

DOROTA ŚLIWA

PREDICATS POLONAIS ET FRANÇAIS RELATIFS
AUX ETAPES DE LA PANIFICATION ET LEUR APPLICATION
POUR ANALYSER LES CORPUS BILINGUES COMPARABLES

POLISH AND FRENCH PREDICATES RELATING
TO THE STAGES OF BREAD MAKING AND THEIR APPLICATION
TO ANALYZE COMPARABLE BILINGUAL CORPORA

Abstract

Comparable bilingual texts are representative for various fields of handicrafts. To meet the needs of writing bilingual dictionaries for the training of professionals, we propose a method of analysis from the Polish and French texts in the field of bakery. The purpose of this analysis is to establish equivalent units to be operated in the applied natural language processing and used for translation, for contrastive analysis and for drafting Polish-French dictionaries. In the first part of this paper, we explain how equivalent units are used in the automatic processing of comparable texts. In the following sections we present the design of cognitive script, the conceptual framework and predicates applied to the extraction of lexical units and we try to highlight significant relationships between them (meronymy and hyponymy).

Key words: comparable bilingual corpus; script; frame; predicate; Polish-French contrastive analysis; translation; dictionaries.

Les premières questions que l'on peut se poser sont celles-ci : pourquoi la panification ? Pourquoi des corpus bilingues comparables ? Comment analyser les prédicats relevés de ces corpus pour les appliquer au traitement automatique des textes bilingues ?

La panification (fabrication du pain) est choisie comme modèle d'analyse d'un domaine et des prédicats relevés d'un corpus bilingue comparable. En

outre, elle fait partie des activités du boulanger qui est un métier relevant de l'artisanat, autour duquel a été organisé le colloque « Termes et relations dans les corpus bilingues des domaines de l'artisanat » les 11-12 mars 2019 à Lyon. Nous avons retenu l'exemple de la baguette parce que ce produit est riche en valeurs culturelles. En Pologne la baguette est connue comme symbole de la France : on l'appelle couramment la « baguette française » (*bagietka francuska*). Or, en France, les meilleures baguettes sont faites à partir de la préfermentation *poolish* (« polonaise ») qui a son origine en Pologne, et ce sont les boulangers viennois qui l'introduisirent en France pour Marie Antoinette (1755-1793). Aujourd'hui, on trouve en France une *baguette poolish*³, fabriquée avec « un mélange de préfermentation d'origine polonaise, la *poolish* », considérée comme « un classique d'Origine ». C'est avec cette préfermentation qu'Antony Teixeira a obtenu le prix de la meilleure baguette de Paris en 2014.

Nul doute que la fabrication du pain, aliment de base, a son importance et son histoire. Dans le contexte actuel en Pologne et en France, la formation des professionnels de la boulangerie éveille toujours l'intérêt des instances officielles. Ainsi, dans le programme ministériel officiel polonais pour les qualifications professionnelles élaboré en 2017⁴, notamment pour la production des produits de boulangerie (T3) et pour la technologie de la boulangerie (T4), les auteurs attirent l'attention des professionnels sur l'apprentissage de la langue spécialisée étrangère et indiquent entre autres les dictionnaires spécialisés comme ressources didactiques. Pour le domaine étudié ici, il n'existe pas de dictionnaires polonais-français de la boulangerie.

Il y a donc de réels besoins et un défi à relever. Pour commencer, nous proposons une ébauche d'une méthode d'analyse des textes bilingues comparables du domaine de la boulangerie. Les activités artisanales sont définies et réglementées par chaque pays et les textes spécialisés de ces domaines sont rédigés sur la base des expériences et des normes propres à chaque pays. Les corpus bilingues comparables sont représentatifs de la majorité des textes dans divers domaines de l'artisanat, d'où l'intérêt de les choisir pour cons-

³ Voir aussi d'autres dénominations que l'on peut retrouver sur des sites de boulangerie : « la pouliche ou *poolish* », « la baguette polka », etc.

⁴ Voir la bibliographie de la source P1, le programme élaboré sur la base juridique suivante : Rozporządzenie Ministra Edukacji Narodowej z dnia 11 stycznia 2012 r. w sprawie kształcenia ustawicznego w formach pozaszkolnych (t.j. Dz. U. z 2014 r., poz. 622), Rozporządzenie Ministra Edukacji z dnia 7 lutego 2012 r. w sprawie podstawy programowej kształcenia w zawodach (Dz. U. 2012, poz. 184, z późn. zm.), Rozporządzenie Ministra Edukacji Narodowej z dnia 15 grudnia 2010 r. w sprawie praktycznej nauki zawodu (Dz. U. 2010, Nr 244, poz. 1626, z późn. zm.).

tituer une base de données de nature à permettre l'extraction des phrases parallèles pour la traduction automatique, pour les études contrastives ou pour l'élaboration d'un dictionnaire bilingue spécialisé.

En un premier temps, nous présenterons quelques tentatives précédentes pour établir des unités d'analyse des corpus bilingues comparables (segments ? script ? cadre ? prédicats ?) afin de les mettre en contextes parallèles. Ensuite seront analysés les prédicats polonais et français relatifs aux étapes de la panification, extraits des textes. Cette description sera suivie de la réflexion sur les relations significatives hiérarchiques en discours sur la panification, notamment les méronymes et les hyponymes.

1. ANALYSES DES CORPUS BILINGUES COMPARABLES ET LE DOMAINE DE LA BOULANGERIE

Généralement, les corpus comparables bilingues ou multilingues sont constitués par des textes qui ne sont pas traduits mais qui relèvent du même domaine, de la même époque, du même type, etc. Il est possible d'isoler dans ces textes des « unités » équivalentes avec des lexèmes et qui peuvent être alignées en vue d'un traitement automatique. L'extraction d'éléments de lexique bilingue à partir de corpus comparables est opérée déjà à partir des années 90, comme le signale Bouamour (2014 : 11) qui cite les chercheurs tels que Rapp, 1995; Fung et Yee, 1998; Chiao et Zweigenbaum, 2002; Gamallo, 2007; Kun et Tsujii, 2009. Citons aussi Déjean & Gaussier (2002) et Morin & Daille (2012) qui précisent les avantages des corpus comparables pour le travail du traducteur et les critères de la constitution de ces corpus.

Quelles méthodes adopter pour parvenir à des phrases parallèles des textes polonais et français du domaine de la boulangerie ? A la recherche des unités parallèles dans les textes comparables nous proposons d'une part, une adaptation des scripts et des cadres comme méthodes pour représenter la structure ontologique de manière organisée, d'autre part, les prédicats et leurs structures au sein d'un cadre conceptuel.

1.1. UNITÉS D'ANALYSE DES CORPUS BILINGUES COMPARABLES

L'identification des unités pour l'analyse contrastive et la traduction à partir des textes comparables nécessite un travail préparatoire du linguiste

qui donne des critères pour isoler ces unités et qui extrait les termes de leur contexte syntaxique en discours.

Une des unités couramment isolées du texte est un **segment**. Selon M. Hurault-Plantet et al. (2006), cité par Sliwa (2012), c'est une unité de discours thématique en « rupture thématique avec le segment précédent et avec le segment suivant ». Dans la constitution des corpus des textes comparables du domaine de la boulangerie, notamment les manuels et les présentations du travail de boulanger, ces unités sont très variées. Le segment relatif à la panification est facilement décomposable en unités lexicales, comme le signale l'énoncé définitoire d'un texte spécialisé français :

La panification est donc l'ensemble des étapes de la fabrication du pain, ponctué par des épisodes actifs, périodes de préparation et de travail de la pâte et des épisodes passifs, périodes pendant laquelle la pâte travaille seule, se repose, fermente ou cuit (F2).

Cependant, seul le critère thématique ne suffit pas. Nous proposons de relier des segments repérés dans les textes à un script cognitif, tel qui a été initié par Schank & Abelson (1977) en tant que formalismes de représentation de connaissance pour être implémentés en programmation.

Un **script cognitif** est, grosso modo, une construction conceptuelle réalisée à partir de la connaissance et de la perception d'une situation (prototypique ou plus ou moins individualisée) évoquée par un mot (les auteurs l'illustrent par l'exemple d'un repas au restaurant). La connaissance d'une situation, commune ou spécialisée, est constituée par des participants et des événements dans leurs relations temporelles (chronologie des événements), spatiales (lieu d'un événement), causales (cause-conséquence) et autres. Cette construction de connaissance dynamique relie les événements (actions) dans l'ordre chronologique (séquences temporelles) et par des liens logiques de cause-conséquence. Elle peut être structurée selon différents aspects (personnes, lieu, etc.).

Un segment isolé du texte analysé est donc relatif au script cognitif PANIFICATION constitué par les actions situées dans l'ordre chronologique et par les « éléments » définis (boulangier, pain, instruments, four, local).

L'ordre chronologique des événements d'un script est marqué par les étapes qui sont organisées en cadres conceptuels, tels qu'ils ont été initiés par Fillmore (1976) dans sa théorie sémantique des cadres et qui sont largement adoptés pour le traitement automatique des textes (voir par exemple

Baker 2009, Pimentel 2015, 2019). Baker (2009 : 32) explique que selon Fillmore « on ne peut comprendre la signification de bien des mots, de manière optimale, qu'en tenant compte du contexte événementiel ou situationnel dans lequel ils s'inscrivent ». Un **cadre conceptuel**, structurant la connaissance d'une étape, est composé par un événement auquel participent des « éléments » avec leurs rôles thématiques. Par exemple, le cadre [*mélange*], désigné par le nom prédicatif, nécessite les acteurs et les circonstances qui ont leur propre rôle thématique : 'agens', 'patients', 'résultat' ; 'instruments', 'lieu'. Ces structures ne sont pas isolées mais forment entre elles un réseau de relations causales (des liens de cause-conséquence intimement liées à l'ordre chronologique des étapes dicté par le script, p.ex. *dosage*→*mélange*→*division*→*formation*, etc.) ou de relations d'ordre hiérarchique, à l'intérieur de chaque cadre, entre les lexèmes désignant les concepts (le cadre est désigné par l'hyperonyme), ou encore de relations liées aux variantes textuelles (notamment entre le savoir commun exprimé par la langue courante et le savoir des professionnels exprimé par la langue spécialisée). L'intérêt d'adopter partiellement la méthode des cadres est de fournir une description organisée et exhaustive de la signification d'un mot qui évoque dans l'esprit des interlocuteurs tout un cadre conceptuel et non seulement un concept isolé.

Dans la présentation des formalismes des connaissances dans le traitement automatique du langage, en partant du script cognitif et passant par les cadres conceptuels, nous arrivons aux unités que sont les prédicats et les primitives, considérées toujours au niveau formel. Krzysztof Bogacki (2015) donne un aperçu des **prédicats** en logique et en linguistique, soulignant que le premier type est une « méthode qui représente un ensemble de critères et décide, pour un objet donné, si ces critères sont remplis ou non ». Le deuxième type, prédicats en linguistique, est « rapproché du schéma syntaxique du verbe ». Ces prédicats qui ont aussi une « caractéristique lexicale » (p.ex. remplacement, suppression, connaissance, perception, etc.) « sont dégagés de différentes manières ». L'auteur développe la présentation du problème par Bachimont (2000) qui inclut les prédicats (à côté des relations, des fonctions, des rôles thématiques, etc.) – mais aussi les « primitives de représentation » – à cet « appareil formel » qui « manipule » la couche « lexicale » (ou « sémantique ») existant à côté de ce niveau formel. Bogacki attire l'attention sur les **primitives** de représentation au niveau formel et sur les primitives sémantiques découvertes par le linguiste « en langue par un série de décompositions lexicales », et postule

d'entreprendre une « tentative de caractérisation » de ces unités élémentaires en vue de la construction d'une « ontologie qui conceptualise le réel ». L'auteur signale que ces ontologies peuvent être construites à partir des corpus de textes aussi, par l'extraction des termes avec leurs relations sémantiques, le plus souvent « la synonymie, la méronymie et l'hypéronymie ».

1.2. ESQUISSE DE LA MÉTHODE D'EXTRACTION DES PRÉDICATS

Les textes spécialisés d'un domaine présentant une connaissance du monde réel et un savoir faire décrivent une réalité qui est ensuite conceptualisée et représentée par les structures ontologiques. Organisées selon les principes de la métaphysique réaliste⁵, ces structures sont censées être reliées aux structures ontiques des entités du réel, et être désignées par des signes « transparents » que sont les unités lexicales de la prédication sur un aspect (une composante) du concept. Dans cette approche basée sur la métaphysique réaliste, le **prédictat** en tant qu'unité lexicale relevant de la prédication sur une propriété d'une entité du réel, a un sens lié à la structure ontologique et contribue à la construction d'un cadre conceptuel. Les prédicats ne sont pas des concepts ni des unités de l'informatique mais des unités lexicales (verbales le plus souvent) désignant directement les concepts et qui ont leur place fixe dans le cadre donné.

La conception des cadres conceptuels de Fillmore, telle qu'elle est adoptée en programmation, ne tient pas compte de la distinction des niveaux d'analyse en linguistique, notamment le niveau de la langue et le niveau des réalisations discursives inscrites dans les textes. Or, à l'étape de l'extraction, les unités d'un texte sont des unités discursives résultant des transformations syntaxiques (nominalisation ou passivation, le plus souvent). La distinction en niveaux d'analyse est opérée par d'autres chercheurs qui développent la conception de l'analyse distributionnelle-transformationnelle de Harris (1976)⁶ et que nous avons adoptée pour l'analyse des prédicats en tant

⁵ La conception intégrale du signe linguistique basée sur la métaphysique réaliste (Aristote – saint Thomas d'Aquin – Gilson – Krapiec – Maryniarczyk) en tant qu'une proposition pour l'approche cognitive en linguistique, a été présentée dans Sliwa 2017 et 2018.

⁶ En 1974 Harris a publié le livre *Notes on English Transformational Grammar*, Université de Paris VIII. Il a été traduit en 1976 par Maurice Gross: *Notes du course de syntaxe*, Paris, Editions du Seuil.

qu'expressions d'un concept (Śliwa 2017). Cette analyse permet de relier différentes variantes morphosyntaxiques de lexèmes en discours à leur phrase-source du niveau métalinguistique exprimant un aspect du cadre conceptuel. Un **cadre**, construction formelle avec des rôles thématiques, est rempli par des phrases-sources qui sont des unités de langue liées aux unités du discours et qui ont leur propres structures prédicatives-subjectivales (liées au contexte), à partir desquelles peuvent être élaborées des structures prédicatives-argumentales (indépendantes du contexte). A l'intérieur de chaque cadre les prédicats forment les phrases sources restituées du discours et désignent les mêmes acteurs et les mêmes circonstances, même s'ils relèvent des différents types de textes. Pour le domaine de la boulangerie, les rôles thématiques principaux que nous allons présenter sont 'Agens' action 'Patients/Résultat'. D'autres rôles 'Instrument, Local, etc.' constituant le contexte de la réalisation de l'action principale pourraient être l'objet d'autres analyses.

Dans les textes relatifs au script PANIFICATION et à ses étapes, nous notons de nombreux noms déverbaux, parmi lesquels il convient de distinguer les noms prédicatifs et les nominalisations des phrases. Le nom *panification* est un nom prédicatif, donc celui qui – pour le définir – nécessite un cadre conceptuel avec d'autres éléments conceptuels instanciant les rôles thématiques (sémantiques). Il est défini par TLFi⁷ comme 'ensemble des opérations consistant à transformer la farine en pain'. Il est, certes, dérivé du verbe dénominal *panifier*, lui-même dérivé du nom latin *panis* ('pain') avec le suffixe *-ifier*, mais il est difficile de le substituer par une phrase correspondante. Par contre, le synonyme de *panification* qu'est *fabrication du pain* illustre plutôt l'exemple de la nominalisation, car le nom *fabrication*, dérivé du verbe *fabriquer*, forme un syntagme avec le SP (*du pain*) qui correspond au rôle 'Patients' dans la phrase de base : *fabriquer le pain*.

Les noms prédicatifs sont plutôt d'ordre conceptuel et désignent un script ou un cadre, et gardent ainsi le statut d'un nom hyperonyme qui a le sens générique 'ensemble des opérations consistant à'. Les nominalisations sont des structures prédicatives transformées, et lorsqu'elles sont restituées avec tous les lexèmes instanciant les rôles thématiques du cadre, elles le complètent. Le cadre conceptuel donne la possibilité d'y insérer les prédications variées (selon les types de textes et les langues) sur un même aspect du concept, présentant ainsi les synonymes possibles. Les différents lexèmes au

⁷ Trésor de la Langue Française informatisé, <https://www.cnrtl.fr/definition/panification>, <https://www.cnrtl.fr/definition/panifier> (consulté en février 2019).

sein d'un cadre sont reliés entre eux par des relations de méronymie et d'hyponymie.

2. LES PRÉDICATS POLONAIS ET FRANÇAIS DÉSIGNANT LES ÉTAPES DE LA PANIFICATION

La conception de la construction de l'ontologie de la fabrication du pain telle que nous venons de la présenter vise à la relier au réel décrit et aux unités de langue, et cela de manière organisée. L'extraction préliminaire des prédicats des différents fragments des textes et ensuite l'application de la méthode d'analyse selon les cadres conceptuels nous a permis de proposer les noms prédicatifs polonais et français désignant les cadres organisés selon l'ordre chronologique. Et c'est par référence à ces cadres que nous pouvons voir les difficultés qui apparaissent pour l'extraction terminologique automatique, à savoir différentes étapes dans les textes où l'ordre chronologique est déplacé.

2.1. ELABORATION DES CADRES CONCEPTUELS DE LA PANIFICATION SELON L'ORDRE CHRONOLOGIQUE

Les étapes (*fazy – étapes*) organisées en cadres conceptuels qui sont marqués par un nom prédicatif désignant une étape du script sont caractérisées par les structures actanciennes des prédicats verbaux qui expriment la même action. Les rôles thématiques étudiés sont ici limités aux rôles les plus concernés : 'Agens' action 'Patients → Résultat'. Le rôle 'Agens' est instancié par deux lexèmes 'Agens : *boulangier, machine*', les rôles 'Patients → Résultat' par des lexèmes 'Patients → Résultat : *ingrédients, pâte, pâtons, pain*'. Le cadre conceptuel, qui contient les éléments nécessaires pour comprendre la signification de mots, est désigné par un nom prédicatif (p.ex. *mélange*). Il est composé par des acteurs et les circonstances qui ont leur rôle thématique dans la prédication sur la structure prédictive-actancielle. A partir de l'énoncé :

La première étape sert à mélanger les différents ingrédients qui composeront la future pâte à pain. Appelée aussi le frasage par les professionnels, ce mélange s'effectue ... (F5)

nous pouvons restituer la phrase-source ‘Agens’ *mélanger* ‘les ingrédients’. Cette phrase est reprise par la nominalisation *mélange* comme le signale le déterminant anaphorique *ce* précédant le nom, laissant dans l’implicite les lexèmes instanciant le rôle ‘Patiens : *les ingrédients*’. Le lexème synonyme *frasage* est introduit dans la prédication sur la même propriété par la glose *appelé aussi*. La mise en parallèle sémantique et syntaxique *mélange* = *frasage* nous autorise à restituer la phrase-source ‘Agens’ *fraser* ‘les ingrédients’. Ainsi, la même action est désignée par deux prédicats français : *mélanger* et *fraser*. C’est à l’intérieur d’un cadre qu’il sera possible d’analyser les dénominations d’une langue (polonais) suivant le type de texte et les mettre en parallèle avec les dénominations d’une autre langue (français).

A partir de la source polonaise (P1) qui est un programme de formation professionnelle de référence nous proposons les cadres [A – N] pour le script PANIFICATION

[A <i>dobieranie i dozowanie – sélection et pesage</i>]	‘Agens’	V	‘Patiens’
[B <i>mieszanie – mélange</i>]	‘Agens’	V	‘Patiens : <i>ingrédients</i> →Résultat : <i>pâte</i> ’
[C <i>wyrabianie – pétrissage</i>]	‘Agens’	V	‘Patiens : <i>pâte</i> ’
			↓
			[D <i>wyrastanie – pointage</i>] ‘Agens’ V
[E <i>dzielenie i odważanie – division et pesage</i>]	‘Agens’	V	‘Patiens : <i>pâte</i> →Résultat : <i>pâtons</i> ’
[F <i>formowanie – formation</i>]	‘Agens’	V	‘Patiens’
			↓
			[G <i>odprężenie – détente</i>] ‘Agens’ V
[H <i>kształtowanie – façonnage</i>]	‘Agens’	V	‘Patiens’
			↓
			[I <i>rozrost – apprêt</i>] ‘Agens’ V
[J <i>nacinanie – scarification</i>]	‘Agens’	V	‘Patiens : <i>pâton</i> /Résultat : <i>grigne</i> ’
[K <i>wkładanie do pieca – enfournement</i>]	‘Agens’	V	‘Patiens : <i>pâton</i> →Résultat : <i>pain</i> ’
			↓
			[L <i>pieczenie – cuisson</i>] ‘Agens’ V
[M <i>wyjmowanie z pieca – defournement</i>]	‘Agens’	V	‘Patiens : <i>pain</i> ’
			↓
			[N <i>schładzanie, wysychanie – ressuage</i>] ‘Agens’ V

Nous remarquons aussi l'alternance de la thématization lorsque le lexème instanciant le rôle 'Agens' désigne le boulanger ou la machine dans les étapes « actives », et lorsque le lexème désignant la pâte devient 'Agens'. Ce changement de rôle 'Patients' → 'Agens' attribué au lexème *pâte* est signalé suite aux actions du boulanger lors des étapes précédentes et se fait dans les étapes (D, G, I, L, N) pendant lesquelles « la pâte travaille seule, se repose, fermente ou cuit » (cf. F2), dites « passives » car l'agent humain n'exerce aucune action sur la pâte. Les spécialistes regroupent aussi les étapes dites de « fermentation »: D – *spoczynek /wyrastanie – pointage*, G – *odprężenie – détente*, I – *rozrost – apprêt*.

En principe, la panification est expliquée par les professionnels selon les étapes. Cependant, la lecture des énoncés des spécialistes laisse parfois le lecteur perplexe.

D'une part, dans la description d'une étape on peut finalement distinguer deux autres (cf. *rozrost* et *wypiek* dans (P3)) ou constater que certaines étapes sont omises, par exemple les étapes [H et I] dans l'explication de *rozrost* et [L] dans l'explication de *wypiek*.

L'ordre chronologique est conservé dans les textes polonais. Mais dans d'autres cas, les chronologies sont déplacées (cf. F4):

Avant de [K] mettre le pain au four, on peut [J] scarifier le pain pour éviter son éclatement et ainsi permettre son développement. (...) On [K] enfourne ensuite le pain dans un four ... (F5)

L'étape [J] précède chronologiquement l'étape [K], ce qui est signalé par le connecteur *avant que*, et ce qui montre la mise en relief de [J] pour en détailler les actions. Après cette explication le connecteur *ensuite* signale la chronologie rétablie. Le même changement de la chronologie est noté dans l'exemple (F4), avec les connecteurs *avant de* et *au moment de*.

Dans le travail d'extraction des lexèmes des textes spécialisés et leur mise en cadre, il est donc important d'être attentif aux connecteurs de temps, notamment dans les textes français.

2.2. ETAPES DE LA PANIFICATION

DÉTERMINÉES PAR LES LINGUISTES VS LES SPÉCIALISTES

La comparaison de la chronologie des actions établie dans les cadres conceptuels par les linguistes avec la chronologie des étapes donnée dans les

textes spécialisés prouvent qu'il y a d'autres difficultés rencontrées lors de l'extraction des termes, ce que montre le tableau 2 ci-dessous :

PANIFICATION		
Cadres	Source (F2)	Source (F5)
[A <i>sélection et pesage</i>]	1. pesage et prise des températures	–
[B <i>mélange</i>]	2. frassage	1. mélange des ingrédients
[C <i>pétrissage</i>]	3. pétrissage	2. pétrissage de la pâte
[D <i>pointage</i>]	5. pointage en masse ou piquage	3. pointage
[E <i>division et pesage</i>]	5. pesage – divisage	–
[F <i>formation</i>]	6. boulage	4. formation des pâtons et des boules
[G <i>détente</i>]	7. détente	5. détente
[H <i>façonnage</i>]	8. façonnage	6. façonnage
[I <i>apprêt</i>]	9. apprêt (2e fermentation)	7. apprêt et la seconde formation
[J <i>scarification</i>]	–	–
[K <i>enfournement</i>]	–	–
[L <i>cuisson</i>]	10. cuisson	8. la cuisson du pain
[M <i>defournement</i>]	–	–
[N <i>ressuage</i>]	–	–

Dans les cadres conceptuels peuvent être regroupés les prédicats synonymes selon le type de texte, à titre d'exemple :

- [B] *mélanger/ fraser des ingrédients* (cf. F5)
- [D] *pointer/pousser/piquer* (cf. F1)
- [F] *diviser/découper/fractionner/détailler les pâtons + former des boules/bouler* (F1, F5)
- [H] *façonner/donner la forme finale aux pâtons* (F2)

Les lexèmes synonymes sont par excellence liés au type du texte qu'il serait opportun de signaler dans le cadre donné.

3. LES RELATIONS SIGNIFICATIVES HIÉRARCHIQUES EN DISCOURS SUR LA PANIFICATION

Les relations significatives hiérarchiques relevées de l'analyse de notre corpus sont d'un côté la méronymie et l'holonymie, et d'un autre côté – l'hyponymie.

3.1. LES MÉRONYMES ET LES GLISSEMENTS MÉTONYMIQUES

La *méronymie* envisagée dans l'approche sémasiologique est définie en sémantique lexicale comme la relation entre un lexème dénotant une partie d'un tout et un lexème dénotant le tout dans lequel entre cette partie. Dans l'approche onomasiologique, il est question de la *partition*, c'est-à-dire de la relation entre une partie et un tout située au niveau ontologique⁸. Appliquée pour le traitement automatique par Winston et al. (1987) au sein des théories des réseaux sémantiques et des frames, elle est exploitée selon les six catégories de relations sémantiques de méronymie organisée autour de l'addition qui est une relation centrale. Nous avons constaté (Śliwa 2012) que ces « modèles de la logique formelle de la description de ces propriétés intervenant dans la relation méronymique, aussi bien dans la dimension hiérarchique (cf. héritage) que dans la dimension horizontale (cf. addition/connexion) », ne sont pas suffisants. Adoptant la logique aristotélicienne, nous avons placé la relation fondamentale de la **méronymie** désignant les relations ontologiques partitives (entre partie et tout) dans la structure de l'entité conceptuelle et dans sa logique (les relations internes entre les composantes d'une entité et les relations externes de cette entité avec les autres) assurant l'intégralité d'une entité. Étant donné qu'Aristote tenait compte de l'activité cognitive du sujet parlant, il nous a paru nécessaire d'« envisager la relation méronymique dans deux perspectives : partie → tout (les parties qui composent le tout) ou tout → partie (le tout qui est divisé en parties) » (Śliwa 2012).

Un des critères étudiés suite à la conception de Wilson et al. (1987) est celui de « séparabilité » que nous plaçons seulement dans la deuxième perspective, tandis que dans la première, le critère pertinent que nous introduisons est celui de « compositionnabilité ». La question fondamentale est ici

⁸ Les deux approches ont été distinguées dans Śliwa (2012) qui donne un historique des études sémantiques et terminologiques et d'où nous relevons les idées principales.

« pourquoi », posée dans l'une ou l'autre perspective, qui justifie la relation cause-effet.

Dans la perspective partie → tout, le critère de « compositionnabilité » opère dans deux dimensions : celle des étapes [A – M] qui composent l'entité de la panification (étapes → panification) et celle des composants (ingrédients avec leurs propriétés actives → pâte) soumis aux activités de l'agent dans les étapes [A – B], désignés par le méronyme (*skladniki – ingrédients*) en fonction de 'Patiens' qui composent un tout désigné par l'holonyme (*ciasto – pâte*) en fonction de 'Résultat'. Les relations dans cette perspective ont le trait « non homéomère ».

Dans la perspective tout → partie, est concernée seulement la nouvelle entité qui subit des transformations résultant de l'activité de l'agent (*ciasto → kęsy → chleb – pâte → pàtons → pain*). Regardant à chaque étape les méronymes en fonction de 'patiens/résultat', nous constatons que la « séparabilité » est pertinente seulement dans l'étape [E] *ciasto – pâte → kęsy – pàtons*. A la question (*pourquoi couper la pâte?*) posée dans cette perspective, nous obtenons la réponse ... *pour obtenir des pàtons*, donc, des entités nouvelles qui vont être soumises à des activités dans les étapes suivantes [F – K] et qui deviendront produit de la panification ([L] *kęsy – pàtons → chleb – pain*). L'entité qui a subi le processus de cuisson [M] est par la suite considérée comme un aliment, et la séparabilité est située dans un autre cadre, donc après la panification, entre le pain et la personne humaine : *couper le pain pour donner à manger aux personnes humaines*. Les relations dans cette perspective ont le trait « homéomère », ce qui facilite des glissements métonymiques (très fréquents dans les textes du domaine de la boulangerie) opérés sur les dénominations de parties dans les étapes qui sont récapitulées ci-dessous, selon le rôle sémantique du méronyme :

[A] 'Patiens : *skladniki – ingrédients*'

[B] 'Patiens : *skladniki – ingrédients*' → 'Résultat : *ciasto – pâte*'

[C] 'Patiens : *ciasto – pâte*'

[D] 'Agens : *ciasto – pâte*'

[E] 'Patiens : *ciasto – pâte*' → 'Résultat : *kęsy – pàtons*'

[F] 'Patiens : *kęsy – pàtons*'

[G] 'Agens : *kęsy – pàtons*'

[H] ‘Patiens : *kęsy – pátons*’

[I] ‘Agens : *kęsy – pátons*’

[J] et [K] ‘Patiens : *kęsy – pátons*’

[L] ‘Patiens : *kęsy – pátons*’ → ‘Résultat : *chleb – pain*’

[M] ‘Patiens : *chleb – pain*’

[N] ‘Agens : *chleb – pain*’

L’analyse des textes sources a permis de dégager quelques types de glissements métonymiques selon les dénominations de ‘Patiens/Résultat’.

En polonais :

[F → M] *formowanie kęsów* → *formowanie form chlebów* (P4)

[I → M] *rozrost kęsów* → *bochenki wyrosły* [K → M] *wkładanie kęsów do pieca* → *wsadzenie bochenków do pieca* (P3)⁹.

En français:

[G ← C] *détente du páton* ← *de la pátte* (F5)

[H → M] *façonnage du páton* → *façonner le pain* (F5)

[J → M] *scarification du páton* → *scarifier les pains* (F1)

[K → M] *enfournement du páton* → *enfourner le pain dans un four* (F5).

Ces analyses ont montré que le glissement métonymique le plus courant est opéré sur les dénominations *kęs* → *chleb – páton* → *pain* et qu’il est caractéristique plutôt dans les textes de vulgarisation (présentation de la fabrication du pain dans une boulangerie, recette, etc.). Les glissements métonymiques sont absents dans les textes des professionnels (P5). Dans l’adaptation de ces textes pour le traitement automatique, ces glissements situés au niveau de cadres conceptuels devraient être pris en compte, car ils font partie intégrante de l’activité discursive où entrent en jeu des facteurs illocutoires aussi (voir sur ce point Pietrzak-Porwicz 2006).

⁹ Le lexème polonais *bochenek chleba* a le sens de ‘forme du pain’ et désigne une unité du pain. Dans (P3) *bochenek* a lui-même un sens métonymique, celui de ‘pain’.

3.2. LES HYPONYMES ET LES SPÉCIFICATIONS SELON UN ASPECT DE L'ACTIVITÉ DU BOULANGER

L'étude des hyponymes nous situe sur l'axe vertical des relations hiérarchiques, celui de l'inclusion d'un concept spécifique dans un concept générique. Après de nombreuses discussions en linguistique sur le statut de l'hyperonyme/l'hyponyme (voir Georges Kleiber, Pierre Lerat, Robert Martin, rapportés par Dorota Śliwa 2017), il n'y a aucun doute que l'hyponyme est d'ordre métalinguistique et que l'hyperonymie/l'hyponymie vise la relation hiérarchique conceptuelle générique/spécifique. A ce titre, elle a sa place dans les études terminologiques (voir Teresa Cabré, Loïc Depecker, rapportés par Śliwa 2017, 2018). Dans notre analyse, l'**hyponyme** désigne le statut conceptuel du lexème dans la hiérarchie conceptuelle sous-tendant les lexèmes qui ont le même rôle sémantique au sein du même cadre. Une telle précision permet de voir la spécialisation dénominative du même concept, avec les moyens lexicaux dont dispose la langue donnée.

Le domaine de la panification comprend l'activité du boulanger qui est spécifiée par les techniques de panification, elles-mêmes conditionnées par le type de pain. Le lexème *pain* est un hyperonyme pour les hyponymes : *pain de froment*, *pain de seigle*... – selon le type de la farine; *pain français*, *pain polka* – selon la recette d'origine française, polonaise et aussi d'autres pays; *baguette*, *ficelle*, ... – selon la forme, etc. Ces relations significatives hiérarchiques sont placées dans les paradigmes des lexèmes qui ont le rôle 'Patients' et demandent une analyse plus approfondie de tous les cadres. Tenant compte des limites de cet article et de l'objectif d'esquisser une ébauche de la méthode d'analyse, nous signalons seulement les prédicats hyponymes dans les cadres [F] et [J].

Les prédicats hyponymes dans le cadre [F] expriment les liens hypero/hyponymiques selon la forme des pâtons. Ces liens sont signalés dans les textes par les anaphoriques ou par les énoncés définitoires.

En polonais, le lexème générique *formowanie* est défini par référence à une catégorie supérieure (*proces obróbki piekarniczej* ('processus d'opération de panification')) qui est ensuite spécifiée par deux actions spécialisées dont les dénominations *werkowanie* ('action de donner une forme ronde') et *łęgowanie* ('action de donner une forme longue') sont introduites par la glose *składający się z* ('qui est composé de'). Le lien entre l'hyperonyme (*formowanie*) et l'énoncé définitoire est signalé par l'anaphorique *to* ('cela') :

Po uzyskaniu odpowiedniej konsystencji masy piekarniczej piekarnia dokonuje tak zwanego formowania kęsów. Jest to proces obróbki piekarniczej, składający się z werkowania i lęgowania kęsów (P4).

Le deuxième texte illustre les liens entre les hyponymes synonymes reliant l'hyperonyme *ręczna obróbka* ('opération de panification manuelle') aux hyponymes *zaokrąglanie* ('action de donner une forme ronde') et *wydłużanie* ('action de donner une forme longue') par la glose explicative *polegająca na* ('qui consiste à'). Les hyponymes de langue générale sont accompagnés des lexèmes spécialisés mis entre parenthèses :

Gdy ciasto osiągnie pożądaną konsystencję, piekarze przystępują (...) do ręcznej obróbki, polegającej na **zaokrągleniu** (werkowaniu) i **wydłużaniu** (lęgowaniu) kęsów (P3).

En français, les lexèmes hyponymes sont en position de terme à définir dans un glossaire :

Boulage : Formation d'une boule de pâte régulière d'aspect (F2).

Dans un autre type de texte qui explique les étapes de la panification, le nom prédicatif *formation* désignant cette étape est suivi de la description de cette opération. A partir de la nominalisation *formation des pâtons et des boules* nous retrouvons la structure de la phrase-source dans laquelle les lexèmes *pâton* et *boule* en fonction de 'Patiens' sont définis (*Les pâtons sont des petits morceaux de la pâte. Les pâtons sont boulés et deviennent des boules*). Le prédicat verbal hyponyme *bouler* a subi la transformation de nominalisation (*boulage*) et se trouve dans l'énoncé définitoire avec la glose *consiste à*, et la transformation passive (*les pâtons seront boulés*) :

Formation des pâtons et des boules : La pâte est ensuite découpée en petit morceaux appelés aussi pâtons [...]. Vient ensuite ce que l'on nomme le boulage qui consiste à resserrer un peu les pâtons et à en faire comme son nom l'indique de petites **boules**. [...] les pâtons seront boulés plus ou moins serrés (F5).

Les analyses des fragments de textes relatifs au cadre [F] sont visualisées de la manière suivante :

[F] 'Agens' *formowanie* – *formation* 'Patiens' : *kęsy* – *pâtons* →
'Résultat' : *kule* – *boules* – hyponymes selon la forme

Formować kęsy – *Former des pâtons et des boules*

zaokrąglać/werkować kęsy → *kule* – *bouler les pâtons* → *boules*
wydłużać/lęgować kęsy – *allonger les pâtons*

A partir de cet échelon des textes sources, nous pouvons constater que les dénominations des ‘Patiens’ ont des hyponymes seulement en français, mais que la dénomination des prédicats est plus spécialisée en polonais.

Les prédicats hyponymes dans le cadre [J] expriment les liens hypero/hyponymiques selon les spécifications d’une action. Ils sont obtenus par la dérivation des verbes à partir des noms précisant l’instrument (*scarifier les pâtons avec une lame* → *lamer les pâtons*, cf. (F4)) ou le résultat (*faire des grignes sur les pâtons* → *grigner les pâtons*, cf. (F1)). Ils ont pour l’hyperonyme le prédicat *scarifier* (‘inciser superficiellement’). L’équivalent polonais *nacinać* (‘inciser superficiellement’) n’a pas d’hyponymes.

Les prédicats hyperonymes et hyponymes du cadre [J] sont visualisés de la manière suivante :

[J] ‘Agens’ *nacinianie* – *scarification* ‘Patiens : *kęsy* – *pâtons*’
 ‘Résultat : /-/ – *scarifications, grignes*’

Nacinać kęsy – *scarifier les pâtons avec une lame*
lamer (‘scarifier les pâtons avec une lame’ N ‘instrument’)
grigner (‘faire les grignes’ N ‘résultat’)

Les différences dans les dénominations des actions dans ce cadre, à savoir une plus grande diversité en français, sont conditionnées par le contexte culturel où l’on attache une plus grande importance aux valeurs artistiques du travail du boulanger en France qu’en Pologne, comme le prouve l’énoncé :

Avant de mettre le pain au four, on peut scarifier le pain [...]. Ces scarifications se nomment des grignes¹⁰ et jouent également un rôle dans l’esthétique du produit fini (F5).

L’analyse détaillée des segments que nous venons de voir permet d’identifier les liens entre les prédicats hyperonymes et hyponymes: les synonymes entre parenthèses (dans les textes polonais), les anaphores (indéfini *to* en polonais, déterminant démonstratif *ce* en français) ou les prédicats définitoires (*appeler, se nommer*, etc.).

¹⁰ Voir aussi <https://cuisine-facile.com/grignage-des-pains.php>.

ELÉMENTS DE LA CONCLUSION

Au terme de la présentation d'une ébauche de la méthode d'analyse proposée, nous avons prouvé qu'elle permet d'établir des équivalents français pour les termes polonais de manière précise et adéquate pour le contexte de communication professionnelle ou autre. Elle peut donner aussi des outils pour les informaticiens qui extraient les unités à partir des séquences parallèles dans la situation où il n'y a pas de textes parallèles. Elle permet aux linguistes d'affiner les analyses contrastives des unités lexicales tenant compte des transformations morphosyntaxiques des prédicats en discours et des liens logiques entre eux en fonction de leur place dans la structure ontologique de l'entité du réel représentée ici sous forme dynamique d'un script cognitif. Enfin, nous avons montré une méthode de recherche des équivalents à partir des textes et documents authentiques pour chaque langue, ce qui garantit une haute qualité de la traduction et de la communication.

Bien évidemment, cette ébauche de méthode d'analyse est limitée aux structures prédicatives avec des rôles essentiels, mais nous avons ouvert aussi la perspective pour développer par l'analyse d'autres rôles et par des expressions des liens cause-conséquence suivant l'ordre chronologique des actions dans le script.

BIBLIOGRAPHIE

- Afli Haïthem, Barrault Loïc & Schwenk Holger, 2012, « Traduction automatique à partir de corpus comparables: extraction de phrases parallèles à partir de données comparables multimodales », [in :] *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*, vol. 2: TALN, 447-454.
- Bachimont Bruno, 2000, « Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances », [in :] *Ingénierie des connaissances. Evolutions récentes et nouveau défis*, Charlet, Jean et al., eds., Eyrolles, 305-323. On-line: https://www.utc.fr/~bachimon/dokuwiki/_media/fr/ontologie-icbook.pdf, consulté le 19.07. 2015.
- Baker Collin F., 2009, « La sémantique des cadres et le projet FRAMENET : une approche différente de la notion de « valence » », *Langages*, n°176, 32-49.
- Bogacki Krzysztof, 2015, « Quelques réflexions sur les primitives, les prédicats, le sens des termes dans les ontologies et dans le lexique standard », *Roczniki Humanistyczne* 63, z. 8, 37-54.
- Bouamor Dhouha, 2014, *Constitution de ressources linguistiques multilingues à partir de corpus de textes parallèles et comparables*, Thèse de doctorat, Paris-Sud.
- Déjean Hervé & Gaussier Éric, 2002, *Une nouvelle approche à l'extraction de lexique bilingues à partir de corpus comparables* [consulté en mars 2019, <http://archive.wikiwix.com/cache/?url=http%3A%2F%2Fwww.cavi.univ-paris3.fr%2Ffleximetrika%2Fthema%2Fthema6%2FDejean.pdf>]

- Fillmore Charles J., 1976, « Frame Semantics and the Nature of Language », *Annals New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280, 20-32.
- Hurault-Plantet Martine, Jardino Michèle & Berthelin Jean-Baptiste, 2006, « Ajustement des frontières de segments thématiques détectés automatiquement », [in :] *Actes du 2ème dé fouilles de textes (DEFT)*, Suisse, Fribourg [consulté en novembre 2012, <http://perso.limsi.fr/Individu/jbb/limsi-def06.pdf>].
- Morin Emmanuel & Daille Béatrice, 2012, « Compositionnalité et contextes issus de corpus comparables par la traduction terminologique », [in :] *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*, vol. 2: TALN, 141-154.
- Pietrzak-Porwisz Grażyna, 2006, « Metonymia w ujęciu kognitywnym », *Prace Językoznawcze*, n°8, 23-39.
- Pimentel Janine, 2015, « Using Frame Semantics to Build a Bilingual Lexical Resource on Legal Terminology », [in :] Kockaert Hendrik & Steurs Frida [eds.], *Handbook of Terminology (HoT)*, vol. 1, Amsterdam, John Benjamins, 427-450.
- Pimentel Janine, 2017, « Specialized verbs and specialized uses of verbs in a comparable corpus of judgments produced in Canada, Portugal and Brazil », [in:] Drouin Patrick, Francoeur Aline, Humbley John, Pinçon Aurélie, (org.). *Terminology and Lexicography Research and Practice*, Amsterdam–Philadelphia, John Benjamins Publishing Company, 109-130.
- Schank Roger & Abelson Robert P., 1977, *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*, New Jersey: Erlbaum.
- Śliwa Dorota, 2012, « La méronymie dans un segment du guide touristique – vers une analyse cohérente des corpus comparables bilingues », *Roczniki Humanistyczne* 60, z. 8, 81-112.
- Śliwa Dorota, 2017, « Prédicats en tant qu'expressions des relations du concept ? Etude des termes juridiques français et polonais du 'mariage' », *Roczniki Humanistyczne* 65, z. 8, 57-71. DOI: <http://dx.doi.org/10.18290/rh.2017.65.8-5>.
- Śliwa Dorota, 2018, « Termin i jego relacje w ujęciu poznawczy », *Orbis Linguarum* 49, 205-216. DOI: 10.23817/olin.50-1.

LES SOURCES POLONAISES (CONSULTÉES EN FÉVRIER 2019)

- P1** : Anna Krawczyk, Angelika Kruszka, Dorota Molendowska & Sylwia Wasilewska-Dąbrowska, 2017, *Program nauczania dla kwalifikacji* (t. 3, *Produkcja wyrobów piekarskich*, t. 4, *Technologia piekarstwa*), Kwalifikacyjny kurs zawodowy, ORE, Krajowy Ośrodek Wspierania Edukacji Zawodowej i Ustawicznej.
- P2** : Szymon Konkol, 2012, *Almanach Cukierniczo-Piekarski*, t. 5, „Produkcja piekarska”, *Materiały dydaktyczne dla ucznia i nauczyciela*, <http://www.mamz.pl/almanach/>.
- P3** : http://piekarniarochna.pl/artykuly/jak_to_jest_zrobione/produkcja_chleba.html.
- P4** : <https://www.pkt.pl/artykul/jak-wyglada-produkcja-tradycyjnego-chleba-10607>.
- P5** : Adam Mełkowski, 2017, *Pieczywo żytnie na kwasie, część 4 – technologia*, <http://m.mistrzbranzy.pl/artykuly/pokaz/id/3712>.

LES SOURCES FRANÇAISES (CONSULTÉES EN FÉVRIER 2019)

- F1** : https://fr.wikipedia.org/wiki/Fabrication_du_pain.
- F2** : Centre de Recherche et d'Etude de la Boulangerie et de ses compagnonnages, <http://www.compagnons-boulangers-pâtisseries.com/crebesc/les-etapes-de-la-panification/>.

F3 : Centre de Recherche et d'Etude de la Boulangerie et de ses compagnonnages http://levainbio.com/cb/crebesc/elaboration-du-pain/?internal_redirect=1.

F4 : <http://www.lalosparis.com/la-fabrication-du-pain/>.

F5 : <https://www.dmoz.fr/loisirs/gastronomie/la-fabrication-du-pain-principes-et-terminologie-de-boulangier/>.

PRÉDICATS POLONAIS ET FRANÇAIS RELATIFS AUX ÉTAPES DE LA PANIFICATION ET LEUR APPLICATION POUR ANALYSER LES CORPUS BILINGUES COMPARABLES

Résumé

Les textes bilingues comparables sont représentatifs pour différents domaines de l'artisanat. Pour répondre aux besoins de la rédaction de dictionnaires bilingues pour la formation des professionnels, nous proposons dans cet article une méthode d'analyse à partir de textes polonais et français du domaine de la boulangerie. Le but de cette analyse est d'établir des unités équivalentes qui seront exploitées et appliquées dans le traitement automatique des langues et qui serviront pour la traduction, pour les analyses contrastives et pour la rédaction des dictionnaires polonais-français. Dans la première partie de l'article nous expliquons d'abord les unités équivalentes utilisées dans le traitement automatique de textes comparables. Dans les parties suivantes nous présentons la conception du script cognitif, du cadre conceptuel et des prédicats appliquée pour l'extraction des unités lexicales et pour déterminer les relations significatives entre elles (méronymie et hyponymie).

Mots-clé : corpus bilingues comparables; script; cadre; prédicat; analyse contrastive polonais-français; traduction; dictionnaires.

POLSKIE I FRANCUSKIE PREDYKATY ZWIĄZANE Z ETAPAMI WYPIEKU CHLEBA ORAZ ICH ZASTOSOWANIE W ANALIZIE DWUJĘZYCZNYCH TEKSTÓW PORÓWNYWALNYCH

Streszczenie

Teksty porównywalne są reprezentatywne dla różnych dziedzin rzemiosła. Wobec zapotrzebowania na dwujęzyczne słowniki z tych dziedzin w kształceniu specjalistów, zaproponowana została w tym artykule metoda analizy na przykładzie polskich i francuskich tekstów z dziedziny piekarstwa. Celem tej analizy jest ustalenie jednostek ekwiwalentnych, które będą zastosowane w językoznawstwie komputerowym na potrzeby w zakresie tłumaczeń, analiz kontrastywnych oraz tworzenia słowników polsko-francuskich. W pierwszej części wyjaśnione są jednostki ekwiwalentne stosowane w językoznawstwie komputerowym, a w następnych koncepcja skryptu poznawczego, ram konceptualnych oraz predykatów zastosowana do ekstrakcji jednostek leksykalnych oraz określenia ich relacji znaczeniowych (meronimia i hiponimia).

Słowa kluczowe: dwujęzyczne teksty porównywalne; skrypt; ramy; predykat; analiza kontrastywna polsko-francuska; tłumaczenie; słowniki.