réalisée dans la LCC FR, montre que les deux lexèmes ont des collocatifs différents : pour *journée*, nous observons notamment des collocatifs adjectivaux *mondial*, *national*, *international*, par contre *jour* est surtout employé dans les collocations à la structure prépositionnelle Nde(Dét)N comme *Jour de l'An*, *le Jour du Seigneur* et *le Jour de la Terre* ce que nous pouvons observer dans la figure 3 :

Figure 3. Représentation graphique de la LCC



Source : French mixed corpus 2012 Leipzig Corpora Collection.

L'analyse contrastive des schémas ci-dessus permet donc assez rapidement de « voir » que la synonymie entre *jour* et *journée* n'est que partielle, ce qui pourrait avoir une utilité didactique. À titre d'exemple, en classe, nous pourrions proposer une activité fondée sur ces deux schémas pour que les apprenants relèvent quelques particularités de chaque lexème qui leur permettront de mieux comprendre la différence sémantico-syntaxique entre *jour* et *journée*. Cette activité peut être suivie p.ex. d'une recherche, dans le même corpus ou sur Internet, d'exemples comprenant des séquences affichées dans les graphes comme *journée internationale, mondiale, jour de l'an, jour de la terre*, etc. afin d'observer des contextes bien concrets de l'emploi des deux lexèmes.

5.3.2. La LCC LT : *diena*

Dans le corpus lituanien de la LCC 2011, le graphique ne se marque pas par le mot *diena* au sens 'fête'. Cependant, dans la liste des cooccurrences significatives, nous les trouvons surtout dans le voisinage gauche.

Exemples : *Valentino diena* 'Saint-Valentin' (2837) ; *Gimimo diena* 'Anniversaire' (1141) ; *Kauno dienos* '**Journées** de Kaunas' (1026) ; *Vilniaus dienos* '**Journées** de Vilnius' (932) ; *Moters diena* '**Journée** des femmes' (749) ; *Motinos diena* '**Fête** des Mères' (709) ; *Tėvo diena* '**Fête** de père' (431) ; *Gynėjų diena* '**Journée** de défenseurs' (379) ; *Muzikos diena* '**Fête** de la musique' (292) ; *Genocido diena* '**Journée** de la mémoire de génocide' (260) ; *Vienybės diena* '**Jour** d'unité' (259) ; *Valstybės diena* '**Journée** d'Etat' (251) ; *Pergalės diena* '**Jour** de la Victoire' (240) etc.

Le corpus permet aussi de relever les collocations prépositionnelles : *Tarptautinė diena be automobilio* 'Journée mondiale sans voiture', *diena be plastiko* 'Journée sans plastique'.

5.3.3. La LCC PL : *dzień*

Les corpus polonais de la LCC *pol_newscrawl_2011* et *pol_news_2020* ont fourni les exemples suivants (les chiffres sont cités d'après la catégorie « Voisinage droit ») :

Dziecka '**Journée** mondiale de l'enfance' (4832) ; *Kobiet* '**Journée** des femmes (l'ONU), **Journée** des droits des femmes' (2255) ; *Matki* 'Fête des Mères' (1367) ; *Edukacji Narodowej* '**Journée** de l'Éducation Nationale' (1493) ; *Pamięci* 'Journée qui commémore les victimes du crime de Katyń, de l'Insurrection de Varsovie' (867) ; *Niepodległości* 'Fête Nationale' (835) ;

Babcy 'Fête des Grands-Mères' (989) ; *Ojca* 'Fête des Pères' (359) ; *Zaduszny* '**Jour des morts**' (793) ; *Światowy Dzień Chorego* '**Journée** mondiale des malades' (510) ; *Dziadka* 'Fête des grands-pères' (366) ; (Zwycięstwa 'Fête) (l'armistice) le 8 mai' (363) ; *Światowy Dzień Pokoju* '**Journée** internationale de la paix' (323) ; *Ziemi* '**Jour** de la Terre' (236) etc.

Nous trouvons aussi les cooccurrences suivantes : *Dzień Wolności podatkowej* '**Jour** de la libération fiscale', *Europejski Dzień bez Samochodu* '**Journée** européenne **sans** voiture', *Dzień Świętości Życia* 'la **Journée** européenne de la vie consacrée', *Dzień Europy* 'la **Journée** de l'Europe'.

Nous observons que les mots *dzień* et *diena* dans les collocations extraites ne se traduisent pas en français uniquement par *journée*, par exemple : *journée mondiale de l'enfance* ou *journée des femmes* et par *jour*, par exemple : *jour de la Victoire*, mais également par *fête*, par exemple *Fête des mères* ou par une autre construction, par exemple : *le 8 mai*.

Cette étape de l'analyse révèle également la structure prépositionnelle (A)NPrépN de la collocation (*europski*) *dzień bez samochodu* 'journée sans voiture' qui sera ensuite analysée en détail pour tous les quatre lexèmes. En français, il s'agira du modèle (A)NPrép(Dét)N.

Notre recherche sera poursuivie afin de mieux appréhender la *collocationality* des lexèmes en question, en suivant les axes révélés par l'étape initiale. Il serait pourtant possible de proposer en classe d'explorer uniquement ce premier corpus, étant donné qu'il fournit déjà des données pertinentes pour une réflexion didactique.

5.4. Analyse collocationnelle de jour et journée basée sur les données de Sketch Engine et Frantext

5.4.1. Collocatifs adjectivaux

Une analyse plus approfondie des collocatifs de *journée* dans Frantext et dans Sketch Engine, a confirmé la présence des adjectifs *mondial*, *international* et *national*. Les collocatifs adjectivaux comme *européen*, *festif* et *commémoratif* ont été extraits de Sketch Engine (sur 42 collocatifs adjectivaux), et *saint*, *glorieux*, *historique*, *solennel* et *familial* de Frantext (sur 1400 collocatifs adjectivaux).

Aucun collocatif adjectival de *jour* au sens de 'fête' n'a été identifié dans Frantext parmi les dix collocatifs adjectivaux les plus fréquents (le premier adjectif trouvé est *heureux* – 188^{ème} parmi 800 collocatifs analysés).

Dans French Web 2017 de Sketch Engine, la situation est pareille excepté la collocation *jour férié*, *férié* étant le premier collocatif adjectival du lexème *jour*, ce qui dévoile son sens de ‘fête’.

Dans Frantext, nous avons extrait les collocatifs adjectivaux de *journée* au sens de ‘fête’ présentés dans le Tableau 1.

Tableau 1. Collocatifs adjectivaux de *journée* dans Frantext

Collocatifs adjectivaux (sur 1400, fréq. > 5)	Fréquence en antéposition (-1)	Fréquence en postposition (+1)
saint	183	-
glorieux	27	-
historique	-	30
national	-	27
international	-	10
solennel	-	6
mondial	-	6
(dimanche) familial	-	1

Source : Frantext.

Dans Sketch Engine, nous avons identifié six collocatifs associés au mot *journée* dans le sens de ‘fête’ parmi les 42 adjectifs examinés : *mondial* (3^{ème} position sur la liste), *international* (5), *national* (7), *festif* (10), *euro péen* (36), *commémoratif* (42). Trois collocatifs – *national*, *international* et *mondial* – apparaissent dans les deux corpus.

Ainsi, force est de constater que *jour* et *journée* au sens de ‘fête’ sont accompagnés de très peu d’adjectifs, ce qui sera aussi valable pour le polonais et le lituanien.

5.4.2. Collocatifs verbaux de *journée* et *jour*

L’analyse des collocatifs verbaux de *jour* et *journée* dans Frantext et dans les corpus français de Sketch Engine a fourni les verbes comme *célébrer*, *fêter* et *commémorer* qui suggèrent leur caractère festif.

Tableau 2. Collocatifs verbaux de *journée* et *jour* dans French Web 2017 (Sketch Engine)

Verbes avec <i>journée</i> comme COD		Verbes avec <i>jour</i> comme COD	
1	passer	1	voir
2	enseigner	2	ouvrir
3	organiser	3	passer
4	souhaiter	4	pouvoir
5	commencer	5	avoir
6	consacrer	6	arriver
7	clôturer	7	durer
8	célébrer	8	travailler

34	fêter	34	célébrer
41	commémorer		

Source : frTenTen17 Sketch Engine.

5.4.3. Patron collocationnel NPrép(Dét)N de *jour* et *journée*

La recherche des collocations des mots *jour* et *journée* de type Nde(Dét)N dans Frantext et des corpus de Sketch Engine a permis d'identifier les collocatifs suivants :

Jour (Frantext) – Nde(Dét)N : *fête* (497), *Noël* (225), *Pâques* (282), *mariage* (202), *saint* (73), *noces* (119), *anniversaire* (91), *congé* (109), *Dieu* (62), *Toussaint* (90), *Pentecôte* (105), *communion* (25), *victoire* (64), *Ascension* (58), *paix* (37), *liberté* (40), *cérémonie* (40), *Résurrection* (39), *libération* (37), *triomphe* (34), *Assomption* (35), *noces* (34), *armistice* (31), *Saint-Louis* (24), *Chandeleur* (25), *Saint-Jean* (24), etc.

Jour (Sketch Engine) sur 42 collocations Nde(Dét)N : *jour de fête* (position 1), *jour de son anniversaire* (pos. 7), *jour de/du mariage* (pos. 9), *jour de Noël* (pos. 11), *jour de l'An* (pos. 15), *jour du Seigneur* (pos. 22), *jour de (sa, des) mort(s)* (pos. 25), *jour de Pâques* (pos. 37), *jour de la Pentecôte* (pos. 39).

Journée (Sketch Engine) sur 42 collocations Nde(Dét)N : *du patrimoine, journée de sensibilisation, d'action, de solidarité, de dimanche, de la femme, d'information*.

Les collocations de *jour* relevées dans la LCC ont été confirmées par Sketch Engine. Ce dernier a fourni encore des combinaisons comme *jour de fête, jour de son anniversaire, du mariage* et des combinaisons avec les noms de fête comme *jour de Noël, jour de Pâques* ou *jour de la Pentecôte*. Ce qui est notable, c'est que nous trouvons dans Frantext de nombreuses collocations du mot *jour* qui sont surtout liées aux fêtes religieuses. On observe, par conséquent, une convergence entre *fête* et *jour*, en termes de motifs syntaxico-lexicaux fournis par Frantext. De l'autre côté, *journée* au sens de 'fête' est un phénomène qui apparaît rarement dans Frantext, en dehors de combinaisons assez particulières comme *journée de Noël* et *journée de Pâques*. Elles ne sont pas confirmées par TLFi, mais relativement fréquentes sur Internet : la séquence *journée de Noël* apparaît 3 640 000 fois sur Google (14.11.2022) vs *jour de Noël* (191 000 000), *journée de Pâques* apparaît 839 000 fois sur Google (14.11.2022) vs *jour de Pâques* (3 230 000). Les combinaisons *journée de Noël* et *journée de Pâques* sont utilisées dans le sens temporel, ce qui est illustré par les exemples suivants :

(1) *J'espère que vous avez passé une belle journée de Noël.* (www 1)

(2) *Durant toute cette journée de Pâques, petits et grands vont se retrouver dans le jardin familial afin de retrouver tous les œufs en chocolat cachés.* (www 2)

ou désignent des célébrations liées à Noël :

(3) *Journée de Noël (marché et animations), Dimanche 11 décembre 2022 – 10h à 18h.* (www 3)

(4) *Le 9 décembre [2021], de 11h à 15, l'UFR EGASS propose à tous les étudiants et personnels de l'université d'Artois une Journée de Noël placée sous le signe de la solidarité et de la convivialité.* (www 4)

Il convient également de mentionner une série de combinaisons de type Nde(Dét)N caractéristiques du lexème *journée* : *journée de sensibilisation, d'action, de solidarité, d'information, d'hommage ou du refus*, qui sont des subséquences des chaînes plus longues comme *Journée de Sensibilisation au Syndrome d'Alcoolisation fœtal*.

L'analyse des lignes de la concordance fournit également des exemples de règles syntaxiques que les apprenants pourraient également analyser eux-mêmes afin de tirer des conclusions sur l'emploi des déterminants. Dans la construction collocationnelle de *jour* avec son collocatif *fête*, il est indispensable que ce dernier soit accompagné d'un complément du nom pour être précédé d'un article défini (ce qui est illustré par la capture d'écran de Sketch Engine).

Figure 4. Capture d'écran de Sketch Engine

Left context	KWIC	Right context
<s> Le 13 juillet, c'est	jour	de fête , mon Android est tombé en pa
) le 8 décembre est le	jour	de la fête de la cathédrale, dédiée à l'lr
s) </s><s> Le 21 juin,	jour	de la fête de la musique, nous sommes
nie ! </s><s> Mais les	jours	de fête , les femmes nouent sur leur têt
e en moi. </s><s> Les	jours	de fêtes étaient gâchés par un évènem
samedi, dimanche et	jour	de fête sur le site de production. </s><
a 67 ans, puisque du	jour	de la fête des SS. </s><s> Apôtres Pie
bles, fêtant ce 8 mars,	jour	de la fête des femmes. </s><s> Ici, on
es, mais ponctués de	jours	de fête , de mariages, des petits bonhe
éfilés durant ces deux	jours	de fête . </s><s> Le Prophète (paix et l
s> 27 décembre 2009	jour	de fête pour la Conrégation entière. P

Source : frTenTen17 Sketch Engine.

Sketch Engine offre une fonctionnalité très efficace permettant de comparer deux lexèmes – Word Sketch Difference, qui visualise les différences les plus significatives concernant deux lexèmes. La figure 5 ci-dessous présente les substantifs dans les patrons collocationnels prépositionnels NcontreDétN.

Figure 5. Substantifs dans les collocations prépositionnelles du type NcontreDétN

"journée/jour" contre			
spécisme	20	0	...
Montpellier	16	0	...
Stade	12	0	...
Nantes	12	0	...
intimidation	11	0	...
harcèlement	42	0	...
FC	15	0	...
misère	17	0	...
PSG	11	0	...
répression	15	0	...
SIDA	11	0	...
précarité	17	0	...
sida	19	0	...
discrimination	11	0	...
racisme	40	11	...
violence	153	49	...
guerre	15	12	...
cancer	16	19	...
équipe	13	24	...
réforme	14	28	...
loi	11	24	...
deux	12	32	...
Israël	0	32	...
armée	0	24	...
cinq	0	13	...

Journée
 Jour

Source : frTenTen17 Sketch Engine.

Pour analyser les patrons prépositionnels comprenant la préposition *contre*, il convient d'analyser plutôt le lexème *journée*, puisque, d'après la figure 5, c'est avec ce lexème que nous observons le plus de cooccurrences comme celles des exemples suivants :

- (5) Une *journée* contre les *discriminations* raciales
- (6) *Journée* contre le *harcèlement* scolaire / à l'école / sexuel (en Egypte)
- (7) à l'occasion de la *journée* contre la *violence* faite aux femmes, éducative.⁵

Parmi toutes les occurrences suivant le patron N(de lutte)contreDétN, qui sont nombreuses (921 occurrences), nous distinguons celles liées :

- à des maladies, par exemple : *journée (de lutte) contre le SIDA* ;
- à des problèmes sociaux : *journée contre la peine de mort, contre les violences sexistes, contre le harcèlement scolaire* ;
- au climat : *journée contre le réchauffement climatique*.

Nous avons également relevé deux types de cooccurrences très nombreuses : NpourDétN (13316) comme *journée pour la santé mentale, pour la biodiversité* et NsansN (5050) : *journée sans tabac, sans voiture, sans achat, etc.* Notons l'emploi des articles définis après les prépositions *contre* et *pour*, dans les exemples analysés, ce qui n'est pas le cas de la préposition *sans*. Les patrons collocationnels contenant les prépositions *contre*, *sans* et *pour* sont très productifs puisqu'ils s'appliquent à la construction de termes désignant des événements dont le rôle est de sensibiliser la société à des problèmes sociaux urgents.

L'extraction des patrons prépositionnels grâce à des outils de Sketch Engine et leur analyse sémantique et syntaxique, notamment l'observation de l'emploi des articles définis, constituent une activité applicable en classe.

5.5 Analyse collocationnelle de *dzień* dans NKJP et Sketch Engine

5.5.1. Les collocatifs nominaux de *dzień* dans NKJP

L'analyse des données de NKJP a permis de relever 70 collocatifs nominaux de *dzień*, soit 10% de tous les collocatifs nominaux. Cette analyse

⁵ Selon les données de ce corpus, nous constatons la nette dominance de « femmes » en tant que victime de violence.

a confirmé que NN était le patron dominant dans NKJP, mais également dans la LCC PL, où nous observons aussi la récurrence de certains collocatifs :

święto (115), *dziecko* (72), *kobieta* (71), *kultura* (64), *niepodległość* (34), *urodziny* (29), *wiosna* (28), *walka* (26), *ziemia* (17), *wyzwolenie* (16), *inauguracja* (13), *strażak* (13), *Wielkanoc* (12), *szkoła* (12), *podpisać* (11), *patron* (11), *zmartwychwstać* (11), *młodzież* (10), *muzyka* (10), etc.

Parmi les 10 collocatifs les plus fréquents du mot *dzień* présentés dans le tableau 3, les plus nombreux sont accompagnés en français par le mot *journée*.

Tableau 3. Collocatifs nominaux les plus fréquents en polonais

Collocatif	Collocatif en fr.	Fréquence
<i>święto</i>	fête	115
<i>dziecko</i>	(journée) enfant	72
<i>kobieta</i>	(journée) femme	71
<i>kultura</i>	(journée) culture	64
<i>niepodległość</i>	indépendance (fête nationale, Armistice)	34
<i>urodziny</i>	anniversaire	29
<i>(pierwszy dzień) wiosna</i>	(le premier jour du) printemps	28
<i>walka</i>	(journée) lutte	26
<i>ziemia</i>	(jour) terre	17
<i>Wyzwolenia (Warszawy,...)</i>	libération (par ex. de Varsovie,...)	16

Source : Narodowy Korpus Języka Polskiego (NKJP).

Le patron NA, qui a été ensuite analysé, est peu présent dans NKJP. Parmi ces collocatifs, on note *zaduszny* 'jour des morts' – 15^e position sur 294 collocatifs adjectivaux, 29^e position sur tous les collocatifs, fréq. 20, $\chi^2 = 3646$ et *święteczny* 'festif' – 43^e position sur 294 collocatifs adjectivaux, 86^e sur tous les collocatifs, fréq. 64, $\chi^2 = 863$.

Le lexème *dzień* est souvent précédé des adjectifs : *światowy* 'mondial' (25 946), *międzynarodowy* 'international' (8308), *europejski* 'européen' (4614). En français, ces adjectifs apparaissent comme des voisins droits.

L'analyse de *dzień* dans NKJP et Sketch Engine a confirmé que le patron VN est également relativement peu varié mais fréquent (100). Les collocatifs verbaux suivants ont été relevés : *obchodzić* 'célébrer', *świętować* 'fêter', *uczcić* 'commémorer', *zorganizować* 'organiser', *ustanowić* 'établir'⁶.

Certains noms de fêtes sont particulièrement intéressants à cause des nuances terminologiques et traductologiques suivantes :

1. *Dzień Niepodległości* – le nom officiel de cette fête est la Fête de l'Indépendance (de la Pologne) ou la Fête Nationale. En France, il s'agit de l'Armistice. Dans NKJP, à part ce sens, nous trouvons quelques occurrences de *Dzień Niepodległości* au sens de la fête américaine *Independence Day*.
2. *Dzień Wyzwolenia* – le Jour de Libération est lié à la libération de différents lieux pendant la Seconde Guerre mondiale, comme le Jour de Libération de Varsovie.
3. *Dzień Zwycięstwa*, combinaison relevée uniquement par le concordancier de NKJP (94 occurrences dans le corpus équilibré). La fête est fixée pour le 8 mai en Pologne, en France et en Lituanie. Ce n'est plus un jour férié ni en Pologne ni en Lituanie, il l'était pourtant à l'époque communiste, et dans ce cas-là, il était célébré le 9 mai.
4. En polonais, la fête appelée *Dzień Zaduszny* ou *Zaduszki* 'le jour des morts' est célébrée le 2 novembre pour commémorer tous les défunts.⁷ En français et en polonais aussi bien qu'en lituanien l'appellation de la fête *le jour des morts* est souvent confondue avec la Toussaint (pl. *Wszystkich Świętych*), qui est pourtant célébrée le 1^{er} novembre pour rendre hommage à tous les saints.

5.5.2. *Dzień* – collocations prépositionnelles (NKJP, Sketch Engine)

Plusieurs motifs prépositionnels ont été distingués dans NKJP et Sketch Engine. Premièrement, notons le motif *dzień bez* 'journée sans' (83) : *dzień bez*

⁶ On note la présence d'un verbe *święcić* au sens de 'świętować' qui n'est utilisé que dans le troisième Commandement – *dzień święty święcić* 'sanctifier les fêtes'.

⁷ C'est une fête religieuse succédant à une tradition populaire slave préchrétienne appelée *Dziady* 'Aïeux' célébrée dans la nuit du 31 octobre au 1^{er} novembre sur le territoire polonais, lituanien et biélorusse. Le nom *Dziady* est très connu en Pologne puisque c'est le titre d'une œuvre de l'un des plus grands poètes d'expression polonaise Adam Mickiewicz, né sur le territoire lituanien et pour cette raison aussi reconnu en Lituanie, sous le nom d'Adomas Mickevičius.

papierosa ‘cigarette’, *futra* ‘fourrure’, *samochodu* ‘voiture’, *przemocy* ‘violence’, *kupowania/zakupów* ‘achats’, *prasy* ‘presse’, *alkoholu* ‘alcool’.

Ensuite, il y a le motif *dzień walki z/przeciw* ‘journée de lutte contre’ (177), qui, comme en français et en lituanien, est souvent précédé des adjectifs suivants : *światowy* ‘mondial’, *międzynarodowy* ‘international’, *europski* ‘européen’. Parmi ces collocations, nous avons observé deux catégories semblables à celles en français et lituanien :

- celles relevant des maladies : *dzień walki z cukrzycą* ‘diabète’, *z depresją* ‘dépression’, *z AIDS* ‘SIDA’, *z rakiem* ‘cancer’, *z białaczką* ‘leucémie’ ;
- celles relevant des problèmes sociaux : *dzień walki z rasizmem* ‘racisme’, *z piractwem* ‘piratage’, *z bezrobociem* ‘chômage’, *z ubóstwem* ‘pauvreté’, *przeciw dyskryminacji osób niepełnosprawnych* ‘contre la discrimination des personnes handicapées’ ;

et une catégorie différente liée à la dépendance : *dzień walki z nikotynizmem* ‘journée de lutte contre nicotisme’, *z narkotykami* ‘drogue’. En polonais, le corpus fournit également un motif collocationnel *dzień walki o* ‘journée de lutte pour’, par ex. *dzień walki o równe prawa* ‘droits égaux’.

D’après le corpus Sketch Engine, le patron *dzień bez* ‘journée sans’, également très productif, fournit des collocations telles que⁸ : *bez samochodu* ‘sans voiture’ (921), *bez papierosa* ‘sans cigarette’ (192), *bez telefonu* ‘sans portable’ (114), *bez przemocy* ‘sans violence’ (72), *bez ofiar śmiertelnych na drogach* (35) ‘sans victimes de la circulation routière’.

5.6 Analyse collocationnelle de *diena* fondée sur DLKT et Sketch Engine

Les données du Corpus du lituanien contemporain ‘*Dabartinės lietuvių kalbos tekstynas*’, marqué plus loin dans le texte par le sigle (DLKT), illustrent dans sa majorité les collocations de *diena* au sens de *fête* sous forme nominale du type NN (*Ndiena*). De toutes les 3049 cooccurrences de *diena*, on a relevé 1783 collocations au sens de *festivité*, dont la structure est du type *Ndiena*. 45 collocations du contenu différent de ce type ont été trouvées.

⁸ Les chiffres sont présentés à titre indicatif, car ils représentent aussi des collocations non pertinentes : *Z kolei co szósta zapytana osoba [...] twierdziła, że dzień bez internetu to dzień stracony* (bizneston.pl) ‘En revanche, une personne interrogée sur six a déclaré qu’une journée sans Internet était une journée perdue.’

Tableau 4. Collocatifs les plus fréquents en lituanien

Collocation	En français	Fréquence
<i>Kauno</i>	Fête de Kaunas (fête de la ville)	999
<i>Gimimo</i>	L'anniversaire	190
<i>Valentino</i>	(Le jour de) la Saint-Valentin	157
<i>Nedarbo</i>	Jour férié	98
<i>Motinos</i>	Fête des Mères	70
<i>Viešpaties</i>	Fête Dieu	37
<i>Valstybės</i>	Fête nationale	28
<i>Tėvo</i>	Fête des pères	22
<i>Visų Šventųjų</i>	La Toussaint	15
<i>Laisvės gynėjų</i>	Journée des défenseurs de la liberté	15
<i>Laisvės</i>	Journée de la liberté	13
<i>Tarptautinė darbo</i>	Fête du Travail	13
<i>Švento Lozorjaus</i>	Saint Lazare	10
<i>Lietuvos globėjo šv. Kazimiero</i>	Saint Casimir	10

Source : DLKT et ItTenTen14 Sketch Engine.

Dans le tableau 4, la liste des collocations les plus fréquentes de *diena* représente quatre types d'équivalents en français :

- équivalent identique, contenant le mot *jour* : *Valentino diena* '(Le jour de) la Saint Valentin' ; *nedarbo diena* 'Jour férié' ;
- équivalent contenant le mot *fête* : *Kauno diena* 'Fête de Kaunas (fête de la ville)' ; *Motinos diena* 'Fête des Mères' ; *Viešpaties diena* 'Fête Dieu' ; *Valstybės diena* 'Fête nationale' ; *Tėvo diena* 'Fête des pères' ;
- équivalent contenant le mot *journée* : *Laisvės gynėjų diena* 'Journée des défenseurs de la liberté' ; *Laisvės diena* 'Journée de la liberté' ;
- équivalent différent : *Gimimo diena* 'L'anniversaire' ; *Visų Šventųjų diena* 'La Toussaint' ; *Švento Lozorjaus diena* 'Saint Lazare' ; *Lietuvos globėjo šv. Kazimiero diena* 'Saint Casimir'.

Les équivalents contenant le mot *fête* prévalent, les collocations au sens de *jour* sont inférieures en nombre. Elles sont présentes dans les trois pays, il s'agit plutôt des événements internationaux. Les fêtes à l'honneur d'un saint portent seulement son nom.

5.6.1. *Diena* – Sketch Engine

La caractéristique collocationnelle du mot *diena*, telle qu'elle est représentée dans le corpus de Sketch Engine est marquée par d'autres formes syntaxiques, nous notons des constructions adjectivales aussi bien que verbales. Les adjectifs se posent des deux côtés du nom d'après leur sens.

(A)N(A)

- (Aright) : solennel *iškilmingas*, international *tarptautinis*, officiel *oficialus*, d'Etat *valstybinis*, extraordinaire *neeilinis*.
- (Aleft) : festif *šventinis*, international *tarptautinis*, mondial *pasaulinis*, sacré *šventas*, bonjour *labas*, officiel *oficialus*.

En général, les adjectifs les plus fréquents ne se distinguent pas par le sens de *fête*, il s'agit plutôt de la nature, du quotidien ('ensoleillé' *sauleta*, 'fatidique' *lemtinga*, 'courte' *trumpa*, etc.).

Les collocations verbales sont relativement peu variées, mais fréquentes. Il s'agit de verbes qui indiquent plutôt l'organisation ou une certaine attente d'une fête. Exemples :

(Vright) : passer *praleisti*, célébrer *švęsti*, proclamer *paskelbti*, commémorer *(pa)minęti*, organiser *surengti*, venir *ateiti*, *sukakti* ;

(Vleft) : célébrer *švęsti*, durer *trukti*, passer *praleisti*, arriver *išaušti*, *ateiti*, souvenir *atminti*, commémorer *minęti*, *vykti*, attendre *laukti*, annoncer *paskelbti*, passer *įvykti*, tenir *surengti*, *paminęti*.

6. CONCLUSION

Grâce aux corpus à notre disposition, nous avons extrait certaines caractéristiques combinatoires propres aux lexèmes analysés en trois langues. Il est à noter que, même si la définition lexicographique ne contient pas de signification de festivité pour les mots *journée* et *jour*, elle s'avère vraiment

pertinente pour ces mots à la lumière des résultats de l'analyse de corpus. Les données extraites des corpus nous ont permis d'identifier les particularités collocationnelles et sémantiques des collocations associées aux mots *jour* et *journée* et de les comparer à leurs équivalents en polonais et en lituanien, ce que nous pouvons résumer ainsi :

- les collocations aux constructions adjectivale et verbale, peu variées, sont semblables pour les trois langues, sauf le mot *jour* qui accepte moins d'adjectifs que le mot *journée*. La situation est inverse avec les collocations de type Nde(Dét)N en français, équivalentes à des constructions génitives NN en polonais et en lituanien : *journée* est plus fréquent dans ce cas que *jour* ;
- parmi les collocations de *journée*, la dominance du caractère *civil* des festivités est très nette, alors que *fête* s'associe plus souvent à des célébrations religieuses. Ce caractère civil très particulier s'exprime surtout dans les combinaisons prépositionnelles caractéristiques, introduisant une nouvelle dimension sémantique des lexèmes *journée*, *dzień* et *diena*, inconnue du lexème *fête* ;
- les résultats ont révélé une plus grande ressemblance entre les lexèmes *fête* et *journée* ;
- une corrélation plus visible entre le français et le polonais révélée pour *fête* ne se manifeste pas pour *dzień* et *diena* ; nous observons plus de ressemblances entre le polonais et le lituanien.

En ce qui concerne les besoins didactiques, les corpus de textes analysés présentent un potentiel didactique vu certaines fonctionnalités comme Word Sketch et des représentations schématiques de relations syntaxiques et sémantiques dans la LCC (Figure 3) et Sketch Engine (Figure 2). Ce dernier semble particulièrement efficace pour l'apprentissage des collocations de type Nde(Dét)N. Sketch Engine offre aussi la fonctionnalité contrastive Word Sketch Difference (Figures 1, 5) susceptible d'être intégrées en classe en vue de passer au crible des lexèmes considérés comme synonymiques, comme *jour* et *journée*, souvent problématiques pour les apprenants, d'après notre expérience.

Il s'agirait de les aider à retenir simultanément le sens et la combinaison lexicale en FLE, ainsi que de fixer les structures qu'ils connaissent déjà partiellement. Les démarches didactiques fondées sur ces fonctionnalités s'inscriraient dans la DDL qui permet à l'apprenant d'agir de manière inductive.

BIBLIOGRAPHIE

- André V., 2020, *Faire de la linguistique de corpus avec des apprenants de français langue étrangère*, « La didactisation du français vernaculaire », pp. 37–66, <https://hal.science/hal-02993341/document> (consulté le 2 juin 2023).
- Boulton A., 2010, *Data-driven learning : the perpetual enigma*, dans : *Explorations across Languages and Corpora*, réd. S. Goźdź-Roszkowski, B. Lewandowska-Tomaszczyk, Frankfurt, pp. 63–580, <https://hal.archives-ouvertes.fr/hal-00528258v1/document> (consulté le 10 septembre 2022).
- Boulton A., Tyne H., 2014, *Des documents authentiques aux corpus : démarches pour l'apprentissage des langues*, Paris.
- Chirobocea O., 2017, *The good and the bad of the corpus-based approach (or data-driven learning) to ESP teaching*, “Naval Academy Scientific Bulletin” XX, 1, pp. 364–371, https://www.anmb.ro/buletinstiintific/buletine/2017_Issue1/FCS/364-371.pdf (consulté le 10 septembre 2022).
- Cholewa J., 2008, *Image encyclopédique et linguistique du chat et du chien en français et en polonais contemporains*, Białystok.
- Debono M. (2018), « *Corpus didactiques : enjeux du traitement des ‘traces’ numériques dans l’enseignement distanciel (des langues)* » dans : *Penser les diversités linguistiques et culturelles. Francophonies, formations à distance, migrations*, réd. I. Pierozak, M. Debono, V. Feussi, V. Huver, Limoges, 432, pp. 121–134, <https://dumas.ccsd.cnrs.fr/AO-LINGUISTIQUE/hal-02149723> (consulté le 2 juin 2023).
- Firth J.R. et al., 1962 (1957), *Studies of linguistic analysis*, Oxford.
- Fabre C., 2015, *Sémantique distributionnelle automatique : la proximité distributionnelle comme mode d’accès au sens*, « Éla. Études de linguistique appliquée » 4(180), pp. 395–405.
- Granger S., Paquot M., 2008, *Disentangling the phraseological web*, dans : *Phraseology : An interdisciplinary Perspective*, réd. S. Granger, F. Meunier, Amsterdam-Philadelphia, pp. 27–49.
- Johns T., 1991, *Should you be persuaded: Two samples of data-driven learning materials*, dans : *ELR Journal*, réd. T. Johns, P. King, 4, pp. 1–16.
- Kazlauskienė V., Dryjańska A., 2022, *Le sens de fête en polonais, en lituanien, en français et sa (non)coïncidence collocationnelle*, „Annales Universitatis Paedagogicae Cracoviensis. Studia linguistica” 17, pp. 20–42.
- Koselak A., 2003, *La sémantique naturelle d’Anna Wierzbicka et les enjeux interculturels*, « Questions de communication » 4, pp. 83–95.
- Lafontaine F., 2021, *Proposal to use the study corpus for contemporary French in Didactics of French as a Foreign Language*, „Jazykovedný časopis” (« Journal of Linguistics ») 72(3), pp. 951–966.

- Maziarz M., 2015, *O pochodzeniu słowiańskich i germańskich nazw pór dnia*, „Acta Universitatis Vratislaviensis”, pp. 35-57.
- Mejri S., 2018, *Les pragmatèmes et la troisième articulation du langage*, « Verbum » XL, 1, pp. 7–19.
- Meličuk I., 1998, *Collocations and Lexical Functions*, in: *Phraseology. Theory, Analysis and Applications*, éd. A. P. Cowie, Oxford, pp. 23–53, <http://olst.ling.umontreal.ca/pdf/CowieLFs.pdf> (consulté le 10 septembre 2022).
- Ruggia S., Gaillat T., 2023, *Les corpus numériques pour la didactique des langues : de la formation des enseignants à l'élaboration de dispositifs d'apprentissage*, « Corpus » 24.
- Szita S., 2023, *Nous reviendrons avec plaisir. Construction de mini-corpus pédagogiques thématiques et identification du vocabulaire-clé*, « Corpus », 24, <http://journals.openedition.org/corpus/7536> (consulté le 30 mai 2023).
- Śliwa D., 2011, *Cœur et sentiments : une étude cognitive des collocations françaises et polonaises*, dans : *Parler des émotions : entre langue et littérature*, éd. A. Krzyżanowska, R. Jakubczuk, Lublin, pp. 191–199.
- Tognini-Bonelli E., 2001, *Corpus Linguistics at Work*, Amsterdam.
- Tremblay O., 2014, *Les collocations : des mots qui font la paire*, « Québec Linguistique Appliquée » VII, 1, pp. 7–26.
- Tutin A., 2005, *Le dictionnaire de collocations est-il indispensable?*, « Revue Française de Linguistique Appliquée », X, 2, https://www.researchgate.net/publication/32230388_Le_dictionnaire_de_collocations_est-il_indispensable (consulté le 10 septembre 2022).
- Tutin A., 2008, *For an Extended Definition of Lexical Collocations*, « Euralex », pp. 1453–1460, https://www.euralex.org/elx_proceedings/Euralex2008/067_Euralex_2008_Agnes%20Tutin_For%20an%20Extended%20Definition%20of%20Lexical%20Collocations.pdf (consulté le 10 septembre 2022).
- Tutin A., Grossmann F., 2002, *Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif*, « Revue Française de Linguistique Appliquée » 7(1), pp. 7–25, <https://hal.archives-ouvertes.fr/hal-01336278/document> (consulté le 10 septembre 2022).

DICTIONNAIRES

Antidote 8, 2018 *Druide informatique inc.*

Brückner A., 1927, *Słownik etymologiczny języka polskiego*, Kraków.

Le Robert, <https://dictionnaire.lerobert.com> (consulté le 10 septembre 2022).

- LKEŽ: *Lietuvių kalbos etimologinio žodyno duomenų bazė*, <https://etimologija.baltne-xus.lt> (consulté le 10 septembre 2022).
- SJP PWN: *Słownik języka polskiego PWN*, <https://sjp.pwn.pl> (consulté le 10 septembre 2022).
- TLFi: *Trésor de la langue Française informatisé*, <http://www.atilf.fr/tlfi> (consulté le 10 septembre 2022).
- WSJP: *Wielki słownik języka polskiego PAN*, red. P. Źmigrodzki, <https://wsjp.pl> (consulté le 10 septembre 2022).

CORPUS

- DLKT: *Corpus du lituanien contemporain*, <http://tekstynas.vdu.lt/tekstynas> (consulté le 10 septembre 2022).
- Frantext: *Analyse et Traitement Informatique de la Langue Française (ATILF)*, <https://www.frantext.fr> (consulté le 10 septembre 2022).
- Leipzig Corpora Collection, https://corpora.uni-leipzig.de/fr?corpusId=fra_mixed_2012 (consulté le 10 septembre 2022).
- NKJP: *Narodowy Korpus Języka Polskiego*, <http://nkjp.pl> (consulté le 10 septembre 2022).
- Sketch Engine, <https://www.sketchengine.eu> (consulté le 10 septembre 2022).

SOURCES

- www 1, https://m.facebook.com/CharcuterieJoelJoly/photos/a.18278056540150-05/3369820293146859/?type=3&paipv=0&eav=AfZm76e_iRHYp7JiVkn-CHvIWZSw9taoTbqjh5Builhc--9HVPjjqrP5BQPC1FL74e_I&_rdr (consulté le 10 septembre 2022).
- www 2, <https://www.rtl.fr/actu/debats-societe/paques-2022-pourquoi-cherche-t-on-des-oeufs-en-chocolat-7900144058> (consulté le 10 septembre 2022).
- www 3, <https://www.festinoel.com/agenda-4005-2022-10-25.html> (consulté le 10 septembre 2022).
- www 4, <https://www.univ-artois.fr/toutes-les-actualites/journee-de-noel-solidaire-de-legass> (consulté le 10 septembre 2022).

**THE SENSE OF CELEBRATION OF THE WORDS *JOUR*, *JOURNÉE*
AND THEIR COLLOCATIONAL EQUIVALENTS
IN POLISH AND LITHUANIAN: CHALLENGES
FOR FRENCH LANGUAGE TEACHING**

Summary

This study is a continuation of the previous collaborative work regarding the analysis of the linguistic worldview of *fête* and its equivalents in Polish and Lithuanian, based on text corpora in the three languages. It revealed that, to understand the concept of celebration in these cultures, the linguistic worldviews of *jour*, *journée* and their equivalents in Polish – *dzień* and in Lithuanian – *diena* should also be explored. Moreover, a partial synonymy between *fête*, *jour* and *journée* may result in their incorrect use of mainly word patterns, both lexical and syntactic, which can be a potential source of difficulties for French language learners. These observations were a starting point for the present study aiming to identify the linguistic worldviews of *jour* and *journée* in the sense of celebration and to compare them with their Polish and Lithuanian equivalents. This research is multidisciplinary, its objective being mainly didactic, yet the methodology is at the crossroads of corpus linguistics and cultural linguistics. Therefore, the paper starts with a focus on the integration of text corpora in foreign language teaching and with the definition of collocation being a key notion of corpus analysis. It is followed by a presentation of the results of our empirical study. It shows that the word *journée* is more frequently used than *jour* while talking about celebrations and that most of the events denoted by the analysed words are civil, not religious, as in the case of the words *fête*, *święto* and *šventė*. The corpus analysis provided word patterns of *jour* and *journée* that are applicable to French language teaching.

Keywords: day, holiday, collocation, French language teaching

**JOUR, JOURNÉE W ZNACZENIU ŚWIĘTA I ICH KOŁOKACYJNE
ODPOWIEDNIKI W JĘZYKU POLSKIM I LITEWSKIM:
WYZWANIA W NAUCZANIU JĘZYKA FRANCUSKIEGO**

Streszczenie

Niniejsze badanie stanowi kontynuację wcześniejszych wspólnych prac obejmujących analizę językowego obrazu *święta* w językach francuskim, polskim i litewskim na podstawie korpusów tekstowych w tych trzech językach. Wykazało ono, że aby zrozumieć koncept świętowania w tych kulturach, należy również zbadać językowy obraz świata *jour* i *journée* oraz ich ekwiwalentów w języku

polskim *dzień* i litewskim *diena*. Częściowa synonimia pomiędzy *fête*, *jour* i *ournée* może skutkować niepoprawnym użyciem, a w szczególności stosowaniem nieprawidłowych wzorców łączliwości, zarówno leksykalnych, jak i składniowych, w konsekwencji być źródłem trudności dla Litwinów i Polaków uczących się języka francuskiego. W związku z tym niniejsze badanie opiera się na analizie i porównaniu językowych obrazów *jour*, *ournée*, *dzień* i *diena* w znaczeniu *święta*. Ma ono charakter interdyscyplinarny. Jego podstawowym celem jest dostarczenie rezultatów, które mogłyby znaleźć zastosowanie w nauczaniu języka francuskiego, natomiast metodologia została osadzona w językoznawstwie korpusowym i kulturowym. Na wstępie prezentowane są podstawy integracji korpusów tekstowych w nauczaniu języków obcych oraz definicja kluczowego terminu analizy korpusowej, jakim jest *kolokacja*. W części empirycznej przedstawione są natomiast wyniki badania. Wykazało ono, że słowo *ournée* jest częściej używane w badanej dziedzinie niż *jour* oraz że większość wydarzeń określanych analizowanymi słowami ma charakter cywilny, a nie religijny, jak w przypadku słów *święto*, *fête* et *šventė*. Analiza korpusowa dostarczyła także wzorców łączliwości *jour* i *ournée*, które mogą być wykorzystane w nauczaniu języka francuskiego.

Słowa kluczowe: dzień, święto, kolokacja, nauczanie języka francuskiego