

NEOPHILOLOGICA

37



UNIWERSYTET ŚLĄSKI
WYDAWNICTWO

NEOPHILOLOGICA

volume 37

*Sémantique, lexique, système :
Hommage à Jurij Apresjan
(1930–2024)*

sous la rédaction de
Svetlana Krylosova, Wiesław Banyś et Beata Śmigielska

Président d'honneur / Honorary President : **Gaston GROSS** † (Université Sorbonne Paris Nord, France)
Rédacteur en chef / Editor-in-Chief: **Wiesław BANYŚ** (Université de Silésie, Pologne)
Rédacteur en chef adjoint / Deputy Editor-in-Chief: **Beata ŚMIGIELSKA** (Université de Silésie, Pologne)

COMITÉ DE RÉDACTION / EDITORIAL BOARD

Denis APOTHÉLOZ	Université Nancy 2, Nancy, FR
Olivier AZAM	École Normale Supérieure, Paris, FR
Vincent BÉNET	INALCO, Paris, FR
Xavier BLANCO ESCODA	Universitat Autònoma de Barcelona, Barcelona, ES
Christine BONNOT	INALCO, Paris, FR
Anna BONOLA	Università Cattolica del Sacro Cuore, Milano, IT
Laura CALABRESE	Université libre de Bruxelles, Bruxelles, BE
Eric CORRE	Université Sorbonne Nouvelle – Paris 3, Paris, FR
Michael DANIEL	Collegium de Lyon, Université de Lyon, Lyon, FR
Jean-Pierre DESCLÉS	Université Paris-Sorbonne, Paris, FR
Giovanni DOTOLI	Università degli Studi di Bari Aldo Moro, Bari, IT
Marco FASCIOLO	Sorbonne Université, Paris, FR
Paolo FRASSI	Università degli Studi di Verona, Verona, IT
Gaston GROSS †	Université Sorbonne Paris Nord, Villetaneuse, FR
Francis GROSSMANN	Université Grenoble Alpes, Grenoble, FR
Zlatka GUENTCHÉVA	CNRS, Paris, FR
Georges KLEIBER	Université de Strasbourg, Strasbourg, FR
Mihail KOPOTEV	University of Helsinki, Helsinki, FI
Svetlana KRYLOSOVA	INALCO, Paris, FR
Anna KRZYŻANOWSKA	Uniwersytet Marii Curie-Skłodowskiej, Lublin, PL
Katarzyna KWAPISZ-OSADNIK	Uniwersytet Śląski, Katowice, PL
Denis LE PESANT	Université Paris Nanterre, Nanterre, FR
Sébastien MARENGO	Université de Montréal, Montréal, CA
Fabrice MARSAC	Université de Strasbourg, Strasbourg, FR
Salah MEJRI	Université Paris 13, Villetaneuse, FR
Igor MELČUK	Université de Montréal, Montréal, CA
Jasmina MILIĆEVIĆ	Dalhousie University, Halifax, CA
Teresa MURYN	Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej, Kraków, PL
Franck NEVEU	Sorbonne Université, Paris, FR
José A. PASCUAL RODRÍGUEZ	Real Academia Española, Madrid, ES
Hélène de PENANROS	INALCO, Paris, FR
Alain POLGUÈRE	Université de Lorraine, Nancy, FR
Michele PRANDI	Università di Bologna, Bologna, IT
Dan van RAEMDONCK	Université libre de Bruxelles, Bruxelles, BE
Dejan STOSIC	Université Toulouse – Jean Jaurès, Toulouse, FR
Monika SUŁKOWSKA	Uniwersytet Śląski, Katowice, PL
Marcela ŚWIĄTKOWSKA	Uniwersytet Jagielloński, Kraków, PL
Stéphane VIELLARD	Sorbonne Université, Paris, FR
Joanna WILK-RACIĘSKA	Uniwersytet Śląski, Katowice, PL

CORRECTION LINGUISTIQUE / LANGUAGE EDITORS

Paweł GOLDA (français), Marzena WYSOCKA-NAREWSKA (anglais)

SECRÉTARIAT DE RÉDACTION / EDITORIAL SECRETARY

Anna CZEKAJ anna.czekaj@us.edu.pl

Institut de Linguistique, Université de Silésie à Katowice
ul. Grot-Roweckiego 5, 41-205 Sosnowiec, Pologne

Accessible sous forme électronique en accès libre / Available in open access electronic form:

www.journals.us.edu.pl/index.php/NEO/

Central and Eastern European Online Library www.ceeol.com

Śląska Biblioteka Cyfrowa www.sbc.org.pl

TABLE DES MATIÈRES

Svetlana KRYLOSOVA : Présentation

Igor BOGUSLAVSKIJ, Leonid IOMDIN¹, Leonid KRYVIN : *Jurij Apresjan (1930–2024)*

Igor Mel'čuk : So-called independent infinitives in Russian: The construction of the form *Ètogo mne ne ponjat'* 'It is impossible for me to understand this'

Lidija IORDANSKAJA : A Pioneering Lexicographic Work: "New Explanatory Dictionary of Russian Synonyms" by Jurij Apresjan

Alina ISRAELI : Impact of the meanings of the verb *moč'* on its ability to form past tense and perfective

Alexei SHMELEV : The Russian Words for Anxiety Revisited

Magdalena DANIELEWICZOWA : Unqualifiable meta-predicative adverbials in Polish

Lucie BARQUE : Is metonymy really more regular than metaphor? A first investigation into the French Lexicon

Sergey SAY, Ilja A. SERŽANT : A Frequency-Based Algorithm for Argument Extraction from Russian Treebanks

Elena URYSON : Sur les verbes multiplicatifs russes (étude pilote)

Valentina APRESJAN, Vladimir PLUNGIAN, Ekaterina RAKHILINA : *Behind the door*: how Russian, Kazakh and Armenian encode spatial relations with double-fronted objects

Irina LEVONTINA : Noun as a Verb

Dorota SIKORA : Le verbe polonais *tesknić* : portraits lexicographiques d'une polysémie

Irina KOR CHAHINE : When The Moment Does Not Last: A Few Words On *Moment*-Constructions in Russian

Alexander LETUCHIJ : Partial observation and nonstandard tense uses: Russian perfective verbs in progressive interpretation

Tilmann REUTHER : Remarks on Ju. D. Apresjan's lexicographic terminology: *leksema* and *vokabula* from a multilingual perspective

Rémi CAMUS : *Poskakali!* « Filons! » / « Assez couru! » : construire l'instabilité (les verbes *nesti*, *nosit'*, *skakat'* et le préfixe *po-*

Svetlana KRYLOSOVA : Définir pour apprendre et apprendre à définir : repenser la lexicographie pour l'enseignement des langues

¹ Leonid Iomdin, ami et collaborateur de longue date de Jurij Apresjan, est décédé le 18 septembre 2024, peu avant la finalisation de ce volume.

CONTENTS

Svetlana KRYLOSOVA: Presentation

Igor BOGUSLAVSKIJ, Leonid IOMDIN¹, Leonid KRYSIN: *Jurij Apresjan (1930–2024)*

Igor Mel'čuk: So-called independent infinitives in Russian: The construction of the form *Ètogo mne ne ponjat'* 'It is impossible for me to understand this'

Lidija IORDANSKAJA: A Pioneering Lexicographic Work: "New Explanatory Dictionary of Russian Synonyms" by Jurij Apresjan

Alina ISRAELI: Impact of the meanings of the verb *moč'* on its ability to form past tense and perfective

Alexei SHMELEV: The Russian Words for Anxiety Revisited

Magdalena DANIELEWICZOWA: Unqualifiable meta-predicative adverbials in Polish

Lucie BARQUE: Is metonymy really more regular than metaphor? A first investigation into the French Lexicon

Sergey SAY, Ilja A. SERŽANT: A Frequency-Based Algorithm for Argument Extraction from Russian Treebanks

Elena URYSON: On Russian multiplicative verbs (a pilot study)

Valentina APRESJAN, Vladimir PLUNGIAN, Ekaterina RAKHILINA: *Behind the door*: how Russian, Kazakh and Armenian encode spatial relations with double-fronted objects

Irina LEVONTINA: Noun as a Verb

Dorota SIKORA: The Polish verb *tęsknić*: lexicographic portraits of a polysemy

Irina KOR CHAHINE: When The Moment Does Not Last: A Few Words On *Moment*-Constructions in Russian

Alexander LETUCHIJ: Partial observation and nonstandard tense uses: Russian perfective verbs in progressive interpretation

Tilmann REUTHER: Remarks on Ju. D. Apresjan's lexicographic terminology: *leksema* and *vokabula* from a multilingual perspective

Rémi CAMUS: *Poskakali!* "Let's Flee!" / "Enough Running!": Constructing Instability (The verbs *nesti*, *nosit'*, *skakat'* and the prefix *po-*)

Svetlana KRYLOSOVA: Defining to Learn and Learning to Define: Rethinking Lexicography for Language Teaching

¹ Leonid Iomdin, a long-time friend and collaborator of Jurij Apresjan, passed away on September 18, 2024, shortly before the completion of this volume.



Présentation*

À la mémoire de Jurij Derenikovič Apresjan (1930–2024)

I.

Le 12 mai 2024, Jurij Apresjan s’est éteint à Moscou. Avec lui, c’est toute une époque de la linguistique qui s’achève — une époque qu’il a marquée par son exigence intellectuelle, sa clarté conceptuelle et sa fidélité profonde à la langue. Sa disparition a suscité une vive émotion parmi trois générations de linguistes. Pour beaucoup, il fut un maître, un guide, un modèle, mais aussi un ami intime et fidèle — un ami pour la vie. Pour d’autres, il a demeuré une voix entendue à distance, mais dont la précision, la rigueur et l’élégance ont laissé une empreinte durable. Et pour tous, il restera un grand savant, dont l’œuvre, faite de découvertes majeures, continue de nourrir la pensée linguistique.

Dès ses premières années de formation, Apresjan s’est imposé par une rigueur intellectuelle hors du commun. Refusant tout compromis, il s’est souvent heurté aux logiques idéologiques et institutionnelles soviétiques, sans pourtant céder jamais sur ses principes. Doté d’une intuition linguistique rare et d’une exigence scientifique intransigeante, il a mené ses recherches avec une précision exemplaire, animé par un sens aigu du détail et de la responsabilité scientifique. Il vivait et pensait « d’une manière incroyablement scrupuleuse, sans jamais faire

* Les paragraphes introductifs de ce texte (Partie I) ont été rédigés en étroite collaboration avec Igor Mel’čuk et Lidija Iordanskaja; cette partie s’appuie sur deux textes qu’Igor Mel’čuk a consacrés à Apresjan : Mel’čuk (2024a) et Mel’čuk (2024b).

de concessions sur ce qu'il tenait pour juste — ni dans la science, ni dans la vie » (Buras, 2022 : 25).

Ce numéro de *Neophilologica*, dédié à Jurij Apresjan, rassemble des contributions très diverses : celles de ses collègues proches, compagnons de pensée et de travail pendant des décennies ; celles des chercheurs qui l'ont côtoyé plus brièvement, mais qui se reconnaissent comme ses disciples ; et celles, enfin, des linguistes qui ne l'ont jamais rencontré, mais dont le parcours intellectuel a été profondément influencé par son œuvre. Tous partagent une même dette, à la fois intellectuelle et humaine.

Parmi ses réalisations les plus marquantes figure la fondation d'un domaine entier de la linguistique : la sémantique lexicale. Son ouvrage *Sémantique lexicale : les moyens synonymiques de la langue* (1974 [en russe], 1980 et 2000 [en polonais], 1992 [en anglais]) a façonné le développement de ce champ pendant plus d'un demi-siècle. Encore aujourd'hui, *Sémantique lexicale* demeure une source inépuisable d'idées et de données linguistiques. Apresjan y affirme que la synonymie — autrement dit, la *paraphrase* — constitue le phénomène central de la sémantique linguistique, socle de toute théorie et de toute description rigoureuse du sens langagier. Cette thèse a été au cœur de ses recherches dès les débuts, comme en témoignent son dictionnaire de synonymes de l'anglais (Apresjan, 1979) et le *Nouveau dictionnaire explicatif des synonymes du russe* (Apresjan, 2023b), sans équivalent dans aucune autre langue (voir l'article de Lidija Iordanskaja dans ce volume).

Pour résumer — de manière nécessairement simplifiée — sa pensée linguistique, on peut mettre en avant un postulat fondamental et deux principes directeurs qui ont guidé son activité de chercheur :

Postulat fondamental : le lexique doit être au centre des recherches linguistiques. Cette orientation allait à contre-courant des paradigmes dominants dans les années 1960–1970, marqués par l'hégémonie de la syntaxe (« Rien que la syntaxe ! »). Malgré les résistances, Apresjan a su recentrer la linguistique russe sur le lexique. L'étude approfondie, systémique et pluraliste du vocabulaire est devenue emblématique de l'école linguistique russe contemporaine.

Principe 1 : *La langue doit être décrite comme un tout.*

Le mot célèbre de Saussure — « La langue est un système où tout se tient » — doit être pris au pied de la lettre. Apresjan insistait sur la nécessité d'une description intégrale de la langue, dans laquelle la grammaire (syntaxe et morphologie) est en harmonie avec le lexique, qui, à son tour, doit être décrit de façon rigoureusement formalisée. C'est cette conviction qui a conduit Apresjan à élaborer le concept de *lexicographie systémique*, où le fonds lexical est décrit comme un système structuré.

Principe 2 : *Chaque lexème est un univers en soi.*

Pour chaque lexème (c'est-à-dire chaque mot pris dans une acception bien déterminée), Apresjan exigeait un *portrait lexicographique* (Apresjan, 1992b, 1992c) — une description formelle et détaillée de tous ses emplois et de toutes ses particularités sémantiques, syntaxiques, prosodiques et pragmatiques. Une telle entreprise repose sur une analyse fine d'un immense corpus de données, fondée non sur des statistiques, mais sur l'intuition, la logique et l'expérience du linguiste. Sur ce plan, Apresjan n'avait pratiquement pas d'égal — sinon Igor Mel'čuk et Anna Wierzbicka (voir, par exemple, les textes d'Irina Kor-Chahine, d'Aleksej Shmelev et de Dorota Sikora dans ce volume).

Apresjan était lui-même un lexicographe accompli. Le *Grand dictionnaire anglais-russe* (Apresjan, 1993–1994), qu'il a dirigé, est considéré comme le meilleur de son genre. Fort de cette expérience à la fois théorique et pratique, il a lancé l'un des projets les plus ambitieux de sa carrière : le *Dictionnaire actif de la langue russe* (Apresjan, 2010, 2014a, 2014b, 2017, 2023a, 2024 ; Babaeva et al., 2025¹). Ce dernier s'inscrit à l'intersection de deux traditions : d'une part, celle des dictionnaires actifs européens, conçus pour répondre aux besoins pratiques, où la sélection et la présentation des informations sont guidées par des objectifs pédagogiques ; d'autre part, celle du *Dictionnaire explicatif et combinatoire* (Mel'čuk & Žolkovskij, 1984 [pour le russe] ; Mel'čuk et al., 1995 [pour le français]), qui offre une base théorique solide pour l'analyse systémique du lexique, grâce notamment à des définitions lexicographiques hautement structurées, à des schémas de régime élaborés et à l'usage des fonctions lexicales (voir l'article de Svetlana Krylosova).

Le *Dictionnaire actif* reprend cet appareil dans l'organisation interne de ses articles, tout en le maintenant « en coulisses » : il guide la description lexicographique sans en alourdir la lecture. Ce compromis rend possible une conciliation rare entre richesse descriptive, cohérence théorique et lisibilité, y compris pour des usagers non spécialistes. Ce monument de cohérence théorique et de précision descriptive illustre à la perfection une lexicographie moderne, animée par un amour profond de la langue — condition première, selon Apresjan, de toute véritable linguistique.

À côté de ces réalisations majeures, Apresjan a aussi proposé une série d'idées et de concepts qui ont profondément influencé la discipline, par exemple :

¹ Après le décès de Ju. D. Apresjan, le travail sur le *Dictionnaire actif de la langue russe* s'est poursuivi sous la direction de ses collaborateurs. Le projet demeure fondé sur la conception théorique, les principes lexicographiques et l'architecture générale élaborés par Apresjan.

• **La notion de polysémie régulière**, systématiquement décrite par Apresjan. Il s'agit de la description des relations sémantiques récurrentes entre les différentes acceptions d'un mot qui ne relèvent pas de coïncidences isolées, mais de mécanismes productifs faisant partie du système de la langue. Ces régularités, analogues à celles des structures dérivationnelles, permettent de prédire la coexistence de certaines significations à partir de schémas sémantiques stables. Ainsi, les noms de récipients (*seau*, *verre*, *tasse*, etc.) présentent, à côté du sens d'objet fabriqué destiné à contenir un liquide, un sens de type quantitatif: *Il a bu un verre d'eau* signifie 'il a bu une quantité d'eau correspondant au contenu typique d'un verre'. Ce type de relation, attesté de manière régulière pour d'autres unités lexicales, manifeste une organisation non arbitraire du lexique fondée sur des correspondances sémantiques structurées (voir l'article de Lucie Barque).

• **Les composantes sémantiques non triviales « prêtes à utiliser »**. Pourquoi la phrase *Une touriste japonaise était fière du saule pleureur au bord de notre rivière* paraît-elle étrange, voire légèrement comique ? Ce n'est pas que l'on ne puisse être fier d'un arbre — mais ici, l'association entre *être fier* et un élément aussi extérieur au sujet semble artificielle. L'explication, comme l'a proposé Apresjan, réside dans un composant sémantique implicite du verbe russe *gordit'sja* 'être fier' : l'objet de la fierté doit appartenir à la *sphère personnelle du sujet* (Apresjan, 1985, 1995 : 636–648). Cette contrainte doit être prise en compte dans la description lexicographique du verbe *gordit'sja* (et de son équivalent français) ainsi que sa valence ou ses restrictions aspectuelles. Voici, à titre d'illustration, la définition proposée dans le *Dictionnaire actif de la langue russe* (Apresjan, 2014 : 389, traduction nôtre, ici et dans les citations suivantes) :

VYGOVARIVAT' 3, IMPF, vieillissant : 'La personne A1 exprime à la personne A3, qui occupe un statut social inférieur ou est d'un âge inférieur, son mécontentement concernant l'action A4 accomplie par A3 ou par quelqu'un appartenant à la sphère personnelle de A3, en lui adressant les paroles A2'.

Vygovarivat' prisluge za ploxo vymytyj pol 'réprimander les domestiques pour le sol mal lavé' | *Vygovarivat' povaru za perežarennye kotlety* 'réprimander le cuisinier pour des boulettes trop cuites'

La notion de **sphère personnelle** est implémentée par une composante sémantique réutilisable, que l'on retrouve dans de nombreuses définitions lexicographiques.

On peut également mentionner une autre composante sémantique de ce type proposée par Apresjan : '**l'observateur**' — une personne fictive postulée par le narrateur et servant de point de référence dans la structuration de l'espace, qui

est représentée dans le discours (Apresjan 1995 : 639–644 ; voir, dans ce numéro, les textes de Valentina Apresjan, Vladimir Plungian & Ekaterina Rakhilina et d'Alexander Letuchy). Des prépositions comme *devant* ou *derrière*, des adverbes comme *au loin*, *à droite*, *à gauche*, ou encore des verbes comme *apparaître* ou *se profiler*, supposent un point de vue orienté. Que veut dire la phrase *Devant la montagne s'étendait un lac* ? Elle implique, selon Apresjan, tout d'abord, que le lac se trouve entre la montagne et l'observateur présumé dans cette phrase. Mais cela ne suffit pas : il faut aussi que, du point de vue de l'observateur (ou du Locuteur), le lac et la montagne soient perçus comme de taille comparable. Si cette condition n'est pas remplie, l'énoncé devient pour le moins étrange — comparez, par exemple *'Devant le bâtiment de l'usine se trouvait un pois chiche*. Une autre condition tient à la distance apparente entre les deux éléments : même si, en réalité, la montagne est éloignée de plusieurs kilomètres, il suffit qu'elle semble proche du lac pour que l'énoncé soit interprété comme naturel. Un exemple tiré du *Dictionnaire actif du russe* illustre clairement la place de l'observateur dans la définition d'un adverbe spatial (Apresjan, 2017 : 17) :

DALEKO 1.1. ('loin') ADV. *Daleko ot A2* ('Loin de A2') : 'à une grande distance du repère spatial A2 ou de l'observateur'.

La distinction rigoureuse entre sens et connotation

Les mots *osël* ('âne') et *išak* ('âne') partagent le même sens², mais se distinguent nettement par leurs connotations : *rabotat' kak osël* signifie 'travailler bêtement', tandis que *rabotat' kak išak* évoque un travail excessif, pénible et résigné. Apresjan insiste sur la nécessité de distinguer, dans les définitions lexicographiques, *le contenu sémantique objectif* (le sens) et *la valeur expressive ou culturelle* (la connotation) associée au lexème par la langue. Cette distinction est essentielle à l'établissement d'un portrait lexicographique précis (Apresjan, 1992d).

Apresjan a revalorisé le rôle de la *connotation* dans la structuration du lexique, en particulier dans l'analyse de la polysémie. Il montre que la connotation du sens de base peut servir de pont sémantique pour dériver un nouveau sens. Ainsi, le mot *petux* ('coq') peut désigner, par extension, un 'homme agressif' — non pas en vertu d'une composante sémantique explicite, mais par la connotation de combativité associée au coq. Ce trait évaluatif est culturellement stabilisé et doit, selon Apresjan, être intégré de manière explicite dans la définition du sens figuré : *petux2* = 'homme agressif, comme le coq1 par sa connotation de combativité'.

² IŠAK désigne plus spécifiquement un âne domestique, tel qu'il est traditionnellement présent en Asie centrale.

En reconnaissant ainsi un statut linguistique aux connotations, Apresjan permet de traiter un grand nombre de cas de polysémie, de dérivation ou de phraséologisation sans recourir à des entités artificielles ou à des homonymes injustifiés, tout en restant au plus près de l'intuition des locuteurs. On peut à ce titre citer ici les définitions proposées par le *Dictionnaire actif du russe des lexèmes* VORONA1 ('corbeau') et VORONA2 ('personne distraite'), voir Apresjan, 2014b : 257.

VORONA1 ('corbeau') NOM COMMUN : 'oiseau sauvage de taille légèrement inférieure à celle d'une poule, au bec puissant, au plumage noir et gris, omnivore, vivant à proximité des habitations humaines en ville et à la campagne, émettant des sons caractéristiques'.

Connotations : distraction, inattention.

VORONA2, employé comme prédicat syntaxique, figuré, familier : 'personne distraite et inattentive, qui perd ou fait souvent tomber des objets' (par analogie avec les connotations de distraction et d'inattention associées à VORONA1).

• **Le lien entre l'anomalie linguistique et certains phénomènes grammaticaux ou « sémantiques profonds »** (comme la présupposition ou le cadre modal). L'établissement de ce lien permet de distinguer l'anomalie linguistique des contradictions logiques ou des absurdités de sens. Par exemple, l'énoncé **Voici un chevaux* est agrammatical, car le numéral *un* exige un nom au singulier. En revanche, une reformulation comme *un ensemble de chevaux, au nombre d'un* — bien qu'inhabituelle et même absurde — est grammaticalement correcte. De même, la phrase *Ce n'est pas vrai que c'est un livre* est grammaticalement acceptable, tandis que **Ce n'est pas vrai que voici un livre* est agrammaticale. Cela tient au fait que le phrasillon (terme de L. Tesnière) *voici* comporte une composante énonciative impliquant le Locuteur (la 1^{re} personne du singulier, c'est-à-dire, 'moi') — « je t'indique que Y est ici » — qui entre en conflit avec le cadre modal introduit par *ce n'est pas vrai que*. L'incompatibilité ne relève donc pas d'une simple erreur formelle, mais d'une contradiction plus profonde dans l'organisation du sens (Apresjan, 1978, 1995 : 598–622).

• **L'intégration des données sur la prosodie lexicalisée dans les dictionnaires**, autrement dit, la prise en compte des phénomènes prosodiques systématiquement associés à certaines unités lexicales. Apresjan distingue plusieurs cas où la prosodie n'est pas un simple effet de contexte, mais un élément sémantique à part entière qui participe pleinement à la structuration du sens (voir les articles d'Alina Israeli et de Magdalena Danielewiczowa). Apresjan en tire la conclusion qu'une description lexicographique rigoureuse d'une lexie ne peut faire l'éco-

nomie des faits prosodiques, dès lors qu'ils sont réguliers et porteurs de sens (Apresjan, 2014b : 245–246).

VOOBŠČE1.3, ADV., généralement avec une négation exprimée ou implicite. *Voobšče ne A1* : 'On aurait pu s'attendre à ce que, sous certaines conditions, dans un certain sens ou dans une certaine mesure, A1 ait lieu ou puisse avoir lieu ; le locuteur indique que A1 n'a lieu en aucune condition, sous aucun rapport et en aucune mesure'. *Porte toujours un accent de phrase.*

- On ↓voobšče ne byl v Pariže.

Il n'est ↓absolument jamais allé à Paris.

VOOBŠČE3, PARTICULE

Voobšče A1 : 'A1 ; le locuteur se réserve le droit d'apporter des restrictions ou des précisions à une affirmation qu'il considère trop générale'.

Ne porte jamais un accent de phrase.

- Ja voobšče prosto tak prišel.

Je suis venu sans raison particulière, en fait.

• **Le critère de Green–Apresjan** est un outil particulièrement utile pour diagnostiquer la polysémie, en la distinguant de la disjonction dans la définition lexicographique (Green, 1969 ; Apresjan, 1974). Il repose sur l'idée qu'il y a véritablement deux sens bien distincts plutôt qu'un sens disjonctif lorsque deux interprétations d'un même mot ne peuvent pas coexister dans un même énoncé sans produire un effet d'incompatibilité ou de dissonance interprétative. Si deux lectures supposées d'une unité lexicale ne peuvent pas être activées simultanément dans un même contexte syntaxique (la coordination étant une mise à l'épreuve fréquente), c'est qu'il s'agit de deux lexèmes distincts. L'effet absurde, voire comique, que produit l'énoncé *'Il a envoyé un colis et ses meilleurs vœux à ses parents* suggère précisément que les deux sens du verbe *envoyer* ne peuvent pas coexister dans une même définition lexicographique : c'est là un symptôme clair de polysémie. Ce critère permet ainsi de tester la disjonction des sens non pas seulement par introspection, mais aussi par l'impossibilité de leur combinaison dans une structure syntaxique partagée.

• Le concept de quark sémantique

Suivant la proposition d'Apresjan, ce terme désigne un élément minimal de sens, plus fin encore qu'un sémantème (c'est-à-dire que le signifié d'une lexie de la langue), et qui donc n'est pas lexicalisé dans la langue considérée. Il joue cependant un rôle essentiel dans la description des sens lexicaux. Un quark sémantique ne correspond à aucun mot en usage — il n'a pas de forme autonome —, mais

il permet de rendre compte des distinctions de sens subtiles entre des lexèmes proches, voire entre des constructions grammaticales. Ces unités infra-lexicales, à la fois insaisissables et indispensables, permettent d'expliquer des régularités dans le comportement syntaxique ou prosodique de certains verbes (Apresjan, 1994)³. Un exemple d'un *quark sémantique* cité par Apresjan est la **stativité**. En effet, les verbes statifs comme *videt'* 'voir' ou *xotet'* 'vouloir' et beaucoup d'autres qui désignent des états homogènes, des propriétés ou des relations⁴, se comportent de manière largement similaire :

- a) Ils présentent un *paradigme grammatical déficient* : ils n'ont ni formes impératives, ni formes passives. On ne peut pas dire en russe **Xoti uexat'!* 'Veuille partir !', **Kartina videlas' imi* 'Le tableau se voyait par eux'.
- b) Ils présentent un *paradigme sémantique déficient* : ils ne peuvent pas être utilisés au sens duratif actuel, prophétique ou dans certains autres usages de l'imperfectif. Ainsi, on ne peut pas dire **On vošël, kogda ja slyšal kakie-to strannye zvuki v uglu* 'Il est entré quand j'entendais des sons étranges dans un coin'.
- c) Ils ne se combinent pas avec des circonstants de but, ni avec des circonstants de temps inclusifs, ni avec la plupart des circonstants de manière, et ne peuvent pas être subordonnés à des verbes comme *zanimat'sja* ou *delat'* 'faire'. Cf. : **S cel'ju postupit' v universitet, on znaet matematiku* 'Afin d'entrer à l'université, il connaît les mathématiques'; **On postepenno sčitat' eti pretenzii nesostojatel'nymi* 'Il considérerait progressivement ces revendications comme infondées', **On zanimal'sja tol'ko tem, čto gordilsja svoimi uspehami* 'Il ne faisait qu'être fier de ses succès'.

Il est évident que cette similarité dans les comportements des verbes statifs est sémantiquement motivée. Nous devons supposer qu'il y a dans leurs sens un *noyau sémantique commun*, mais si réduit qu'il ne peut pas se matérialiser sous la forme d'un mot de la langue naturelle dans leurs définitions. Il suffit de jeter un coup d'œil rapide à la liste de verbes statifs (cf. note 4 de bas de page) pour se convaincre qu'aucune explication raisonnable du comportement commun de ces verbes ne peut contenir des composantes sémantiques communes. La solution est exactement le quark de stativité dont sont munis tous ces verbes.

³ L'article examine les similitudes et les différences entre les écoles sémantiques de Moscou et de Pologne dans leur approche du métalangage sémantique.

⁴ Par exemple, *slyšat'* 'entendre', *želat'* 'désirer', *znat'* 'savoir', *sčitat'* 'considérer', *dumat'* (čto) 'penser (que)', *gordit'sja* 'être fier', *stydit'sja* 'avoir honte', *zavidovat'* 'envier', *stoit'* 'valoir', *vesit'* 'peser', *izobražat'* 'représenter' [cf. *Le tableau représente une forêt d'hiver*], *otnosit'sja* 'appartenir à' [cf. *Les pumas appartiennent à la famille des félins*], *vysit'sja* 'se dresser', *belet'* 'apparaître blanc', etc.

Et ce n'est là qu'un bref aperçu. Il faudrait encore évoquer la théorie de l'antonymie d'Apresjan, ses travaux sur les quasi-synonymes, sur les marqueurs communicatifs, sur l'aspect verbal... Il est tout simplement impossible d'en dresser ici une liste exhaustive.

Le style scientifique d'Apresjan se distingue par une clarté cristalline, une rigueur impeccable, une concision sans compromis et une élégance presque artistique. Ses textes se lisent avec aisance et plaisir, quelle que soit leur complexité. Il avait un mot pour désigner cette qualité — *neprinuzhdenost'* (« aisance, naturel ») — et il écrivait avec cette grâce fluide qui emporte l'adhésion sans jamais forcer.

Dans les années 1970, en pleine stagnation brejnévienne, Apresjan fut licencié de l'Institut de la langue russe de l'Académie des sciences de l'URSS pour avoir signé une lettre publique de soutien aux écrivains Andreï Siniavski et Ioulii Daniel, emprisonnés pour avoir publié leurs œuvres à l'étranger. Refusant de se plier aux diktats idéologiques, il rejoignit un institut ministériel (« Informlektro »), où il fonda un laboratoire de linguistique pour les proscrits de l'Académie. Ce groupe, affecté à la traduction automatique, permit l'élaboration de descriptions inédites de la syntaxe du russe. Une fois encore, même dans les conditions les plus absurdes, Apresjan transforma l'épreuve en réussite scientifique.

Avec la *perestroïka*, il revint à l'Académie des sciences. Il y dirigea de nouveaux projets lexicographiques, s'employa à combattre les résidus du soviétisme et poursuivit inlassablement sa tâche de linguiste. Jusqu'à la fin de sa vie, il resta fidèle à son engagement civique : en février 2022, il fut l'un des premiers signataires de la lettre ouverte « Nous exigeons l'arrêt immédiat de tous les actes de guerre dirigés contre l'Ukraine », publiée dans *Le Monde* et cosignée par 664 scientifiques et chercheurs russes. Par cet acte, Apresjan confirma, une fois encore, son refus de toute compromission avec les autorités totalitaires⁵.

Jusqu'au bout, il aura porté la même exigence de rigueur, d'engagement et d'amour de la langue.

Il a « érigé un monument plus durable que le bronze » : son œuvre immense, ses quatre petits-enfants, et un royaume invisible d'admiration et d'affection dans des centaines de cœurs.

⁵ Appel de 664 chercheurs et scientifiques russes : « Nous exigeons l'arrêt immédiat de tous les actes de guerre dirigés contre l'Ukraine » (2022). *Le Monde*. Publié le 25 février 2022 [https://www.lemonde.fr/idees/article/2022/02/25/appele-de-664-chercheurs-et-scientifiques-russes-nous-exigeons-l-arret-immediat-de-tous-les-actes-de-guerre-diriges-contre-l-ukraine_6115263_3232.html].

Ce recueil se veut une manière, parmi d'autres, de prolonger la présence de Jurij Apresjan dans notre discipline. Il témoigne de ce que son œuvre a semé : des recherches, des vocations, des interrogations, des fidélités. Et nous espérons qu'il fera entendre, à travers la diversité des voix réunies ici, quelque chose de l'exigence tranquille et profonde qui fut la sienne.

II.

Le volume s'ouvre sur un texte commémoratif signé par Igor Boguslavskij, Leonid Iomdin⁶ et Leonid Krysin, *Jurij D. Apresjan (1930–2024)*, qui rend hommage à la trajectoire scientifique, humaine et intellectuelle de Jurij Derenikovič Apresjan. Ce texte retrace les grandes étapes de sa carrière, depuis ses débuts dans la linguistique structurale jusqu'à la fondation de la Sémantique intégrale et de la Lexicographie systématique. Les auteurs soulignent notamment le rôle central d'Apresjan dans la constitution de l'École sémantique de Moscou, son engagement civique dans les années 1970, son influence dans le développement des systèmes de traduction automatique, ainsi que la portée théorique de ses dictionnaires. Ce témoignage collectif, nourri d'une longue proximité personnelle et intellectuelle avec Apresjan, met en lumière à la fois l'ampleur de son héritage scientifique et les qualités humaines qui ont marqué plusieurs générations de linguistes.

Ce texte est suivi d'un ensemble de contributions portant sur les trois domaines chers à Apresjan : la lexicographie, la sémantique et la syntaxe. Ces travaux témoignent, chacun à leur manière, de la fécondité des questions qu'il a posées et de l'héritage qu'il laisse à la linguistique contemporaine.

L'article « So-called Independent Infinitives in Russian: The Construction of the Form *Ètogo mne ne ponjat'* » d'Igor Mel'čuk offre une description lexicographique formelle de deux constructions russes de type « nom au datif + NE + infinitif », exprimant l'impossibilité : *Mne ne vstat'* 'Il m'est impossible de me lever' et *Mne ne rabotat' tam* 'Il me sera impossible de travailler là-bas'. Contrairement à l'approche courante, qui considère ces infinitifs comme indépendants, Mel'čuk soutient qu'il s'agit d'une locution syntaxiquement discontinue (sans lien syntaxique direct entre leurs composantes) : [X-u] 'BYT' ... NE⁻¹I [Y-nut'], où

⁶ Leonid Iomdin, ami et collaborateur de longue date de Jurij Apresjan, est décédé le 18 septembre 2024, peu avant la finalisation de ce volume.

l'infinitif dépend syntaxiquement du verbe *быть*. L'article présente une modélisation complète de ces constructions aux niveaux sémantique, syntaxique profond et syntaxique de surface, accompagnée de notices lexicographiques détaillées. Entre autres, ce travail plaide pour une reconnaissance lexicographique à part entière des locutions, au même titre que les lexèmes ordinaires.

Dans « *A Pioneering Lexicographic Work: New Explanatory Dictionary of Russian Synonyms* by Jurij Apresjan », Lidija Iordanskaja propose une lecture analytique du *Nouveau dictionnaire explicatif des synonymes du russe*, conçu par Apresjan à la fin des années 1990, à la lumière des principes théoriques de la lexicographie systématique. L'article met l'accent sur la portée scientifique du projet, en particulier sur les innovations méthodologiques introduites par Apresjan : notions de *portrait lexicographique*, de *dominante d'une série synonymique*, etc. Iordanskaja montre comment ces outils permettent de décrire de façon rigoureuse les relations synonymiques, en intégrant des dimensions jusqu'alors absentes des dictionnaires classiques, comme le statut référentiel et communicatif des lexèmes et la possibilité d'emploi en énoncé autonome. Sans se limiter à une présentation de contenu, l'article propose une véritable évaluation critique du modèle lexicographique élaboré par Apresjan, en soulignant à la fois sa puissance théorique, sa richesse descriptive et son accessibilité pour le lecteur. Cette contribution est d'autant plus précieuse que ce dictionnaire, malgré son importance, demeure relativement peu connu des linguistes non russophones.

Dans son article « *The Meanings of the Verb мо́чь* », Alina Israeli propose une analyse détaillée des valeurs sémantiques du verbe modal russe *мо́чь* 'pouvoir', en examinant ses principaux emplois et en montrant dans quels cas ils peuvent s'articuler avec le perfectif *смо́ж*. L'étude revient sur une question à la fois classique et complexe : la polysémie de *мо́чь*, comparable à celle des verbes modaux analogues dans d'autres langues indo-européennes. L'auteure s'attache à clarifier les différentes acceptions de ce verbe en contexte, en mettant en relation ses usages présents avec ceux de *смо́ж*. Sept acceptions fondamentales sont distinguées : une acception d'aptitude interne, trois acceptions d'aptitude externe ou de permission, et trois acceptions épistémiques. Chaque cas est analysé en fonction de son contexte d'emploi, de sa structure syntaxique et de son schéma prosodique propre. L'article met aussi en lumière les disparités dans la distribution temporelle de ces acceptions : certaines sont compatibles avec le passé ou le futur, d'autres non, et la relation entre *мо́чь* et *смо́ж* ne suit pas toujours un strict parallélisme aspectuel. L'un des apports majeurs du travail est d'aller au-delà de la seule description sémantique pour intégrer l'analyse de l'intonation, ce qui ouvre vers une approche pragmatique attentive aux corrélations entre modalité, aspect, prosodie et structure des énoncés.

Intitulée « The Russian Words for Anxiety Revisited », l'étude d'Aleksej Shmelev offre une approche lexicographique des mots russes désignant les états psychologiques/affectifs d'inquiétude et d'anxiété. Elle porte plus précisément sur les familles lexicales de *bespokoit'* 'inquiéter' et *trevoga* 'angoisse, anxiété', à partir des sous-corpus parallèles russe-anglais et anglais-russe du *Corpus national de la langue russe* (ruscorpora.ru). L'auteur prolonge ici les analyses d'Apresjan, en confrontant les définitions issues du *Nouveau dictionnaire explicatif des synonymes* (Apresjan, 2023b) et du *Dictionnaire actif du russe* (Apresjan, 2014a) aux données empiriques issues de la traduction. L'étude met en évidence une distinction récurrente : les mots du champ de *bespokoit'* 'inquiéter' impliquent une évaluation rationnelle du cours des événements, tandis que ceux relevant de *trevoga* 'angoisse' renvoient à une émotion plus diffuse, plus intense et plus difficilement contrôlable. L'analyse des équivalents anglais (*worry*, *anxious*, *uneasy*, *alarm*, etc.) illustre la manière dont la traduction fait ressortir des composantes sémantiques distinctes. L'un des apports méthodologiques essentiels de cet article réside ainsi dans l'exploitation des corpus parallèles, qui permet de faire émerger des traits parfois invisibles dans les textes originaux. L'auteur souligne en particulier la polarité négative du verbe *bespokoit'sja*, fréquent à la forme négative ou à l'impératif, et discute le rôle des stratégies de traduction (concordance vs sens contextuel) dans l'analyse sémantique. Cette contribution éclaire les mécanismes de lexicalisation des états affectifs en russe, tout en mettant en lumière la complexité des facteurs – sémantiques, contextuels et pragmatiques – qui influencent leur traduction.

Dans « Unqualifiable Meta-Predicative Adverbials in Polish », Magdalena Danielewiczowa propose une description approfondie d'un sous-ensemble particulier d'adverbes polonais en *-o* / *-e*. Ces adverbes, tels que *iscie królewski* 'vraiment royal', *dosłownie (idiotyczny)* 'littéralement idiotique' ou *jawnie (wrogi)* 'manifestement hostile', se distinguent à la fois des adverbes de manière ordinaires et des adverbes énonciatifs ou métatextuels. L'auteure les qualifie de métaprédicatifs non qualifiables, car ils qualifient une lexie prédicative sans pour autant pouvoir être eux-mêmes modifiés, intensifiés ou niés. L'analyse repose sur une série de tests syntaxiques, prosodiques et sémantiques permettant de circonscrire cette classe particulière. Ces adverbes sont caractérisés par leur impossibilité d'avoir des degrés, d'être coordonnés ou niés, ainsi que par leur position fixe dans l'énoncé. L'étude examine également la place de ces unités dans la structure thématique et rhématique des énoncés, et leur lien avec les expressions épistémiques. Un des mérites du travail est d'articuler les niveaux phonétique, syntaxique, sémantique, informationnel et contextuel, afin de proposer une catégorisation fine de ces adverbiaux, tout en posant des questions plus larges sur leur

statut par rapport aux particules et aux adverbes métatextuels. L'article met ainsi en lumière un phénomène peu étudié mais central pour la compréhension des interactions entre syntaxe, prosodie et pragmatique.

Lucie Barque consacre son article « Is Metonymy Really More Regular Than Metaphor? A First Investigation into the French Lexicon » à une étude quantitative sur la polysémie régulière dans le lexique français. Depuis le travail fondateur d'Apresjan (1974), la tradition linguistique associe généralement la métonymie à la régularité, et la métaphore à l'irrégularité. L'auteure met cette hypothèse à l'épreuve en analysant un échantillon de près de 3000 noms français et en identifiant les alternances de sens impliquant un changement de catégorie sémantique (par exemple *mouton* 'animal' / 'personne crédule', ou *pied* 'partie du corps' / 'pied de table'). L'étude montre que les schémas métonymiques réguliers sont plus nombreux que les schémas métaphoriques, mais que les deux figures produisent, en moyenne, des proportions comparables de mots polysémiques. Ce travail propose ainsi un éclairage nuancé sur la relation entre figure lexicale et polysémie régulière. Il met à disposition du lecteur un ensemble de données factuelles, tout en ouvrant des perspectives méthodologiques : affiner les classifications sémantiques et diversifier les annotations permettra de mieux comprendre la manière dont les locuteurs perçoivent les liens entre les sens d'un même mot.

Sergey Say et Ilja Seržant présentent dans « A Frequency-Based Algorithm for Argument Extraction from Russian Treebanks » un algorithme permettant d'identifier automatiquement les arguments verbaux (= les actants) dans les corpus arborés du russe (Universal Dependencies, UD). Leur approche s'inscrit dans l'héritage théorique d'Apresjan, en combinant méthode distributionnelle et ancrage sémantique, et s'appuie notamment sur le *Dictionnaire actif de la langue russe*. L'analyse des données russes met en lumière des problèmes d'annotation, notamment l'étiquetage incohérent des dépendances dans les corpus UD. Les résultats révèlent une bonne précision, avec un faible taux de faux positifs : par exemple, dans l'énoncé *On načal besedu* 'Il a commencé une conversation', l'algorithme identifie correctement *besedu* comme objet direct de *načal*, contrairement à certaines annotations UD erronées. Le travail illustre une intégration réussie entre cadre théorique linguistique et méthodologie computationnelle. Au-delà du russe, cette méthode offre un outil prometteur pour l'analyse du corpus dans des langues peu dotées en ressources lexicographiques. Elle s'inscrit ainsi dans une perspective typologique fondée sur l'analyse empirique des données, attentive à la variabilité des constructions actantielles à travers les langues.

Dans son article « Sur les verbes multiplicatifs russes (étude pilote) », Elena Uryson propose une redéfinition de cette classe verbale et examine l'ambiguïté de certains énoncés qui les contiennent. Inscrite dans la tradition de l'École

sémantique de Moscou, l'étude montre que les multiplicatifs se caractérisent par le focus qu'ils placent sur la répétition d'actes de même nature — c'est la manière dont s'accomplit l'action qui est mise en avant. Cela les distingue de verbes comme *begat'* 'courir' ou *kosit'* 'faucher [l'herbe]', où le but ou le résultat priment sur la série des actes. L'analyse repose sur un ensemble de verbes tels que *barabanit'* 'tambouriner [avec les doigts sur la table]', *bul'kat'* 'glouglouter', *morgat'* 'cligner de l'œil', *maxat'* 'agiter [la main]' ou *xlopat'* 'applaudir'. L'article distingue deux sous-classes: les multiplicatifs stricts (*mercat'* 'scintiller'), qui désignent nécessairement une pluralité d'actes et n'ont pas de forme semelfactive correspondante, et les multiplicatifs non stricts (*stučat'* 'toquer', *maxat'* 'agiter [la main]'), qui admettent deux interprétations ('une fois' ou 'plusieurs fois') et forment souvent un couple avec un semelfactif comme *stuknut'* ou *maxnut'*. L'un des apports majeurs de cette étude est de montrer que la variabilité d'interprétation de ces verbes est une propriété lexicale stable, et qu'elle permet de mieux comprendre leur place dans le système aspectuel du russe.

Valentina Apresjan, Vladimir Plungian et Ekaterina Rakhilina s'interrogent dans « *Behind the door: How Russian, Kazakh and Armenian Encode Spatial Relations with Double-Fronted Objects* » sur la manière dont les langues encodent la frontalité des objets à double façade, tels que les portes ou les fenêtres, dans les constructions spatiales. Selon la définition classique d'Apresjan, la frontalité n'est pas une simple propriété géométrique, mais une caractéristique fonctionnelle, liée à l'usage typique des objets par les humains. Pourtant, contrairement à ce que l'on pourrait supposer, cette propriété cognitive ne se traduit pas partout de la même façon dans les systèmes linguistiques. L'étude en question s'appuie sur des données des trois langues, et montre que les langues varient quant à la conceptualisation de tels objets. En russe, les constructions sont loin d'être unificables: *pered dver'ju* 'devant la porte' peut désigner aussi bien l'espace intérieur qu'extérieur, tandis que *za oknom* 'derrière la fenêtre' désigne systématiquement l'espace extérieur. En revanche, les syntagmes *za dver'ju* et *pered oknom* sont ambigus, et leur interprétation peut dépendre du point de vue du locuteur. En kazakh et en arménien, d'autres systèmes sont en place: en kazakh, la postposition *aldynda* sert à désigner l'espace « devant », quel que soit le contexte; en arménien, l'alternance entre *dimac* 'face de' (avec frontalité) et *araj* 'devant' (plus neutre) permet de préciser si l'objet est traité comme fronté ou non. Cette étude montre que la frontalité des objets à double façade relève de systèmes typologiques distincts, où interagissent la structure grammaticale, les conventions pragmatiques et les cadres cognitifs activés par les locuteurs.

Irina Levontina consacre son article « *Noun as a Verb* » à l'analyse d'un phénomène pragmatique et syntaxique particulier du russe: l'emploi de noms

pour réaliser des actes de parole performatifs, en particulier dans les requêtes. L'exemple emblématique est la construction *Našedšego pros'ba vernut'* 'Que la personne qui a trouvé (les clés) les rende', litt. 'Trouveur_{ACC} prière rendre', largement utilisée, bien que non standard. Dans les requêtes, l'usage de *pros'ba* 'prière' permet d'éviter l'emploi direct des formes à la première ou à la deuxième personne, jugées trop intrusives dans certains contextes. L'article décrit en détail les schémas syntaxiques associés à cet emploi. Le nom *pros'ba* 'prière' peut introduire l'allocutaire :

- au datif : *Pros'ba passažiram zanjat' svoi mesta*, litt. 'Prière aux passagers_{DAT} prendre leurs places' ;
- par une subordonnée : *Esli est' anesteziolog, pros'ba projti v salon*, litt. 'S'il y a un anesthésiste, prière se rendre dans la cabine' ;
- par une adresse explicite : *Tovarišči, pros'ba razojtis'*, litt. 'Camarades, prière vous disperser' ;
- et — fait le plus remarquable — à l'accusatif : *Kolleg_{ACC} pros'ba ne šutit'*, litt. 'Prière les collègues ne pas plaisanter', configuration exceptionnelle en russe, vraisemblablement héritée du verbe transitif *prosit'* 'demander'.

Cette particularité révèle un lien inattendu entre performativité et comportement syntaxique des noms, montrant comment les actes de parole peuvent influencer directement la structure grammaticale. Plus largement, l'étude éclaire la manière dont la langue articule politesse, distanciation et forme grammaticale.

Dans son article intitulé « Le verbe polonais *tęsknić* : portraits lexicographiques d'une polysémie », Dorota Sikora propose une réévaluation de la description lexicographique de ce verbe central du lexique affectif polonais. Le verbe *tęsknić* 'éprouver de la nostalgie de', 'se languir de', 'ressentir le manque de' couvre un spectre sémantique très large, allant du regret affectif à l'aspiration vers un objet désiré ou attendu. Contrairement aux dictionnaires contemporains, qui n'en recensent que deux acceptions principales (l'une affective, l'autre volitive), l'analyse menée ici sur corpus conduit à distinguer quatre acceptions distinctes : deux relèvent du champ des sentiments (par ex. *tęsknić za domem* 'avoir la nostalgie de sa maison'), une troisième traduit une attitude épistémique (comme lorsqu'on regrette de ne pas savoir : *tęsknić za odpowiedzią* 'regretter de ne pas avoir la réponse'), et une quatrième exprime une visée prospective ou volitive (par ex. *tęsknić do wakacji* 'aspirer aux vacances'). L'étude s'inscrit dans la tradition de la lexicographie systémique telle que conçue par Apresjan : elle articule la recherche des traits communs aux verbes de sentiment à l'observation minutieuse des propriétés propres à *tęsknić*, en prenant en compte ses liens syntagmatiques et paradigmatiques. L'enquête montre notamment que la forme du complément (avec *za* + instrumental ou *do* + génitif) est révélatrice de la

configuration sémantique activée, entre nostalgie affective, désir de retour ou attente tournée vers le futur. Au-delà du cas de *tęsknić*, ce travail ouvre des perspectives méthodologiques pour la description contrastive du lexique des sentiments et pour l'analyse fine de la polysémie.

Irina Kor-Chahine consacre son article « When the Moment Does not Last: A Few Words on Момент-Constructions in Russian » à l'évolution sémantique et syntaxique des constructions russes avec le substantif *moment* 'moment'. Si ce terme désignait à l'origine un instant bref ou un point précis dans le temps (au sens du latin *momentum*), son usage contemporain s'est progressivement élargi, au point de devenir l'un des noms les plus fréquents du russe actuel. L'étude repose sur une analyse diachronique, menée sur les données du *Corpus national de la langue russe*, du XVIII^e siècle à nos jours. L'auteure distingue quatre constructions principales : (1) (PRÉP) *moment* ГЭН. (*s momenta roždenija* 'depuis le moment de la naissance') ; (2) V / Na DÉT *moment* (*v dannyj moment* 'à l'instant présent') ; (3) VERBE *moment* (*ne upustit' moment* 'ne pas laisser passer l'occasion') ; et (4) ADJ *moment* (*opasnyj moment* 'moment critique'), qui peut appartenir aux constructions de type (2) ou (3), mais où *moment* développe un sens particulier. Ces constructions révèlent une évolution marquée : le sens originel 'bref instant' tend à reculer, au profit d'usages plus abstraits ou argumentatifs, où *moment* fonctionne comme synonyme partiel de *aspekt*, *ètap*, *vopros*, voire *problema*. On observe également une augmentation nette des emplois non littéraires, avec une forte productivité dans les textes médiatiques et institutionnels du XXI^e siècle. Ce travail offre ainsi un « portrait lexicographique » de *moment*, illustrant les processus de figement, de lexicalisation et de glissement sémantique dans le russe contemporain. Il éclaire plus largement les mécanismes par lesquels certains noms abstraits comblent un vide lexical dans le vocabulaire discursif.

Dans son article « Partial Observation and Nonstandard Tense Uses: Russian Perfective Verbs in Progressive Interpretation », Alexander Letuchy s'intéresse à un usage particulier du passé perfectif russe, observé notamment avec certains verbes de mouvement préfixés — comme *poexat'* 'commencer à rouler' ou *poletet'* 'commencer à voler'. Dans des contextes de discours direct, ces formes fonctionnent comme des marqueurs de commentaire immédiat : le locuteur utilise un passé perfectif pour désigner un événement en cours d'observation, sans en avoir perçu le début ni anticiper la fin (*Direktor pošël* 'Le directeur est en train d'y aller'). Il ne s'agit pas pour autant d'un véritable « passé progressif » : l'étude montre que cet usage ne partage pas les propriétés combinatoires typiques du progressif (incompatibilité avec les modificateurs du présent ou les subordonnées de simultanéité). Le phénomène s'explique par le rôle particulier des verbes de mouvement, dont la sémantique directionnelle permet cette lecture « en direct »,

alors que d'autres verbes inceptifs (comme *zapet'* 'se mettre à chanter') ne l'autorisent pas. A. Letuchy décrit ce mécanisme comme un processus métonymique : le début de l'observation est interprété comme le début de la situation elle-même. Enfin, l'étude élargit cette observation à d'autres emplois non standards des temps verbaux en russe, notamment les emplois résultatifs et évidentiels, et met en évidence le rôle crucial de la notion d'« observation partielle » dans la structuration aspectuo-temporelle du discours.

L'article « Remarks on Ju. D. Apresjan's Lexicographic Terminology: *leksema* and *vokabula* in a Multilingual Perspective » de Tilmann Reuther revisite deux notions clés de la terminologie lexicographique élaborée par Jurij Apresjan et Igor Mel'čuk : *leksema* et *vokabula*. Alors que le terme *leksema* s'est largement diffusé dans différentes traditions linguistiques – sous les formes *lexeme*, *lexème*, *Lexem*, *lexema* –, le terme *vokabula*, pourtant central dans la théorie d'Apresjan, est demeuré marginal, y compris dans la lexicographie russe contemporaine. T. Reuther analyse cette dissymétrie à la lumière de l'histoire et des spécificités des traditions terminologiques nationales, en examinant le destin contrasté des mots apparentés (*vocable*, *vocable*, *Vokabel*, *vocablo*), dont les usages antérieurs ont freiné l'adoption du sens technique proposé par Apresjan et Mel'čuk. L'étude ne se limite pas au constat : elle envisage les conditions d'une éventuelle réhabilitation de *vokabula* dans la tradition russe et propose, dans une perspective multilingue, des solutions terminologiques pour l'allemand et l'anglais (par ex. *lemma* ou *keyword*). L'article offre ainsi une contribution importante à la réflexion théorique sur la lexicographie et souligne l'actualité et la fécondité du legs scientifique d'Apresjan et de Mel'čuk.

Dans « Poskakali ! 'Filons !' / 'Assez couru !' : construire l'instabilité (les verbes *nesti*, *nosit'*, *skakat'* et le préfixe *po-*) », Rémi Camus analyse l'ambivalence sémantique du préfixe *po-* en russe, traditionnellement décrit comme porteur de deux valeurs principales : l'inchoation ('se mettre en mouvement') et la délimitation ('agir pendant un temps limité'). À partir d'une étude détaillée de la paire *nesti* / *nosit'* et du verbe *skakat'*, l'auteur montre que ces deux valeurs ne s'opposent pas simplement comme deux lectures aspectuelles concurrentes, mais s'articulent autour d'un invariant plus profond : le sens de l'instabilité, qu'il s'agisse du mouvement d'un objet, d'un déplacement animal ou d'usages figurés. L'article met aussi en lumière le rôle décisif de la morphologie : les verbes de déplacement déterminés (comme *nesti*) et les verbes indéterminés (comme *nosit'*), dont la distinction relève d'un processus morphologique bien défini, n'expriment pas l'instabilité de la même manière, ce qui explique la répartition des valeurs de *po-*. Enfin, l'auteur souligne la diversité des contextes syntaxiques et interprétatifs où ces interprétations se manifestent, invitant à dépasser l'explication en termes de

simple « déplacement » pour envisager une conception plus générale et unifiée du fonctionnement du préfixe *po-* dans le système du verbe russe.

Dans son article « Définir pour apprendre et apprendre à définir : repenser la lexicographie pour l'enseignement des langues », Svetlana Krylosova explore le potentiel didactique du *Dictionnaire actif de la langue russe*, conçu par Ju. D. Apresjan et son équipe. Plutôt que de voir dans cet ouvrage un simple outil de référence, elle montre qu'il peut servir de véritable matrice méthodologique pour l'enseignement des langues étrangères. Après avoir rappelé les principes qui structurent la définition lexicographique dans la Lexicologie explicative et combinatoire de Mel'čuk et dans la Lexicographie active d'Apresjan, l'auteure propose d'en dériver des « définitions pédagogiques », adaptées aux besoins des apprenants. Elle illustre cette démarche à travers des activités centrées sur la distinction de quasi-synonymes, la combinatoire lexicale et les interactions entre lexique et grammaire, montrant ainsi comment l'analyse lexicographique peut se transformer en outil de production et de réflexion linguistique en classe. L'article met en évidence que « définir pour apprendre » constitue une voie privilégiée pour « apprendre la langue ».

Nous sommes très reconnaissants à toutes les personnes qui ont contribué à l'évaluation des articles, pour leurs lectures rigoureuses, leurs remarques précieuses et l'attention fidèle qu'elles ont portée à l'œuvre et à la mémoire de Jurij Apresjan : Gueorgui Armianov (France), Olivier Azam (France), Vincent Bénét (France), Natalia Bernitskaïa (France), Xavier Blanco (Espagne), Bohdan Krzysztof Bogacki (Pologne), Christine Bonnot (France), Anna Bonola (Italie), Tatiana Bottineau (France), Andrzej Charciarek (Pologne), Eric Corre (France), Michael Daniel (France), Nina Dobrushina (France), Valentina Fedchenko (France), Paolo Frassi (Italie), Olga Inkova (Suisse), Anetta Kopecka (France), Mihail Kopotev (Finlande), Katarzyna Kwapisz-Osadnik (Pologne), Sébastien Marengo (Canada), Jasmina Milićević (Canada), Irina Nesterenko (France), Elena Nikishina (Russie), Alain Polguère (France), Dmitri Sichina-va (Allemagne), Dejan Stosic (France), Monika Sułkowska (Pologne), Valentina Zhukova (Norvège). Nos remerciements vont à l'ensemble des auteurs de ce numéro — celles et ceux qui ont partagé un travail achevé, mais aussi celles et ceux qui, sans pouvoir finaliser leur texte, ont tenu à s'associer à cet hommage. Leur engagement témoigne, lui aussi, de la vitalité de l'héritage intellectuel de Jurij Apresjan. Enfin, nous adressons nos remerciements à Marie Starynkevitch, qui a assuré la traduction en français de deux textes de ce volume, ainsi qu'au Centre de recherches Europes-Eurasie (CREE) de l'Inalco — Paris pour son précieux soutien.

Puissent ces pages faire entendre, au-delà des analyses qu'elles proposent, un écho fidèle à l'esprit de précision, de générosité et de profondeur qui a animé l'œuvre et la personne de Jurij Derenikovič Apresjan.

Svetlana Krylosova

Références citées

- Apresjan, Ju. (1971). O reguljarnoj mnogoznačnosti. *Izvestija AN SSSR. Otdelenie literatury i jazyka*, t. XXX(6), 509–523.
- Apresjan, Ju. (1978). Jazykovaja anomalija u logičeskoe protivorečie. Dans M. R. Mayenowa (éd.), *Tekst. Język. Poetyka, Zbiór studiów* (127–151). Ossolineum.
- Apresjan, Ju. (éd.) (1979). *Anglo-russkij sinonimičeskij slovar'*. Russkij jazyk.
- Apresjan, Ju. (1985). Ličnaja sfera govorjaščego i naivnaja model' mira. *Semiotičeskie aspekty formalizacii intelektual'noj dejatel'nosti. Škola-seminar «Kutaisi-1985». Tezisy dokladov i soobščenij* (263–268). Moskva.
- Apresjan, Ju. (1992a). *Lexical Semantics. A Guide to Russian Vocabulary*. Trad. A. Lehman, N. and L. Longan & A. Eulenberg. Karoma Publishers.
- Apresjan, Ju. (1992b). Leksikografičeskie portrety (na primere glagola BYT'). *Naučno-texničeskaja informacija, serija* 2(3), 20–33.
- Apresjan, Ju. (1992c). Lexicographic Portraits and Lexicographic Types. Dans M. Guiraud-Weber, Ch. Zaremba (éds), *Linguistique et slavistique. Mélanges offerts à Paul Garde* (360–376). Université de Provence.
- Apresjan, Ju. (1992d). Konnotacii kak čast' pragmatiki slova (leksikografičeskij aspekt). *Russkij jazyk. Problema grammatičeskoj semantiki i ocenočnye faktory v jazyke* (45–64). Nauka.
- Apresjan, Ju. (1994). O jazyke tolkovanij i semantičeskix primitivax. *Voprosy jazykoznanija* 4, 3–16.
- Apresjan, Ju. (1995). *Integral'noe opisanie jazyka i sistemnaja leksikografija. Izbrannye trudy, t. II. Jazyki russkoj kul'tury*.
- Apresjan, Ju. (éd.) (2010). *Prospekt aktivnogo slovarja russkogo jazyka*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (éd.) (2014a). *Aktivnyj slovar' russkogo jazyka. T. 1. Jazyki slavjanskoj kul'tury*.
- Apresjan, Ju. (éd.) (2014b). *Aktivnyj slovar' russkogo jazyka. T. 2. Jazyki slavjanskoj kul'tury*.

- Apresjan, Ju. (éd.) (2017). *Aktivnyj slovar' russkogo jazyka*. T. 3. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (éd.) (2023a). *Aktivnyj slovar' russkogo jazyka*. T. 4(1). Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (éd.) (2023b). *Novyj ob"jasnitel'nyj slovar' sinonimov russkogo jazyka, 2e ed.* Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (éd.) (2024). *Aktivnyj slovar' russkogo jazyka*. T. 4(2). Jazyki slavjanskoj kul'tury.
- Apresjan, Ju., Mednikova È. & Petrova A. (éds) (1993–1994). *Novyj bol'soj anglo-russkij slovar'*. Russkij jazyk.
- Babaeva E., Galaktionova I., Levontina I., Krylova T. (éd.) (2025). *Aktivnyj slovar' russkogo jazyka*. T. 5(1). MCNMO.
- Buras, M. (2022). *Lingvisty, prišedšie s xoloda*. AST.
- Green, G. (1969). On the Notion 'Related Lexical Entry'. *Papers from the Fifth Regional Meeting* (76–88). Chicago Linguistic Society.
- Mel'čuk, I. (2024a). An Epoch Has Ended: to Apresjan's Dear Memory. *Russian Journal of Linguistics* 28(2), 480–483.
- Mel'čuk (2024b). *Pro Apresjana* [texte inédit].
- Mel'čuk, I., Clas A. & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Duculot.
- Mel'čuk, I. & Žolkovskij, A. (1984). *Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka: opyty semantiko-sintaksičeskogo opisanija russkoj leksiki*. Wiener Slawistischer Almanach.



Igor Boguslavskij
Leonid Iomdin †
Leonid Krysin

Jurij Apresjan
(1930–2024)*

Le 12 mai 2024 s'est éteint, dans sa 95^e année, le savant et académicien russe Jurij Derenikovitch Apresjan. Avec sa disparition, la science a perdu non seulement un linguiste de génie, investigateur insatiable de la langue russe, mais aussi un homme d'une immense culture au charisme incontesté.

L'influence de ses travaux et de sa personnalité hors du commun sur la linguistique de la deuxième moitié du XX^e et du début du XXI^e siècle ainsi que sur l'atmosphère générale du monde de la recherche en Russie n'est plus à démontrer. Plusieurs de ses ouvrages sont devenus des classiques dès leur parution ; ainsi de *Idées et méthodes de la linguistique structurale contemporaine* [1] (1966)¹, *Étude expérimentale de la sémantique du verbe russe* [2] (1967), *Sémantique lexicale* [3] (1974), *Description intégrale de la langue et lexicographie systématique* [4] (1995), *Études en sémantique et lexicographie* [5] (2009). À l'étranger, les travaux d'Apresjan sont connus tant par les nombreuses traductions de ses textes que par l'ouvrage *Systematic Lexicography* [6], écrit pour répondre à une commande d'Oxford University Press, publié en 2000 et réédité en 2008.

Apresjan est l'un des rares, ou peut-être le seul théoricien de la langue dont le nom soit connu de centaines de milliers, voire de millions de personnes fort éloignées des sciences du langage ; exemple éloquent, le tirage de son *Grand*

* Le texte a été traduit du russe à l'occasion de la préparation du numéro 37 de *Neophilologica* par Marie Starynkevitch, avec le soutien du Centre de recherche Europes-Eurasie (Inalco, Paris). La version originale a paru dans *Izvestija RAN. Studies in Literature and Language*, 2024, 83(3), 136–140.

¹ NdT : Par souci de lisibilité, nous indiquons dans le corps du texte la traduction des ouvrages d'Apresjan mentionnés. Le lecteur trouvera à la fin de l'article la liste complète des titres originaux, par ordre d'apparition.

dictionnaire anglais-russe [7] en trois tomes, à la rédaction duquel il prit une part très active au début des années 1970, et qui fut réédité sous sa direction à partir de la troisième édition à la fin des années 1980. Ces dernières années, ce dictionnaire a été significativement renouvelé ; il est en train de passer à un nouveau format interactif. Le *Dictionnaire anglais-russe des synonymes* [8] paru sous sa direction en 1979 connut lui aussi plusieurs rééditions.

Apresjan est né à Moscou. Il achève en 1953 des études d'anglais à l'Institut pédagogique national des langues étrangères de Moscou (par la suite Institut Maurice Thorez, aujourd'hui Université linguistique d'État de Moscou), où il enseigne l'anglais pendant plusieurs années tout en effectuant des recherches en vue d'un doctorat d'État ayant pour sujet *Synonymes phraséologiques en anglais contemporain* [9], qu'il soutient en 1958.

En 1960, Apresjan rejoint la section de linguistique structurale de l'Institut de la langue russe de l'Académie des sciences d'URSS, où il travaille jusqu'à l'été 1972 en tant que « jeune collaborateur scientifique ». En juin 1972, le conseil scientifique de l'Institut de la langue russe « ne considère pas comme possible » de prolonger son contrat. Apresjan est pourtant déjà un spécialiste reconnu en linguistique structurale, auteur de deux ouvrages (dont l'un est traduit dans plusieurs langues) et d'un grand nombre d'articles. Une telle décision de la part de la direction de l'établissement est exclusivement dictée par des considérations non scientifiques : Apresjan était persécuté pour sa pensée dissidente (*inakomyslie*). Il ne dissimulait pas son opposition à la politique du pouvoir soviétique, mue par des dogmes idéologiques. Il avait plusieurs fois pris la défense de personnes condamnées injustement, telles Iouli Daniel et André Siniavski, Aleksandr Ginzburg, Jurij Galanskov ; il avait pris fait et cause pour son ami Konstantin Babitski, condamné à plusieurs années d'exil pour avoir participé à une manifestation de protestation sur la Place rouge contre l'entrée des troupes soviétiques en Tchécoslovaquie le 25 août 1968.

De 1972 à 1985, Apresjan travaille à l'Institut de recherche appliquée Informelectro, où il rassemble un groupe de linguistes et d'informaticiens qui s'emploie sous sa direction à concevoir un système de traduction automatique de textes scientifiques et techniques, d'abord du français vers le russe, puis de l'anglais vers le russe. En 1985, il rejoint, avec quelques-uns de ses collaborateurs, l'Institut des problèmes liés à la transmission de l'information (IPPI) de l'Académie des sciences de l'URSS (rebaptisé depuis Institut Xarkevič des problèmes liés à la transmission de l'information). Il y occupe d'abord le poste de collaborateur scientifique en chef avant de diriger le laboratoire de linguistique informatique de 1989 à 1994. En 1995, il confie la direction du laboratoire à l'un de ses disciples et compagnons d'armes pour reprendre le poste de directeur de recherches

du même laboratoire. Sans interrompre son travail à l'IPPI, Apresjan retourne à l'Institut Vinogradov de la langue russe, où il prend la tête d'un collectif de scientifiques qui fondera ensuite le domaine de la sémantique théorique.

La science linguistique soviétique officielle ne se pressa pas de reconnaître les mérites incontestables d'Apresjan. Pis, elle lui mit des bâtons dans les roues : le chercheur fut à maintes reprises pris à parti dans différents réunions et congrès et convoqué au KGB ; il fut interdit de voyage à l'étranger lorsque les établissements de recherche l'invitèrent. On le priva du droit d'enseigner à l'université, tandis que ses articles étaient refusés à la publication et que les références à ses travaux dans les publications d'autres chercheurs soviétiques étaient effacées. En 1973, le manuscrit de l'ouvrage *Sémantique lexicale* [10] qu'il avait remis aux éditions Nauka fut délibérément « perdu ». Par la suite, lorsque son auteur parvint miraculeusement à le reconstituer et à le publier, cet ouvrage devait poser les jalons d'une nouvelle approche de la sémantique théorique et de la lexicographie. Ni ces disciplines, ni l'enseignement de la linguistique ne seraient ce qu'ils sont aujourd'hui sans cet ouvrage. Ce n'est pourtant qu'en 1984 et à Minsk, loin de la science officielle professée à Moscou, qu'Apresjan présenta sa monographie *Sémantique lexicale* au titre d'un doctorat d'État.

L'atmosphère dans laquelle baignait la linguistique soviétique de la fin des années 1960 et des années 1970 a été bien décrite par Apresjan lui-même dans sa préface à l'ouvrage en deux volumes *Textes choisis* [11] (1995). Il y évoque non seulement ses propres tribulations, mais aussi les persécutions subies par d'autres chercheurs. En dépit de ces conditions hostiles, Apresjan travailla pendant toutes ces années sur toutes sortes de sujets touchant aussi bien à la linguistique théorique ou appliquée qu'aux études russes ; il est l'auteur d'études qui exercèrent une influence cruciale sur le développement des sciences du langage aussi bien en Russie qu'à l'étranger.

L'idée d'une description intégrée (c'est-à-dire unique pour le lexique et la grammaire) du langage naturel, qu'il met en application dans de nombreux articles et dictionnaires, dont le *Nouveau dictionnaire explicatif des synonymes de la langue russe* [12] (2004) et le *Dictionnaire actif de la langue russe* [13], dont l'élaboration a commencé en 2010 à l'initiative d'Apresjan et sous sa direction, est désormais solidement ancrée dans la conscience des linguistes. Elle paraît aujourd'hui aller de soi, à tel point qu'on peine à s'imaginer à quel point elle était novatrice à son époque.

Le séminaire permanent fondé par Apresjan, d'abord en linguistique appliquée, puis en linguistique théorique, poursuivit ses activités pendant 46 ans. Y présenter ses travaux était considéré comme un honneur par de nombreux linguistes russes et étrangers. Le séminaire représenta la « ruche » dans laquelle

se forma progressivement l'École sémantique de Moscou dont Apresjan était la figure tutélaire, et qui était déjà bien établie au début du XXI^e siècle. Apresjan prisait particulièrement le séminaire comme lieu d'échanges informels. En dépit de l'autorité dont il était investi dans le milieu de la linguistique russe, Apresjan avait une préférence pour le dialogue et la discussion d'égal à égal avec ses collègues et disciples. Il savait écouter son interlocuteur sans lui imposer son propre point de vue.

Au début des années 1990, les mérites de Ju. Apresjan furent enfin reconnus par la science académique : en juin 1992, il fut élu membre actif de l'Académie des sciences russe (sans passer par le statut de membre-correspondant). Il va sans dire que cette nomination ne changea pas son attitude et n'entama pas le moins du monde ses qualités d'homme et de chercheur : il poursuivit son travail sans relâche pour la linguistique, tout en se montrant toujours aussi accessible et simple dans l'échange, toujours aussi attentif à ses collègues qu'il l'avait été par le passé.

En 2004, l'Académie des sciences russe lui remit la Médaille d'or Dahl. Apresjan devint membre étranger de l'Académie nationale des sciences d'Arménie, professeur éminent à l'Université d'État de Moscou (MGU), lauréat du prestigieux prix Humboldt (*Humboldt-Forschungspreis*), docteur honoris causa de l'université Saint-Clément-d'Ohrîd de Sofia, docteur honoris causa de l'Université de Varsovie.

Jurij Apresjan est connu bien au-delà des frontières russes pour être le père de l'École sémantique de Moscou. C'est dans le cadre de cette école qu'il créa une vaste théorie sémantique très élaborée fondée sur un corpus linguistique immense et varié. En cela, elle se distingue déjà de la plupart des cadres théoriques actuels, qui reposent bien souvent sur un corpus très limité. Apresjan s'était toujours senti lexicographe : c'est bien pourquoi il éprouvait la pertinence de ses théories en lisant de volumineux dictionnaires (russes et anglais) de la première à la dernière page (par exemple, il avait lu les quatre tomes du *Petit dictionnaire académique de la langue russe* [14]). Les travaux d'Apresjan influencèrent profondément de nombreuses études qui ne s'inscrivaient pas formellement dans le cadre de l'École sémantique de Moscou, et firent sensiblement évoluer le raisonnement scientifique de bien des chercheurs.

Apresjan se distinguait par sa faculté exceptionnelle à donner corps à des concepts théoriques généraux à travers des ouvrages lexicographiques concrets. Cette faculté se manifesta dans tous les dictionnaires sur lesquels il travailla : le *Dictionnaire explicatif et combinatoire de la langue russe* [15] (élaboré par Jurij Apresjan en collaboration avec Igor Mel'čuk et Aleksandr Žolkovskij), le dictionnaire *Le verbe russe et le verbe hongrois. Régime et cooccurrences*

(avec È. Pall) [16], le *Grand dictionnaire anglo-russe* [7], le *Dictionnaire anglo-russe des synonymes* [8], et naturellement le *Nouveau dictionnaire explicatif des synonymes de la langue russe* [12], qui intégra presque toutes les nouveautés de la pensée théorique de l'École sémantique de Moscou. Les trois derniers dictionnaires étaient des ouvrages collectifs ; Apresjan joua toutefois un rôle déterminant dans leur élaboration, aussi bien en tant qu'auteur qu'à titre de rédacteur minutieux, vigilant et avisé.

Les dictionnaires créés par Apresjan et sous sa direction scientifique se veulent des dictionnaires actifs : la notion de dictionnaire actif implique qu'aux unités lexicales soient associées les informations nécessaires non seulement à la compréhension des textes, mais aussi à l'utilisation fluide de la langue par ses locuteurs (ses locuteurs natifs aussi bien que ceux qui l'étudient comme langue étrangère). Ainsi une entrée de dictionnaire doit-elle contenir des renseignements exhaustifs sur toutes les propriétés essentielles de l'unité lexicale considérée. Dans la pensée d'Apresjan, l'ensemble de ces renseignements, avec les règles suivant lesquelles se construisent les énoncés, constitue l'équivalent formel de ce que l'on appelle la compétence langagière des locuteurs. Les idées concernant cette notion de dictionnaire actif prennent corps dans le travail qui se fait actuellement sur le dictionnaire en plusieurs tomes intitulé *Dictionnaire actif de la langue russe* (dont sont parus quatre tomes à ce jour). En dépit d'un état de santé difficile dans la dernière période de sa vie, Apresjan prit une part active à ce travail.

Une autre source de stimulation pour le développement de la théorie linguistique fut le travail mené par Apresjan sur les systèmes de traitement automatique du langage naturel, en particulier sur le système de traduction automatique intitulé ÈTAP (*ÈlektroTexničeskij Avtomatičeskij Perevod*, « Traduction automatique électro-technique »), auquel il se dévoua corps et âme pendant des années. Apresjan, ainsi que les autres créateurs de ce système, le percevaient dans une large mesure comme une sorte de terrain d'expérimentation permettant d'éprouver leurs théories linguistiques. C'est, au fond, précisément ce travail qui amena Apresjan à formuler un principe majeur de la description intégrée de la langue, à savoir le principe de convergence de la grammaire et du lexique. Le système ÈTAP représenta quant à lui l'application pratique de ce principe.

Jurij Apresjan faisait preuve d'une exigence extrême concernant la qualité des textes qu'il faisait paraître. Cette exigence s'appliquait également aux exemples illustrant des théories ou des entrées de dictionnaires. S'il s'agissait d'un exemple littéraire, il était emprunté à un chef-d'œuvre. Si c'était un exemple élaboré par le chercheur, il était édifiant au sens premier du terme, évoquant un fait historique ou une vérité scientifique (pas nécessairement linguistique !) ou bien brossant un tableau optimiste du monde, dans lequel agissaient des êtres

bons, quoique non exempts de faiblesses et d'émotions humaines communément partagées.

Par la suite, Apresjan aborda un autre sujet, tout aussi important. Il élaborà au terme de recherches approfondies une classification fondamentale des prédicats. S'écàrtant des classifications existantes conçues par Jurij Maslov, Zeno Vendler et d'autres chercheurs, il proposa une ontologie élaborée et cohérente des prédicats, dont beaucoup furent ainsi présentés sous cet angle pour la première fois. Il y distinguait, entre autres, les prédicats d'activité (commercer, faire la guerre), d'attitude (faire des bêtises, faire des caprices, gouailler), d'occupation (jouer, se reposer), d'existence (exister, être, être présent) et d'interprétation (se tromper, séduire). Comme le soulignait Apresjan, une telle classification est indispensable à la formulation des règles linguistiques les plus diverses, syntaxiques, sémantiques, communicatives ou relatives aux cooccurrences.

Apresjan mena un travail sur un thème connexe à la classification fondamentale des prédicats, en mettant au point un système cohérent des traits sémantiques des mots et des rôles sémantiques des prédicats. Le premier est utilisé à titre expérimental dans le corpus de textes russes annoté avec précision Sin-TagRus, constitué par les collègues les plus proches d'Apresjan à l'IPPI.

Apresjan consacra les dernières années de sa vie à deux projets d'envergure. Il s'engagea très activement dans l'élaboration d'un *Dictionnaire actif de la langue russe* [13] : son rôle ne se limita pas à celui de concepteur du dictionnaire, de rédacteur d'un grand nombre de ses entrées, de directeur d'ouvrage des tomes déjà parus et de directeur du projet dans son ensemble. Il continua à mener inlassablement un travail d'étayage théorique du dictionnaire et à diffuser les résultats obtenus dans de nombreuses publications.

Si le travail sur le *Dictionnaire actif de la langue russe* [13] est mené par un collectif bien établi, le deuxième projet relève quant à lui de la création personnelle de Jurij Apresjan. Il consacra plus de dix ans à un sujet qui l'occupait depuis l'époque de la publication de son *Dictionnaire des synonymes anglais-russe* [8] (1979), à savoir la lexicographie comparée de l'anglais et du russe. La lexicographie russe a rassemblé un grand nombre d'observations des plus diverses concernant les divergences entre les deux langues. Ce n'est toutefois que dans le cadre du travail mené par Apresjan que les observations de cette nature furent le résultat d'une étude systématique du lexique des deux langues – démarche que l'on peut à bon droit considérer comme la marque de fabrique d'Apresjan. Pour obtenir un tel résultat, il réalisa un immense travail préparatoire. Il constitua un dictionnaire raisonné et combinatoire bilingue anglais-russe, qui comportait plus de 8 000 entrées. Chacune réunissait des informations factuelles sur les différentes propriétés du terme, y compris, à un degré de précision

sans commune mesure avec les travaux précédents, sur ses fonctions lexicales (telles que comprises selon la Théorie Sens-Texte). À partir de ce dictionnaire fut menée une vaste étude des divergences entre l'anglais et le russe. Les résultats majeurs de cette expérience concernent aussi bien les similitudes que les divergences relevant des cooccurrences lexicales et sémantiques, de la sémantique, mais aussi de l'ordre des mots et de l'organisation communicative des énoncés.

Une conscience professionnelle rare reposant sur un sens moral élevé, une exigence stricte envers ses propres recherches aussi bien que les travaux des autres chercheurs : ces qualités réunies chez Apresjan s'associaient au respect avec lequel il considérait les autres approches scientifiques, y compris celles qui divergeaient fortement de ses propres théories.

La disparition d'un homme et savant tel que Jurij Apresjan est une perte douloureuse. Tous ceux qui l'ont connu ou qui, sans le connaître, ont étudié ses travaux, en sont sincèrement affligés.

Références citées (par ordre d'apparition)

Sauf mention contraire,

les références citées sont toutes des ouvrages d'Apresjan

1. *Idei i metody sovremennoj strukturnoj lingvistiki* [Idées et méthodes de la linguistique structurale contemporaine] (1966).
2. *Èksperimental'noe issledovanie semantiki russkogo glagola* [Étude expérimentale de la sémantique du verbe russe] (1967).
3. *Leksičeskaja semantika* [Sémantique lexicale] (1974).
4. *Integral'noe opisanie jazyka i sistemnaja leksikografija* [Description intégrale de la langue et lexicographie systématique] (1995).
5. *Issledovanija po semantike i leksikografii* [Études en sémantique et lexicographie] (2009).
6. *Systematic Lexicography* (2000).
7. *Bol'soj anglo-russkij slovar'* [Grand dictionnaire anglais-russe].
8. *Anglo-russkij sinonimičeskij slovar''* [Dictionnaire anglais-russe des synonymes] (1979).
9. *Frazeologičeskie sinonimy v sovremennom anglijskom jazyke* [Synonymes phraséologiques en anglais contemporain] (thèse de doctorat, 1958)
10. *Leksičeskaja semantika* [Sémantique lexicale] (1984).
11. *Izbrannye trudy* [Textes choisis] (1995).

12. *Novyj ob''jasnitel'nyj slovar' sinonimov russkogo jazyka* [Nouveau dictionnaire explicatif des synonymes de la langue russe] (2004).
13. *Aktivnyj slovar' russkogo jazyka* [Dictionnaire actif de la langue russe] (publication toujours en cours ; premier tome paru en 2014).
14. Evgen'eva A. (dir.) *Slovar' russkogo jazyka v četyrëx tomax (Malyj akademičeskij slovar')* [Dictionnaire de la langue russe en quatre tomes – Petit dictionnaire académique de la langue russe].
15. Mel'čuk, I., Žolkovskij, A. et al. [Apresjan fit partie des contributeurs] (1984). *Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka. Opyty semantiko-sintaksičeskogo opisanija russkoj leksiki*. [Dictionnaire explicatif et combinatoire de la langue russe. Essais de description sémantico-syntaxique du lexique russe].
16. Apresjan, Ju. & Pall, È. (1982). *Russkij glagol – vengerskij glagol. Upravlenie i sočetaemost'* [Le verbe russe et le verbe hongrois. Régime et cooccurrences].



Igor Mel'čuk

Observatoire de linguistique Sens-Texte
Université de Montréal
Canada

 <https://orcid.org/0000-0002-4520-0554>

So-called independent infinitives in Russian: The construction of the form *Ètogo mne ne ponjat'* 'It is impossible for me to understand this'

Abstract

The paper considers the Russian construction of the form “N_{DAT} BYT' NE V_{INF}”—one of several so-called “independent infinitive constructions.” This construction represents two syntactically discontinuous idioms: [X-*u*] 'BYT'I.1 ... NE'¹ [Y-*nut'*] 'It is/was/will be impossible that X has done Y' and [X-*u*] 'BYT'I.1 ... NE'² [Y-*at'*] 'It will be impossible that X does Y'. The paper describes two illustrative sentences containing these idioms, giving and discussing their formal representations on the semantic, deep-syntactic and surface-syntactic levels; it also presents the lexical entries for both idioms with detailed linguistic comments. The problems of the “independent infinitives” in Russian and of phraseologically bound lexemes are discussed.

Keywords

Russian, syntax, phraseology, independent infinitive, discontinuous idioms, lexical entries of idioms, phraseologically bound lexemes

To the dear memory of my friend Jura

On May 12, 2024, one of the greatest linguists of our time—the Russian linguist Jurij Apresjan—completed his earthly journey. It is not appropriate to try to present here his scientific legacy and its importance; suffice it to say that the main domain of Apresjan's endeavor was lexicology and lexicography. May these pages, which present the lexical entries of two Russian idioms supplied with

linguistic comments, be a modest contribution to the body of a huge lexicographic project he launched: *Active Dictionary of Russian*, of which the first six issues have been published so far (A—L): Apresjan, Ju., ed. 2014–2025.

1 The paper's tasks

One of the hot topics in Russian linguistics is the so-called independent infinitive. An independent infinitive is considered to be the syntactic head of its clause, i.e., the top node of the corresponding syntactic structure. Thus, in the sentence *Ėtogo mne ne ponjat'* lit. 'This to.me not to.have.understood' = 'It is impossible for me to understand this' the infinitive *ponjat'* is taken to be syntactically the top node. This construction is extensively discussed in a lot of works; fortunately, there is a relatively recent study dedicated to the phenomenon of Russian independent infinitives—Padučeva (2017), which presents a compact picture of the domain: numerous examples from *Russian National Corpus* (<https://ruscorpora.ru>), an exhaustive classification of the types of independent-infinitive constructions, a careful examination of their semantic, syntactic and morphological properties, as well as a good overview of the relevant literature. I will be using Padučeva's data, although I do not share her syntactic treatment of the construction.

Only one type of the independent-infinitive constructions is considered here: the infinitive with *NE* 'not' that expresses impossibility (Padučeva 2017: 2.1.4); other types—9 of these!¹—are left out. The present paper is a first step toward the formal description of the Russian independent-infinitive constructions in the Meaning-Text framework and, as a consequence, it is rather sketchy: it lacks a general embracing frame.

The constructions treated below—referred to as dative-infinitive constructions—can be illustrated with sentences (1) and (2):

- (1) *Ne poguljat' mne Ø_{PRES}^{BYT'1.1} /bylo/budet po Moskve*
 lit. 'Not to.go.for.a.walk to.me is/was/will.be in Moscow' =
 'It is/was/will be impossible for me to go for a walk in Moscow'.

¹ Such as the "unavoidability infinitive" (*Krizisu byt' nepremenno* lit. 'To crisis to.be necessarily' = 'A crisis is unavoidable'), the "obligation infinitive" (*Mne zavtra rano vstavat'* lit. 'To.me tomorrow early to.stand.up' = 'I have to stand up early tomorrow'), etc.

(2) *Ne guljat' mne Ø_{PRES}^{BYT'I.1} po Moskve*

lit. 'Not to.go.for.walks to.me is in Moscow' =

'It will be impossible for me to go for walks in Moscow'.

- 1. Ø_{PRES}^{BYT'I.1} stands for the zero wordform of the verb BYT'I.1 ('be' the copula) in the present tense, as in *Ivan sčastliv* 'Ivan is happy'.
2. A dot between words in a literal English gloss indicates that these words correspond to one word in the Russian example.
3. All abbreviations and symbols are explained in the table at the end of the paper, p. 18.

The three claims of the present paper are as follows:

- The infinitive appearing in (1) and (2) is by no means syntactically independent: it depends on the verb BYT'I.1 'be' as its copular-attributive complement. **NB:** The status of the verb BYT'I.1 in this role will be made more precise below—**Comment 4**, p. 8.
- There are two dative-infinitive constructions expressing "impossibility"; they are semantically and syntactically very close, but distinct: one, illustrated in (1), means 'it **is**, **was** or **will be** impossible that X has done Y' (the fact Y **is completed**); the other one, illustrated in (2), means 'it **will be** impossible that X will be doing Y' (the fact Y **is not completed**).²
- In both sentences (1) and (2) we are dealing with idioms: in (1), 'BYT'I.1 ... NE¹1 expresses the meaning 'impossible'; and in (2), 'BYT'I.1 ... NE¹2 expresses the meaning 'will be impossible'. Each one of these idioms fully determines the construction it implements, so that the description of this syntactic construction is reducible to the lexicographic entries of the corresponding idioms.³

- 1. The top corners ' ' enclose an idiom.
2. The three dots in the name of an idiom indicate its syntactic discontinuity: there is no direct syntactic link between the idiom's lexical components; see **Comment 5** on p. 9.

In conformity with these claims, the tasks of the present paper are:

- | | |
|--|---|
| | 1. to present the semantic, deep-syntactic and surface-syntactic structures of both dative-infinitive constructions in question—that is, in fact, of both idioms; |
| | 2. to present the lexical entries for both idioms. |

² On 'completed facts', see **Comment 3**, p. 8.

³ The idioms 'BYT'I.1 ... NE¹1 and 'BYT'I.1 ... NE¹2 are homonymous with the idiom 'BYT'I.1 ... NE², as in *Mne bylo ne do obeda* lit. 'To.me it was not up.to lunch' = 'I could not think about the lunch, because I was preoccupied with other things'.

2 The idiom $\text{BYT}'\text{I.1} \dots \text{NE}'1$

The sentence (3) contains an instance of the idiom $\text{BYT}'\text{I.1} \dots \text{NE}'1$:

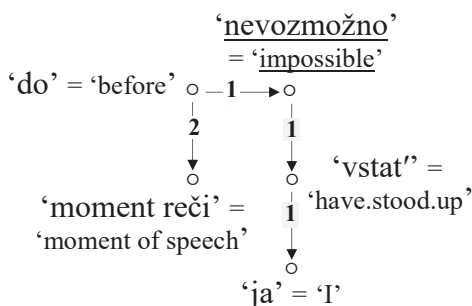
(3) {*Samomu*} *mne bylo ne vstat'* lit. 'To.myself to.me it was not to.have.stood.up'
 ≈ 'It was impossible for me to stand up by myself'.

- The curly brackets in examples enclose the elements that are irrelevant for the topic discussed, but needed for the naturalness of the example; these elements do not appear in the formal representations of the respective sentences.

Here are the semantic [Sem-], deep-syntactic [DSynt-] and surface-syntactic [SSynt-] structures [-S] of sentence (3).

Figure 1

SemS of sentence (3)



Literal reading of this semantic structure:

'before the moment of speech is.impossible that I have.stood.up' =
 'It was impossible for me to stand up'

Figure 2

DSyntS of sentence (3)

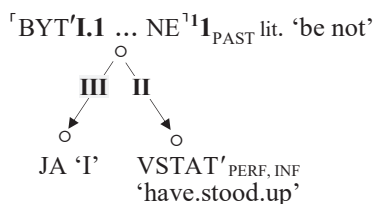
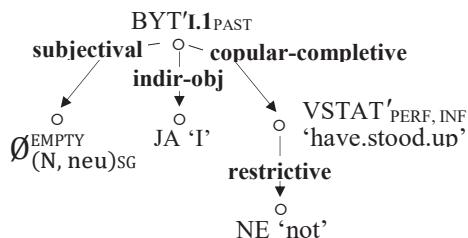


Figure 3

SSyntS of sentence (3)



The idiom 'BYT'I.1 ... NE'1 lacks DSynt-actant I—it is impersonal; see **Comment 7** on p. 10.

$\emptyset_{(N, \text{neu})\text{SG}}^{\text{EMPTY}}$ is a zero lexeme of Russian: a semantically empty, a.k.a. impersonal, nominal pronoun (3rd person, neuter gender, singular), equivalent to such impersonal pronouns as Eng. *IT*, Fr. *IL* and Ger. *ES*; it is a dummy SSynt-subject that ensures the person-number-gender agreement of the Main Verb (in this case, 'BYT'I.1 ‘be [copula]’).

The semanteme ‘nevožmozno’ = ‘impossible’ is expressed in sentence (3) by the idiom 'BYT'I.1 ... NE'1 (its lexicographic number 1 is due to the existence of another idiom of the same vocable— $\text{'BYT'I.1 ... NE'12}$, see Section 3).

The lexical entry for the idiom 'BYT'I.1 ... NE'1 follows.

'BYT'I.1 ... NE'1 , verbal idiom, impersonal; **no** IMPER, INF,⁴ PART(inciple), CONVERB; the imperative 2.SG form *bud'* is marginally possible only for the expression of the irrealis.

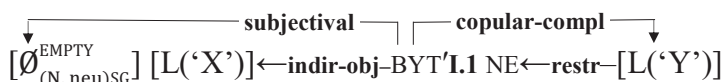
- The underscoring of a semanteme in a Sem-structure indicates its communicatively dominant role; the square brackets enclose the actants of the idiom; L(‘X’) stands for “lexeme L that expresses the meaning ‘X’.”

Signified = Definition

$\text{'[X-u] } \emptyset_{\text{PRES}}^{\text{BYT'I.1}} \text{ ne [Y-nut]'}$ =

Textual version ‘<u>nevožmozno</u> [čtoby X Y-nul]’ = ‘impossible [that X has done Y]’ Formal version ‘<u>nevožmozno</u>’ = ‘impossible’ $\circ -1 \rightarrow \circ -1 \rightarrow \circ$ [‘Y’]	1) ‘Y’ includes the semanteme ‘completed’; 2) ‘Y is desirable for X’

Signifier [= the idiom’s SSynt-structure]



⁴ The infinitive in the citation form of this idiom (as well in that of the $\text{'BYT'I.1 ... NE'12}$ idiom) is, so to speak, artificial; it cannot be used in a natural Russian sentence.

Syntactics

Government pattern

—	'Y' ⇔ II	'X' ⇔ III
1. Ø ^{EMPTY} _{(N, neu)SG}	1. V _{PERF, INF} obligatory	1. N _{DAT}

1) If 'X' does not include 'human', then L('X') [= DSynt-actant III] cannot be elided (see **Comment 6**, p. 9).

Rebĕnku tam Ø^{BYT'1.1}_{PRES}/bylo/budet ne zasnut' lit. 'To.child there it is/was/will.be not to.fall.asleep.' = 'It is/was/will be impossible for the child to fall asleep there.'

Pri takom šume tam Ø^{BYT'1.1}_{PRES}/bylo/budet ne zasnut' lit. 'With such racket there it is/was/will.be not to.fall.asleep.' = 'With such a racket it is/was/will be impossible to fall asleep there.'

Lexical functions

Conv_{23D} : *byt' 'ne po zubam'* lit. 'be not for teeth' ≈ 'be such that one is unable to handle it'

Conv_{23N} : *'ne svetit'* ≈ 'be not to be expected'

Conv_{32N} : (*vsě*) *nikak ne* [V_{Y-PERF, FUT}] 'in no way' [*Ivan nikak ne zasnĕt* 'Ivan can in no way fall asleep.']

Anti_D : *byt' 'po silam'* lit. 'be within strengths' ≈ 'be doable',⁵ *byt' 'pod silu'* lit. 'be under strength' ≈ 'be doable', *byt' 'v* [A_(poss)(N_X)] *silax'*² lit. 'be in X's strengths' ≈ 'be doable'

Anti_{012N} : *poluĉit'sja* ≈ 'succeed'; *udat'sja* ≈ 'manage'

Anti_{021N} : *byt' 'v silax'*¹ lit. 'be in strengths' ≈ 'be capable', *byt' 'v sostojanii'* lit. 'be in state' ≈ 'be in a position'; *moč'* 'can'

Magn : *'ni za ĉto'* lit. 'not for anything', *nikogda* 'never', *nikak* 'in no way', *'nikoim obrazom* 'in no way' | ←V_{INF}

☞ The expression " | ←V_{INF}" means that a Magn adverbial syntactically depends on the verb in the infinitive—i.e. on DSynt-actant II of the idiom.

⁵ See the description of the idioms 'PO SILAM', 'POD SILU', 'V SILAX'1 (*Ja v silax sdelat' èto* lit. 'I am in strengths to do this.') and 'V SILAX'2 (*Sdelat' èto – v moix silax* lit. 'To.do this is in my strengths'.) in Iomdin (2010: 171–173).

Examples

Ivanu ètu zadaču ne rešit' ni za čto 'It is absolutely impossible for Ivan to solve this problem.'

Nam budet nikak ne ujtí nezametno 'It will be absolutely impossible for us to go away unnoticed.'

Ottuda ne ujtí nezametno 'It is impossible to go away from there unnoticed.'

S ètim nam bylo by ne spravít'sja, esli by ne podospela pomošč' 'It would be impossible for us to manage this, if it weren't for the help that arrived.'

Comments

Linguistic comments accompanying this lexical entry are presented in four groups: comments concerning the idiom's signified, the idiom's signifier and the idiom's syntactics, as well as the idiom itself as a whole.

The signified [= the meaning] of the idiom

Comment 1

[X-*u*] 'BYT'I.1 ... NE'¹I means 'impossible [for X to have done Y]' rather than '[X is] unable [to do Y]'; cf. (4a):

(4) a. *Ivanu_x ne kupít' ètot dom: xozjain ne prodast ni za čto*

'It is impossible for Ivan to buy this house: the owner will never sell.'

This sentence can be uttered in case when Ivan is absolutely able to buy the house: he has the money and there are no legal obstacles.

b. *Ivanu_x ne byt' izbrannym* 'It is impossible for Ivan to be elected.'

Sentence (4b) is also by no means about Ivan's abilities—it is rather about his chances.

Comment 2

This lexical entry (and that of 'BYT'I.1 ... NE'¹2) uses a "double-barreled" lexicographic definition: in accordance with the proposal in Mel'čuk & Polguère (2018: 2.3.3), it is presented both in traditional, i.e., textual, form and as a semantic network. The textual form is more accessible to the linguistic intuition of the reader while the semantic network is better for formal reasoning.

Comment 3

The conditions on the signified of the idiom do the following.

- Condition 1 requires that the fact 'Y' be 'completed'. This means that the semanteme configuration 'Y' contains along the semanteme 'Y', which identifies the fact in question, the semanteme 'completed', which bears on 'Y': 'Y' = 'Y'←1-completed'. The semanteme 'Y' itself can be any type of dynamic fact that admits 'completion': an action, an event, a process. Thus, Rus. 'bolet' = 'be sick' cannot be 'completed', but 'zabolet' = 'fall sick' can be both 'completed' and not 'completed'; 'probolet' [period of time T] = 'be sick [during the period of time T]' can be only 'completed'. The semanteme 'completed' denotes the achievement of an inherent limit, after which the fact 'Y' cannot last because of its nature. This semanteme corresponds to one of the senses of the Russian perfective verbal aspect.
- Condition 2 indicates that 'fact Y is desirable for X' (from the viewpoint of the Speaker). Thus, sentence (5) tells us that Ivan wants to be killed in a war:

(5) *Ivanu ne byt' ubitym na vojne* 'It is impossible for Ivan to be killed in a war'.⁶

The semantic component 'fact Y is desirable for X' is also present as a condition on the definition of the idiom 'BYT'I.1 ... NE'I.2.

The signifier [= the SSynt-structure] of the idiom

The idiom's signifier, that is, its SSynt-structure, calls for two comments.

Comment 4

The verb BYT'I.1 'be' is used in both idioms in quite a specific way. BYT'I.1, which in the general case functions as a 100% copula verb, is in these idioms not copular at all. Semantically speaking, it does not link its SSynt-attribute to the SSynt-subject, since its SSynt-subject is semantically empty. Syntactically, this BYT'I.1 governs a dative (= indirect) object, which is impossible for a genuine copula. Morphologically, as indicated above, it lacks the imperative and all non-finite forms.

BYT'I.1 of our idioms retains only the "outer shell" of the genuine BYT'I.1. (On the verb BYT' 'be' in Russian, see the classical work Apresjan (2013), as well as Mel'čuk (2019: 11–16 and 2023b); on the lexicographic treatment of phraseologically bound lexemes—in particular, of the verb BYT'I.1, see **Conclusions**, Item 2, p. 16.)

⁶ Sentence (5) can equally express a different speech act—namely, a prediction.

Comment 5

The idiom $\text{'BYT'I.1} \dots \text{NE}^{\text{'11}}$ is syntactically discontinuous: its lexical components— 'BYT'I.1 and NE —are not linked syntactically to each other, but are connected via DSynt-actant II (= L('Y')); see Mel'čuk (2023a: 40–41). The idiom $\text{'BYT'I.1} \dots \text{NE}^{\text{'12}}$ is also discontinuous.

The syntactics [= the co-occurrence properties] of the idiom

Semantic co-occurrence

Comment 6

The Sem-actant 'Y' of the idiom is necessarily a fact, and it must be a 'completed fact'. An inverse constraint is imposed on the Sem-actant 'Y' of the idiom $\text{'BYT'I.1} \dots \text{NE}^{\text{'12}}$: here the fact must not be 'completed' (see **Comment 3** above). The Sem-actant 'X' of $\text{'BYT'I.1} \dots \text{NE}^{\text{'11}}$, that is, the Agent, can remain unexpressed—under the condition, though, that it is human.⁷ (NB: The Sem-actant 'X' of both idioms is not only the Agent: it can also be the Experiencer, the Patient, etc., according to the semantic nature of 'Y'; the term "Agent" is used here as a convenient abbreviation.) Russian allows the Speaker to not elaborate on the identity of the human Agent of the idiom $\text{'BYT'I.1} \dots \text{NE}^{\text{'11}}$: if 'X' is not expressed, fact Y can be impossible for people in general, for a particular person identified by the context, for the Speaker himself, or for the Addressee. Sentence (6), just as its translation, is vague; it can mean any Agent implied by the situation:⁸

- (6) *Ėtot tekst srazu ne ponjat'* lit. 'This text at.once is not to.understand'. =
'It is impossible to understand this text at once'.

By contrast, Sem-actant 'X' of the idiom $\text{'BYT'I.1} \dots \text{NE}^{\text{'12}}$ cannot remain unexpressed:

- (7) a. *Tam ne poguljat'* 'It is impossible to go for a walk there.' [$\text{'BYT'I.1} \dots \text{NE}^{\text{'11}}$]
vs.
b. **Tam ne guljat'* 'It will be impossible to go for a walk there.' [$\text{'BYT'I.1} \dots \text{NE}^{\text{'12}}$] ~
Tam nikomu/ljudjam ne guljat' 'It will be impossible for anyone/for people to go for walks there.'

⁷ A non-human Agent must be explicitly expressed: *Mašine_x tam bylo ne proexat'* lit. 'To.a.car there was not to.pass' = 'A car could not pass there', *Vinogradu_x zdes' ne vyzret'* lit. 'To.grapes here is not to.ripen' or *Daže slonu_x budet ne spravit'sja* lit. 'Even to.an.elephant will.be not to.manage'. If the Agent is not expressed, it means that it is human (L. Iomdin, personal communication).

⁸ See Zimmerling 2021 for a detailed analysis of Russian dative-predicative constructions in which the Experiencer remains unexpressed.

Syntactic co-occurrence

Comment 7

The idiom $\text{'BYT'I.1 ... NE}^1\text{'1}$ syntactically corresponds to an impersonal verb: its SSynt-subject is the empty pronoun $\emptyset_{(N, \text{neu})\text{SG}}^{\text{EMPTY}}$, which in Russian (and in several other Pro-Drop languages, such as Serbian, Polish, Spanish or Italian) is also a zero lexeme. This idiom is parallel to such Russian impersonal verbs as TOŠNIT' 'nauseate' (*Menja tošnit* lit. 'It nauseates me.') or ZNOBIT' 'make shiver' (*Menja znobit* lit. 'It shivers me.'). At the DSynt-level, just like these verbs, $\text{'BYT'I.1 ... NE}^1\text{'1}$ has DSynt-actants **II** and **III**—without DSynt-actant **I**. The same is true about the idiom $\text{'BYT'I.1 ... NE}^1\text{'2}$.

Comment 8

DSynt-actant **II** of the idiom $\text{'BYT'I.1 ... NE}^1\text{'1}$ —that is, the dependent infinitive $\text{L('Y')}_{\text{INF}}$ —must be in the perfective aspect, while DSynt-actant **II** of the idiom $\text{'BYT'I.1 ... NE}^1\text{'2}$, also an infinitive, must be in the imperfective. This corresponds to the semantic constraints imposed on the Sem-actant 'Y'—a completed fact in $\text{'BYT'I.1 ... NE}^1\text{'1}$ and anything but a completed fact in $\text{'BYT'I.1 ... NE}^1\text{'2}$. (As indicated in **Comment 3** above, 'completed' is one of the regular meanings expressed by the Russian perfective aspect.)

Morphological co-occurrence

The idiom $\text{'BYT'I.1 ... NE}^1\text{'1}$ is used in all three tenses of the indicative and in the conditional-subjunctive mood, while the idiom $\text{'BYT'I.1 ... NE}^1\text{'2}$ appears only in the present, see below.⁹ $\text{'BYT'I.1 ... NE}^1\text{'1}$ has no non-finite forms and, for obvious semantic reasons, it cannot express the grammeme of the imperative; the same is true about $\text{'BYT'I.1 ... NE}^1\text{'2}$. However, $\text{'BYT'I.1 ... NE}^1\text{'1}$ has the imperative form *bud'* that expresses the irrealis, albeit its use is not quite natural (while still being grammatical):

(8) *Bud' Ivanu ne spravit'sja, on by skazal ob ètom*

'Had it been impossible for Ivan to manage [the situation], he would have said so.'

⁹ There is a similar expression pattern involving the so-called *BY*-infinitive (Rus. *BY* is a clitic particle that expresses the conditional-subjunctive mood), as in the sentence *Ne xodit' by Ivanu tuda* 'It would be better for Ivan not to go there', which expresses a mild negative opinion about X doing or having done Y rather than impossibility of X's doing Y. This is one of genuine independent infinitives—the optative infinitive.

This marginal use is worth mentioning because for $\text{'BYT'I.1 ... NE'12}$ it is absolutely excluded: **Bud' Ivanu ne guljat' po Moskve...* 'Had it been impossible for Ivan to go for walks in Moscow...'

The idiom as a whole

Comment 9

The idiom $\text{'BYT'I.1 ... NE'11}$ is an approximate converse antonym with respect to the lexemes $\text{POLUČIT'SJA} \approx \text{'succeed'}$ and $\text{UDAT'SJA} \approx \text{'manage'}$, as well as to the collocations $\text{byt'I.1 'v sostojanii'}$ and $\text{byt'I.1 'v silax'1}$ (both meaning $\approx$ 'be able'):

$\text{'BYT'I.1 ... NE'11} = \text{Anti}_{012n}(\text{polučit'sja}) = \text{Anti}_{012n}(\text{udat'sja})$;

$\text{'BYT'I.1 ... NE'11} = \text{Anti}_{021n}(\text{byt'I.1 'v sostojanii'}) = \text{Anti}_{021n}(\text{byt'I.1 'v silax'1})$;
cf. (9):

(9) a. $\text{Mne}_{\text{III}} \text{'bylo ne' ponjat'_{II} ètogo}$

lit. 'To.me it was not to.understand this.' $\cong$

$\text{U menja}_{\text{II}} \text{ne polučilos' ponjat'_{I} èto}$

lit. 'At me not obtained to.understand this.' $\equiv$

$\text{Mne}_{\text{II}} \text{ne udalos' ponjat'_{I} èto}$

lit. 'To.me not succeeded to.understand this.'

b. $\text{Mne}_{\text{III}} \text{'bylo ne' ponjat'_{II} ètogo}$

lit. 'To.me it was not to.understand this.' $\cong$

$\text{Ja}_1 \text{ byl ne 'v sostojanii' ponjat'_{II} èto}$

lit. 'I was not in state to.understand this.' $\equiv$

$\text{Ja}_1 \text{ byl ne 'v silax'1 ponjat'_{II} èto}$

lit. 'I was not in strengths to.understand this.'

Comment 10

The approximate semantic equivalences based on the synonyms, the antonyms and the conversives of $\text{'BYT'I.1 ... NE'11}$ can be illustrated with the sentences in (10); all of these mean $\approx$ 'I cannot solve this problem':

(10) $\text{Èta zadača mne ne po zubam.}$ $\text{Ja ne v sostojanii rešit' ètu zadaču.}$

$\text{Èta zadača mne ne pod silu.}$ $\text{Ja ne mogu rešit' ètu zadaču.}$

$\text{Ja (vsë) nikak ne rešu ètu zadaču.}$ $\text{Mne ne po silam rešit' ètu zadaču.}$

$\text{Ja nesposoben rešit' ètu zadaču.}$ $\text{Mne ne pod silu rešit' ètu zadaču.}$

$\text{Ja ne v silax rešit' ètu zadaču.}$ $\text{Ne v moix silax rešit' ètu zadaču.}$

Comment 11

The idiom 'BYT'I.1 ... NE'¹ has two rather infrequent variants.

- A colloquial variant without negation (Mets, ed. 1985: 205), possible in the context of strong contrast; cf. (11):

(11) *Ivanu ne vstat', a mne vstat'!* lit. 'To.Ivan it is not to.stand.up, but to.me it is to.stand.up!' =

'It is impossible for Ivan to stand up, but for me it is possible!'

- A variant where the negation NE 'not' is replaced with one of the two synonymous idioms 'EDVA LI' and 'VRJAD LI', both meaning 'hardly' (such replacement is possible practically in all contexts):

(12) *Ivanu edva li <vrjad li> vstat'* 'It is hardly possible that Ivan stands up.'

For the idiom 'BYT'I.1 ... NE'¹ such variants are less acceptable.

Comment 12

The expression of the form *byt' ne V_{PERF, INF}* can be part of several idioms, which, of course, are different from our two idioms; for instance:

[X-u] 'POČEMU (BY)/ČEGO (BY) BYT'I.1 ... NE' [Y-nut']?

lit. 'To.X why/of.what it is not to.be to.do.Y?' = 'Why does X not do Y?' [a suggestion].

Čego (by)/Počemu (by) *Ivanu ne pojti poguljat'*? 'Why wouldn't Ivan go for a walk?'

[X-u] 'BYT'I.1 NE OTKAZAT'' [v Y-e] lit. 'To.X be not to.refuse in Y' = 'X really does have Y'.

V izvestnoj xitrosti Ivanu bylo ne otkazat' lit. 'In certain cunningness to. Ivan it was not to.refuse' = 'Ivan really had certain cunningness'.

[X-u] 'ČTO BY BYT'I.1 ... NE' [Y-nut']?

lit. 'To.X what it would not to.be to.do.Y?' = 'Why wouldn't X do Y?' [a suggestion].

A čto by Ivanu ne pojti poguljat'? 'And why wouldn't Ivan go for a walk?'

3 The idiom $\text{BYT}'1.1 \dots \text{NE}'12$

For this idiom the illustrative sentence is (13):

- (13) *Mne* $\emptyset_{\text{PRES}}^{\text{BYT}'1.1}$ *ne rabotat'* {*v Moskve*} lit. 'To.me it is not to.work in Moscow' =
'It will be impossible for me to work in Moscow.'

The semantic, deep-syntactic and surface-syntactic structures of this sentence are as follows:

Figure 4

SemS of sentence (13)

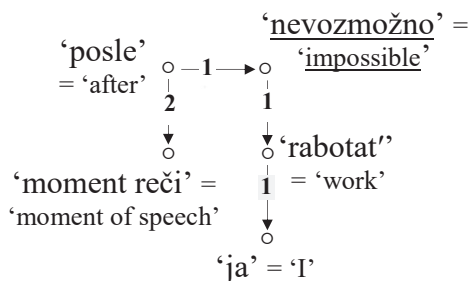


Figure 5

DSyntS of sentence (13)

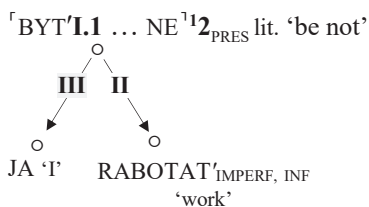
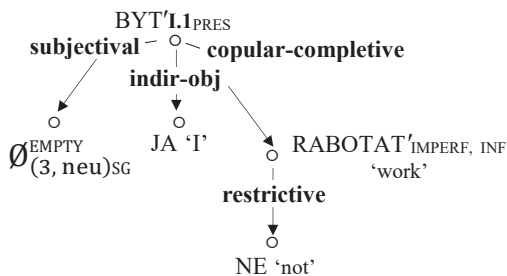


Figure 6

SSyntS of sentence (13)

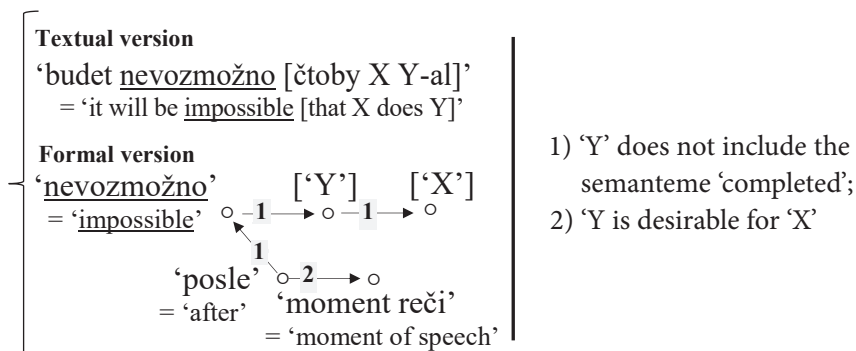


The configuration of semantemes 'will be impossible' is expressed in sentence (13) by the idiom $\text{BYT}'1.1 \dots \text{NE}'12$. Here is the lexical entry of this idiom.

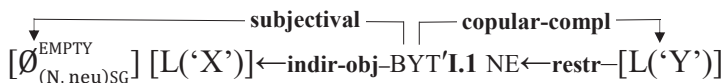
BYT'I.1 ... NE¹2, verbal idiom, impersonal; **only** PRES, **no** IMPER, INF, PART(icle), CONVERB.

Signified = Definition

‘[X-u] Ø^{BYT'I.1}_{PRES} ne [Y-at’]’ = .



Signifier [= the idiom’s SSynt-structure]



Syntactics

Government pattern

	‘Y’ ⇔ II	‘X’ ⇔ III
Ø ^{EMPTY} _{(N, neu)SG}	1. V _{IMPERF, INF} obligatory	1. N _{DAT} obligatory

Ne pit' Ivanu ètot kon'jak lit. ‘Not to drink to Ivan this cognac’. =
‘It will be impossible for Ivan to drink this cognac’.

Lexical functions

Conv_{23D} : ‘*ne svetit'*’ ≈ ‘not to be expected’

Anti_{23N} : *polučit'sja* ≈ ‘succeed’; *udat'sja* ≈ ‘manage’

Magn : *nikogda* ‘never’ | ←V_{INF}

Examples

Net, ne vstrečat'sja im v dal'nejšej žizni 'No, it will be impossible for them to meet in the future life.'

Ne slyšat' mne bol'se nikogda golos Ivana! 'It will be impossible for me to hear Ivan's voice ever again!'

Comments

The comments to the lexical entry of the idiom $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$ contain much information about $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$; therefore, the comments for the latter idiom are shorter.

The signified [= the meaning] of the idiom

Unlike the meaning of the idiom $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$, the meaning of this idiom includes the reference to the time: the impossibility will take place in the future; this particularity has a repercussion on the morphological properties of the idiom, see below.

The signifier [= the SSynt-structure] of the idiom does not require any additional indications.

The syntactics [= the co-occurrence properties] of the idiom

Semantic co-occurrence

The Sem-actant 'X' of the idiom $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$ must be obligatorily specified—it cannot remain unexpressed (in contrast to $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$, where the expression of 'X' is not obligatory). The Sem-actant 'Y' of $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$ can be any fact that is not 'completed'.

Syntactic co-occurrence

DSynt-actant II of the idiom $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$ —that is, the dependent infinitive $\text{L('Y')}_{\text{INF}}$ —must be in the imperfective aspect. This corresponds to the semantic constraint imposed on Sem-actant 'Y': it can be anything but a completed fact.

Morphological co-occurrence

The idiom $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$ appears only in the present tense. Its instances contain only the zero wordform of BYT'I.1 , the presence of which is postulated by analogy with the idiom $\text{'BYT'I.1} \dots \text{NE}^1\text{I}$.

The idiom as a whole

The expression of the form “*byt' ne V_{IMPERF, INF}*” can manifest other idioms, for instance:

- [X-*u*] ‘BYT’I.1 NE PRIVYKAT’ [k Y-*u*] lit. ‘To.X to.be not to.get.accustomed to Y’ = ‘Y is normal/habitual for X’.
Ivanu bylo k žare ne privykat’ ‘Heat was {something} normal for Ivan.’
- [X-*u*] ‘BYT’I.1 NE ZANIMAT’ [Y-*a*] lit. ‘To.X be not to.borrow Y’ = ‘X has a lot of Y’.
Tebe budet krasoty ne zanimat’ ‘You will have a lot of beauty’.

Unlike ‘BYT’I.1 ... NE’¹², these idioms are used in the three tenses.

Conclusions

The following three conclusions can be formulated: the first is a general lexicographic principle, the second presents a particular lexicographic recommendation, and the third concerns a specific question of Russian syntax.

1. An idiom must have its own complete lexicographic entry in the lexicon—just like any ordinary lexeme, and not be buried in the lexical entry of one of its components. Idioms were included as separate entries in the first published *Explanatory Combinatorial Dictionaries* of Russian (Mel’čuk & Zholkovsky 1984, Mel’čuk & Žolkovskij 2016) and French (Mel’čuk et al., 1984–1999), and the corresponding principle was explicitly stated and discussed in Mel’čuk (2013: 268–273). The above lexical entries demonstrate once again that an idiom necessitates, in the general case, rich lexicographic information, and therefore it must be entered in a lexicon as the headword of a full-fledged dictionary article, on the same footing as ordinary lexemes. Nothing but a dusty tradition stands in the way of this proposal. (A reminder: idioms and lexemes make up the class of lexical units).

2. A lexeme that is part of an idiom can feature some particular syntactic and/or morphological properties that are absent when this lexeme is used outside of the idiom; these properties are to be indicated in the lexical entry of the idiom. In other words, such a lexeme does not receive a special lexical entry describing it as this idiom’s component. Lexemes of this type are called deviant (Mel’čuk 2023a: 46). Thus, the verb BYT’I.1 used as a component

of our two idioms is a deviant lexeme; it should not receive a separate lexical entry representing it as this component: all the particularities this pseudo-BYT'I.1 shows inside the idioms are specified in the lexical entries of the idioms themselves.

Attention: The situation is different with respect to collocates. A collocate, even a unilexeme (= lexeme that appears in one or two collocations only), can have its own lexical entry characterizing it as a collocate.

3. The infinitive with the idioms 'BYT'I.1 ... NE'I.1 and 'BYT'I.1 ... NE'I.2—that is, their Sem-actant 'Y'—is not syntactically independent: as has been shown, it functions as copular-completive dependent of the verb BYT'. Thus, the list of controversial “independent infinitives” in Russian has one element less. The present paper, however, should by no means be construed as a global rejection of Russian independent infinitives: the language does have them. At least four types of the independent infinitive are obvious:

- Imperative infinitive (Padučeva 2017: 2.1.5): *Vstat'!* 'Stand up!'
- Nominating infinitive (in Russian *nazyvnoj infinitiv*; Padučeva 2017: 2.1.7):
– *Da čto ty govoriš? Exat' tuda sejčas?* 'But what are you saying? To go there right away?'
- Optative infinitive (Padučeva 2017: 2.2): *Tebe by snačala posmotret'!* 'You should first have had a look.'
- Narrative infinitive (Plungjan 2013): *A ty srazu kričat'!* 'And you up and started screaming right away!'

Future studies will confirm or refute the existence of other independent infinitives in Russian.

Acknowledgments

The draft of the paper was read and criticized by Leonid Iomdin, Lidija Iordan-skaja, Svetlana Krylosova, Sébastien Marengo, Jasmina Milićević and Alain Polguère. Their criticisms and suggestions helped me significantly improve the exposition. I ask them to kindly receive here the most heartfelt expression of my gratitude.

Abbreviations and symbols

DSynt- : deep-syntactic	∅ ^{EMPTY} _{(N, neu)SG}	: semantically empty zero nominal pronoun lexeme—dummy SSynt-subject; cf., e.g., Eng. IT in <i>It is night</i>
L : lexeme	∅ ^{BYT'I.1} _{PRES}	: zero present tense wordform of the verb BYT'I.1
L('X') : lexeme that expresses the meaning 'X'	[]	: the square brackets enclose the actants of the headword
-S : structure	{ }	: the curly brackets enclose the words added to the example for naturalness
Sem- : semantic	⌞ ⌟	: the top corners enclose an idiom
SSynt- : surface-syntactic	...	: the three dots in the name of an idiom indicate its syntactic discontinuity
‘ <u>σ</u> ’ : communicatively dominant semanteme		

References

- Apresjan, Ju. (2013). Slovarnaja stat'ja glagola BYT': ispravlennoe i dopolnennoe izdanie [The lexical entry for the verb BYT': a corrected and expanded edition]. In Gura, A., Belova, O., Berezovič, E., eds, *Slavica Svetlanica. Jazyk i kartina mira. K jubileju Svetlany Mixajlovny Tolstoj*, pp. 9–30, Indrik, Moskva. [See also: Apresjan, Ju., ed. (2014–2025, vol. 1: pp. 394–401).]
- Apresjan, Ju., ed. (2014–2025). *Aktivnyj slovar' russkogo jazyka* [Active Dictionary of Russian, vols. 1–5, A–L]. Jazyki slavjanskoj kul'tury/MCNMO, Moskva.
- Iomdin, L. (2010). Sintaksičeskie frazemy: meždu leksikoj i sintaksisom [Syntactic phrasemes: between the lexicon and the syntax]. In Apresjan, Ju., Boguslavskij, I., Iomdin, L., Sannikov, V., *Teoretičeskie problemy russkogo sintaksisa. Vzaimodejstvie grammatiki i slovarja* (pp. 141–190). Jazyki slavjanskix kul'tur, Moskva.
- Mel'čuk, I. (2013). *Semantics: From Meaning to Text*. Vol. 2. John Benjamins, Amsterdam/Philadelphia.
- Mel'čuk, I. (2019). Russian *U Y-a est'/- X* constructions. *Russkij jazyk v naučnom osveščeenii*, № 1 (№ 37), 7–45.

- Mel'čuk, I. (2023a). *General Phraseology: Theory and Practice*. John Benjamins, Amsterdam/Philadelphia.
- Mel'čuk, I. (2023b). Snova pro glagol БЫТ' – tri novye leksemy i koe-čto eščë [Once again about the verb БЫТ' 'be': three new lexemes and something more]. In Davidjuk, T., Isaev, I., Mazurova, Ju., Tatevosov, S., Fëdorova, O., eds, *Jazyk kak on est': sbornik statej k 60-letiju Andreja Aleksandroviča Kibrika* (pp. 388–397). Buki Vedi, Moskva. See also: https://olst.ling.umontreal.ca/static/pdf/%d0%9c%d0%b5%d0%bb%d1%8c%d1%87%d1%83%d0%ba_2023_%d0%9a%d0%b8%d0%b1%d1%80%d0%b8%d0%ba.pdf
- Mel'čuk, I. et al. 1984–1999. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques, Vols. I–IV*. Les Presses de l'Université de Montréal, Montréal.
- Mel'čuk, I., Polguère, A. (2018). Theory and practice of lexicographic definition. *Journal of Cognitive Science*, 19: 4, 417–470. See also: https://olst.ling.umontreal.ca/static/pdf/Mel'c%cc%8cuk_Polgu%c3%a8re_2018.pdf
- Mel'čuk, I., Zholkovskij, A. (1984). *Explanatory Combinatorial Dictionary of Modern Russian*, Wiener Slawistischer Almanach, Vienna.
- Mel'čuk, I., Žolkovskij, A. (2016). *Tolkovo-kombnatornyj slovar' russkogo jazyka. 2oe izdanie, ispravlennoe* [Explanatory Combinatorial Dictionary of Russian. Second edition, corrected]. Jazyki slavjanskix kul'tur, Moskva.
- Mets, N., ed. (1985). *Praktičeskaja grammatika russkogo jazyka dlja zarubežnyx prepodavatelej-rusistov* [Practical Grammar of Russian for Foreign Teachers of Russian]. Russkij jazyk, Moskva.
- Padučeva, E. (2017). Konstrukcija s nezavisimym infinitivom [Construction with an independent infinitive]. In *Materialy dlja proekta korpusnogo opisanija russkoj grammatiki*, http://rusgram.ru/Конструкция_с_независимым_инфинитивом
- Plungjan, V. (2013). "Narrativnyj infinitiv" v russkom jazyke: k klassifikacii kontekstov [The "narrative infinitive" in Russian: towards a classification of contexts]. In Kibrik, A. ed., *Lingvističeskij bespredel – 2. Sbornik naučnyx trudov k jubileju A. I. Kuznecovoj* (pp. 352–360). Izd-vo MGU, Moskva. See also: <https://www.academia.edu/36675139>
- Zimmerling, A. (2021). Russkie predikativy i ontologija sostojanij [Russian predicatives and the ontology of states]. *Komp'juternaja lingvistika i intellektual'nye texnologii*, 20, 738–748.



Lidija Iordanskaja

University of Montreal
Canada

A Pioneering Lexicographic Work: “New Explanatory Dictionary of Russian Synonyms” by Jurij Apresjan*

Abstract

The paper describes an innovative lexicographic product published in Moscow in the late 1990s—Ju. Apresjan’s *New Explanatory Dictionary of Russian Synonyms* [NEDRS]. NEDRS is essentially based on Apresjan’s theory of systemic lexicography and integral linguistic description. His introduction to NEDRS proposes several new important linguistic notions, such as lexicographic type, lexicographic portrait (of a lexeme), the dominant of a synonym series, etc. The main characteristic of NEDRS is the wealth and precision of the linguistic information supplied: thus, it specifies the referential and communicative status of the lexeme under consideration and its actants, the type of the speech act expressed, the capacity of constituting an independent utterance, etc.; in particular, it offers meticulous and exact analyses of semantic distinctions between synonyms. Such data were not provided before in other dictionaries. At the same time, NEDRS is extremely reader-friendly.

Keywords

Russian, semantics, lexicography, synonyms, synonym series, lexicographic notions

In his “Postface” to the *English-Russian Synonym Dictionary* (Apresjan 1979), Ju. Apresjan formulated some fundamental principles for the lexicographic description of synonym series; these principles were fully embodied in the said

* This article is an updated version of a previous Russian-language publication (*Russkij jazyk v naučnom osveščanii*, 2001, № 2, 240–249); it has been prepared especially for this issue of *Neophilologica* [Editors].

dictionary, where after a series of English synonymous lexemes a detailed Russian explication is supplied: the common part of their meanings is presented and the semantic distinctions between synonyms are stated and commented upon.

Apresjan's "New Explanatory Dictionary of Russian Synonyms"—NEDRS (NOSS—*Novyj Ob'jasnitel'nyj Slovar' Sinonimov russkogo jazyka*—in Russian) follows and develops the above principles. NEDRS had its two first issues published in 1997 and 2000 (Apresjan et al., 1997 and 2000).¹ These issues were compiled by a team of linguists headed by Jurij Apresjan. The first issue contains 132 lexical entries (= synonym series), and the second, 118. The vast majority of the words treated belong to the world of humans: the human mind and body, feelings, mental states, speech acts, etc. Each lexical entry is dedicated to one **synonym series** and features the following structure:

1) **The headword zone**, which includes

- all synonyms of the series
- the common meaning of these synonyms, that is, the intersection of their meanings
- lexicographic examples—specially constructed sentences to illustrate the typical contexts for the use of the synonyms in question
- the indication of semantically related series (not obligatory)

2) **The semantic zone**, which includes

- the list of semantic features that distinguish the synonyms of this series
- a detailed comparison of the synonyms

An example will help the reader better represent the sophistication and fineness of the description proposed. The series *NAMERENIE* 'intention' includes the synonyms *namerenie* 'intention', *umysel* 'aforethought/intent', *zamysel* 'idea/design', *zadumka* 'idea/point', *plan* 'plan', *proekt* 'project' and *prožekt* 'irrational project'; their shared semantic content is 'a thought of a person X about what he wants to do and the readiness to deploy the necessary efforts in order to do this'; the semantic features underlying the distinctions between these synonyms are: 1) the correlation of the thought and the will: *namerenie* 'intention' is, in the first place, the will, *plan* 'plan' being rather the thought; 2) the type of X: in X's *umysel* 'X's aforethought/intent', X can be only a person, and he has to execute his *umysel* himself, while in X's *plan* 'plan', X can well be an organization or a social body and X's *plan* can be realized by somebody else; 3) the character and scale of what

¹ The third issue of NEDRS was Apresjan et al., 2003; it contains 105 lexical entries. In 2004, the Apresjan team published a unified volume (Apresjan, Ju., ed., 2004), with 354 entries (= synonym series); this is a corrected and expanded edition. For obvious reasons, these later works are not covered in the present review.

is to be done: in *namerenie to do Y*, Y can be anything, but in *zadumka to do Y*, Y is something creative ($\approx$ artistic) and on a minor scale; etc. Five and a half large format pages of additional comments follow, supplying a real wealth of semantic information on the lexemes involved.

- 3) **The morphological zone** specifies morphological information about synonyms
- 4) **The syntactic construction zone**, which includes syntactic information, in particular, data about government and linear position
- 5) **The cooccurrence zone**, covering lexical and morphological cooccurrence
- 6) **The linguistic illustration zone**
- 7) **The semantic (quasi-)equivalent zones**
 - phraseological synonyms
 - analogs, that is, lexemes having meanings close enough to the common meaning of the series
 - conversives, exact and approximate
 - antonyms, exact and approximate
 - derivatives
- 8) **Bibliography** (not in all series).

NEDRS is a unique lexicographic work that embodies the results of Apresjan's many-year theoretical and practical research in semantics, lexicology and lexicography (Apresjan 1995a, b). Each entry of NEDRS achieves such depth and precision of linguistic description that, in point of fact, it can be considered a full-fledged scientific paper. Apresjan's brilliant introduction to the “Project of NEDRS” (Apresjan 1995c) gives an exhaustive presentation of this dictionary and shows how it differs from other Russian synonym dictionaries. To characterize NEDRS, we cannot avoid repeating (sometimes simplifying) central ideas of this introduction; to this, we will add our reader's impressions and some considerations and suggestions.

1 NEDRS is an innovative lexicographic product underpinned by a solid theoretical basis

1.1 Systemic integral description of language

NEDRS is developed following Apresjan's conception of systemic lexicography and integral description of languages (Apresjan 1995b).

The main principle of systemic lexicography is quite simple: the lexicon of the given language must be described as a system. The key notion here is a **lexicographic type**—a group of lexemes that feature multiple common semantic and syntactic properties and, as a consequence, require a uniform lexicographic description.

Integral linguistic description presupposes a tight coordination/collaboration of grammar (= system of general rules) and lexicon (= system of lexical units). Each lexeme gets a comprehensive characterization: a **lexicographic portrait**. As far as the exhaustiveness of the information about each lexeme is concerned, NEDRS is incomparable: thus, no dictionary we know indicates communicative, prosodic or referential properties of the lexeme and/or of its actants. The practical lexicography is, as a rule, far behind achievements of theoretical linguistics. But NEDRS does not have this defect: it is a direct practical embodiment of Apresjan's theoretical studies, as well as of those of other modern Russian linguists—such as I. Boguslavskij, M. Glovinskaja, L. Iomdin, E. Padučeva, etc. All linguistic notions involved are rigorously defined in a glossary of terms (found in the Introduction). This glossary has an independent value demonstrating the linguistic conception of Apresjan and, more generally, of the Moscow Linguistic School; thus, it presents such notions introduced by Apresjan as the 'Speaker's personal sphere', 'the Observer's frame', 'interpretational verbs' ("an interpretational verb does denote a specific action, but serves to for an evaluative interpretation of another, quite specific action": Issue 2, p. xxv; for instance, *MEŠAT'* 'interfere' or *SOBLAZNJAT'* 'seduce'), etc.²

The traditional view of **synonyms** is essentially refined. Firstly, the synonyms in a series must share many semantic components in the assertive part of their definitions, and the main components of the assertion (most often, it is the generic component) must be identical. Thus, the verbs *gordit'sja* 'be proud', *xvastat'sja* 'boast' and *risovat'sja* 'show off' are not synonyms, since their generic components are different: *gordit'sja* 'be proud' denotes a feeling, *xvastat'sja* 'boast' denotes a speech act, and *risovat'sja* 'show off', a type of behavior.³

Secondly, the differences between synonyms can be not only semantic or stylistic, but also concern their pragmatic properties, connotations, lexical cooccurrence, prosody, possible syntactic constructions, etc.

A new important notion is introduced: the **dominant of a synonym series**. This is the synonym that semantically is the least specialized and stylistically

² NEDRS's second issue contains a new, improved version of the terminological glossary. The third issue (2003) adds a score of new terms and offers corrections to several definitions.

³ The glosses proposed for Russian lexical units are quite approximate—they are meant only to help distinguish these units.

the most neutral. The dominant is used as the name of the series (that is, of the corresponding lexical entry) and thus determines the place of the series in the dictionary—synonym series are presented in the alphabetical order of the dominants.

1.2 Synonym series

In conformity with the principle of systemic description of the lexicon, the first zone of the majority of lexical entries contains the so-called **preamble**, which lists different semantically related synonym series and specifies the meaning differences between these and the head series (i.e., the series described in the given entry). At the same time, the meaning shared by all these series is indicated.

The next zone considers the semantic differences between the members of the head series. In this way, the treatment of lexical meanings is hierarchical: from the more general distinctions to the more specific ones. For instance, the preamble of the series *NADOEST'* ‘cause to be fed up’ (Issue 2 of NEDRS) contains other similar series related according to the component ‘reaction to a repetitive situation’ and establishes the main semantic differences between these series:

- The synonyms of the *NADOEST'* ‘cause to be fed up’ series (*naskučit'* ‘become boring’, *ostočertet'* ‘make sick and tired’, *opostylet'* ‘grow hateful’, *priest'sja* ‘become unpleasant because of frequent repetitions’) all indicate that the Experiencer wants to cease the contact with the fact that is repeated.
- The series *dosaždat'* ‘annoy’, *donimat'* ‘pester’, *dokučat'* ≈ ‘pester’, *presledovat'* ‘pursue/persecute’ puts into focus the impact of an unpleasant situation on the Experiencer.
- In the series *oprotivet'* ‘become disgusting’ and *omerzet'* ‘become loathsome’ the main semantic component is ‘disgust’.
- In the converse series *privyknut'* ‘get accustomed’, *prisposobit'sja* ‘adapt [one-self]’, *priterpet'sja* ‘adjust by being patient’ and *svyknut'sja* ‘get used’ the situation that repeats itself is presented as the norm.

The zone following the preamble presents the oppositions that distinguish the synonyms of the head series. All the semantic oppositions between different synonym series and inside of each series are formulated in the most uniform way—to make visible the systemic character of the lexicon organization. Thus, the preamble of the entry *TOLSTYJ* (Issue 2) establishes two classes of synonym series denoting the body volume different from the average: too big volume vs. too small volume. The lexical units of both series contrast according to the same features: the degree of deviation from the norm, the evaluation by

the Speaker, the type of the subject qualified by the adjective, and the reason for the deviation from the norm (there are in addition some more specific differences). The work on NEDRS can result in sets of relevant semantic oppositions for various lexicographic types.

The systemic orientation of the dictionary does not interfere with detailed descriptions of individual lexical units [= LUs] as separate entities with characteristic properties. These properties are presented in the fullest way possible, so that each LU gets its exhaustive lexicographic portrait.

1.3 New types of lexicographic information

In conformity with the principle of integral description of the lexicon, the characterization of each LU must take into account all general types of grammar rules. Therefore, the description of synonyms in NEDRS includes information that is new with respect to what is normally presented in traditional synonym dictionaries—for instance, communicative and prosodic properties. Here are a few examples.

- **Referential status** of the lexeme or of its actants. In the series AVTORITET ‘authority’ (Issue 2) actant I of the noun *prestiz* ‘prestige’ must be a generic term—in contrast to *avtoritet*, which does not have this constraint (*avtoritet Ivanova* vs. **prestiz Ivanova*, cf. *prestiz vrača* ‘the prestige of a doctor’).

- **Communicative status** of the lexeme. The adjective *blizkij* ‘close’ (series BLIZKIJ.1, Issue 2) can be a rheme, while its synonym *bližnij* ‘near’ cannot (*Bereg kažetsja blizkim* ‘The shore seems close’ vs. **Bereg kažetsja bližnim* ‘The shore seems proximate’). The adverb *poblizosti* ‘nearby’ (series BLIZKOI.1, Issue 2) can be a clause adverbial, but its synonym *blizko* ‘close’ cannot (*Poblizosti byl posëlok zolotoiskatelej* ‘Nearby there was a prospector settlement’ vs. **Blizko byl posëlok zolotoiskatelej*). The adverb *vdrug* ‘suddenly’ (series NEOŽIDANNO ‘unexpectedly’, Issue 2) tends to be a theme, while its synonym *vrasplox* ‘unawares’ prefers the rhematic position.

- **Focus of attention.** In the series UBEŽDAT’ ‘convince’ (Issue 2) the verbs *ubeždat* and *uverjat* ‘assure’ put into focus the addressee of the verb, while *dokazyvat* ‘prove’ focuses on the thesis. Consequently, the first two have the direct object corresponding to the addressee and the indirect object expressing the thesis: *ubeždat/uverjat kogo* ‘whom’ *v čëm* ‘of what’; with *dokazyvat* things are the other way around: *dokazyvat komu* ‘to whom’ *čto* ‘what’.

- **The type of speech act.** The verb *deržit* ‘talk back’ (the series GRUBIT’ ‘be rude’, issue 2) implies ‘answering rudely’, while its synonyms *grubit* and *xamit* ‘be very rude’ do not carry such an implication.

• **Semantic scope.** The adverb *neožidanno* 'unexpectedly' (series NEOŽIDANNO 'unexpectedly', Issue 2) has both the broad scope, characterizing the situation as a whole (*Neožidanno on vernulsja* 'Unexpectedly he came back', and the narrow scope, characterizing only some particular aspects of the situation (*Neožidanno on vernulsja ne odin* ⟨*očen' bystro/poveselevšim*⟩ 'Unexpectedly he came back not alone ⟨quite quickly/more joyful⟩'. Its synonyms *vdrug* 'suddenly' and *vnezapno* 'all of a sudden' cannot have a narrow scope (*Vdrug on vernulsja* 'Suddenly he came back' vs. **Vdrug on vernulsja ne odin* 'Suddenly he came back not alone').

• **Two new parameters for the description of the discursive expressions:**

—Capability to be used as an independent utterance. Thus, in the series VOVSE NE 'absolutely not' (Issue 2) we have *Vovse net!* 'Not at all!' and *Otnjud' net!* 'Absolutely not!', but not **Daleko net!* 'Far from not!'. Similarly, the expressions *Konečno* 'Of course' and *Razumeetsja* ≈ 'Obviously' (series KONEČNO 'of course', Issue 2) can be used inside a clause, as well as self-standing statement, while *bessporno* 'no doubt' is used mainly inside a clause.

—Factivity vs. putativity. The synonyms *bezuslovno* 'definitely', *nesomnenno* 'undoubtedly', *bessporno* 'beyond debate' and (to a lesser degree) *konečno* 'of course' can be within the scope of a putative verb, while their factive synonyms *razumeetsja* 'obviously' and *estestvenno* 'naturally' cannot: *Ja sčitaju, čto on nesomnenno* ⟨*bessporno, bezuslovno, konečno*⟩ *vyputaetsja iz ètoj istorii* 'I think that he undoubtedly ⟨beyond debate, definitely, of course⟩ will extricate himself from this incident' vs. **Ja sčitaju, čto on razumeetsja* ⟨*estestvenno*⟩ *vyputaetsja iz ètoj istorii*.

• **Fine-grained semantics of morphological forms.** The entry UBEŽDAT' 'convince' (Issue 2) indicates that the forms of the past imperfective of different synonyms have different meanings: contrary to *uverjat'* 'assure', the past imperfective forms of the verbs *ubeždat'* 'convince' and *dokazyvat'* 'prove' have multiple resultative meaning: the Subject managed several times to get his/her thesis accepted. The aspectual characteristics are essential in other parts of the dictionary entry as well: thus, the semantic distinctions between the synonyms of the UBEŽDAT' series are weakened and partially neutralized in the imperfective aspect.

• **Differences in metaphorical uses.** The adjective *nežnyj* 'tender' (the NEŽNYJ series, Issue 1) cannot be replaced by its synonym *laskovyj* 'affectionate' if it used in a figurative sense: *nežnyj* ⟨**laskovyj*⟩ *zapax/ton/vkus* 'smell/tone/taste'. In the series RISOVAT' 'draw', only *risovat'* allows for metaphorical use, but not *zarisovyvat'* 'sketch', *pisat'* 'paint', *malevat'* 'paint poorly': *Ego povest' risuet* ⟨**zarisovyvaet*/**pišet*/**maljuet*⟩ *kartinu npravov jazyčeskoj Rusi* 'His novel draws a picture of mores of pagan Rus' or *Moroz risoval* ⟨**zarisovyval*/**pisal*/**maleval*⟩ *zatejlivye uzory na stekle* 'The frost was drawing intricate patterns on the window glass'.

- **Data on polysemy.** The NEDRS discusses in sufficient detail the polysemy of some synonyms, for instance, of *vybrat'* 'choose' (the VYBRAT' series, Issue 2) or of *ljubit'* 'love' (the LJUBIT' series, Issue 2).

- **Conditions of neutralization of semantic distinctions between synonyms** (information absent from traditional synonym dictionaries). For instance, the IZBAVIT' 'deliver [Y from Z]' series (Issue 2) indicates that if the oblique object (deliver from what?) of the verb is an undesirable internal state of a human, the semantic distinctions between *izbavit'* and *spasti* 'save' can be neutralized: *On spas (izbavit') eë ot golovnyx bolej točnym massażem* 'He saved her from headaches by punctual massage'; for a different oblique object this distinction remains operational: *Tonkie steny baraka ne spasali (*ne izbavljali) ot vystrelov* 'The barack's thin walls did not protect against bullets'.

- **The system of usage labels** in NEDRS is much finer and more precise than in other dictionaries: for instance, such labels as **uxodjaščee** 'obsolescing' and **neobixodnoe** 'not everyday'.

2 NEDRS is a reader-friendly—"humane"—dictionary

The unique property of this dictionary is combining the exhaustiveness and precision of the information supplied with a very "humane" form in which this information is presented. An entry of NEDRS contains much more data than traditional dictionaries intended for people; in this respect NEDRS is closer to formal dictionaries exploited in computer text processing. However, unlike such formal dictionaries as the *Explanatory Combinatorial Dictionary of Russian* (see Melčuk & Žolkovskij 1984), which address a well-prepared user, NEDRS is aimed at a mass reader. Its authors minimize the use of a special metalanguage and symbols, while favoring extended explanations "in prose," supported by numerous well-chosen examples. This characteristic of NEDRS is manifested in its following features.

- A NEDRS lexical entry starts with the presentation of a synonym series with the indication of its dominant and a relatively formal specification of the meaning shared by all synonyms of the series. Then follow a few simple constructed examples ("lexicographic sayings") and the list semantic contrasts typical of the given series. After this the reader gets a detailed description of the similarities and differences between the series synonyms—in a free, but always transparent form. This information is offered instead of full lexicographic definitions

of synonyms, which formally would be sufficient to show these similarities and differences. However, in this case, the comparison of synonyms and extraction of the relevant knowledge would be incumbent on the reader, and this would be an “unfriendly” act by the authors, who want the dictionary to be at the disposal of average readers. Moreover, Apresjan et al. 1995: 111 shows that it is theoretically legitimate and justified to describe the meanings of synonyms by indicating their semantic similarities and differences; such a description, just as rigorous lexicographic definitions of the *Explanatory Combinatorial Dictionary* type, models metalinguistic practices of native speakers.

- A formal dictionary tries to cover **all** the uses of a given lexeme, disregarding different frequency and standardness; a definition can include a disjunction whose members are unequal with respect to the number and importance of the linguistic phenomena covered. In other words, a formal lexicographic definition does not reflect the native speaker’s intuition about the most typical uses of the lexeme. A formal dictionary’s only goal is to ensure the transition from an explicit representation of meaning to actual lexemes (and vice versa), which is necessary and sufficient for automatic text synthesis and analysis. But for a human user the information about the most standard cases is highly useful. Therefore, NEDRS is oriented towards a real picture of lexeme use, widely recurring to marks such as **obyčno** ‘as a rule’, **časče vsego** ‘most frequently’, **inogda** ‘sometimes’, **imeet tendenciju** ‘tends to’ and concentrating on the most typical uses of the lexeme under consideration.

- NEDRS is especially “humane” in what concerns the form in which information is presented. Firstly, the quantity of special symbols is minimized. Secondly, artificial metawords, such as **KAUZIROVAT’** ‘to cause’, are avoided. Thirdly, the dictionary is not limited to concise formulations that, strictly speaking, are logically sufficient, but pedagogically wanting. The user-friendliness of NEDRS is exactly in its explanatory power. After similarities and differences between the synonyms of the series are compactly represented in the synopsis, each section of the synopsis is described again—in free form, with justifications and developed explanations, supplied with convincing linguistic illustrations.

- NEDRS tries to explain—as much as this is possible—the phraseologized expressions including the lexeme under discussion by the metaphors induced by this lexeme. Thus, the differences in restricted lexical cooccurrence of synonyms *talant* ‘talent’ and *spособnosti* ‘abilities/gifts’ (TALANT, Issue 1) are linked to the fact that *talant*, unlike *spособnosti*, is perceived by Russian speakers as capital: one can *rastratit’* ‘squander’ or *sberegat’* ‘save’ *talant*, and *talant issjakaet* ‘gets exhausted’. *Zakat* ‘sunset’ (ZAKAT, Issue 2) is likened in Russian to *ogon’*

‘fire’: *Zakat vspyxivaet* ‘flares up’, *gorit* ‘burns’, *polyxaet* ‘blazes’, *dogoraet* ‘burns out’. The information of this type is useless for computer processing of phrasemes (since all the phrasemes must be listed in the lexical entry anyway), and it is not found in a formal dictionary. But for a human user it is quite valuable—exactly because of its explanatory power: for instance, it helps a language learner remember the corresponding phrasemes.

The restricted lexical cooccurrence is also explained through encyclopedic information or the lexeme’s meaning. Thus, *gostepriimstvo* ‘hospitality’ is considered in Russian as a specific ethnic feature, so that the phrases *russskoe/gruzinskoe gostepriimstvo* are more standard than *russskoe/ gruzinskoe radušie* ‘Russian/Georgian cordiality’; at the same time, since *radušie* denotes a particular behavior that manifests the joy of receiving a guest, this lexeme co-occurs naturally with adjectives characterizing behavior: *iskrennee/fal’sivoe/pokaznoe radušie* ‘sincere/false/ostentatious cordiality’.

- One of the obvious advantages of NEDRS is an extraordinary abundance of examples, cited in generous contexts: they really animate the lexeme illustrated. The dictionary’s reader (and we mean “reader”: this dictionary can be read as a book) gets esthetic pleasure not only from fine semantic analyses, but from examples as well.

- Finally, NEDRS seems to be as a declaration of love to the Russian language. The authors have a crush on their data: they analyze the meanings lovingly, they select their examples affectionately. Their approach reminds one of Haj Ross (1983) who, while preaching about “humane linguistics,” exclaims: “Love your data!” NEDRS is a perfect specimen of amorous lexicography.

It is worth mentioning that an impeccable mastery of language materials is characteristic of all works of NEDRS editor, Jurij Apresjan. We do not know of any other theoretical linguist whose publications are to such an extent saturated with concrete linguistic data and who could process it with such ease, while “feeling” each word in such a fine way.

3 Possible uses of NEDRS

People can use the dictionary with different interests and pursuing the most different goals.

- A Russian speaker who is interested in the language can take interesting trips through the dictionary making fabulous discoveries, concerning, in particular,

"some not obvious and difficult to grasp differences between synonyms" (Apresjan et al., 1995: 82)

- A teacher of Russian finds in NEDRS a huge lot of useful knowledge—such as, for instance, the indication of prototypical speech situations in which a particular synonym is to be used.

- A learner of Russian can receive all the information corresponding to his/her level of Russian: from the most basic to the most fine-grained.

- A lexicographer can learn quite a few important things by diving into NEDRS.

—Some **heuristic techniques appropriate for establishing semantic differences**. Thus, what is known as "the inner form" of the word (roughly, etymological structure) may prove useful. For instance, *metel'* 'blizzard' and *v'juga* '≈ blizzard with turning snow whirls' differ by the character of snow movements, which is reflected in the etymology: *metel'* *metët* lit. 'a blizzard is sweeping' (*met-* is the root of the verb *mesti* 'sweep'), that is, the snow is blowing parallel to the earth surface, but *v'juga kružit* lit. 'a blizzard is turning', that is, it is producing turning whirls. The verb *ostolbenet'* 'be rooted to the spot' (*stolb* is 'pole') implies the immobility of a person who is standing—'be turning into a pole'; its synonyms *ocepenet'* 'become numb', *okamenet'* 'be petrified' and *zameret'* 'freeze [figurative]' do not have such an implication. The meaning of the noun *radušie* 'cordiality' includes the components '*rad*+*ost*' = joy', because *radušie* denotes joy from being with pleasant visitors; *gostepriimstvo* 'hospitality' emphasizes the readiness for receiving *gosti* 'guests'.

—**A trove of data on lexical cooccurrence**. As a general rule, each synonym is supplied with typical phrases containing it; many of these phrases are in fact the collocations described by syntagmatic lexical functions. Several paradigmatic lexical functions are represented in NEDRS as well: conversives, antonyms and derivatives are indicated for many synonyms.

—**Valuable information for lexicographic definitions**. One of the most efficient methods for finding the relevant semantic components to go into the definition of a lexical unit is to compare this unit with semantically close units—and this is exactly what NEDRS does.

- NEDRS is useful for specialists of Natural Language Processing [= NLP], more precisely, for automatic text synthesis and analysis, as well as for text generation.

—In conformity with the principle of integral linguistic description, NEDRS presents the lexical information rigorously aligned to the grammar; therefore, it is a good foundation for a formal (computer) dictionary.

—NEDRS supplies some material for the construction of a concept hierarchy, which is paramount for several tasks of NLP, such as logical reasoning (≈ theorem

proving) and paraphrasing. One can indicate the replacement of a specific name by the generic term in order to avoid tedious repetitions: *Zdes' rastët mnogo berëz. Èto derevo...* 'Many birches grow here. This tree...'

—NEDRS offers the conditions for neutralization of semantic differences between synonyms, and these conditions are necessary in the paraphrasing rules that involve approximate synonyms; such rules are unavoidable in any sufficiently rich system of automatic synthesis. Note that paradigmatic correlates of a given synonym—antonyms, conversives, derivatives, analogs—are also quite handy for paraphrasing.

- A linguist finds in NEDRS a good deal of precious data related to various linguistic problems. As is to be expected, different lexicographic types contribute to solving different problems.

—NEDRS contains rich information concerning the relationship between linguistic meaning and extralinguistic reality, that is, concerning criteria usable to distinguish between semantic and encyclopedic knowledge. Here is an interesting example. The synopsis of the entry *UBEŽDAT'* 'convince' (Issue 2) says that the verbs *ubeždat'* and *uverjat'* 'assure' differ from *dokazyvat'* 'prove' by the prosody with which they are uttered. At first blush, you can think that this indication is not about the meaning, but rather about the description of the world: as stated in the synopsis, the situations "CONVINCE" and "ASSURE" presuppose the Addressee's emotional resistance, so that the Speaker has to deploy emotional efforts in order to overcome this resistance; hence a particular prosody. But you have to change your opinion when you see in the derivative zone the adjective *ubeditel'nyj* 'convincing' = 'expressing the desire to convince'; more than that, there is the phrase *ubeditel'nyj ton* 'convincing tone'. This shows that the indication of specific prosody of the verbs in question is linguistically justified.

—NEDRS offers some materials for the study of the following problem: Is it possible to explain different metaphorical uses of synonyms by their semantic differences? Take the entry *NEŽNYJ* 'tender' (Issue 1). If somebody is *nežnyj*, his actions are refined: they presuppose sweetness, carefulness and care, a special attention to trifles of everyday life—that is, the ability to perceive something non-obvious and difficult to grasp. The meaning of the adjective *laskovyj* 'affectionate', a synonym of *nežnyj*, does not carry these semantic components. Therefore, *nežnyj* develops another wordsense: 'tonkij = fine' = 'such that perceiving it requires an elevated ability to distinguish' (with the antonym *grubyj* 'rough'), while *laskovyj* lacks such a wordsense: *nežnyj/*laskovyj zapax* 'smell' (olfaction), *nežnyj/*laskovyj cvet* 'color' (vision), *nežnoe/*laskovoe mjaso* 'meat' (taste).

—Many observations found in NEDRS are, in point of fact, *sui generis* linguistic riddles, each of which warrants a special study. Here is an interesting example.

The adverb *poblizosti* 'nearby' (in the series *BLIZKO*1.1, Issue 2) can characterize the absence of some objects around a given reference point, while a semantically close adverb *nepodalěku* 'not far away' cannot: *Mašin poblizosti ne bylo* 'There were no cars nearby' vs. **Mašin nepodalěku ne bylo* 'There were no cars not far away'. Two other adverbs of this series are also unable to appear in a negative utterance, and both have the negative prefix *NE-*: *nedaleko* and *nevdaleke*, both meaning roughly 'not far away': **Mašin nedaleko/nevdaleke ne bylo* 'There were no cars not far away' (cf. *blizko* 'close' and *vblizi* 'nearby': *Mašin blizko/vblizi ne bylo* 'There were no cars close {to us}/ nearby'). The prefix *NE-* seems to add to the adverb a meaning that cannot enter into the scope of negation. The prefix *NE-* expresses in many adjectives and adverbs the opposite quality, but with a nuance of moderation. The adjectives *nekrasivyj* lit. 'non-beautiful', *nedobryj* lit. 'non-nice', *nevysokij* lit. 'non-high' are used instead of *urodlivyj* 'ugly', *zloj* 'mean', *nizkij* 'low' if the Speaker wants to soften a negative evaluation or to avoid excessive definiteness. The lexicographic definition of a lexeme with the prefix *NE-* could include the indication of the Speaker's attitude towards his/her evaluation, something along the following lines: 'the Speaker does not want this evaluation to be too definite' (see Apresjan 1995a: 312 on the necessity of such a component). The component of "intended indefiniteness" may block the use of *NE-*-prefixed adverbs in negative utterances. Cf.: *On mozet sdelat' èto xorošo/neploxo* 'He can do it well/lit. non-badly' vs. *On ne mozet sdelat' èto xorošo/*neploxo* 'He cannot do it well/non-badly'; or: *Xotja ego prosili govorit' tixo/negromko, on ne govoril tixo/*negromko* 'Although he was asked to speak quietly/non-loudly, he did not speak quietly/non-loudly'. A similar property of so-called weakly definite pronouns was stated by E. Padučeva (1985: 212–24): "A weakly definite pronoun cannot be within the scope of negation," as in *Ja tebe priněs koe-čto* 'I have brought you something' vs. **Ja tebe ne priněs koe-čego* 'I haven't brought you something'. Padučeva defines weak definiteness in the following way: a pronoun is weakly definite if its referent is known to the Speaker, but unknown to the Addressee; a particular case of this is a situation when the Speaker deliberately avoids definiteness. Therefore, the adverbs of the *nepodalěku*, *nedaleko*, *nevdaleke* type could be called weakly definite.

4 Shortcomings of NEDRS

The genre of review requires from its author to look for shortcomings in the publication under review. In this case, this task is quite complex: I cannot find in

NEDRS essential drawbacks, but see only some objective difficulties unavoidable in any study of this type. It goes without saying that such a huge work cannot be 100% free from minor inaccuracies and simple oversights. And although it is ungentlemanly to go nit-picking in a product of such a high quality, *le genre oblige*, and I will indicate a few problematic spots.

In the entry ТАЈНО ‘secretly’ (Issue 2) the synopsis lacks the indication of an important semantic feature that is used in the entry: ‘presence of the Observer’. According to this feature, the synonym series is divided into two sets: the presence of the Observer is relevant for *nezametno* ‘unnoticed’, *potixon’ku* ‘noiselessly’ and *ukradkoj* ‘furtively’; it is irrelevant for *tajno* ‘secretly’, *tajkom* ‘in secret’, *ispod-tiška* ‘on the sly’ and *vtixomolku* ‘on the quiet’.

The entry УЛЫБКА ‘smile_N’ (Issue 2) presents the shared meaning of the nouns *ulybka* ‘smile’, *usmeška* ‘grin’ and *uxmylka* ‘smirk’ as ‘stretched lips, which normally expresses that something is pleasant or hilarious for the subject’; does this mean that the generic meaning for these lexemes is ‘stretched lips’? It seems that this is not quite accurate: all types of smile are characterized by a general facial expression, where the whole face participates.

The entry УЛОВКА ‘ruse’ (Issue 2) says that the synonyms *xitrost’* ‘trick’ and *manëvr* ‘maneuver’ can both be used in the instrumental case to denote the manner in which an action is performed, while having different referential statuses: generic for *xitrost’* and specific-referential for *manëvr*. As illustration the following examples are given: *On xitrost’ju* ⟨**ëtoj xitrost’ju*⟩ *otvlëk eë vnimanie* ‘He detracted her attention by ruse (*by this ruse)’ vs. *Ëtim lovkim manëvrom on ustranil svoego sopernika* ‘He removed his rival by this deft maneuver’. However, the sentence *On ëtoj xitrost’ju otvlëk eë vnimanie* seems quite normal. It would be more accurate to say that *xitrost’* can be used in this role with any of the two statuses—generically or specific-referentially, but *manëvr* allows only for specific-referential use.

In some cases, one has the impression that two synonym series are separated without sufficient grounds. Thus, NEDRS presents the following two series:

- XUDOŠČAVYJ ‘lean’ (Issue 2): *xudoščavyj* ‘lean’, *podžaryj* ‘wiry’, *suxoj* lit. ‘dry’; ‘such that the volume of his body is **less** than average, because it contains no fat and **no extra** flesh’
- XUDOJ ‘thin [a living being]’ (Issue 2): *xudoj* ‘thin’, *toščij* ‘skinny’, *kostljavyj* ‘bony’; ‘such that the volume of his body is **much less** than average, because it contains no fat and **little** flesh’

It seems that the differing components do not impose the separation of these two series.

One of the main problems for the authors of NEDRS was polysemy. They had to isolate and describe the wordsense pertaining to the given series, and this requires many difficult decisions about the system of lexemes (= wordsenses) that compose the vocable under consideration. And the description of a multilexemic vocable raises the problem “a separate lexeme vs. a use of the same lexemes.” Under ‘a particular use of a lexeme’ the NEDRS authors understand such a use where the lexeme shows a minor semantic change with respect to its inherent meaning in specific contextual conditions. On many occasions, the authors speak of different ‘uses’ of a lexeme under description, since they obviously do not want to give this unit the status of a separate lexeme. It is quite probable that some solutions proposed with respect to real or potential polysemy will change in the process of further studies. It would be highly convenient for the dictionary’s user to have an index for already established wordsenses—since the present index of lexical units treated in NEDRS often does not make clear which wordsense is considered.

To conclude, let us repeat what we have started with: NEDRS is a huge step forward in the theoretical lexicology and practical lexicography. But not only in the lexicology/lexicography: it is a highly significant contribution to Russian studies in general and the whole of theoretical linguistics. (Remember, in particular, the system of formal linguistic notions introduced by Apresjan.) NEDRS, indeed, turned out to be a testing ground for a far more ambitious project: *Active Dictionary of Russian* (the first four volumes: Apresjan, Ju., ed. 2014—2023). This dictionary opens a new era in lexicology and lexicography—but here begins a new, different story.

References

- Apresjan, Ju. (1979). *Anglo-russkij sinonimičeskij slovar'*. Russkij jazyk, Moskva.
- Apresjan, Ju. (1995a). *Izbrannye trudy. T. I: Leksičeskaja semantika*. Jazyki russkoj kul'tury, Moskva.
- Apresjan, Ju. (1995b). *Izbrannye trudy. T. II: Integral'noe opisanie jazyka i sistemnaja leksikografija*. Jazyki russkoj kul'tury, Moskva.
- Apresjan, Ju., Boguslavskaja, O., Levontina, I., Uryson, E. (1995) *Novyj ob'jasnitel'nyj slovar' sinonimov russkogo jazyka: Prospekt*. Russkie slovari, Moskva.
- Apresjan, Ju., Boguslavskaja, O., Levontina, I., Uryson, E., Glovinskaja, M., Krylova, T. (1997). *Novyj ob'jasnitel'nyj slovar' sinonimov russkogo jazyka. Pervyj vypusk*. Jazyki russkoj kul'tury, Moskva.

- Apresjan, Ju., Boguslavskaja, O., Krylova, T., Levontina, I., Uryson, E., Apresjan, V., Babaeva, E., Glovinskaja, M., Grigor'eva, S., Ptencova, A. (2000). *Novyj ob''jasnitel'nyj slovar' sinonimov russkogo jazyka. Vtoroj vypusk*. Jazyki russkoj kul'tury, Moskva.
- Apresjan, Ju., Apresjan, V., Boguslavskaja, O., Krylova, T., Levontina, I., Uryson, E., Babaeva, E., Galaktionova, I., Glovinskaja, M., Grigor'eva, S., Iomdin, B., Ptencova, A., Sannikov, A. (2003). *Novyj ob''jasnitel'nyj slovar' sinonimov russkogo jazyka. Tretij vypusk*. Jazyki slavjanskoj kul'tury, Moskva.
- Apresjan, Ju., ed. (2004). *Novyj ob''jasnitel'nyj slovar' sinonimov russkogo jazyka. Vtoroe izdanie, ispravlennoe i dopolnennoe*. Jazyki slavjanskoj kul'tury — Wiener Slawistischer Almanach, Moskva—Vena.
- Apresjan, Ju., ed. (2014—2023). *Aktivnyj slovar' russkogo jazyka, Toma 1—4, A—J* [Active dictionary of Russian, Vols. 1—4, A—J]. Jazyki slavjanskoj kul'tury—MCNMO, Moskva.
- Melčuk, I., Žolkovskij, A. (1984). *Tolkovo-kombinatornyj slovar' russkogo jazyka*. Wiener Slawitischer Almanach, Wien.
- Padučeva, E. (1985). *Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju*. Nauka, Moskva.
- Ross, H. (1983). Human linguistics, *Forum Linguisticum*, 7: 3, 211–230.

This is an expanded and corrected English version of the publication in Russian—*Russkij jazyk v naučnom osveščanii*, 2001, №2, 240–249.



Alina Israeli

American University
USA

 <https://orcid.org/0009-0005-5345-879X>

Impact of the meanings of the verb *moč'* on its ability to form past tense and perfective

Abstract

It has long been noted that *moč'* – *smoč'* ‘can, be able to, may’ is an unusual aspectual pair, yet *moč'* has been commonly treated as having three meanings: alethic, deontic and epistemic. The article abandons this traditional treatment and identifies seven distinct meanings of the verb *moč'*: internal ability, permissive, epistemic, warning, perceived right, permission to begin, and external possibility. Their distinctions are not only semantic but also prosodic and in their ability to have a past tense form and/or to form a perfective based on *smoč'*.

Keywords

modality, possibility, aspect, verb classes, intonation

Introduction

The modal verb *moč'* ‘can, be able to, may’, just like its English counterparts *may*, *must*, *should*, represents a polysemic entity, and the goal of the present article is to identify its meanings, to disambiguate them and to determine how they correlate with tense and aspect.

Many studies discuss Russian modality in very general terms and include various expressions of modality, such as adverbials, conjugated modal verbs,

adjectives, etc. (Bulygina & Šmelëv (1997), Beljaeva (1990)). Van der Auwera & Plungian (1998) discuss possibility and necessity together and advance the distinction of participant-internal modality vs. participant-external modality. Petrušina & Li (2015) take the Van der Auwera & Plungian model verbatim and apply it to Russian future. Both of these studies have a very general approach to modality and attempt to create a very complex structure of modality, similar to a Periodic Table of Elements, and then to fill in individual boxes.

The traditional, top-down approach to modality as alethic, deontic and epistemic (Bulygina & Šmelëv (1997), Sokolova (ms), Petrova (2005), while applicable to Russian, does not capture the full range of meanings. Beljaeva (1990) and Choi (1999) examine ‘possibility (probability)’ and the verbs *moč’* – *smoč’* respectively thus narrowing the scope. Most works on *moč’* intermix present tense and past tense examples, for example Beljaeva (1990). Dictionaries treat *moč’* – *smoč’* as an aspectual pair, but Choi (1999) questions this assumption, while Janda (2018) emphasizes the unusual status of *smoč’* and its growing frequency. As will be shown later, some meanings of *moč’* can have a perfective variant and some cannot.

Strictly speaking, one has to explore each iteration of *moč’* separately:

možet V ⁱ¹	ne možet V ⁱ	smožet V ⁱ	ne smožet V ⁱ	možet ne V ⁱ
možet V ^p	ne možet V ^p	smožet V ^p	ne smožet V ^p	možet ne V ^p
mog V ⁱ	ne mog V ⁱ	smog V ⁱ	ne smog V ⁱ	mog ne V ⁱ
mog V ^p	ne mog V ^p	smog V ^p	ne smog V ^p	mog ne V ^p

For example, *Ona moglaⁱ zvonitⁱ¹* may mean ‘She could have called (and we did not know)’ or ‘She would (repeatedly) call’, a repeated action in the past. Meanwhile *Ona smogla^p zvonitⁱ²* is unlikely and hence hard to interpret. At the same time *Ona smogla^p govoritⁱ* ‘She could speak’ is perfectly correct. In other words,

¹ There are a number of abbreviations used in this article: V means the verb; superscripts p and i mean ‘perfective’ and ‘imperfective’ respectively; p preceded by a number, for example 1 p., means ‘first person’; sg. means ‘singular’, pl. means ‘plural’.

² The anonymous reviewer suggested that *Ona smogla^p zvonitⁱ čašče* ‘She managed to call more often’ is possible. I was unable to find a single example of “*smog* zvonitⁱ čašče*” (* signifies truncation) either in the RNC or on the web. However, there are multiple examples of *smogu/smogut zvonitⁱ čašče* ‘I / they will be able to call more often.’ which only points to the fact that this modal verb functions differently in different tenses and indicates that additional investigation is needed.

the combination $\text{MOG}^{\text{P}3} + \text{V}^{\text{P}}$ is possible but not for all verbs. But the present tense meanings of *moč'* introduce a number of meanings that are not represented by the past tense and vice versa. And not even all persons can be used with all meanings. For example, one of the meanings is 'granting permission' *On mozet voiti* 'He may enter', which by definition cannot have 1 p. sg. The reason for this is that in order to grant permission, there should be a verbal or non-verbal (conventional) prohibition of the action, for example one cannot enter a doctor's office or an official's office without permission. Yet it is hard to imagine circumstances where the prohibition comes from the speaker herself to herself, and where she can cancel the ban by granting herself permission.

Previous approaches to the verb in question have not identified all possible meanings of *moč'*. This empiric study attempts to remedy this lacuna. I will identify additional meanings⁴ (numbers in subscript will identify each type) and explore their correlation with the past tense and the perfective counterpart. As will be shown, not all meanings can have a perfective counterpart. I will also present the intonation pattern used with each meaning. Methodologically, the study is akin to an enhanced lexicographic entry, since it will expand the existing entry in BTS, which only lists four meanings. This modal verb is a classic case of polysemy (*mnogoznačnost'*) as discussed in Apresjan (1995). An analysis of the twenty morphologico-syntactic patterns listed above is beyond the scope of this article. Examples are drawn either from the Russian National Corpus (www.ruscorpora.ru hereafter RNC) or from internet sources. For readability, citations from the internet are uniformly marked as [WEB].

Let us examine the meanings of the verb *moč'* based on present tense usage.

I. Speaker (S) states that the Participant (P) has the ability to perform V (*moč'*₁)

Here the Speaker only assesses the Participant's ability. For skills, both impf (activity) and pf (accomplishments) are possible. The intonation is neutral.

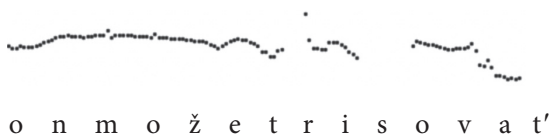
³ *Mog* is the masculine past tense form of *moč'*. Here MOG stands for any past tense form, masculine, feminine or neuter, sg. or pl.

⁴ Since it is impossible to prove the negative, I cannot prove that there are no other meanings, but in three years of collecting data I have not come across any additional meanings. Being a native speaker of Russian, I used my own intuition as well as the data I came across in addition to searching the RNC.

- (1) *On mo^žet risovat'ⁿⁱ.*

'He can draw.' (has the general skill)

Here is the intonation contour⁵ of (1):



It is very similar to IK-1 of Bryzgunova's (1977) classification; however, as the other six intonation contours below will suggest, her classification is limited and schematic and does not account for semantic variations.

The listed meaning of (1) is one of the possible meanings of this utterance, the other being permission, which I will discuss later. But what is important here is that *mo^žet risovat'* may mean the skill.

- (2) *On mo^žet [?]risovat'ⁿⁱ / narisovat'^p golovu l'va.* (has the specific skill)

'He can draw the head of a lion.'

- (3) *Čto **mo^žet risovat'**ⁿⁱ rebēnok v poltora goda, kogda ešče net nužnoj motoriki?*

'What can a one year and a half old child draw when he has none of the necessary hand motor skills?' [D. Rubina. Na solnečnoj storone ulicy (1980–2006)]

- (4) *Ona **mo^žet risovat'**ⁿⁱ kartiny, dobavila ona, no ona **ne mo^žet narisovat'**^p ni odnogo ieroglifa, daže ieroglifa, izobražajuščego ee sobstvennuju familiju.*

'She can draw/paint paintings, she added, but she cannot draw a single [Japanese] character, even the one designating her own last name.' [K. Simonov. Rasskazy o japonskom iskusstve (1958)]

With verbs of motion (VOMs), one would expect impf indeterminate infinitives.

- (5) *On mo^žet xodit' / plavat' / katat'sja na kon'kax.*

'He can walk / swim / skate.'

⁵ Many thanks to Paul R. Bennett for his help with the technology.

When the object is bounded (telic), only pf is licensed, as in (6), while in the absence of scope limitation represented by plural of the nouns, we find impf, as in (7):

(6) *On možet dokazat'*^{pf} / **dokazyvat'*^{im} *ětu teoremu.*

'He can prove this theorem.' (has the ability)

(7) *On možet dokazyvat'*^{im} *teoremy, rešat' zadači.*

'He can prove theorems, solve problems.'

On the other hand, (6') with the imperfective V is the permission to begin the action of proving and belongs to subtype VI.

(6') *On možet₆ dokazyvat'*^{im} *ětu teoremu.*

'He may begin proving this theorem.'

Example (8) shows scope limitation:

(8) *Nakonec, esli emu ne verjat, on sam lično možet dokazat'*^{pf}, *i daze sejšas, siju minutu, čto on, Saška, nyrnět i probudet pod vodoj rovno desjat' minut.*
'Finally, if they don't believe him, he himself personally can prove, even now, this minute, that he, Sashka, can dive and stay under water exactly ten minutes.' [A. Kuprin. *Listrogony* (1911)]

The past tense of this particular meaning means the loss of the skill. It could be the consequence of some event, such as an accident, aging or death, as in the next example, about the late artist Uemura Shōen or the late actor Cameron Boyce:

(9) *Ona mogla₁ risovat' obyčnyx ženščin točno takimi že, kak princess i naoborot.*
'She could draw regular women exactly like a princess and vice versa.'
[Ženščiny glazami japonskoj xudožnicy Uëmury Sëen, WEB (17.7.2025)]

(10) *On mog₁ pet', tancevat', igrat', on delal vsě èto s bleskom...*

'He could sing, dance, act, he did everything brilliantly...' [Poslednjaja rol' Këmerona Bojsa v grjaduščem seriale «Rajskij gorod» — kak smotret', WEB (17.7.2025)]

Negative construction

The negative construction of the type *on ne možit / ne mog risovat'* does not convey the meaning of lack of skill but rather of physical or emotional inability:

- (11) *On sovsem postarel. Ne mog xodit'. Ja byl k nemu očen' privjazan...*
(internal physical inability)
'He got quite old. He could not walk. I was very attached to him...' [S. Dovlatov. Naši (1983)]
- (12) *Posle smerti Georgija Aleksandroviča ja ne mogu xodit' v ètot teatr.* (internal emotional inability, not loss of skill but an inability to bear the pain associated with loss)
'After the death of G.A. I cannot go to this theater.' [S. Spivakova. Ne vsě (2002)]

Meanwhile a lack of skill is conveyed by the verb *umet'*:

- (13) *A umenija takogo ne bylo. Kris sovsem ne umel risovat'. Nikak.*
'He did not have such a skill. Chris did not know at all how to draw. Not a bit.' [M. Kantor. Čestnyj angličanin (2011)]
- (14) *Mal'čik eščë ne umeet xodit' — potomu tak devstvenno bely ego bašmački.*
'The boy still does not (know how to) walk — that is why his little shoes are so virgin white.' [does not have the walking skill] [I. Grekova. Fazan (1984)]

Perfective construction

Perfective *smoč'* can combine with the pf infinitive:

- (15) *On smožet^p dokazat'^p ètu teoremu / narisovat'^p golovu l'va.*
'He will be able to prove this theorem / draw a lion's head.'

Example (16) means that even mother will have enough skill to draw this:

- (16) *Èto smožet narisovat'^p daže mama!*
'Even Mom will be able to draw this!' [WEB (17.7.2025)]

The future perfective of *smoč'* conveys either a simple future, as in (17) or a potential conditional, as in (18):

- (17) *On otob'ětsja, **smožet**^P **dokazat**^P, čto èto pustye domysly.*
 'He will defend himself, he will be able to prove that this is all empty conjectures.' [A. Marinina. Illjuzija grexa (1996)]
- (18) *Esli est' malejšee narušenje v oformlenii dokumentov ili uščemleny prava graždan, ljuboj gramotnyj jurist **smožet**^P **dokazat**^P nezakonnost' sdelki.*
 'If there is the slightest violation in filling out of the documents or violation of citizens' rights, any qualified lawyer could prove the illegality of the deal.'
 [A. Zaxarov. Sem' sposobov obmanut' bližnego (2002)]

II. Speaker gives P permission to do V or suggests that P do V (*možet₂ napisat' mne*)

Permission for a specific pf action (19) vs. background impf permission (20):

- (19) *On **možet mne** [?]zvonitⁿⁱ / pozvonit^P.*
 'He can call me.'
- (20) *Možete / on **možet mne** zvonitⁿⁱ / [?]pozvonit^P v ljuboe vremja.*
 'You can / he can call me any time.'

The intonation contour of (19) (with *pozvonit^P*) is as follows:



Permission for an action stipulates a unique act, hence perfective, meanwhile giving permission to perform the action at *any time*. It must be noted that (20') is not incorrect, but rather than permission it means Speaker's assessment of Participant's possible behavior:

- (20') *On **možet mne** pozvonit^P v ljuboe vremja.*
 'He can call me any time [I expect him to call any time, it is possible that he would call any time].'

Permission cannot have past tense since there is no tense concord in Russian. Permission not to do something is also possible:

- (21) — *Načalo **možete ne čitat'***, — *govoril Viktoru Petroviču Kolja*.
 “‘You don’t have to read the beginning,’ Kolya would say to Viktor Petrovich.’

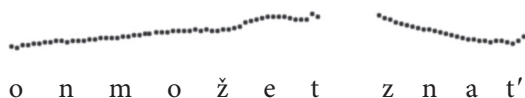
This type also has a neutral intonation. [B. Sarnov. Jura Krasikov tvorit čude-sa (1969)]

III. Speaker surmises Participant’s state (in Vendler’s sense) or whereabouts (*možet₃ guljat'*), the epistemic meaning

The Speaker’s assumption covers the realm of possibilities regarding the Participant, based on Speaker’s knowledge regarding that same Participant:

- (22) *On **možet znat'***.
 ‘He may know.’

Here is the intonation contour for (22):



- (23) *On **možet guljat'** / byt'ⁿ na rabote*.
 ‘He may be walking (be outside) / at work.’
- (24) *On **možet pozvonit'**^p / priiti^p v ljubuju minutu*.
 ‘He may call / come at any minute.’
- (25) *On **možet₃ uznat'**₁^p (find out)*.
 ‘He may find out.’

Example (25) deserves additional comment. Example (25') means that *he* has the ability to find out the issue at hand, while (25) means that the knowledge may befall him.

- (25') *On možet₁ uznat'₁^P* (find out).
'He can find out.'

All of these assessments can be used in the past tense:

- (26) *Vdrug est' kakie-to izvestija, ved' Sokolov **mog pozvonit'**^P...*
'What if there is some news, Sokolov indeed could call...' [V. Kungurceva. Vedogoni, ili Novye poxoždenija Vani Žitnogo (2009)]

Not all the English MAY- constructions lend themselves to the Russian *možet*:

- (27) He may be sleeping. — '*On možet spat'*. / *On možet byt' spit*.

Although it works well for *sleeping consciousness*:

- (28) *Vot tak i sovest': ež možno zagnat' v ugol, eju možno ne pol'zovat'sja, **ona možet spat'** — no v takom slučae ona eščė sposobna probudit'sja.*
'It is the same with consciousness: you can corner it, not use it, it may be asleep, but in this case it may still be awakened.' [Protoierej Dimitrij Smirnov. Propovedi (1984–1989)]

- (29) a. *On mog znat'^{ti}*.
'He may have known.'
b. *On mog ne znat'^{ti}*.
'He may not have known.'
c. *On mog pozvonit'^P*.
'He could call.'
d. *On mog priiti^P v ljubuju minutu.*
'He could come at any minute.'

Perfectivisation is possible for pf verbs while not for impf:

- (30) *On smožet^P uznat'^P*.
'He is capable of recognizing.' or 'He could find out.'
(31) **On smožet^P znat'^{ti} / guljat'^{ti} / spat'^{ti}*.
'He will be able to know / to stroll / be asleep.'

In both readings of (30), the Speaker implies that *he* will inadvertently be able to recognize the object or to find out something of which he has not been informed.

Negation is also possible:

(32) a. *On mo^žet ne znat'^{ri}.*

'He may not know.'

b. *On mo^žet ne ponjat'^{ri}.*

'He may not understand.'

While the intonation so far has been neutral, here there is a difference in emphatic stress (marked by bold letters):

(33) *On m**o**žet₂ katat'sjaⁱ.*

'He may skate.'

(34) *On mo^žet₃ kat**a**t'sjaⁱ.*

'He may be skating.'

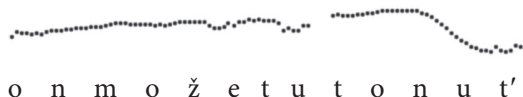
IV. Speaker fears Participant performing V and possible, likely negative, consequences or outcome (*mo^žet₄ utonut'^p*)

This meaning manifests itself in cases when the action V is highly undesirable:

(35) *On mo^žet₄ utonut'^p / pogibnut'^p / zabolet'^p / ozjabnut'^p.*

'He may drown / perish / fall ill / get cold.'

Here is the intonation contour for (40) with *utonut'*:



(36) *On mo^žet₄ ploxo o nas podumat'^p.*

'He may get a bad opinion about us.'

(37) *On mo^žet₄ nas uzna^t'₂^p* (recognize).

'He may recognize us.'

(38) *On mo^žet₄ ne pri^jti^p*.

'He may not come.'

(39) *Volodja zna^l, čto on-to sygra^et, no čto orkestr mo^žet₄ ne potjanu^t'^p*.

'Volodya knew that *he* would be able to play (it), but that the orchestra might not be up to the task.' [S. Spivakova. Ne vsě (2002)]

Past tense is possible:

(40) *On mog₄ utonu^t'^p / pogibnu^t'^p / zabolet'^p / ozjabnu^t'^p*.

'He could drown / perish / fall ill / get cold.'

(41) *On mog₄ ploxo o nas poduma^t'^p*.

'He could get a bad opinion about us.'

(42) "*On-to zna^l, čto èto označae^t!*" "*Mog₄ ne uspe^t'^p...*" — *neuverenno pred-položil Myškin*.

'"He (of all people) knew what that meant!" "He may not have had enough time..." haltingly surmised Myshkin.' [V. Belousova. Vtoroj vystrel (2000)]

Perfectivization is impossible, as if there were some effort on the part of the Participant to perform the negative action and achieve the negative result, which were in fact what the Speaker was dreading (both the action and the result):

(43) **On smo^žet utonu^t'^p*.

'He will be able to drown.'

(44) **On smog utonu^t'^p*.

'He was able to drown.'

In the case of *mo^žet₄* there is a slight rise in intonation towards the end; the syntagmatic stress is marked by bold letters.

(45) *On m**o**žet₂ \ ne prixodit'ⁱ*. — permission

'He doesn't have to come.'

- (46) *On mo^žet₄ ne prijti^p ↗*. — apprehension, warning
'He may not come.'
- (47) *On mo^žet₂ \ zarezat' barana*. — permission
'I give him my permission to slaughter a ram.'
- (48) *On mo^žet₄ zare^zat' ↗ barana*. — apprehension, warning
'He may kill the ram.'
- (49) *On mo^žet₄ zare^zat' barana ↗*. — apprehension, warning, fear
'I fear, he may kill the ram.'

V. Speaker states that Participant thinks that he has the right to do V (*mo^žet₅ nakazat'*)

Unlike *mo^žet₄* which signifies inadvertent damage (and the intonation goes up on the damaging action), *mo^žet₅* signifies Speaker's assessment of Participant's possible damaging action due to Participant's belief in his right to behave this way (and the intonation goes down on the damaging action). The Speaker is not in favor of the damaging or negative action V.

- (50) *On mo^žet₅ nakazat'^p *.
'He can punish.'

Here is the intonation contour for (53):



o n m o ž e t n a k a z a t'

- (51) *On mo^žet₅ izdevat'sjaⁱ *.
'He may taunt.'
- (52) *On mo^žet₅ nasmexat'sjaⁱ *.
'He may ridicule.'
- (53) *On mo^žet₅ nakričat'^p \/ nagrubit'^p *.
'He may scold, be rude.'

Past tense is possible:

(54) *On mog₅ nakazat'^p √.*

'He could punish.'

(55) *On mog₅ izdevat'sjaⁱ √.*

'He could taunt.'

(56) *On mog₅ nasmexat'sjaⁱ √.*

'He could ridicule.'

(57) *On mog₅ nakričat'^p √/ nagrubit'^p √.*

'He could scold, be rude.'

Perfective is impossible since this is an epistemological meaning and no effort on the part of the Participant is possible:

(58) **On smožet₅ nasmexat'sjaⁱ √.*

'He will be able to ridicule.'

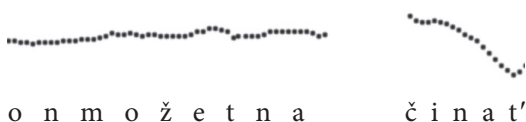
VI. Speaker gives the Participant permission directly or indirectly to begin V (*možet₆ načinat'ⁱ*)

The intonation here is very interesting: it goes down and then slightly up over the same vowel:

(59) *On možet₆ načinat'ⁱ √.*

'He may begin.'

Here is the intonation contour for (59):



(60) *On možet₆ razdavat'ⁱ √/ načat'^p razdavat'ⁱ √.*

'He may begin distributing.'

(61) *On mozet₆ vojti^P* ㄩ.

‘He may come in now.’

(62) *On mozet₆ nalivat^{ri}* ㄩ.

‘He may start pouring (wine).’

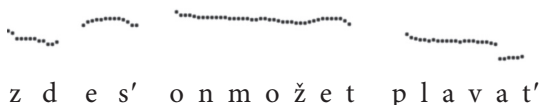
VII. Speaker assesses the external reasons for the Participant’s ability (or more often inability — *on ne mozet₇ prijti*) to perform the action V

The external obstacles and overcoming them are more visible in negative and perfective constructions, that is when the action was prevented from taking place or where the Participant was or will be unable to overcome the obstacle. Choi (1999: 49) discusses a person’s ability to swim; one has to have the skill, the inner ability (*mozet₁*), but also there must be a swimming pool where that person can exercise his or her ability:

(63) a) *Zdes’ on mozet₇ plavat^{ri}* ㄩ.

‘He can swim here.’

Here is the intonation contour of (63a):



b) *On zdes' smozet₇ plavat^{ri}*.

‘He will be able to swim here.’

c) *On mogⁱ₇ zdes' plavat^{ri}*.

‘He could swim here’

d) **On smogⁱ zdes' plavat^{ri}*.

Similarly to type I, impf past tense, as in (63c), signals that some major life event took place which caused further inability to swim, or that the person moved away and the place designated as *here* is no longer available to him. In

example (63d) pf *smog* would mean that he overcame some single obstacle, yet impf *plavat'* indicates a process, consequently there is a clash. It can be made acceptable by replacing the activity with a pf verb:

d') *On smogⁱ zdes' poplavat'^p.*
'He managed to have a swim here.'

This example underscores the fact that there is no mechanical correlation between the aspectual pair *moč'* – *smoč'*.

Another important point is that not all verbs combine with every meaning of *moč'*, including the three additional meanings. Some meanings are more restrictive than others: 'granting permission', 'permission to begin an action', and 'surmising the state or action' are considerably more restrictive.

In conclusion, I have described seven meanings of *moč'* based on present tense usage. They can be grouped as follows: One meaning is internal (internal ability), three are external (external conditions for V, permission to perform V, and directive to Begin V), and three are epistemic (whereabouts of P or engagement of P in some activity V, apprehension of P's perceived right to perform a negative V, and fear that P may accidentally perform an action with a negative outcome). I have thus subdivided Petrova's alethic group, which included both internal ability and external conditions. I have also enlarged her epistemic group by two types and her deontic group by one additional type.

I have shown that some meanings by the nature of their action (for example directive to begin) cannot have either past tense or a perfective counterpart or both. This explains the limitation of perfective *smoč'*.

Additionally, I have shown the importance of intonation for the disambiguation of modal constructions with *moč'*. It is worth noting that the intonation contours in the traditional (Bryzgunova) model are inadequate for categorizing the intonations used with *moč'*.

References

- Apresjan, Ju. (1995). *Leksičeskaja semantika*. (2-e izd.) Vostočnaja literatura RAN, Moskva.
- Van der Auwera, J., & Plungian, V. (1998). Modality's semantic map. *Linguistic Typology*, 2, 79–124.
- Beljaeva, E. (1990). Vozmožnost'. In A.V. Bondarko (Ed.), *Teorija funkcional'noj grammatiki. Temporal'nost'. Modal'nost'* (s. 126–142). Nauka, Moskva.

- Bryzgunova, E. (1977). *Zvuki i intonacija ruskoj reči*. Russkij jazyk, Moskva.
- Bulygina, T., & Šmel'ev A. (1997). *Jazykovaja konceptualizacija mira (na materiale ruskoj grammatiki)*. Jazyki ruskoj kul'tury, Moskva.
- BTS —Kuznetsov S. (Ed.) (1998) *Bol'soj tolkovyj slovar' russkogo jazyka*. Norint, Sankt-Peterburg.
- Choi, S. (1999). Semantics and Syntax of *moč'* and *smoč'*: Their "Aspectual" Relationship. *Russian Linguistics*, 23(1), 41–66.
- Janda, L. (2018). A Stranger in the Lexicon: The Aspectual Status of Russian *smoč'* 'be able, manage (to)'. In S. Dickey, & M. Lauersdorf (Eds.). *V zeleni drželi zeleni breg: Studies in Honor of Marc L. Greenberg* (pp. 105–125). Slavica Publishers, Bloomington.
- Padučeva, E. (2016). Modal'nost'. In V. Plungjan (Ed.), *Materialy k korpusnoj grammatike russkogo jazyka* (pp. 18–94). Nestor-istorija, St. Petersburg.
- Petrova, M. (2005). Semantičeskaja specifika form SV russkix glagolov so značeniem vozmožnosti. In A. Kibrik, *Komp'juternaja lingvistika i intellektual'nye texnologii: Trudy meždunarodnoj konferencii "Dialog 2005"* (pp. 420–424), RGGU, Moskva.
- Petruxina, E., & Li, Čž. [= Joohong Lee]. (2015). Modal'nost' glagol'nyx form buduščego vremeni v russkom jazyke. *Vestnik Moskovskogo universiteta. Serija 9. Filologija*, 4, 72–86.
- Sokolova, S. (ms) On the Interaction between the Categories of Aspect and Modality: A Case Study of the Russian Verbs *moč'* – *smoč'*.



Alexei Shmelev

 <https://orcid.org/0000-0002-5019-1525>

The Russian Words for Anxiety Revisited

Abstract

This paper deals with the Russian words for worry and anxiety. The analysis focuses on the word families of *bespokoit'* and *trevoga*. It is based on data from the Russian-English and English-Russian subcorpora of the *Russian National Corpus*. These data support the hypothesis that the words in the *bespokoit'* family involve a greater degree of intellectual assessment of the situation, whereas the words in the *trevoga* family point to a more intense and less uniform emotional experience. At the same time, the study helped verify certain details of earlier descriptions that had been obtained without parallel corpora.

The choice of a translation equivalent in each particular case depends not only on semantics but also on the overall translation strategy and the individual preferences of the translator. Nevertheless, certain general patterns can be observed. For example, with the word *bespokoit'sja*, as with the English word *worry*, it is important that the subject is thinking over the situation and considering that it may take a turn for the worse. In the case of *trevoga*, as with *anxiety*, a key element is that the subject lacks important knowledge about the situation and therefore does not know what can be done to prevent negative developments. The words *vstrevožit'* and *vstrevožit'sja*, like the English word *alarmed*, are characterized by a sudden realization that something bad might happen and that some action must be taken in response.

Key words

Russian language, semantics, translation, parallel corpora, lexicographic portrait

1. Preliminary remarks

The intended object of the present paper is the family of Russian words for uneasiness over some impending or anticipated event: *bespokojstvo*, *trevoga*, *bespokoit'(sja)*, *trevožit'(sja)*, *bespokojnyj*, *trevožnyj* as well as their prefixal derivatives. Most of them are polysemous.

The development of corpus linguistics and the appearance of electronic parallel corpora of the Russian language have made it possible to study the semantics of Russian lexical units under the assumption that one may regard translation equivalents of a linguistic unit extracted from real texts as a source of information about its semantics. In what follows, I seek to validate, disprove or improve the semantic analysis of the words under consideration obtained without recourse to parallel corpora.

It should be remembered that the choice of translation counterparts for key words of a given text strongly depends on the strategy followed by the translator. Individual translator solutions may be based on different considerations depending on the goals that the translator sets themselves. Generally speaking, one can distinguish between two approaches to translation. One is predominantly source-oriented and tries to bring across everything that the author 'meant' ('had in mind', 'wanted to say'). To this end, a translator may insert notes, commentaries, etc. Such a translation gives the readers a more or less complete understanding of the semantic content of the original yet does not give them a direct impression of the text. The other approach is predominantly oriented towards the readers and strives to make the translation create the same impression on the foreign-language speaker that the original has on the native speaker (this approach is used by certain translators of fiction).

The choice of strategy is determined by the translator's aims, and there is no universal strategy that is applicable in all cases. This is particularly true of literary translations (not only fiction in the strict sense of the word: also drama, verse, non-fiction and many non-literary genres). When a translator comes across an expression that is repeated several times throughout the text, they can go by its general meaning and try to select an equivalent that would best bring across the meaning of the original in each particular context ("contextual meaning strategy"), or they can choose a single equivalent of the expression and try to use it in all contexts wherever contextual usage warrants ("form strategy", or the "formal equivalence approach"¹). The result of the form strategy is a concordance between the source text and the translation with respect to a given expression, which takes place when this expression in the original is translated in each of its occurrences with the same expression; therefore, one may also refer to the form strategy as concordance strategy.

In the case of the contextual meaning strategy, there often arises a variety of different translations of a single word, especially if the latter has no direct equiv-

¹ Formal equivalence approach, a term coined by Eugene Nida, tends to emphasize fidelity to the lexical details and grammatical structure of the original language (Nida & Taber, 1969, p. 200).

alent in the target language and the translator must choose each time an equivalent that best expresses, in their opinion, the aspects of meaning that are the most significant in the given context. This may break the reader's impression especially as the constant repetition of a word might be part of the author's intent.²

Yet translations following the contextual meaning strategy may be of great use when one seeks to analyze the semantics of the corresponding expression. In each individual case, the translator has to decide what implicit aspects of its meaning can be sacrificed and what aspects are important for the meaning of the sentence or the text as a whole. These shifts lead to the emphasis of specific aspects of meaning and help to see relevant semantic components of the expression under consideration.

The concordance strategy often allows expressing author's intention in a better way than the contextual meaning strategy (although forced concordance, where the condition "Wherever contextual usage warrants" is ignored and normal usage is violated, gives the translation an "unnatural" and "wooden" sound³). However, it is less useful for the purpose of semantic analysis since it does not emphasize different semantic components of the original expression.

In addition, it should be taken into account that some translators choose unusual equivalents, which may be considered their individual stylistic trait.

Having this in mind, we can analyze the data from the RNC (<https://ruscorpora.ru/en>).

2. Apresjan on the Russian words for 'worry' and 'anxiety'

Apresjan made a subtle analysis of the above-mentioned Russian words. He proposed the following explication of the verbs denoting the emotions in ques-

² It is significant that Constance Garnett in her translation of Chekhov's short story *Toska* (included in the *Russian National Corpus*, RNC) renders the Russian language-specific word *toska* the same throughout (although she follows the contextual meaning strategy in most cases). Here, this emotion is at the center of attention (in this short work, the word occurs six times, not counting the title), and a variation of translations would have disrupted the author's intentions. Garnett translated the story's title as *Misery*, and this is also the translation that she gives in five cases out of six (in the remaining case, she slightly changes the syntactic structure and uses the related adjective *miserable*). Her translations of *toska* in other works show a great variance of equivalents: *yearning*, *despondency*, *anguish*, *misery*, *depression*, *dejection*, *distress*, *agony*, *restlessness*, *trouble*, *misgivings*, etc. (Shmelev, 2020, p. 663).

³ See (Beekman & Callow, 1989, p. 152).

tion, namely, *bespokoit'sja* ⁴, *volnovat'sja* ¹ and *trevožit'sja*: ‘To have unpleasant feeling that is usually felt when a person does not know something important about the situation regarding this person and this person fears (*on opasaetsja*) that the situation has changed or may change for the worse’ (Apresjan, 2004). His explication of *bespokoit'sja* ¹ in (Apresjan, 2014) is virtually the same: *A1 bespokoitsa iz-za A2* ‘A1 (a person) has unpleasant feeling as is usually felt when a person does not know something important about A2 (a situation) regarding this person and this person fears (*on opasaetsja*) that the situation has changed or may change for the worse.’

Some other explications go along the same lines: *bespokoit' 1* and *trevožit' 1*: ‘To cause an unpleasant feeling that is usually felt when a person does not know something important about the situation regarding this person and this person fears (*on opasaetsja*) that the situation has changed or may change for the worse’ (Apresjan, 2004); *A1 bespokoit A2*: ‘A1 (a situation) causes an unpleasant feeling in A2 (a person), the feeling that is usually felt when a person does not know something important about the situation regarding this person and this person fears (*on opasaetsja*) that the situation has changed or may change for the worse’ (Apresjan, 2014); *bespokoistvo 1*: ‘An unpleasant feeling that is usually felt when a person does not know something important about the situation regarding this person and this person fears (*on opasaetsja*) that the situation has changed or may change for the worse’ (Apresjan, 2014).

We can notice that these explications contain the verb *opasat'sja* ‘to fear’. In accordance with observations in (Zaliznjak & Mikaëljan, 2023), *opasenie* (the substantive form of the verb *opasat'sja*) involves a mental attitude rather than emotional state. I will illustrate this point with the following example. On the evening of the day of Resurrection, the disciples were meeting together behind locked doors “for fear of the Jews” (according to the Synodal translation, *iz opaseniija ot Iudeev*). They clearly understood what they were afraid of and what can be done to avoid bad things that might happen (they locked the doors). The words *trevo-ga* or *bespokoistvo* (instead of *opasenie*) would be totally inappropriate in this context because their focus is on the emotional state rather than mental attitude and the subject does not know what could be done to avoid bad things that might happen. However, as the above explications suggest, they are also linked with an on-going thinking process.

Let us also cite explications Apresjan suggested for some other lexical meanings of the words under consideration:

⁴ The numerals (1, 2, etc.) indicate distinct lexemes, not senses of a single lexeme, following Apresjan's lexicographic notation.

- **Bespokoit'sja 2:** *A1 bespokoitsa ob A2*
'A1 (a person) thinks about A2 (a situation) or the condition of A2 (a person) and try to ensure that they meet normal standards' (Apresjan, 2014).
- **Bespokoit' 2 and trevožit' 2:**
'to cause somebody (a minor) inconvenience or to disturb somebody's peace' (Apresjan, 2004).
- **Bespokoit' 2.1:** *A1 bespokoit A2 delaja A3*
'A1 (a person) disturbs the peace or comfort of A2 by the actions A3 because these actions suggest that A2 has to react somehow' (Apresjan, 2014).
- **bespokojstvo 2:**
'A situation that requires a person who is in this situation to do actions that are inconvenient for this person or disturb this person's peace' (Apresjan, 2014).

3. Russian-English subcorpus

In what follows, I will discuss the evidence provided by the Russian-English and English-Russian subcorpora of the RNC with special reference to the former.⁵ I will begin by looking at different ways of translating the words under consideration when these words appear in the original Russian text. In addition to the data of the Russian-English subcorpus of the RNC, I will refer to the text of *Cancer Ward* by Alexander Solzhenitsyn although it is not included into the RNC. The reason is that the words for 'worry' and 'anxiety' are of great importance for *Cancer Ward* (they may be called key words of the novel). Suffice it to say that Chapter 4 and Chapter 5 of *Cancer Ward* are entitled *Patients' Worries* (*Trevogi bol'nyx*) and *Doctors' Worries* (*Trevogi vračej*), respectively.

I will begin with the verb *bespokoit'sja* (in various lexical meanings). The verb is relatively frequent in the Russian-English subcorpus (592 samples in the corpus of 17,934,957 words as of February 2025, that is, 33 ipm). Translations of this word vary greatly, and the variation of its translations not only shows its high language specificity, but also helps to verify its semantic analysis. The inevitable loss of meaning during the translation of this word into English leads to the emphasis of specific aspects of meaning. The most common equivalent in the subcorpus

⁵ The expressions under consideration in the examples extracted from RNC (as well as from *Cancer Ward*) are boldfaced throughout this article.

(used 250 times out of the 592 occurrences of the verb *bespokoit'sja* in the original Russian texts) is the verb *to worry* (and *worried*). This word suggests a long-term thinking process related both to present and future. The subject notices certain signs in present that cause them to think that something bad may be happening or may happen in the future. It is often linked to thinking about someone else as in (1) and (2); however, the verb *bespokoit'sja* may also be used when one is supposed to think about oneself as in (3) and (4):

- (1) Ja mučitel'no načala za **nego bespokoit'sja**.
'I started **to worry about him**.' [Fedor Dostoevsky. *Netochka Nezvanova* (Jane Kentish, 1985)]
- (2) – **Ne bespokojsja za neë**, gospoža. S takoj farmakis (koldun'ej) ničego ne slučitsja.
'**Do not worry about her**, Mistress. Nothing can happen to a pharmakis, a sorceress, like her.' [Ivan Yefremov. *Thais of Athens* (Maria K., 2011)]
- (3) Ty že mlad, Kolja. A **za operaciju ne bespokojsja**. Ved' u našego doktora vse vyxodit xorošo.
'You're a brick, lad. And **don't worry about the operation**. Our doctor is a magician.' [Dmitry Medvedev. *Stout hearts* (David Skvirsky, 1961)]
- (4) – **O den'gax ne bespokojtes'!** – dobavila Margarita Tixonovna.
"'**Don't worry about money!**' Margarita Tikhonovna added.' [Mikhail Elizarov. *The Librarian* (Andrew Bromfield, 2015)]

Other translation equivalents that one can see in the Russian-English subcorpus of the RNC are *uneasy*, *anxious* (one could add to it constructions with the words *anxiously* and *anxiety*), *alarmed*, *afraid*, *(to) fear*, *to concern oneself* (and *concerned*), *to disturb oneself* (and *disturbed*), *nervous*, etc. These words emphasize different aspects of the meaning of *bespokoit'sja*. E.g., the word *anxious* (used 36 times as an equivalent of *bespokoit'sja*) emphasizes the component 'A1 does not know something important about the situation' as in (5), (6), (7):

- (5) — Kak žal', čto ja eë ne predvedomil! Očen' budet **bespokoit'sja** ...
'What a pity I did not tell her! She will be dreadfully **anxious**...' [Fedor Dostoevsky. *The Brothers Karamazov* (parts 3–4) (Constance Garnett, 1912)]

- (6) — V čem delo? — pričityvala Lida. — Počemu ego net, ja **bespokojus'**.
 'What on earth is the matter,' Lydia kept **worrying**. 'Why doesn't he come?
 I'm **anxious**.' [Vladimir Nabokov. *Despair* (Dmitri Nabokov, Vladimir Nabokov, 1965)]

- (7) Mama do six por ne prišla, ja **bespokojus'**.
 'My mother still hasn't come back. I'm quite **anxious**.' [Vasily Grossman. *Life and fate*. Part 2 (Robert Chandler, 1985)]

The word *uneasy* is even a more common translation equivalent of *bespokoit'sja* (it occurs as an equivalent of *bespokoit'sja* 86 times); in particular, Constance Garnett uses it consistently. The word *uneasy* emphasizes the component 'A1 has unpleasant feeling' as in (8) and (9):

- (8) — Da čto s toboj, Vasja? — zakričal on nakonec, edva dogonjaja ego. — Neuželi ty tak **bespokoish'sja?**..

"Why, what is the matter with you, Vasya?" he cried at last, hardly able to keep up with him. "Can you really be so **uneasy?**" [Fedor Dostoevsky. *A Faint Heart* (Constance Garnett, 1918)]

- (9) Èto **bespokoit** muža, **volnuet**, deržit v postojannoj **trevoge**; on ljubit svoju mat', očen' ljubit. Nu, i ja **bespokojus'**. **Duša bolit**.

'That **worries** my husband; it **troubles** him and keeps him in constant **agitation**; he loves his mother, loves her dearly. So I am **uneasy**, too, **my soul is in pain**.' [Anton Chekhov. *The New Villa* (Constance Garnett, 1900–1930)]

An important feature of *bespokoit'sja* is its tendency towards negative polarity. In more than half cases (344), the verb immediately follows the negative particle *ne*. In addition, there are examples with some distance between the particle *ne* and the verb (*ne nado bespokoit'sja*, *ne stoit bespokoit'sja*, *ne očen' bespokoil'sja*, etc.), e.g.:

- (10) — **Ne stoit bespokoit'sja**, prokurator, — skazal Afranij ...
 "No need to worry, Procurator," said Aphranius... [Mikhail Bulgakov. *Master and Margarita* (Richard Pevear, Larissa Volokhonsky, 1979)]

- (11) — Ja **ne bespokojus'**, — veselo skazal Mostovskoj, — ja **ne sobirajus' bespokoit'sja**.

‘I’m **quite calm**,’ said Mostovskoy brightly. ‘I’ve **got nothing to worry about**.’ [Vasily Grossman. Life and fate. Part 2 (Robert Chandler, 1985)]

It is worth noting that there is no clear borderline between *bespokoit'sja 1* and *bespokoit'sja 2* especially under negation. Such constructions deny that the subject (A1) is in the process of ongoing thinking about the situation in question; that is why A1 neither feels uneasy because of the situation (as *bespokoit'sja 1* would suggest) nor tries to improve the situation (as *bespokoit'sja 2* would suggest); so the distinction *bespokoit'sja 1* and *bespokoit'sja 2* is being blurred. In other words, *Ne bespokojtes'* ‘Don’t worry’ may imply either ‘Don’t be uneasy’ or ‘No need to act’ (or both), and it is often unclear which reading is meant (so, it is rather vagueness than ambiguity). However, in some cases, the difference becomes quite clear. In *The Dream* (filmed by Mikhail Romm in 1941), Anna (who works as a servant for madam Rosa Skorokhod, a Jewish shop and lodging-house owner) wrote in her matrimonial ad, “The Jews need not worry (*Evreev prosjat ne bespokoit'sja*),” and Rosa Skorokhod (played by Faina Ranevskaya) says, “The Jews are not worried (*Evrei ne bespokojatsja*).” In the ad, it is clearly *bespokoit'sja 2*, but Rosa prefers to interpret it as *bespokoit'sja 1*.

In half cases (296 out of 592), the verb *bespokoit'sja* is used in the imperative form immediately following the negative particle *ne*. This reinforces the observation that *bespokoit'sja 1* is often used in negative sentences in the imperative form (Apresjan, 2014). The most common translation of that construction (regardless of the lexical meaning of the verb) is *Don’t worry* (100 samples). Sometimes translators choose other equivalents, e.g., *Don’t disturb yourself*. Needless to say that Constance Garnett often translates the construction as *Don’t be uneasy* (38 samples)⁶, e.g.:

- (12) — **Ne bespokojtsja**, nynešnj raz ničego ne proizojdět.

‘**Don’t be uneasy**, nothing will happen this time.’ [Fedor Dostoevsky. The Brothers Karamazov (parts 3–4) (Constance Garnett, 1912)]

- (13) — A teper’ **ne bespokojtes'**, ja i sama ne budu gljadet’ na vas, a budu vniz smotret’.

‘But **don’t be uneasy**, I won’t look at you now. I’ll look down.’ [Fedor Dostoevsky. The Possessed, or The Devils (Constance Garnett, 1913)]

⁶ Sometimes other translators choose this equivalent as well.

- (14) — **Ne bespokojtes'**, Rodion Romanovič, esli b ja xlopotal v svoju vygodu, to ne stal by tak prjamo vyskazyvat'sja...

'**Don't be uneasy**, Rodion Romanovitch, if I were working for my own advantage, I would not have spoken out so directly.' [Fedor Dostoevsky. *Crime and Punishment* (Constance Garnett, 1914)]

However, sometimes (not so often though) she uses other constructions, e. g. *Don't worry (yourself)*, *Don't trouble (yourself)*, *Don't disturb yourself*, *Don't be frightened*, *Don't be anxious*, *Don't distress yourself*, etc. (her choice of an equivalent in each particular case seems arbitrary).

One should also mention the construction with the verb *izvolit'* 'to desire' (a little obsolete) in the imperative form, namely, *Ne izvol'te bespokoit'sja* (literally, 'Do not desire to worry'). Common translations are again *Don't worry* with some marker of politeness, e.g.:

- (15) — **Ne izvol'te bespokoit'sja**, messir, — otozvalsja Azazello...

'**Kindly do not worry**, Messire,' responded Azazello...' [Mikhail Bulgakov. *Master and Margarita* (Richard Pevear, Larissa Volokhonsky, 1979)]

- (16) Gosudarynja, **ne izvol'te bespokoit'sja** — svernëm my šeju ètomu Artamoše.

'Your Highness, **please don't worry** — we'll throttle that Artamosha.' [Vladimir Sorokin. *Day of the Oprichnik* (Jamey Gambrell, 2011)]

Again, Constance Garnett often renders *Ne izvol'te bespokoit'sja* as *Don't be uneasy*.

It should be observed that there are two prefixed derivatives semantically linked to *bespokoit'sja* 1 and *bespokoit'sja* 2, namely, *obespokoennyj* 'worried' (passive participle derived from the causative perfective verb *obespokoit'* <X-a> 'to cause <X> to start worrying'), which means roughly the same as *bespokoit'sja* 1 (that is, it refers to unpleasant feeling caused by thinking about some situation, which may change to the worse)⁷, and *pobespokoit'sja* <o čëm-libo> 'to take care <of smth>' (that is, the verb is semantically linked to *bespokoit'sja* 2).

Various forms of the verb *obespokoit'* (mainly the participle *obespokoennyj*) occur in the Russian-English subcorpus of the RNC 85 times. A common translation is *to worry (worried)*; it appears in the subcorpus 13 times, e.g.:

⁷ It should be mentioned that this word, contrary to *bespokoit'sja*, does not show negative polarity.

- (17) Vo vremja razgovora u starogo Genri byl očen' **obespokoennyj** vid.
 'Throughout the conversation old Henry looked very **worried**.' [Ilya Ilf, Evgeny Petrov. Little Golden America (Charles Malamuth, 1944)]
- (18) Vstretivšie nas molodye i zagorelye rostovskie čekisty byli **obespokoeny**.
 'The young, tan Rostov Chekists who met us were **worried**.' [Vladimir Sorokin. Bro (Jamey Gambrell, 2011)]

However, *to disturb* (*disturbed*) is even more frequent (32 samples), e. g.:

- (19) Tol'ko teper', **obespokoennyj** krasotoj i tajnoj, ogljadevšis', obnaručil Sëmka, što meždu stenami i polom ne prjamoj ugol, a strogoe, pravil'noe zakruglenie želobom vnutr'.
 'Only now, **disturbed** by the beauty and mystery, did Syomka look round and discover that instead of a right angle between the walls and the floor, there was an even line of masonry curving inwards.' [Vasili Shukshin. A Master Craftsman (Kathleen Mary Cook, 1980)]
- (20) On byl potrjasën i **obespokoën**.
 'He was in shock and **disturbed**.' [Vladimir Sorokin. Bro (Jamey Gambrell, 2011)]

Translators often choose that equivalent when the feeling is caused by some external situation.

One should also mention *trouble* and *troubled* (it occurs as a translation equivalent of *obespokoit'* and *obespokoennyj* 10 times). Other translation equivalents occur much less frequently.

As for the verb *pobespokoit'sja*, it turned out that its frequency in the Russian-English subcorpus was very low, which was contrary to my expectations since the word is commonly used in modern Russian discourse (consider: *Sko-ro zima, nado pobespokoit'sja o drovax* 'Winter is coming, we must think about the firewood'). One can find there three examples only, of which only one corresponds to its conventional meaning in modern Russian:

- (21) Kak-to v xolodnyj zimnij den' ja zametil, što Ivanov, tol'ko što pribyvšij v lager', vse žmëtsja k kostru, droža ot xoloda. Ja podošël k nemu i uvidel, što pidžak ego byl nadet na goloe telo. Tixij, skromnyj čelovek, on ne ščadil svoego zdorov'ja radi vypolnenija zadanija. Posle ètogo slučaja my,

konečno, **pobespokoilis'** ob odežde dlja našego zdolbunovskogo svjaznogo.

'One frosty day in winter I noticed that Ivanov, who had just arrived in the camp, was huddled up with cold at one of the fires. I went up to him and saw that he wore nothing beneath his jacket. Quiet and unassuming he did not stint his health where the fulfilment of an assignment was concerned. Naturally, we **saw to it** that our Zdolbunovo messenger was given adequate clothes.' [Dmitry Medvedev. Stout hearts (David Skvirsky, 1961)]

As for the verb *trevožit'sja*, it is much less frequent in the Russian-English subcorpus than *bespokoit'sja* (93 samples, that is, only a little more than 5 ipm). Common translation equivalents of *trevožit'sja* are virtually the same as the equivalents of *bespokoit'sja*, namely, *worry* (including the form *worried*), *alarm* (including the form *alarmed*), *anxious* (plus *anxiously* plus *anxiety*). The examples are not numerous; but in general, the translators chose the equivalents along the same line as for *bespokoit'sja*. E.g., *anxious*, *anxiously* or *anxiety* are sometimes used to indicate the subject does not know something important about the situation as in (22):

- (22) — Pëtr Lavrent'evič, — tože vpolgolosa sprosila Štrum, — kak tam Mad'jarov, blagopolučen? Pišet on vam? Ja inogda očen' **trevožus'**, sam ne znaju otčego.

'Pyotr Lavrentyevich,' asked Viktor in the same hushed voice, 'how's Mad'yarov? Is he all right? Have you heard from him? Sometimes I get very **anxious**. I don't know why.' [Vasily Grossman. Life and fate. Part 2 (Robert Chandler, 1985)]

The verb *vstrevožit'sja* and the passive participle *vstrevožennyj* (derived from the causative verb *vstrevožit'*) deserve special attention.

The verb *vstrevožit'sja* refers to the starting point of the emotional state *trevožit'sja*. A typical scenario assumes that the subject has just become aware of some event and thinks that it is necessary to act immediately, otherwise something bad may happen; however, the subject does not know exactly what can be done. One might expect that a common translation equivalent of *vstrevožit'sja* would be *alarm* or *alarmed* since the semantic features of *vstrevožit'sja* correspond to the essential elements of the emotional *alarm*: the suddenness, the mobilization to action, the awareness of an impending danger, the need to act immediately, and an element of uncertainty since the alarmed person does not

have a clear plan of action (Wierzbicka, 1999, p. 81). This hypothesis is only partially correct: *vstrevožit'sja* is often translated as *alarm* or *alarmed* (18 times out of 43), but this is not the only translation equivalent. As a rule, *alarm* or *alarmed* are chosen when the starting point of the emotional state is in focus (sometimes, this idea of starting point is strengthened by such words as *to become*, *suddenly*, etc.), e.g.:

- (23) Dobrodušnaja fel'dšerica Praskov'ja Fëdorovna navestila poëta vo vremja grozy, **vstrevožilas'**, vidja, čto on plačet, zakryla štoru, čtoby molnii ne pugali bol'nogo, listki podnjala s polu i s nimi pobežala za vračom.
 'The good-natured nurse Praskovya Fyodorovna visited the poet during the storm, **became alarmed** on seeing him weeping, closed the blinds so that the lightning would not frighten the patient, picked up the pages from the floor, and ran with them for the doctor.' [Mikhail Bulgakov. Master and Margarita (Richard Pevear, Larissa Volokhonsky, 1979)]
- (24) Rebjata, vidimo, rasskazali o žurnale doma, roditeli **vstrevožilis'**, i razraziljsja skandal.
 '...some of the boys talked about the magazine at home, their parents grew **alarmed** and a row erupted.' [Vladimir Bukovsky. To build the castle. My life as a dissenter (Michael Scammell, 1979)]

However, some translators regard the words like *to become* as sufficient means to convey the idea of starting point:

- (25) Verxovnaja žrica vynula iz čaški nemnogo temnoj mazi, krestoobrazno pročertila eju Tais meždu grudej, pod nimi i obvela soski. Na kože vystupila sinevataja okraska. Tais **vstrevožilas'**, ne ostanutsja li pjatna.
 'The high priestess took a little bit of dark ointment from the cup and drew a cross between Thais' breasts, then underlined the breasts and circled the nipples. The skin became bluish. Thais **became worried** about whether the stains would last.' [Ivan Yefremov. Thais of Athens (Maria K., 2011)]

As for the passive participle *vstrevožennyj*, it means roughly the same as *trevožit'sja* (that is, it refers to the same emotional state, maybe with a slight accent on the starting point of the state). However, it occurs in the Russian texts from the Russian-English subcorpus a little more frequently (namely, there are 128 samples of *vstrevožennyj* in the subcorpus). The words *alarm* and *alarmed*

are used as translation equivalents 30 times (they are the most frequent equivalents). Beside them, other equivalents occur: *anxious* (and *in... anxiety*), *agitated* (and *agitation*), *disturbed*, *frightened*, *worried*, *troubled*, *upset*, *distressed*. As one can see, most of them presuppose an external cause of the emotional state.

The causative verbs *bespokoit'* and *trevožit'* are polysemous as one can see from the above explications of *bespokoit' 1*, *trevožit' 1* and *bespokoit' 2*, *trevožit' 2*. The verbs *bespokoit'* and *trevožit'* are semantically very close in both lexical meanings; however, they behave slightly differently in different contexts.

The verbs *bespokoit' 1* and *trevožit' 1* differ in perspective. The verb *bespokoit' 1* puts emphasis on the resulting emotional state whereas *trevožit' 1* draws attention to the cause of the emotional state. Accordingly, *bespokoit' 1* is used less frequently than *bespokoit'sja 1* (the latter refers to the emotional state directly) whereas *trevožit' 1* (which refers to the causation) is used more frequently than *trevožit'sja*; cf. (Apresjan, 2004).

The translation equivalents of the verbs *bespokoit' 1* and *trevožit' 1* are generally the same as those of the verbs *bespokoit'sja 1* and *trevožit'sja*. A relatively common translation for *bespokoit' 1* is *worry*, and for *trevožit' 1* – *worry* and *alarm*. However, overall, the materials from the Russian-English subcorpus of the *Russian National Corpus* (RNC) do not provide reliable additional information about the semantics of the verbs under consideration.

Much more interesting are the Russian-English subcorpus data related to the verbs *bespokoit' 2* and *trevožit' 2*.

When someone deliberately and briefly interrupts another person's peace or activity in order to ask for a small favor, the verb *bespokoit'* is more commonly used (according to the numbering of meanings adopted in [Apresjan, 2014], it is *bespokoit' 2.1*). Although Apresjan (2004) writes that in this type of usage *bespokoit'* and *trevožit'* are completely synonymous, the texts show that *trevožit'* is not commonly used in this context. In *Cancer Ward*, the orderly Elizaveta Anatolyevna says to Dr. Dontsova (head of the radiology department):

- (26) Ljudmila Afanas'evna! — skazala ona, čut' izgibajas' v izvinenie, kak èto byvaet u povyšenko vežlivyx ljudej. — Mne očen' nelovko **bespokoit'** vas po melkomu povodu, no ved' prosto berët otčajanie! – ved' net že trjapok, sovsem net!

“Ludmila Afanasyevna,” she said, slightly bowing her head in apology, as excessively polite people sometimes do, “I’m sorry **to trouble** you with such a small thing, but it really is enough to drive one to despair. There are no cleaning rags, absolutely none.”

The verb *trevožit'* instead of *bespokoit'* would seem completely inappropriate to many native speakers of Russian in this context. It would be perceived not as a disruption of peace, but as the causation of an emotional state of anxiety (i.e., as *trevožit' 1*). This is also true for many examples in the RNC in which the verb *bespokoit' 2.1* is used, e.g.:

- (27) Emu kazalos', čto delo, po kotoromu on prišël, nastol'ko neznačitel'no, čto iz-za nego sovestno **bespokoit'** takogo vidnogo sedogo graždanina, kakim byl Ippolit Matveevič.

'He felt his business was so insignificant that it was shameful **to disturb** such a distinguished-looking grey-haired citizen as Vorobyandinov.' [Ilya Ilf, Evgeny Petrov. *The Twelve Chairs* (John Richardson, 1961)]

- (28) ...ja bojalsja tebjja **bespokoit'**.

'...I didn't dare **to disturb** you.' [Aleksandr Kuprin. *Olesya* (Stepan Apresyan, 1982)]

- (29) Prostit, – skazala Sof'ja Aleksandrovna, – očen' neudobno vas *bespokoit'*, no sročnoe delo... Vy ne mogli by vyzvat' k telefonu Polinu Aleksandrovnu?

"'I'm very sorry **to disturb** you," Sofya Alexandrovna said, "but it's urgent... I wonder if you could please ask Polina Alexandrovna to come to the phone?" [Anatoly Rybakov. *Children of the Arbat* (Harold Shukman, 1989)]

That said, it is possible that intuitions differ among native speakers. In some texts included in the Russian-English subcorpus of the RNC, the verb *trevožit' 2* is used, for example:

- (30) Ja posetitelja takogo, kak vy, bez sekretarja doložit' ne mogu, a k tomu že i sami, osoblivo daveča, zakazali ix ne **trevožit'** ni dlja kogo, poka tam polkovnik, a Gavrila Ardalionyč bez doklada idët.

'No, no! I can't announce a visitor like yourself without the secretary. Besides the general said he was not to be **disturbed** — he is with the Colonel C—. Gavrila Ardalionovitch goes in without announcing.' [Fedor Dostoevsky. *The Idiot* (Eva Martin, 1915)]

But there is a situation where replacing *bespokoit' 2.1* with *trevožit' 2* is entirely impossible. This concerns the use of *bespokoit' 2.1* as a formula of politeness

when addressing someone, especially on the telephone; this often carries a shade of servility (cf. Apresjan, 2004; 2014)., e.g.:

- (31) No Dubenkov okazalsja uslužlivym i dobrodušnym čelovekom. On zvonil Štrumu po telefonu i govoril: — **Bespokoit** Dubenkov. Ja ne pomešal, Viktor Pavlovič?

‘In fact, and Dubyonkov was very decent obliging. When he phoned Viktor, he said: “Dubyonkov **speaking**. I hope I’m not disturbing you, Viktor Pavlovich?” [Vasily Grossman. *Life and Fate*, Part 3 (Robert Chandler, 1985)]

At the same time, the perfective correlate of *trevožit’ 2*, namely, *potrevožit’*, is quite acceptable, e.g.:

- (32) — My **potrevožili** vas, madam Duncil’, — otnessja k dame konferans’e, — vot po kakomu povodu: my xoteli vas sprosit’, est’ li eščë u vašego suprugaja valjuta?

‘We **have troubled** you, Madame Dunchil,’ the master of ceremonies adverted to the lady, ‘with regard to the following: we wanted to ask you, does your husband have any more currency?’ [Mikhail Bulgakov. *The Master and Margarita* (Richard Pevear, Larissa Volokhonsky, 1979)]

Finally, one can turn to the nouns *bespokojstvo* and *trevoga*. Both nouns are polysemous. According to Apresjan (2014), the word *bespokojstvo* has two meanings: *bespokojstvo 1* and *bespokojstvo 2*, corresponding to *bespokoit’ 1* and *bespokoit’ 2.1*, respectively. The word *trevoga* has a meaning corresponding to the verb *trevožit’ 1* (in Apresjan, 2004, this is labeled *trevoga 1*), but it lacks a meaning corresponding to *trevožit’ 2*. However, *trevoga* has a distinct meaning realized in expressions like *vozdušnaja trevoga* ‘air alert’, *požarnaja trevoga* ‘fire alarm’, and idioms such as *bit’ trevogu* ‘to sound the alarm.’ This polysemy must be kept in mind when analyzing the words *bespokojstvo* and *trevoga* in the Russian-English subcorpus and their translation equivalents.

Both *bespokojstvo* in its two lexical meanings and *trevoga 1* exhibit a wide range of translation equivalents. This is already apparent from the translation of typical phrases involving these words. For instance, the phrase *sil’noe bespokojstvo* (meaning *bespokojstvo 1*) occurs four times in the RNC and has four different translations: *intense uneasiness*, *great restlessness*, *strong anxiety*, and, within the expression *v sil’nom bespokojstve*, it is rendered as *much agitated*.

With regard to these words, translators generally follow a context-sensitive strategy⁸. Even Pevear and Volokhonsky, in translating *trevoga* and *bespokojstvo* in *The Master and Margarita*, consistently rely on a contextual meaning approach, although they usually follow a concordance-based strategy when translating key words of a literary work. For example, the word *toska* in *Crime and Punishment* and in *The Master and Margarita* is regularly rendered as *anguish*, despite the frequent semantic shift this entails (see Shmelev, 2020, pp. 663–664). Thus, the word *trevoga*, which appears 14 times in the novel (specifically in reference to an internal state), is translated sometimes as *alarm* (more often), sometimes as *anxiety*. The word *bespokojstvo* (seven occurrences) is also translated in various ways. Also, it is worth considering their translations of the highly characteristic phrase *vpast' v bespokojstvo* from *The Master and Margarita*:

- (a) *Ivan vpal v bespokojstvo* – *Ivan became anxious*
- (b) *A vy ne vpadete v bespokojstvo?* – *You're not going to get upset?*
- (c) *Ivan ne vpal v bespokojstvo, kak i obeščal...* – *Ivan did not get upset, as he had promised...*
- (d) *...neizvestno otčego vpal v tosku i bespokojstvo* – *...for some unknown reason, lapsed into anxiety and uneasiness*
- (e) *Ivanuška vpal v bespokojstvo* – *Ivanushka fell into anxiety*

4. English-Russian subcorpus

Generally speaking, the analysis of the occurrence of the words under investigation in translated texts is often even a more effective research tool than the analysis of the translation equivalents of these words. The translation of language-specific words in the original is a problem that the translator, as a rule, reflects on, and thus the translator's individual preferences play a great role in the solution of such problems. Moreover, the translator's grounds for choosing a particular solution “may be superficial understanding, the influence of bilingual dictionaries, or even the desire to give the translated text a veneer of ‘foreignness’ to make the reader perceive the text as a translation” (Mixajlov, 2005, p. 381). Thus, when the original language is Russian, the parallel corpus sometimes gives us insufficient informa-

⁸ The terms *contextual meaning strategy* and *context-sensitive strategy* are treated as synonymous throughout this paper.

tion about the semantics of Russian lexical units. In contrast, the appearance of linguistic units in a translation most often results from an unconscious decision of the translator that reflects his spontaneous linguistic activity. In other words, translation into Russian is close to everyday linguistic activity of speakers of Russian. An occurrence of a given Russian expression in translation, more often than not, reflects a 'naïve' choice of words. By studying why a translator uses a given word, we are often able to uncover some of the latter's semantic characteristics that went unnoticed during the analysis of original texts.

Indeed, not all material in the English-Russian subcorpus of the RNC is relevant for such analysis. When a translator follows the concordance strategy, the choice of the Russian word is determined by the usage in the original rather than by the semantic features of the Russian words. For example, the subcorpus includes Lemony Snicket's *The Ersatz Elevator*. In this novel, considerable attention is given to metalinguistic reflection on the word *anxious* and its semantic differences from *nervous*. The importance of the word *anxious* in the text compels the translator to follow the concordance strategy and to consistently choose the same word — *trevoga* or its derivatives — to translate *anxious* (often contrasted with *nervous*). Apparently, *trevoga* and its derivatives are perceived by the translator (and likely by many other Russian native speakers) as the closest equivalents to English *anxious*. Abandoning the concordance strategy in favor of a contextual semantic one would have been impossible here, as it would completely undermine the basis for the metalinguistic reflection that plays an important role in the novel.

The verb *bespokoit'sja* appears 1,060 times in the English-Russian subcorpus of the RNC, which is about the same frequency as in the Russian-English subcorpus. In exactly half the examples (530 occurrences), it is used as a translation of the English verb *worry*. This confirms the hypothesis that *bespokoit'sja* is associated with the subject's thoughts that something bad may happen.

In approximately half the cases, the verb directly follows the negative particle *ne* or is separated from it by a few words. This supports the hypothesis of a rudimentary negative polarity in the verb *bespokoit'sja*.

The English-Russian subcorpus data show that the description in Apresjan (2014), according to which only *bespokoit'sja* 2 (in which the prepositional phrase with *o* marks the object of effort) can govern a prepositional-case group with *o*, and not *bespokoit'sja* 1, is not entirely accurate.⁹ There are quite a few

⁹ Such government was actually noted for the verb *bespokoit'sja* 1 in (Apresjan, 2004, p. 27): "All three synonyms [*bespokoit'sja* 1, *trevožit'sja*, *volnovat'sja* 1] in their imperfective forms also govern prepositional phrases of the type *o kom-libo* or *o čem-libo* with a similar meaning; cf. — *Varen'ka*, *Varen'ka*, — *laskovo zagovorila Sof'ja Aleksandrovna*, — *obo mne bespokoiš'sja, dobraja*

examples in the subcorpus where it is specifically *bespokoit'sja* **1** (as seen from the English original) that governs a prepositional phrase with *o*, and in these cases, the phrase indicates the cause of the emotional state. Compare, e.g.:

- (33) And then he began **to worry about the time**.

‘No zatem on načal **bespokoit'sja o tom, skol'ko vremeni prošlo**.’ [Clive Staples Lewis. The Chronicles of Narnia. The Voyage of the ‘Dawn Treader’ (G. A. Ostrovskaja, 1991)]

- (34) “I hope he’s all right,” said Christopher Robin. “**I’ve been wondering about him**. I expect Piglet’s with him...”

‘— Ja nadejus’, on živ i zdrav, — skazal Kristofer Robin. — **Ja nemnogo bespokojus’ o nĕm**. Interesno, Pjatačok s nim ili net?’ [A. A. Milne. Winnie-the-Pooh (B. Zaxoder, 1960)]

- (35) I **am a little worried about** the extra expense of this, but if you do not think it justified for the sake of safety, let me know and I will alter my arrangements accordingly.

‘Ja **nemnožko bespokojus’ o** dopolnitel’nyx rasxodax, no esli vy ne sočtete, čto oni opravdany soobraženijami predostorožnosti, dajte mne znat’, i ja izmenju svoi plany.’ [Dorothy L. Sayers. Strong Poison (M. M. Lanina, 2007)]

- (36) I **am getting quite uneasy about** him, though why I should I do not know, but I do wish that he would write, if it were only a single line.

‘Očen’ **bespokojus’ o** nĕm, xotja, sobstvenno, ne znaju, počemu: xorošo bylo by, esli by on napisal xot’ odnu stročku.’ [Bram Stoker. Dracula (N. Sardrova, 1912)]

- (37) I **had worried about** how I’d go about undressing, how, if he wasn’t going to help, I’d do what so many girls did in the movies, take off my shirt, drop my pants, and just stand there, stark-naked, arms hanging down, meaning: This is who I am, this is how I’m made, here, take me, I’m yours. But his move had solved the problem.

ty duša. Obo mne bespokoit'sja ne nado. ‘– Varenka, Varenka,’ Sofya Alexandrovna began tenderly, ‘you’re worrying about me, you kind soul. There’s no need to worry about me.’ (A. Rybakov. Children of the Arbat).”

- ‘Ja **bespokoilsja o tom**, kak budu razdevat’sja, esli on ne pomožet, kak budu delat’ to, čto stol’ko devušek delali v kino, snimu futbolku, stjanu trusy, i budu stojat’ tam, v čem mat’ rodila, opustiv ruki po bokam, kak by govorja: Èto ja, tak ja sozdan, vot, beri menja, ja tvoj. No ego dejstvija rešili problemu.’ [André Aciman. Call me by your name (V. P. Dimirova, 2019)]
- (38) The MGB reported the anxieties of the Jews in Soviet Ukraine, who understood that the policy must come from the top, and **worried that** “no one can say what form this is going to take.”
 ‘MGB raportoval o straxax evreev v Sovetskoj Ukraine, kotorye ponimali, čto ukazaniya, dolžno byt’, spuščeny sverxu, i **bespokoilis’ o tom, čto** “nikto ne znaet, kakuju formu èto primet”. [Timothy D. Snyder. Bloodlands. Europe between Hitler and Stalin (L. Zurnadži, 2015)]
- (39) He says they **were all very dubious about** what sort of reception they would get.
 ‘Džordž Xarrison rasskazyvaet, čto rebjata **očen’ bespokoilis’ o tom**, kak ix vstretjat.’ [Davies Hunter. The Beatles (A. Gryzunova, 2017)]
- (40) Despite putting a brave face on things for everyone else’s benefit at Garden Lodge, privately I **began to get very anxious about** my own health. I thought I could be HIV positive as well.
 ‘Pered obitateľjami Garden Lodž ja xrabrilsja, no v glubine duši **bespokoilsja o** sobstvennom zdorov’e. Ja podozreval, čto tože mogu byt’ VIČ-položitel’nym. [Jim Hutton. Mercury and Me (V. Gerstle, 2019)]
- (41) Wilson’s **thoughts were with** the little girl rushed to emergency in critical condition.
 ‘Uilson **bespokoilas’ o** malen’koj devočke vsego devjati let, ktoruju uvezli na skoroj pomošči v kritičeskom sostojanii... [Rosie Dimanno. ‘There’s a man with a gun!’ (Inosmi.ru, 2018)]

There are also numerous examples of the causative verb *obespokoit’*, most often in the participial form *obespokoennyj*, 229 in total. In fewer than half (around a third — 75), the direct stimulus was the English word *worry*, but *worry* remains the most frequent stimulus.

Of note is the fact that the verb *pobespokoit’sja* (semantically related to *bespokoit’sja* 2) appeared 14 times in the English-Russian subcorpus (compared

to only three times in the Russian-English subcorpus). The examples correspond well to the meaning of *bespokoit'sja* 2:

- (42) If one gentleman under a cloud **is not to put himself a little out of the way** to assist another gentleman in the same condition, what's human nature?

‘Esli odin džentl'men, popavšij v bedu, **ne pobespokoitsja nemnogo**, čto-by pomoč' drugomu džentl'menu, naxodjaščemusja v takom že položenii, čego stoit čelovečeskaja priroda!’ [Charles Dickens. The Posthumous Papers of the Pickwick Club (A. V. Krivcova, E. L. Lann, 1933)]

- (43) With a shock he realized that this was the smoke from his little cooking-fire, which he had **neglected** to put out.

‘S ispugom on ponjal, čto èto dym ot ego sobstvennogo malen'kogo kostra, kotoryj on **ne pobespokoilsja** pogasit’.’ [J. R. R. Tolkien. The Lord of the Rings: The Two Towers (M. Kamenkovič, V. Karrik, 1994)]

- (44) Obviously, Father Abbot **was seeing to it** that they all cooled their heels.

‘Očevidno, otec abbat **pobespokoilsja** o tom, čtoby oni oxladili svoj pyl.’ [Walter M. Miller, Jr. A Canticle For Leibowitz (S. Borisov, 1999)]

This raises a question: why such a disproportion (14 examples of *pobespokoit'sja* in the English-Russian subcorpus vs. only three in the Russian-English) exist? Certainly, the English-Russian subcorpus is larger, but by less than a factor of two. Do translators use this verb significantly more often than authors who write originally in Russian?

One might hypothesize that *pobespokoit'sja* is a relatively new verb (as the Russian originals in the Russian-English subcorpus are significantly older than the translated texts in the English-Russian subcorpus). But the Main corpus of RNC shows that its frequency has grown only slightly since the 19th century. Thus, the reason for the discrepancy remains unclear.

The verb *trevožit'sja* occurs far less frequently in both subcorpora compared to *bespokoit'sja*. The stimuli for its appearance are generally the same (*worry* being the most frequent). Yet it is interesting to examine in which cases translators prefer this verb as an equivalent.

As mentioned, *bespokoit'sja* is chosen as a translation of *worry* 530 times, while *trevožit'sja* appears 123 times. In addition, there are 121 examples with the causative verb *bespokoit'*, 103 with *bespokojstvo*, 75 with *obespokoit'* (usually in the form *obespokoennyj*), 20 with *zabespokoit'sja*, 10 with *obespokoit'sja*, and

isolated cases with *pobespokoit'sja* and *pobespokoit'*. Also: 101 examples with *trevoga*, 61 with *trevožit'*, 57 with *vstrevožit'* (including *vstrevožennyj*), and 13 with *vstrevožit'sja*. Altogether, translations of *worry* using words from the *bespokoit'* and *trevoga* word families account for almost half of all translation variants, with a clear predominance of the *bespokoit'* family. This likely reflects the fact that these words contain an element of intellectual evaluation of the situation, which better conveys the specific meaning of *worry*, implying reflection and concern about a possible negative outcome.

For the words *anxious*, *anxiously*, and *anxiety*, the situation is different. Here, although the difference is small, the *trevoga* family predominates. This is likely due to the fact that the English words emphasize a lack of important knowledge about a situation, which gives rise to an intense emotional experience.

For the word *alarm*, the predominance of the *trevoga* family is more significant. Notable translation equivalents include *vstrevožit'* (including *vstrevožennyj*) and *vstrevožit'sja*, likely reflecting the idea of a sudden realization of impending danger.

Overall, the analysis of English-to-Russian translations shows that *bespokoit'(sja)* and *trevožit'(sja)* are close synonyms, and the choice between them often depends on the translator's individual preference. It can be observed that in describing internal states, *bespokoit'sja* is used significantly more often than *trevožit'sja*, while in describing the cause of such states, the causative verb *trevožit'* is used nearly as often as *bespokoit'*. Observations from Apresjan (2004) are largely confirmed: *bespokoit'sja* involves a greater degree of intellectual evaluation, while *trevožit'sja* points to a more intense and less uniform emotional experience. At the same time, data from the parallel subcorpus of the RNC were insufficient to verify whether external signs mentioned by Apresjan such as uncontrolled facial expressions and small motor movements are indeed characteristic of *trevožit'sja*.

The difference between *bespokoit'(sja)* and *trevožit'(sja)* is especially striking in contexts involving negation: *ne bespokoit'sja* may indicate indifference, while *ne trevožit'sja* indicates the absence of anxiety. It is no coincidence that expressions like *Don't worry*, *Never mind*, etc., are often translated into Russian as *ne bespokoj'sja* or *ne bespokojtes'*. This explains why *ne bespokoj'sja* and *ne bespokojtes'* occur significantly more often (267 instances) than *ne trevož'sja* and *ne trevož'tes'* (39 instances) in the English-Russian subcorpus.

5. Conclusions

The use of data from the Russian-English and English-Russian subcorpora of the RNC made it possible to test the descriptions of words from the *bespokoit'* and *trevoga* word families as presented in the works of Apresjan. A significant portion of the observations made in those works was confirmed. At the same time, it was possible to supplement and refine these descriptions. In particular, the rudimentary negative polarity of *bespokoit'sja* was identified. Possible syntactic patterns were clarified. Overall, the analysis demonstrated the usefulness of parallel corpus data in lexicographic portraying.

6. Cited references

- Apresjan, Ju. (2004). BESPOKOIT' 1, TREVOŽIT' 1; BESPOKOIT' 2, TREVOŽIT' 2; BESPOKOIT'SJA 1, TREVOŽIT'SJA, VOLNOVAT'SJA 1. In Ju. Apresjan (Ed.) *Novyj ob''jasnitel'nyj slovar' sinonimov russkogo jazyka*. Jazyki slavjanskoj kul'tury; Wiener Slawistischer Almanach, Moskva, Vena.
- Apresjan, Ju. (2014). BESPOKOIT'; BESPOKOIT'SJA. In Ju. Apresjan (Ed.) *Aktivnyj slovar' russkogo jazyka*. Jazyki slavjanskoj kul'tury, Moskva.
- Beekman, J., & Callow, J. (1989). *Translating the Word of God*. Zondervan Publishing House.
- Mixajlov, M. [= Mihailov, M.] (2005). Častica i celoe: k voprosu o poiske sootvetstvij služebnyx slov v parallel'nom korpuse tekstov. *Computational Linguistics and Intellectual Technologies*, 377–381.
- Nida, E., & Taber C. (1969). *The Theory and Practice of Translation, With Special Reference to Bible Translating*. Brill, Leiden.
- Shmelev, A. (2020). Language-Specific Words in the Light of Translation: The Russian *Toska*. *Computational Linguistics and Intellectual Technologies*, 19 (26), 658–669.
- Wierzbicka, A. (1999). *Emotions across Languages and Cultures*. Cambridge University Press; Éditions de la Maison des Sciences de l'Homme, Cambridge, Paris.
- Zaliznjak, Anna, & Mikaëljan, I. (2023). TREVOGA, OPASENIE i STRAX v russkom jazyke: popytka razgraničeniya blizkix značenij. In S. Pereverzeva, I. Itkin & L. Xesed (Eds.), *Grani estestvennogo jazyka i kinesiki. Sbornik statej k 75-letiju G. E. Krejdlina* (p. 120–132). Diskurs, Moskva.



Magdalena Danielewiczowa

University of Warsaw
Poland

<https://orcid.org/0000-0003-1371-1867>

Unqualifiable meta-predicative adverbials in Polish*

Abstract

The article aims to highlight specific features that enable the identification of a certain class of Polish adverbial expressions ending in -o/-e, which do not belong (unlike regular adverbs) to the object level of language, nor represent metatext in the manner that particles or sentential adverbs do. The paper thus discusses units such as *iście* [królewski] ‘truly royal’, *dosłownie* [idiotyczny] ‘literally idiotic’, *jawnie* [wrogi] ‘openly hostile’, etc. Although based on linguistic data from Polish, research hypotheses are more universal. The author introduces a set of tests to ensure precision in distinguishing unqualifiable meta-predicative adverbials from those units that are superficially similar. The following parameters are taken into account: (i) prosodic features, especially the relationship of a word and the sentence (non-contrasting) stress; (ii) position in the linear word order; (iii) possibility of being negated; (iv) capacity to combine with common adverbs in coordinate constructions; (v) capacity to combine with intensifiers, weakeners, or any other predicates in subordinate constructions; (vi) relationship to the grammatical category of the degree; (vii) range of syntactic and semantic combinability; (viii) ability to be used in a context with an ordinary adverb of the same shape. An important aim of this work is the examination of (a) relationships between the epistemic character of the expressions in question, (b) their belonging to the metalevel of language, and (c) their place in the theme-rheme structure of potential utterances where they may be functioning.

Keywords

semantics, adverbials, adverbs, metatext, meta-predicates

* The ideas suggested in this work were previously discussed in the presence of, among others, J. D. Apresjan and L. L. Iomdin at one of the meetings of the Commission on the Grammatical Structure of Slavic Languages, and then presented at the conference of the Slavic Linguistics Society in Tokyo 2022.

I. Preliminary remarks

As is well known, the class of expressions commonly referred to as adverbs is remarkably heterogeneous. Despite advanced research in terms of general linguistic theory or in the framework of descriptions of individual languages (see e.g. Greenbaum 1969, Bartsch 1972, Thomason & Stalnaker 1973, Grzegorzczkowska 1975, Mørdrup 1976, Nølle 1990, 1990a, Molinier 1990, Cinque 1999, Apresjan 2003, adverbial entries, Nilsen 2004, McNally & Kennedy 2008, Morzycki 2008, Maienborn & Schäfer 2012, Boguslavskaya & Boguslavsky 2014), this class still hides some unexplored aspects.

It will be claimed that a special group of adverbial expressions should be distinguished from similar, but not identical, units. We shall refer to them by calling them “unqualifiable meta-predicative adverbials” (henceforth UMA). There is little doubt that one encounters a universal aspect of language here, although, in this article, we shall focus on the Polish data from the National Corpus of Polish¹.

The scope of the data to be investigated covers exclusively the adverbials ending in *e* or *o*, which are semantically related to the corresponding adjectives (in some cases merely by origin). There are at least three groups of expressions determined as above. One of these groups includes the most common adverbs like *szybko* ‘quickly’, *zrećźnie* ‘deftly’, or *wesoło* ‘cheerfully’. The other group is heterogeneous and it contains words such as *oczywiście* ‘of course’, *przeważnie* ‘mostly’, *właściwie* ‘actually’, *mianowicie* ‘namely’, or *wyłącznie* ‘exclusively’. They bear only a superficial similarity to what we take to be the most usual adverbs. In contemporary Polish, these words play different roles in utterances so, depending on the adopted research perspective, they are termed differently as “sentential adverbs” (Jackendoff 1972, Martin 1973, 1974, Bellert 1977, Lang 1979), “speaker-oriented adverbs” (Thomas 2009) “particles”, “modal operators”, “discourse words”, “metatextual comments” (Ducrot 1980: 37–39, Nølle 1990ab, Wajszczuk 1997, 2005), etc. Beyond these more or less salient subclasses of adverbials, some more special functional groups of expressions have to be isolated. Several such groups can be pointed out, e.g.:

- relativizing adverbs (domain adverbs) like *formalnie niedopracowany* ‘formally underdeveloped’, *teoretycznie poprawny* ‘theoretically correct’, *praktycznie nie do użycia* ‘practically not usable’ (Běličová-Křížková 1980, Danielewiczowa 2021, 2021a);

¹ See: <https://nkjp.pl/>.

- adverbs locating states of affairs in time and space, e.g. *następnie* ‘then’, *ostatecznie* ‘finally’, *później* ‘later’, *blisko* ‘close to’, *daleko od* ‘far from’;
- a microsystem of words which point to some contrast in sentences, e.g., *przeciwnie* ‘on the contrary’, *odwrotnie* ‘vice versa’, *inaczej* ‘differently’, *odpowiednio* ‘respectively’;
- adverbs that are a part of the speaker’s metatextual glosses about his/her main text, e.g.: *paradoksalnie* ‘paradoxically’, *ogólnie* ‘generally’, *krótko* <mówiąc> ‘briefly’, *konkretnie* ‘specifically’, and many others.

II. UMA as a special semantic-syntactic category.

Introduction

None of the groups listed above will become the target object of our direct interest in this work. As there exists a special category of adverbials, we shall try to characterize them by contrasting this group of words with the adverbs indicated in our preliminary remarks. The group’s name “unqualifiable meta-predicative adverbials” reflects the two consecutive traits of the adverbials, meta-predicativity and unqualifiability.

To begin with, let us make some observations on a paradigmatic example with the contrast to be elucidated. This example is represented by a pair of sentences in which contrastive adverbial items are present. Each of them appears in the same diacritic form. The functional difference between the two occurrences is also readily visible due to the corresponding English translation. In this example, the UMA occurrence and the non-UMA occurrence are in the following order: the UMA occurrence is placed on the right-hand side of the notation, and the non-UMA occurrence is placed on the left-hand side of it.

- (1) *Zdobywcy z podbitą ludnością rozprawili się bezwzględnie. Vs Miałeś bezwzględnie rację.*

‘The conquerors suppressed the aborigines ruthlessly.’ Vs ‘You were absolutely right.’

Let us dwell on the paradigmatic example for a moment to show several traits of how our UMA operators function. Notice that *bezwzględnie* in the first sentence of example (1) can remain under the sentence stress, the second is inherently unstressed. We can put only a corrective or emphatic accent on it, but this

kind of accent may be placed, as is well known, on any part of the sentence, so it is irrelevant. Moreover, the first *bezwzględnie* in (1) can be negated: *nie bezwzględnie, lecz miłosiernie* ‘not ruthlessly, but mercifully’, while the second, as inherently unstressed, thus non-rhematic, cannot be denied:

- (2) * *Miałeś nie bezwzględnie rację.*
 ‘*You were not absolutely right.’

It is also worth noting that the unstressed *bezwzględnie*, unlike the stressed one, cannot take the last position in a sentence, usually reserved for rhematic words.

There is also another significant difference between stressed and unstressed *bezwzględnie*: the former combines with modifiers, e.g., *dość bezwzględnie* ‘quite ruthlessly’, *nie dość bezwzględnie* ‘not ruthlessly enough’, while the latter does not. It is unqualifiable, that is to say, it undergoes no further characterization; cf.

- (3) *Zdobywcy z podbitą ludnością rozprawili się raczej bezwzględnie.* Vs **Miałeś raczej bezwzględnie rację.*
 ‘The conquerors suppressed the aborigines rather ruthlessly.’ Vs ‘*You were rather absolutely right.’

It should be mentioned, moreover, that *bezwzględnie* in the first sentence from example (1) is gradable:

- (4) *Zdobywcy z podbitą ludnością rozprawiali się coraz bezwzględniej.*
 ‘The conquerors suppressed the aborigines more and more ruthlessly.’

The UMA *bezwzględnie* in the second sentence from example (1) cannot be graduated:

- (5) * *Miałeś bezwzględniej / najbezwzględniej rację.*
 ‘*You were more absolutely / the most absolutely right.’

Summarizing this part of our study, let us repeat that stressed *bezwzględnie* represents ordinary adverbs of manner, while the unstressed is an example of the UMA.

III. A survey of selected UMA operators

The analogical opposition between two different adverbials with a similar form can also be observed in many other Polish sentences. Consider the following examples in which the first sentence contains an adverb of manner, while the second, a corresponding UMA:

- (6) *Uczeń musi ubierać się czysto. Vs To są rozważania czysto teoretyczne.*
'The student must dress cleanly.' Vs 'These are purely theoretical considerations.'
- (7) *Wypowiadałeś się zbyt dosłownie. Vs Kiedy przemawia publicznie, czuje się dosłownie jak ryba wyjęta z wody.*
'You were speaking too literally.' Vs 'When speaking publicly, he literally feels like a fish out of water.'
- (8) *Twój syn zachowuje się bardzo dziwnie. Vs Odczuwanie dźwięku w tych dziwnie głuchych, wąskich i zamkniętych przestrzeniach nieuchronnie prowadzi do skojarzeń z klaustrofobiczną trumną.*
'Your son behaves very strangely.' Vs 'The sensation of sound in these strangely deaf, narrow, and closed spaces leads inevitably to associations with a claustrophobic coffin.'
- (10) *Musimy działać bardziej zdecydowanie. Vs Niska aktywność uniwersytetów i centrów badawczych wpływa na zdecydowanie słabe efekty.*
'We need to act more decisively / firmly.' Vs 'Low activity of universities and science-research centers decisively / undoubtedly is responsible for poor effects.'
- (11) *Ta aktorka była ubrana nie całkiem kompletnie. Vs Była za to kompletnie pijana.*
'This actress was not fully dressed.' Vs 'But she was completely drunk.'
- (13) *Będziemy głosować jawnie, przez podniesienie ręki. Vs Kiedy te reakcje są szczególnie intensywne, pacjent może stać się przygnębiony, wycofany lub jawnie wyzywający.*
'We shall vote openly, by a show of hands.' Vs 'When these reactions are particularly intense, the patient can become morose, withdrawn, or openly defiant.'

- (14) *Wycieczka zakończyła się szczęśliwie. Vs Uszkodzenie było szczęśliwie niewielkie.*
 ‘The trip ended happily.’ Vs ‘The damage was, luckily, slight.’
- (15) *Musisz ćwiczyć bardziej regularnie / regularniej. Vs On z nas regularnie zakpił.*
 ‘You need to exercise more regularly.’ Vs ‘He openly mocked us.’
- (16) *Dziecko spało spokojnie. Vs W pudełku zmieści się spokojnie dwadzieścia kredek.*
 ‘The child was sleeping calmly.’ Vs ‘Twenty crayons easily fit in the box.’
- (17) *Zachowywał się całkiem zwyczajnie. Vs Ten chłopak jest zwyczajnie nieodpowiedzialny.*
 ‘He behaved quite normally.’ Vs ‘This boy is just plain irresponsible.’

Questions arise about the linguistic status of expressions represented in the pairs of sentences above: a) are we dealing in this case with two applications of the same monosemic word?; b) are they different meanings of a polysemous word?; c) or are they rather two independent units of language? Everything indicates that the last of these questions should be answered positively. It is, namely, about separate linguistic entities, and not only about a change of context, associated with a change in the shade of meaning, which does not cause the loss of the identity of the expression in question. For the present moment, let us notice that in most of the examples above, two different English words are necessary to cover different meanings, while in Polish, different meanings are covered by very similar (though not identical, as we shall see later) forms.

IV. UMA operators in their contrast with other adverbials. A systematic account

In this section, the central topic of the paper will be elaborated on in detail. What makes it possible to clearly distinguish expressions such as those presented in the second sentence of each of the examples (1, 6–17) from the corresponding adverbs of manner is the fact of their divergent characteristics obtained by using the following criteria (i–viii):

- (i) prosodic features, especially the relationship of a word and the sentence (non-corrective) stress;
- (ii) position in the linear word order;
- (iii) possibility of being negated;
- (iv) compatibility with intensifiers, weakeners, or other predicates in subordinate constructions;
- (v) ability to be concatenated with common adverbs in coordinate constructions;
- (vi) relationship to the grammatical category of the degrees of comparison;
- (vii) range of syntactic and semantic combinability;
- (viii) ability to be used in a context with an ordinary adverb of the same shape.

The abovementioned criteria have already been partially applied to the example (1) analysis. They will be discussed more systematically, with an eye on other examples later in the article.

(i) Prosodic features: stressed vs inherently unstressed units

UMA operators are inherently unstressed, which means that the non-corrective stress cannot rest on them. On the contrary, ordinary adverbs are either currently stressed or unstressed. (They could be stressed or unstressed if only the components of the content of the utterance in which they are used were organized differently.) See the following example where stressed words are capitalized and an asterisk indicates an unacceptable accent:

- (18) *Zachowywał się całkiem ZWYCZAJNIE.* Vs *Ten chłopak jest zwyczajnie NIE-ODPOWIEDZIALNY.* Vs **Ten chłopak jest ZWYCZAJNIE nieodpowiedzialny.*
 ‘He behaved quite NORMALLY.’ Vs ‘This boy is plain IRRESPONSIBLE.’
 Vs ‘This boy is just PLAIN irresponsible.’

The situations when UMAs are isolated as independent utterances are not counterarguments because the phrasal stress accompanies them automatically:

- (19) – *Czy on jest odpowiedzialny za ten wypadek?* – *ZDECYDOWANIE.*
 (20) ‘– Is he responsible for this accident? – DEFINITELY.’

(ii) Different positions in the linear word order

The UMA usually precedes the word for which it opens a valency position; it is utterly impossible to move it to the final position in the sentence. This place is normally stressed when other prosodic factors do not undermine a given word order. In examples (21, 22), incorrect sentences have been marked with asterisks:

- (21) *Twój syn zachowuje się bardzo DZIWNIE.* Vs *Twój syn był dziwnie SPOKOJNY.*
Vs **Twój syn był spokojny DZIWNIE.*

‘Your son behaves very strangely.’ Vs ‘Your son was strangely calm.’ Vs ‘*Your son was calm strangely.’

- (22) *Będziemy głosować przez podniesienie ręki, to znaczy JAWNIE.* Vs *Była to propozycja jawnie KORUPCYJNA.* Vs **Była to propozycja korupcyjna JAWNIE.*

‘We will vote by a show of hands, in other words, openly.’ Vs ‘It was an openly corrupt proposal.’ Vs ‘*It was a proposal corrupt openly.’

In this regard, UMA operators differ, among others, from adverbial intensifiers, which in Polish can be moved to the last, stressed position:

- (24) *Zależy mu na tym BARDZO.*

‘He cares about it very much.’

- (25) *Mądra jesteś NIESŁYCHANIE.*

‘You are wise, extremely.’

(iii) Possibility vs impossibility of being negated

Since UMA operators do not take sentence stress, they cannot, unlike adverbs of manner, function as rhemes; accordingly, they cannot be negated; cf. e.g.:

- (26) *Piotr przegrał, bo działał niezdecydowanie.* Vs **Piotr był niezdecydowanie szybszy.*

‘Peter lost because he acted indecisively / irresolutely.’ Vs ‘*Peter was irresolutely faster.’

- (27) *Dziecko spało niespokojnie.* Vs **W pudełku zmieści się niespokojnie dwadzieścia kredek.*

‘The child was sleeping uneasily.’ Vs ‘*Twenty crayons will fit uneasily in the box.’

It is noteworthy that the form *nieszczęśliwie*, which is a negative counterpart of the adverb *szczęśliwie*, also covers another UMA, lexically independent of the UMA *szczęśliwie*, e.g.:

- (28) *Zakochał się szczęśliwie.* *Zakochał się nieszczęśliwie.*

‘He fell in love happily.’ ‘He fell in love unhappily.’

- (29) *Szczęśliwie zakochał się w spokojnej dziewczynie.*

‘Luckily, he fell in love with a quiet girl.’

- (30) *Nieszczęśliwie zakochał się w kłótniwej dziewczynie.*

‘Unluckily, he fell in love with a quarrelsome girl.’

(iv) Compatibility vs incompatibility with intensifiers, weakeners, or any other predicates in subordinate constructions

In contrast to adverbs of manner, UMA cannot subordinate intensifiers, weakeners, or any other predicates. Combinations such as *działać na wpół jawnie* ‘act semi-overt’ or *formułować zarzuty bardziej zdecydowanie* ‘make accusations more firmly’ are correct, while **na wpół jawnie korupcyjny* ‘*semi-overt corrupt’, **był bardziej zdecydowanie pierwszy* ‘*he was more firmly first’ are totally unacceptable. It is a very important, definitional feature of the UMA class.

(v) Ability vs disability to be concatenated with common adverbs in coordinate constructions

UMA cannot be combined with regular adverbs in coordinate constructions. Other words can be attached to them (something which is nevertheless a sporadic phenomenon), this is conditional on units belonging to the metalevel of the language; cf. *silnik pracuje spokojnie i czysto* ‘the engine runs quietly and cleanly’ vs **w pudełku ładnie i spokojnie mieści się dwadzieścia kredek* ‘twenty crayons fit nicely and easily into the box’ vs *w pudełku spokojnie, a nieoczekiwanie mieści się dwadzieścia kredek* ‘twenty crayons fit easily, though unexpectedly, into the box’.

(vi) Relationship to the grammatical category of the degrees of comparison: gradable vs ungradable units

UMA, unlike adverbs of manner, are not subject to graduation, cf. e.g.:

- (31) *Ten student zachowuje się dziwnie // dziwniej niż tamten // najdziwniej ze wszystkich.* Vs *Mama coś dziwnie się spóźnia* / **dziwniej* / **najdziwniej się spóźnia*.

‘This student behaves strangely // more strangely than that one // the strangest of all.’ Vs ‘Mom is strangely late.’ // ‘Mom, a strange thing, is late.’ Vs ‘Mom is *more strangely / *the most strangely late.’

Even if some UMA operators take the extreme values, they do not apply to an ordinary superlative. There is always another UMA or a product of an operational unit behind this lexicalized form (Danielewiczowa 2022); see also (Grochowski 2008). Insofar as just a *positive* or *superlative*, but not a *comparative* form, it cannot play the role of a grammatical degree of comparison; cf.

- (32) *A rajska jabłoń spokojnie / najspokojniej / *spokojniej grzała w słońcu swoje owoce.*

‘And the apple tree of paradise was calmly warming its fruits in the sun.’

See also the example (33) where a lexicalized comparative or superlative is used:

- (33) *Lepiej / najlepiej / *dobrze nic nie mów.*
‘Better say nothing.’

and (34) where a special Polish operation is applied (see Danielewiczowa 2022):

- (34) *Spotkamy się najpóźniej za rok.*
‘We’ll meet next year at the latest.’

(vii) Range of syntactic and semantic combinability

The range of syntactic and semantic combinability of UMA differs from common adverbs, which combine primarily with verbs, as the name suggests. The units belonging to UMA combine with verbs, adjectives, adverbs, and even nouns,

but place severe semantic constraints on the accompanying expressions. Each UMA has its selective restrictions; therefore, just signaling the problem, let alone exhausting it, would take hours. A detailed semantic analysis of the selected UMA was presented in Danielewiczowa (2012).

Now, let us say only that, for example, *dosłownie* 'literally' dislikes "literal" contexts, which is paradoxical. You cannot say in Polish:

- (35) **Wczoraj dosłownie umyliśmy nasz samochód.*
'*Yesterday we literally washed our car.'

The word *dosłownie* chooses metaphorical or quotative phrases:

- (36) *Czuję się dosłownie zdruzgotany.*
'I feel literally devastated.'
- (37) *On jest dosłownie mną!*
'He is literally me!'

The UMA *jawnie* 'openly' first, negatively qualifies characteristics it precedes and, second, points to the accumulation of data as a factor enabling a given assertion; cf.

- (38) *Zachowywała się jawnie wrogo.* Vs **Zachowywała się jawnie przyjaźnie.*
'She behaved in an openly hostile way.' Vs 'She behaved openly friendly.'

The UMA *spokojnie* 'calmly', in turn, juxtaposes two alternative states of affairs, each of which is accompanied by a certain amount of knowledge on the one hand and some amount of ignorance on the other. Knowledge in favor of the negative component of the alternative is outweighed by data in favor of the positive one; it is therefore not a question of assessing, as is the case with the meaning of *jawnie* 'openly', but of weighing the factors that make up a certain result at stake.

(viii) Ability to be used in a context with an ordinary adverb of the same shape

The necessity of separating UMA from ordinary adverbs is also confirmed by the possibility of realizing non-pleonastic utterances in which expressions belonging to both lexical classes are used. The following examples show that in

the same Polish sentence, one can correctly, and not tautologically, use a regular adverb and an UMA having a similar form:

- (39) *Zdecydowanie musimy działać bardziej zdecydowanie.* ‘We decisively have to act more firmly.’
- (40) *Nasze dzieci szczęśliwie ułożyły sobie życie dość szczęśliwie.* ‘Luckily, our children have made their own lives quite happy.’
- (42) *Ta organizacja jawnie działa tylko na wpół jawnie.* ‘This organization openly acts only semi-overtly.’

V. Outstanding problems

Now, we can return to the question about the status of the words in the pairs represented in examples (1, 6–17) and, respectively, (39–42). Are we dealing with polysemy, homonymy, or independent units differing simultaneously in form and meaning?

Even if the boundaries between polysemy and homonymy are difficult to determine, we can certainly say that both relationships require the identity of the forms (polysemous or homonymous words have the same forms and different meanings). This basic definitional condition is not met in the case of the word class under consideration. Since natural language is primarily realized in a spoken form, the expression plane of a given lexical unit includes both segmental and suprasegmental features (Danielewiczowa 2011). For this reason, on the one hand, the adverb *jawnie*₁, which takes sentential stress, and, on the other hand, the inherently unstressed UMA *jawnie*₂ are two different elements of the Polish lexicon, in their form as well as in meaning. We shall argue that such a relationship occurs between the adverbials in all the other pairs of sentences represented in the examples shown above.

It should be clarified what exactly the term *unqualifiable meta-predicative adverbial* means. Adverbs of manner refer to relations in the world, so they belong to the object level of language (Bogusławski 2005). This fact has its formal expression in the possibility of being stressed, i.e., taking on a rhematic non-corrective sentence accent. This also amounts to remaining within the reach of the actual negation. Adverbs of manner meet those requirements. Moreover, they can be characterized by other predicates, e.g.:

(43) *Odpowiedział na pytania bardzo INTELIGENTNIE / bardzo NIEINTELIGENTNIE.*

‘You answered the questions very INTELLIGENTLY / very UNINTELLIGENTLY.’

Things are different with UMA. When uttered in Polish sentences, they cannot be further specified or modified by any expressions. Their peculiarity consists in introducing neither additional information, called „characteristics of manner”, nor metatextual comments related to the basic content of a given utterance (as a whole). They qualify a predicate (also in its nominalized form) to which they are immediately attached as valid in their current relation to the epistemic object being referred to (Grochowski 1986a, Bogusławski 2022). The qualification of the predicate consists of the speaker’s attention to certain features of the circumstances of the current application of this predicate. The meta qualification is always thematized. It is also worth emphasizing that from another point of view, UMA operators are, for the most part, words belonging to the broader class of epistemic units (Danielewiczowa 2002, Bogusławski 2007). They refer to epistemic circumstances that prompt the speaker to attribute the respective characteristics to states of affairs as objects in the discourse.

It should be noticed that in the field under consideration there are four- and not two-stage oppositions, based respectively on a) ordinary adverbs of manner, b) the expressions having the form *w taki a taki sposób* ‘in such and such a way’, c) UMA, and d) sentential adverbs or particles, e.g.:

- *wycieczka zakończyła się szczęśliwie* ‘the trip ended happily’ – *w szczęśliwy sposób udało mu się zapamiętać adres* ‘in a lucky way, he managed to memorize the address’ – *szczęśliwie nie zapomniał adresu* ‘he didn’t, luckily, forget the address’ – *na szczęście nie zapomniał adresu* ‘fortunately, he didn’t forget the address’;
- *zachowuje się trochę dziwnie* ‘he behaves quite strangely’ – *w dziwny sposób uniknął obrażeń* ‘in a strange way, he avoided injuries’ – *jestem o to dziwnie spokojny* ‘I am strangely calm about it’ – *o dziwo, zdał egzamin / ku memu zdziwieniu zdał egzamin* ‘Unexpectedly enough, he passed the exam’ / ‘to my surprise, he passed the exam’.

The following problems require further reflection: How is this four-scale opposition implemented in other Polish expressions? Which of its members are present and which are not? Do the language mechanisms present in Polish have their analogies in different language systems? How does the analytic expression *w _ sposób* ‘in a _ way’ differ from an adverb of manner, on the one hand, and an UMA, on the other?

At this point, it seems more important to explain the relationship between the class of UMA and what in the literature on the subject is referred to as sentential adverbs or particles. Most of the tests examined above give the same result for the UMA and the particles. Meanwhile, if it is impossible to show the difference between the two classes of words, the postulate of separating UMA from other linguistic facts may be threatened.

Wajszczuk's classification proposals (1997, 2000, 2005: 105–121) come in handy here. The author describes particles and conjunctions as operators ordering the thematic-rhematic structure of an utterance, rather than its syntactic structure. In addition to their formal and semantic liberal behavior, a characteristic property of these operators is that they never occupy the valency positions of other expressions: they do not impose any restrictions on the words they are linked to. They merely never co-occur with expressions of the same rank.

Unlike extremely liberal particles, unqualifiable meta-predicative adverbials are sensitive to grammatical and semantic context. This feature distinguishes them from words such as *mianowicie* 'namely', *rzeczywiście* 'actually', *oczywiście* 'obviously', *pozornie* 'seemingly', *przeważnie* 'mostly', *prawdopodobnie* 'presumably', *niewątpliwie* 'undoubtedly', which, being adverbs by origin, have evolved up to the point of becoming similar to the original particles or sentential adverbs. In contemporary Polish, they can co-occur with any kind of words (even proper names) that take a non-corrective stress and can play the role of rhematic expressions. This is not the case with UMA. Cf.:

- (44) *Antykościelne hasła wykrzykiwał przypuszczalnie Janek.* Vs *Antykościelne hasła wykrzykiwał *iście / *czysto / *dosłownie Janek.*
 'Anti-church slogans were shouted, probably by Johnny.' Vs 'Anti-church slogans were shouted *truly / *purely / *literally by Johnny.'

Epistemic particles or sentential adverbs of adverbial origin, in contrast to UMA, can be prosodically, and thus with using punctuation, separated from the rest of the utterance, e.g.:

- (45) *Stefan, przypuszczalnie, skłamał.* Vs **Stefan, zwyczajnie, skłamał.*
 'Steven, probably, lied.' Vs 'Steven, normally, lied.'
- (46) *To był, istotnie, królewski obiad.* Vs **To był, iście, królewski obiad.*
 'It was, indeed, a royal dinner.' Vs. 'It was a truly royal dinner.'

The lack of such a possibility in the case of UMA favors the hypothesis that they are involved in syntactic relations rather than in the thematic-rhematic organization of utterances.

An important argument in favor of distinguishing UMA from particles on the one hand and sentential adverbs on the other is a kind of semantic subordination of the former by the latter. The point is that the UMA, together with the predicate whose characteristic it carries, remains within the range of influence of a possible particle, rather than vice versa, e.g.:

- (47) *To była rzeczywiście czysto antypolska wypowiedź.*

‘It was actually a purely anti-Polish statement.’

- (48) *Janek niewątpliwie bezwzględnie zasługuje na nagrodę.*

‘Undoubtedly, Johnny absolutely deserves the award.’

- (49) *Tym razem na szczęście zdecydowanie mają rację.*

‘Fortunately, this time, they are definitely right.’

- (50) *Chłopcy niestety dziwnie milczą o całej sprawie.*

‘Unfortunately, the boys are strangely silent about the whole thing.’

Conclusions

Unqualifiable meta-predicative adverbials cannot be reduced to standard adverbs of manner (cf. Grzegorzczkova 1975), particles, or sentential adverbs (Grochowski 1986: 57, 1997: 25, Żabowska 2008). This group has to be assigned a special, prominent status. What makes this an urgent task is the phenomenon of a clear numerical expansion of the class. The first symptom of a word abandoning the domain of standard adverbs is an extension of its combinability with adjectives, intensifiers, prepositional phrases, or even nouns. The widening of this process is associated with a change in the stress pattern and the position provided for a given word in the linear organization of the utterance (the adverbial tends to be placed in an unstressed anteposition). The next stage of evolution is predictively complete freedom of combinability, which accompanies the passage of the word from syntax to the area of the thematic-rhematic structure. In this way, the word acquires the nature of a particle or sentential adverb.

References

- Apresjan, J. D. (ed.). (2003). *Novyj ob''jasnitel'nyj slovar' russkogo jazyka*. Škola "Jazyki slovjanskoj kul'tury".
- Bartsch, R. (1972). *Adverbialsemantik. Die Konstitution logisch-semantischer Repräsentationen von Adverbialkonstruktionen*. (Linguistische Forschungsberichte 6). Athenäum Verlag. English Translation: *The Grammar of Adverbials. A Study in the Semantics and Syntax of Adverbial Constructions*. North-Holland Publishing Company 1976.
- Běličová-Křížková, H. (1980). K adverbialnímu určení zřetele v ruštině a v češtině. *Československá Rusistika*, 25 (4), 149–155.
- Bellert, I. (1977). On semantic and distributional properties of sentential adverbs. *Linguistic Inquiry*, 8, 337–351. JSTOR, <http://www.jstor.org/stable/4177988>. Accessed 11 Nov. 2024.
- Boguslavskaya, O. J. & Boguslavsky, I. (2014). Emotion and inner state adverbials in Russian. In: *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)* (pp. 38–47). Uppsala University. <https://aclanthology.org/W15-2106>. Accessed 8 Nov. 2024.
- Bogusławski, A. (2005). O operacjach przysłówkowych. In: M. Grochowski (ed.) *Przysłówki i przyimki. Studia ze składni i semantyki języka polskiego* (pp. 15–44). Wydawnictwa Uniwersytetu Mikołaja Kopernika.
- Bogusławski, A. (2007). *A Study in the Linguistics-Philosophy Interface*. BelStudio.
- Bogusławski, A. (2022). On the status of „attestation adverbials” in bipolar questions. *Linguistica Copernicana*, 19, 33–40. DOI:10.12775/LinCop.2022.003. Accessed 8 Nov. 2024.
- Cinque, G. (1999). *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. Oxford University Press.
- Greenbaum, S. (1969). *Studies in English Adverbial Usage*. Longmans.
- Grochowski, M. (1986). *Polskie partykuły. Składnia, semantyka, leksykografia*. Zakład Narodowy im. Ossolińskich.
- Grochowski, M. (19896a). O metapredykatywnej funkcji niektórych wyrażeń partykułowo-przysłówkowych w strukturze tekstu. In: T. Dobrzyńska (ed.) *Teoria tekstu. Zbiór studiów* (pp. 139–148). Zakład Narodowy im. Ossolińskich.
- Grochowski, M. (1997). *Wyrażenia funkcyjne. Studium leksykograficzne*. Instytut Języka Polskiego Polskiej Akademii Nauk.
- Grochowski, M. (2008). Operatory metatekstowe o kształcie superlativu przysłówka. *Južnoslovenski Filolog*, 64, 61–72.
- Grzegorzczkova, R. (1975). *Funkcje semantyczne i składniowe polskich przysłówków*. Zakład Narodowy im. Ossolińskich.

- Danielewiczowa, M. (2002). *Wiedza i niewiedza. Studium polskich czasowników epistemicznych*. Katedra Lingwistyki Formalnej UW.
- Danielewiczowa, M. (2011). Wieloznaczność – skaza na języku czy na jego opisie. In: M. Bańko & D. Kopcińska (eds.) *Różne formy, różne treści. Tom ofiarowany Profesorowi Markowi Świdzińskiemu* (pp. 37–47). Wydział Polonistyki UW.
- Danielewiczowa, M. (2012). *W głąb specjalizacji znaczeń. Przysłówkowe metapredykaty atestacyjne*. BelStudio.
- Danielewiczowa, M. (2021). *Aspekt tematyczny w informacyjnej strukturze wypowiedzi. Integracja i rozszerzanie wiedzy*. Wydawnictwa Uniwersytetu Warszawskiego.
- Danielewiczowa, M. (2021a). Tematičeskij aspekt i ego pokazateli v pol'skom jazyke. *Slovnia časopis pro slovanskou filologiju*, 90 (2), 129–137.
- Danielewiczowa, M. (2022). Adverbial superlative forms outside the degree opposition: lexical and operational units. *Studies in Polish Linguistics*, 17 (3), 115–143. DOI:10.4467/23005920SPL.22.006.16732. Accessed 8 Nov. 2024.
- Ducrot, O. (1980). Analyse pragmatique. *Communication* 32, 11–16. https://www.persee.fr/doc/comm_0588-8018_1980_num_32_1_1481. Accessed 8 Nov. 2024.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press.
- Lang, E. (1979). Zum Status der Satzadverbiale, *Slovo a Slovesnost*, 40, 200–213. <http://sas.ujc.cas.cz/archiv.php?art=2620>. Accessed 8 Nov. 2024.
- Martin, R. (1973). Le mot *puisque*: notion d'adverbe de phrase et de presupposition sémantique. *Studia Neophilologica*, 45 (1), 104–114.
- Martin, R. (1974). La notion d'adverbe de phrase: essai d'interprétation en grammaire générative. In: von H. E. Brekle, H. J. Heringer, Ch. Rohrer, H. Vater & O. Werner (eds.) *Actes du colloque franco-allemand II* (pp. 66–75). Max Niemayer Verlag.
- Maienborn, C. & Schäfer, M. (2012). Semantics of adjectives and adverb(ial)s. In: K. von Eusinger, C. Maienborn & P. Portner (eds.) *Semantics. An International Handbook of Natural Language Meaning II* (vol. II, pp. 1390–1419). Walter de Gruyter GmbH & Co. KG.
- McNally, L. & Kennedy, Ch. (eds.). (2008). *Adjectives and Adverbs: Syntax, Semantics, and Discourse*. Oxford University Press.
- Morzycki, M. (2008). Adverbial modification of adjectives: Evaluatives and a little beyond. In: J. Dölling, T. Heyde-Zybatow & M. Schäfer (eds.). *Event Structures in Linguistic Form and Interpretation* (pp. 103–126). De Gruyter. <https://doi.org/10.1515/9783110925449.103>.
- Molinier, Ch. (1990). Une classification des adverbes en *-ment*. *Langue Française*, 88, 28–40. https://www.persee.fr/doc/lfr_0023-8368_1990_num_88_1_5750. Accessed 8 Nov. 2024.
- Mørdrup, O. (1976). *Une analyse non-transformationnelle des adverbs en -ment*. Akademisk Forlag. (= *Revue Romane*, numero spéciale 11).

- Nilsen, Ø. (2004). Domain for adverbs. *Lingua*, 114, 809-847. DOI:10.1016/S0024-3841(03)00052-4.
- Nølke, H. (1990a). Les adverbiaux contextuels: problèmes de classification. *Langue Française*, 88, 12-27. https://www.persee.fr/doc/lfr_0023-8368_1990_num_88_1_5749. Accessed 8 Nov. 2024.
- Nølke, H. (1990b). Recherches sur les adverbs: bref aperçu historiques des travaux de classification. *Langue Française*, 88, 117-122. https://www.persee.fr/doc/lfr_0023-8368_1990_num_88_1_5757. Accessed 8 Nov. 2024.
- Thomas, R. (2009). Speaker-oriented adverbs. *Natural Language and Linguistic Theory*, 27, 497-544. JSTOR, <http://www.jstor.org/stable/40270294>. Accessed 11 Nov. 2024.
- Thomason, R. H. & Stalnaker, R. C. (1973). A semantic theory of adverbs. *Linguistic Inquiry*, 4, 195-220. <https://www.jstor.org/stable/4177764>.
- Wajszczuk, J. (1997). *System znaczeń w obszarze spójników polskich*. Katedra Lingwistyki Formalnej UW.
- Wajszczuk, J. (2000). Can a division of lexemes according to syntactic criteria be consistent? *Biuletyn Polskiego Towarzystwa Językoznawczego*, 55, 20-38.
- Wajszczuk, J. (2005). *O metateksie*. Katedra Lingwistyki Formalnej UW.
- Żabowska, M. (2008). *Polskie partykuły epistemiczne nieprzesądzające prawdziwości sądu*. (Unpublished Ph.D. diss. Uniwersytet Mikołaja Kopernika).



Lucie Barque

Université Sorbonne Paris Nord
& Laboratoire de Linguistique Formelle
France

 <https://orcid.org/0000-0002-1806-848X>

Is metonymy really more regular than metaphor? A first investigation into the French Lexicon

Abstract

Regular polysemy is considered more intrinsically linked to metonymy than to metaphor in both linguistic and psycholinguistic literature. Yet, on the one hand, metonymy exhibits varying degrees of regularity, and on the other hand, metaphor is also subject to polysemic regularity. This paper aims to assess the extent to which metonymy and metaphor exhibit regularity in the lexicon. Based on the analysis of lexicographic descriptions of a sample of nearly 3,000 French nouns, this study compares the regularity of polysemy patterns—defined as recurring sense alternations—within the two types of figures. While the findings confirm theoretical assumptions by identifying more regular metonymic patterns than regular metaphoric ones, they reveal, on average, no significant difference between metonymy and metaphor in terms of regularity. The resulting dataset is publicly available to allow further exploration of how regular polysemy emerges in the lexicon and its correlation with lexical figures.

Keywords

Regular polysemy, metonymy, metaphor, French lexicon

1. Introduction

Since Jurij Apresjan's foundational paper on regular polysemy (1974), both linguistic and psycholinguistic literature have consistently associated metonymy with regularity and metaphor with irregularity. Theoretical studies on regular

polysemy have primarily focused on highly regular metonymic alternations, such as the animal-meat or the container-content relations (Nunberg & Zaenen, 1997; Pustejovsky, 1995, among others). In contrast, metaphor has received relatively little attention in discussions of regular polysemy, despite also exhibiting patterns of regularity (e.g., Copestake & Briscoe, 1995; Vicente & Falkum, 2017).

In this paper, we aim to investigate to what extent metonymy is more regularly observed than metaphor in the lexicon, both in terms of number of sense alternations and in terms of number of polysemous words associated with each of them. To do so, we analyse a large set of sense alternations observed within a broad sample of French nouns and provide regularity scores for metonymy and metaphor patterns. In addition to empirically testing the theoretical assumption of a strong correlation between figure and regularity, this study aims to advance the understanding of the various factors that contribute to the regularity of sense alternations. It also seeks to provide data that can be used in psycholinguistic studies, for instance, to examine the distinct effects of figure and regularity on the cognitive processing of polysemous words and their mental representation¹.

The remainder of the paper is organized as follows. Section 2 provides a brief review of previous research on the relationship between lexical figures and regular polysemy, as well as past attempts to assess scalar regularity using lexical resources or corpora. Section 3 details the methodology used for data selection and analysis. Section 4 presents the main findings, which confirm that metaphor also plays a central role in the formation of regular polysemy, although generating fewer patterns than metonymy. However, the proportion of polysemous words it produces is comparable to that of metonymic patterns.

2. Related work

2.1. Regular polysemy and lexical figures

According to Apresjan (1974), a sense extension exhibited by a word can be considered regular if at least one other word within the same semantic class undergoes the same type of alternation. Both theoretical and empirical studies in linguistics tend to associate regular polysemy with metonymy (Nunberg & Zaenen,

¹ The data are available at <https://osf.io/txvhw/>.

1997; Pustejovsky, 1995; Dölling, 2020). For instance, among the 50 nominal regular polysemy patterns listed in Apresjan (1974), 47 are metonymic. Similarly, Srinivasan and Rabagliati (2015) examine the cross-linguistic occurrence of regular polysemy by testing 27 patterns, 24 of which are metonymic. Psycholinguistic studies on the processing and mental representation of ambiguous words also suggest a correlation between figurative meaning and regularity: metonymy is seen as involving more closely related senses than metaphor (e.g., Klepousniotou et al., 2008; Lopukhina et al., 2018; Yurchenko, 2020) and more closely related senses tend to exhibit greater regularity (e.g., Xu et al., 2020).

The connection between figure and regularity can be partly explained by the structural characteristics of the two figures. Metonymy relies on a contiguity relationship between the referent of the derived sense and the referent of the source sense, which results in a referent closely tied to the one denoted by the source. In contrast, metaphor is based on an analogy between two referents, leading to the creation of a referent distinct from the original. This difference generally means that two senses linked by metonymy are more closely related than those linked by metaphor, especially when contiguity involves referential inclusion—i.e., when the referent of the derived sense is part of the referent of the source sense (e.g., in the metonymy “animal > meat”). As a result, the target sense is generally more predictable. Thus, interpreting or producing a metonymic sense, which is generally closer to its source meaning and more predictable than a metaphorical sense, requires less cognitive effort, which helps explain why metonymy is more regularly used.

2.2. Identifying polysemy patterns

The use of structured large-scale lexical resources has enabled the automatic detection of lemmas exhibiting similar sense alternations. This approach, primarily applied to English WordNet data (Buitelaar, 1998; Peters, 2006, a.o.), has led to inventories of polysemy patterns with varying degrees of completeness and precision. For example, Buitelaar (2000) leveraged the synset hierarchy in WordNet to identify 1,341 sense clusters that potentially represent regular polysemy across 3,336 English nouns. For French, Barque et al., (2018) identified around a hundred metonymy and metaphor patterns from two lexical resources. While these studies have provided initial data on regular alternations in the lexicon, the results—extracted automatically—have not been fully validated through manual annotation, nor were they obtained from a systematic sampling of the lexicon.

Two lexical resources hold promise for a more precise identification of polysemy patterns, as they provide hand-crafted annotations of relations between word senses. The RL-fr (*Réseau Lexical du Français*) is a French lexical network in which nodes correspond to semantically specified lexical units, and arcs represent paradigmatic or syntagmatic relations. One particular paradigmatic relation is co-polysemy (Polguère, 2018). The RL-fr is still under development and currently contains 9,659 co-polysemy relations, which can be leveraged to infer regular patterns of sense alternation (Polguère, 2022). As for English, Chain-Net is a recently released database in which relations between 22,178 senses of 6,500 words from Princeton WordNet 3.0 (Fellbaum, 1998) have been annotated for metonymy, metaphor, or homonymy (Maudslay et al., 2024). However, this second set of studies has not yet undertaken a comprehensive inventory of regular alternations.

Our study aims to bridge this gap by offering a rigorous identification and classification of regular polysemy patterns attested in the French lexicon.

2.3. Assessing polysemy patterns regularity

Calculating the degree of regularity of a polysemy pattern requires, at least, evaluating the proportion of words that truly fit the pattern in relation to the total number of words that could potentially do so. It also requires information about the frequency of these words and, ideally, the distribution of their meanings. The more polysemous words a pattern generates in the language relative to its generative potential, the more regular it is. This regularity increases if the words belonging to the pattern are commonly used and if the polysemous words it generates have balanced meanings. For example, nouns like *bœuf* ‘beef’ and *autruche* ‘ostrich’ should not contribute equally to calculating the regularity of the “animal>meat” pattern, as *bœuf* is more frequent than *autruche*, and the meanings ‘animal’ and ‘meat’ are more balanced for the former than for the latter. Accordingly, at least four variables should be considered when determining the regularity of a polysemy pattern $A \rightarrow B$:

1. The number of words attested with sense A,
2. The number of words attested with both senses A and B,
3. The frequency of these words in the corpus,
4. The relative frequency of senses A and B for polysemous words.

Lombard et al., (2023) proposed several measures that account for a varying number of parameters depending on the availability of the data. These measures yield regularity scores that range from 0 (indicating irregularity) to 1 (indicating

systematicity), enabling comparisons between patterns within the same language and across different languages. Since the information related to parameter 4 is unavailable for a large portion of the lexicon, both for French and other languages, we will use the following formula, which relies on the first three parameters to compute the regularity degree of a pattern $P_{A \rightarrow B}$. In this formula, n represents the number of polysemous words instantiating P , N denotes the number of words having source meaning A in the dataset, and f corresponds to the word's frequency. Frequencies are log-transformed to minimize discrepancies in absolute frequency between words.

$$Reg(P_{A \rightarrow B}) = \frac{\sum_{i=1}^n \log(f_i)}{\sum_{j=1}^N \log(f_j)}$$

2.4. This study

For practical reasons, we restrict our analysis to sense alternations involving meanings that do not belong to the same coarse semantic class, such as the animal–person metaphor (e.g., *mouton* ‘sheep/follower’) or the animal–meat metonymy (e.g., *mouton* ‘sheep/mutton’). By contrast, polysemic extensions within the same semantic class, such as the person–person metaphor (e.g., *boucher* ‘butcher/killer’), are not considered here. Nevertheless, it could be the case that one of the two figures of speech under study is more likely than the other to generate new meanings within the same semantic class as the source meaning. This suggests that our account of the predominance of metonymy in regular polysemy should ultimately be complemented by an analysis of such within-class extensions.

The study is based on the analysis of lexicographic data to infer properties of regular polysemy in the lexicon. We know that lexicographic data offer an imperfect view on the lexicon of a given language, defined as the theoretical sum of words and word senses known and/or used by the speakers of that language. This is particularly problematic for assessing polysemy regularity, since the more a polysemy is regular, the more the meaning it produces is predictable (from the source meaning), so the less it is probable for these senses to be systematically consigned in the dictionary. For instance, in theory, any noun that primarily refers to a cooking container can also be used denote the quantity of food it holds (e.g., *she ate two plates of rice*). Another well-known limitation of lexicographic data is that they do not provide information on sense frequency, however a crucial parameter for assessing the regularity of polysemy patterns.

Despite these limitations, our study will give a first empirical assessment of how metonymy and metaphor contribute to regular polysemy in the French nominal lexicon. It also provides a large set of manually curated data for conducting psycholinguistic experiments on the intricate relationships between lexical figures and regularity.

3. Data

In this section, we first expose the methodology used to select a sample of nouns representing common French and to extract regular sense alternation based on their lexicographic description (Section 3.1). Then we detail how we manually analysed these data in order to identify regular sense extensions as well as their instances in the selected data (Section 3.2).

3.1. Data selection

3.1.1. A semantically annotated version of the French Wiktionary

As noted above, most previous studies that aimed to identify polysemy patterns from lexicographic data used WordNet (Buitelaar, 1998; Peters, 2006, a.o.). Since French lacks WordNet-like resources that meet the two conditions of coverage and quality, we leverage data from the *SuperWikt-fr* lexicon, a recently released version of the French Wiktionary in which every nominal sense has been automatically annotated with its general semantic class (Angleraud et al., 2025). The resource is broad coverage, with more than 300,000 annotated nominal senses, and the quality of the automatic semantic classification allows for an extraction of sense alternation patterns that we will be further analysed manually. The semantic classification was performed using supervised classifiers trained and tested on a large set of manually curated data. The classification reached about 85% precision, with performance varying across semantic categories. The easiest senses to classify were unsurprisingly those having concrete referents, such as Person (F-score: 96%), Body (F-score: 85%) or Artifact (F-score: 85%), while the most challenging cases mostly correspond to abstract meanings, such as Cognition (66%) or State (62%). Moreover, the use of supersenses as semantic tags (e.g., Person, Act, State) allows for a certain interoperability with other studies investigating regular polysemy with WordNet-like resources for other languages.

3.1.2. A representative sample of familiar French nouns

The objective of this study is to examine to what extent metonymy exhibits greater regularity than metaphor in the French lexicon. For this purpose, we need to define a representative sample of French that is large enough to capture most of the recurring sense alternations while being of reasonable size to allow for a detailed manual analysis. To achieve this, we leverage data from *Lexique 3*, a resource that provides information on words found in a large French corpus, including both textual frequency and subjective familiarity estimates (New et al., 2004). More precisely, we focus on the 5,989 nouns reported in *Lexique 3* as being known by 100% of respondents. This constraint ensures that the analysed set consists of words generally familiar to adult French speakers. It should be noted, however, that this criterion does not specify which senses of these words are commonly known—an issue to which we return below.

Since our interest lies in regular polysemy, we exclude from this set the 1,579 nouns that are monosemous, i.e., associated with a single sense in *SuperWikt-fr*. Among the remaining 4,410 multisemous nouns, 1,494 are mono-categorical, meaning all their different senses belong to the same broad semantic class (or supersense). Because our analysis is restricted to sense alternations that involve a shift in semantic category, these nouns are also set aside. Table 1 presents the distribution of the 2,916 remaining nouns according to the number of distinct supersenses their meanings have been automatically associated within the *SuperWikt-fr*.

Table 1.
Repartition of the 2,926 nouns of the dataset according to their ambiguity level (supersense granularity)

Ambiguity level	2	3	4	5	6	7	8	9	10	11	
Nb of nouns	1,510	750	353	165	68	42	16	5	4	3	2,916

The distribution of the 8,455 supersense instances in the dataset is given in Table 2, in descending order. The table provides a raw characterisation of each supersense label.

Table 2*Distribution of supersenses in the dataset*

Supersense	Definition	Nb
Artifact	Manufactured physical object (includes devices, buildings, constructed place, etc)	1,350
Act	Action caused by an agent	1,074
Cognition	Mental object / informational object and other products of mental activity (includes fields of knowledge, doctrines and methods)	950
Person	Humans and imaginary characters with prominent human attributes	768
Attribute	Constitutive property of an entity or of a situation	464
State	Non-constitutive property of an entity, resulting states, or state of affairs	346
Animal	Non-human living being	338
Object	Natural physical object (includes microorganisms and natural places)	329
Food	Natural food or prepared food	319
Event	Dynamic situation without an agent (natural event, accidental event)	295
Body	Body or body part or organ of animals or humans	276
Institution	All types of organizations, created with a purpose	248
Substance	Physical material (liquid, solid, or gas) plus chemical compounds, whether manufactured or not	220
Plant	Entity of the plant kingdom or a part of this entity	215
Artifact*Cognition	Document, informational medium	213
Quantity	Abstract numerical object (includes numbers, units of measurement, proportions, etc.)	181
Act*Cognition	Communication act	150
Possession	Abstract entities of the type intangible assets or financial assets (includes money, intellectual property, financial assets)	145
GroupxPerson	Group of people	127
Communication	Linguistic objects, languages, communication medium (e.g., radio, internet)	123
Feeling	Transitory psychological or physiological state	117
Time	Time period or time reference	107

Table 2 (Continuation)

Supersense	Definition	Nb
Phenomenon	Entity observed at a given moment, subject to a form of temporality but not occurring (natural forces, sensory manifestations)	100
		8,455

3.1.3. Extracting candidates for polysemy patterns

The algorithm for automatically extracting groups of nouns with similar supersense alternations operates in two steps. First, following the order in which senses are described, we group together those that belong to the same supersense. For instance, the noun *punaise* ‘bug’, ‘wretch’, ‘drawing pin’ has four senses in *Wiktionary*, with senses 2 and 3 falling under the same semantic category, *Person*, and thus being grouped together (1).

(1) *punaise* ‘bug’, ‘wretch’, ‘drawing pin’

Animal [1. Nom désignant de nombreuses sortes d’insectes, souvent de forme plate, qui émettent une odeur désagréable lorsqu’ils sont effrayés ou tués. ‘A name referring to many types of insects, often flat in shape, that emit an unpleasant odor when frightened or crushed.’]

Person [2. (Par analogie) (Péjoratif) Personne petite, malingre, peu imposante. ‘(By analogy) (Pejorative) A small, scrawny, unimposing person.’ 3. (Familier) (Par extension) Personne malveillante. ‘(Informal) (By extension) A malicious person.’]

Artifact [4. (Par analogie de forme) Petit clou court, à tête plate et large, dont on se sert pour fixer au mur ou sur une planche des feuilles de papier, des images, des affiches, etc. ‘(By analogy of shape) A short, small nail with a flat, wide head, used to fasten sheets of paper, pictures, posters, etc., to a wall or a board.’]

Then, we examine the relationship between the first supersense and each of the subsequent supersenses. For a noun whose senses are grouped into *n* distinct supersenses, *n-1* supersense alternations are extracted. Returning to *punaise* in (1), only the semantic alternations of type Animal-Person and Animal-Artifact are analysed. This approach calls for several comments. In the specific case of *punaise*, it makes sense not to compare the two Person-type senses with the Artifact sense, as they are not directly related. However, it is often the case that a derived sense does not stem directly from the primary meaning of the word but

rather from another derived sense. It would therefore have been more satisfactory to analyse all possible ordered combinations of distinct supersenses. However, the reality of the amount of data to be analysed leads us to limit ourselves to this first type of extension, leaving the other cases for future study. Furthermore, we assume that the first listed sense corresponds to the primary meaning, from which all other senses derive, either directly or indirectly. We are aware that this proxy has its limitations, especially concerning morphologically complex words, which will be discussed later.

Following this methodology, we identified a total of 354 recurring supersense alternations, with the number of nouns per pattern ranging from 2 to 184. Table 3 presents the 15 most frequently observed supersense alternations in the selected data, each illustrated by a noun with at least two senses annotated with these supersenses.

Table 3.

Most frequent supersense alternations automatically extracted from the dataset

Sense alternation	Nouns	Example
Artifact – Cognition	184	<i>tube</i> ‘tube / hit’
Act – Artifact	178	<i>chargement</i> ‘loading / cargo’
Artifact – Act	137	<i>équerre</i> ‘square / L-sit’
Person – Artifact	128	<i>cycliste</i> ‘cyclist / cycling shorts’
Act – Event	118	<i>arrêt</i> ‘stop something / something stops’
Act – Cognition	110	<i>révélation</i> ‘disclosure’
Animal – Person	108	<i>mouton</i> ‘sheep’
Artifact – Body	91	<i>trompe</i> ‘horn / trunk’
Artifact – Object	89	<i>bouton</i> ‘button / pimple’
Person – Animal	72	<i>ouvrière</i> ‘worker / worker bee’
Cognition – Act	72	<i>signature</i> ‘signature / signing’
Animal – Artifact	71	<i>souris</i> ‘mouse’
Act – State	69	<i>isolement</i> ‘isolation’
Act – Attribute	62	<i>organisation</i> ‘organization / set up’
Body – Artifact	61	<i>pied</i> ‘foot / (table) leg’

3.2. Data analysis

Not every supersense alternation automatically extracted from the dataset instantiates a given polysemy pattern. The analysis therefore involves identifying which ones do and which ones do not.

3.2.1. Discarding unknown meanings

The first step of the analysis is to determine whether the annotator—the author of the paper in this case—is familiar with the meanings associated with a given supersense pair. We know that the selected nouns are well-known, given the way they were selected from *Lexique 3* (cf. section 3.1.2), but *Wiktionary* provides a detailed account of polysemy, including many specialized meanings. Since analysing relationships between senses that are at least partially unknown introduces significant uncertainty, such senses must be excluded from our analysis. The examples below illustrate cases where the author is not familiar with the two senses described in (2b) and (3b). Linguistic labels are often included in definitions (in brackets, such as [mécanique]), which could potentially be used as indicators for automatically identifying rare meanings. However, since these labels are not consistently applied to mark specialized or uncommon senses, as seen in (2b), and can also appear for common meanings, as shown in (3a), this step was carried out manually.

(2) *polka* ‘polka’, ‘wheelbarrow’

- a. Danse de salon d’origine tchèque ou polonaise.
(A ballroom dance of Czech or Polish origin.)
- b. Brouette utilisée dans les champignonnières.
(A wheelbarrow used in mushroom farms)

(3) *rouage* ‘component / toll’

- a. [mécanique] Chacune des pièces qui concourent au fonctionnement d’une machine.
([mechanics] Each component that contributes to the operation of a machine)
- b. [histoire, Droit d’Ancien Régime] Péage pour circuler sur les routes.
([History, Old Regime Law] Toll for travelling on roads)

As explained in the previous section, analysed supersense alternations occur between groups of senses (e.g., the *Person* supersense associated with two definitions in example (1)). A supersense alternation *A–B* associated with

a given noun is discarded if all senses of *A* or all senses of *B* are unknown to the annotator.

3.2.2. Identifying semantic classification errors

For each noun exhibiting a supersense alternation A–B, we must first verify whether the two senses involved indeed correspond to semantic type A and B respectively. Since nominal senses have been automatically tagged based on their lexicographic description (cf. section 3.1.1), some errors are inevitably present in the data. The two examples below illustrate such false positives. In (4), the two senses of *boulangier* ‘baker’ does not instantiate a Person–Artifact supersense alternation, as the target meaning has been misclassified—likely because the noun *tenant* used in its definition can also refer to an artifact. In (5), the source meaning of *nounours* ‘teddy bear’ has been classified as Animal instead of Artifact, probably due to the fact that the lexicographic definition uses *ours* (bear) as hypernym rather than explicitly stating that the noun refers to a toy.

(4) *boulangier* (baker)

- a. Artisan dont le métier est de fabriquer ou de vendre le pain. → **Person**
(Someone who bakes and sells bread)
- b. Tenant d’une boulangerie. → **Artifact [Person]**
(The owner of a bakery)

(5) *nounours* (teddy bear)

- a. Ours en peluche. → **Animal [Artifact]**
(Teddy bear, litt. Plush bear)
- b. Person très caline. → **Person**
(Cuddly person)

Once corrected, these pairs of meanings are either filtered out—if both meanings belong to the same broad class, as in *boulangier* ‘baker’ – or reclassified into the appropriate supersense alternation group of supersense alternation, as with *nounours* ‘teddy bear’ in the Artifact–Person alternation.

3.2.3. Focusing on regular metonymical and metaphorical relations

The next step is to determine whether two meanings of a given noun can be considered directly related through a regular metonymic or metaphorical link. First, we must establish whether a given sense alternation represents a case of metonymy, metaphor, or another type of relationship. Then, for metonymies and metaphors, we need to assess whether they follow a regular pattern. In line with

Apresjan (1974), we consider a sense extension to be regular if it is instantiated by at least two polysemous words. For example, the metonymic shift from a noun denoting an attribute to a noun denoting a person characterized by that attribute is regular, since it applies, among others, to *nullité* ‘worthlessness / a nobody’ and *laideur* ‘ugliness / unattractive person’.

Since our analysis focuses on pairs of senses involving a supersense shift, the data should not include cases of specialization or generalization. The four possible values describing the semantic link between meanings A and B are as follows:

1. Homonymy: the two meanings are not semantically related. For example, the two meanings of the noun *palais* ‘palate / palace’ have no semantic connection.

(6) *palais* ‘palate / palace’

- a. Paroi supérieure qui sépare la fosse nasale de la bouche. → **Body**
(Upper wall / roof that separates the nasal cavity from the mouth)
- b. Vaste demeure urbaine d’un souverain, d’un prince, d’un chef d’État [...].
→ **Artifact**
(A large urban residence of a sovereign, a prince, a head of state [...])

2. Metaphor: the two meanings are related through a metaphoric extension, where the referent of B can be seen as similar to the referent of A. In (7), the target meaning of *girouette* refers to a person whose behaviour resembles that of a weather vane.

(7) *girouette* ‘weather vane / fickle person’

- a. Pièce de fer-blanc ou d’autre matière mince et légère, ordinairement en forme de flèche empennée ou de coq, qu’on pose sur un pivot en un lieu élevé, de manière qu’elle tourne au gré du vent et qu’elle en indique la direction. → **Artifact**
(A piece of tinplate or another thin and lightweight material, usually in the shape of a feathered arrow or a rooster, placed on a pivot in a high location so that it turns with the wind and indicates its direction.)
- b. Personne qui change souvent d’avis, de sentiment, de parti. → **Person**
(A person who often changes their mind, feelings, or allegiance.)

3. Metonymy: The two meanings are related through a metonymic extension, where the referent of B is in a contiguity relationship with the referent of A. For example, *taxi* can refer to a type of vehicle or to the person who drives that vehicle.

(8) *taxi* ‘taxi / taxi driver’

- a. Véhicule automobile terrestre privé, conduit par un chauffeur, destiné au transport payant de passagers et de leurs bagages, de porte à porte. →

Artifact

(A private land motor vehicle, driven by a chauffeur, intended for the paid transportation of passengers and their luggage, door to door.)

- b. Chauffeur de taxi. → **Person**

(Taxi driver)

The examples in (7) and (8) are prototypical instances of metaphor and metonymy respectively. However, it is common for a meaning relationship to involve both figures (e.g., Goossens, 2003). For example, *araignée* ‘spider / spider strainer’ refers to a kitchen tool that is not similar to the animal itself, but rather to the web the spider constructs. Cases where an analogy is built upon a contiguity relationship are classified as metaphor. Conversely, cases where a contiguity extension is built upon an analogy relationship are classified as metonymy.

4. Undecidable: There is a relationship between the two meanings, unlike in cases of homonymy, but this relationship is undecidable. Such indeterminacy typically arises in cases of morphological co-derivation, where meanings A and B, both resulting from a polyfunctional morphological process, interact with semantic extensions (e.g., Salvadori, 2024). For instance, the French suffix *-ier* can derive nouns denoting both persons (e.g., *rentier* ‘male annuitant’, *aventurière* ‘female explorer’) and artifacts (*dentier* ‘dentures’, *gouttière* ‘gutter’). This polyfunctionality often applies to a single formal base, leading to nominal ambiguity. A clear example is *cuisinière* ‘cook / kitchen stove’, which can refer to both a person and an artifact. The two senses of *cuisinière* are likely derived in parallel from the same root (*cuisine* ‘kitchen’). However, theoretically, one could posit a contiguity relationship (the kitchen stove is what the cook typically uses to cook) or an analogy relationship (the kitchen stove has the same function as the cooker), both of which are attested for simplex nouns. Cases like this are classified as undecidable.

- (9) *cuisinière* ‘cook / kitchen stove’
- a. Celle qui cuisine, qui prépare, qui cuit la nourriture → **Person**
(A woman who cooks, prepares, and bakes food)
 - b. [Électroménager] Fourneau de cuisine servant à chauffer ou faire cuire les aliments, souvent muni d’éléments chauffants sur sa surface de travail. → **Artifact**
([Appliance] A kitchen stove used for heating or cooking food, often equipped with heating elements on its surface)

4. Results

The analysis protocol described in the previous section was manually applied to a reduced set of 60 supersense alternations, each associated with at least 20 candidate nouns (mean = 50), resulting in a total of 2,960 instances of supersense alternations to be analysed. The results of this analysis are displayed in Table 4.

Table 4.
Distribution of the 2,992 analysed instances of supersense alternations

Discarded supersense alternations	Because of unknown meaning	1,327
	Because of the absence of supersense shift	86
Classified supersense alternations	Homonymy	97
	Metaphor	638
	Metonymy	728
	Undecidable	84
		2,960

4.1. Discarded sense alternations

During the first step of the analysis, 1,327 instances of supersense alternation (45%) were discarded because at least one sense in the pair was unknown to the author (cf. Section 3.2.1). Table 5 lists the 60 supersense alternations along with the percentage of instances discarded due to unknown meanings. This high proportion of rare meanings can be attributed to Wiktionary’s extensive coverage and its collaborative nature, where some contributors may focus on specialized domains. For instance, the dataset includes 142 senses related to different types

of charges on a heraldic shield. The Animal-Cognition supersense alternation, which has the highest rate of discarded instances due to unknown target meanings (98%), primarily involves such specialized senses, as illustrated in (10) with *crapaud* (toad).

(10) *crapaud* ‘toad’

- a. Batracien qui ressemble à la grenouille, aux pattes plus courtes, aux mœurs plus terrestres et dont le corps est couvert de pustules venimeuses. → **Animal**
(A batrachian resembling a frog, with shorter legs, more terrestrial habits, and a body covered in venomous pustules)
- b. [héraldique] Meuble représentant l’animal du même nom dans les armoiries. Il est généralement représenté dans sa forme naturelle (non stylisée). → **Cognition**
(A heraldic charge representing the animal of the same name. It is usually depicted in its natural (non-stylized) form)

Table 5 shows that supersense alternations highly vary on that respect. For instance, only 18% of the Animal-Person sense alternations (first column, third line) were filtered out. By comparison, 70% of the Person-Animal supersense alternation (third column, 13th line) were discarded, primarily because many of the animals belong to unfamiliar sub-species, such as *républicain* (republican), which refers to a species of passerine birds that live in communities in large suspended nests.

Table 5.
Percentage of discarded instances of supersense alternations due to unknown meanings

Plant-Food	11	Feeling-Act	35	Person-Artifact	48
Act-Feeling	17	Food-Act	36	State-Act	48
Animal-Person	18	Institution-Artifact	36	Body-Person	50
Act-Institution	21	Act-Cognition	37	Artifact-Cognition	51
Act-Artifact*Cognition	22	Artifact-Quantity	37	Object-Artifact	53
Animal-Food	25	Plant-Artifact	39	Person-Cognition	55
Cognition-Person	26	Attribute-Artifact	39	Body-Cognition	55
Act-Possession	27	Act-GroupxPerson	40	Artifact-Body	56
Food-Person	27	Artifact-Act	42	Artifact-Object	56

Table 5 (Continuation)

Attribute-Act	28	Artifact-State	42	Animal-Artifact	59
Artifact-GroupxPerson	29	Quantity-Artifact	42	Act-Object	64
Cognition-Act	30	Act-Person	42	Act-Attribute	65
Attribute-Person	31	Body-Artifact	43	Person-Animal	70
Artifact-Possession	31	Cognition-Artifact	43	Food-Artifact	70
Artifact-Person	32	Substance-Artifact	43	Person-Food	70
Attribute-Cognition	32	Object-Cognition	43	Body-Plant	76
Artifact-Attribute	32	Artifact-Food	44	Person-State	78
Act-Artifact	33	Animal-Substance	45	Artifact-Plant	82
Artifact-Institution	34	Act-State	47	Artifact-Animal	93
Cognition-Institution	35	Food-Body	47	Animal-Cognition	98

Additionally, 86 instances of supersense alternations (3%) were discarded due to semantic classification errors. After correction, the pair of meanings no longer imply a supersense shift (cf. example (4)).

4.2. Classified sense alternations

The analysis of the remaining 1,547 supersense alternations is presented in the lower part of Table 4. Among these alternations, 5% are instances of homonymy, 41% corresponds to metaphors, 47% to metonymies and 5% were classified as undecidable. Of the 1,366 pairs involving metaphor or metonymy, 78% were found to follow a regular pattern as defined in the previous section. This result provides a first assessment of the prevalence of regular polysemy among cases of (raw) sense alternations in a representative set of the lexicon.

Tables 6 and 7 present the full list of metonymic and metaphoric polysemy patterns that were extracted from the data, along with the number of their attestations (*n*) and the number of words that could theoretically instance each pattern (*N*). For example, in Table 6 (line 1), of the 138 nouns of the dataset that primarily refer to an animal, 59 exhibit a metaphorical meaning referring to a person. The regularity score of the pattern (*Reg*) was calculated using the formula provided in section 2.3, with word frequency estimates from *Lexique 3* (New et al., 2004).

Table 6.

Metaphorical patterns extracted from the dataset, ranked in descending order of regularity score. Column “n” indicates the number of words that enter the pattern A–B. Column “N” indicates the number of words whose first meaning is of type A.

Metaphorical patterns	n	N	Reg	Example
1. Animal-Person	59	138	0.48	<i>mouton</i> ‘sheep’
2. Body-Artifact	25	73	0.39	<i>pied</i> ‘foot / (table) leg’
3. Plant-Artifact	12	50	0.37	<i>feuille</i> ‘leaf / sheet of paper’
4. Institution-Artifact	3	20	0.21	<i>prison</i> ‘prison / cage’
5. Food-Artifact	12	78	0.20	<i>banane</i> ‘banana / bum bag’
6. Food-Body	11	78	0.19	<i>miche</i> ‘bun / cheek or boob’
7. Animal-Artifact	19	138	0.18	<i>souris</i> ‘mouse’
8. Food-Person	12	78	0.18	<i>légume</i> ‘vegetable’
9. Artifact-Cognition	63	504	0.15	<i>bagage</i> ‘luggage / background’
10. Food-Act	8	78	0.14	<i>pêche</i> ‘peach / punch’
11. Person-Artifact	29	247	0.10	<i>explorateur</i> ‘explorer’
12. Person-Animal	15	247	0.09	<i>ouvrière</i> ‘worker / worker bee’
13. Artifact-Person	36	504	0.07	<i>bulldozer</i> ‘bulldozer’
14. Artifact-Body	32	504	0.07	<i>disque</i> ‘disk’
15. Artifact-Object	28	504	0.07	<i>cuvette</i> ‘basin’
16. Object-Artifact	5	65	0.07	<i>bloc</i> ‘block’
17. Body-Plant	4	73	0.07	<i>pied</i> ‘foot / (vine) plant’
18. Cognition-Artifact	7	137	0.06	<i>trapèze</i> ‘trapezoid / trapeze’
19. Plant-Food	3	50	0.06	<i>tomate</i> ‘tomato plant / Pastis&grenadine’
20. Artifact-Act	24	504	0.05	<i>planche</i> ‘board / floating on your back’
21. Object-Cognition	2	65	0.05	<i>croûte</i> ‘scab / tacky painting’
22. Body-Person	3	73	0.05	<i>membre</i> ‘limb / member’
23. Artifact-Food	15	504	0.04	<i>chausson</i> ‘slipper / (apple) turnover’
24. Act-Feeling	9	400	0.03	<i>torture</i> ‘torture / agony’
25. Cognition-Act	3	137	0.02	<i>épopée</i> ‘epic / epic journey’
26. Animal-Food	2	138	0.02	<i>souris</i> ‘mouse / lamb chank’
27. Body-Cognition	2	73	0.02	<i>ossature</i> ‘skeleton / structure’
28. Artifact-Plant	6	504	0.01	<i>aiguille</i> ‘needle’

Table 6 (Continuation)

Metaphorical patterns	n	N	Reg	Example
29. Artifact-Attribute	3	504	0.01	<i>façade</i> ‘front / appearance’
30. Artifact-Quantity	5	504	0.01	<i>plafond</i> ‘ceiling’
31. Artifact-GroupxPerson	4	504	0.01	<i>brochette</i> ‘skewer / bunch’
32. Artifact-State	4	504	0.01	<i>impasse</i> ‘dead end / deadlock’
33. Person-Food	2	247	0.01	<i>religieuse</i> ‘nun / religieuse’
Mean	14	249	0.1	

Table 7.

Metonymic patterns extracted from the dataset, ranked in descending order of regularity score. Column “n” indicates the number of words that enter the pattern A–B. Column “N” indicates the number of words whose first meaning is of type A.

Metonymy patterns	n	N	Reg	Example
1. Feeling-Act	12	20	0.71	<i>plaisir</i> ‘pleasure’
2. Substance-Artifact	20	46	0.43	<i>verre</i> ‘glass’
3. Act-Artifact	109	400	0.27	<i>armement</i> ‘arming / weapons’
4. Plant-Food	17	50	0.25	<i>tomate</i> ‘tomato tree / tomato’
5. Quantity-Artifact	5	21	0.22	<i>décimètre</i> ‘decimeter / 10-centimeter ruler’
6. Animal-Food	19	138	0.18	<i>canard</i> ‘duck’
7. Attribute-Act	26	118	0.17	<i>politesse</i> ‘courtesy’
8. Institution-Artifact	5	20	0.17	<i>menuiserie</i> ‘carpentry workshop / woodwork’
9. Act-Cognition	49	400	0.16	<i>essai</i> ‘attempt / essay’
10. State-Act	5	27	0.15	<i>équilibre</i> ‘balance / handstand’
11. Body-Artifact	7	73	0.13	<i>jambe</i> ‘leg / pantleg’
12. Body-Person	6	73	0.13	<i>cerveau</i> ‘brain’
13. Cognition-Act	15	137	0.10	<i>disco</i> ‘disco music / dance’
14. Act-Institution	35	400	0.10	<i>administration</i> ‘administration / administrators’
15. Attribute-Cognition	12	118	0.10	<i>absurdité</i> ‘absurdity / nonsense’

Table 7 (Continuation)

Metonymy patterns	n	N	Reg	Example
16. Attribute-Person	11	118	0.10	<i>laideur</i> 'ugliness / unattractive person'
17. Animal-Substance	11	138	0.10	<i>crocodile</i> 'crocodile / crocodile leather'
18. Plant-Artifact	4	50	0.09	<i>gazon</i> 'grass / lawn'
19. Food-Act	6	78	0.09	<i>méchoui</i> 'mechoui'
20. Artifact-Act	39	504	0.08	<i>barre</i> '(high) bar'
21. Act-Artifact*Cognition	27	400	0.07	<i>convocation</i> 'summons / summons letter'
22. Attribute-Artifact	7	118	0.06	<i>résistance</i> 'resistance / resistor'
23. Object-Artifact	2	65	0.05	<i>terrain</i> 'ground / building plot'
24. Act-Possession	19	400	0.05	<i>emprunt</i> 'loan'
25. Act-GroupxPerson	13	400	0.05	<i>défense</i> 'defense / defenders'
26. Artifact-Person	18	504	0.04	<i>violon</i> 'violin / violinist'
27. Artifact-Institution	13	504	0.04	<i>bibliothèque</i> 'bookshelf / library'
28. Artifact-GroupxPerson	11	504	0.04	<i>cabinet</i> 'office / cabinet'
29. Person-Artifact	7	247	0.03	<i>ballerine</i> 'ballerina / ballet slippers'
30. Act-State	19	400	0.03	<i>isolement</i> 'isolation'
31. Person-State	4	247	0.03	<i>gendarme</i> 'military police officer / status'
32. Act-Person	8	400	0.03	<i>apparition</i> 'appearance / apparition'
33. Act-Object	8	400	0.02	<i>projection</i> 'projection / ejection'
34. Act-Feeling	6	400	0.02	<i>excitation</i> 'excitement'
35. Cognition-Institution	5	137	0.02	<i>pharmacie</i> 'pharmacy'
36. Artifact-Cognition	3	504	0.01	<i>carillon</i> 'carillon / chime'
37. Artifact-Body	3	504	0.01	<i>maillot</i> 'swimsuit / pubic area'
38. Artifact-Object	4	504	0.01	<i>fond</i> 'bottom / seabed'
39. Act-Attribute	6	400	0.01	<i>séduction</i> 'seduction / appeal'
40. Cognition-Artifact	2	137	0.01	<i>mécanique</i> 'mechanics / mechanism'
41. Artifact-Food	4	504	0.01	<i>raclette</i> 'scraper / raclette'
42. Artifact-Quantity	6	504	0.01	<i>cuillère</i> 'spoon / spoonful'
43. Artifact-Possession	4	504	0.01	<i>enveloppe</i> 'envelope / cash envelope'
Mean	14	270	0.1	

First, considering the number of supersense alternations associated with polysemy patterns, we observe that metonymies (43 supersense alternations) occur more frequently than metaphors (33 supersense alternations), thus empirically supporting the claim that metonymy is more characteristic of regular polysemy than metaphor. However, when examining the regularity scores, the results reveal no significant difference in the mean scores between metonymy and metaphor (0.1 in both cases). This unexpected outcome warrants at least two important observations.

The first one concerns the semantic granularity used to describe regular polysemy. Supersenses were selected because they have been used to describe large sets of lexical data across several languages. However, supersenses are not uniform in terms of semantic granularity: some correspond to broader semantic classes (e.g., Act, Artifact), while others are more specific (e.g., Animal, Feeling). As a result, some supersense alternations encompass multiple patterns, while others do not. For instance, the Act → Artifact metonymy is linked to several patterns, depending on the role of the referent of the artifact sense within the situation denoted by the act sense. These patterns include:

- Act → Artifact, instrument of the action
(e.g., *équipement* ‘equipping/equipment’)
- Act → Artifact, location where the action takes place
(e.g., *arrêt* ‘stopping/stop’)
- Act → Artifact, result of the action
(e.g., *tatouage* ‘tattoing/tattoo’)

This fixed semantic granularity to define the list of polysemy patterns in the dataset has a clear impact on their estimated degree of regularity, which will need further refinement in future analyses. For instance, the Artifact-Quantity metonymy (Table 7, line 42) appears to have a low degree of regularity (0.01) because it is evaluated across a broad set of 504 nouns primarily referring to manufactured objects. However, the observed regular sense extension is more specifically associated with the Container → Quantity pattern (e.g., *cuillère* ‘spoon / spoonful’, *seau* ‘bucket / bucketful’). In contrast, metaphors appear to involve less specific source meanings.

The second observation concerns the properties of the target meaning in metonymy compared to metaphor. In metonymy, target meanings tend to be generally more frequent and less marked than in metaphor, where a significant proportion of meanings created through analogical extension belong to familiar language. This aspect will also require further exploration in future analyses.

Despite these two observations suggesting that the method used to calculate regularity scores may underestimate the regularity of metonymy patterns compared to metaphor ones, it remains true that, on average, regular metonymy does not tend to produce significantly more polysemy than regular metaphor. This result should at least prompt us to reassess with greater caution the assumed strong correlation between lexical figure and polysemy regularity.

Conclusion

This paper provides a first quantitative assessment of how much more regular metonymy is compared to metaphor. The study is based on a carefully defined, representative sample of nominal sense alternations in French. The results show that metonymy is significantly more regular than metaphor in terms of the number of alternation patterns observed in the analysed sample but that, on average, these patterns produce comparable proportions of polysemic words. The greater propensity of metonymy for regularity is thus nuanced by these results, which call for a measured evaluation of the correlation between figure and degree of regularity. We identify several promising avenues for further investigation to enhance the assessment of regular polysemy. From a methodological perspective, we have highlighted the limitation of relying on judgments from a single annotator. The next step is to collect multiple annotations to measure inter-annotator agreement—both on word sense familiarity, to determine which sense alternations are perceived as natural by speakers, and on identifying the underlying figure behind these alternations, to assess how easily speakers recognize semantic relationships between known senses of the same word. Furthermore, because supersenses provide a level of granularity that is often too coarse, they do not allow for a sufficiently precise analysis of polysemy patterns. Building on the descriptions provided in this study, a more fine-grained analysis will therefore be developed in future work to obtain more precise regularity scores. Finally, we have highlighted the complex relationship between polysemy and morphological derivation, leaving open the question of whether stronger regularities stem from this connection.

References

- Apresjan, Ju. (1974). Regular polysemy. *Linguistics*, 14, 5–32.
- Angleraud, N., Barque, L., & Candito, M. (2025). Annotating the French Wiktionary with supersenses for large scale lexical analysis: a use case to assess form-meaning relationships within the nominal lexicon. In *Proceedings of the 31st International Conference on Computational Linguistics* (pp. 5321–5332). <https://aclanthology.org/2025.coling-main.356/>.
- Barque, L., Haas, P., & Huyghe, R. (2018). Polysémie régulière et néologie sémantique: constitution d'une ressource pour l'étude des sens nouveaux. *Neologica: revue internationale de la néologie*, 12, 91–108. https://hal.science/hal-03823528v2/file/Neologica_polyReg.pdf.
- Buitelaar, P. (1998). *CoreLex: systematic polysemy and underspecification*. Phd Thesis, Brandeis University, Waltham.
- Buitelaar, P. (2000). Reducing lexical semantic complexity with systematic polysemous classes and underspecification. In *NAACL-ANLP 2000 Workshop: Syntactic and Semantic Complexity in Natural Language Processing Systems*, Vol. 1 (pp. 14–19). <https://aclanthology.org/W00-0103.pdf>.
- Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of semantics*, 12(1), 15–67. https://www.researchgate.net/profile/Ted-Briscoe/publication/2609962_Semi-Productive_Polysemy_and_Sense_Extension/links/0912f51370d570a560000000/Semi-Productive-Polysemy-and-Sense-Extension.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail&rtd=e30%3D.
- Dölling, J. (2020). Systematic polysemy. In L. Matthewson, C. Meier, H. Rullmann, T. E. Zimmermann & D. Gutzmann (Eds), *The Wiley Blackwell Companion to Semantics*. Wiley Online Library, Hoboken.
- Falkum, I., & Vicente, A. (2015). Polysemy: Current perspectives and approaches. *Lingua*, 157, 1–16. <https://philarchive.org/archive/FALPCP>.
- Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. MIT press, Cambridge.
- Goossens, L. (2003) 'Metaphtonymy: The interaction of metaphor and metonymy in expressions for linguistic action'. In R. Dirven & R. Pörings (Eds), *Metaphor and metonymy in comparison and contrast*. (pp. 349–377). Mouton de Gruyter, Berlin, New York.
- Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: the comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1534–1543.
- Lombard, A., Huyghe, R., Barque, L., & Gras, D. (2023). Regular polysemy and novel word-sense identification. *The Mental Lexicon*, 18(1), 94–119.

- Lopukhina, A., Laurinavichyute, A., Lopukhin, K., & Dragoy, O. (2018). The mental representation of polysemy across word classes. *Frontiers in psychology*, 9, 192.
- Maudslay, R. H., Teufel, S., Bond, F., & Pustejovsky, J. (2024). ChainNet: Structured Metaphor and Metonymy. In WordNet. *Proceedings of LREC-COLING 2024* (pp. 2984–2996). ELRA and ICCL, Torino. <https://arxiv.org/pdf/2403.20308>.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2 : A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516–524. <https://link.springer.com/content/pdf/10.3758/BF03195598.pdf>.
- Nunberg, G., & Zaenen, A. (1997). La polysémie systématique dans la description lexicale. *Langue française*, 12–23.
- Peters, W. (2006). In Search for More Knowledge: Regular Polysemy and Knowledge Acquisition. In P. Sojka, K.-S. Choi, C. Fellbaum & P. Vossen (Eds.), *Proceedings of GWC 2006* (pp. 245–251). Masaryk University, Brno.
- Polguère, A. (2018). A lexicographic approach to the study of copolysemy relations. *Russian Journal of Linguistics*, 22(4), 788–820.
- Polguère, A. (2022). Fishing out regular polysemy with the Gener (generic) relation. *Lexique*, 27–54. <https://www.peren-revues.fr/lexique/856>.
- Pustejovsky, J. (1995). *The Generative Lexicon*, MIT Press, Cambridge.
- Salvadori, J. (2024). *L'ambiguïté des noms déverbaux en français. Une étude quantitative du sens construit*. Phd Thesis, University of Fribourg, Fribourg.
- Srinivasan, M., & Rabagliati, H. (2015). How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua*, 157, 124–152. https://escholarship.org/content/qt69z8f7qf/qt69z8f7qf_noSplash_426afa3377f6db5afe571954392fa2ff.pdf.
- Vicente, A., & Falkum, I. L. (2017). Polysemy. In: M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press, Oxford.
- Xu, Y., Duong, K., Malt, B. C., Jiang, S., & Srinivasan, M. (2020). Conceptual relations predict colexification across languages. *Cognition*, 201, 104–280.
- Yurchenko, A., Lopukhina, A., & Dragoy, O. (2020). Metaphor is between metonymy and homonymy: Evidence from event-related potentials. *Frontiers in Psychology*, 11, 2113.



Sergey Say

University of Potsdam
Germany

 <https://orcid.org/0000-0001-8066-9166>

Ilja A. Seržant

University of Potsdam
Germany

 <https://orcid.org/0000-0002-8066-9251>

A Frequency-Based Algorithm for Argument Extraction from Russian Treebanks

Abstract

Arguments, unlike adjuncts, are typically understood as verb-specific dependents, which includes the fact that the morphosyntactic devices used for argument encoding are determined by individual verbs. Building on this observation, we operationalize arguments as dependents whose encoding device occurs with a given verb at a significantly higher-than-average frequency. We apply an argument extraction algorithm to a dataset of 132,221 verb dependents from Russian treebanks available in the Universal Dependencies (UD) platform. To evaluate the algorithm's performance, we compare its results to a manually annotated subset, informed by *The Active Dictionary* and a detailed semantic understanding of argumenthood. The frequency-based algorithm achieves acceptable precision (approx. 0.83), with particularly few false positives, making it a promising tool for cross-linguistic applications in typologically diverse languages with UD treebanks. Theoretically, we argue that a quantitative distributional approach to valency—originally proposed in Ju. D. Apresjan's early pioneering work—broadly aligns with the in-depth semantic analyses of individual verbs and their meanings found in his later works, including *The Active Dictionary*.

Keywords

Russian, argument, treebank, frequency, case, preposition, dependency

1. Objectives and Approach¹

As early as 1965, Jurij Derenikovich Apresjan proposed a hypothesis stating that “there is a regular correspondence between the syntactic properties of words and their semantic features” (Apresjan, 1965, p. 51). To the present day, this idea remains a cornerstone in the study of verbs and their valency across various approaches and perspectives (Helbig & Schenkel, 1983, pp. 61–62; Levin, 1993; Lazard, 1994, p. 133; Levin & Rappaport Hovav, 2005; Malchukov & Comrie (eds.), 2015, etc.).

Despite the apparent appeal of this claim, it immediately presents a challenge in selecting analytical tools. One extreme is to rely on the syntactic properties of verbs, such as their combinatorial potential, to infer claims about their meanings. For example, any Russian verb that can occur with *v* ‘in’ plus the accusative case might be interpreted as a kind of motion verb. This approach benefits from a potentially solid empirical foundation, including frequency data, but its drawback is that such inferences must be treated with caution and ideally verified through independent semantic analysis. The other extreme is to make meticulous judgments about verb meanings and trace the mechanisms by which syntactic patterns relate to semantic nuances. The main issue with this approach is that meanings are not directly observable and inevitably remain a theoretical construct.

While numerous insightful studies fall somewhere along the spectrum between these two extremes, Apresjan’s contribution to the field is exceptional in that he made influential advances spanning the entire range of possible approaches to valency throughout his career. In his early work, he explored the extraction of semantic information from verbs’ distributional properties, employing a wide array of quantitative techniques (Apresjan, 1965; 1967). Over time, he shifted toward the in-depth semantic end of the continuum, placing increasing emphasis on thought experiment (*myslennyj èksperiment*, see Apresjan & Páll, 1982, p. 39) and semantic decomposition (*tolkovanie*, Apresjan, 1974; 1995). This shift was also reflected in his extensive work on dictionaries (Mel’čuk et al., 1984; Apresjan (Ed.), 2004), culminating in *The Active Dictionary*, an endeavor launched

¹ The study reported here was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 317633480 – SFB 1287. We are indebted to the two anonymous reviewers, as well as to Peter Arkadiev, Maria Ovsjannikova, Anastasia Panova, Dmitri Sitchinava, Aigul Zakirova, and all participants of the Slavic Linguistics Department colloquium at the University of Potsdam for their comments on an earlier version of this paper, which we took into account during the revision. The usual disclaimers apply.

in 2014 and continuing to the present day, even after Jurij Derenikovich's passing (Apresjan (Ed.), 2014-).

This shift likely reflects Apresjan's growing dissatisfaction with the computational distributional approach of his early work, leading him to conclude that detailed semantic analysis is more accurate, insightful, and ultimately superior. While this may be true, lexicographic approaches to valency based on semantic decomposition have two inherent limitations (Sarkar & Zeman, 2000, p. 691). First, no dictionary can cover all verbs and their variable usage in actual texts—a limitation that has become even more apparent with the advent of large corpora. Even for well-documented languages like Russian, dictionaries inevitably have a restricted scope. Second, an approach relying on nuanced intuitions is typically feasible only for a researcher's native language and is impractical for most of the world's languages, especially those without strong lexicographic traditions.

Given these considerations, combining the quantitative distributional perspective with the semantics-oriented lexicographic approach remains essential for advancing the study of verb-dependent relationships. Our paper follows this approach to address the longstanding problem of distinguishing arguments from adjuncts — often considered central to the automatic extraction of valency frames from corpora (Sarkar & Zeman, 2000, pp. 691–692). Specifically, we propose a co-occurrence-based quantitative technique for automatically differentiating arguments from adjuncts and apply it to Russian treebanks from the Universal Dependencies (UD) project (de Marneffe et al., 2021), thereby aligning our approach with Apresjan's early methodology. We then compare the results with the treatment of argumenthood in *The Active Dictionary* (Apresjan (Ed.), 2014-), the hallmark of his later approach.

The results presented below are valuable in their own right and, hopefully, contribute to our understanding of valency from a token-based perspective. More importantly, however, the technique introduced here paves the way for its application to languages that lack detailed, semantically oriented accounts of their verbal lexica but have UD treebanks available. In this sense, our study serves as a preparatory step for a larger research project aimed at cross-linguistic comparison of valency class systems from a token-based perspective.

The paper is structured as follows. Section 2 examines the argument–adjunct distinction and key argumenthood criteria. Section 3 outlines our data and methodology, detailing the algorithm for distinguishing arguments from adjuncts based on co-occurrence frequencies in treebanks. Section 4 evaluates the algorithm's performance and theoretical implications. Finally, Section 5 provides a brief summary and outlook.

2. The argument-adjunct distinction: state of the art

Since at least Tesnière's "Éléments de syntaxe structurale" (1959), it has been widely recognized that some verb dependents are more closely associated with the verb than others (Lazard, 1994; Dixon, 2009, pp. 97–128). A textbook example of this distinction appears in (1), where *ona* 'she' and *našim otdelom* 'our department' represent participants inherent to the meaning of *rukovodit* 'manage', while the adverbial phrase *s sentjabrja* 'since September' is not essential to the verb's meaning.

(1) *Ona rukovodit našim otdelom s sentjabrja.*

'She has been managing our department since September.'

While intuitively appealing, this distinction is far from unproblematic, even in terminology. The common English terms *arguments* and *adjuncts* are not universally accepted (see Frajzyngier et al., 2024 for an overview of alternatives), and mapping terms across linguistic traditions—e.g., *actants* vs. *circonstants* in French or *aktanty* vs. *sirkonstanty* in Russian—further complicates the picture. The real challenge, however, lies not in the terminology but in identifying suitable criteria for distinguishing verbal dependents.² The numerous formal and semantic criteria used to separate arguments from adjuncts often yield conflicting classifications (see the discussion of problematic patterns in Russian in Plungjan & Raxilina, 1998; Muravenko, 1998 and in other articles in the same issue of *Semiotika i informatika*). As a result, some scholars argue that a rigid two-way classification of verbal dependents is neither feasible nor necessary in typological research (Jacobs, 1994; Haspelmath, 2014, p. 9). Despite these caveats, we will use *arguments* and *adjuncts* as the standard terms. Before presenting our perspective, we briefly highlight key distinctions from the literature.

The most widely cited contrast between arguments and adjuncts, in both the Moscow Semantic School and beyond, is that arguments correspond to a predicate's inherent participants—essential to its meaning in terms of semantic decomposition (*tolkovanie*) (Apresjan, 1974; Boguslavskij, 1996, p. 23). Variations on this idea can be found in Van Valin (2001, p. 93) and Frajzyngier et al. (2024). A classic example is Apresjan's decomposition of *arendovat*'

² In this text, we use the term *dependents* as a cover term for arguments and adjuncts. In corpus linguistics, a similar notion is sometimes referred to more broadly as *complements* (Schulte im Walde 2009), although this usage may differ from traditional syntactic definitions.

'rent': "A rents C" means that in exchange for compensation D, person A acquires from person B the right to use property C for a period T" (Apresjan, 1995, Vol. 1, p. 120). While this approach often yields clear results, it is not without issues—for instance, the verb *arendovat'* involves five potential arguments, but not all are equally essential (e.g., property C must be identifiable, while compensation D may remain unspecified). Since semantic decomposition is a theoretical construct, criteria for determining a verb's inherent elements can vary (Testelec, 2001, p. 168–178).

Apart from the central but elusive contrast rooted in the semantic decomposition of verb meaning, argumenthood criteria and the corresponding approaches fall into two broad categories: semantic and syntactic. A common semantic strategy is to classify dependents by roles (agents, patients, experiencers, locations, etc.³). While some correlations exist — agents are typically arguments, while places and causes often are not—this is not a panacea. For instance, instruments can behave as either arguments or adjuncts, forming a continuum (Koenig et al., 2008; Bohnemeyer, 2007). More nuanced distinctions have been proposed, such as Jolly's (1993) view that adjuncts, as modifiers, are hybrids: they express semantic roles like arguments but also predicate information of their own. While such insights are valuable, they are difficult to operationalize, making them less useful for tasks like data annotation and cross-linguistic comparison.

Syntactic criteria might seem more reliable, with obligatoriness being an obvious candidate. While arguments are generally more obligatory than adjuncts, this is not a universal test, as strict syntactic obligatoriness is largely illusory (Helbig & Schenkel, 1983, pp. 35–36). Most languages allow omitting participants under certain conditions (see Ljaševskaja & Kaškin, 2015, p. 502 for Russian), and even non-obligatory prepositional phrases can function as arguments when they are present (Jolly, 1993, p. 283). Moreover, languages vary greatly in how freely they permit omission, further limiting obligatoriness as a cross-linguistic criterion.

Closely tied to argumenthood criteria is the distinction between *core* and *oblique* dependents. While no universal definition exists,⁴ "oblique" typically

³ A key problem with this approach is the lack of a universally accepted set of semantic roles, let alone a reliable method for classifying verbal dependents into discrete roles. See Bickel et al. (2014) for an overview of the challenges and an empirical alternative, which results in fuzzy clustering that doesn't apply to some types of dependents.

⁴ In Role and Reference Grammar, "oblique" refers to core arguments marked by an adposition or "oblique case" (a morphological term that is, by the way, not typologically unproblematic), such as *to Bill* in *Harry gave the key to Bill* (Van Valin, 1993, pp. 40–41; Beavers, 2010). Others use "oblique" more broadly, to refer to adjuncts or adverbial modifiers (Dryer & Gensler, 2013).

refers to dependents that, despite being inherent to a verb's meaning, are morphosyntactically similar to adjuncts — such as prepositional phrases in *graduate from Harvard* or *apologize for the mistake*. This approach is useful for language-specific analysis but poses challenges for cross-linguistic comparison due to variation in case systems, adpositional usage, and verb indexing strategies.

Thus, defining the argument–adjunct distinction through simple formal contrasts is problematic (Haspelmath, 2014). However, deeper syntactic contrasts may be more useful for cross-linguistic studies. For instance, Helbig & Schenkel (1983, p. 38) highlight how German verbs like *wohnen* ‘live, reside’ and *sterben* ‘die’ interact differently with locative phrases. While *Er wohnte in Dresden* ‘He lived in Dresden’ and *Er starb in Dresden* ‘He died in Dresden’ appear structurally similar, their internal organization differs, as shown in (2) and (3), taken from (Helbig & Schenkel, 1983, p. 38).

(2) **Er wohnte, als er in Dresden war*, literally ‘He lived when he was in Dresden’.

(3) *Er starb, als er in Dresden war* ‘He died when he was in Dresden’.

Helbig & Schenkel (1983, p. 38) classify locative phrases with *wohnen* ‘live’ as “tightly bound verb complements” (*enge Verbergänzung*) and with *sterben* ‘die’ as “free verb complements” (*freie Verbergänzung*), aligning with the argument–adjunct distinction. While insightful, such syntactic tests are not universally applicable (Haspelmath, 2014).

Apart from morphosyntactic differences, arguments are inherently “verb-specific and thus have to be learned together with each verb, whereas the use of adjuncts is independent of particular verbs” (Haspelmath, 2014, p. 5; see also Beavers, 2010, p. 842). A key aspect of argument verb-specificity is that verbs typically require specific coding devices, such as cases and adpositions (Testelec, 2001, p. 187; see also Jacobs, 1994 on formal specificity as a dimension of valency). This property, known as “subcategorization” (syntactic combinability), coexists with “selection” (semantic combinability). Its lexical nature is evident in (4), where each Russian verb, despite semantic similarities, follows a distinct valency pattern: a scheme that links semantic participants to syntactic slots marked by specific encoding devices.

(4) a. *Petja simpatiziruet Maše*.

‘Petja likes Masha’ (the nominative + dative pattern; here and below, the coding device for the experiencer is mentioned first).

- b. *Petja vosxiščaetsja Mašej.*
'Petja admires Masha' (the nominative + instrumental pattern).
- c. *Petja vľjubilsja v Mašu.*
'Petja fell in love with Masha' (the nominative + v 'in' accustive pattern).
- d. *Petja ľjubit Mašu.*
'Petja loves Masha' (the nominative + accusative pattern).
- e. *Petja razočarovalsja v Maše.*
'Petja is disappointed in Masha' (the nominative + v 'in' locative pattern).
- f. *Pete nadoela Maša.*
'Petja is fed up with Masha' (the dative + nominative pattern).

In contrast, the encoding of adjuncts is determined by their own meaning and lexical properties, with the head verb playing little role. This is shown in (5), where each adjunct follows a distinct coding pattern but can combine not just with *videt'sja* 'see each other' but with any Russian verb compatible with temporal adverbials (*priexat* 'arrive', *ženit'sja* 'marry', etc.).

- (5) a. *My videlis' vtorogo sentjabrja.*
'We saw each other on September 2nd' (the genitive pattern).
- b. *My videlis' vo vtornik.*
'We saw each other on Tuesday' (the v 'in' + accusative pattern).
- c. *My videlis' na prošloj nedeľe.*
'We saw each other last week' (the *na* 'on' + locative pattern).
- d. *My videlis' na Pasxu.*
'We saw each other on Easter' (the *na* 'on' + accusative pattern).
- e. *My videlis' prošlym letom.*
'We saw each other last summer' (the instrumental pattern).

Since arguments are verb-specific, unlike adjuncts, their documentation and analysis have largely been a lexicographic task, as seen in numerous dictionaries by Apresjan and colleagues (Apresjan & Páll, 1982; Mel'čuk et al., 1984;

Apresjan (ed.), 2004; Apresjan (ed.), 2014–). While theoretical linguistics sometimes offers gradual approaches to argumenthood with various criteria, lexicographers take a practical stance: dictionaries list verbs with a discrete set of dependents and their encoding patterns, sometimes grouped into larger sets (Zolotova 2006). This also applies to lexical databases like FrameNet and its derivatives, which focus on arguments rather than adjuncts. A case in point is FrameBank, based on Russian data (Ljaševskaja & Kaškin, 2015, p. 502).

Most approaches view the link between verbs and specific coding devices, such as cases and adpositions, as a **typical property** of arguments rather than their defining feature. Since dictionaries typically lack frequency data and do not capture usage variability, they do not emphasize this link. However, it plays a key role in the automatic identification of arguments based on their co-occurrence with verbs in corpora (see Korhonen et al., 2000; Schulte im Walde, 2009 and references cited there).

To summarize, approaches to argumenthood vary widely, and no agreed-upon procedure exists for distinguishing arguments from adjuncts. This paper does not aim to propose a new or “best” definition of arguments. Instead, we take a more modest approach: we explore argument-coding specificity as a relatively accessible, corpus-quantifiable property of verbal dependents and compare our frequency-based extraction technique to established semantics-based lexicographic approaches to Russian.

3. Data and methods

3.1. Data extraction and preannotation

The data for this study come from Universal Dependencies (UD), a collection of treebanks covering about 150 languages (Nivre et al., 2020; Zeman et al., 2022). Here, we focus on Russian, though the methodology applies to any UD treebank. All the spreadsheets mentioned in Section 3, as well as the code used to analyze the data, are available as the Supplementary materials at https://github.com/serjzhka/Algorithm_Argument_Extraction_Neophilologica.

UD treebanks have several properties that make them well-suited for quantitative valency studies. First, they are based on naturalistic texts (fiction, blogs, news, etc.). Second, the Russian UD treebanks are large (see below for details) and continue to grow. Third, UD treebanks provide deep morphological and

syntactic annotation, consistently applied across diverse languages, facilitating token-based typological analysis. Specifically, UD annotation includes lemmatization, allowing automatic tracing of usage patterns across morphological forms and syntactic contexts. Finally, as their name suggests, UD treebanks explicitly mark dependency relations, categorized into a concise set of universal types (e.g., “nominal subject”, “indirect object”, “adverbial modifier”).⁵

UD treebanks also have inherent limitations. The most significant one may affect our study’s planned typological extension: UD is heavily biased toward “major” languages, limiting typological diversity. “Minor” languages are under-represented, and existing corpora vary in size, content, and quality.

Additionally, UD annotation presents technical challenges for studying verbal arguments. The framework treats function words like adpositions and auxiliaries as dependents of their associated content words (de Marneffe et al., 2021, p. 269), diverging from standard dependency grammar. For instance, in *give the toys to the children*, *give* takes *toys* as a direct object (“obj”) and *children* as an oblique object (“obl”), with *to* analyzed as a dependent of *children* under a “case” relation. This approach enhances cross-linguistic comparability between case-rich languages and those relying on adpositions. However, it complicates the automatic extraction of argument-encoding devices associated with specific verbs, as discussed below.

As a first step, we extracted all dependents tagged as direct objects (“obj”), indirect objects (“iobj”), or oblique objects (“obl”) of finite verbs in all Russian treebanks available as part of the UD collection (“Taiga”, “SynTagRus”, “GSD”, and “PUD”)⁶. We used UD version v2.11, published on November 15, 2022 (Zeman et al., 2022). The total size of the corpus we used is 1744K tokens. For data extraction, we used a Python script developed by Ivan Seržant. The script generates a spreadsheet where each row represents a nominal or pronominal dependent annotated for 27 parameters. The key parameters for this analysis are:⁷

⁵ An important question is the exact procedures used for these annotations. Unfortunately, the UD documentation is not fully explicit, and minor inconsistencies suggest the procedures varied across treebanks.

⁶ We did not extract canonical nominative subjects (tagged as, e.g., “nsubj” in UD), as they are unequivocally arguments and thus irrelevant for our analysis. Notably, UD takes a conservative approach to subjecthood: dependents often described in the literature as non-canonical subjects (e.g., accusative NPs in clauses like *menja tošnit* ‘I feel nauseous’) are analyzed as objects or obliques—and therefore were included in our dataset. The algorithm skipped adverbial expressions like *tuda* ‘there (directional)’, tagged as “advmod”, even if distributionally similar to *v gorod* ‘to the city’.

⁷ Additional parameters related to word order, verb morphology, and animacy are not relevant here but are available in the Supplementary materials and will be used elsewhere.

- “sentence”: the full sentence with the target dependent and head verb.
- “verb lemma”: infinitive form of the head verb.
- “object form”: inflected form of the dependent.
- “object dependency relation”: UD treebank tag for the dependent, e.g. “iobj”.
- “object case”: case marking of the dependent (e.g., “Acc”, “Dat”, “Gen”).
- “adposition lemma”: any adposition identified as a dependent of the target dependent.

Table 1 presents the annotation of a single raw entry from our spreadsheet as an example (displayed as a column rather than a row for technical reasons).

Table 1.

Sample spreadsheet entry (8 selected parameters)

entry no	23959
corpus	ru_syntagrus-ud-dev.conllu
id	2020_Corpus2_0Khochu_byt_negrom.xml_304
sentence	Я смотрю на круглое личико, и мне кажется, что это – она.
verb lemma	смотреть
finite verb	смотрю
object dependency relation	obl
object form	личико
object case	Acc
adposition lemma	на

The spreadsheet generated by the data extraction script contained 132,221 entries, where each entry corresponds to a verb dependent token. Next, we annotated the encoding devices associated with these dependents. By default, this was a combination of case form and adposition (if present), e.g., “naACC” for the entry in Table 1. However, due to UD annotation errors—such as misdisambiguation of homonymous case forms or incorrect lemmatization of prepositions—parts of the process were manual.⁸ Ultimately, we identified 92 distinct encoding devices, twice the number found in Apresjan (1965, p. 46), largely due to the presence of rare prepositions in UD treebanks, such as *posredine* ‘in the middle (of)’ or *vzamen* ‘in exchange (for)’. The resulting spreadsheet

⁸ At this stage, we either reannotated or excluded all examples tagged as nominative in the original UD annotations, as these were either annotation errors or irrelevant for our purposes (e.g. when used in citations).

(*Russian_data_with_encoding*), which served as the input for the subsequent analysis, is available in the Supplementary materials.

As the next step, we removed entries where the encoding device could not be identified (e.g., indeclinable elements), marking them with <NA> tags. We also excluded nominals that were not dependents at the verb phrase level, such as the expletive *èto* ‘this’ in sentences like *No èto my opjat’ sil’no zabegaem* ‘But here we’re getting ahead of ourselves again’, marking them with the “EXPL” tag. After filtering out these entries, the number of tokens available for analysis was reduced to 122,551 entries.

Table 2 presents partial annotations, including the “encoding device”, for three entries from the sentence *Posle ètogo ljudi rasskazyvali zalu o rezul’tatax* ‘After that, people told the audience about the results.’

Table 2.
Partial annotations of three sample entries, including “encoding device” tags

entry no	verb lemma	object dependency relation	object form	object case	adposition lemma	encoding device
2386	рассказывать	obl	этого	Gen	после	posleGEN
2387	рассказывать	iobj	залу	Dat		DAT
2388	рассказывать	obl	результатах	Loc	о	oLOC

As noted earlier, UD explicitly rejects the argument-adjunct distinction, opting instead for the core-oblique distinction. The UD documentation justifies this choice: “... the argument/adjunct distinction is subtle, unclear, and frequently argued over ... the best practical solution is to eliminate it ... The core-oblique distinction is ... both more relevant and easier to apply cross-linguistically than the argument-adjunct distinction” (<https://universaldependencies.org/u/overview/syntax.html>; retrieved March 1, 2025). In UD, this distinction is based on the morphosyntactic encoding of dependents (de Marneffe, 2021, p. 268), partially overlapping with parameters like morphological case and disregarding whether a dependent is obligatorily selected by a specific verb.

While reluctance to provide discrete argument vs. adjunct annotation is understandable, distinguishing “obj”, “iobj”, and “obl” dependents is not entirely satisfactory either. The UD documentation offers only vague definitions, such as “iobj” as a “nominal core argument of a verb that is not its subject or (direct) object” or “obl” as “a nominal functioning as a non-core (oblique) modifier of a predicate” (de Marneffe, 2021, p. 266). Beyond their vagueness, these categories are inconsistently applied. The “obj” tag typically marks accusative prepositionless

objects, yet 2,369 of 37,691 “obj” entries involve diverse encoding strategies with unclear rationale. Often, these cases involve verbs sometimes classified as transitive due to features like passive formation (Kamynina, 1999, p. 146; but see Fowler, 1996 for a different approach). However, annotation remains inconsistent—e.g., the instrumental *rukami* ‘(with) arms’ as a dependent of *maxat* ‘wave’ appears variably as “obj”, “iobj” and “obl”.

The “obj” vs. “obl” distinction is more meaningful for accusative prepositionless dependents: “obj” mainly marks clear direct objects, while “obl” (and its subtype “obl:tmod”) typically indicate adverbial time or frequency modifiers, such as *vsju nedelju* ‘the whole week’ or *každyju minutu* ‘every minute.’ Though not entirely consistent, this distinction helps differentiate direct objects from adverbials, a contrast not captured by other parameters.

3.2. Argumenthood annotation algorithm

In the previous section, we showed that UD-internal tags do not distinguish arguments from adjuncts. This challenge forms the basis of our analysis, which aims to develop an algorithm for automatic argumenthood annotation.

The core principle of our algorithm, introduced in Section 2, is that argument-encoding devices—such as cases and prepositions in Russian—are verb-specific and quantitatively distinct from adjuncts, whose distribution is governed by semantic compatibility and internal structure. Based on this, we operationalize argumenthood as follows: a dependent is interpreted as an argument if the relative frequency of its encoding device with a given verb exceeds a certain baseline, as specified below. This approach treats argumenthood as a gradable phenomenon but can be converted into a discrete distinction by setting a cut-off point for the frequency difference.

This approach closely resembles Apresjan’s concept of “government strength” (*sila upravljenija*), defined as “the proportion of cases in which a given verb occurs with a specific (prepositional-)case form out of the total occurrences of the verb” (Apresjan, 1965, pp. 51–52), where adjuncts are dependents with very low government strength. Comparing observed frequency to a baseline is a standard method in subcategorization frame acquisition (Korhonen et al., 2000; Schulte im Walde, 2009). The main challenge is determining this baseline. A simple approach sets a fixed threshold (e.g., >20% of a verb’s occurrences), while we use a more refined method, comparing observed frequency to expected frequency, which is calculated under the assumption that the encoding device occurs equally often with all verbs in the corpus (see Sarkar & Zeman, 2000 for a similar idea as applied to Czech

data and Schulte im Walde 2009, p. 957 for a brief overview of available methods). Our approach avoids “penalizing” low-frequency coding devices, aiming instead to detect all cases where individual verbs attract specific constructions. For example, the combination *nad* ‘above’ with the instrumental case occurs 38 times with the verb *rabotat* ‘work’ (as in *rabotat nad proektom* ‘work on a project’) — just 8% of the 467 clauses with this verb. Yet this is high relative to the baseline frequency of *nad* + instrumental, which appears with only about 0.4% of all verb tokens. This discrepancy reflects the intuitively correct idea that *nad* + instrumental is a lexically specified pattern for expressing the domain or object of work.

This concept was implemented in R (R Core Team, 2021). The procedure was as follows. Based on the raw dataset from Section 3.1 (*Russian_data_with_encoding*), we created a spreadsheet (*valency_frames*) where each row represents a finite verb token in the treebank, and columns correspond to the 92 attested encoding devices (e.g., “ACC”, “INS”, “naACC”). Cells contain “0” or “1” depending on whether the specific verb token has a dependent in the specified (prepositional-) case form. For example, the three entries in Table 2 are merged into one row with “1” in the columns for “posleGEN”, “DAT”, and “oLOC”. The *valency_frames* spreadsheet includes 87,797 entries, matching the number of verb tokens in the raw data. It was also used to calculate verb lemma prevalence; for instance, *rasskazyvat* ‘tell (IPFV)’, shown in Table 2, appears 153 times in the dataset.

Next, we created a summary spreadsheet (*verbs_and_encoding_devices*) grouping co-occurrences with encoding devices by verb lemmas. The 7,538 rows represent distinct verb lemmas, while the 92 columns indicate the prevalence of encoding devices co-occurring with each verb. For example, for *rasskazyvat* ‘tell (IPFV)’, which appears 153 times in the data, 76 instances involve a dependent encoded by o ‘about’ + locative, 58 by the dative case, and only 2 by *posle* ‘after’ + genitive (as seen in Table 2), along with additional dependent types.

The figures in this spreadsheet represent observed frequencies. Expected frequencies were calculated by determining the overall relative prevalence of each encoding device across all verb tokens in the corpus and multiplying it by the total number of tokens for each verb lemma. For example, dependents in the prepositionless dative case appear in 7,929 out of 87,797 clauses, yielding a ratio of 0.09. Under the null hypothesis that *rasskazyvat* ‘tell (IPFV)’ selects the dative case at the same rate as other verbs, we would expect about 13.8 occurrences ($= 153 \times 0.09$). The observed count (58) far exceeds this, indicating a stronger-than-average preference for the dative case with *rasskazyvat*.

As the final stage in the analysis, we needed a procedure to infer argumenthood status from the differences between observed and expected frequencies. We used a two-step algorithm:

- If the expected frequency of a verb-encoding combination was ≥ 5 , we applied the χ^2 -test. The combination was classified as an argument relation if the observed frequency exceeded the expected frequency and the p-value was < 0.05 .
- Otherwise, we used Fisher's exact test with the same p-value threshold (< 0.05) and an additional condition that the observed count was at least 3.

This procedure is largely similar to two subcategorization frame acquisition methods discussed by Korhonen et al. (2000). The key difference is the use of Fisher's exact test to address issues with low-frequency items, which make the χ^2 -test unreliable (Korhonen et al., 2000, p. 205).⁹ The additional requirement that the observed frequency exceed 2 helps prevent premature argumenthood classification for rare combinations.

Using this procedure, we assigned a binary argument vs. non-argument status to each verb-encoding combination, following the standard discrete approach to argumenthood (Levin & Rappaport Hovav, 2005). For example, *rasskazyvat'* 'tell (IPFV)' was classified as selecting arguments marked by the dative case, o 'about' + locative, and *pro* 'about' + accusative.

Argumenthood judgments are summarized in a spreadsheet (*argumenthood_df*), where "1" denotes arguments and "0" adjuncts. Like *verbs_and_encoding_devices*, it contains 7,538 rows (verb lemmas) and 92 columns (encoding devices). However, this format is not very clear visually. Instead of presenting it directly, Table 3 shows an excerpt combining data from two spreadsheets: it shows observed frequencies of some verb-encoding combinations from *verbs_and_encoding_devices*, with boldface indicating "1"s from *argumenthood_df*, marking algorithmically identified arguments.

Table 3.

Selected verb – encoding device combinations frequencies and their argumenthood

		INS	ACC	DAT	vLOC	sGEN	nadINS	dljaGEN
<i>представлять</i>	'represent'	33	244	28	22	1	0	14
<i>иметь</i>	'have'	7	794	1	102	10	1	24
<i>посоветовать</i>	'advise'	0	22	15	1	0	0	1
<i>называть</i>	'call'	109	273	0	33	2	0	0

⁹ The main downside of Fisher's exact test is that it is computationally demanding, which makes it poorly suited for across-the-board application.

Table 3 (Continuation)

		INS	ACC	DAT	VLOC	sGEN	nadINS	dljaGEN
<i>подвергнуться</i>	‘undergo’	0	1	21	3	1	0	0
<i>вынудить</i>	‘force’	0	13	0	1	1	0	0
<i>закончить</i>	‘finish’	5	50	0	14	0	1	0
<i>работать</i>	‘work’	38 ¹⁰	19	0	193	22	38	5
<i>становиться</i>	‘become’	256	2	1	34	5	0	20

The spreadsheet *argumenthood_df*, which annotates each verb-encoding combination based on the quantitative argumenthood test, is valuable in its own right and can serve as a reference for future analyses.

However, at this stage, we encountered a systematic complication, exemplified by the verb *rasskazyvat’* ‘tell’ discussed above. Intuitively, this verb is transitive, with accusative dependents like *istorii* ‘stories’ functioning as clear arguments.¹¹ Indeed, *rasskazyvat’* appears with an accusative dependent 28 times in our data. Yet, under our frequency-based algorithm, accusative dependents of *rasskazyvat’* did not meet the quantitative argumenthood criteria: their frequency was not significantly higher than expected, reflecting the verb’s diverse valency patterns (see Table 2). This negative result is purely mathematical—since transitive verbs are common, the expected co-occurrence rate with an accusative object is high, complicating algorithmic identification (Korhonen et al., 2000, p. 204).

This issue is even more pronounced in languages with many labile verbs, where overt direct objects appear less frequently than in strictly transitive verbs. To address this, we made an exception when integrating algorithmic argumenthood judgments into our raw data: every accusative dependent was initially classified as an argument (see Section 4.3 for an additional correction). This decision reflects the special status of the basic transitive pattern in most languages, encompassing a broad range of verbs centered on action (Lazard, 1994, pp. 134–135). Many

¹⁰ The frequency of *rabotat’* with an instrumental-case dependent did not surpass the argumenthood threshold, contrary to standard views: many such combinations involve professions (e.g., *rabotat’ povarom* ‘work as a cook’) and would count as arguments in decomposition-based approaches. In principle, the algorithm could be refined by factoring in grammatical animacy — *rabotat’* with animate instrumentals is unusually frequent — but we deliberately avoided overfitting, opting for a simpler, cross-linguistically more transferable model.

¹¹ The issue of temporal adverbials expressed by prepositionless accusatives, such as *vsju nedelju* ‘the whole week’, will be discussed separately in Section 4.3.

linguists argue that case assignment in canonical transitives is structural rather than lexical (Yip et al., 1987, p. 222; see also de Marneffe et al., 2021, p. 267).

In any event, our binary annotation remained fully automatic and content-agnostic—aside from accusative dependents, which were assigned “1” by default, all other combinations were tagged “1” or “0” based on the frequency-based algorithmic test.

The resulting spreadsheet (*data*) integrates UD treebank annotations, Python-generated annotations, semi-manual encoding device annotations, and algorithmic binary argumenthood classifications. As discussed in Section 3.1, the number of verb-dependent tokens suitable for further analysis is 122,551. In Section 4, we evaluate the quality of our argumenthood classification.

3.3. Incorporating lexicographic annotations

To assess the algorithm’s performance (Section 3.2), we compared its results with annotations based on a qualitative item-by-item analysis of raw entries, distinguishing arguments from adjuncts (this is a standard evaluation procedure in automatic acquisition of verb frames from corpora, see Schulte im Walde, 2009, p. 958). For this, we relied on two dictionaries: Rozental’s *Upravlenie v russkom jazyke* (*Government in the Russian Language*) (Rozental’, 1986) and *The Active Dictionary*, initiated by Ju.D. Apresjan (Apresjan (Ed.), 2014–).¹² These sources differ in scope, target audience, and approach to argumenthood.

Rozental’s dictionary is a practical, prescriptive reference that lacks detailed analysis of argumenthood. It focuses on verbs with variable or problematic valency, omitting frequent, prototypically transitive verbs like *ubivat’* ‘kill’, while including more complex derivatives like *ubivat’sja* ‘grieve, mourn bitterly.’ For covered verbs, it provides typical valency patterns, often with brief explanations of relevant meanings.

In contrast, *The Active Dictionary* is a comprehensive academic work that reflects Apresjan and colleagues’ views on argumenthood, semantic decomposition, and combinatorial potential. Each entry provides detailed information on meaning, syntax, and usage, including a standardized representation of government (*upravlenie*). Importantly, government patterns are provided separately for each of the verb’s meanings. To compare the two sources, we included one sample dictionary entry from each of them in the Appendix at the end of the article.

¹² The annotations discussed in this section were manually prepared by Vera Arbieva Pais, to whom we express our sincere gratitude.

The dictionary-based annotations relied on whether the dictionaries listed the relevant encoding devices for the specific meaning of the given verb. Since this process was manual and required semantic analysis, it was time-consuming and impractical for the entire dataset. Ultimately, we obtained annotations for the first 5,423 entries in our *data* spreadsheet, ordered alphabetically by verb lemma. Unlike the algorithmic approach, this manual procedure was token-based — meaning that the same combination of a verb lemma and an encoding device could be classified as an argument in one entry and as an adjunct in another, depending on the meanings of the verb and its dependent in the specific sentence.

The annotations distinguish between “1” (the usage pattern is listed in the verb’s government pattern and considered an argument), “0” (not listed and not an argument), and <NA> (verb not included in the dictionary). Additional tags, detailed in the Supplementary materials, appear in separate columns. The most important is “s” (for *semantic argument*), used when a syntactic dependent fulfills a semantic valency in the verb’s decomposition but the specific form is not listed in its government pattern. An example is entry 44925, shown in (6).

- (6) *V otvet Irina Vasil'evna rasstegnula vorotnik u Semëna Petroviča i bryznula na nego vodoj.*

‘In response, Irina Vasilievna unbuttoned Semyon Petrovich’s collar and sprayed him with water.’

The Active Dictionary states that in this meaning, *bryzgat'* ‘spray, sprinkle’ has three semantic arguments, one of which expresses the endpoint of spraying. We infer that *na nego*, literally ‘onto him’ functions as an argument in Apresjan’s framework (hence the main tag “1”). However, the exact syntactic form (*na* ‘on, onto’ + accusative) is not specified in the verb’s government pattern (*model' upravlenija*), warranting the additional tag “s” in our dataset, which reflects the morphosyntactic flexibility of the encoding pattern.

Manual dictionary-based argumenthood annotations are included in the final spreadsheet (*data_with_dictionary_annotations*).

4. Results

In this section, we compare the performance of our argument extraction algorithm with a manually annotated data subset based on dictionaries. Our goals

are twofold: first, to identify weaknesses in the algorithm with a view to future improvements; second, to highlight quantitative insights that may be overlooked in lexicography, which typically ignores text frequency. That said, we do not seek to replace the standard, semantically grounded concept of argumenthood with a simplistic, black-box algorithm. Rather, we aim to compare the two perspectives for their mutual benefit.

4.1. Rozental’s dictionary

Before evaluating our algorithm, we quantitatively compare two lexicographic approaches to argumenthood—those of Rozental’s dictionary (Rozental, 1986) and *The Active Dictionary* (Apresjan (Ed.), 2014–)—as shown in Table 4. The counts reflect manual annotations of a subset of UD entries, as discussed in Section 3.3.¹³

Table 4.
*Argumenthood: Rozental’s dictionary vs. The Active Dictionary*¹⁴

		The Active Dictionary			total
		yes	no	<NA>	
Rozental’s dictionary	yes	984	47	65	1,096
	no	696	510	37	1,243
	<NA>	2,267	646	171	3,084
	total	3,947	1,203	273	5,423

As the number of <NA>s in Table 4 shows, *The Active Dictionary* covers 95% (5,150 of 5,423) of the manually annotated dataset, far surpassing Rozental’s 43%. Overall, the two sources align, with matching annotations (boldfaced in the table) more common than discrepancies. However, Apresjan’s approach is more inclusive, cf. 696 tokens classified as arguments in *The Active Distionary* but not in Rozental’s dictionary as opposed to just 47 tokens of the opposite type.

¹³ Apart from directly identifying the necessary verbs in dictionaries, we sometimes inferred them from aspectual pairs or reflexive counterparts (see the discussion of additional tags in the Supplementary materials). We used <NA> only when none of these approaches yielded a satisfactory result.

¹⁴ In this and the subsequent tables, “yes” indicates dependents mentioned in the government patterns in the dictionaries ($\approx$ arguments), while “no” refers to all other dependents ($\approx$ adjuncts).

This difference underscores the lack of a principled consensus on argumenthood in naturalistic data as compared to preselected examples. In Rozental’s approach, fewer than half of (pro)nominal dependents in the text are arguments, compared to about 77% in Apresjan’s. This discrepancy should be taken into account when evaluating our algorithm’s performance. Table 5 cross-tabulates our algorithmized annotations with those based on Rozental’s dictionary.

Table 5.
Argumenthood: Rozental’s dictionary vs. frequency-based algorithm

		frequency-based algorithm		
		yes	no	total
Rozental’s dictionary	yes	1,007	89	1,096
	no	682	561	1,243
	<NA>	2,193	891	3,084
	total	3,882	1,541	5,423

As noted earlier, more than half of UD entries contain verbs missing from Rozental’s dictionary, making algorithmic approaches inherently superior for corpus analysis—even aside from the time-consuming nature of dictionary-based annotations. More importantly, the two approaches handle arguments asymmetrically: nearly all dependents listed in Rozental’s government patterns (92%, see Table 4, first row) are also classified as arguments by our algorithm. The reverse does not hold — Rozental’s non-arguments are more often identified as arguments than as adjuncts by our algorithm. This again underscores the restricted and conservative nature of Rozental’s approach to argumenthood. Manual inspection shows that these additional arguments identified by the algorithm, such as the dative NP in (7), typically do meet standard semantics-based criteria.

(7) *Vot ètim parnjam ja verju!*
‘These guys I do trust!’

Rozental’s dictionary does not list human dative arguments as part of the verb *verit’* ‘believe’ government pattern, likely because it focuses on variable cases, while the dative use in (7) is regular and predictable from the verb’s meaning. We conclude that our algorithm is better suited for detecting argumenthood in corpora than Rozental’s dictionary. This is not a critique of Rozental’—his dictionary was intended for contexts unrelated to automatic annotation. The rest of

the paper compares our results with *The Active Dictionary*, which has a broader scope and a strong theoretical foundation.

4.2. Apresjan’s dictionary: an overview

Table 6 summarizes our algorithm’s performance relative to *The Active Dictionary*, as Table 5 did for Rozenal’s dictionary.

Table 6.
Argumenthood: The Active Dictionary vs. frequency-based algorithm

		frequency-based algorithm		
		yes	no	total
<i>The Active Dictionary</i>	yes	3,409	538	3,947
	no	329	874	1,203
	<NA>	144	129	273
total		3,882	1,541	

Based on Table 6, our algorithm performs sufficiently well if the published volumes of *The Active Dictionary* are taken as the gold standard. Of the 3,738 entries identified as arguments by the algorithm, 3,409 are also classified as such in *The Active Dictionary* (excluding verbs missing from the dictionary), which yields a precision of approximately 0.91 for the algorithmic positives (see Korhonen et al., 2000 for the definition of precision). Precision for negatives is lower (≈ 0.62),¹⁵ resulting in an overall precision of ≈ 0.83 . We refer to these two kinds of mismatches as “false positives” (Section 4.3) and “false negatives” (Section 4.4). This standard terminology does not imply that the dictionary is always correct and the algorithm is wrong; as we show below, the two approaches sometimes capture different aspects of argumenthood.

¹⁵ Sarkar & Zeman (2000, p. 696) observe a similar disparity in the performance of the best-performing algorithm among the three they tested for automatic verb frame extraction in Czech.

4.3. False positives

As seen in Table 6, false positives are rare, accounting for 329 out of 5,423 entries. One possible source is our decision (see Section 3.2.) to classify all prepositionless accusatives as arguments, including cases like (8).

- (8) *Želudok bolit ne prekraščaja pjatyj den’.*
‘The stomach has been hurting nonstop for the fifth day’.

The noun phrase *pjatyj den’* ‘for the fifth day’ is clearly an adjunct, traditionally analyzed as a circumstantial (*obstoitel'stvo*) in Russian grammar. Thus, classifying all prepositionless accusatives as arguments inevitably leads to errors in such cases. The UD distinction between dependents bearing “obj” and “obl” (or occasionally “obl:tmod”) relations might seem helpful here. Table 7 cross-tabulates this distinction with the manual annotation based on *The Active Dictionary* for all prepositionless accusatives.

Table 7.
Prepositionless accusative dependents: UD-internal annotations vs. The Active Dictionary

		UD-internal annotations			total
		“obj”	“obl”	“obl:tmod”	
<i>The Active dictionary</i>	yes	1,068	53	6	1,127
	no	6	19	7	32
	<NA>	73	0	1	74
	total	1,147	72	14	1,233

As seen in Table 7, the “obl” and especially “obl:tmod” UD tags are slightly more common among adjuncts (“no” in the table) than arguments (“yes” in the table) when distinguished based on *The Active Dictionary*. To refine our algorithm, we adjusted the final annotation (the “argumenthood_corrected” column in the *data* spreadsheet) to rely on UD-inherent tags for accusative dependents: “obj” was classified as an argument, while all others were adjuncts. This adjustment may seem questionable for Russian, as accusative dependents tagged as “obl” often correspond to arguments in *The Active Dictionary*. However, we applied it anyway, since it helps with languages where object-like dependents frequently express adverbial meanings and had little effect on our algorithm’s accuracy for Russian. In any case, Russian accusative dependents contribute minimally to

false positives. Since we applied a special rule for prepositionless accusatives, we exclude them from further discussion of our algorithm's performance, focusing on the remaining 4,190 entries.

As noted above, false positives are generally rare, and there are few consistent patterns behind their occurrence. However, a few scenarios can still be illustrated — one involves polysemous encoding devices, as shown in examples (9) and (10).

- (9) *S ego pomošč'ju možno uznat', čto **varitsja v staleplavil'noj peči**.*

'With its help, one can find out what is **being brewed in the steelmaking furnace**.'

- (10) *V drevnem Kitae **varilos'** pivo iz proroššego risa.*

'**In ancient China, beer was brewed** from sprouted rice.'

Our algorithm detected a statistical association between *varit'sja* 'be brewed' and the encoding device *v* 'in' with the locative case. In (9), this correctly classifies the boldfaced phrase as an argument, aligning with *The Active Dictionary*. However, the algorithm generalizes this pattern to all occurrences of the same encoding device with *varit'sja*, leading to an error in (10), where the phrase denotes general localization rather than the vessel or container inherent in the verb's semantics.

While cases like (10) reflect clear errors in the algorithm's output, it occasionally captures patterns that are overlooked even in the detailed analysis of *The Active Dictionary*. One such case is shown in (11).

- (11) ***Dlja pozitivista est'** tol'ko fakty i različnye sposoby ix vzaimouvjazki.*

'**For a positivist, there are** only facts and various ways to connect them.'

The algorithm classified *dlja* 'for' phrases as arguments of *byt'* 'be' based on their frequent co-occurrence, though *The Active Dictionary* does not list this among the verb's many meanings and government patterns. While the boldfaced phrase in (11) is not a typical argument, it is not a standard adjunct either. Adjuncts typically narrow a clause's truth-conditional reference —for instance, *She baked a cake for him* necessarily entails *She baked a cake*. In (11), however, there is no such entailment that only facts exist and can be connected in various ways. Instead, the *dlja*-phrase effectively introduces a new meaning of *byt'* akin to 'seem' or 'be considered,' where it functions as an argument.

Another case where the algorithm offers semantic insights is shown in (12).

- (12) *No kogda oni gus'kom breli po lesnoj tropinke ... on uže tvrdo znal, čto ničego ne budet.*

‘But when they **trudged** in single file **along the forest path** ... he already knew for sure that nothing would happen.’

Based on frequent co-occurrence, the algorithm classified *po* ‘along, by’ phrases and instrumental-case dependents as arguments of *bresti* ‘trudge’. While *The Active Dictionary* does not include such phrases in the verb’s government pattern, their frequent corpus occurrence aligns with its semantic decomposition: ‘A person (A1) slowly moves toward place (A2) from place (A3), struggling to lift their feet due to weakness, fatigue, or difficult conditions—water, deep snow, or mud’ (Apresjan (ed.), 2014, p. 351). Though the path is not treated as a variable like A1, A2, or A3, which defines arguments, its typical properties (“water, deep snow, or mud”) are explicitly mentioned, making it more integral to the verb’s meaning than a standard adjunct.

To summarize, cases we conventionally label as “false positives” partly result from the algorithm’s inability to distinguish different senses of the same encoding device. However, some of these so-called false positives actually reveal insights that go beyond standard lexicographic perspectives, showing that they are not always truly false.

4.4. False negatives

False negatives arise when the algorithm fails to classify a verb’s dependent as an argument, even though the dictionary does. In our manually annotated dataset, they are more frequent (538 entries) than false positives (329). Two main causes account for false negatives: data sparsity and verb polysemy.

The first scenario involves rare verbs. The algorithm requires statistically significant differences and sets an absolute threshold of at least three occurrences for a verb-plus-encoding-device combination. Under these conditions, verbs that occur in the dataset only once or twice cannot, in principle, be classified as taking arguments by the algorithm, even when the dependent in question is clearly an argument in terms of content, as illustrated by (13).

- (13) *Novikov uže ballotirowsja v mèry.*
‘Novikov has already run for mayor.’

In (13), *ballotirovat'sja* ‘run for office’ appears with *v* (‘in’) and the rare “2nd accusative”.¹⁶ This encoding device, specific to verbs denoting a change in social status, clearly marks arguments. However, with only one occurrence of *ballotirovat'sja* in our dataset, the algorithm could not extract its argument-encoding pattern.¹⁷

Thus, data sparsity is the primary cause of false negatives—an issue with no perfect solution. While our algorithm currently produces binary annotations (argument vs. adjunct), a more realistic approach might be to filter out low-frequency verbs and assign <NA>s in unclear cases instead of (potentially false) negatives.

The second major source of false negatives occurs with highly frequent, syntactically versatile verbs (see Tao et al., 2024 on the correlation between verb frequency and valency diversity). Verbs like *bit* ‘beat’, and *vesti* ‘lead’ appear in numerous valency patterns, some of which are relatively infrequent. As a result, the algorithm fails to identify certain argument structures. This issue was already noted in Apresjan (1965), where polysemy was disregarded when extracting semantic information from frequency distributions.

Apart from the two main causes of false negatives, we also expected them to arise with verbs exhibiting “flexible” government patterns compatible with the same meaning. As noted in Apresjan and Páll (1982), many Russian verbs do not require a fixed preposition but can combine with various preposition-case pairs expressing a shared semantic role. For instance, *vernut'sja* ‘return’ can take different prepositions depending on the noun rather than the verb, e.g., *vernut'sja iz otpuska* ‘return from vacation’, *ot druga* ‘from a friend’s place’, *s koncerta* ‘from a concert’. In Apresjan and Páll’s framework, flexible patterns are categorized as P1 (source), as exemplified by examples with *vernut'sja* ‘return’, as well as P2 (goal), P3 (static location), and P4 (path). Similar observations appear in Helbig & Schenkel (1983, p. 43), Ljaševskaja & Kaškin (2015, p. 482), and *The Active Dictionary* (Apresjan (Ed.), 2014–).

We accounted for flexible government patterns in our manual annotations based on *The Active Dictionary* (see Section 3.3), marking verbs with such patterns with an additional “s” tag, as discussed in Section 3.3. This distinction is illustrated in (14), which follows a rigid government pattern, and (15), which

¹⁶ The second accusative is a form of Russian animate nouns whose exponent coincides with the nominative (not the usual accusative), but which appears after prepositions that normally require the accusative.

¹⁷ The algorithm extracted this pattern only for three verbs: *pojti* ‘become X’ in such contexts, *vyjti* ‘make one’s way to X’, and *vybit'sja* ‘rise to be an X’.

exhibits a flexible one—despite both involving the same preposition-case combination.

- (14) *Ušakov **vgljadyvalsja v temnotu** trezvimy glazami.*
‘Ushakov peered into the darkness with sober eyes’.
- (15) *Timorcy prinosjat v žertvu byka – zakalyvajut i s pesnjami **brosajut v more.***
‘The Timorese sacrifice a bull—slaughter it and throw it into the sea with songs’.

Since our argument extraction algorithm relies on the frequency of specific encoding devices rather than their groups, verbs with flexible government patterns were expected to yield more diffuse distributions, reducing the algorithm’s efficiency. Table 8 presents the relevant data: a breakdown of 2820 entries identified as arguments in *The Active Dictionary*, excluding those marked by prepositionless accusative (handled separately; see Section 4.3). As noted above, 538 of these were not recognized as arguments by the frequency-based algorithm (“false negatives”). Table 8 tests whether the false negative rate correlates with the rigidity or flexibility of the government pattern in *The Active Dictionary*.

Table 8.
Frequency-based argument extraction algorithm: rigid vs. flexible government patterns

	frequency-based algorithm		total
	yes	no	
rigid	837	221	1,058
flexible	1,445	317	1,762
	2,282	538	2,820

An unexpected result emerges from the data in Table 8: false negatives occur at similar rates for rigid (21%) and flexible (18%) government patterns identified in *The Active Dictionary*. This suggests that, despite their theoretical compatibility with various encoding devices, verbs with flexible government patterns tend to favor specific ones in actual usage, making their distributional behavior resemble that of verbs with rigid patterns. For example, *vernut'sja* ‘return’ favors *iz* ‘from’ plus the genitive (as in *vernut'sja iz komandirovki* ‘return from a business trip’), although other ablative prepositions can also occur with this verb (e.g., *vernut'sja s zanjatij* ‘return from class’ or *vernut'sja ot vsenoščnoj* ‘return from an all-night

vigil'). The frequency-based algorithm captures this asymmetry by classifying dependents introduced by *ot* plus the genitive, but not those with the other two prepositions, as arguments of *vernut'sja* 'return'.

5. Summary and outlook

The argument–adjunct distinction is complex, spanning formal and semantic dimensions that create interrelated but not identical contrasts. This complexity is evident in the evolution of Ju. D. Apresjan's views on argumenthood, from his early quantitative distributional approach to the in-depth semantic analyses of his later work.

This paper aims to reconcile different approaches to argumenthood using Russian data. Our primary goal is not to uncover new facts about Russian grammar and lexicon, but to establish a new methodological avenue for token-based typology, using Russian as a starting point. Specifically, we propose an algorithm for extracting arguments from UD treebanks and evaluate its performance against semantically nuanced lexicographic sources, primarily *The Active Dictionary* (Apresjan (ed.), 2014–). The key idea is that verb arguments typically require specific encoding forms, such as prepositions and cases in Russian, leading to frequency peaks in the distribution of encoding devices across verb lemmas. In contrast, adjuncts are not lexically determined and show flatter distributions. Based on this principle, the algorithm identifies verb–encoding device combinations that occur significantly more often than expected under a flat zero hypothesis.

Evaluated against *The Active Dictionary*, the algorithm performed well, achieving an overall precision of about 0.83. False positives were particularly rare, indicating that dependents identified as arguments by the algorithm generally align with those recognized in semantic-based lexicography. The asymmetry between infrequent false positives and more common false negatives likely reflects an inherent distinction between semantic valency and formal government: while formal government presupposes semantic valency, the reverse is not necessarily true (Helbig & Schenkel, 1983, p. 44).

The low rate of false positives is promising for the practical goals of this study. This paper is part of a broader effort to analyze argumenthood and valency cross-linguistically in typologically diverse languages. For Russian, such analysis—though time-consuming—can, in principle, rely on semantic pattern analysis, given the availability of high-quality lexicographic resources. However, for

most languages, such resources are missing and unlikely to emerge soon. In this context, an algorithm that reliably extracts arguments from UD treebanks, even if incomplete, provides a valuable starting point for a quantitative, token-based study of valency patterns. We hope that the UD platform will continue to expand in favor of lesser-documented languages and become increasingly suitable for full-scale typological research.

While the current algorithm achieves a reasonably good accuracy rate as a starting point, further improvement is likely necessary. Two main avenues appear promising in this regard. First, the algorithm should be adapted to better handle rare verbs. Even a seemingly defeatist solution—such as refraining from assigning argumenthood labels to rare verbs altogether—may be more useful in practice than producing false negatives. Second, in its current form, the algorithm processes each verb-dependent token in isolation. In reality, a verb's arguments typically occur as part of a larger frame, with dependents interacting with one another. Such interaction is crucial both theoretically—being a cornerstone of Construction Grammar (Goldberg, 1995), for instance—and practically, as it can significantly improve the efficiency of verb frame extraction algorithms (Sarkar & Zeman, 2000).

Beyond its practical applications, the frequency-based argument extraction algorithm also yields insights into the theoretical understanding of argumenthood. (i) Verbs can exhibit co-occurrence frequency peaks with dependents not traditionally considered full-fledged arguments, yet their presence can alter verb meaning and affect argument structure (e.g., *byt' dlja* + Genitive, 'be for someone'; see Section 4.3). (ii) While canonical arguments correspond to variables in a verb's semantic decomposition, frequency peaks also occur with dependents that represent key semantic components, revealing complex links between syntactic and semantic structure (e.g., *bresti* 'trudge'; see Section 4.3). (iii) A quantitative, token-based perspective suggests that distinctions between rigid government patterns, which require a specific encoding form, and more flexible patterns, where verbs allow multiple competing encoding devices with similar spatial meanings, may not be entirely clear-cut (see Section 4.4).

References

- Apresjan, Ju. (1965). Opyt opisanija značenij glagolov po ix sintaksičeskim priznakam (tipam upravlenija). *Voprosy jazykoznanija*, 5, 51–66.

- Apresjan, Ju. (1967). *Èksperimental'noe issledovanie semantiki russkogo glagola*. Nauka, Moskva.
- Apresjan, Ju., & Páll, E. (1982). *Russkij glagol — vengerskij glagol. Upravlenie i sočetaemost'* (Vols. 1–2). Tankönyvkiadó, Budapest.
- Apresjan, Ju. (1974). *Leksičeskaja semantika: Sinonimičeskie sredstva jazyka*. Nauka, Moskva.
- Apresjan, Ju. (1995). *Izbrannye trudy* (Vols. 1–2). Jazyki russkoj kul'tury, Moskva.
- Apresjan, Ju. (Ed.). (2004). *Novyj ob'jasnitel'nyj slovar' sinonimov russkogo jazyka* (2nd edition). Jazyki russkoj kul'tury, Moskva.
- Apresjan, Ju. (Ed.). (2014–). *Aktivnyj slovar' russkogo jazyka* (Vol. 1, 2014; Vol. 2, 2014; Vol. 3, 2017; Vol. 4, Part 1, 2023; Vol. 4, Part 2, 2024). Jazyki slavjanskoj kul'tury Moskva.
- Barðdal, J. (2011). Lexical vs. structural case: A false dichotomy. *Morphology*, 21, 619–654. <https://doi.org/10.1007/s11525-010-9174-1>
- Beavers, J. (2010). The structure of lexical meaning: Why semantics really matters. *Language*, 86(4), 821–864.
- Bickel, B., Zakharko, T., Bierkandt, L., & Witzlack-Makarevich, L. (2014). Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment. *Studies in language*, 38(3), 485–511.
- Boguslavskij, I. (1996). *Sfera dejstvija leksičeskix edinic*. Jazyki russkoj kul'tury, Moskva.
- Bohnemeyer, J. (2007). Morpholexical transparency and the argument structure of verbs of cutting and breaking. *Cognitive Linguistics*, 18(2), 153–177.
- de Marneffe, M. C., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal Dependencies. *Computational Linguistics*, 47(2), 255–308.
- Dixon, R. M. W. (2009). *Basic linguistic theory. Volume 1: Methodology*. Oxford University Press, Oxford.
- Dryer, M. S. (with Gensler, O. D.). (2013). Order of object, oblique, and verb. In M. S. Dryer & M. Haspelmath (Eds.), *WALS Online* (v2020.4) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.13950591> (Available online at <http://wals.info/chapter/84>, accessed on 2025-02-25.)
- Fowler, G. (1996). Oblique passivization in Russian. *The Slavic and East European Journal*, 40(3), 519–545.
- Frajzyngier, Z., Gurian, N., & Karpenko, S. (2024). Minimal participant structure and the emergence of the argument/adjunct distinction. *Studies in Language*, 48(1), 181–227. <https://doi.org/10.1075/sl.22029.fra>.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press, Chicago.
- Haspelmath, M. (2014). Arguments and adjuncts as language-particular syntactic categories and as comparative concepts. *Linguistic Discovery*, 12(2), 3–11.

- Helbig, G., & Schenkel, W. (1983). *Wörterbuch zur Valenz und Distribution deutscher Verben*. Tübingen: Max Niemeyer.
- Jacobs, J. (1994). *Kontra Valenz*. Trier: Wissenschaftlicher Verlag Trier.
- Jolly, J. A. (1993). Preposition assignment in English. In R. D. Van Valin Jr. (Ed.), *Advances in role and reference grammar* (pp. 275–310). Benjamins, Amsterdam.
- Kamynina, A. (1999). *Sovremennyy russkij jazyk. Morfologija*. MGU, Moskva.
- Koenig, J.-P., Mauner, G., Bienvenue, B., & Conklin, K. (2008). What with? The anatomy of a role. *Journal of Semantics*, 25(2), 175–220.
- Korhonen, A., Gorrell, G., & McCarthy, D. (2000). Statistical filtering and subcategorization frame acquisition. In *Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora* (pp. 199–205). Hong Kong, China.
- Lazard, G. (1994). *L'actance*. Presses Universitaire de France, Paris.
- Ljaševskaja, O., & Kaškin, E. (2015). Tipy informacii o leksičeskix konstrukcijax v sisteme Frejmbank. *Trudy Instituta russkogo jazyka im. V. V. Vinogradova*, 6, 464–555.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago Press, Chicago.
- Levin, B., & Rappaport Hovav, M. (2005). *Argument realization*. Cambridge University Press, Cambridge.
- Malchukov, A., & Comrie, B. (Eds.). (2015). *Valency classes in the world's languages* (Vols. 1-2). De Gruyter Mouton, Berlin, Boston.
- Mel'čuk, I., Žolkovskij, A., & Apresjan, Ju. (1984). *Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka: Opyty semantiko-sintaktičeskogo opisanija russkoj leksiki*. Wiener Slavistischer Almanach, Wien.
- Muravenko, E. (1998). O slučajax netrivial'nogo sootvetstvija semantičeskix i sintaksičeskix valentnostej glagola. *Semiotika i informatika*, 36, 71–81. Jazyki russkoj kul'tury, Moskva.
- Nivre, J., de Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S., Schuster, S., Tyers, F., & Zeman, D. (2020). Universal dependencies v2: An ever-growing multilingual treebank collection. In *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 4034–4043). European Language Resources Association, Marseille.
- Plungjan, V., & Raxilina, E. (1998). Paradoksy valentnostej. *Semiotika i informatika*, 36, 108–119. Jazyki russkoj kul'tury, Moskva.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. Retrieved from <https://www.R-project.org/> (accessed May 11, 2024).
- Rozental', D. (1986). *Upravlenie v russkom jazyke* (2nd edition). Kniga, Moskva.

- Sarkar, A., & Zeman, D. (2000). Automatic extraction of subcategorization frames for Czech. In *Proceedings of the 18th International Conference on Computational Linguistics* (pp. 691–697). Saarbrücken, Germany.
- Schulte im Walde, S. (2009). The induction of verb frames and verb classes from corpora. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (pp. 952–972). Walter de Gruyter, Berlin.
- Tao, S., Donatelli, L., & Hahn, M. (2024). More frequent verbs are associated with more diverse valency frames: Efficient principles at the lexicon-grammar interface. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers), 11795–11810.
- Testelec, Ja. (2001). *Vvedenie v obščij sintaksis*. RGGU, Moskva.
- Tesnière, L. (1959). *Éléments de syntaxe structurale*. Klincksieck, Paris.
- Van Valin, R. D. Jr. (1993). A synopsis of Role and Reference Grammar. In R. D. Van Valin Jr. (Ed.), *Advances in Role and Reference Grammar* (pp. 1–166). Benjamins, Amsterdam.
- Van Valin, R. D. Jr. (2001). *An introduction to syntax*. Cambridge University Press, Cambridge.
- Yip, M., Maling, J., & Jackendoff, R. (1987). Case in tiers. *Language*, 63(2), 217–250.
- Zeman, D., et al. (2022). *Universal dependencies 2.11*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. Retrieved from <http://hdl.handle.net/11234/1-4923>.
- Zolotova, G. (2006). *Sintaksičeskij slovar'. Repertuar elementarnyx edinic russkogo jazyka*. 3rd edition. URSS, Moskva.

Appendix

The entries for *bazirovat'sja* ‘base oneself, to be based’ in (I) Rozental’s dictionary (1986) and in (II) *The Active Dictionary* (Apresjan (Ed.), 2014–)

(I) базироваться на чём и на что. 1. на чём (основываться на чем-л. в своих суждениях или действиях). Базироваться на данных опыта. Нравственное воспитание коммунистического человека, прежде всего, базируется на воспитании его способностей, на развитии его сил, его созидательного, творческого актива (Макаренко). 2. на что и на чём (опираться на что, иметь что-л. в качестве базы). Самолёты базировались на новый аэродром (на новом аэродроме).

(II) БАЗИРОВАТЬСЯ, ГЛАГ.; -руюсь, -руется; НЕСОВ; СОВ нет.

базироваться 1

Базироваться на фактах <на опыте>.

ЗНАЧЕНИЕ. *А1 базируется на А2 'В своих действиях или деятельности А3 человек А1 использует информацию А2 или информацию о явлении А2, считая, что она необходима для успешности А3'.*

УПРАВЛЕНИЕ 1.

*А1 * редк. ИМ: (В своих прогнозах) метеорологи базируются (на анализе спутниковых данных).*

*А2 * на ПР: базироваться на анализе спутниковых данных.*

*А3 * в ПР: базироваться в своих выводах <в своих прогнозах>.*

УПРАВЛЕНИЕ 2.

*А3 * ИМ: Прогнозы экономистов базируются (на оптимистической оценке роста цен на углеводороды).*

*А2 * на ПР: базироваться на опти*

СОЧЕТАЕМОСТЬ. *Базироваться на других принципах <на теории Дарвина, на новой системе подсчёта голосов>; базироваться на чьих-л. оценках <на устаревших представлениях>.*

☐ *Экономическое обоснование такой стратегии базируется на отчётах аналитиков («Computerworld», № 25, 2004). На этом и будут базироваться наши дальнейшие рассуждения («Информационные технологии», № 8, 2004). А она [власть] базируется не столько на силе, богатстве, благостности, мудрости etc. властителя, сколько [...] на готовности подданных к повиновению (М. Соколов). Все эти программы эффективнее предшествующих, так как базируются на новых объективных данных о путях оптимизации работы мозга взрослого и ребёнка («Вопросы психологии», 2004.10.12). Ведь в англосаксонском праве всё базируется на прецеденте (К. Амелин).*

СИН: *основываться (на чем-л.), опираться (на что-л.); АНА: исходить из чего-л.; КОНВ: базировать, основывать; ДЕР.: база.*

базироваться 2

На этом аэродроме базировалось несколько эскадрилий.

ЗНАЧЕНИЕ. *А1 базируется на А2: 'Организация или войска А1 расположены на территории или в месте А2, где они обеспечиваются всем необходимым для их нормального функционирования'.*

УПРАВЛЕНИЕ.

*А1 * ИМ: Фирма базируется (в Лондоне).*

*А2 * ГДЕ: базироваться в Токио <на Кольском полуострове, под Москвой>.*

☐ *Попал я морчасти погранвойск, базирующиеся в дальневосточной Находке (Л. Вертинская). Огрызкову сообщили, куда они должны прибыть,*

[...] – урочище Боговизна, где базировалось руководство партизанской зоны (В. Быков). Организация экономического сотрудничества и развития [...] базируется в Париже (В. Ключников). На самом деле там базировалась разведшкола одной из спецслужб (В. Морозов). ДЕР: базирование; база (эскадрилий). [Ю.А.]



Elena Uryson

 <https://orcid.org/0000-0002-9391-173X>

Sur les verbes multiplicatifs russes (étude pilote)*

On Russian multiplicative verbs (a pilot study)

Abstract

The object of the paper is Russian multiplicative verbs, that is verbs denoting repeating homogenous acts, cf. *šagat'* («to step»), *maxat'* («to wave»), *ikat'* («to hiccup»), etc. My goal is to specify the definition of this class of verbs and to interpretate from the semantic point of view a well known systemic ambiguity of some sentences with such a verb. The study is carried out within the frame of Moscow semantic school. The problem is that there are many verbs which denote processes consisting of repeating homogenous acts, and still such verbs are never classified as multiplicative; cf. *idti* («to walk»); *bežat'* («to run»); *kosit' travu* («to mow the grass»), etc. The point is that in the focus of the meaning of a multiplicative verb are repeating homogenous acts while in the focus of *idti* 'to walk', *bežat'* 'to run', *kosit' travu* 'to mow the grass', etc. is purpose of action. I also demonstrate that meaning of many multiplicative verbs contains semantic component 'once or more than once', and because of that in some contexts they denote not series of homogenous acts but a single act. I call such multiplicative verbs non-strict as opposed to strict multiplicatives which always denote repeating homogenous acts (cf. *drožat'* «to tremble for some time»). From this point of view I discuss aspectual pairs like *šagat'* – *šagnut'*, *maxat'* – *maxnut'*. The semantic interpretation of the verbs under discussion allows us to reveal the direction of derivation in the pair multiplicative – semelfactive.

Keywords

Russian, multiplicative verb, semelfactive verb, aspectual pair, semantics

* L'auteure exprime sa profonde gratitude à S.G. Krylosova pour ses précieuses remarques éditoriales.

1. Objet et objectif de l'étude

La présente étude porte sur les verbes **multiplicatifs** russes, tels que

- *barabanit'* (*pal'cami po stolu*) 'tambouriner [avec ses doigts sur la table]'¹
- *bul'kat'* 'glouglouter, bruire [pour l'eau]'
- *celovat'* 'embrasser'
- *česat'sja* 'se gratter'
- *ikat'* 'hoqueter'
- *maxat'* 'agiter [la main]'
- *morgat'* 'cligner de l'œil'
- *šagat'* 'marcher'
- *xlopat'* (*v ladoši*) 'taper [dans les mains], applaudir', etc.

L'analyse s'attache à la sémantique lexicale de ces verbes, tout en abordant une question aspectologique particulière : la corrélation aspectuelle des verbes dans des paires telles que *bul'kat'*^{nPF} – *bul'knut'*^{nPF} 'glouglouter / bruire', *maxat'*^{nPF} – *max-nut'*^{nPF} (*rukoj*) 'agiter [la main]'.

Toute description d'un groupe verbal suppose d'abord de préciser sa place dans une classification des prédicats. Nous retiendrons ici la classification simplifiée suivante :

- a. **états** : *Selo stoit na beregu ozero* 'litt. Le village se tient sur la rive du lac'.
- b. **événements** : *On napisal pis'mo* 'Il a écrit une lettre' ; *Al'pinisty dostigli veršiny* 'Les alpinistes ont atteint le sommet'.
- c. **procès** : *On pišet pis'mo* 'Il écrit une lettre' ; *Putešestvenniki podnimalis' na pereval* 'Les voyageurs montaient vers le col' ; *Šel sil'nyj dožd'* 'Il pleuvait fort'.

Selon cette terminologie, les verbes examinés ici sont rattachés à la classe des procès.

Knjazev (2007 : 439) propose de définir les verbes multiplicatifs comme l'expression de « procès composés d'une multitude indéterminée d'actes de même nature, se répétant à une périodicité relativement élevée ». Ainsi, *maxat' flažkom* 'agiter un drapeau' signifie, en simplifiant : 'lever et baisser plusieurs fois la main tenant un drapeau, sans interruption ou avec seulement de brèves pauses'. De même, *ikat'* (dans un énoncé comme *Sidel i ikal* 'Il était assis et il hoquetait') signifie 'sous l'effet d'un trouble respiratoire, produire involontairement des sons

¹ NdT : Les verbes cités en russe sont référencés dans l'index alphabétique présenté en fin d'article.

gutturaux, brefs et saccadés, dus à des mouvements convulsifs de la poitrine ou de la gorge, sans interruption ou avec de brèves pauses’.

Nous montrerons ci-dessous que cette définition, bien que généralement admise, n’est pas pleinement adéquate et proposerons une définition plus précise du multiplicatif.

REMARQUE.

Il est facile de voir que les multiplicatifs se rapprochent sémantiquement des verbes dont le sens implique la répétition d’une certaine situation, par exemple, *xaživat’ (v gosti)* ‘passer de temps en temps rendre visite’ (*My k nim xaživali po večeram* ‘Le soir, nous avions l’habitude d’aller chez eux’). Ces verbes qui indiquent la répétition d’une situation globale sont appelés itératifs². Voici quelques exemples :

Tableau 1.
Quelques exemples de verbes itératifs

verbe itératif	traduction approximative	exemple
<i>počityvat’ (gazety)</i>	‘lire des journaux de temps à autre’	<i>Detektivy ne očen’ ljublju, no vremja ot vremeni počityvaju</i> ‘Je n’aime pas beaucoup les romans policiers, mais j’en lis de temps en temps’.
<i>zaxaživat’ (v gosti)</i>	‘passer de temps en temps, rendre visite’	<i>On k nam zaxažival.</i> ‘Il passait nous voir de temps en temps’.
<i>edat’ (delikatesy)</i>	‘manger de temps en temps des mets gastronomiques’	<i>Ja v Pariže ljagušek edal</i> ‘À Paris, il m’est arrivé de manger des grenouilles’.
<i>dirat’ (za uši)</i>	‘tirer à plusieurs reprises par les oreilles’	<i>Ja za uši ego dirala, tol’ko malo</i> [A. S. Griboedov] ‘Je lui tirais les oreilles à plusieurs reprises, mais pas assez’.

La différence entre la répétition d’actes constituant une situation unique (*ikat’, maxat’*) et la répétition d’une situation complète dans le temps (cf. les exemples du Tableau 1) a été décrite à maintes reprises dans la littérature linguistique (voir la synthèse de Šluinskij, 2006). Nous n’insisterons pas sur ce point.

² Nous employons les termes *multiplicatif* et *itératif* dans le sens défini par Ju. Maslov (2004). Notons cependant que, dans certains travaux, le terme *multiplicatif* (*mul’tiplikativ* [NdT]) est employé dans une autre acception. Ainsi, chez Mel’čuk (1998), relèvent de cette classe les verbes russes formés avec le suffixe *-yva/-iva* (*xaživat’, siživat’, edat’*), qui sont généralement considérés comme des verbes itératifs.

Un verbe désignant un acte, un *quantum* dans la succession de ces « actes de même nature », peut être associé au multiplicatif. Un tel verbe est généralement pourvu du suffixe *-nu* (ou de sa variante stylistiquement marquée *-anu*). Ainsi, on trouve des couples tels que

- *bul'kat'*^{IPF} / *bul'knut'*^{PF} 'glouglouter, bruire'
- *ikat'*^{IPF} / *iknut'*^{PF} 'hoqueter'
- *lizat'*^{IPF} / *liznut'*^{PF} et *lizanut'*^{PF} 'lécher'
- *morgat'*^{IPF} / *morgnut'* 'cligner de l'œil'
- *xlopat'*^{IPF} / *xlopnut'*^{PF} (*v ladoši*) 'taper [dans les mains]'

Dans la grammaire russe, des couples comme *morgat'* / *morgnut'*, *ikat'* / *iknut'* sont traditionnellement considérés comme **aspectuels**. Dans le présent travail, ce point de vue sera discuté: nous montrerons que de telles paires ne satisfont pas pleinement à la définition d'un couple aspectuel. L'objectif principal de cette étude est de dégager un schéma d'interprétation du multiplicatif et de décrire certaines particularités sémantiques de ces verbes, qui n'ont jusqu'à présent pas attiré l'attention des chercheurs. On étudiera également la sémantique du semelfactif, dans la mesure où cela est nécessaire à la description des multiplicatifs³.

La définition du multiplicatif donnée ci-dessus (Knjazev, 2007: 439) est en fait une description généralisée de son sens. Elle suppose que les verbes rattachés à cette classe ont déjà été définis, ou qu'il a du moins été montré que chacun d'eux désigne effectivement une série d'actes similaires. Or, la linguistique ne dispose pas aujourd'hui d'une telle description. En pratique, lorsqu'il classe un verbe dans le groupe des multiplicatifs, le chercheur s'appuie avant tout sur son intuition linguistique et sur une analyse préalable – souvent implicite – du sens du verbe. En même temps, cette interprétation, même provisoire, **se trouve vérifiée**: le chercheur examine la situation que le verbe décrit, observe son comportement en contexte et met en évidence les traits expliqués par la sémantique⁴. Il est évident qu'une telle analyse sémantique préalable ne satisfait pas les chercheurs habitués à se fonder sur des critères morphologiques et syntaxiques. C'est pourquoi, dans certains travaux, la définition de la classe des multiplicatifs ne repose pas sur la sémantique du verbe, mais sur un autre critère, plus simple: le verbe multiplicatif y est défini par ses relations

³ Désignant un procès (mais non un événement), le multiplicatif est un verbe imperfectif. Le semelfactif, en revanche, désigne un événement et est un perfectif. Une littérature scientifique assez importante est consacrée à la morphologie et à la sémantique du semelfactif, opposé au multiplicatif correspondant (Dickey & Janda, 2009; Janda & Makarova, 2009; Gorbova, 2016).

⁴ Les principes de vérification des définitions des unités lexicales ont été exposés en détail par Ju. Apresjan (1974: 95–118).

paradigmatiques, autrement dit comme un verbe ayant un semelfactif qui lui est sémantiquement corrélé⁵.

Toutefois, il convient de souligner que le système lexical d'une langue⁶ n'est pas toujours parfaitement cohérent : il peut comporter des lacunes difficiles, voire impossibles, à expliquer du point de vue du système. En russe, par exemple, la désignation d'un ensemble ou d'une classe d'éléments de très petite taille est souvent associée à celle d'un élément unique de cet ensemble (ou de cette classe) :

- *pesok* 'le sable' – *pesčinka* 'le grain de sable'
- *pyl'* 'la poussière' – *pylinka* 'le grain de poussière'
- *soloma* 'la paille' – *solominka* 'le brin de paille.'
- *trava* 'l'herbe' – *travinka* 'le brin d'herbe'
- *žemčug* 'les perles' – *žemčuzina* 'la perle'

Cependant, certains couples ne disposent pas d'un dérivé permettant de désigner un élément unique de l'ensemble :

- *gravij* 'le gravier' – ?
- *ščeben'* 'le cailloutis' – ?
- *seno* 'le foin' – ?

De la même façon, il arrive que la langue n'ait pas de semelfactif associé à un multiplicatif. Par exemple, on trouve le couple *šagat'* / *šagnut'* 'marcher' / 'faire un pas', mais le verbe *semenit'* 'trotter' n'a pas d'équivalent semelfactif⁷. Dès lors, de nombreux verbes échappent à la description, alors même qu'il serait naturel de les inclure dans la classe des multiplicatifs, comme, par exemple :

- *barabanit'* (*pal'cami po stolu*) 'tambouriner [des doigts sur la table]'
- *česat'* (*spinu*) 'se gratter [le dos]'
- *česat'sja* 'se gratter' (*Sobaka češetsja* 'Le chien se gratte')
- *gladit'* (*sobaku*) 'caresser [un chien]'
- *mercat'* 'scintiller'

⁵ C'est précisément sur ce critère de sélection que repose la description la plus complète des verbes multi-actes (multiplicatifs) russes (Birjulin, 2024).

⁶ Nous partons des représentations théoriques suivantes sur l'organisation des sens lexicaux à l'échelle de l'ensemble du lexique : a) L'ensemble des sens lexicaux d'une langue naturelle est organisé de manière systémique, quoique dans une moindre mesure que, par exemple, sa grammaire ; b) Le système lexical possède des aspects à la fois classificatoire et opérationnel ; c) En tant que système classificatoire, le lexique se présente comme une hiérarchie souple, comprenant des classes et sous-classes sémantiques qui se recoupent à de multiples reprises ; d) En tant que système opérationnel, il se caractérise par des règles d'interaction des sens dans le texte (Apresjan, 2006 : 69).

⁷ L'analogie entre les couples du type *maxat'* – *maxnut'* (multiplicatif – semelfactif), d'une part, et les couples du type *fasol'* – *fasolina* (substantif désignant un matériau – singulatif) de l'autre, a été relevée notamment dans les travaux de Mehlig (1994) et de Xrakovskij (1998 : 488).

- *sučit'* (*nogami*) 'bouger [les jambes] nerveusement et rapidement'
- *vertet'* (*golovoj*) 'tourner [la tête]', etc⁸.

En outre, un acte unique d'un procès multiplicatif peut être exprimé par un verbe qui n'est pas formé avec le suffixe *-(a)nu*. Ainsi, à côté de *kusat'* 'mordre', on trouve *ukusit'* 'mordre une fois' : *Sobaka kusaet vora* 'Le chien mord le voleur' – *Sobaka ukusila vora* 'Le chien a mordu le voleur'. Dans de tels cas, on parle d'un mode d'action unique (*semelfactif*) au sens large, le semelfactif marqué par le suffixe *-(a)nu* n'en constituant qu'un cas particulier. Toutefois, de nombreux travaux consacrés aux multiplicatifs ne prennent en compte que les semelfactifs en *-(a)nu*.

Dans le présent travail, la classe des multiplicatifs est envisagée indépendamment de l'existence d'un semelfactif correspondant.

Remarques terminologiques

La majorité des verbes d'une langue naturelle sont polysémiques. Dans notre analyse, nous considérerons toujours un verbe dans son sens concret, attesté en contexte. Dans la tradition de l'École sémantique de Moscou, un mot pris dans sa signification concrète et contextuelle est appelé *lexème*. Dans ce travail, lorsque nous parlerons d'un lexème verbal concret, nous emploierons, par simplification, le terme *verbe*.

Selon notre définition, un verbe multiplicatif désigne un procès ; par conséquent, seul un verbe imperfectif peut être multiplicatif. Toutefois, il existe des verbes perfectifs dérivés d'un verbe multiplicatif exprimant un procès. Par exemple :

- *počesat'*^{PF} *spinu* 'gratter le dos'
- *pogladit'*^{PF} (*sobaku*) 'caresser un peu [le chien]'
- *pokukarekat'*^{PF} 'pousser un "cocorico"
- *poxixikat'*^{PF} 'ricaner, glousser un peu / pendant un moment'
- *zabarabanit'*^{PF} *pal'cami po stolu* 'se mettre à tambouriner des doigts sur la table'
- *zamaxat'*^{PF} *rukami* 'se mettre à agiter les mains'

Dans l'énoncé *On zakašlja^{PF} i potom rasčixalsja^{PF}* 'Il s'est mis à tousser puis à éternuer', il y a eu au minimum plusieurs quanta de toux et au minimum plusieurs quanta d'éternuement. Il est clair que la sémantique d'un tel verbe perfectif renvoie à un procès multiplicatif (ce qui a été remarqué notamment

⁸ Pour une liste détaillée de tels verbes, voir Birjulin (2024 : 11).

par Xrakovskij (1987 : 130)). Dans la suite de cet article, par souci de concision, nous appellerons multiplicatif non seulement le verbe imperfectif désignant un procès composé d'actions répétées, mais aussi le verbe perfectif renvoyant à un tel procès.

2. Les classes de multiplicatifs : essai de description sémantique

2.1. Vers une définition du multiplicatif et du semelfactif

Un point mérite ici une attention particulière (Tatevosov, 2016 : 82). Il existe en russe des verbes qui désignent des procès composés d'actes de même nature répétés de manière périodique, mais qui, malgré cela, ne sont jamais rangés parmi les multiplicatifs. Par exemple :

- *bežat'* 'courir', déterminé et *begat'* 'courir', indéterminé
- *idti* 'aller', déterminé et *xodit'* 'aller', indéterminé
- *letet'* 'voler', déterminé et *letat'* 'voler', indéterminé
- *plyt'* (*brassom*) 'nager [la brasse]', déterminé, etc.

En effet, les verbes *idti*, *xodit'*, *bežat'*, *begat'*, *letet'*, *letat'*, *plyt'* (*krolem*) 'nager [le crawl]' et d'autres verbes de déplacement désignent des actions composées de mouvements périodiques des jambes (ou des pattes d'un animal) ou de battements d'ailes. Ces verbes semblent parfaitement correspondre à la définition du multiplicatif. Pourtant, dans aucune des études que nous connaissons ces verbes ne sont rangés parmi les multiplicatifs : ce choix repose sur une intuition qu'il convient d'explicitier. Il ne s'agit pas seulement du fait qu'ils ne forment pas de couple sémantique avec un semelfactif. Comme nous l'avons vu plus haut, c'est aussi le cas de nombreux verbes dont la nature multiplicative ne fait aucun doute. Par exemple,

- *semenit'* 'trotter, marcher à petits pas'
- *barabanit'* (*pal'cami po stolu*) 'tambouriner [des doigts sur la table]'
- *gremet'* (*kastrjuljami*) 'faire du bruit [avec des casseroles]'
- *iskrit'* (*Provod iskrit*) 'faire des étincelles' (*Le câble fait des étincelles*)
- *kolotit'* (*nogoj v dver'*) 'frapper [du pied contre la porte]', etc. (voir Birjulin, 2024).

La raison en réside dans la structure sémantique des verbes du type *idti*, *xodit'*.

On admet bien volontiers que le simple fait qu'un verbe renvoie à des actes de même nature répétés périodiquement ne suffit pas à en faire un multiplicatif.

Il faut encore que ces actes attirent l'attention de l'observateur, c'est-à-dire qu'ils se trouvent au focus, qu'ils soient mis en relief par un accent sémantique⁹. Prenons des verbes désignant le mode de déplacement le plus courant, le plus naturel pour l'homme ou pour un autre être vivant, tels que *idti*, *xodit'*, *letet'*, *letat'*, etc. Ils désignent indiscutablement des mouvements périodiques répétés et homogènes d'une partie du corps, grâce auxquels l'être se déplace. Toutefois, ces mouvements étant constitutifs du déplacement lui-même et allant de soi, il n'y aurait pas de sens à attirer spécialement l'attention sur eux.

En ce qui concerne le multiplicatif, c'est précisément le caractère des quanta qui est mis en relief dans sa signification. Comparons les verbes suivants :

- *idti* 'aller, marcher'
- *šagat'* 'marcher à grands pas'
- *semenit'* 'trotter, marcher à petits pas'

Ils désignent tous le même mode de déplacement de l'homme à la surface (du sol ou du plancher), et ne diffèrent que par l'indication du caractère du mouvement des jambes – son amplitude et sa fréquence. Mais avec le verbe *idti* ces mouvements ne sont pas au focus : par défaut, on suppose qu'ils sont tout à fait ordinaires. Avec *semenit'* et *šagat'*, en revanche, sont au focus les « quanta répétés du mouvement », et c'est pour cette raison que les linguistes considèrent ces verbes comme des multiplicatifs. Nous allons maintenant argumenter ce point de vue.

Il est bien établi que lorsqu'une partie de l'énoncé est au focus, mise en relief par un accent sémantique (éventuellement contrastif) et, par conséquent, prosodique, c'est cette partie focale qui est niée. Ainsi, dans *On ne sdelal buterbrod S ŠEMGOJ* – 'Il n'a pas fait un sandwich avec du saumon' – ce qui est nié n'est pas le fait qu'il ait préparé un sandwich, mais bien que le sandwich préparé soit au saumon (Chafe, 1976). Il en va de même pour l'élément focal dans le sens d'une lexie : si une composante du sens est mise en relief par un accent sémantique (c'est-à-dire placé au focus), ce n'est pas le sens du mot qui est niée, mais bien cette composante précise.

Prenons nos verbes *idti*, *semenit'*, *šagat'*.

- (1) *On šěl*^{passé du verbe *idti*} *po koridoru* 'Il **marchait** dans le couloir'.
- (2) *On semenil po koridoru* 'Il **trottinait** dans le couloir'.
- (3) *On šagal po koridoru* 'Il **marchait à grands pas** dans le couloir'.

Dans l'énoncé (1), ce qui est affirmé – ou, dans une phrase négative, ce qui est nié –, c'est le fait même du déplacement du sujet :

⁹ La notion de focus et celle d'accent sémantique sont exposées dans (Apresjan, 2004a).

- (4) — *On šel po koridoru, ty dolžen byl ego vstretit'*
 'Il marchait dans le couloir, tu aurais dû le croiser'.
 — *Net, on ne šel po koridoru (ja ego ne vstretil)*
 'Non, il ne marchait pas dans le couloir (je ne l'ai pas croisé)'.

Il en va différemment avec les multiplicatifs *semenit'* et *šagat'*. Dans l'énoncé (5), ce qui est au focus de l'attention du locuteur n'est pas tant le fait du déplacement du sujet que sa manière particulière de marcher – en l'occurrence les petits pas (c'est cette composante sémantique qui est mise en relief dans le sens du verbe *semenit'*). C'est pourquoi la phrase (6) paraît au minimum étrange, sinon ironique : elle viole le postulat de pertinence de Grice, puisque la manière de marcher n'a aucun rapport avec la possibilité de croiser le sujet dans le couloir.

- (5) *On semenil po koridoru* 'Il trottinait dans le couloir'.
 (6) ? *On semenil po koridoru, ty dolžen byl ego vstretil'* 'Il trottinait dans le couloir, tu aurais dû le rencontrer'.

La phrase (7) est tout aussi maladroite, voire davantage. Dans cet énoncé, la négation porte sur la composante du sens placée au focus – les mouvements spécifiques des jambes –, mais on ne comprend pas en quoi l'absence d'une façon particulière de marcher serait liée à la possibilité de rencontrer ou non le sujet.

- (7) ? *On ne semenil po koridoru, ja ego ne vstretil* 'Il ne trottinait pas dans le couloir, et je ne l'ai pas rencontré'.

Il en va de même avec le verbe *šagat'* :

- (8) ? *On šagal po koridoru, ty dolžen byl ego vstretit'* 'Il marchait à grands pas dans le couloir, tu aurais dû le rencontrer'.
 (9) ? *On ne šagal po koridoru, ja ego ne vstretil* 'Il ne marchait pas à grands pas dans le couloir, et je ne l'ai pas rencontré'.

À première vue, à l'instar du verbe *šagat'*, *semenit'*, sont formés des verbes qui désignent des actions physiques courantes, souvent liées aux travaux agricoles ou aux activités domestiques. Par exemple :

- *kosit'* (*travu*) 'faucher [l'herbe]'
- *molotit'* (*cepom pšenicu*) 'battre [le blé au fléau]'
- *pilit'* (*brevno*) 'scier [une poutre]'

- *rubit' (drova)* 'fendre [du bois]'
- *ryt' (kanavu)* 'creuser [un fossé]'
- *teret' (na těrke jabloko)* 'râper [une pomme, à la râpe]'
- *tkat' (polotno)* 'tisser [une toile]'
- *vjazat' (sviter)* 'tricoter [un pull]'
- *žat' (pšenicu)* 'moissonner [le blé]'

Les actions désignées par ces verbes impliquent des mouvements relativement fréquents et réguliers des mains tenant un outil déterminé¹⁰. *A priori*, de tels verbes devraient constituer des représentants exemplaires de la classe des multiplicatifs. Pourtant, les linguistes ne les y rattachent pas. Dans les dictionnaires (Evgen'eva, 1981–1984; Ušakov, 1935–1940), ces lexèmes sont définis par rapport à une action générale, envisagée uniquement sous l'angle de son but. Par exemple :

- *kosit' 'faucher' et žat' 'moissonner' → srezát' 'couper'*
- *rubit' 'couper à la hache' → razdeljat' na časti 'diviser en parties'*
- *tkat' 'tisser' → izgotavlivat' polotno putëm spletenija nitej 'fabriquer une étoffe par l'entrelacement de fils'*

Or, les dictionnaires ne précisent pas que ces lexèmes renvoient à des mouvements périodiques des mains (ou, avec le développement des techniques, à des mouvements périodiques de parties d'un mécanisme). Il est possible que cette indication ne fasse pas partie du sens même du lexème : ces mots désignent l'action et son but, sans égard pour le mode concret de sa réalisation. Le moyen habituel de l'exécution de l'action est généralement sous-entendu et ne retient pas l'attention ; il ne fait donc pas partie de leur sens. C'est pourquoi les verbes désignant une activité orientée vers un but ne sont pas perçus comme multiplicatifs, même lorsqu'ils se réalisent à travers une série de mouvements rythmés du même type. Tel est le cas, par exemple, de *pilit' 'scier', ryt' 'creuser', skoblit' 'racler', strogat' 'raboter', čistit' 'nettoyer'* (Knjazev, 2007 : 446).

Si, au contraire, le verbe renvoie à un mouvement rythmé, même produit par l'homme, mais que la référence au but ne fasse pas partie de son sens (bien que ce dernier soit souvent déductible du contexte), il est naturel de considérer ce verbe comme appartenant à la classe des multiplicatifs. Par exemple :

- *kačat' (kolybel')* 'balancer [un berceau]'
- *ukačivat' (reběnka)* 'bercer [un enfant]'
- *trjasti (jablonju)* 'secouer [un pommier]'

¹⁰ Ces outils sont (ou étaient) universellement répandus : la faux, la faucille, la scie, la hache, le fléau, la pelle, les aiguilles à tricoter ; et, par exemple, le tissage implique que l'on déplace à la main la navette du métier à tisser.

- *vstrjaxivat'* (*probirku*) 'secouer [une éprouvette]'
- *maxat'* (*rukoj*) 'agiter [la main], faire signe'
- *podtjagivat'sja* (*na turnike*) 'faire des tractions [sur une barre]'
- *gresti* (*veslami*) 'ramer'
- *boltat'* (*nogami*) 'remuer [les jambes]'
- *xlopat'* (*v ladoši*) 'taper des mains, applaudir'
- *viljat'* (*xvostom*) 'remuer [la queue]'
- *časat'* (*spinu*) 'se gratter [le dos]'
- *bryzgat'* (*vodoj*) 'asperger [d'eau]' et ainsi de suite.

On voit donc qu'il ne suffit pas, pour qu'un verbe soit considéré comme multiplicatif, qu'il désigne un procès composé d'actions répétées de manière périodique ; il faut encore que ces actions soient placées au centre de l'attention. Si le verbe exprime une action orientée vers un but, accomplie de la manière la plus commune, la référence à son mode de réalisation devient superflue : celui-ci va de soi. L'assertion de tels verbes porte alors sur le but et non sur le moyen.

Les actes répétés de même nature attirent naturellement l'attention lorsqu'ils ne sont pas tout à fait communs. Cf. *šagat'* 'marcher d'un pas appuyé' et *semit'* 'trotter' vs *idti*, *xodit'* 'marcher'. De tels actes accompagnent souvent un procès ordinaire impliquant un sujet humain, mais ils sont néanmoins perçus comme un écart par rapport à la norme. C'est le cas, par exemple, des actions répétées d'émission de sons « inconvenants » qui accompagnent des procès ordinaires tels que manger, respirer, marcher. Par exemple :

- *čavkat'* 'manger bruyamment'
- *xlebat'* 'boire à grosses gorgées'
- *sopet'* 'renifler'
- *šmygat'* (*nosom*) 'renifler'
- *smorkat'sja* 'se moucher'
- *xrapet'* 'ronfler'
- *šarkat'* 'traîner les pieds en faisant du bruit'
- *topat'* 'taper du pied, marcher bruyamment' (cf. *Ne topaj, xodi tiše* 'Ne fais pas autant de bruit, marche plus doucement'), etc.

Pour une large part, on trouve dans la classe des multiplicatifs des verbes désignant l'émission répétée de son ou de lumière¹¹, phénomènes qui attirent naturellement l'attention. C'est le cas, par exemple, de certains procès d'ordre naturel ou non, en tout cas non contrôlés :

¹¹ Ce phénomène est bien connu et a fait l'objet d'analyses approfondies depuis longtemps. On en trouvera une présentation détaillée, accompagnée d'une bibliographie substantielle, chez Knjazev (2007 : 443–446).

- *mercat'* 'scintiller'
- *sverkat'* 'briller'
- *pobleskivat'* 'scintiller'
- *žurčat'* 'murmurer (pour l'eau)'
- *kapat'* 'goutter, couler'
- *zvenet'* 'tinter, bruire'
- *treščat'* 'crépiter'
- *skripet'* 'grincer'
- *drebezžat'* 'trembler'
- *gromyxat'* 'gronder'
- *ljazgat'* 'cliqueter'
- *šuršat'* 'bruire, froufrouter'
- *šelestet'* 'bruire, froufrouter'

Sont également concernés des verbes (souvent idéophones) qui désignent des sons émis par des animaux :

- *kukarekat'* 'faire cocorico'
- *mjaukat'* 'miauler'
- *kukovat'* 'coucouler'
- *karkat'* 'croasser'
- *xrjukat'* 'grogner'
- *lajat'* 'aboyer'
- *tjavkat'* 'japper'

Rappelons, comme nous l'avons déjà noté plus haut, qu'une classification sémantique détaillée des multiplicatifs russes est proposée dans Birjulin (2024)¹².

Notons certains détails sémantiques qui ne semblent pas avoir retenu l'attention des chercheurs.

Premièrement, la notion de « périodicité relativement élevée » des actes ne fait pas partie de la sémantique de tous les multiplicatifs. Ainsi, le verbe *migat'* 'clignoter' ou le verbe *kukarekat'* 'chanter (pour un coq)' n'impliquent pas de périodicité régulière.

(10) *U nas po večeram svet migaet* 'Le soir, la lumière clignote chez nous.'

(11) *Petux xodil po dvoru i kukarekal* 'Le coq marchait dans la cour et chantait.'

¹² À la différence de Birjulin (2024), nous ne considérons pas comme multiplicatifs les verbes qui désignent le déplacement régulier d'un objet autour d'un cercle : le tour complet accompli par l'objet, en l'absence de pause au point de départ, n'est pas, de prime abord, perçu comme un « quantum » de déplacement.

Ce qui importe, c'est que les actes de même nature qui composent le procès se répètent dans un même « round d'observation »¹³, même lorsque ces répétitions ne sont pas régulières.

Ensuite, les actes (les quanta) d'un procès multiplicatif peuvent se distinguer les uns des autres avec plus ou moins de netteté. Par exemple, *mjaukat'* 'miauler' et *murlykat'* 'ronronner'.

(12) *Koška mjaukala i tērlas' o nogi* 'Le chat miaulait et se frottait contre nos jambes'.

(13) *Koška murlykala i izgibala spinu* 'Le chat ronronnait et courbait le dos'.

Le verbe *mjaukat'* réfère à des miaulements assez clairement séparés les uns des autres. *Murlykat'* suggère plutôt un ronronnement continu, néanmoins décomposable en actes isolés : le quantum d'une telle succession est un son prolongé dont la hauteur et l'intensité varient, la frontière entre deux quanta correspondant à l'intervalle où est produit le son le plus grave et le plus faible.

On pourra distinguer de manière analogue les verbes *bul'kat'* 'glouglouter' et *žurčat'* 'bruire', ou encore *vsxrapyvat'* 'ronfler (par moments)' et *xrapet'* 'ronfler'.

(14) *Voda bul'kala, vylivajas' iz butylki* 'L'eau glougloutait en sortant de la bouteille'.

(15) *Voda žurčit v aryke* 'L'eau bruit dans le canal d'irrigation'.

(16) *Vo sne on vsxrapyval* 'Il ronflait [par moments] dans son sommeil'.

(17) *Vo sne on xrapel* 'Il dormait en ronflant'.

Il est naturel de supposer que, si les actes composant le procès multiplicatif sont bien distincts les uns des autres, la langue sera pourvue d'un semelfactif, c'est-à-dire d'un verbe qui désigne un seul de ces actes.

Prenons les verbes *migat'* 'clignoter' et *mercat'* 'scintiller', dans des contextes où ils renvoient à l'émission de lumière :

(18) *My stoim i smotrim, kak migajut zvezdy* 'Nous contemplons les étoiles qui clignent¹⁴. [V. Nekrasov]

(19) *Zvezdy ele merkali na blednom nebe* 'Les étoiles scintillaient à peine dans le ciel pâle'. [Ju. Dombrovskij]

¹³ La notion de « round d'observation » a été introduite par Ju. Apresjan (2004b).

¹⁴ NdT : Pour la clarté de l'analyse, nous traduisons systématiquement *migat'* par 'clignoter' et *mercat'* par 'scintiller', même si ces équivalents ne conviennent pas parfaitement dans tous les contextes.

- (20) *Plaval i migal v uglu krasnyj ogoněk lampady* 'La petite flamme rouge de la veilleuse vacillait et clignotait dans le coin.' [A. Zajcev]
- (21) *Na uglu kakogo-to pereulka slabo mercal ogoněk v kioske* 'À l'angle d'une ruelle scintillait faiblement une lumière dans un kiosque.' [M. Bulgakov]
- (22) *Nad golovoj migali i potreskivali lampy dnevnogo sveta.* 'Au-dessus de [nos] têtes, les néons clignotaient et crépitaient.' [S. Dovlatov]
- (23) *Sleva mercala krasnaja lampočka pul'ta* 'À gauche scintillait une petite lumière rouge de la télécommande.' [S. Dovlatov]

Les deux verbes désignent une lumière inégale, plus ou moins vive ; avec *migat'*, les intervalles entre les pics de luminosité sont plus longs¹⁵, et il se peut aussi que *migat'* implique des éclairs plus vifs. Autrement dit, les deux verbes renvoient à un procès multiplicatif d'émission de lumière, mais *migat'*, à la différence de *mercat'*, suppose des quanta bien distincts les uns des autres.

C'est pourquoi le multiplicatif *migat'* forme un couple aspectuel avec le semelfactif *mignut'* (24) alors que *mercat'* n'a pas de semelfactif :

- (24) *Lampa mignula i pogasla* 'La lampe a clignoté [une fois] puis s'est éteinte.'

On trouvera une différence analogue entre les verbes *šagat'* 'marcher d'un pas appuyé' et *semenit'* 'trotter'. Le quantum du procès décrit par le premier suppose une certaine amplitude dans le mouvement de la jambe, et donc une certaine durée ainsi qu'une certaine distance entre les pas. Les quanta du procès décrit par *šagat'* apparaissent ainsi nettement distincts les uns des autres. Dans le cas de *semenit'*, l'amplitude du mouvement de la jambe, et par conséquent la durée du mouvement et la longueur du pas, sont moindres. Les quanta du procès décrit par *semenit'* se distinguent donc moins nettement les uns des autres que ceux du procès décrit par *šagat'*. C'est pourquoi *šagat'* forme un couple aspectuel avec le semelfactif *šagnut'*, tandis que *semenit'* n'a pas de semelfactif correspondant.

Ainsi, l'absence de semelfactif associé à un multiplicatif peut s'expliquer par des facteurs relevant du système de la langue.

Cette hypothèse pourrait être confirmée ou infirmée par l'analyse, sous cet angle, de l'ensemble des verbes russes classés – par tous les chercheurs, ou du

¹⁵ R. Frumkina a montré expérimentalement que le choix entre *migat'* et *mercat'* dépend de la fréquence des impulsions lumineuses : les sujets de l'expérience optent pour *migat'* lorsque l'intervalle entre les impulsions est sensiblement plus long (Frumkina, 1984).

moins par la majorité d'entre eux – parmi les multiplicatifs. Une telle analyse dépasse toutefois le cadre de la présente étude¹⁶.

2.2. Multiplicatifs stricts et non stricts

Considérons les deux exemples suivants :

(25) *Regulirovščica mašet flažkom* 'La régulatrice agite un drapeau.'

(26) *Devočka robko gladit sobaku* 'La fillette caresse timidement le chien.'

On y rencontre les multiplicatifs *maxat'* (*flažkom*), qui signifie en simplifiant 'lever et baisser la main avec un drapeau à plusieurs reprises et sans interruption', et *gladit'* (*sobaku*), qui signifie 'passer la main sur le haut du corps d'un chien à plusieurs reprises et sans interruption'. Cette représentation doit toutefois être complétée.

En effet, l'énoncé (25), comme l'énoncé (26), admet deux interprétations. Dans la première, le procès décrit se compose effectivement d'une série d'« actes » ou de « quanta » : la régulatrice lève et abaisse plusieurs fois le drapeau ; la fillette passe plusieurs fois sa main sur le corps du chien. Dans la seconde, le procès décrit se compose en tout et pour tout d'un seul acte (un « quantum ») : la régulatrice a agité le drapeau une fois ; la fillette a passé sa paume sur le corps du chien une seule fois¹⁷.

¹⁶ Certains semelfactifs qui ont leur entrée dans le dictionnaire (Birjulin, 2024) ne sont représentés que par un nombre infime d'occurrences dans le *Corpus national de la langue russe* (ruscorpora.ru) et apparaissent de toute évidence comme des occasionnalismes. Ainsi *murlyk-nut'* 'ronronner une fois' (semelfactif formé sur *murlykat'* 'ronronner') n'est attesté que 6 fois, contre 159 occurrences pour *mjauknut'* 'miauler une fois' ; *kudaxtnut'* 'caqueter une fois' (formé sur *kudaxtat'* 'caqueter'), 6 occurrences ; *česanut'* (*zatylok*) 'se gratter une fois [la tête]' (formé sur *česat'* (*zatylok*) 'se gratter [la tête]'), 3 occurrences (sous-corpus des textes littéraires entre 1800 et 2020). Dans ce dernier cas, le verbe a aussi une valeur figurée, avec le sens de 'se tourner les pouces' ou de 'se creuser la tête'. La corrélation entre le caractère des quanta d'un procès composé d'actes répétés et l'existence d'un semelfactif associé au multiplicatif considéré relève donc plutôt d'une tendance que d'une norme stricte. Une autre difficulté rencontrée dans la description des semelfactifs tient au fait que le suffixe *-nu-* est polysémique. Il convient donc de distinguer le véritable semelfactif du verbe à suffixe *-nu-* dépourvu de valeur semelfactive. Cf. *rexnut'sja* 'craquer', *otprjanut'* 'se reculer brusquement', *očnut'sja* 'reprenre connaissance', etc. Certains de ces lexèmes sont en réalité des dérivés sémantiques du semelfactif (Makarčuk, 2021).

¹⁷ Cette ambiguïté du multiplicatif est pertinente d'un point de vue typologique. Dans certaines langues, elle est levée par la forme verbale elle-même. Ainsi, en anglais, *kick* renvoie à une action répétée dans certaines de ses formes, et à une action unique dans d'autres. De même, en

Ainsi, dans l'exemple (25), *Regulirovščica mašet flažkom* 'La régulatrice agite un drapeau', l'énoncé au présent historique peut décrire une situation où la régulatrice n'a effectué qu'un seul mouvement, cf. :

- (27) *Včera idu ja po Elisejskim poljam i vižu: regulirovščica mašet flažkom, i mašiny nemedlenno ostanavlivajutsja*
'Hier, je me promène sur les Champs-Élysées et je vois : la régulatrice agite un drapeau, et les voitures s'arrêtent immédiatement'.

L'énoncé (28) peut être compris de la même manière 'La fillette effleure d'un geste le dos du chien et s'enfuit aussitôt' :

- (28) *Devočka robko gladit sobaku i srazu ubegajet*
'La fillette caresse timidement le chien et s'enfuit aussitôt'.

Ce fait a été décrit dans différents travaux (Bondarko & Bulanin, 1967 ; Jakovlev, 1975 ; Glovinskaja, 1982 ; Xrakovskij, 1989 ; Birjulin, 2001). Nous en concluons que le sens de tels multiplicatifs comprend le trait sémantique 'une fois ou plus d'une fois'. C'est ce trait qui explique la double interprétation des exemples (25) et (26).

Notons que dans certains contextes, cette ambiguïté est levée.

- (29) *On dolgo maxal flažkom* 'Il a agité un drapeau pendant longtemps'.

Dans ce cas, seule la première interprétation de la disjonction est retenue.

Dans notre approche, de tels multiplicatifs relèvent de la classe, relativement étendue, des prédicats porteurs d'une organisation disjonctive du sens (Apresjan, 1974 ; Uryson, 1998).

Toutefois, tous les verbes multiplicatifs n'obéissent pas à ce fonctionnement. Ainsi, des verbes comme *trjasti* 'secouer' (dans ses différents sens : *trjasti probirku* 'secouer une éprouvette', *Avtobus trjasët* 'Le bus cahote'), *drožat* 'trembler' (*Ruki drožat* 'Les mains tremblent'), *kolebat'sja* 'vaciller' (*Plamja kolebletsja* 'La flamme vacille'), *kolyxat'sja* 'onduler', *vibrirovat* 'vibrer', *mercat* 'scintiller', *poskripyvat* 'grincer', etc., désignent des procès toujours constitués d'une multitude d'actes de même nature, répétés avec une périodicité élevée. Les énoncés où ils apparaissent n'admettent donc pas d'autre interprétation.

français, *cligner (des yeux)* à l'imparfait réfère à une succession d'actes, tandis qu'au passé composé, le même verbe désigne un acte unique (Xrakovskij, 1989 : 26).

Ainsi, la classe des verbes multiplicatifs peut être divisée en deux groupes.

Le premier comprend les verbes du type *trjasti* ‘secouer’, *vibrirovat’* ‘vibrer’, *mercat’* ‘scintiller’, qui renvoient nécessairement à un procès composé d’actes répétés, c’est-à-dire d’actes ayant lieu plus d’une fois. Nous qualifierons les verbes de ce type de « multiplicatifs stricts ».

Le second groupe comprend les verbes du type *maxat’* ‘agiter’, *gladit’* (*sobaku*) ‘caresser [un chien]’, qui admettent deux interprétations : le procès peut se composer de plusieurs actes, mais aussi d’un seul.

Nous qualifierons les verbes de ce type de « multiplicatifs non stricts ». Nous établirons la même distinction entre les procès correspondants.

Comme on le voit dans cette définition, les multiplicatifs non stricts se rapprochent, par leur sens, de certains procès ordinaires : un procès ordinaire peut se répéter ou non au cours d’un laps de temps assez court, et il faut être capable de distinguer le procès composé d’un acte unique du procès ordinaire bref (tout comme il faut pouvoir distinguer la simple répétition du procès de la succession d’actes dans un procès multiplicatif).

Prenons par exemple les couples *maxat’*^{IPF} / *maxnul’*^{PF} (*flažkom*) ‘agiter [un drapeau]’ et *sžimat’*^{IPF} / *sžat’*^{PF} (*ruki*) ‘serrer [les mains]’. Dans les deux cas, le verbe perfectif peut apparaître dans un contexte où l’action se produit plusieurs fois :

(30) *Signal’sčik dva raza maxnul’ flažkom* ‘La vigie a agité le drapeau deux fois.’

(31) *On dva raza krepko sžat’ ej ruki* ‘Il lui a serré les mains deux fois.’

Pour autant, il est clair que *maxat’* est un verbe multiplicatif, tandis que *sžimat’* ne l’est pas.

Notons que l’imperfectif *maxat’*, sans soutien contextuel, peut désigner soit un procès composé d’un acte unique, soit un procès multiplicatif :

(32) *Signal’sčik mašet flažkom* ‘La vigie agite le drapeau.’

(S’il s’agit d’un acte unique, cet énoncé est équivalent à *Signal’sčik vsmaxivaet flažkom* ‘La vigie donne un coup de drapeau.’)

En revanche, *sžimat’*, sans soutien contextuel, ne peut désigner qu’un procès ordinaire, non répété :

(33) *On sžimaet ej ruki* ‘Il lui serre les mains.’

Prenons encore un autre exemple : *ulybat'sja* / *ulybnut'sja* 'sourire' et *podtjagivat'sja* / *podtjanut'sja* (*na turnike*) 'faire des tractions [à la barre]'. Dans les deux cas, le perfectif peut, en contexte, se rapporter à un nombre déterminé d'actions :

(34) *Ona tri raza emu ulybnulas'*^{PF} 'Elle lui a souri trois fois.'

(35) *Ona tri raza podtjanulas'*^{PF} *na turnike* 'Elle a fait trois tractions à la barre.'

Cependant, l'imperfectif *podtjagivat'sja*, en dehors d'un contexte de répétition, peut naturellement être compris comme la désignation d'un procès multiplicatif, bien qu'il admette également une interprétation où l'action n'est effectuée qu'une seule fois :

(36) *Podxodit k turniku i načinaet podtjagivat'sja*^{IPF} 'Il s'approche de la barre et commence à faire des tractions / une traction.'

À l'inverse, le verbe imperfectif *ulybat'sja*, sans soutien contextuel, est compris comme la désignation d'un procès ordinaire, non répété :

(37) *Podxodit k zerkalu i načinaet ulybat'sja*^{IPF} 'Il s'approche du miroir et commence à sourire.' (Ses lèvres s'écartent pour former un sourire une seule fois).

On voit donc que *ulybat'sja* n'est pas un multiplicatif, tandis que *podtjagivat'sja* est un multiplicatif non strict.

Il est naturel de considérer qu'un verbe est un multiplicatif non strict dès lors qu'il peut être compris comme désignant un procès multiplicatif sans qu'un contexte particulier vienne soutenir une telle interprétation (cf. Tatevosov, 2016 : 86). Dans le cas contraire, le verbe n'appartient pas à la classe des multiplicatifs.

Remarquons que, si le multiplicatif est par excellence un verbe imperfectif, le semelfactif est, quant à lui, toujours un verbe perfectif. Cf. *ščëlkat'* – *ščëlknut'* (*pal'cami*) 'claquer [des doigts]', *stučat'* – *stuknut'* (*v okoško*) 'toquer [à la fenêtre]', *zevat'* – *zevnut'* 'bâiller', etc.

De plus, d'après le critère de Maslov (Maslov, 2004), le multiplicatif non strict imperfectif (par ex. *maxat'* 'agiter', *podtjagivat'sja* 'faire des tractions') forme un couple aspectuel avec le verbe perfectif semelfactif correspondant (*maxnut'* 'donner un coup de drapeau', *podtjanut'sja* 'faire une traction'). Toutefois, le sens lexical des deux membres du couple n'est pas tout à fait identique : si l'imperfectif peut désigner un ou plusieurs actes du procès, le semelfactif, lui, ne renvoie qu'à un seul acte du même procès.

Il se peut qu'il existe des critères formels permettant de distinguer les multiplicatifs stricts des non stricts en fonction de leur compatibilité avec certains contextes, mais nous n'en avons pas connaissance, pas plus que nous n'avons connaissance de travaux consacrés à ce problème. En outre, les multiplicatifs non stricts peuvent eux-mêmes être subdivisés en différentes classes selon la préférence plus ou moins marquée donnée à leur interprétation itérative (sur le plan typologique, cette question est abordée dans (Xrakovskij, 1989 : 26–27). Une description un tant soit peu exhaustive des verbes multiplicatifs envisagée sous cet angle dépasserait le cadre de la présente étude.

Notons que certains multiplicatifs stricts forment des couples synonymiques avec des multiplicatifs non stricts. La distinction entre ces synonymes tient au fait que le multiplicatif strict désigne nécessairement un procès multiplicatif, tandis que le multiplicatif non strict peut désigner aussi bien un procès multiplicatif qu'un procès composé d'un seul acte. Voici quelques exemples :

- *trjasti* – *vstrjaxivat'* 'secouer' : *Provizor trjaset probirku* 'Le pharmacien secoue l'éprouvette' (le mouvement est effectué plusieurs fois) – *Provizor vstrjaxivajet probirku* 'Le pharmacien secoue l'éprouvette' (le mouvement est effectué plusieurs fois ou bien une seule fois).
- *mercat'* – *migat'* 'scintiller – clignoter' : *Zvezda mercaet* 'L'étoile scintille' (multitude d'actes de même nature) – *Lampa migaet* 'La lampe clignote' (un acte ou plusieurs).

Ainsi, la distinction entre une récurrence des actes obligatoire et facultative peut-elle servir de caractère distinctif entre de tels synonymes.

Les multiplicatifs non stricts représentent un intérêt particulier pour nous. Voici notre hypothèse : le multiplicatif strict n'a pas de semelfactif associé, contrairement au multiplicatif non strict, qui en est *a priori* pourvu.

En effet, le sens du multiplicatif non strict comprend la référence à un quantum du procès en question. De ce fait, ce quantum se détache de la succession des quanta et acquiert une importance pragmatique. Mais s'il en est ainsi, le multiplicatif en question dispose probablement d'un semelfactif associé, avec lequel il forme un couple aspectuel. Le multiplicatif strict, au contraire, ne renvoie jamais à un seul acte du procès, quel que soit le contexte. La raison en est peut-être que les actes d'un tel procès ne sont pas assez distincts les uns des autres, ou qu'un seul acte a peu de chances de revêtir une importance pragmatique pour le locuteur. En tout cas, la langue ne semble pas disposer d'un verbe désignant un seul acte de ce procès : autrement dit, il n'existe pas de semelfactif associé à ce multiplicatif.

2.3. Les différentes désignations possibles du quantum d'un procès multiplicatif

On se rend compte assez vite qu'il existe des verbes multiplicatifs non stricts qui, contre toute attente, ne forment pas de couple aspectuel avec un semelfactif. Prenons le verbe *celovat'* 'embrasser'. Ce verbe imperfectif est un multiplicatif non strict :

- (38) *Korol' videl, kak šut celoval^{IPF} koroleve ruku* 'Le roi vit le bouffon embrasser la main de la reine' (un seul contact des lèvres avec la main de la reine, ou plusieurs).

Mais il n'existe pas, dans la langue littéraire, de verbe *celovanut'*, en dépit du fait que l'énoncé (38) peut être compris comme la description d'un acte unique. On peut supposer que, dans de tels cas, en l'absence de semelfactif perfectif en *-(a)nu-*, la langue dispose d'autres moyens pour désigner un seul acte d'un procès multiplicatif non strict dont les quanta sont bien distincts. En effet, un seul acte d'un tel procès peut être désigné par le verbe perfectif qui forme un couple aspectuel avec l'imperfectif en question.

Un seul quantum de ce procès est exprimé par le perfectif *pocelovat'* :

- (39) *Šut poceloval^{PF} koroleve ruku* 'Le bouffon embrassa la main de la reine.'

Il est en outre possible de préciser le nombre de quanta :

- (40) *Šut triždy poceloval^{PF} koroleve ruku* 'Le bouffon embrassa trois fois la main de la reine.'

Cet énoncé est analogue à :

- (41) *On triždy vzmaxnul^{PF} flažkom* 'Il agita le drapeau trois fois.'

Ainsi, le perfectif *pocelovat'* peut-il être vu comme une sorte d'équivalent du semelfactif en *-(a)nu-*.

On notera que ce moyen de désigner un quantum d'un procès multiplicatif est également possible dans certains cas où le multiplicatif dispose formellement d'un semelfactif associé. Prenons le verbe *kolot' (igloj)* 'piquer [avec une aiguille]' et son semelfactif *kol'nut'*.

- (42) *Koldun'ja kolet^{IPF} princa otravlennoj igloj* ‘La sorcière pique le prince avec une aiguille empoisonnée’ (une fois ou plus d’une fois).

Pour renvoyer à un quantum de ce procès, on emploie généralement le perfectif *ukolot'*:

- (43) *Koldun'ja ukolola^{PF} princa otravlennoj igloj* ‘La sorcière piqua le prince avec une aiguille empoisonnée’.

L’énoncé suivant est également possible :

- (44) *Koldun'ja kol'nula^{PF} princa igloj* ‘La sorcière piqua le prince d’un coup d’aiguille’.

Un quantum isolé du procès *kolot'* ‘piquer’ peut toujours être désigné par *ukolot'*, tandis que *kol'nut'* renvoie à certaines spécificités de l’acte (un contact de l’aiguille avec la peau extrêmement bref ou léger)¹⁸. *Ukolot'*, tout comme *kol'nut'*, peut être employé lorsque le nombre de quanta est précisé :

- (45) *Koldun'ja tri raza ukolola princa otravlennoj igloj* ‘La sorcière piqua trois fois le prince avec une aiguille empoisonnée’.

Toutefois, dans certains cas, le verbe perfectif formant un couple aspectuel avec l’imperfectif multiplicatif admet lui aussi cette double interprétation : il peut renvoyer à un seul acte ou à une succession d’actes. Ainsi de *pogladit'* (*sobaku*) ‘caresser [un chien]’ :

- (46) *Devočka robko pogladila^{PF} <gladila^{IPF}> sobaku* ‘La fillette a timidement caressé <caressait > le chien’ = ‘La fillette a passé sa main sur le chien une fois ou plus d’une fois’.

L’ambiguïté disparaît si le nombre d’actes est précisé :

- (47) *Devočka odin raz <tri raza> robko pogladila^{PF} sobaku* ‘La fillette a timidement caressé le chien une fois <trois fois>’.

¹⁸ La spécificité sémantique des semelfactifs en *-(a)nu* est décrite avec précision dans Zalizniak & Šmelëv (2015 : 125–127).

Les multiplicatifs non stricts *zvonit'* (*v dver'*) 'sonner [à la porte]' et *stučat'* (*v okoško*) 'toquer [à la fenêtre]' ont un fonctionnement similaire : ils peuvent désigner un seul coup aussi bien qu'une série de coups.

(48) *Podxodit'*^{IPF} *k dveri i zvonit* 'Il s'approche de la porte et sonne.'

(49) *Podkradyvaetsja*^{IPF} *k ee oknu i stučit* 'Il se faufile vers sa fenêtre et toque.'

Les perfectifs correspondants conservent cette ambiguïté :

(50) *Podošel k znakomoj dveri i pozvonil*^{PF} 'Il s'est approché de la porte familière et a sonné.'

(51) *Podkralsja k zavetnomu oknu i postučal*^{PF} 'Il s'est faufile vers la fenêtre secrète et a toqué.'

L'ambiguïté disparaît si le contexte explicite le nombre d'actes :

(52) *On pozvonil*^{PF} *v dver' <postučal*^{IPF} *v okno> vsego odin raz* 'Il a sonné à la porte <a toqué à la fenêtre> une seule fois.'

(53) *On tri raza pozvonil*^{PF} *v dver' <postučal*^{PF} *v okno>* 'Il a sonné à la porte <a toqué à la fenêtre> trois fois'¹⁹.

Ainsi, les verbes perfectifs *pogladit'*, *pozvonit'*, *postučat'* peuvent-ils s'employer pour désigner un quantum d'un procès multiplicatif. Notons que *stučat'* (*v okno*) 'toquer [à la fenêtre]' a bien un semelfactif *stuknut'* (*v okno*) 'donner un coup à la fenêtre', mais qu'un seul acte du procès peut néanmoins être également désigné par le perfectif *postučat'*.

L'analyse de ces verbes montre que la classe des multiplicatifs non stricts existe en russe, indépendamment de l'existence d'un semelfactif associé. Les semelfactifs représentent en effet une réalisation particulièrement fréquente du mode unique de l'action, et l'étude des multiplicatifs impose de prendre en compte ce mode dans son ensemble.

¹⁹ Le fonctionnement des verbes multiplicatifs non stricts dans un contexte où le nombre d'actions est déterminé est décrit de manière plus détaillée dans (Uryson, 2025).

3. La dérivation sémantique dans le couple multiplicatif – semelfactif

L'interprétation sémantique que nous proposons des multiplicatifs non stricts permet de répondre à la question de la direction de la dérivation sémantique dans le couple multiplicatif – semelfactif (cf. *maxat'* – *maxnut'*), autrement dit de déterminer « lequel de ces sens est le plus simple et le plus fondamental, et lequel est le plus complexe et dérivé » (Xrakovskij, 1998 : 489). Reprenant une métaphore de Plungjan (2011), le problème peut être reformulé ainsi : un procès multiplicatif est comparable à une chaîne constituée de maillons. Reste à savoir si la chaîne est assemblée à partir de maillons premiers (dans ce cas, le semelfactif est la forme de base, et le multiplicatif la forme dérivée), ou si, à l'inverse, un maillon est extrait d'une chaîne déjà existante (le multiplicatif étant alors premier, et le semelfactif dérivé).

Cette question n'est pas pertinente pour le multiplicatif strict, qui n'a pas de semelfactif associé : dans ce cas, seule la chaîne est désignée, et non le maillon isolé.

Pour le multiplicatif non strict, sa sémantique renvoie simultanément à « un maillon de la chaîne » (un quantum du procès) et à « la chaîne dans son ensemble » (la succession de ces actes). Ces deux valeurs sont unies par une relation de disjonction ('ou'), que l'on peut schématiser ainsi : 'P une fois ou plus d'une fois' (où P désigne le procès indépendamment de sa segmentation en quanta). D'un point de vue logique, les deux composantes – 'maillon' et 'chaîne' – ont le même statut, puisqu'elles sont reliées par 'ou'. La question de la direction de la dérivation dans le couple multiplicatif – semelfactif s'efface donc : aucun des deux sens n'est premier ni dérivé par rapport à l'autre.

Il reste que, dans la langue russe, un verbe spécifique existe pour désigner un seul acte d'un procès multiplicatif : le semelfactif. Cela ne signifie pas que, d'un point de vue logique (et non morphologique), le semelfactif soit dérivé du multiplicatif : il s'agit simplement du moyen par lequel la langue résout l'ambiguïté propre à la sémantique disjonctive.

La situation peut différer dans d'autres langues. Ainsi, en evenk, une sous-classe de multiplicatifs est marquée par un suffixe particulier, tandis que le semelfactif associé reçoit un autre suffixe distinct (Nedjalkov & Sverčkova, 1989).

La description typologique de la sémantique des multiplicatifs stricts et non stricts, ainsi que la typologie morphologique des couples multiplicatif – semelfactif, constituent un sujet à part entière.

Traduit du russe par Marie Starynkevitch

Références citées

- Apresjan, Ju. (1974). *Leksičeskaja semantika*. Nauka.
- Apresjan, Ju. (2004a). *Lingvističeskaja terminologija Slovarja*. In Ju. Apresjan (éd.), *Novyy ob"jasnitel'nyj slovar' sinonimov russkogo jazyka* (22–52). Jazyki slavjanskoj kul'tury – Wiener Slawistischer Almanach.
- Apresjan, Ju. (2004b). Sinonimičeskij rjad PRIVYKNUT' 1. In Ju. Apresjan (éd.), *Novyy ob"jasnitel'nyj slovar' sinonimov russkogo jazyka*. Jazyki slavjanskoj kul'tury – Wiener Slawistischer Almanach.
- Apresjan, Ju. (2006). Osnovanija sistemnoj leksikografii. In Ju. Apresjan (éd.), *Jazykovaja kartina mira i sistemnaja leksikografija* (33–160). Jazyki slavjanskix kul'tur.
- Birjulin, L. (2000). Mul'tiplikativnye vs. semel'faktivnye konstrukcii v russkom jazyke. In A. Barentsen & You. Poupynin (éds), *Functional Grammar: Aspect and Aspectuality. Tense and Temporality. Essays in honour of Alexander Bondarko* (23–42). Lincom Europa.
- Birjulin, L. (2024). *Tolkovo-sintaksičeskij slovar' russkix mul'tiplikativov i semel'faktivov*. ILI RAN.
- Bondarko, A. & Bulanin, L. (1967). *Russkij glagol*. Prosveščenie.
- Chafe, W. (1976). Givenness, Contrastivness, Definiteness, Subjects, Topics, and Points of View. In Ch.N. Li (éd.), *Subject and Topic* (27–55). Academic Press.
- Dickey, S. & Janda, L. (2009). *Xoxotnul, sxitril*: the relationship between semelfactives formed with -nu- and s- in Russian. *Russian Linguistics* 33, 229–248.
- Evgen'eva, A. (éd.) (1981–1984). *Slovar' russkogo jazyka. T. I-IV*. Russkij jazyk.
- Frumkina, R. (1984). *Cvet, smysl, sxodstvo*. Nauka.
- Glovinskaja, M. (1982). *Semantičeskie tipy vidovogo protivopostavljenija russkogo glagola*. Nauka.
- Jakovlev, V. (1975). Mnogoaktnost' kak sposob glagol'nogo dejstvija. *Filologičeskie nauki* 3, 97–105.
- Janda, L. & Makarova, A. (2009). Do It Once: A Case Study of the Russian -nu- Semelfactives. *Scando-Slavica* 55, 78–99.
- Knjazev, Ju. (2007). *Grammatičeskaja semantika: Russkij jazyk v tipologičeskoj perspektive*. Jazyki slavjanskix kul'tur.
- Makarčuk, I. (2021). K tipologii derivacij glagol'noj mery: semel'faktiv i delimitativ. *Voprosy jazykoznaniija* 6, 40–68.
- Maslov, Ju. (2004a). Sistema osnovnyx ponjatij i terminov slavjanskoj aspektologii. In Ju. Maslov, *Izbrannye trudy. Aspektologija. Obščee jazykoznanie* (365–395). Jazyki slavjanskix kul'tur.

- Maslov, Ju. (2004b). Vid i leksičeskoje značenie glagola v sovremennom russkom literaturnom jazyke. In Ju. Maslov (éd.), *Izbrannye trudy. Aspektologija. Obščee jazykoznanie* (71–90). Jazyki slavjanskix kul'tur.
- Mehlig, H. (1994). Gomogennost' i geterogennost' v prostranstve i vremeni: o kategorii glagol'nogo vida v russkom jazyke. *Revue des études slaves* 66, 585–606.
- Mel'čuk, I. (1998). *Kurs obščei morfologii. T. 2. Morfologičeskie značeniya*. In V. Plungjan, V. Percov & E. Savvina (éds), *Jazyki russkoj kul'tury*. Wiener Slawistischer Almanach.
- Nedjalkov, I. & Sverčkova, Ju. (1989). Vyraženie množestvennosti situacij v evenkijskom jazyke. In Xarkovskij (éd.), *Tipologija iterativnyx konstrukcij* (63–72). Nauka.
- Plungjan, V. (2011). *Vvedenije v grammatičeskiju semantiku: grammatičeskie značeniya i grammatičeskie sistemy jazykov mira*. RGGU.
- Ušakov, D. (éd.) (1935–1940). *Tolkovyj slovar' russkogo jazyka. T. 1–4*. OGIZ.
- Šluinskij, A. (2006). K tipologii predikatnoj množestvennosti: organizacija semantičeskoj zony. *Voprosy jazykoznanija* 1, 46–75.
- Tatevosov, S. (2016). *Glagol'nye klassy i tipologija akcional'nosti*. Jazyki slavjanskix kul'tur.
- Uryson, E. (1998). “Nesostojavšajasja polisemija” i nekotorye ee typy. *Semiotika i informatika* 36, 226–262.
- Uryson, E. (2025). Glagoly SOV i NESOV v kontekste mnogokratnosti: k vyjavleniju semantiki vidovyx grammem. *Voprosy jazykoznanija* 1, 78–94.
- Xrakovskij, V. (1989). Semantičeskie typy množestva situacij i ix estestvennaja klassifikacija. In V. Xrakovskij (éd.), *Tipologija iterativnyx konstrukcij* (5–53). Nauka.
- Xrakovskij, V. (1998). Tipologija semel'faktiva. In M. Čertkova (éd.), *Tipologija vida: problemy, poiski, rešenija* (485–490). Jazyki slavjanskix kul'tur.
- Zaliznjak, A. & Šmel'ev, A. (2015). Vvedenije v russkiju aspektologiju. In A. Zaliznjak, I. Mikaëljan & A. Šmel'ev, *Russkaja aspektologija: v zaščitu vidovoj pary* (15–151). Jazyki slavjanskix kul'tur.

Index alphabétique des verbes mentionnés, accompagnés de leur traduction

- *barabanit'* (*pal'cami po stolu*) ‘tambouriner [des doigts sur la table]’
- *begat'* ‘courir’
- *boltat'* (*nogami*) ‘remuer [les jambes]’
- *bryzgat'* (*vodoj*) ‘asperger [d’eau]’
- *bul'kat'* / *bul'knut'* ‘glouglouter’
- *celovat'* / *pocelovat'* ‘embrasser’
- *čavkat'* ‘manger bruyamment’

- *časat'* (*spinu*) 'se gratter [le dos]'
- *časat'sja* 'se gratter'
- *drebezžat'* 'trembler'
- *drožat'* 'trembler'
- *gladit'* / *pogladit'* (*sobaku*) 'caresser [un chien]'
- *gremet'* (*kastrjuljami*) 'faire du bruit [avec des casseroles]'
- *gresti* (*veslami*) 'ramer [avec des rames]'
- *gromyxat'* 'gronder'
- *idti* 'aller à pied [verbe déterminé]'
- *ikat'* / *iknut'* 'hoqueter'
- *iskrit'* 'faire des étincelles'
- *kapat'* 'goutter, couler'
- *karkat'* 'croasser'
- *kačat'* (*kolybel'*) 'balancer [un berceau]'
- *kolebat'sja* 'vaciller'
- *kolotit'* (*nogoj v dver'*) 'taper [du pied dans la porte]'
- *kolyxat'sja* 'onduler'
- *kosit'* (*travu*) 'faucher [l'herbe]'
- *kukarekat'* / *pokukarekat'* 'chanter « cocorico » [pour un coq]'
- *kukovat'* 'coucouler'
- *kusat'* / *ukusit'* 'mordre'
- *lizat'* / *liznut'*, *lizanut'* 'lécher'
- *ljazgat'* 'cliqueter'
- *lajat'* 'aboyer'
- *maxat'* – *zamaxat'* – *maxnut'* – *vzmaxnut'* (*flažkom, rukoj, rukami*) 'agiter [un drapeau, la main, les mains]'
- *mercat'* 'scintiller'
- *migat'* 'clignoter'
- *mjaukat'* 'miauler'
- *molotit'* (*cepom pšenicu*) 'battre [le blé au fléau]'
- *morgat'* / *morgnut'* 'cligner de l'œil'
- *murlykat'* 'ronronner'
- *pilit'* (*brevno*) 'scier [une poutre]'
- *pobleskivat'* 'scintiller'
- *podtjagivat'sja* / *podtjanut'sja* (*na turnike*) 'faire des tractions [à la barre]'
- *poskripyvat'* 'grincer'
- *poxixikat'* 'ricaner, glousser'
- *rubit'* (*drova*) 'fendre [du bois]'
- *ryt'* (*kanavu*) 'creuser [un fossé]'

- *semenit'* 'trotter'
- *skoblit'* 'racler'
- *skripet'* 'grincer'
- *smorkat'sja* 'se moucher'
- *sopet'* 'renifler'
- *stučat'* / *stuknut'* / *postučat'* (*v okoško*) 'toquer [à la fenêtre]'
- *sučit'* (*nogami*) 'taper [des pieds]'
- *sverkat'* 'briller'
- *sžimat'* / *sžat'* (*ruki*) 'serrer [les mains]'
- *šagat'* / *šagnut'* 'marcher'
- *šarkat'* 'faire du bruit en traînant [les pieds]'
- *šelestet'* 'bruire, froufrouter'
- *šmygat'* (*nosom*) 'renifler [du nez]'
- *šuršat'* 'bruire, froufrouter'
- *ščelkat'* / *ščelknut'* (*pal'cami*) 'claquer [des doigts]'
- *teret'* (*na těrke jabloko*) 'râper [une pomme à la râpe]'
- *tjavkat'* 'japper'
- *tkat'* (*polotno*) 'tisser [une toile]'
- *topat'* 'taper du pied, marcher bruyamment'
- *treščat'* 'crépiter'
- *trjasti* 'secouer, cahoter'
- *ukačivat'* (*reběnka*) 'bercer [un enfant]'
- *ulybat'sja* / *ulybnut'sja* 'sourire'
- *vertet'* (*golovoj*) 'tourner [la tête]'
- *vibrirovať'* 'vibrer'
- *viljat'* (*xvostom*) 'remuer [la queue]'
- *vjazat'* (*sviter*) 'tricoter [un pull]'
- *vstrjaxivat'* (*probirku*) 'secouer [une éprouvette]'
- *vsxrapyvat'* 'ronfler [par moments]'
- *xlebat'* 'boire à grosses gorgées'
- *xlopat'* / *xlopnut'* (*v ladoši* / *v ladoni*) 'taper [dans les mains]'
- *xodit'* 'aller de manière habituelle, marcher'
- *xrapet'* 'ronfler'
- *xrjukat'* 'grogner'
- *zevat'* / *zevnut'* 'bâiller'
- *zvenet'* 'tinter, bruire'
- *zvonit'* / *pozvonit'* (*v dver'*) 'sonner [à la porte]'
- *žat'* (*pšenicu*) 'moissonner [le blé]'
- *žurčat'* 'murmurer, bruire [pour l'eau]'



Valentina Apresjan

Dartmouth College
USA

 <https://orcid.org/0000-0003-1348-2224>

Vladimir Plungian

 <https://orcid.org/0000-0002-2393-1399>

Ekaterina Rakhilina

 <https://orcid.org/0000-0002-7126-0905>

Behind the door: how Russian, Kazakh, and Armenian encode spatial relations with double-fronted objects

Abstract

This paper explores how the concept of frontness is linguistically encoded in spatial constructions involving double-fronted objects, such as doors and windows, across three languages: Russian, Kazakh, and Armenian. Our analysis reveals significant cross-linguistic variation in how languages assign and interpret frontness in spatial expressions such as ‘in front of’ and ‘behind.’ Drawing on corpus data, field work with informants, and typological questionnaires, we show that the frontness of double-fronted objects such as doors and windows is not a universal feature, as one might have expected based on their experiential characteristics. Instead, spatial encoding varies across languages and appears to form different typological systems in this regard. Russian treats both doors and windows as fronted objects but only in two constructions – ‘in front of the door’ and ‘behind the window’, both referring to exterior location; Armenian assigns frontness primarily to windows; and Kazakh partially treats windows as fronted, while doors exhibit inconsistent patterns. These findings suggest that the encoding of spatial relations involving double-fronted objects forms a typologically diverse domain.

Key words

Frontness, frames of reference, egocentric, intrinsic, spatial relations, double-fronted objects, Russian, Kazakh, Armenian, deixis, typology

According to the classical definition by Juri Apresjan (1974, 1986), *frontness* is a non-trivial semantic feature that characterizes objects with a single functionally defined side, through which they are typically used. As previously noted by Fillmore (1968), frontness is not merely a geometric property but an experiential and functional one. The inherent front of objects, especially artifacts, is determined by their typical interaction with humans – for example, a house has a clear front where the door is located, while a tree does not.

In languages like English, Russian, and French, frontness is reflected in different interpretations of expressions with spatial prepositions such as ‘in front of’ or ‘behind’. Phrases with fronted nouns and prepositions ‘in front of’ or ‘behind’ are interpreted with an *intrinsic* frame of reference (Levinson 2003, Majid et al., 2004), or a type of spatial coding system where the location of an object is described based on the inherent orientation of another object rather than the observer’s perspective (relative egocentric frame of reference) or fixed environmental landmarks (absolute, or geocentric frame of reference). Thus, phrases like *zelënaja izgorod' pered domom*, *hedge in front of the cottage*, or *la haie devant la maison* all refer to a growth of shrubs next to the front side of the house – i.e., the side where the main entrance door is located. For their understanding, they do not require knowledge about the position of the observer (or about the cardinal directions).

However, when used with non-fronted nouns, spatial prepositions invoke a relative egocentric frame of reference: in order to locate the object which is *pered derevom*, *in front of the tree* or *devant l'arbre* one has to know the position of the speaker-observer. Thus, Marc Chagall’s well-known lithographic poster *Les Amoureux Devant l'Arbre* (‘The lovers in front of the tree’) describes the location of the couple from the viewer’s position: the lovers are between the viewer and the painted tree in the poster.

Since frontness is defined functionally and experientially rather than purely linguistically, one might expect it to display a certain degree of cross-linguistic universality. After all, cultural differences notwithstanding, both humans and many culturally significant objects exhibit fronts that are similarly defined across cultures: the human face, with its eyes and mouth, universally marks the forward-facing side of the body, just as the front of a house is typically where the door, windows, or point of entry is located.

However, even within a single language, denotations of fronted objects do not always behave uniformly in accordance with their asymmetrical nature and intrinsic orientation. For instance, *a bike in front of the building* is likely to be understood as a bike positioned on the entrance side of the building. This interpretation arises for two reasons: when used as a landmark for a relatively small

object, an intrinsic understanding of building is invoked, as the building is probably viewed from a short distance; also, from an encyclopedic perspective, it is more logical to leave bikes near an entrance rather than at the back of a building. However, in the phrase *There was a small white building in front of the huge black one*, a relative, egocentric (speaker-observer) perspective may be invoked. In this case, the exact location of the entrance is irrelevant because the focus is on the cityscape as a whole, and additionally, the distance between the speaker-observer and the two buildings may be too large for the entrance to be visible. So, a building can be treated either as a place where people live or which they enter or holistically, as a non-fronted spatial object. This variation in interpretation suggests that frontness is not an absolute property but a more flexible, context-dependent feature.

In addition, not every object has a simple intrinsic orientation: some objects can have different intrinsic orientations when moving or being static. For example, the sentences *There was a person in front of the train* and *I saw several familiar faces in front of our train* evoke distinct spatial images. In the phrase *There was a person in front of the train*, the person is positioned dangerously on the tracks, directly in the path of the moving train. In this intrinsic orientation, the front of the train is its moving part, or the train operator's car. In contrast, the interpretation of the sentence *I saw several familiar faces in front of our train* is based on a different intrinsic orientation: people are positioned on the platform, beside the train, near the passenger doors – a space conventionally used for boarding rather than for motion-related orientation. This contrast illustrates how the same spatial phrase can shift in meaning depending on whether the object in question is static or moving, and how even intrinsic orientation can be influenced by the egocentric perspective: in the case of the train, whether the speaker adopts the car operator's viewpoint, or the boarding passengers' viewpoint.

The fact that, even within a single language, interpretations of expressions involving fronted objects and spatial prepositions may vary depending on speaker perspective, as well as on whether the object is static or moving, prominent or backgrounded, suggests that frontness is not a fixed spatial category but a flexible, context-sensitive feature. This flexibility raises the possibility of cross-linguistic variation. Our study compares how Russian, Armenian, and Kazakh encode frontness, focusing on the interplay between intrinsic and egocentric frames of reference, and examines whether cross-linguistic patterns reflect shared embodied experience or language-specific spatial conventions. In particular, we build on the observation made by Apresjan & Raxilina (2020) that certain objects, such as doors and windows, are double-fronted, due to their bidirectional use (people pass through doors and look through windows from both sides), and are treat-

ed differently from single-fronted objects in Russian. This raises the question of whether similar distinctions exist in other languages, a question we address in our paper through cross-linguistic comparison.

1. *Dveri* ‘doors’ and *okna* ‘windows’: intrinsic vs. egocentric orientation of double-fronted objects in Russian

In this paper, we examine two types of orientationally ambivalent objects that serve as transitional points between spaces – doors and windows. Although both are used from both sides – doors for entry and exit, windows for looking in or out – our preliminary research on Russian reveals that their linguistic treatment is not straightforward (Apresjan & Raxilina, 2020). We analyzed doors and windows in Russian, focusing on corpus examples containing the prepositions *pered* ‘in front of’ and *za* ‘behind’ with *dver’* ‘door’ and *okno* ‘window’). Our findings revealed a clear asymmetry in their interpretation.

In most corpus examples, *pered dver’ju* ‘in front of the door’ refers to the exterior space outside the building, as in the following line from Nikolai Gumilev’s poem: *Ja stoju pered dver’ju tvoeju, ne dano mne inogo puti* ‘I stand before your door, no other path is given to me’. Thus, the space *pered dver’ju* is typically associated with the outside, reinforcing the conventionally primary function of the door: to allow entry rather than exit. Although doors are used for both entering and exiting, language reflects a goal bias, favoring entry as the default function. This bias suggests that enclosed spaces are conceptualized more as places that people seek to enter, rather than as places from which they seek to escape. As a result, the front of a door is mostly perceived as facing outward, suggesting an asymmetrical orientation.

However, the expression *za dver’ju* ‘behind the door’ does not suggest the same asymmetry. If *pered dver’ju* typically refers to the outside, one might expect *za dver’ju* to consistently mean inside. While this interpretation is possible, *za dver’ju* can also refer to an external location, making it effectively synonymous with its opposite, *pered dver’ju*. These two readings of *za dver’ju* involve two different perspectives of the speaker: *Ona skrylas’ za dver’ju* ‘She disappeared behind the door’ means that the speaker is outside, observing someone going inside, whereas *Za dver’ju stoyal kakoj-to neznakomec* ‘A stranger was standing behind the door’ means that the speaker is inside, referring to someone outside, on the other side of the door.

Thus, while *pered dver'ju* is interpreted intrinsically (based on the object's inherent orientation), *za dver'ju* is interpreted egocentrically, from the deictic perspective of the speaker – whether inside or outside – referring to the opposite side of the door from their position. So, while in *pered dver'ju* the door behaves as a typical fronted object with an outside front, in the expression *za dver'ju* it behaves as a non-fronted object, relying on the position of the speaker-observer to establish reference. This fluid interpretation makes *door* an exception to traditional frontness rules.

On the other hand, *okno* 'window' demonstrates the opposite patterns of interpretation, reflecting its distinct function – not as a point of entry but as a means of viewing the outside. As a result, *pered oknom* 'in front of the window' can refer to either an interior or exterior location, depending on the position of the speaker-observer. It is thus interpreted egocentrically: *Ona sidela pered oknom* 'She was sitting in front of the window' refers to the location inside, with the speaker inside the room, further away from the window. In the sentence *Pered oknom roslo derevo* 'A tree was growing in front of the window', the location is outside, with the speaker either outside and the tree positioned between them and the window, or inside, right beside the window, where the tree blocks their view. By contrast, *za oknom* 'behind the window' is uniformly interpreted in reference to the location outside, indicating an intrinsic frame of reference: *Za oknom rosli derev'ja* 'Behind the window, trees were growing'. However, *za oknom* is unlikely to be used if the speaker is not inside; instead, alternative constructions would typically be used to convey the same idea – e.g., *pod oknom* 'under the window'. If the speaker is looking into the window from the street and sees someone inside, this would be expressed using *v okne* 'in the window'.

This contrast between egocentric interpretation for *pered oknom* and *za dver'ju* and intrinsic interpretation for *za oknom* and *pered dver'ju* reflects the functional asymmetry between doors and windows in linguistic encoding of the Russian language. This pattern suggests a potential point of cross-linguistic variation as different languages may assign frontness and spatial relations to double-fronted objects like doors and windows in distinct ways. In this paper, we consider the data from two genetically distinct languages – Kazakh and Armenian – to explore this variation.

Our data for Kazakh and Armenian comes from a combination of corpus materials and native speaker consultations. For Kazakh, we used the Turkic Web Corpus available on Sketch Engine and supplemented it with judgments from nearly forty informants, thanks to a linguistics class taught to Kazakh-speaking students by one of the authors at Nazarbayev University (Astana) during the period in which this paper was being written. For Armenian, we relied on the Eastern

Armenian National Corpus (EANC) and elicited judgments from a single, highly qualified native speaker. While the resources available for Kazakh and Armenian were not equal in scope, we consider the resulting datasets to be as comparable as possible under the circumstances.

To collect data from informants, we used a typological questionnaire containing Russian phrases that illustrate the various usages and interpretations discussed above, as well as some others (the full questionnaire is in the Appendix). We then conducted interviews with native speakers of Kazakh and Armenian who are bilingual in Russian to analyze their responses.

The questionnaire was based on several parameters, such as interior vs. exterior position of the object, interior vs. exterior position of the speaker-observer, intrinsic vs. egocentric frame of reference, static vs. dynamic predicate, where each parameter represented a potential point of cross-linguistic variation. The questionnaire is presented in the Appendix together with the most frequent Kazakh and Armenian phrases used to translate each context.

In presenting our data, we provide morpheme-by-morpheme glosses only for the relevant portions of each sentence, particularly in longer corpus examples. This allows us to focus on the constructions under study without overloading the reader with unrelated material.

2. *Dveri* ‘doors’ and *okna* ‘windows’: double-fronted objects in Kazakh¹

In Kazakh, the spatial relations under discussion are encoded in two ways: through the case marking of the noun denoting the Landmark (e.g., ‘door’ and ‘window’) and through postpositions, which derive from various orientational nouns meaning ‘front,’ ‘side,’ ‘near,’ ‘back,’ ‘outside,’ ‘other side,’ ‘bottom,’ and ‘base.’ Static and dynamic predicates are distinguished by the locative vs. ablative case of the orientational noun, or the nouns ‘window’ or ‘door.’ The choice of orientational nouns and their forms is not uniform across speakers. Below, we present the most representative contexts from the corpus and the questionnaire along with their glossed translations into Kazakh, including frequency distribu-

¹ We thank the students of *Language and Communication* class, Spring 2025 at Nazarbayev University for their participation in the questionnaire, and Ular Yestay, a Linguistics major student at the same university, for glossing the linguistic examples in Kazakh.

tions where applicable. Most contexts have responses from 39 informants, while some include data from 40 informants. Overall, Kazakh exhibits a different pattern from Russian. In Russian, the phrases ‘in front of the door’ and ‘behind the window’ follow an intrinsic frame of reference, typically referring to exterior locations. However, in Kazakh, the intrinsic “exterior” reading occurs in ‘behind the door’ and ‘behind the window’ constructions, while ‘in front of the door’ and ‘in front of the window’ can refer to either interior or exterior locations, depending on the position of the speaker and pragmatic factors. As a result, Kazakh relies on an egocentric frame of reference or pragmatic factors for the interpretation of these spatial relations.

‘In front of’, regardless of the kind of double-fronted object – door or window – or whether the referenced object (Trajector) is interior or exterior, receives an almost uniform translation with the genitive of ‘door’ / ‘window’ followed by the postposition *aldynda* ‘front-3.POSS-LOC’. In (1), the natural reading is with the exterior position of the Trajector (he):

(1)

<i>Ol</i>	<i>esik-tiń</i>	<i>ald-y-nda</i>	<i>tur-yp</i>	<i>qoñyrau</i>	<i>šal-dy</i>
3SG	door-GEN	front-3.POSS-LOC	stand-CVB	bell	ring-PST.3
‘He/she stood in front of the door and rang the bell’					

In all 39 responses, ‘in front of’ was translated by *aldynda* ‘front-3.POSS-LOC’.

In (2), the natural reading is with the interior position of the Trajector (dog); the translation of ‘in front of the door’ by all 39 respondents is equivalent to (1): with the ‘door’ in the genitive, followed by the preposition *aldynda*:

(2)

<i>It</i>	<i>esik-tiń</i>	<i>ald-y-nda</i>	<i>kilem-de</i>	<i>žat-ty</i>
dog	door-GEN	front-3.POSS-LOC	carpet-LOC	lie-PST.3
‘The dog lay on the rug in front of the door’				

We find the same postposition *aldynda* in the translation of contexts referring to exterior and interior locations relative to the windows, again, by all informants. Context (3) refers to the lilac bushes growing *outside*, while context (4) refers to a ficus plant growing *inside*. In both, ‘in front of the window’ is translated as *terezeniń aldynda* ‘window-GEN front-3.POSS-LOC’:

(3)

<i>Tereze-niń</i>	<i>ald-y-nda</i>	<i>börtergül</i>	<i>ös-ti</i>
window-GEN	front-3.POSS-LOC	lilac	grow-PST.3
‘Lilacs grew in front of the window’			

(4)

<i>Tereze-niń</i>	<i>ald-y-nda</i>	<i>fikus</i>	<i>tur-dy</i>
window-GEN	front-3.POSS-LOC	figus	stand-PST.3
‘A large ficus stood in front of the window’			

Thus, ‘in front of’ in Kazakh is consistently encoded by *aldynda*, regardless of the location of the Trajector, which is determined by pragmatics. We infer spatial orientation based on contextual knowledge – bells are typically located outside, dogs usually lie on rugs indoors, lilacs grow outside, and ficuses are indoor plants. This suggests that doors and windows do not function as inherently fronted objects when used with the postposition *aldynda* ‘in front of’.

Yet the natural corpus data from the Turkic Web Kazakh corpus suggests certain tendencies in the interpretation of both ‘in front of the window’ and ‘in front of the door’. Among the 51 examples of *tereze aldynda* ‘window front-3. POSS-LOC’, only one quarter yielded pragmatically natural exterior interpretations, such as (5), while the majority referred to inside locations, such as (6), suggesting that the property of frontness may be scalar, with the privileged facade of the window facing inside:

(5)

<i>Tereze</i>	<i>ald-y-nda</i>	<i>qus</i>	<i>Úshyp</i>	<i>ótse</i>
window	front-3.POSS-LOC	bird	fly-CVB	pass-COND
‘If a bird flies past in front of the window’				

(6)

<i>Tereze</i>	<i>ald-y-nda</i>	<i>jazu</i>	<i>stoly</i>
window	front-3.POSS-LOC	writing	Desk
‘A writing desk in front of the window’			

The translations of the Russian phrase *za dver’ju* ‘behind the door’ further demonstrate the scalar nature of frontness. In the context *Vxodit’ poka nel’zja. Postojte za dver’ju* ‘Entering is not allowed yet. Please wait behind the door’, where the pragmatically natural reading is ‘outside the door’, 18 informants used the construction *esiktiń aldynda* ‘in front of the door’, suggesting that for some speakers,

this expression functions similarly to its Russian counterpart, referring intrinsically to an outside location. So, one may suggest that *esiktiń aldynda* also possesses a certain degree of frontness, with the privileged facade of the door facing outside. Interestingly, five informants described the exterior location as *esiktiń artynda* ('door-GEN back-3.POSS-LOC'), or 'at the back of the door', which likewise suggests an intrinsic frame of reference, but with the privileged facade facing inwards.

The preference for the exterior localization referenced by 'in front of the door' examples was strongly supported by the available corpus data. We analyzed 40 examples from the Turkic Web Kazakh corpus to match the number of our informants: 20 included the genitive construction *esiktiń aldynda* 'door-GEN front-3.POSS-LOC' and 20 featured the less formal nominative form *esik aldynda* 'door-NOM front-3.POSS-LOC'. Of these 40 examples, 30 unambiguously referred to exterior locations, while only one example clearly referred to an interior setting. Most examples spoke of cars, buildings, cows, trees, or other clearly outdoor objects located in front of the door, as in the following examples:

(7)

<i>Siir-lar-ı</i>	<i>esik</i>	<i>ald-y-nda</i>	<i>tur</i>
cow-PL-POSS.3	door	front-3.POSS-LOC	stand.PRS.3
'Their cows are standing in front of the door'			

(8)

<i>Esik-tiń</i>	<i>ald-y-nda</i>	"GAZ-21" Volga	<i>tur</i>
door-GEN	front-3.POSS-LOC	"GAZ-21" Volga	stand.PRS.3
'In front of the door stands a GAZ-21 Volga'			

14 informants used the construction *esiktiń syrtynnda* 'door-GEN outside-3.POSS-LOC' with the postposition *syrtynnda* which means 'outside', thus preferring to mark the exterior location in a non-ambiguous way:

(9)

<i>Kir-u-ge</i>	<i>äli</i>	<i>bol-ma-y-dy</i>	<i>esik-tiń</i>	<i>syrt-y-nda</i>	<i>küt-iñiz</i>
enter-INF-DAT	yet	be-NEG-NMLZ-ACC	door-GEN	outside-3.POSS-LOC	wait-IMP.2PL
'Entering is not allowed yet. Please wait behind the door'					

Only a minority of the speakers adopted an egocentric frame of reference: three informants described the location as *esiktiń arǵy jaǵynda*, or 'on the other side of the door' (door-GEN other.side-3.POSS-LOC) suggesting that the speaker is positioned inside, while the Trajector is outside.

In contexts (10) and (11), we can see that the Russian phrase *za oknom* ‘behind the window’ is consistently interpreted by Kazakh speakers as referring to outside locations, regardless of pragmatics. Even in situations where their world knowledge suggests an inside location for the Trajector, their translations indicate an outside interpretation. Thus, these combinations are interpreted with an intrinsic frame of reference, and windows exhibit characteristics of fronted objects, in the same way as in Russian. In (10), the natural interpretation of the location is exterior, and the majority of the speakers unambiguously mark it as ‘outside’ in their translations of the Russian original *za oknom* ‘behind the window’:

(10)

<i>Ol</i>	<i>tereze-niń</i>	<i>syrt-y-nda</i>	<i>butaq-ta-ǵy</i>	<i>torgay-dy</i>	<i>kór-di</i>
3SG	window-GEN	outside-3.POSS-LOC	branch-LOC-REL	sparrow-ACC	see-PST.3
‘He/she saw a sparrow on a branch outside the window’					

In (11), the natural interpretation of the location is interior, yet the majority of the speakers also used ‘outside’ in their translations of the Russian *za oknom* ‘behind the window’ despite the pragmatic indications to the interior location of the Trajector:

(11)

<i>Ol</i>	<i>tereze-niń</i>	<i>syrt-y-nda</i>	<i>tigip</i>	<i>otyrgan</i>	<i>qyz-dy</i>	<i>kór-di</i>
3SG	window-GEN	outside-3.POSS-LOC	sew-CVB	PROG-PTCP	girl-ACC	see-PST.3
‘Outside the window he/she saw a girl sewing’						

A rare variant of the translation, featuring the literal rendering ‘*behind the window*’ – *tereze-niń art-y-nda* (*window-GEN back-3.POSS-LOC*) – did not occur at all in the natural data, thus corroborating our informant judgments. The most natural corpus examples for exterior spatial contexts used either *terezeniń syrt-ynda* or the more colloquial *tereze syrtynnda* (with nominative *window*), but these, too, were relatively infrequent: we identified only a handful of instances in the Turkic Web Kazakh corpus.

(12)

<i>Tereze</i>	<i>syrt-ynda</i>	<i>qys</i>
window	outside-3.POSS-LOC	winter
‘Outside the window is winter’		

Thus, in the case of ‘behind the window’, the interpretation aligns in both languages: in Russian and Kazakh, this phrase refers to an outside location, suggesting that the privileged front of the window is inside – consistent with its function of providing a view outward. However, the interpretation of ‘behind the door’ differs. In Russian, the location of the Trajector is largely determined by the position of the speaker, who is on the opposite side of the Trajector, thereby invoking an egocentric frame of reference. In contrast, Kazakh mostly interprets ‘behind the door’ intrinsically, meaning ‘outside’ and indicating that, as with the window, the privileged front of the door is understood to be inside the house.

3. *Dveri* ‘doors’ and *okna* ‘windows’: double-fronted objects in Armenian

In modern standard Eastern Armenian, there are two main ways to express the meaning of ‘in front of X’: the postpositions *dimac’* and *araj*, both of which govern the genitive case of the noun (there is also a more formal variant of the latter, *arjev*, with the difference between the two being purely stylistic).

The postposition *dimac’* ‘in front of; opposite’ is etymologically related to the noun *demk’* ‘face’ and to some extent retains this connection in its semantics: one of its privileged contexts of use is to indicate the face-to-face positioning of two objects, i.e., with their front sides facing each other. In this sense, it can be said that the feature of frontness is directly embedded in the Armenian grammatical system. According to our consultant Victoria Khurshudyan, the contexts typical for *dimac’* largely overlap with those in which the French prepositional phrase *en face de* is used (which also contains a word meaning ‘face’).

As for the postpositions *araj/arjev* ‘in front of; ahead; forward’, they are etymologically related to words meaning ‘front’ and ‘first’; *araj* is also used as a spatial adverb ‘ahead’. In its modern semantics, this pair of postpositions often profile not so much frontness but rather the relative distance between an object and an observer. It seems that the topological characteristics of the object are not as important for the use of these postpositions as its position relative to the observer. Thus, in a context like ‘There was a bottle of wine on the table in front of him’, the most natural choice would be *araj*, since a bottle is not a topologically oriented object:

(13)

<i>Arĵev-i</i>	<i>seġan-in</i>	<i>mi</i>	<i>Šiš</i>	<i>gini</i>	<i>ka-r</i>
in.front-3P	table-DAT	IND	bottle	wine	exist-3SG.PST
‘There was a bottle of wine on the table in front of him’					

Accordingly, if someone is referred to as *im araĵ nstac* ‘sitting before me’, it is more likely that this expression describes seats standing one behind the other (like in a bus) – while a person who is *im dimac’ nstac* is supposed to be sitting across from me (and thus facing me).

Unlike *dimac’*, *araĵ* has developed a secondary temporal meaning ‘before (in time); earlier’, which is in line with its “relational” semantics: thus, *inĵnic’ araĵ* (with the ablative ending!) will mean ‘prior to me’ (cf. genitive-based *im araĵ* ‘before me / in front of me’). The following example is also significant: here, *araĵ* is privileged because the observer (coinciding with the speaker) is immobile, while the trajector is moving in the speaker’s field of vision (and thus its topological orientation is irrelevant).

(14)

<i>Anc’-ir</i>	<i>im</i>	<i>araĵ</i>	<i>or</i>	<i>es</i>	<i>u</i>	<i>barekam-ner-s</i>	<i>noric’</i>
pass-IMP.2SG	1SG.GEN	before	that	I.NOM	and	friend-PL-1P	again
<i>tesne-nk’</i>	<i>k’o</i>	<i>geġec’kut’yun-ə</i>					
see-SUB.1PL.	2SG.GEN	beauty-DEF					

‘Pass **before me** so that my friends and I may see your beauty again’ [Step’an Zoryan]

Due to its broader meaning, *araĵ* can often appear in the same contexts as *dimac’*, while the reverse is less likely to be true.

On the other hand, in contexts like ‘standing/sitting in front of a window’, *dimac’* can be expected, especially when referring to the interior side of the window facing the observer.

Taking the context of the **doors**, it can be said that standing / lying in front of the door privileges *dimac’*, especially when the trajector is close to the door and facing it. A typical corpus example is as follows:

(15)

Azat taracut’yunnerə zbəġec’num ein p’ok’rik banĵaranoc’nerə, haverə k’uĵuĵ ēin anum dĵneri dimac’, mi šun hač’um ēr, oroš tneri tanik’neri vra cux ēr barĵranum.
 ‘The free spaces were occupied by small vegetable gardens, **chickens clucked in front of the doors**, a dog barked, and smoke rose from the roofs of some houses.’

<i>haver-ə</i>	<i>k'ujjuj</i>	<i>ēin</i>	<i>an-um</i>	<i>dř-ner-i</i>	<i>dimac'</i>
hen-PL.DEF	clucking	AUX.3PL.PST	make-CNV.IMPF	door-PL-GEN	in.front

Interestingly, corpus examples with *ařaj* mostly provide a different configuration, when the door is not a point one is facing, but rather a point at some trajectory (passed or to be passed), cf.:

(16)
Gyuği tnerum mekik-mekik kvarven čragnerə, zarmac'ac, k'nat'at'ax mardik knay-en lusamutneric', durs kgan dřneri ařaj imanalı, t'e inč' ē patahel.
'One by one, the lights will be turned on in the village houses, surprised, sleepy people will look out the windows, **come out to the doors** (lit. 'come out before the doors') to find out what has happened.'

<i>durs</i>	<i>k-ga-n</i>	<i>dř-ner-i</i>	<i>ařaj</i>
outside	FUT-go-3PL	door-PL-GEN	Before

(17)
Trdata dřan ařaj nran dimavorec' hazarapet Gnunin mi k'ani palatakannerov ev ařajnordec' t'agugu harkabažinə.
'Before the Trdat Gate (= 'before he entered the gate'), he was met by the commander Gnuni with several courtiers who led the queen's tax collection.'

<i>dř-an</i>	<i>ařaj</i>
door-GEN	before

In our questionnaire (many thanks to Victoria Khurshudyan for consultations), the sentences 1-4 may have both *ařaj* and *dimac'*; *dimac'* seems particularly appropriate in the context (3), because the package is supposed to be left just opposite the door and close to it:

(18)

<i>P'ostatar-ə</i>	<i>canroc'-ə</i>	<i>dř-an</i>	<i>dimac'</i>
postman-DEF	package-DEF	door-GEN	in.front

'The mail carrier left the package in front of the door'
On the contrary, in the sentences (5–6) the most natural choice would be the spatial postposition *mot* with the general meaning 'at; near; next to', and only in the sentence (7) the postposition *hetevum* is used, referred to the location 'behind'.

The linguistic behaviour of Armenian **windows** is slightly more homogeneous. As said before, sitting in front of the window (inside the room) requires *dimac'*, while the situations with an outside observer can be characterized with *araj* (if it is about lilacs growing outside in front of the window) or just *mot* (if it is about a person staying outside).

Interestingly, the transitive situations of looking through the window are uniformly coded with the help of another means, namely, the ablative case (requiring no postpositions). Thus, seeing **through the window** is rendered by the form *patuhanic'* window. ABL.

4. Conclusion

The analysis of the data from the examined languages demonstrates that the frontness of double-fronted objects such as doors and windows is not a universal feature. Instead, their behavior varies across languages and appears to form different typological systems in this regard. Different systems may treat both doors and windows as fronted or only one of these objects. In Russian, both doors and windows are fronted though this is manifested only in half of the constructions – ‘in front of the door’ and ‘behind the window’. In Armenian, only the window is fronted. In Kazakh, the situation is more similar to Armenian: only the window is a fully fronted object, when occurring in the construction corresponding to the Russian *za oknom*. However, the door is also somewhat fronted, yet inconsistently: quite a few speakers in some contexts treat ‘in front of the door’ as referring intrinsically to the exterior location, while a small minority of the speakers treat ‘at the back of the door’ in the same fashion.

Thus, we see cross-linguistic variation in the treatment of double-fronted objects along different axes of the spatial system: (1) which objects are conceptualized as fronted – doors, windows or both; (2) which facade is perceived as privileged – the one facing inwards or outwards; (3) in which constructions frontness is manifested – the constructions with ‘in front of’ or ‘behind’. The system of double-fronted objects can also be complicated by additional parameters. For example, in Armenian, the orientation of the Trajector relative to the double-fronted object plays a role: facing it (*dimac'*) or with one’s back to it (*araj*). Overall, our study demonstrates the high complexity of the frontness systems in double-fronted objects and serves as a preliminary, pilot investigation paving the way for a broader typological study, which may reveal additional parameters of this phenomenon.

5. References

- Apresjan, Ju. (1974). *Leksičeskaya semantika: sinonimicheskie sredstva jazyka*. Nauka, Moskva. [2nd ed.: (1995). *Izbrannye trudy* [Selected works], t. 1. Jazyki russkoj kul'tury, Moskva].
- Apresjan, Ju. (1986). *Dejksis v leksike i grammatike i naivnaya model' mira. Semiotika i informatika*, 28, 5–33. [2nd ed.: (1995). *Izbrannye trudy* [Selected works], t. 2. Jazyki russkoj kul'tury, Moskva].
- Apresjan, V., & Raxilina, E. [= Rakhilina, E.] (2020). *Dve storony fasadnosti: dveri i okna. Trudy IRJA RAN*, 2(24), 136–148. <https://trudy.ruslang.ru/ru/archive/2020-2/136-148>.
- Fillmore, C. (1968). Types of lexical information. *Ohio State University Working Papers in Linguistics*, 2, 56–103.
- Levinson, S. C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge University Press, Cambridge. <https://www.cambridge.org/core/books/space-in-language-and-cognition/D07AD2885A025E00B1C94ED722071D80>.
- Majid A., Bowerman M., Kita S., Haun D. B. M., & Levinson S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, 8(3), 108–114. <https://www.sciencedirect.com/science/article/abs/pii/S1364661304000312>.
- Svorou, S. (1994). *The Grammar of Space* (Typological Studies in Language 25). John Benjamins, Amsterdam.

6. Appendix (questionnaire)

Door, 'in front of'

Outside location, static context

(1) *On stojal pered dver'ju i zvonil – nikto ne otkryval.*

'He stood in front of the door and rang the bell – no one opened it'

Kazakh: есіктің алдында 'in front of'

Armenian: *dʳan aʳaĵ* 'before'

Inside location, static context

(2) *Sobaka ležala na kovrike pered dver'ju.*

'The dog lay on the rug in front of the door'

Kazakh: есіктің алдында 'in front of'

Armenian: *dʳan dimac'* 'in front of'; *dʳan aʳaĵ* 'before' is also possible

Outside location, dynamic context

(3) *Počtal'on ostavil posylku pered dver'ju.*

'The mail carrier left the package in front of the door'

Kazakh: *есіктің алдында* 'in front of'

Armenian: *dʳan dimac'* 'in front of'

Inside location, static context

(4) *Ja stojal pered dver'ju i ne rešalsja ujtj.*

'I stood in front of the door and hesitated to leave'

Kazakh: *есіктің алдында* 'in front of'

Armenian: *dʳan aʳaĵ* 'before'

Door, 'behind'**Outside location, static context**

(5) *Vxodit' poka nel'zja. Postojte za dver'ju.*

'Entering is not allowed yet. Please wait behind the door'

Kazakh: *есіктің алдында* 'in front of'

Armenian: *dʳan mot* 'at, nearby'

Location in the space between the open door and the wall, static context

(6) *Gde u vas venik? Na kuxne za dver'ju.*

'Where is your broom? In the kitchen behind the door'

Kazakh: *есіктің артында* 'behind the door'

Armenian: *dʳan hetevum* 'in the space behind'

Window, 'in front of'**Inside location, dynamic context**

(7) *Ja sela pered oknom s knižkoj.*

'I sat in front of the window with a book'

Kazakh: *терезенің алдында* 'in front of the window'

Armenian: *patuhani dimac'* 'in front of'

Outside location, static context

(8) *Pered/pod oknom rosła siren'.*

'Lilacs grew in front of/under the window'

Kazakh: *терезенің алдында* 'in front of the window'

Armenian: *patuhani aʳaĵ* 'before'

Inside location, static context

(9) *Pered oknom stojal bol'soj fikus.*

'A large ficus stood in front of the window.'

Kazakh: *терезенің алдында* 'in front of the window'

Armenian: *patuhani dimac* 'in front of'

Window, 'behind'

Outside location, static context

(10) *Za oknom šel dožd.*

'It was raining behind the window'

Kazakh: *Терезенің сыртында* 'behind the window'

Armenian: *drsum* 'outside' (the noun is not specified)

Outside location, static context

(11) *On uvidel za oknom / v okne vorob'ja na vetke.*

'He saw a sparrow on a branch outside/through the window'

Kazakh: *Терезенің сыртында* 'behind the window'

Armenian: *patuhanic* 'window.ABL'

Inside location, static context

(12) *Za oknom / v okne on uvidel devočku, kotoraja šila.*

'Outside/through the window, he saw a girl sewing'

Kazakh: *Терезенің сыртында* 'behind the window', *terezeden* 'from the window'

Armenian: *patuhanic* 'window.ABL'



Irina Levontina

 <https://orcid.org/0000-0001-7368-2135>

Noun as a Verb

Abstract

The paper explores the phenomenon of using nouns in place of verbs in performative speech acts, particularly in the context of requests. A key example is the Russian construction *Našedšego ključī pros'ba vernut'* 'Whoever finds the keys, please return them', *lit.* 'a request to the finder to return', which, while non-standard, is widely used in speech. The use of nouns instead of verbs in performative utterances is not unique. For instance, saying *Moi pozdravlenija!* 'My congratulations!' instead of *Pozdravliaju vas* 'I congratulate you' reflects a similar phenomenon. This tendency often arises because speakers aim to avoid the implied intimacy implied by 1st/2nd-person pronouns and verb forms, especially in formal or less intimate communication. The peculiarity of the Russian word *pros'ba* 'request' in contrast to similar words (e.g., *mol'ba* 'plea') is that in performative usage it develops unique government patterns. Specifically, it governs the accusative case to indicate the addressee as a semantic argument of the word *pros'ba*, borrowing this pattern from the transitive verb *prosit'* 'to ask'. This reveals an unexpected connection between performativity and the syntactic behavior of the word, highlighting how performative contexts can influence grammatical structures.

Keywords

performative verbs, government pattern, Russian, directive speech acts, semantic actants, addressee, valency

To my teacher Jurij Derenikovič Apresjan with gratitude and love

0.

Jurij Derenikovič Apresjan greatly enjoyed analyzing the structural peculiarities of individual words. For instance, he enthusiastically discussed the word *sistema* 'system': it is incapable of governing an infinitive —one cannot say *sistema delat' zarjadku* 'system to do exercises'. However, within the phrase *vojti v sistemu*

‘to become a system’, such a possibility does emerge *U nego vošlo v sistemu delat' zarjadku* ‘It became a system for him to do exercises’. Compare, for example, the word *privyčka* ‘habit’ which can take an infinitive complement itself: *vošlo v privyčku delat' zarjadku* ‘It became a habit to do exercises’ and *U nego pojavilas' privyčka delat' zarjadku* ‘He developed a habit of doing exercises’. One might also recall Apresjan’s famous co-authored paper with Leonid Iomdin on the construction *negde spat'* ‘no place to sleep’ (Apresjan & Iomdin, 1989, pp. 3–89). I believe Apresjan would also have been interested in analyzing the unusual grammatical properties of the word *pros'ba* ‘request’.

This work is based on the principles of systematic semantic description developed by Ju. Apresjan (2000), one of the most important of which is the idea of an integral description of language. Integral description presupposes that the dictionary and the grammar are mutually adjusted. The objective of this paper is to analyze the Russian word *pros'ba* ‘request’, demonstrating special syntactic and semantic behavior, particularly in performative contexts. Linguistic material is partially sourced from the corpora (mainly <https://ruscorpora.ru/>) and the Internet, partly from the author’s collection. So statistical conclusions cannot be drawn.

1. Examples

We often come across utterances like:

- *Poezd dal'se ne pojdět, **pros'ba** osvobodit' vagony.*
‘The train will not go further, please vacate the cars, lit. a request to vacate the cars.’
- ***Pros'ba** ne bespokoit'.*
‘Please do not disturb, lit. a request not to disturb.’
- *Poterjany ključ. Našedšego **pros'ba** vernut' na vaxtu.*
‘Keys lost. Whoever finds them, please return to the guard post, lit. a request to the finder to return’, etc.

They are widely used in speech, although some of them, as the last one, are non-standard.

Here are some examples:

- (1) *Pokornejšaja pros'ba ob"jasnit' mne, esli možno, kakim obrazom, za čto i počemu ja podvergnut byl segodnjašnemu obysku?*
 'Most humble request to explain to me, if possible, how, why, and for what reason I was subjected to today's search?' [F. Dostoevskij (1871–1872)]
- (2) *Dnja tri nazad ja poslal v redakciju rukopis' Peškova [...], zaglavie "Staruxa Izergil". [...] Pros'ba: posmotret' i rešit' eë sud'bu, [...], ne v dolgom vremeni.*
 'Three days ago, I sent to the editor the manuscript of Peshkov [...], entitled *The Old Woman Izergil*. [...] Please (lit. a request: to) review and decide its fate, [...], without delay.' [V. Korolenko (1894)]
- (3) *Kakaja nelepost'! Bronzovyj Lermontov udivlënno smotrit nam vsled, uperev podborodok v vysokij vorotnik officerskogo sjurtuka. «Sledujuščaja ostanovka "Kurskij vokzal", konečnaja. Pros'ba ne zabylvat' večši v salone tramvaja». Priexali.*
 'What absurdity! The bronze Lermontov looks after us in surprise, his chin resting on the high collar of his officer's coat. "Next stop: Kursky Station, final stop. Please (lit. a request) do not forget your belongings in the tram." We arrived.' [S. Burlačenko (2019)]
- (4) *Syn vyskočil počti srazu že, budto teleportirovalsja na pervyj etaž, i proskol'znul v priotkrytuju dver' pod"ezda (kogo-to opjat' ždali, podperev dver' kirpičom, no Petrov ne posmotrel kogo — uvidel tol'ko listok v kletočku, karakuli, nadpis' "Ubeditel'naja pros'ba ne zakryvat' dver'"), uvidel poloski skotča na vsex četyrëx uglax listočka).*
 'The son dashed out almost immediately, as if teleported to the first floor, and slipped through the slightly open door of the entrance (someone was being waited for again, propping the door with a brick, but Petrov didn't look to see who—he only saw a checkered piece of paper, scribbles, the inscription "Urgent request: do not close the door," and saw strips of tape on all four corners of the paper).' [A. Sal'nikov (2016)]
- (5) *Kak-to pri vstreče v Moskve on mne poxvastalsja, čto kogda vypadal redkij slučaj vystupit' so stixami, kto-to iz prijatelej šel po rjadam i predupreždal: "Pros'ba ne zapisyvati".*
 'Once, when we met in Moscow, he boasted to me that when the rare opportunity would arise to perform poetry, one of his friends would walk through

the rows and warn: “Please (lit. a request) do not record.” [R. Sokoloskij (2015)]

- (6) *Ubeditel'naja pros'ba ne kormit' životnyx. Direkcija zooparka.*
 ‘Strong request: do not feed the animals. Zoo administration.’ [Signs on Doors and Windows (2003)]
- (7) *Esli že predložennoe vremja ili mesto doktora Smita ne ustraivaet, to pros'ba dogovorit'sja o vstreče s sekretarëm Zamestitelja Rukovoditelja Centra.*
 ‘If the proposed time or place does not suit Dr. Smith, please (lit. a request) arrange a meeting with the secretary of the Deputy Head of the Center’
 [I. Pis'menny (2005)]
- (8) — *Priexali, — soobščil Degtjarëv, — poezd dal'se ne pojdët, pros'ba osvobodit' vagony.*
 “‘We’ve arrived,” Degtyarev announced, “the train will not go further, please (lit. a request) vacate the cars.” [D. Doncova (2004)]
- (9) *Bol'saja pros'ba dopolnit' moi argumenty, i gde vozmožno, usilit'.*
 ‘A big request to supplement my arguments and, where possible, to strengthen them.’ [12 Hours a Day?, forum (2010–2011)]

2. Description of the material

At first glance, *Našedšego pros'ba vernut' lit.* ‘a request to the finder to return’, from the point of view of syntax resembles sentences like:

- *U menja problema <vopros>* ‘I have a problem/question’
- *Nam kryška <xana, kapec>* ‘We’re done for’

However, in the latter cases, the noun functions as a predicative with a zero copula and easily allows a non-zero copula:

- *U menja budet problema <byl vopros>* ‘I will have a problem/had a question’
- *Nam byla by kryška <xana>* ‘We would be done for’

However, in some sentences with the lexeme *pros'ba* in question, the noun acts as the predicate itself, and no copula is present. Incorrect: **Našedšego ključa byla pros'ba vernut'* lit. 'A request to the finder was to return'.

In syntactic terms, our lexeme *pros'ba* is closer to units like interjections: *stop*, *xlop*, *basta*, *xoroš*, *lady*, *marš* 'stop, clap, enough, good, alright, march', etc. With these, no copula is possible; they themselves form the predicate. Compare also the term clausative – "Part of speech of a lexemic expression which is syntactically equivalent to a clause; e.g. , Down [with terrorists!] | Wow! | Not at all." (Melčuk, 2023, p. 247).

However, it should be noted that the range of non-verbal predicate expressions turns to be broader than one might think. For example, the expressions *pas* 'pass' and *vist* 'whist' used in Russian Preference, a game of cards, are commissive speech acts: after the contract is ordered, the other players must declare whether they commit to taking the required number of tricks in the game or not. If a player commits to taking a certain number of tricks, they declare "whist"; if not, "pass."

- (10) *Damy, koroli, tuzy bezprestanno metalis' pered glazami; v glazax to i delo razdavalis' slova: kuplju, pas, vist, remiz, vam remiz!...*
 'Ladies, kings, aces constantly flashed before my eyes; in my ears, the words "I'll buy, pass, whist, remise, you remise!" kept ringing out...' [A. Družinin (1850)]

- (11) «*Pas*», — skazal on i snova vypil. Igrali, verovatno, v preferans. Ty kogda sobi-raeš'sja exat'? — sprosil Viktor i ob"javil «vosem' pervyx». V otvet poslyšalos': «*Vist*», «*Pas*», — i partnëry «legli».
 "Pass," he said and drank again. They were probably playing preference. "When are you planning to leave?" Victor asked and declared "eight first." In response came: "Whist," "Pass," and the partners "lay down." [V. Vysotskij (1969–1970)]

3. Problem statement

All the above uses of the word *pros'ba* are performative. *Našedšego pros'ba vernut'* 'Whoever finds it, please return' is not a description of a request but the request itself (Austin, 1962; Searle, 1974). The ability to be used performatively is an

individual feature of the word *pros'ba*. Neither the word *prikaz* 'order', nor *trebovanie* 'demand', nor *mol'ba* 'plea' are used in this way.

Of course, the word *pros'ba* also has an ordinary descriptive (i.e. non-performative) lexeme. For example, in the following case, the request itself is expressed by an imperative, and the word *pros'ba* merely introduces the corresponding speech act:

- (12) — *Il'ja, — govorju ja drožaščim golosom, — u menja k tebe **pros'ba**; možeš byt', ona pokažetsja tebe glupoj i smešnoj pri takix obstojatel'stvax... no... no... ne brosjaj Fomku.*

“Ilya,” I said in a trembling voice, “I have a request for you; it might seem silly and ridiculous under these circumstances... but... don’t abandon Fomka.” [E. Nagrodskaja (1910)]

In such descriptive uses, the word *pros'ba* is accompanied by a copula, not necessarily zero:

- (13) *I vsě-taki eë neizmennaja **pros'ba** byla — dat' ej bombu i pozvolit' byt' odnim iz metal'ščikov.*

‘And yet her constant request was to give her a bomb and allow her to be one of the throwers.’ [B. Savinkov (1909)]

- (14) *So svojej storony, u menja tože budet **pros'ba** ob ostorožnosti. A v čěm delo? Pust' o moej missii znaet kak možno men'se ljudej. Vy, vaš Erik, kol' skoro vy emu tak bezrazdel'no doverjaete, i vsě.*

‘On my part, I will also have a request for caution. What’s the matter? Let as few people as possible know about my mission. You, your Eric, since you trust him so completely, but no one else.’ [V. Rybakov (1993)]

- (15) *Tol'ko u menja budet **pros'ba**... — Andrej opjat' pomolčal, prodolžal, naxmurivšis': — V silu opredelënnyx obstojatel'stv ja ostalsja bez dokumentov...*

“Only I will have a request...” Andrey paused again, and then continued, frowning: “Due to certain circumstances, I am left without documents...” [N. Dežnev (2002)]

4. The word *pros'ba*: syntactic coding of the addressee

The situation of a request presupposes three semantic actants: the subject, the addressee, and the content (who is asking, whom they are asking, and what they are asking for). Thus, the word *pros'ba* has three semantic arguments.

In its performative use, *pros'ba* displays several interesting features, first of all related to the realization of the addressee valency.

4.1. DAT

First, it can be expressed by the dative case:

- (16) *I vot razdalsja razlučajuščij golos, deskat', ostorožno, poezd otpravljaetsja, **pros'ba** passažiram^{DAT} zanjat' svoi mesta.* 'And then the parting voice rang out, saying, "Careful, the train is departing, passengers are requested to take their seats."' [V. Kungurceva (2009)]

- (17) ***Pros'ba** provožajuščim^{DAT} vyteret' slězy — oni gorju ne pomogut.* 'Request to those seeing off to wipe their tears—they won't help with the grief.' [M. Baru (2013)]

- (18) *Iskrenne pozdravljaem pobeditel'nic, (blizkim^{DAT} **pros'ba** rascelovat' ot našego imeni)!* 'We sincerely congratulate the winners, (their dear ones are requested to kiss them on our behalf)! [Contest Completed! Our Congratulations! (2003)]

- (19) *Studentam^{DAT}, želajuščim oformit' kartočki gosudarstvennogo pensionnogo straxovanija, **pros'ba** obraščat'sja v otdel po rabote so studentami (komnata 226).* 'Students wishing to apply for state pension insurance cards are requested to contact the student affairs office (room 226).' [Signs on Doors and Windows (2004)]

- (20) *Zainteresovannym licam^{DAT} **pros'ba** ostavljat' svoi koordinaty po e-mail xxxxx@land.ru ili icq 000000000.* 'Interested parties are requested to leave their contact details at e-mail xxxxx@land.ru or ICQ 000000000.' [Electronic Announcement (2004)]

- (21) *Ved' ne slučajno mnogie firmy dajut ob"javlenija: priglašajutsja vypuskniki takix-to vuzov, ostal'nym^{DAT} — pros'ba ne bespokoit'sja.*

'It's no coincidence that many firms post ads: graduates of certain universities are invited, others are requested not to worry.' [V. Bolotov (2007)]

As we can see, the use of the dative in this context is quite typical. The dative is the standard case for the addressee valency. Additionally, this might be influenced by constructions like *Vsem^{DAT} molčat^{INF}*, *vsem^{DAT} vyjti^{INF}* 'Everyone, be quiet; everyone, leave.' Cf.:

- (22) *Razdalos' šipenije gromkoj svjazi: "Pros'ba vsem^{DAT} ostavat'sja na svoix mestax."*

'The loudspeaker hissed: "Everyone is requested to remain in their seats."' [A. Iličevskij (2007)]

4.2. Subordinate clause

Another, also quite common, option is the use of a subordinate clause indicating the circle of possible addressees of the request: it goes without saying that the subordinate clause is not a syntactic actant of *pros'ba*:

- (23) *U kogo imeetsja foto mestnosti, do stroitel'stva magazina «VAVILON», okolo tipografii, pros'ba vystavit' zdes'.*

'Those who have photos of the area before the construction of the "Babylon" store near the printing house, please (lit. a request) post them here.' [Demolition of Trees on Krasnye Zori (03.03.2015)]

- (24) *Kogo eščë net v našej zamečatel'noj tabličke, bol'saja pros'ba prislat' na mejl [ssylka] sledujuščie dannye — fotografija životnogo, klička, pol, data roždenija, svoë imja.*

'Those not yet in our wonderful table, please kindly (a big request to) send the following data by email [link]: photo of the animal, name, gender, date of birth, your name.' [Cats Forever! (2008)]

- (25) *Komu ne interesno — bol'saja pros'ba ne xodit'.*

'Those not interested — please kindly do not come (lit. a big request not to come).' [Today in the Top Blogs: The Story of a Teacher, blog (2008)]

- (26) *U kogo imeetsja KIM dosročnogo, ogromnaja pros'ba otpisat' v ls.*
 'Those who have early KIM (tests), a huge request to write in private messages.'

[Exchange of Manuals, Materials, and Books, forum (2012)]

- (27) — *Nu vot, letit on i vdrug slyšit ob"javlenie: esli v samolëte est' anesteziolog, pros'ba projti v salon biznes-klassa.*
 "So, he's flying and suddenly hears an announcement: if there is an anesthesiologist on the plane, please (lit. a request) proceed to the business class cabin." [V. Valeeva (2002)]

4.3. Address

Another way to indicate the addressee of the request is through an address. In this case the indication of the addressee is also syntactically separated from the predicate, and the syntactic connection between them is broken. The address does not fill the syntactic valency of the word *pros'ba*:

- (28) *Iz telecentra pri ètom razdavalsja golos v megafon: «Tovarišči, pros'ba razojtis'».*
 'From the television center, a voice rang through the megaphone: "Comrades, please (lit. a request) disperse..." [L. Danilkin (2007)]
- (29) *Gy, filologi, pros'ba ne davit' intellektom!* 'Heh, philologists, please (lit. a request) don't crush with intellect!') [Languages. Polyglots, Hey!, forum (2006)]
- (30) *Golos v gromkogovorit ele, metalličeskij, sklëpannyj iz obrezkov krovël'nogo žezeza, proiznes s drebežžaniem: «Passažiry, vyletajuščie v Mozdok, pros'ba projti na posadku».*
 'The voice in the loudspeaker, metallic, riveted from scraps of roofing iron, uttered with a rattle: "Passengers flying to Mozdok, please (lit. a request) proceed to boarding." [A. Proxanov (2001)]
- (31) *Ženščina normal'noj vnešnosti, 56/163/60, kv., tel., rabota, nezavisimaja, dlja sozdanija sem'i poznamitsja s mužčinoj 56–63 let, s avto, bez m/p. Mužčiny bez sredstv — pros'ba ne bespokoit'. 680051, Xabarovsk, a/ja 31/8.*

‘A woman of normal appearance, 56/163/60, apartment, phone, job, independent, will meet a man 56–63 years old, with a car, without financial problems, for creating a family. Men without financial means—please (lit. a request) do not disturb. 680051, Khabarovsk, P.O. Box 31/8’) [Amur Meridian (22.12.2004)]

- (32) *Rukovoditeli — pros'ba razobratsja s ètoj naglost'ju, i eščë POMYT' POLY V KONCE KONCOV!!*

‘Administrators—please (lit. a request) deal with this outrage, and also WASH THE FLOORS FINALLY!’ [Reviews of the Sparta Fitness Club (12.06.2013)]

- (33) *I vot nado že... — Provožajuščie, pros'ba pokinut' vagon! Otpravljaemsja.*

‘And here we go... “Those seeing off, please (lit. a request) leave the car! We’re departing.”’ [M. Blexman (2012)]

Notice that there is always some kind of punctuation mark separating the two parts.

4.4. ACC

The most interesting is the following option: the addressee valency can be marked by the accusative case, which is quite unusual for the Russian language. So to speak, *pros'ba* is here a transitive noun:

- (34) *Vystupaet A. L. Kogan: “Po ustanovivšejsja praktike massovogo napadenija na kompozitora pros'ba vse^{ACC} vkladyvat' svoi soobraženija”.*

‘A. L. Kogan speaks: “According to the established practice of mass attacks on the composer, everyone is requested to share their thoughts.”’ [V. Švec (1945)]

- (35) *Ja vot bez osmotra (kolleg^{ACC} pros'ba ne šutit'!) ne stala by davat' odnoznačnogo zaključenija!!!*

‘I, without an examination (colleagues, please (lit. a request) do not joke!), would not give a definitive conclusion!!!’ [Beauty, Health, Rest: Medicine and Health, forum (2005)]

- (36) *Vsex^{ACC}, komu znakoma èta osoba, pros'ba pozvonit' po telefonu.*

‘Everyone who knows this person, please (lit. a request) call the phone number.’ [V. Vojnovič (2000)]

- (37) *Minutočku vnimanija. Odesskaja kinostudija priglašajet artistov dlja učastija v massovyx scenax. Želajuščix^{ACC} **pros'ba** sobrat'sja u glavnogo vxoda v četyre časa.*

‘Attention, please. The Odessa Film Studio invites actors to participate in crowd scenes. Those interested are requested to gather at the main entrance at four o'clock.’ [Iz žizni otdyvujuščix, film (1980)]

- (38) *Ispolnjajuščij objazannosti načal'nika kriminal'noj milicii po Kol'skomu rajonu Denis Lobov podčerk nul: “Ljudej^{ACC}, kotorye vsë videli, slučajnyx očevidec^{ACC} **pros'ba** soobščit'.”*

‘Acting Chief of Criminal Police for the Kola District Denis Lobov emphasized: “People who saw everything, accidental eyewitnesses, are requested to report.”’ [Vandals destroyed over 40 gravestones at a Murmansk cemetery (26.06.2009)]

- (39) *A vy gotovy zakrutit' takoj roman? (**pros'ba** otvečat' tol'ko ženščin^{ACC}).*

‘Are you ready to start such a romance? (only women are requested to answer).’ [Resort romance with a local (06.07.2005)]

- (40) *Interesy raznostoronnije. Sudimyx^{ACC} — **pros'ba** ne bespokoit'. 680021, Xabarovsk, a/ja 18/2, dlja Ž-34, dlja Ol'gi.*

‘Interests are diverse. Convicted—please (*lit.* a request) do not disturb. 680021, Khabarovsk, P.O. Box 18/2, for Zh-34, for Olga.’ [Amur Meridian (22.12.2004)]

As can be seen from the examples, the syntactic construction under consideration is not entirely new, but it has become widespread in recent times.

4.5. Non-canonical and canonical syntactic coding of the addressee

At the same time, all the listed ways of marking the addressee of a request for the word *pros'ba* are non-canonical. The “legal” way is κ + *dative*:

- (41) *Gramota. Otvery sprav očnoj služby vopros № 246452. Zdravstvujte. Kak pravil'no, “našedšego^{ACC} **pros'ba** vernut'” ili “našedšemu^{DAT} **pros'ba** vernut'”? Ili i tak, i tak verno? OTVET. Pravil'no: k našedšemu^{k+DAT} **pros'ba** vernut'. 29 sentjabrja 2008.*

‘Hello. Which is correct, *našedšego*^{ACC} **pros'ba** vernut' (the finder please return) or *našedšemu*^{DAT} **pros'ba** vernut' (lit. to the finder a request to return)? Or are both correct? ANSWER Correct: *k našedšemu*^{k+DAT} **pros'ba** vernut' (to the finder please return. September 29, 2008.' [Reference Service Grammar answers from the reference service question No. 246452 (2.14.2025)]

Indeed, outside performative uses, the correct form is: *Vaša pros'ba k nemu*^{k+DAT} vernut' *ključy byla vpolne razumnoj* ('Your request to him to return the keys was quite reasonable'). Here, *emu*^{DAT} or *ego*^{ACC} are impossible. However, the “correct” government pattern is only freely used in descriptive utterances:

- (42) *Zatem Šimanskij sprosila, est' li u neë pros'by k komandovaniju*^{k+DAT} *gospitalja*. 'Then Shymansky asked if she had any requests for the hospital administration.' [V. Grossman (1960)]
- (43) *Delo prinimalo ser'eznyj oborot, i ja kak rukovoditel' ekspedicii obratilsja s pros'boj k Stepanovu*^{k+DAT} *sxodit' vniz na poisk gruppy Evgenija Očkina*. 'The matter was taking a serious turn, and I, as the expedition leader, made a request to Stepanov to go down and search for Yevgeny Ochkin's group.' [I. Vol'skij (1994)]
- (44) *V pis'me o Pavlu Florenskomu (24 maja 1924 goda) Aleksej Fëdorovič priglašael ego učastvovat' v ètom sbornike [...]. V konce pis'ma pros'ba k o. Pavlu*^{k+DAT} *posetit' suprugov Losevyx i razdelit' s nimi «blagostnye vospominanija» v godovščinu ix svad'by 23 maja po staromu stilju (5 ijunja po n. st.)*. 'In a letter to Fr. Pavel Florensky (May 24, 1924), Alexey Fyodorovich invites him to participate in this collection [...]. At the end of the letter, there is a request to Fr. Pavel to visit the Losev couple and share with them the “blessed memories” on the anniversary of their wedding, May 23 (old style) / June 5 (new style).' [A. Taxo-Godi (2009)]

This also occurs in utterances that are intermediate between descriptive and performative:

- (45) *U menja k tebe*^{k+DAT} *pokornejšaja pros'ba poxodatajstvovat' za menja pered Ivanom*. 'I have a most humble request for you: to intercede on my behalf with Ivan.' [Al. P. Čexovv (1887)]

- (46) *Ja na samoe maloe vremja, i kak tol'ko ustroju, ne posmotru ni na kakie prepjatstvija, ni na vremja, i črez poltora ili dva mesjaca ja na doroge v Rim. K vam^{k+DAT} moja ubeditel'nejšaja pros'ba uznat', mogu li ja vzjat' sestër moix do èkzamena i takim obrazom vyigrat' vremja ili neobxodimo oni dolžny ožidat' vypuska. Dajte mne otvet vaš i, esli možno, skoree.*

'I am here for a very short time, and as soon as I settle things, I will not let any obstacles or time constraints stop me. In a month and a half or two, I will be on my way to Rome. My most urgent request to you is to find out if I can take my sisters before the exam and thus save time, or if they must necessarily wait until graduation. Please give me your answer, and if possible, as soon as you can.' [N. Gogol (1836–1841)]

In performative utterances, however, such constructions, while they do occur, are atypical:

- (47) *Poka že k èvoljucionistam^{k+DAT} vse napravljenij pros'ba byt' poskromnee i nazyvat' svoë učenje gipotez, a ne dokazannoj teorij.*

'For now, evolutionists of all stripes are requested to be more modest and refer to their ideas as hypotheses, not as proven theories.' [Darwinism: PRO and CONTRA, forum (04.03.2007)]

- (48) *K avtoram^{k+DAT} portala: Pros'ba soedinit' èti vospominanija kak-nibud' gde-nibud' ot del'no, bez lišnix replik so storony drugix pol'zovatelej, požalujsta.*

'To the authors of the portal: Please (lit. a request) combine these memories somewhere separately, without unnecessary comments from other users, if possible.' [Memories of Wartime Voronezh, forum (2007)]

In the last example, the author uses a colon and a capital letter to separate the addressee from the predicate, breaking the syntactic connection between them.

5. Syntactic coding of the content of a request

To summarize, the peculiarity of the Russian word *pros'ba* 'request' lies in its performative usage, where it develops special government patterns. In particular, it

governs the accusative case to indicate the addressee, apparently borrowing this government pattern from the transitive verb *prosit'* 'to ask'.

It should be noted that performative usage also limits the ways in which the content valency of *pros'ba* can be filled. In such cases, the content is typically expressed with an infinitive, whereas in descriptive usage, the noun *pros'ba* (like the verb *prosit'*) can also use the construction *o* + *prepositional case*. This is impossible in performative usage:

(49)

- a. **Prosim** mužčin pomoč' s pogruzkoj.
'We ask the men to help with the loading'
- b. Mužčin **poprosili** pomoč' s pogruzkoj.
'The men were asked to help with the loading'
- c. **Pros'ba** k mužčinam pomoč' s pogruzkoj byla proignorirovana.
'The request to the men to help with the loading was ignored'
- d. Mužčiny! **Pros'ba** pomoč' s pogruzkoj.
'Men! Please help with the loading'
- e. **Prosim** mužčin o pomošči s pogruzkoj.
'We ask the men for help with the loading'
- f. Mužčin **poprosili** o pomošči s pogruzkoj.
'The men were asked for help with the loading'
- g. **Pros'ba** k mužčinam o pomošči s pogruzkoj byla proignorirovana.
'The request to the men for help with the loading was ignored'

However, the following variant sounds strange.

- h. *Mužčiny! **Pros'ba**^{PERF} o pomošči^{o+PREP} s pogruzkoj.
'Men! Request for help with the loading'

6. Performative nouns and politeness

The phenomenon of using a noun denoting a speech act instead of a standard verb in performative utterances is not unique to *pros'ba*. Compare:

- (50) *Moi pozdravlenija!* 'My congratulations!' instead of *Pozdravljaju* 'I congratulate' [speech act 'Congratulation'];
- (51) *Moi soboleznovanija!* 'My condolences!' instead of *Soboleznuju* 'I condole' [speech act 'Condolence'].

This is likely due to the avoidance of first and second person forms, especially in formal or distant communication (see Goffman 1967, Brown & Levinson 1987, Larina 2009). First and second person pronouns and verb forms can create a sense of overly close interaction, which may feel uncomfortable in certain contexts. Therefore, speakers often prefer to replace them with more impersonal constructions.

For example, the verb *prosit'* (to ask) in performative usage often appears in the third person plural (see Slavkova 2014):

- (52) *Passažirov **prosjať** proĵti na posadku.*
'Passengers are asked to proceed to boarding.'

Similarly, instead of *Esli xotite* 'If you want', one often hears *esli est' želanie* 'If there is a desire'. This form is practically standard when addressing clients.

7. Speech act of request

The speech act of a request has a feature that distinguishes it from other directive speech acts, such as demands or orders: in a request, the speaker understands that the addressee is not obligated to fulfill their wish. The fulfillment of the speaker's desire is a matter of the addressee's goodwill; cf. *pokornejšaja pros'ba* 'most humble request'.

As defined by M. Glovinskaja in the *Integrative Dictionary*:

“Person X wants P to happen; X, believing that person Y can do P and not believing that Y must do P, tells Y that they want Y to do P; X says this in such a way that the addressee understands that X does not believe that Y must do P” (Glovinskaja, 2003, p. 882; see also Glovinskaja, 1991, p. 63; Glovinskaja, 1993)

This places the requester in a vulnerable position. This vulnerability explains the strong preference for avoiding first-person forms when there is no close relationship.

For this reason, the speech act of a request is characterized by various non-standard ways of formulating the utterance. For example:

- *Passažirov **prosjat** proiti na posadku*
‘Passengers are asked to proceed to boarding’
- But not **Pered passažirami **izvinjajutsja** za zaderžku rejsa*
‘Passengers are apologized to for the flight delay’
- Or: **Passažirov **blagodarjat** za ponimanie*
‘Passengers are thanked for their understanding’

At the same time, the valencies of the addressee and the content must be expressed as clearly as possible for the request to be fully realized. Therefore, performative usage gives rise to unusual constructions with the noun *pros'ba*, which mimics a verb but lacks personal markers.

8. Conclusion

The Russian word *pros'ba* ‘request’ demonstrates special syntactic and semantic behavior, particularly in performative contexts. Its ability to govern the accusative case indicating the addressee, borrowing this pattern from the transitive verb *prosit'* ‘to ask’, is particularly striking. This reveals an unexpected connection between performativity and the syntactic behavior of the word, highlighting how performative contexts can influence grammatical structures. At the same time, this reflects the broader tendency in Russian as well as in other languages

(cf. *Thanks!*, *Auguri!*, *Mein Mitleid*, etc.,) to use nouns to perform speech acts, creating a sense of distance and politeness¹.

References

- Apresjan Ju. (2000), *Systematic lexicography*. Translated by Kevin Windle. Oxford University Press, Oxford.
- <https://www.cambridge.org/core/journals/journal-of-linguistics/article/abs/juri-apresjan-systematic-lexicography-translated-by-kevin-windle-oxford-oxford-university-press-2000-pp-v304/5F407245011C9C692BFFE47C473392C0>
- Apresjan Ju., Iomdin L. (1989). Konstrukcii tipa *negde spat'* v russkom jazyke: sintaksis i semantika. *Semiotika i informatika*, 29, 3–89.
- Austin, J.L. (1962). *How to Do Things with Words*. Clarendon Press, Oxford.
- Brown, P. & Levison S. C. (1987). *Politeness: Some Universals in Language Usage*, Cambridge University Press, Cambridge.
- Glovinskaja M. (1991). Slovarnaja stat'ja glagola *prosit'*. In Ju. D. Apresjana (Ed). *Materialy k integral'nomu slovarju sovremennogo russkogo literaturnogo jazyka (obrazcy slovarnyx statej)* *Semiotika i informatika*, 32, 63–71.
- Glovinskaja, M. (1993). Semantika glagolov reči s točki zrenija teorii rečevyx aktov. In E. Zemskaja (Ed.) *Russkij jazyk v ego funkcionirovanii. Kommunikativno-pragmatičeskij aspekt* (ss. 158–218). Nauka, Moskva. <https://uchitu.ru/articles/m-ya-glovinskaya-semantika-glagolov-rechi-s-tochki-zreniya-teorii-rechevyh-aktov.html> (accessed February 15, 2025).
- Glovinskaja, M. (2003). *Prosit'*. In Ju. Apresjan, O. Boguslavskaja, I. Levontina, et al., *Novyj ob''janitel'nyj slovar' sinonimov russkogo jazyka*. Jazyki slavjanskoj kul'tury, Moskva, Vena.
- Goffman, E. (1967). *Interaction rituals: essays on face-to-face behavior*, Anchor Books, New York.
- Larina, T. (2009). *Kategorija vežlivosti i stil' kommunikacii. Sopostavlenie anglijskih i russkix lingvokul'turnyx tradicij*, Rukopisnye pamjatniki Drevnej Rusi, Moskva.
- Melčuk, I. (2023). *General Phraseology. Theory and Practice*. *Linguisticae Investigationes Supplementa*, 36, Amsterdam: John Benjamins, Amsterdam.
- Searle, J. R. (1974). *Speech Acts, an essay in the philosophy of language*. Cambridge University Press, Cambridge.

¹ The author cordially thanks Igor Melčuk for valuable comments.

Slavkova, S. (2014). Performative use of the verbs to ask and pray in Russian and Bulgarian: the pragmatic role of type and tense. *Scando-Slavica*, 60(2), 231–252. <https://doi.org/10.1080/00806765.2014.984464>



Dorota Sikora

Université du Littoral-Côte d'Opale
France

Le verbe polonais *tęsknić* : portraits lexicographiques d'une polysémie

The Polish verb *tęsknić*: lexicographic portraits of a polysemy

Abstract

The present paper explores the polysemy of the Polish verb *tęsknić*, drawing upon the systemic and systematic semantic analysis and lexicographic description designed by Ju. Apresjan. Contemporary dictionaries document two senses of *tęsknić*: the primary one denotes a feeling, while the second meaning is predominantly volitional rather than emotional. However, an analysis of available corpora reveals that in everyday usage, *tęsknić* appears to manifest semantic contents far beyond the scope of these 'official' meanings.

According to a systematic approach to lexicography, it is essential to allocate equal attention to both the unification of lexical units within a particular lexicographic type and to individual properties of each unit, particularly with regard to their semantically motivated formal characteristics. As to the common properties of the lexical type, the present study draws upon a substantial corpus of crosslinguistic research in the field of feelings and emotions. The main attention is given to individuation, that is to particular semantic components and linguistically relevant properties of *tęsknić*, including syntactic structures of PP complements, perfective derivatives and intensifiers. In this respect, significant differences appear in the behaviour of occurrences identified in the corpus, leading to a hypothesis of a polysemy with four rather than two meanings.

The findings of the study have resulted in the identification of four distinct meanings of *tęsknić*, as opposed to the two previously documented.

Keywords

Systematic lexicography, polysemy, verbes de sentiments

1. Introduction

Dans cette contribution¹, nous nous penchons sur le verbe polonais *tęsknić*. On pourrait objecter que ce vocable, si central dans le paysage lexiculturel² polonais, a déjà inspiré une vaste littérature et que l'on sait désormais que, lexicale-ment parlant, il n'a son pareil ni en français (Sikora, 2008 ; Krzyżanowska, 2014 pour son format nominal), ni en anglais (Wierzbicka, 1992). Mais c'est justement pour cette raison que sa description lexicographique est si importante : l'équivalence sémantique entre les langues étant approximative seulement, comment accèderait-on à son – ou plutôt à ses sens – autrement que par une solide explicitation des composantes sémantiques et de leurs différentes configurations ?

C'est en nous appuyant sur les principes d'une lexicologie telle qu'elle a été conçue et pratiquée par Ju. Apresjan (1992, [2000] 2008), que nous mènerons cette étude, en nous efforçant de montrer qu'outre les deux acceptions de *tęsknić* répertoriées dans les dictionnaires de polonais, ce vocable en a deux autres. Autrement dit, nous essaierons de justifier l'hypothèse selon laquelle la polysémie de ce verbe compte en réalité non pas deux, mais quatre sens copolysèmes³.

Si la démarche lexicographique développée par Ju. Apresjan est à la fois systématique et systémique⁴, c'est parce qu'elle repose sur un ensemble de méthodes rigoureuses visant à considérer le système lexical tant dans sa dimension sémantique

¹ Nous remercions chaleureusement les évaluateurs anonymes pour leur relecture attentive et bienveillante. Leurs remarques nous ont permis d'améliorer l'exposé de notre travail. Certaines suggestions, pour diverses raisons, n'ont pas pu être suivies dans ce texte, mais elles nous sont très utiles dans la poursuite de cette réflexion.

² Nous nous permettons ce néologisme, en reprenant sous une forme adjectivale le terme de *lexiculture* proposé par Galisson (1991).

³ Les notions de lexicologie et la terminologie dont nous nous servons sont celles de la Lexicologie Explicative et Combinatoire (LEC ; voir Mel'čuk *et al.*, 1995 ; Polguère, 2016), reprises par la Lexicologie des systèmes lexicaux (Polguère, 2014). Une lexie (unité lexicale) en constitue l'unité de base, point de départ de toute réflexion et l'objet de description. C'est un signe linguistique à trois facettes : signifiant, signifié et caractéristiques combinatoires. On distingue deux types de lexies : un lexème est une acception particulière d'un mot (ce qui revient à admettre qu'avec une polysémie à quatre sens différents, nous avons affaire à quatre lexèmes *tęsknić*) ; une locution associe un signifiant polylexical et un signifié généralement non compositionnel. Le terme de *vocable* est un label – un signifiant – associé à une ou, dans l'immense majorité des cas, à plusieurs acceptions (lexèmes). Polguère (2018) qualifie alors ces dernières de copolysèmes, en insistant sur les relations sémantiques entre les acceptions d'un vocable polysémique. Notons que les deux approches, la LEC et les systèmes lexicaux, doivent beaucoup aux travaux de Ju. Apresjan (voir Mel'čuk, 2024).

⁴ L'ouvrage original connu sous forme d'une traduction révisée *Systematic Lexicography* à l'origine porte bien le titre *Integral'noe opisanie jazyka i sistemnaia leksikografija* (Moscou, 1995).

tique que fonctionnelle, notamment dans ses articulations à la grammaire. Le chercheur est ainsi invité à adopter une double perspective : l'une conduit à plonger dans le macrocosme linguistique à la recherche de types lexicographiques. Un type lexicographique est un groupement de lexies partageant une ou plusieurs propriétés auxquelles s'appliquent les mêmes règles prosodiques, morphologiques, syntaxiques, sémantiques, et qui, de ce fait, doivent être décrites de manière homogène. Une telle vision macro permet de rendre l'organisation systémique du lexique en tant qu'ensemble de lexèmes appartenant à des classes bien définies sur la base de caractéristiques communes à tous leurs membres. C'est ainsi que l'on peut assurer la cohérence d'un dictionnaire.

Notre objectif n'étant pas de construire un dictionnaire, ni même de décrire une classe de lexèmes, mais d'étudier la polysémie d'un vocable, le versant systémique reposant sur une approche macrolinguistique s'en trouvera quelque peu inversé. Plutôt que d'étudier plusieurs lexèmes susceptibles de former une classe, nous exploiterons les caractéristiques sémantiques communes aux lexies du champ sémantique des sentiments pour observer comment celles-ci s'articulent dans le sémantisme de *tęsknić*. Pour ce faire, nous nous appuyons sur des recherches menées selon plusieurs perspectives. Les travaux de Wierzbicka (1992, 1995), les études menées dans le cadre du projet *Emolex* (Novakova & Tutin, 2009 ; Diwersy *et al.*, 2014), les analyses sur corpus de Dziwirek et Lewandowska-Tomaszczyk (2010), la sémantique de la voix médiane (Dobrowolska-Pigoń, 2020), ainsi que les discussions, souvent contrastives, permettant de situer *tęsknić* et *tęsknota* dans le champ sémantique des sentiments en polonais (Nowakowska-Kempna, 1995 ; Pajdzińska, 2003 ; Krzyżanowska, 2005 ; Sikora, 2008, 2023) font effectivement apparaître un type lexicographique dont on pourrait difficilement saisir les caractéristiques sans cette diversité de regards et de méthodes.

Parallèlement, en adoptant une perspective individualisante, le lexicographe s'attache à broser ce que l'on pourrait considérer comme portrait intime de chacun des sens particuliers d'un mot. Une telle entreprise est possible uniquement en menant de « meticulous studies of separate word senses in all of their linguistically relevant aspects » (Apresjan, 2000 : xi). Il s'agit de repérer et de décrire ce qui est particulier, spécifique à un sens donné. Cette double focalisation à la fois sur les traits qu'une lexie partage avec les autres lexèmes de son type, et sur ceux qui lui sont propres doit s'articuler sur la dimension grammaticale au sein d'une approche intégrée et unifiée.

En adoptant cette double perspective sémantico-lexicale, notre réflexion est exposée en six étapes. Tout d'abord, en nous appuyant sur des recherches antérieures, nous réunissons des traits sémantiques communs aux lexèmes du champ

des sentiments⁵ pour passer, dans la section 3, aux traits habituellement retenus pour le verbe *tęsknić*. La section 4 passe en revue ses caractéristiques linguistiquement pertinentes dont l'examen nous permettra de proposer, en 5, des portraits lexicographiques de quatre acceptions qui forment la polysémie de *tęsknić* et de conclure dans la section 6.

2. La sémantique des sentiments : traits commun d'un type lexicographique

L'objectif de la présente recherche définit la démarche à suivre : afin de paraphraser les différents sens du verbe *tęsknić*, qui n'ont pas d'équivalents lexicaux en français, il nous faudra recourir à des lexies, notamment classifiantes⁶, dont dispose le français. Or, si la réalité physiologique et psychique de ce que nous ressentons peut être considérée comme commune à tous les humains indépendamment de leurs langue et culture, le découpage conceptuel et le matériau lexical disponible pour la désigner diffèrent d'une langue à l'autre. Expliciter le sens de lexèmes polonais en français implique donc une attention particulière aux composantes définitionnelles, et plus précisément, à leurs dénominations.

2.1. Hyperonymes classifiant le vocabulaire affectif

Pour les deux langues, la terminologie comprend plusieurs expressions génériques tant lexémiques que polylexicales : *stan psychiczny* 'état psychique', *stan wewnętrzný* 'état intérieur', *uczucie* (≈ 'sentiment', cf. Krzyżanowska, 2014), *emocja* 'émotion', *postawa* 'attitude', *dyspozycja* 'disposition' en polonais ; *senti-*ment, *émotion*, *affect*, *état affectif*, *état psychique*, *état intérieur*, *attitude mentale* en français. Si cette pluralité dénomminative témoigne de la complexité des configurations psychiques, cognitives et comportementales des phénomènes affectifs eux-mêmes, certains termes appartiennent non pas à la langue générale mais au

⁵ Le terme *champ sémantique* n'est pas synonyme de *type lexicographique*. Nous admettons qu'au sein d'un champ sémantique, on peut distinguer plus d'un type lexicographique.

⁶ Une lexie classifiante constitue la composante centrale d'une définition lexicographique telle qu'elle est postulée au sein de la LEC (cf. Mel'čuk & Polguère, 2016). Elle exprime le genre prochain du défini. Par exemple, 'animal aquatique' est la composante classifiante de la paraphrase définitoire de *poisson* 1, 'sentiment' – de *peur*, d'*amour* et de bien d'autres.

vocabulaire disciplinaire. En effet, *état psychique*, *affect* ou encore *attitude mentale* semblent peu, voire pas du tout adaptés pour traduire les conceptualisations naïves, non-scientifiques du monde et leurs lexicalisations.

Les classifieurs les plus fréquents dans la littérature de langue française sont incontestablement les lexèmes *sentiment* et *émotion*, même si leurs sémantismes présentent des particularités frappantes. Pour Schmid (2000), *sentiment* est un nom-coquille, sémantiquement très plastique, qui prend sa véritable signification en contexte – raison pour laquelle il a souvent besoin d'un complément (Blumenthal, 2009). Ce complément de forme *de SN* contient fréquemment des lexèmes qui ne relèvent pas du domaine d'affectivité. Augustyn et Grossmann (2014) qualifient *sentiment* de *translateur* ou de *convertisseur psychologique* conférant un caractère psychique et affectif à des lexèmes qui en sont dépourvus. Effectivement, les collocations telles que *sentiment de sécurité* ou – au contraire – *d'insécurité*, *sentiment de révolte*, *sentiment d'incompréhension*, *sentiment d'appartenance*, *sentiment de cohérence* et bien d'autres sont largement attestés en corpus et sur la Toile.

Cette combinatoire pourrait s'expliquer par le triple attachement de *sentiment*, qui renvoie à la fois aux domaines d'affectivité, des sens et de l'intellect⁷ (Novakova *et al.*, 2018). Comparé au sens très général de *sentiment*, *émotion* apparaît comme un nom général bien plus solidement inscrit dans le domaine affectif, se comportant sur le plan combinatoire comme un item sémantiquement plein. C'est sans doute parce que, comme le montre Polguère (2013) en développant les discussions exposées dans Wierzbicka (1992) et Blumenthal (2009), *émotion* est un quasi-synonyme sémantiquement plus riche de *sentiment*. On peut en effet définir *émotion* comme sentiment d'un certain type (alors que l'inverse – **sentiment signifie 'émotion d'un certain type'* – n'est pas vrai) : plutôt intenses et ponctuelles, les émotions surgissent comme réactions physiologiques à des stimuli (Apresjan, 2008).

Les analyses de Krzyżanowska (2014) menées dans une perspective contrastive permettent de conclure cette étape de notre réflexion en retenant comme hyperonymes *uczucie* et *sentiment*, même si l'on ne saurait admettre leur équivalence sémantique. En effet, « *[u]czucie* se distingue par une composante affective et physique (Wierzbicka, 1995), tandis que *sentiment* comporte une composante affective et cognitive », renvoyant en français à des « états psychiques conscients, qui trouvent leurs réalisations linguistiques dans des lexèmes spécifiques » (*Ibidem*, 119) que l'on exprime en polonais par des lexèmes tels que *odczucie*, *wycucie*, *poczucie*⁸.

⁷ Cette triple appartenance pourrait à son tour expliquer la plasticité sémantique du lexème *sentiment* observée par Schmid (2000).

⁸ Ces trois lexèmes polonais n'ayant pas d'équivalents exacts en français, nous n'en proposons pas de traduction. Leurs analyses et paraphrases sont proposées dans Krzyżanowska (2014).

Si les hyperonymes *sentiment*⁹ et *uczucie* servent surtout à classer le vocabulaire nominal, nous utiliserons surtout le premier d'entre eux pour décrire les sens affectifs du verbe *tęsknić*.

2.2. Sémantique des sentiments : en format nominal ou verbal ?

Les études consacrées à l'expression linguistique des sentiments adoptent des clés d'entrée différentes selon les positions théoriques adoptées. Lorsque l'on s'intéresse à la réalisation syntaxique des structures argumentales des prédicats affectifs, souvent au sein de la classe plus vaste des prédicats psychologiques, l'attention se focalise sur le lexique verbal. Les nominalisations sont alors étudiées sous l'angle de l'héritage possible des arguments verbaux et plus loin, de la structure temporelle et aspectuelle des prédicats. L'idée sous-jacente est alors de repérer des contenus sémantiques similaires à partir de structures syntaxiques, en les interprétant en termes de rôles assignés aux arguments. Les régularités syntaxiques conduisent à considérer que les verbes tels que *craindre* et *aimer* appartiennent à la même classe – classe I proposée par Belletti et Rizzi (1988) – car l'argument expérienceur y figure en position de sujet, alors que l'argument thème est réalisé par un complément accusatif. Cependant, pour faire apparaître les composants sémantiques communs, il faut décomposer le sens en recourant à des prédicats très généraux tels que ÉTAT pour les deux verbes cités, CAUSE et ÉTAT, lorsqu'il s'agit de la classe II dont relèvent par exemple *effrayer*, *attrister*, etc. Notons d'ailleurs que *tęsknić* échappe à ce classement : si l'expérienceur est bien porteur du cas nominatif, le SP n'est ni à l'accusatif, ni au datif.

En sémantique lexicale et dans les travaux lexicologiques, c'est au contraire le vocabulaire nominal qui semble être privilégié, les analyses des verbes de sentiments restant relativement rares. Wierzbicka (1997) en propose une de *tęsknić*, comparé successivement à *to miss* et à *to pine*, mais dans son ouvrage consacré à la conceptualisation des émotions dans différentes langues (Wierzbicka, 1992), ainsi que dans de nombreux autres travaux (Wierzbicka, 2011, 2013, 2019), la réflexion est construite à partir des noms. Dans la discussion de la structure des entrées dictionnaires (Apresjan, 2008), le rôle de la métaphore dans la description du vocabulaire des émotions est également expliqué à partir de lexèmes nominaux. Polguère (2013) en fait autant pour l'organisation du champ

⁹ Dans la suite de ce texte, nous employons souvent *sentiment* et *émotion* de manière alternative. Il ne s'agit pas de gommer les particularités sémantiques que nous venons de discuter, mais d'éviter les répétitions. Que le lecteur veuille bien nous excuser cette entorse stylistique à la rigueur terminologique.

lexical des sentiments en français. En quête des émotions premières communes à tous les humains et de leurs lexicalisations dans des langues particulières, Johnson-Laird et Oatley (1989) proposent une typologie des noms. Plus récemment, les neuf champs sémantiques d'émotions définis et étudiés dans la perspective comparée du projet *Emolex* sont également explorés essentiellement à travers les noms qui en font partie (cf. Diwersy *et al.*, 2014, ainsi de très nombreux travaux présentant les résultats des recherches réalisées dans la cadre de ce projet, notamment Novakova & Tutin, 2009). Pour le sentiment polonais qui nous intéresse tout particulièrement dans cet article, le format nominal *tęsknota* retient plus souvent l'attention (Nowakowska-Kempna, 1995 ; Pajdzińska, 2003 ; Krzyżanowska, 2005).

Or, les locuteurs ne partagent pas nécessairement cette préférence pour les noms. Wierzbicka (1995) remarque que pour parler de ce qu'ils ressentent, les polonophones natifs recourent le plus souvent aux verbes, contrairement aux locuteurs d'anglais préférant les adjectifs. Ces résultats sont confirmés par les analyses comparées réalisées sur des corpus anglais et polonais par Dziwirek et Lewandowska-Tomaszczyk (2010). Pour notre part, nous avons interrogé *Narodowy Korpus Języka Polskiego* (NKJP) et la base de données conversationnelles *Spokes* pour vérifier si cette orientation vers le vocabulaire verbal des sentiments se confirme pour *tęsknić* et pour *tęsknota*. Les résultats sont sensiblement différents selon le type de textes. Nous avons soumis plusieurs requêtes d'abord sur NKJP équilibré (désormais NKJPé), puis intégral (NKJPi). Les deux collections de textes présentent une différence de taille trop importante (240,192,461 mots dans NKJPé versus 1,524,696,745 dans la version intégrale) pour prendre en compte les fréquences absolues. Il a fallu calculer la fréquence par million de mots, et c'est sur cette donnée que nous nous appuyons.

Tant dans la version équilibrée qu'intégrale, l'écart entre *tęsknić* et *tęsknota* est faible : dans NKJPé, la fréquence relative est de 16,67 pour le verbe, de 16,73 pour le nom ; dans NKJPi, respectivement de 10,58 pour le premier et de 9,81 pour le second. Nous avons donc lancé une deuxième série de requêtes sur NKJPi, en distinguant entre les textes écrits et oraux. Les données issues du corpus écrit sont similaires aux premiers résultats : 10,74 pour *tęsknić*, 9,81 pour *tęsknota*. En revanche, à l'oral, l'écart devient significatif en faveur du lexème verbal dont la fréquence relative est de 13,6 contre 4,04 seulement pour le nom. Les résultats d'une requête soumise à *Spokes* affichent la même tendance : 14,48 face à 2,66. Il s'avère ainsi que les locuteurs de polonais emploient plus souvent le verbe *tęsknić* que le nom *tęsknota*.

2.3. Un type lexicographique complexe

Les lexèmes du champ sémantique des sentiments renvoient à des états complexes qui, comme l'expliquent Ju. Apresjan et V. Apresjan (2000), engagent à des degrés divers sept différents systèmes. La perception (i) à travers les organes de notre corps s'accompagne de sensations qui relèvent de la physiologie (ii), en déclenchant des réactions motrices et des actions physiques (iii). Notre volonté et nos désirs (iv), liés aux jugements (v) portés sur les choses telles qu'elles sont perçues (i), vont de pair avec des ressentis (vi) et avec des réactions verbales (vii). Les états affectifs consistent ainsi dans toute une série d'activités se traduisant par les verbes *percevoir, sentir, faire, vouloir, penser, éprouver, parler*¹⁰. Les composantes sémantiques relatives à ces diverses facettes sont lexicalisées selon des configurations différentes, avec une saillance variable, tant pour les noms que pour les verbes.

Nous nous tournons à présent vers le verbe *tęsknić* que les locuteurs polonophones semblent préférer au nom correspondant, lorsqu'ils parlent de leur ressenti face à une séparation.

3. *Tęsknić* dans les dictionnaires et dans le discours

Les dictionnaires polonais sont unanimes, lorsqu'il s'agit de décrire la polysémie de *tęsknić* : elle comprend deux acceptions copolysèmes. Nous reprenons ci-dessous les paraphrases définitionnelles proposées dans *Słownik Języka Polskiego* (*Dictionnaire de la langue polonaise*, désormais SJP) disponible en ligne¹¹.

TĘSKNIĆ 1 : 'odczuwać żal, być smutnym z powodu czyjejś nieobecności, braku kontaktu z kimś lub z czymś'
ressentir du regret, être triste à cause de l'absence de quelqu'un, de l'absence de contact avec quelqu'un ou avec quelque chose

TĘSKNIĆ 2 : 'mocno pragnąć pozyskać, osiągnąć coś'
désirer fortement obtenir quelque chose, souhaiter atteindre quelque chose

¹⁰ Cf. Ju. Apresjan & V. Apresjan (2000 : 208) : *to perceive, to sense, to do, to want, to think about, to feel, to speak*.

¹¹ sjp.pwn.pl/szukaj/tesknic.html, consulté le 28 décembre 2024.

Les traductions aussi proches que possible de ces définitions apparaissent en italique. La première paraphrase, celle qui rend le sens du *tęsknić 1* situe ce lexème dans le champ sémantique des sentiments, alors que la seconde présente *tęsknić 2* comme quasi-synonyme de *vouloir*. Certains emplois du verbe *tęsknić* semblent conformes à cette répartition sémantique : parmi les exemples présentés ci-dessous, (1) et (2) illustrent bien le premier des deux sens, tandis que la seconde acception apparaît dans les énoncés (6) et (8). Les définitions sont suffisamment générales pour subsumer également (3) et (4) sous *tęsknić 1*, alors que (5) à (8) pourraient être considérés comme des exemples de *tęsknić 2*.

- (1) *Bardzo tęsknię za swoją córką.*

bardzo_{ADV} tęsknić_{V.PR.1} za_{PRÉP} swoja_{PRON.INS.SG} córka_{N.INS.SG}¹²

‘Ma fille me manque beaucoup.’

- (2) *Tęskniła za domem, za Suzanne, za małą Zuzią...*

tęsknić_{V.PST.3} za_{PRÉP} dom_{N.INS.SG}, za_{PRÉP} Suzanne_{Np.INS.SG}, za_{PRÉP} mała_{ADJ.INS.SG}

Zuzia_{Np.INS.SG}

‘La maison, Suzanne, la petite Zuzia lui manquaient.’

- (3) *Biedne dziecko, Maciek tęskni bardzo do matki.*

biedny_{ADJ.NOM} dziecko_{N.NOM} Maciek_{Np.NOM} tęsknić_{V.PR.3} bardzo_{Adv} do_{PRÉP}

matka_{N.GÉN}

‘Pauvre enfant, sa mère manque beaucoup à Maciek.’

- (4) *Boże, jak ja strasznie/okropnie tęsknię do tego kraju.*

Bóg_{N.VOC.SG}, jak_{PRON} ja_{PRON.I.NOM} strasznie_{Adv}/okropnie_{Adv} tęsknić_{V.PR.1} do_{PRÉP}

ten_{PRON} kraj_{N.GÉN}

‘Mon Dieu, qu’est-ce que ce pays me manque terriblement/atrocement’.

¹² Nous employons les abréviations des gloses selon Leipzig Glossing Rules compilées par le Laboratoire de linguistique formelle (UMR 7110 CNRS et Université Paris Cité), disponibles en ligne (http://www.llf.cnrs.fr/sites/llf.cnrs.fr/files/statiques/Abreviations_gloses-fra.pdf, consulté le 20 février 2025). Les abréviations utilisées sont développées à la fin du présent article. Pour des raisons de lisibilité de nos exemples (que nous avons préféré ne pas abrégé), nous nous permettons quelques libertés dans l’usage des gloses. Les abréviations apparaissent en indice. Pour éviter d’étendre les exemples tels que (5), (6) et (7) sur plusieurs lignes, avec leurs gloses intercalées, nous n’avons pas aligné les lexèmes des exemples et ceux des gloses pour préserver une lecture linéaire.

- (5) *Ludzie mogą tęsknić za socjalnymi zabezpieczeniami dawnego reżimu.*

ludzie_{N.PL.NOM} móc_{V.PR.6} tęsknić_{INF} za_{PRÉP} socjalny_{ADJ.INS.PL} zabezpieczenie_{N.INS.PL}
dawny_{ADJ.GÉN.SG} reżim_{N.GÉN.SG}

‘Les gens peuvent regretter les garanties sociales de l’ancien régime.’

- (6) *Ciągła obecność hałaśliwego rodzeństwa, nadmiar głosów i gestów powodowały, że Aniela tęskniła za cichą samotnością.*

ciągły_{ADJ.NOM} obecność_{N.NOM.SG} hałaśliwy_{ADJ.GÉN.SG} rodzeństwo_{N.GÉN.SG},
nadmiar_{N.NOM.SG} głosy_{N.GÉN.PL} i_{CONJ} gesty_{N.GÉN.PL} powodowały_{V.PST-6} że_{CONJ}
Aniela_{NP.NOM} tęsknić_{V.PST.3} za_{PRÉP} cicha_{ADJ.INS.SG} samotność_{N.INS.SG}

‘La présence constante de la fratrie bruyante, la surabondance des voix et des gestes faisaient que Aniela se languissait d’une solitude tranquille.’

- (7) *Dzieci tęsknią do komputera.*

dziecko_{N.NOM.PL} tęsknić_{V.PR.6} do_{PRÉP} komputer_{N.GÉN.SG}

‘L’ordinateur manque aux enfants.’

- (8) *Ja tęsknię do spokoju, bo pozostał mi rok do emerytury – powiedział pan Radziszewski.*

ja_{PRON.1.NOM} tęsknić_{V.PR.1} do_{PRÉP} spokój_{N.GÉN.SG} bo_{CONJ} pozostać_{V.PST.3} ja_{PRON.1.DAT}
do_{PRÉP} emerytura_{N.GÉN.SG} – powiedzieć_{V.PST.3} pan_{N.NOM.SG} Radziszewski_{ADJ.NOM.SG}

‘J’ai hâte d’être tranquille, et il me reste un an jusqu’à la retraite – dit Monsieur Radziszewski.’

Cependant, plutôt que d’essayer de rattacher tous les exemples aux deux acceptions officiellement reconnues, il nous semble intéressant de réexaminer la polysémie de *tęsknić*. Ne serait-elle pas, en réalité, plus riche ? Pour le vérifier, conformément aux principes de la lexicologie apresjanienne, nous avons étudié ce que l’on peut, à notre sens, considérer comme des traits linguistiquement pertinents de ce verbe.

4. Caractéristiques linguistiquement pertinentes de *tęsknić*

La métaphore de portrait lexicographique est en effet parlante : un portraitiste observe les traits et les mouvements visibles en surface pour les associer avec ceux de la personnalité de son modèle. De même, le travail de lexicographe repose sur

une observation méticuleuse de caractéristiques propres à chaque lexie, avec une attention particulière portée aux composantes sémantiques susceptibles d'expliquer son comportement grammatical, morphologique, sémantique, syntaxique, etc. Un trait est considéré comme linguistiquement pertinent, si l'on peut établir un lien (par exemple de motivation) entre lui et une règle générale – grammaticale, morphologique, sémantique, etc.¹³ – régulant le fonctionnement linguistique d'une lexie donnée.

Nos analyses se sont focalisées sur trois faits de langue qui peuvent s'expliquer par des différences de sens associés au signifiant *testnić*. Nous admettons que i) la forme du complément prépositionnel exprimant le second actant, ii) le potentiel dérivationnel de chacun des lexèmes *testnić* et iii) les modificateurs d'intensification autorisés constituent leurs traits linguistiquement pertinents sur le plan syntaxique, combinatoire et morphologique.

4.1. Compléments prépositionnels du verbe *testnić*

S'il existe un accord sur la polysémie de *testnić*, un certain flottement apparaît à propos de la forme du complément exprimant le second actant de chacun des deux copolysèmes. Dans les exemples (1) à (8) ci-dessous, l'actant Y vers lequel s'oriente le sentiment de l'expérienceur, s'exprime par un syntagme prépositionnel. Certaines ressources précisent que pour *testnić* 1, seul le complément *za* *SN_{Ins}* réalise syntaxiquement le second actant, *do* *SN_{Gén}* étant réservé pour l'acception 2. Ce sont notamment les préconisations du SJP. Cette même ressource renvoie cependant vers un commentaire linguistique de M. Bańko¹⁴ indiquant que, pour l'acception 1, les deux formes de complément sont acceptées *modulo* une différence de sens. En ce qui concerne le copolysème 2, seul le SP introduit par *do* est possible. *Wielki Słownik Języka Polskiego*¹⁵ (désormais *WSJP*) admet en revanche un double régime syntaxique *za* *SN_{Ins}* et *do* *SN_{Gén}* pour chacune des deux acceptions.

Il est intéressant de remarquer que ces deux compléments sont formellement identiques à ceux qui expriment des relations spatiales. Ainsi, les syntagmes *za* *SN_{Ins}* des exemples (1)–(2) et (5)–(6) ont la même forme que le complément

¹³ Apresjan (2008 : xvi) : "A certain property is considered to be linguistically relevant if there is a rule of grammar or some other sufficiently general rule (semantic rules included) that accesses this property."

¹⁴ <https://sjp.pwn.pl/poradnia/haslo/testnic;1561.html>, consulté le 28 décembre 2024.

¹⁵ *Grand Dictionnaire de la Langue Polonaise* (https://wsjp.pl/haslo/do_druku/10732/testnic, consulté le 8 février 2025).

possessif *za domem rolnika* qui, dans (9), permet de situer les pièces de monnaie en argent par rapport à la maison. Les compléments *do SN_{Gén}* des énoncés (3)–(4) et (7)–(8) sont formellement identiques au complément allatif *do tego kraju* explicitant un terme de déplacement dans (10).

- (9) *Srebrne monety leżały w ziemi, za domem rolnika.*

srebrny_{ADJ.NOM.PL} moneta_{N.NOM.PL} leżeć_{V.PST.6} w_{PRÉP} ziemia_{N.ABL.SG} za_{PRÉP}
dom_{N.INS.SG} rolnik_{N.GÉN.SG}

‘Les pièces de monnaie en argent étaient enterrées derrière la maison du fermier.’

- (10) *Tysiące inżynierów z całego świata przybywa do tego kraju.*

Tysiąc_{NUM.NOM.PL} inżynier_{N.GÉN.PL} z_{PRÉP} cały_{ADJ.GÉN.SG} świat_{N.GÉN.SG}
przybywać_{V.PR.3} do_{PRÉP} ten_{DÉT.GÉN.SG} kraj_{N.GÉN.SG}

‘Des milliers d’ingénieurs arrivaient dans ce pays.’

Associé à un prédicat statif, *za SN_{Ins}* décrit une localisation¹⁶. *Do SN_{Gén}* participe au contraire de relations dynamiques aussi bien lorsqu’il s’agit d’un changement de position dans l’espace que d’un mouvement fictif¹⁷. On remarque par ailleurs qu’il n’est pas possible de factoriser le contenu sémantique de ces compléments par une conjonction. Les phrases telles que (11) et (12) sont plutôt bancales, ce qui conduit à penser qu’une conjonction des compléments *za SN_{Ins}* et *do SN_{Gén}* pour une même occurrence de *tęsknić* ne peut pas représenter deux instances différentes de second actant sémantique.

- (11) ? *Tęsknię za swoją córką i do matki.*

? tęsknić_{V.PR.1} za_{PRÉP} swoją_{PRON.INS.SG} córka_{N.INS.SG} i_{CONJ} do_{PRÉP} matka_{N.GÉN.SG}

- (12) ? *Aniela tęskniła za cichą samotnością i do komputera.*

? Aniela_{Np.nom} tęsknić_{V.PR.3} za_{PRÉP} cicha_{ADJ.INS.SG} samotność_{N.INS.SG} i_{CONJ} do_{PRÉP}
komputer_{N.GÉN.SG}

¹⁶ La préposition *za* peut être employée pour décrire le site d’un prédicat dynamique, mais elle gouverne alors un SN à l’accusatif.

¹⁷ Un *mouvement abstrait* (Langacker, 1987) ou un *mouvement fictif* (Talmy, 2000) n’en est pas un, en réalité. Il s’agit d’un transfert conceptuel où une situation non spatiale, en l’occurrence affective, est présentée comme si elle relevait de ce domaine d’expérience.

4.2. Complément prépositionnel et dynamique

4.2.1. Des sentiments dynamiques ?

Le complément sélectionné par *tęsknić* constitue un trait linguistiquement pertinent car on peut le mettre en rapport avec le trait de dynamique présent dans certaines acceptions.

Les verbes de sentiments endogènes (Anscombre, 1995), c'est-à-dire ceux pour lesquels l'origine « se confond avec le lieu psychologique », se caractérisent par une certaine dynamique. Talmy (2000) l'attribue à la position de sujet accordée à l'actant X prototypiquement associée à l'agentivité. L'expérienceur apparaît ainsi comme instigateur, voire d'initiateur de son état affectif.

La dynamique de certains prédicats affectifs peut s'associer à d'autres composantes sémantiques traduisant l'implication des systèmes discutés dans la section 2.3 ci-dessous¹⁸. Premièrement, les lexies affectives font partie de celles pour lesquelles l'intensité est une caractéristique définitoire : comme le remarque Kleiber (2013), elle représente l'un de leurs paramètres constitutifs, exactement comme la taille ou le volume permettent de décrire des objets matériels. Or, l'intensité d'un sentiment peut varier dans le temps, en provoquant des réactions physiologiques, physiques, verbales.

Deuxièmement, si, prototypiquement, les états affectifs ne sont pas le résultat d'une démarche intentionnelle à proprement parler, ils n'excluent pas une certaine forme de participation active. Même s'il ne les contrôle, l'expérienceur n'en est pas pour autant un récepteur passif – il vit des situations affectives en s'y engageant consciemment. Wierzbicka (1995) considère d'ailleurs que cette participation active, manifeste entre autres en polonais, est caractéristique des cultures slaves.

Troisièmement enfin, outre le trait d'intensité, les prédicats de sentiment possèdent également une dimension volitionnelle. Or, la volonté, souvent impliquée dans un état affectif, signifie que l'expérienceur est prêt à agir pour changer l'état de choses. En aimant quelqu'un, on souhaite le contact avec cette personne : ce désir fait partie intégrante de l'amour. De même, dans l'exemple (3), *Maciek tęskni do matki*, Maciek veut se retrouver de nouveau en présence de sa mère. La composante volitionnelle y est nettement plus marquée que dans les exemples (1) et (2).

4.2.2. Complément prépositionnel, dynamique et orientation d'un sentiment

La forme du complément prépositionnel semble avoir une relation directe avec le caractère plus ou moins dynamique d'un sentiment. Celui-ci peut être saisi

¹⁸ La dynamique semble ainsi inscrite dans le type lexicographique des lexies de sentiments, notamment parce qu'elle est indissociable des prédicats *to do*, *to speak*, *to think about*, *to want*. Cf. Apresjan (2000 : 208) et la section 2.3 *supra*.

comme un procès ou comme un état. Dans les phrases (1)–(2) et (5)–(6), où le second actant s'exprime par *za SN_{Ins}*, la situation décrite par *teżknić* est stativale, puisque la structure sémantique du verbe s'organise autour des composantes relatives au ressenti négatif, fait notamment de tristesse et de regret dans le cas de (1) et (2), et au jugement positif porté sur la présence de Y dont X est séparé au moment de référence. Dans la section 5, nous reviendrons sur les caractéristiques sémantiques de *teżknić* dans (1)–(2) et (5)–(6), mais notons dès à présent leur présupposé commun : on peut *teżknić za SN_{Ins}* uniquement si l'on a été auparavant en présence de Y, auquel on est attaché par un lien affectif ou social. Ainsi, focalisé sur le ressenti provoqué par l'éloignement de Y, l'état est mis en relation avec un intervalle temporel passé.

Les exemples (3)–(4) et (7)–(8) illustrent deux sens différents du verbe *teżknić*, qui ont néanmoins des traits sémantiques communs. Tout d'abord, dans ses emplois avec *do SN_{Gén}*, *teżknić* ne présuppose pas de contact préalable avec Y, que ce soit une personne, un lieu ou une situation. L'exemple (13) montre que le sentiment peut porter sur un second actant que l'on n'a jamais connu :

(13) *Tęsknię do ojca, którego nigdy nie miałam.*

teżknić_{V.PR.I} do_{PRÉP} ojciec_{N.GÉN.SG} który_{PRON.GÉN.SG} nigdy_{ADV} nie_{NÉG} mieć_{V.PST.I}
 'Je me languis d'un père que je n'ai jamais eu.'

En observant les énoncés (3) et (4), on remarque ensuite qu'employé avec *do SN_{Gén}*, *teżknić* présente les traits de dynamicité et d'intentionnalité. On ne pourrait, certes, lui accorder une valeur actionnelle, mais l'état affectif orienté vers Y s'accompagne de la volonté, voire de l'intention de retrouver Y et d'être au contact avec lui. Le ressenti s'associe ainsi à une réaction visant un rapprochement, ce qui autorise une interprétation en termes de mouvement fictif. À l'opposé de (1) et de (2), le sentiment est ainsi orienté vers le futur.

4.3. Potentiel dérivationnel

Par *potentiel dérivationnel* nous entendons la capacité d'un verbe imperfectif de sélectionner un certain type de préfixes pour dériver des perfectifs. Comparé avec celui des verbes de déplacement par exemple, *teżknić* a un potentiel dérivationnel relativement faible : ses perfectifs sont formés à l'aide des préfixes *za-* et *po-*. *Zateżknić*, suivi aussi bien des compléments *za SN_{Intr}* et *do SN_{Gén}*, a une valeur inceptive. Stawnicka (2015 : 91) classe de telles formations parmi les dérivés initiaux et augmentatifs conférant à la phase inchoative qu'ils désignent une

étendue temporelle de courte durée. Employé également avec les deux compléments, *potęsknić* a, toujours en suivant Stawnicka, une valeur déterminative et diminutive, puisqu'il dénote une situation délimitée dans le temps, qui couvre un intervalle de courte durée. Notons qu'avec certains modificateurs adverbiaux (*czasem* 'parfois', *od czasu do czasu* 'de temps en temps', etc.), *potęsknić* acquiert une lecture itérative. L'intensité du sentiment qu'il exprime est nettement plus faible que celle de *zateęsknić*.

- (14) On ma za mną/do mnie *zateęsknić*.

on_{PRON.3.MASC.NOM} mieć_{V.PR.3} za_{PRÉP} ja_{PRON.1.INS} /do_{PRÉP} ja_{PRON.1.GÉN} *zateęsknić*_{INF}
 'Il doit commencer à ressentir que je lui manque.'

- (15) *Fajnie jest czasem potęsknić za kimś/do kogoś trochę.*

fajnie_{ADV} być_{V.PR.3} czasem_{ADV} *potęsknić*_{INF} za_{PRÉP} ktoś_{PRON.INS.SG}/do_{PRÉP} ktoś_{PRON.GÉN.SG} trochę_{ADV}

'C'est sympathique de ressentir de temps à autre que quelqu'un nous manque un peu.'

Le cas du perfectif *stęsknić się*, formé à l'aide du préfixe s- et du clitique *się* est plus complexe¹⁹. Il a une valeur résultative en ce que l'expérimenteur apparaît comme affecté par l'accumulation que représente la charge affective ressentie. *Stęsknić się* dénote une sorte de réaction à l'état affectif²⁰. La présence du clitique *się* le rapproche des verbes de sentiment à construction médiane (ou moyenne) qui forment une large classe dans le lexique polonais (Dobrowolska-Pigoń, 2020).

D'après les dictionnaires consultés²¹, *stęsknić się* accepte les deux formes de complément prépositionnel. Cependant, notre requête sur corpus a retourné deux occurrences seulement de *stęsknić się do SN_{Gén}*. Dans l'immense majorité de ses emplois, *stęsknić się* sélectionne plutôt *za SN_{Ins}*, comme dans (16) et (17). De même, une vingtaine de locuteurs dont nous avons sollicité l'avis se sont montrés unanimes en faveur de cette deuxième construction, en qualifiant les exemples (18) et (19) d'artificiels, voire d'inacceptables.

¹⁹ Plusieurs dictionnaires font de *stęsknić się* une entrée, en admettant qu'il s'agit d'un lexème à part entière dont on peut paraphraser le sens sans recourir à celui de sa base de dérivation. Ainsi, le SJP en propose la définition suivante : 'odczuć dotkliwie brak kogoś lub czegoś' ('ressentir fortement l'absence de quelqu'un ou de quelque chose').

²⁰ Cf. Wilczewska (1966 : 54) citée dans Dobrowolska-Pigoń (2020 : 124).

²¹ Voir, à titre d'exemple : https://wsjp.pl/haslo/do_druku/77736/stesknic-sie.

(16) *Stęskniłam się za domem.*

*stęsknić*_{V.PST.I} *się*_{PRON} *za*_{PRÉP} *dom*_{N.INS.SG}

(17) *Stęskniłam się za dziećmi.*

*stęsknić*_{V.PST.I} *się*_{PRON} *za*_{PRÉP} *dziecko*_{N.INS.PL}

(18) ? *Stęskniłam się do domu.*

*stęsknić*_{V.PST.I} *się*_{PRON} *do*_{PRÉP} *dom*_{N.GÉN.SG}

(19) ? *Stęskniłam się do dzieci.*

*stęsknić*_{V.PST.I} *się*_{PRON} *do*_{PRÉP} *dziecko*_{N.GÉN.PL}

Une explication possible serait que *stęsknić się* montre une nette préférence pour *za* *SN_{Ins}*, parce que sa valeur résultative de réaction à l'état qui affecte l'expérienceur s'accorde mieux avec l'orientation de ce complément vers le passé. Or, avec *do* *SN_{Gén}*, le sentiment décrit comprend un élément volitionnel qui tend vers le futur, c'est-à-dire au-delà de la borne temporelle de cette forme perfective.

4.4. Intensification

Les collocatifs intensificateurs²² permettent d'observer les congruences avec la base, et plus précisément, avec les composantes sémantiques sur lesquelles ils portent. Nous reprendrons ici la double distinction que nous avons utilisée pour analyser la combinatoire du lexème *manquer* qui, dans certaines conditions, peut traduire le contenu sémantique de *tęsknić* (Sikora, 2023).

La première distinction est celle que propose Kleiber (2013) entre deux types d'intensité. Il y a, en effet, celle évoquée plus haut, qui constitue un trait sémantique lexicalisé de certaines lexies. Cette intensité est l'une des propriétés définies des sentiments, des couleurs, des odeurs, etc. En paraphrasant le sens d'*amour* et d'*aimer*, nous devons spécifier cette composante, car c'est entre autres selon ce paramètre que diffèrent par exemple *aimer* et *adorer*. Mais dans d'autres cas, l'intensité se définit comme le degré de possession d'une propriété. Ainsi, pour les verbes de déplacement, la vitesse constitue l'un des paramètres définis. Dans une phrase telle que *Yann a l'habitude de marcher très vite*, c'est ce

²² En Lexicologie Explicative et Combinatoire, ce sont ceux qui constituent la valeur de la fonction lexicale syntagmatique *Magn*. Pour une mise à jour récente du répertoire des fonctions lexicales, voir Mel'čuk et Polguère (2021).

trait qui se trouve intensifié pour indiquer un degré considéré comme supérieur à la normale.

La deuxième distinction est celle entre deux modes d'intensification décrits par Molinier (1990) et Molinier et Levrier (2000) dans leur classement des adverbiaux. Le premier repose sur une évaluation quantitative et s'exprime par des adverbiaux tels que *beaucoup*, *très*, *un peu*, *quelque peu* en français ; *bardzo*, *trochę*, *ogromnie* en polonais. Cette intensification par quantification massive est, selon Kleiber (2013), caractéristique de l'intensité du premier type, celle-là même qui constitue une propriété des lexèmes du champ sémantique des sentiments.

Les adverbes français *terriblement*, *cruellement*, *horriblement*²³, etc., ainsi que *strasznie*, *okropnie*, *koszmarnie*, etc. en polonais intensifient la prédication par évaluation appréciative. Dans le cas des verbes de sentiment, ils qualifient le ressenti par un jugement porté sur ce paramètre (voir ci-dessus, les composantes impliqués dans un état affectif définies par Ju. Apresjan et V. Apresjan, 2000). Dans les exemples (1) à (4), *tęsknić* peut être intensifié aussi bien par quantification massive exprimée par *bardzo* et *trochę* que par évaluation appréciative à l'aide de *strasznie* et *okropnie*. Une modification intensifiante est en revanche problématique dans (5)–(6) et dans (7)–(8).

Il est temps à présent de vérifier dans quelle mesure ces trois caractéristiques – la forme du complément sélectionné, le potentiel dérivationnel et le mode d'intensification – contribuent aux portraits lexicographiques des lexèmes qui forment la polysémie de *tęsknić*.

5. Une galerie de portraits polysémiques

La double perspective orientée à la fois vers les traits communs des lexèmes affectifs et vers les propriétés particulières du verbe *tęsknić* conduit à l'hypothèse d'une polysémie à quatre acceptions. Deux d'entre elles, *tęsknić* 1.1 et *tęsknić* 1.2, relèvent du champ sémantique des sentiments. Les deux autres lexicalisent

²³ Molinier et Levrier (2000) qualifient ces adverbes d'intensifs-appréciatifs, en considérant qu'ils ont un effet intensifiant auprès des adjectifs et une valeur appréciative, lorsqu'ils modifient les verbes. Sikora (2023) montre qu'en cas de verbes de sentiment, ces adverbiaux expriment une intensité directement liée à l'appréciation du ressenti. En ce qui concerne le polonais, on remarque cependant qu'ils réalisent une évaluation appréciative auprès de prédicats à polarité négative. Ailleurs, par exemple dans *Ona strasznie go kocha* 'Elle l'aime terriblement', l'appréciation semble neutralisée au profit de l'intensification.

respectivement un attitude épistémique (*tęsknić II*) et une forte volonté d'obtenir quelque chose (*tęsknić III*).

Les portraits lexicographiques esquissés dans cette section ne prennent pas la forme de définitions lexicographiques. Leur objectif est de présenter les configurations particulières de composantes sémantiques dans leurs articulations aux propriétés syntaxiques (la forme de leurs compléments), morphologiques (le potentiel dérivationnel) et combinatoires (le type de modificateurs intensifiants). Chemin faisant, nous reprenons les exemples (1) à (8) pour illustrer chacune des acceptions.

TĘSKNIĆ I.1

- (1) *Bardzo tęsknię za swoją córką.*

bardzo_{ADV} tęsknię_{V.PR.1} za_{PRÉP} swoją_{PRON.INS.SG} córką_{N.INS.SG}
 'Ma fille me manque beaucoup.'

- (2) *Tęskniła za domem, za Suzanne, za małą Zuzią...*

teńsknił_{V.PST.3} za_{PRÉP} dom_{N.INS.SG}, za_{PRÉP} Suzanne_{Np.INS.SG}, za_{PRÉP} mała_{ADJ.INS.SG}
 Zuzia_{Np.INS.SG}
 'La maison, Suzanne, la petite Zuzia lui manquaient.'

Tęsknić I.1 est un verbe de sentiment dont le second actant s'exprime par un complément *za* SN_{Ins}. Le ressenti de l'expérimenteur trouve sa source dans la perception de deux situations. L'une, située dans le passé, où Y était physiquement et affectivement proche, est jugée positive, en contraste avec la période actuelle au moment de référence. Cette dernière, marquée par la séparation, est perçue comme quelque chose de mauvais. La perception de la séparation d'avec Y est la véritable raison d'un ressenti qui s'apparente à une douleur inscrite dans la durée. Orienté vers le passé, *tęsknić I.1* n'a pas de composante volitionnelle et ne présente aucun caractère dynamique.

Les deux modes d'intensification sont possibles, en renforçant par des modificateurs de quantification massive le trait d'intensité caractéristique de l'état affectif lui-même ou bien en caractérisant le ressenti pas un jugement appréciatif.

Quant à son potentiel dérivationnel, *tęsknić I.1* permet de créer des dérivés perfectifs en sélectionnant les préfixes *ZA-*, *PO-*, ainsi que la forme médiane *s-* *SIĘ* avec les valeurs présentées dans la section précédente.

TĘSKNIĆ 1.2

L'acception *tęsknić* 1.2, dont le second actant se réalise sous forme de complément *do* SN_{GÉN}, appartient également au champ des sentiments, mais elle se distingue de la première par un caractère nettement plus dynamique. Les exemples (3) et (4) rappelés ci-dessous en sont des illustrations :

- (3) *Biedne dziecko, Maciek tęskni bardzo do matki.*

biedny_{ADJ.NOM} dziecko_{N.NOM} Maciek_{Np.NOM} tęsknić_{V.PR.3} do_{PRÉP} matka_{N.GÉN}
 'Pauvre enfant, sa mère manque beaucoup à Maciek.'

- (4) *Boże, jak ja strasznie/okropnie tęsknię do tego kraju.*

Bóg_{N.VOC.SG}, jak_{PRON} ja_{PRON} strasznie_{ADV}/okropnie_{ADV} tęsknić_{V.PR.1} do_{PRÉP} ten_{PRON} kraj_{N.GÉN}
 'Mon Dieu, qu'est-ce que ce pays me manque terriblement/atrocement.'

Ce lexème ne présuppose pas de contact préalable entre X et Y : l'exemple (13) discuté plus haut montre que l'expérienceur peut ne jamais avoir rencontré la personne vers laquelle son sentiment s'oriente.

- (13) *Tęsknię do ojca, którego nigdy nie miałam.*

tęsknić_{V.PR.1} do_{PRÉP} ojciec_{N.GÉN.SG} który_{PRON.GÉN.SG} nigdy_{ADV} nie_{NÉG} mieć_{V.PST.1}
 'Je me languis d'un père que je n'ai jamais eu.'

L'élément volitionnel – le désir de (re)trouver Y – tient une position centrale au même titre que le ressenti négatif de l'expérienceur, auquel la présence de Y pourra mettre un terme. Le sentiment décrit par *tęsknić* 1.2 s'accompagne ainsi d'une certaine attitude, ce qui d'une part suppose une intensité plus importante, et de l'autre, un dynamisme lié à une visée actionnelle vers une relation interpersonnelle, certes souhaitée seulement, mais dont l'intention semble indissociable de l'état affectif.

Comme pour l'acception de base, une modification intensifiante est possible tant par quantification massive que par évaluation appréciative portant respectivement sur les composantes 'intensité' et 'ressenti' lexicalisées dans ce verbe.

Tęsknić 1.2 permet de dériver des perfectifs en PO et ZA- : *potęsknić* et *zatęsknić* ont en commun de découper un segment temporel dont la borne finale n'a pas un caractère tétique. Autrement dit, ces perfectifs ont un contour temporel net, sans être articulés sur un état résultant. On remarque notamment que ces dérivés ne permettent pas de former de participe à valeur résultative (**potęskniony*, -a, -e ; **zatęskniony*, -a, -e). Ils s'opposent en cela à *stęsknić się* qui, lui, donne lieu

à la forme adjectivale *śęskniony*, *-a*, *-e*. Comme nous l'avons indiqué dans la sous-section 4.3, la construction *śęsknić się do SN_{Gén}* a soulevé des réticences chez nos informateurs probablement parce que ce verbe introduit une borne finale de nature télique qui s'accorde mal avec l'orientation de *śęsknić* 1.2 vers le futur, c'est-à-dire vers l'intervalle situé au-delà de cette borne, à droite de l'axe de temps.

ŚĘSKNIĆ II :

La troisième acception de *śęsknić*, illustrée dans les exemples (5) et (6), s'écarte du champ sémantique des sentiments pour se rapprocher des verbes épistémiques.

- (5) *Ludzie mogą śęsknić za socjalnymi zabezpieczeniami dawnego reżimu.*

ludzie_{N.PL.NOM} móc_{V.PR.6} śęsknić_{INF} za_{PRÉP} socjalny_{ADJ.INS.PL} zabezpieczenie_{N.INS.PL}
dawny_{ADJ.GÉN.SG} reżim_{N.GÉN.SG}

'Les gens peuvent regretter les garanties sociales de l'ancien régime.'

- (6) *Ciągła obecność hałaśliwego rodzeństwa, nadmiar głosów i gestów powodowały, że Aniela śęskniła za cichą samotnością.*

ciągły_{ADJ.NOM} obecność_{N.NOM.SG} hałaśliwe_{ADJ.GÉN.SG} rodzeństwo_{N.GÉN.SG},
nadmiar_{N.NOM.SG} głosy_{N.GÉN.PL} i_{CONJ} gesty_{N.GÉN.PL} powodowały_{V.PST-6} że_{CONJ}
Aniela_{NP.NOM} śęsknić_{V.PST.3} za_{PRÉP} cicha_{ADJ.INS.SG} samotność_{N.INS.SG}

'La présence constante de la fratrie bruyante, la surabondance des voix et des gestes faisaient Aniela se languir d'une solitude tranquille.'

Le second actant de *śęsknić* II est de nature situationnelle et il s'exprime par des compléments *za SN_{Ins}*. Comme le sens de base 1.1, ce verbe présuppose une expérience ou un usage préalables de Y par X, dont ce dernier est privé au moment de référence. Les composantes dominantes de cette acception sont ainsi liées à l'expérience d'une situation passée, révolue, et aux jugements de X qui l'opposent au *status quo* actuel : *les garanties sociales de l'ancien régime* et *une solitude tranquille* apparaissent comme préférables, sur un fond de regret, à l'état de choses en vigueur au moment de référence. Le ressenti négatif est ainsi le résultat d'un processus essentiellement cognitif et épistémique.

Plutôt qu'un sentiment, ce lexème dénote une préférence issue d'une comparaison de deux situations. Un état de choses passé au moment de référence est jugé positivement par rapport à la situation actuelle. On remarque d'ailleurs que l'on pourrait reformuler les phrases (5) et (6) en remplaçant *śęsknić* II par son quasi-synonyme plus général *woleć* (préférer).

Rarement attestée, l'intensification de ce lexème est néanmoins plus fréquente sur le plan quantitatif qu'appréciatif. Le nombre d'occurrences de *śęsknić* II modi-

fiées par *bardzo* anté- et postposé au verbe est faible : en réponse à notre requête sur corpus, nous avons obtenu 242 occurrences, dont 11 seulement illustrent ce sens. L'exemple (20) compte parmi elles.

(20) *Ja pracuję w firmie budowlanej. Bardzo tęsknię za kopalnią.*

ja_{PRON.I.NOM} pracować_{V.PR.I} w_{PRÉP} firma_{N.ABL.SG} budowlany_{ADJ.ABL.SG} bardzo_{ADV}
 tęsknić_{V.PR.I} za_{PRÉP} kopalnia_{N.INS.SG}
 'Je travaille dans une entreprise BTP. La mine me manque beaucoup.'

Quant à la modification appréciative, 9 concordances attestent de la construction *strasznie tęsknić II* parmi les 95 résultats retournés. Il est intéressant de remarquer cependant que *strasznie* semble désémantisé dans ces emplois, au point de ne plus exprimer une appréciation²⁴, mais de marquer un degré élevé de possession d'un trait. Parmi les autres marqueurs appréciatifs gardant un certain contenu lexical à polarité négative (*straszliwie, okropnie, piekielnie, potwor- nie*, etc.), seul *okropnie* apparaît deux fois, dont une entre guillemets. On en conclut que le contenu cognitif et épistémique du lexème *tęsknić II* se prête peu à des modifications d'intensité. L'exemple (5), dans lequel nous avons introduit un intensificateur quantitatif nous semble discutable en (21) :

(21) ? *Ludzie mogą bardzo tęsknić za socjalnymi zabezpieczeniami dawnego reżimu.*

ludzie_{N.PL.NOM} móc_{V.PR.6} bardzo_{ADV} tęsknić_{INF} za_{PRÉP} socjalny_{ADJ.INS.PL}
 zabezpieczeni_{N.INS.PL} dawny_{ADJ.GÉN.SG} reżim_{N.GÉN.SG}
 'Les gens peuvent regretter beaucoup les garanties sociales de l'ancien régime.'

De même, le potentiel dérivationnel de *tęsknić II* est plus réduit par rapport aux deux acceptions précédentes dans la mesure où ce troisième sens permet de dériver uniquement des perfectifs inchoatifs en *ZA-*.

(22) *Ludzie zatęsknią za spokojnymi rządami Pawlaka.*

ludzie_{Npl.NOM} zatęsknić_{V.FUT.6} za_{PRÉP} spokojny_{ADJ.INS.PL} rządy_{Npl.INS} Pawlak_{Np.GÉN.SG}
 'Les gens se mettront à préférer les temps tranquilles du gouvernement Pawlak.'

²⁴ À la différence de *okropnie* et *straszliwie*, l'adverbe *strasznie* est souvent employé avec des prédicats à polarité positive tels que *kochać*, ce que nous interprétons comme un signe d'une désémantisation plus avancée.

TĘSKNIĆ III :

Construite par extension métaphorique de *tęsknić* I.2, l'acception III ne se situe pas non plus dans le champ sémantique des sentiments car elle décrit une relation entre un individu et la situation dont il voudrait avoir le bénéfice : *komputer* – 'ordinateur' dans l'exemple (7) et *spokój* – 'tranquillité' dans (8). Certes, dans le cas de (7), le complément *do* SN_{Gén} renvoie à un objet physique, mais il s'agit d'un emploi métonymique : ce que souhaitent les enfants, ce sont les activités qu'ils peuvent pratiquer grâce à un ordinateur.

(7) *Dzieci tęsknią do komputera.*

dziecko_{N.NOM.PL} tęsknić_{V.PR.6} do_{PRÉP} komputer_{N.GÉN.SG}
'L'ordinateur manque aux enfants.'

(8) *Ja tęsknię do spokoju, bo pozostał mi rok do emerytury – powiedział pan Radziszewski.*

ja_{PRON.I.NOM} tęsknić_{V.PR.1} do_{PRÉP} spokój_{N.GÉN.SG} bo_{CONJ} pozostać_{V.PST.3} ja_{PRON.I.DAT.SG}
do_{PRÉP} emerytura_{N.GÉN.SG} – powiedzieć_{V.PST.3} pan_{N.NOM.SG} Radziszewski_{ADJ.NOM.SG}
'J'ai hâte d'être tranquille, car il me reste un an jusqu'à la retraite – dit Monsieur Radziszewski.'

Tęsknić III dénote en effet un état psychique plutôt qu'affectif. Le sens de ce lexème est dominé par la composante cognitive que l'on peut décrire en renvoyant aux activités intellectuelles et à un jugement. Sans avoir forcément expérimenté Y auparavant, X pense beaucoup et de manière récurrente à Y, en jugeant que l'usage qu'il pourra en faire lui procurera du plaisir et du bien-être. Or, comme le remarque Będkowska-Kopczyk (2014), l'intellectualisation d'un état confère à ce dernier un caractère volitionnel : Y constitue ainsi un objet sur lequel porte la volonté du premier actant. L'acception *tęsknić* III décrit donc un procès intellectuel associé à une volonté d'action en vue de (re)trouver Y. On remarque d'ailleurs que les exemples (7) et (8) peuvent être reformulés à l'aide de collocations des SN qui suivent *do* placés dans la portée du modalisateur *vouloir*, respectivement : *dzieci chcą używać komputera* – 'les enfants veulent utiliser un ordinateur' ; *chcę spokoju* – 'je veux être tranquille'.

Le caractère cognitif et intellectuel du prédicat se manifeste sur le plan syntaxique par la possibilité d'employer ce verbe avec un complément propositionnel. Ces structures apparaissent surtout à l'oral ; parmi toutes les acceptions de *tęsknić*, seul le sens III les autorise. L'énoncé (23) est l'une des 14 occurrences que nous en avons repérées en corpus.

(23) *Kraść nie kradnę i tylko tęsknię, żeby jakiś pracę znaleźć.*

kraść_{INF} nie_{NÉG} i_{CONJ} tylko_{ADV} tęsknić_{V.PR.1}, żeby_{CONJ} jakiś_{DÉT.ACC} pracan_{ACC.SG}
znaleźć_{INF}

‘Voler – non, je ne vole pas. Je voudrais simplement trouver un travail.’

Le lexème *tęsknić* *III* accepte l'intensifieur quantitatif *bardzo* – on pourrait l'insérer sans problème dans l'exemple ci-dessus – pour renforcer la composante volitionnelle. Dans (23), il indiquerait alors le haut degré du souhait de trouver un emploi. Quant aux dérivés perfectifs, *tęsknić* *III* permet de forger uniquement l'inchoatif *zatęsknić*, comme dans l'exemple (24) :

(24) *Ale ludzie, nie przyzwyczajeni do samodzielności, zatęsknili do państwa opiekuńczego.*

ale_{CONJ} ludzie_{Npl.NOM} nie_{NÉG} przyzwyczajony_{ADJ.NOM.PL} do_{PRÉP}
samodzielność_{N.GÉN.SG} zatęsknić_{V.PST.6} do_{PRÉP} państwo_{N.GÉN.G}
opiekuńczy_{ADJ.GÉN.SG}

‘Mais les gens, non coutumiers de liberté de choix, se sont mis à souhaiter le retour de l'état protecteur.’

6. Conclusions

En décrivant un lexème, le chercheur se doit d'accorder autant d'attention aux traits que celui-ci partage avec les autres unités de son type lexicographique, qu'à ceux qui lui sont particuliers. Les analyses reposent ainsi sur un mouvement unifiant et sur une démarche individualisante (cf. Apresjan, 2008 : xvi, *unification et individuation*). Ces deux aspects ont été pris en compte dans notre étude, même si l'on pourrait nous reprocher un certain déséquilibre au profit des propriétés particulières du verbe *tęsknić*. Nous avons construit notre recherche en nous appuyant sur de nombreux travaux antérieurs qui ont mis en évidence un certain nombre de traits communs aux lexèmes du champ lexical des sentiments, notamment en ce qui concerne les composantes sémantiques relatives aux sept systèmes impliqués dans les expériences affectives. Forte de cet apport, nous avons pu nous concentrer sur les structures sémantiques de chacune des acceptions de *tęsknić* en lien avec leurs caractéristiques linguistiquement pertinentes : la complémentation, le potentiel dérivationnel et les collocatifs intensifiants. L'observation de ces traits particuliers nous conduit à proposer une hypothèse de polysémie avec

quatre acceptions de *tęsknić*, dont deux seulement s'inscrivent dans le champ des sentiments. Les deux autres sens, *tęsknić II* et *tęsknić III*, s'organisent autour des composantes cognitive et volitionnelle.

Une fois notre réflexion exposée, on pourrait nous reprocher – non sans raison – de nous arrêter là où le travail lexicographique par excellence devrait commencer. En effet, si les caractéristiques sémantiques et les traits linguistiquement pertinents de chacune des quatre acceptions ont bien été réunies et discutées, nous n'en proposons pas de transpositions sous forme de définitions. Ce n'est pourtant pas l'envie qui manque – plutôt la place, puisque chacune des paraphrases constitue en fait une étude à part. Il convient notamment de vérifier comment ces copolysèmes s'intègrent dans les types lexicologiques respectivement des lexèmes de sentiments, d'opinion et de volition. Ensuite, il faudrait rattacher à ce réseau lexical les dérivés perfectifs des différents lexèmes *tęsknić*, avec leurs propres polysémies. Sans oublier une discussion expliquant les choix des dénominations retenues pour les composantes définitionnelles en cas de paraphrases en français. Pour nous y être déjà attelée ailleurs (Sikora, 2020) juste à titre d'illustration et sans aboutir à des résultats satisfaisants, nous pensons que l'on est encore loin d'en avoir fini avec le verbe *tęsknić*.

Abréviation des gloses

Parties du discours : ADJ – adjectif, ADV – adverbe, CONJ – conjonction,
DÉT – déterminant, N – nom, N_p – nom propre, N_{pl} – *plurale tantum*,
NUM – numéral, PRÉP – préposition, PRON – pronom, V – verbe,
INF – verbe infinitif; NÉG – opérateur de négation;
catégories grammaticales : MASC – masculin, SG – singulier, PL – pluriel;
catégories flexionnelles : 1 – 1^{re} personne (du singulier),
3 – 3^e personne (du singulier), 6 – 6^e personne (3^e personne du pluriel),
ABL – ablatif, NOM – nominatif, GÉN – génitif, DAT – datif, INS – instrumental,
VOC – vocatif

Références citées

- Anscombre, J.-C. (1995). Morphologie et représentation événementielle : Le cas des noms de sentiment et d'attitude. *Langue Française* 105, 40–54.
- Apresjan, Ju. (1992). Systemic Lexicography as a Basis of Dictionary-making. *Journal of the Society of North America* 14, 79–87.
- Apresjan, Ju. (2008[2000]). *Systematic Lexicography* (K. Windle, Trad.). Oxford University Press.
- Apresjan, Ju. & Apresjan, V. (2000). Metaphor in the Semantic Representation of Emotions. In Ju. Apresjan, *Systematic Lexicography* (203–214). Oxford University Press.
- Augustyn, M. & Grossmann, F. (2014). Entre hyperonymie et spécification : Un drôle de sentiment. In P. Blumenthal, I. Novakova, & D. Siepmann (éds), *Les émotions dans le discours—Emotions in Discourse* (123–134). Peter Lang.
- Belletti, A. & Rizzi, L. (1988). Psych-Verbs and θ -Theory. *Natural Language & Linguistic Theory* 6(3), 291–352.
- Będkowska-Kopczyk, A. (2014). Verbs of emotion with se in Slovene : Between middle and reflexive semantics. A cognitive analysis. *Cognitive Studies* 14, 203–218.
- Blumenthal, P. (2009). Les noms d'émotion : Trois systèmes d'ordre. In I. Novakova & A. Tutin (éds), *Le lexique des émotions* (41–64). ELLUG.
- Diwersy, S., Goossens, V., Grutschus, A., Kern, B., Kraif, O., Melnikova, E. & Novakova, I. (2014). Traitement des lexies d'émotions dans les corpus et les applications d'EmoBase. *Corpus* 13, 269–293.
- Dobrowolska-Pigoń, M. (2020). Strona pośrednia a polskie czasowniki uczuć. *LaMi-CuS* 4(4), 122–150.
- Dziwirek, K. & Lewandowska-Tomaszczyk, B. (2010). *Complex emotions and grammatical mismatches : A contrastive corpus-based study*. De Gruyter Mouton.
- Galisson, R. (1991). *De la langue à la culture par les mots*. Clé international.
- Johnson-Laird, N. & Oatley, K. (1989). The language of emotions : An analysis of a semantic field. *Cognition and Emotion* 3(2), 81–123.
- Kleiber, G. (2013). À la recherche de l'intensité. *Langue Française* 177, 63–76.
- Krzyżanowska, A. (2005). O polskiej « tęsknocie » i francuskiej « nostalgie ». *Poradnik Językowy* 1, 50–59.
- Krzyżanowska, A. (2014). Les termes génériques du vocabulaire affectif : Le cas de « sentiment » et « uczucie ». In P. Blumenthal, I. Novakova, & D. Siepmann (éds), *Les émotions dans le discours – Emotions in Discourse* (105–120). Peter Lang.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar—Theoretical Prerequisites*. Stanford University Press.

- Melčuk, I. (2024). An Epoch Has Ended: To Apresjan's Dear Memory. *Russian Journal of Linguistics* 28(2), 480–483.
- Melčuk, I., Clas, A. & Polguère, A. (1995). *Introduction à la Lexicologie Explicative et Combinatoire*. Éditions Duculot.
- Melčuk, I. & Polguère, A. (2016). La définition lexicographique selon la Lexicologie Explicative et Combinatoire. *Cahiers de Lexicologie* 2(109), 61–91.
- Melčuk, I. & Polguère, A. (2021). Les fonctions lexicales dernier cri. In S. Marengo (éd.), *La théorie Sens-Texte. Concepts-clés et application* (73–153). L'Harmattan.
- Molinier, C. & Levrier, F. (2000). *Grammaire des adverbes. Description des formes en -ment*. Droz.
- Narodowy Korpus Języka Polskiego. <https://nkjp.pl>.
- Novakova I. & Tutin, A. (éds). (2009). *Le lexique des émotions et sa combinatoire syntaxique et lexicale*. ELLUG.
- Nowakowska-Kempna, I. (1995). *Konceptualizacja uczuć w języku polskim*. Wyższa Szkoła Pedagogiczna Tow. Wiedzy Powszechnej w Warszawie.
- Pajdzińska, A. (2003). Obraz tęsknoty w polszczyźnie. *Zeszyty Naukowe WSHE w Łodzi* 3(41), 7–17.
- Polguère, A. (2013). Les petits soucis ne poussent plus dans le champ lexical des sentiments. In F. Baidier & G. Cislariu (éds), *Cartographie des émotions. Propositions linguistiques et sociolinguistiques* (21–42). Presses Sorbonne Nouvelle.
- Polguère, A. (2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography* 27(4), 396–418.
- Polguère, A. (2016). *Lexicologie et sémantique lexicale – Notions fondamentales*. Les Presses de l'Université de Montréal.
- Polguère, A. (2018). A Lexicographic Approach to the Study of Copolysemy Relations. *Russian Journal of Linguistics* 22(4), 788–820.
- Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells*. De Gruyter Mouton.
- Sikora, D. (2008). Éléments d'une sémantique pour la description d'un sentiment. *Arena Romanistica* 1, 124–142.
- Sikora, D. (2020). Sémantique lexicale et didactique des langues. Synthèse des travaux présentée en vue de l'obtention de l'Habilitation à diriger des recherches. En ligne : <https://hal.science/tel-04014291v1>.
- Sikora, D. (2023). Combinatoire lexicale et polysémie: Le cas de MANQUER. In L. Meneses Lerín, S. Mejri & J. Goes (éds), *La combinatoire lexicale : théories, langues et productions langagières* (181–198). L'Harmattan.
- Słownik Języka Polskiego. Państwowe Wydawnictwo Naukowe. <https://sjp.pwn.pl>.
- Spokes. *Conversational Data Search*. <https://spokes.clarin-pl.eu>.
- Stawnicka, J. (2015). Czasownikowe formanty modyfikacyjne w języku polskim. *Studia z Filologii Polskiej i Słowiańskiej* 45, 81–97.

- Talmy, L. (2000b). *Toward a Cognitive Semantics. Typology and Process in Concept Structuring* (Vol. 2). MIT Press.
- Wierzbicka, A. (1969). Struktura semantyczna słów nazywających uczucia (T. Lisowski, Trad.). *Poznańskie Studia Polonistyczne. Seria Literacka* 34(54), 275–289.
- Wierzbicka, A. (1992). *Semantics, Culture, and Cognition. Universal Human Concepts in Culture-Specific Configurations*. Oxford University Press.
- Wierzbicka, A. (1995). Emotion and Facial Expression: A Semantic Perspective. *Culture & Psychology – CULT PSYCHOL* 1, 227–258.
- Wierzbicka, A. (2011). What's wrong with "happiness studies"? The cultural semantics of *happiness, bonheur, Glück, and scastě*. *Word and Language (Slovo i Jazyk)* 1, 155–171.
- Wierzbicka, A. (2013). Human Emotions and English Words: Are Anger and Disgust Universal? In A. Wierzbicka (éd.), *Imprisoned in English: The Hazards of English as a Default Language* (68–86). Oxford University Press.
- Wilczewska, K. (1966). *Czasowniki zwrotne we współczesnej polszczyźnie*. Towarzystwo Naukowe w Toruniu.
- Żmigrodzki, P. (éd.). *Wielki Słownik Języka Polskiego*. <https://wsjp.pl>.



Irina Kor Chahine

University Cote d'Azur, CNRS, BCL
France

 <https://orcid.org/0000-0001-7564-9996>

When The Moment Does Not Last: A Few Words On *Moment*-Constructions in Russian

Abstract

The article is devoted to constructions with the noun 'moment' (MC) in modern Russian. Our corpus study shows that Russian *moment* is mainly used in certain constructions. We focus on its four main types; each of them determines the semantics of the noun *moment*. While the word originally designated an instant, a precise moment (*moment vzryva* 'moment of explosion'), it also designates a long temporal period with very vague boundaries (*na dannyj moment* 'currently, at today's time') or loses its temporal meaning altogether in order to designate a particular aspect of the situation (*ličnyj moment* 'personal moment') which places it in the domain of argumentative vocabulary. Some of these constructions are characterized by a certain degree of fixity. The paper describes both syntactic and semantic features of each of the four models through the three centuries of evolution of the MC in modern Russian.

Keywords

Russian, semantics, lexis, constructions, lexicalization, diachrony, corpus study

1. Introduction

Under the term *construction*, we understand a syntactically linked lexical phrase denoting a certain semantics. One usually speaks of verbal constructions, with a verb and its arguments (such as *to do my best*), or of prepositional ones (e.g. *in the future*) or also of adjectival (e.g. *happy about it*) constructions. Less common are nominal constructions (such as *no way*). This may be because nouns generally have fewer arguments and are usually used themselves as arguments of a verb or a preposition.

Since constructions have their own semantics, the semantics of nouns making them up can vary depending on close context. It is possible for a noun to take on a meaning which deviates quite far from its initial meaning. This is particularly true of constructions with Russian *moment*.

The noun *moment* is one of the most used nouns at the beginning of this century (Ljaševskaja & Šarov 2009)¹. Its frequent usage caused it to progressively extend its semantics. *Bol'šoj tolkovyj slovar'* (BTS 2000) explains its usage in three meanings: i) the noun *moment* with a noun in Genitive case designates the time at which an action or a process begins (*nabljudat' moment vosxoda solnca* 'to observe the sunrise'); ii) with Genitive or Determinative, *moment* designates a time period of development of something (*tekuščij, nastojaščij, dannyj moment* 'the present, current, this moment'), and iii) *moment* points to a particular aspect of an event or a circumstance (*Pri provedenii reformy ne učteny ves'ma suščestvennye momenty* 'The reform failed to take into account several highly significant factors') (BTS 2000, *moment*).

Based on its Latin origin (*momentum*), dictionaries define *moment* as designating an instant, a very brief time period, which places *moment* very close to Russian *mig* and *mgnovenie* ('instant'), and their proximity is often mentioned (see also Ožegov & Švedova). However, the semantics of this trio is not identical. These nouns have not yet received an in-depth description in terms of synonymy like items described by the *New Dictionary of Synonyms* (Apresjan 2004).

We intend here to make a detailed description of syntactic constructions with Russian *moment* as well as to follow its dynamics through the Russian National Corpus from the 18th to the 21st century.

2. Materials and methods

Our study is based on the Main corpus of *Russian National Corpus* (henceforth Russian corpus). This corpus represents an overview of texts belonging to diverse genres and styles manifesting the usage of words in contemporary Russian. Moreover, it may reveal possible functional changes over the duration of the last three centuries. The working corpus was extracted by automated random selection both from the beginning (18th and 19th centuries; 785 occurrences) and the ending (the 21st century; 575 occurrences) samples of the Main Corpus.

¹ It ranks at the 99th position (of 1001st nouns) in the Russian Dictionary of frequency.

Table 1.*Numbers of occurrences in a working corpus*

Time period	1700–1799	1800–1869	...	2015–2018
occurrences	49	736	—	575

The paper is structured as follows: Section 3 discusses semantic particularities of each of four *moment*-constructions (MC) presented before. Section 4 is devoted to quantitative data and to the dynamics of the semantics of MC over time. Section 5 presents conclusions.

3. Semantics of moment-constructions

To provide a deeper analysis of the noun *moment* in modern Russian, let us examine syntactic constructions where it usually appears. Our data processing identified four main constructions most frequent with Russian *moment*:

- (i) *moment* enters into a prepositional phrase with *v* and *na* and a determinative;
- (ii) *moment* is followed by a Genitive;
- (iii) *moment* enters in a relation with a verb;
- (iv) *moment* is accompanied by an adjective.

3.1. V / Na Det moment (*v dannyj moment* ‘at the moment’)

According to our data, the most frequently used construction with *moment* is a prepositional phrase. A complementary search on collocations of 3-grams “prep + adj + *moment*” in Russian corpus shows that, with an exception with *do* in *do poslednego momenta* ‘until the last moment’, in the twenty first positions we exclusively find prepositions *v* and *na* with Accusative case: *v / na dannyj moment* ‘at the moment’, *v poslednij, nastojaščij, nužnyj, sledujuščij, opredelennyj m.* ‘at the last, present, right, next, definite moment’, *na segodnjašnjij m.* ‘for the time being’, *v kritičeskij, rešajuščij, izvestnyj, rešitel'nyj m.* ‘at a critical, decisive, certain, crucial moment’, *na tekuščij, nastojaščij m.* ‘now, currently’, *v kritičeskie, raznye momenty* ‘at critical, different moments’, and so on. In such constructions *moment* mostly has a temporal meaning, like in other constructions with

Accusative case (*v ètu sekundu* ‘at this very second’, *na sledujuščee voskresen’e* ‘the following Sunday’, *v tot den’* ‘on that day’).

Table 2.

Corpus search by 3-grams “prep + adj + moment” (top 10)

№	Occurrences	Ratio n-gram, %	Documents	3-grams “prep + adj + moment”	Translation
1	2257	23.24	1486	v dannyj moment	<i>at the moment</i>
2	1252	12.89	940	v poslednij moment	<i>at the last moment</i>
3	1017	10.47	835	v nastojaščij moment	<i>at present, currently</i>
4	757	7.80	642	na dannyj moment	<i>for the moment</i>
5	469	4.83	417	v nužnyj moment	<i>at the right moment</i>
6	205	2.11	154	v sledujuščij moment	<i>in the next moment</i>
7	201	2.07	189	do poslednego momenta	<i>until the last moment</i>
8	156	1.61	134	v opredelënnij moment	<i>at a certain point</i>
9	132	1.36	108	na segodnjašnjij moment	<i>for the time being</i>
10	131	1.35	124	v kritičeskij moment	<i>at a critical time</i>

Table 2 presents the top ten of 3-grams collocations ‘prep + adj + moment’ used in the Russian corpus. This sample shows a possible “specialization” of the prepositional MC. Almost half of the uses (43%) are represented by four models: *dannyj* and *na dannyj moment* (31%), *v nastojaščij moment* (10%) and *na segodnjašnjij moment* (1%). All these constructions are semantically very close. Furthermore, they manifest an evident tendency to lexicalization. While *dannyj* and *ètot* have very close semantics and can be viewed as synonymous (as in *dannyj / ètot predmet*), in MC they are not. For example: we can say *V ètot moment on vsë ponjal* but not **v dannyj moment on vsë ponjal* ‘At that moment, he understood everything’. In conjunction with *moment*, *ètot* leads the noun to its initial semantics of an instant while *dannyj* conducts it to a time period interpretation. Therefore, *v dannyj moment* becomes semantically close to *v nastojaščeje vremja* pointing out a current time period. The *moment* becomes consequently synonymous to *vremja* ‘time’.

Thus, the components of these constructions semantically combine perfectly well; the adjectives *nastojaščij*, *segodnjašnjij* and *dannyj* denoting the present moment just as the noun *moment* denotes a moment that can really refer only to the present instant. The semantics of these MCs does not point out “the moment”, a *punctual instant*, but has a more extended meaning of a longer time

period. In terms of Constructional Grammar, we deal here with purely idiomatic “constructions” whose meaning cannot be deduced from the combination of meanings of its components and which are characterized by a certain degree of fixity.

Nevertheless, in prepositional constructions *moment* does not always carry on the semantics of a longer period. In similar contexts, with the adjective *poslednij* ‘last’, *moment* is not synonymous to *vremja* ‘time’ and indicates on the contrary a short instant (*v poslednij moment* ‘at the last moment’ vs. *v poslednee vremja* ‘last time’).

The semantics of a short, punctual instant can appear with other determinants, like adjectives (*poslednij* ‘last’, *nužnyj* ‘necessary’, *sledujuščij* ‘next’, *opredelennyj* ‘certain’, *kritičeskij* ‘critical’) and pronouns (*ëtot* ‘this’, *takoj* ‘such’, *kakoj-to* ‘some moment’ ‘moment’).

3.2. (Prep) moment G (*s momenta roždenija* ‘from the moment of birth’)

When *moment* is followed by a genitive noun, this construction can denote either an event, a state, or a process. Depending on the case, the interpretations of *moment* will be different. For example, when the construction denotes an event, we find mostly deverbal nouns formed from perfective verbs. Here are some examples from the 21st century: *moment probuždenija* ‘the moment of awakening’, *prosvetlenija* ‘of enlightenment’, *pojavlenija* ‘of emergence’, *vstreči* ‘of encounter’, *otkrytija* ‘of discovery’, *zapuska* ‘of launch’, *pokupki* ‘of purchase’, *sotvorenija* ‘of creation’, *začatija* ‘of conception’, *roždenija* ‘of birth’, *publikacii* ‘of publication’, *oplaty* ‘of payment’, *snjatija* ‘of withdrawal’, *sokrytija* ‘of concealment’, *sžatija* ‘of compression’, *vzryva* ‘of explosion’, *ubijstva* ‘of murder’...

In this sample, most nouns, and in particular the deverbal nouns ending by *-ie*, denote an instantaneous event and often bear inchoative semantics (cf. inchoative prefixes *po-* and *za-*). The words *probuždenie* ‘awakening’, *publikacija* ‘publication’ or *zaključenie* ‘conclusion’ mark change in a process where a new state is established after a previous state. So, one is awakened after a sleep, a text is published after a manuscript version, or a conclusion is reached after a process of deliberation that precedes it. The same semantics also characterizes such nouns as *vsreča* ‘meeting’, *pokupka* ‘purchase’, *ubijstvo* ‘murder’, *načalo* ‘beginning’, *vychod* ‘exit’, *perexod* ‘transition’, *vzryv* ‘explosion’. In combination with the words of this group, *moment* focuses on a brief time period of transition from one state to another. It is often a punctual instant. The semantics of *moment* can thus be schematized as a point from which a new state begins.

Moment can also be used in collocation with other nouns. With words denoting states, *moment* designates the time of a certain state. There are collocations like: *moment myšlenija* ‘a moment of reflection’, *bolezni* ‘of illness’, *pereživaniya* ‘of anxiety’, *uspokoeniya* ‘of tranquility’, *tišiny i garmonii* ‘of stillness and harmony’, and so on. With these nouns, there is no indication of the beginning or of a punctual instant specific to the original moment semantics. Here *moment* is close to non-telic nouns such as *sostojanie* ‘state’ (*sostojanie myšlenija* ‘thinking state’, *sostojanie bolezni* ‘state of disease’, *sostojanie tišiny i garmonii* ‘state of calm and harmony’). Here, the meaning of *moment* implies a longer duration, the time period of a certain state.

With event nouns, MC denotes the duration of a certain process: *moment prazdnika* ‘during celebration’, *m. obsuždenija* ‘during discussion’, *m. literaturnyx sporov* ‘during literary debate’, *m. trenirovki* ‘during exercise’, *m. tvorčestva* ‘during creative process’, *m. sozidanija* ‘during creation’, *m. opasnosti* ‘moment of danger’, *m. povestvovaniya* ‘during narrative’, *m. krizisa* ‘during crisis’, *m. čtenija* ‘during reading’... where the ‘moment’ will last the duration of this process. As a result, *moment* still deviates from its initial semantics of punctual instant.

Thus, these semantic groups of nouns in Genitive constructions connect *moment* to its temporal meaning of a short instant or of a longer time period, similar to what we could observe in prepositional constructions. However, there are some nouns with which temporal interpretation is less obvious. They indicate by themselves a certain process with an obvious temporal dimension, like chronological time, lifetime, or duration of a story: *moment vremeni* ‘moment of time’, *žizni* ‘in life’, *prostoya* ‘of downtime’, *moment rasskaza* ‘moment of a story’, *čexovskoj «Čajki»* ‘of Chekhov’s “The Seagull”’, *čtvërtoj časti* ‘of the fourth part of...’, etc. Here *moment* focuses on a particular point or points in the process and might be close to the meaning of an *aspect*, a particular instance of something, especially when it is used in plural. This can be observed on the top of bi-gram search in the Russian corpus: *momenty žizni* ‘moments of life’, *vremeni* ‘of time’, *istorii* ‘of history’, *boja* ‘of battle’, *razvitija* ‘of development’, *inercii* ‘of inertia’, *tvorčestva* ‘of creativity’, *otčajanija* ‘of despair’, *opasnosti* ‘of danger’, *perexoda* ‘of transition’, *dviženija* ‘of movement’, *vдохновения* ‘of inspiration’, *bytija* ‘of being’, *dejstvija* ‘of action’, *krizisa* ‘of crisis’, and so on. Though the temporal interpretation is still possible, the non-temporal meaning for some occurrences (like *momenty istorii* ‘moments of history’) are preferable (cf. “certain points, questions of the history”).

The semantic deviation of *moment* from its temporal to a non-temporal meaning can also be observed in the verbal constructions below.

3.3. V moment (*nastupil moment* 'the moment has come')

The third type of constructions considering here involves verbs. In these constructions, *moment* can also have a temporal meaning (an instant or a time period), and can appear in position of a complement or a subject. Only a dozen collocations are frequently used in these constructions.

Constructions occurring with *moment* can be organized in three groups. **The first group** includes verbs that indicate that the word *moment* is understood as a linear and continuous process, like a time (as already mentioned previously), which we find in the following constructions:

- 1) *V moment*: *vyždat' m.* 'to bide your time', *ždat' momenta* 'to wait for the moment', *(ne) upustit' m.* '(to seize) to miss the moment', *propustit' m.* 'to miss the mark'.

Semantically speaking, here *moment* is getting closer to the words like *vremja* 'time', *slučaj* 'occasion': cf. *upustit' slučaj* 'to miss the chance' / *vremja* 'time', *ždat' (udobnogo) slučaja* 'to wait for an opportunity', *vyždat' vremja* 'to bide time'.

In subject position, *moment* is usually found in postposition to the predicate. In the predicate function, we often find the verbs *nastupat' / nastupit' / nastat'* 'to come' and *podospet' / podospet'* 'to arrive' with inchoative temporal semantics:

- 2) *Moment V*: *nastupit m.* 'there will come a time', *nastal m.* 'the moment has come', *nastupaet takoj m.* 'there comes a point', *podospel m.* 'the moment has arrived'.

Such word order with a predicate in anteposition brings constructions with *moment* closer to constructions of the type *nastupila vesna* 'the spring has arrived', *vzošlo solnce* 'the sun has risen', *vernulsja Petja domoj* 'Peter came home'... which introduce some event, and an event usually contains a temporal dimension. Canonical order with a subject anteposition (SVO) is very unusual here: cf. *??moment nastupit, nastal, nastupaet, podospel*... In these constructions, *moment* is often accompanied by a modifier or a subordinate with *kogda*:

- 3) *Nastupit moment, kogda vse ob ètom zabudut.*
'There will come a day when everyone will forget about it.'

The **second group** brings the semantics of *moment* closer to the nouns of **material objects**, of artefacts. Thus, *moment* can be grasped, chosen, touched, stopped or even returned to:

- 4) *V moment: sxvatit' m.* 'to grasp', *ulovit' m.* 'to catch', *vybrat' m.* 'to select', *prikosnut'sja k ètomu m.* 'to touch', *ostanovit'sja na ètom m.* 'to dwell on', *vozvraščat'sja k ètomu m.* 'to return to the moment'.

And finally, the **third group** treats the noun *moment* as **an aspect of a fact**, often the object of a discussion or mental operation:

- 5) *V moment: obsuždat' ètot m.* 'to discuss this point', *progovarivat' takoj m.* 'to talk about this point', *raz"jasnjat' nekotorye m.* 'to clarify some points', *vosprinimat' takie m.* 'to perceive these points', *otmetit' momenty* 'to note points', *pomnit' m.* 'to remember this moment', *perežit' ètot m.* 'to experience this fact', *sčest' m.* *podxodjaščim* 'to find this moment appropriate', *ètot m. ne otrabotali* 'this moment was not worked out'.

And these "moments" carry positive connotations:

- 6) *V moment: radovat'sja momentu* 'to enjoy the moment', *ljubovat'sja momentom* 'to admire the moment', *naslaždat'sja momentom* 'to savor the moment'; *èti momenty zavoraživajut menja* 'these moments fascinate me'.

Puskat' nekotorye problemy na samotëk, drugie rešat', no otvlečënno. Načat' radovat'sja momentu, smotret' na žizn' spokojno, a ne kak na boleznnennoe sostojanie. [kollektivnyj. «Veseluxa» na 11 tysjač znakov (2015.03.14)]

'To let some problems go, to solve others, but in a distracted way. Begin to *enjoy the moment*, to look at life calmly, not as a painful condition.' [Collective. "Fun" for 11 thousand characters (2015.03.14)]

3.4. Adj moment (*opasnyj moment* 'dangerous point')

Finally, the fourth type of constructions includes an adjective. Here we find again the temporal semantics, as an instant (especially with *è* + Accusative phrase) and as a time period as a part of prepositional and verbal phrases which we studied above as well as other constructions. Some of them, like *perelomnyj* 'turning point' and *ključevoj moment* 'key point', tend to lexicalization.

It is very interesting to notice that in adjective constructions, *moment* develops other, non-temporal, meanings. Its semantics becomes close to the meaning of ‘aspect’ or ‘side’, in particular with determinative (*dannyj* ‘this’) or modal (*nužnyj* ‘necessary’) adjectives and some evaluative adjectives like *interesnyj* ‘interesting’, *suščestvennyj* ‘essential’, *nastojščij* ‘real’, *vospitatel’nyj* ‘educational’, *položitel’nyj* ‘positive’ and some other.

But larger context is often needed to determine what the exact meaning of *moment* is. Compare:

- 7) *Èto proizošlo v samyj neožidannyj moment.* → ‘instant’
‘It came at the most unexpected moment’
- 8) *Neožidannym momentom bylo ego priznanie.* → ‘point, episode’
‘The unexpected point was his confession’

Regarding these examples, the role of predicates is very important. In various contexts, *moment* accompanied by an adjective tends to define a large spectrum of abstract meanings. Among them, we can find the following:

- *Meaning of ‘aspect’*: **ponimat’** religioznye, duxovnye *m.* ‘understand the religious, spiritual aspects’, **tam est’** očen’ važnye *m.* ‘there are some very important points’...
- 9) *Narod, kotoryj sozdal takuju pogovorku, rassmatrival žizn’ za predelami svoego fizičeskogo suščestvovanija — i èto odin iz važnejšix momentov čelovečeskogo bytija voobščë.* [A. Danilova, G. Gupalo. Jurij Norštejn: Opjat’ zabyli čeloveka // 2016]
‘The people who created such a saying considered life beyond their physical existence – and this is *one of the most important aspects of being human* in general.’ [A. Danilova, G. Gupalo. Yuri Norštein: Forgotten Man Again // 2016]
- *Meaning of ‘question’*: **otsylka k** dalěkomu istoričeskomu *m.* ‘reference to a distant historical question’, **prihodim k** očen’ suščestvennomu *m.* ‘come to a very significant point’...
- 10) *Èto krajne važnyj moment, na kotoryj Burd’e obraščaet vnimanie.* [A. Bikbov, S. Naranovič, “Gosudarstvo — èto illjuzija” (2017.05.31).] *This is an extremely important point to which Bourdieu draws attention.* [A. Bikbov, S. Naranovich, “The State is an Illusion.” (2017.05.31)]

- Meaning of 'event, episode': *u menja est' interesnyj ličnyj moment* 'I had an interesting event in my life'; *samyje važnye m. prošedšego goda* 'the most important events of the past year'...
- 11) *Čto čuvstvovali oni na vokzale, kogda im prišlos' uspokaivat' ego — tak on rydal. Ne raz lovil on sebja na tom, čto vo vremja dramatičnyx momentov vdruk sovsem kak postoronnij slyšal svoe xixikan'e, smex.* [«Volga», 2016]
'What they felt at the train station when they had to comfort him – how much he was sobbing. More than once he caught himself that *during dramatic events* suddenly quite as an outsider he heard his giggle, his laughter.' [“Volga”, 2016]
- Meaning of 'phase, period': *informacija o samyx rannix m. suščestvovanija Vselennoj* 'information about the earliest periods of the universe's existence', *vyxod ěnciklopedii otnetil očen' važnyj m.* 'the release of the encyclopedia marked a very important phase'...
- 12) — *Tovarišči pisateli! Nastal trudnyj moment v vašej žizni.* [A. Titov. Obščeziti pisatelej // «Volga», 2015]
'— Comrade writers! A *difficult period in your life has come*.' [A. Titov. Dormitory of Writers // “Volga”, 2015].
- Meaning of 'case, occasion':
- 13) *Xotja byvajut interesnye momenty: v 2015 godu ja povetil kartinku «Lolita» na Nevskom (...)* [Afiša Daily, 2018]
'Although *there are interesting cases*: in 2015, I hung a picture of Lolita on Nevsky (...) [Afiša Daly, 2018]

Nevertheless, this list of possible meanings of *moment* is far to be exhaustive. In some cases, the semantics of *moment* puts it apart from the meaning of existing abstract items in Russian, and it becomes very difficult to find any close synonym. This is what we see in the following examples:

- 14) — *Teper' Ol'ga obed ne skoro nam prigotovit, — i lomaja ves' vospitatel'nyj moment, ja zaržal.* [A. Klepakov. Opekun // «Volga», 2016]
'“Now Olga won't soon make us lunch,” and *breaking the whole educational 'moment'*, I laughed.' [A. Klepakov. The guardian // “Volga”, 2016].
- 15) *Odna polovina govorila, čto v transgendere voobšče net nikakogo osvoboždajuščego momenta, on davit na ženščinu eščë sil'nee, potomu čto, kak izvestno, transgendernye ljudi často gendernyj displej kak by utrirujut.*

[N. Nazarova. «Byl li Lenin feministom ili vse-taki seksistom». (23.03.2017)]
 'One half said that *there is no liberating 'moment' at all in being transgender*, it puts even more pressure on women because, as we know, transgender people often a little exaggerate gender display.' [N. Nazarova. "Was Lenin a feminist or still a sexist". (23.03.2017)]

- 16) *Okolačivalsja to v odnoj, to v drugoj komnate. (...) v konce koncov, Lukač prijutilsja v kabinete Marksa, gde zavedoval Rjazanov. **Političeskaja pozicija nesomnenno s pravymi momentami** <...>. Kak filosof on izvesten kak ukлонist.* [M. Nesterenko. *Ideologija i filologija*. (2017.05.16)]
 'He hung around in one room or the other. (...) eventually Lukács took refuge in Marx's office, where Ryazanov was in charge. *The political position is undoubtedly with right-wing 'moments'* <...>. As a philosopher he is known as an evader. (M. Nesterenko. *Ideology and philology*. (2017.05.16))
- 17) *Otveču i eščë na odin vopros: počemu my ob"javljaem o partii imenno sejčas, za tri dnja do vyborov. Potomu čto **pol'zuemsja medijnym momentom**. Sejčas daže Pervyj kanal ne smožet ob ètom promolčat' (...).* [«Snob», 2018]
 'I will also answer another question: why we are announcing the party now, three days before the elections. Because *we are taking advantage of the media 'moment'*. Now even Channel One will not be able to keep silent about it (...)' [«Snob», 2018]

In these constructions, the noun *moment* has obviously extended its semantics by gaining more abstract value. Contrarily to syncretic process guided by economy principle (cf. *vannaja komnata* → *vannaja*), the *moment* constructions go in the revert direction by decomposing and focusing some abstract notion denoting by an adjective.

See the following examples found in our corpus:

- 18) *Vytjagivat' samye interesnye momenty* → *samoe interesnoe*
Pereživat' perelomnyj moment → *perelom*
Ètapnyj moment → *ètap*
Vospitatel'nyj moment → *vospitanie...*

In such analytical collocations occurring in the 21st century, *moment* is currently filling a semantic gap in Russian abstract lexicon. Here, the initial meaning of time has completely moved on the periphery.

4. Discussion and conclusions

In this study, we have taken into consideration the four most used constructions with *moment*, which are:

1. **(Prep) moment G** (*s momenta roždenija* ‘from the moment of birth’), which indicates the starting point of a phenomenon marked by the genitive;
2. **V / Na Det moment** (*v dannij moment* ‘at the moment’) with a determinative pronoun or adjective, which indicates a particular instant;
3. **V moment** (*nastupil moment* ‘the time has come’, *ne upustit' moment* ‘to seize the moment’), where *moment* serves as an argument of the verb (subject or complement) and also denotes a period of time;
4. **Adj moment** (*opasnyj moment* ‘dangerous moment’), which may belong to Type 2 or Type 3 constructions above, but where *moment* develops a particular meaning.

According to the NKRJa, *moment*-constructions have been gradually undergoing changes over the past three centuries. Thus, if, at the beginning of the 18th century. Type 1 constructions with the genitive indicating the starting point of a phenomenon clearly prevailed (48% of all uses), at the beginning of the 21st century, Type 2 constructions became dominant, along with *V Det moment* (49%) and *Na Det moment* (15%) which make up the majority of uses (64%). Type 2 constructions, as well as those of Type 4 with an adjective (*ličnyj moment* ‘personal moment’) appeared in the corpus only since the twenties of the 19th century, probably under the influence of the French language². In these constructions we observed the rapid loss for *moment* of its original temporal meaning (Lat. *momentum*) in favor of a more general abstract sense. The noun *moment* became a part of argumentative vocabulary, and in this vein, Type 2 constructions *v / na nastojaščij moment* are gradually tending to the lexicalization.

Semantically speaking, the not-temporal meaning of *moment* brought it closer to nouns such as *aspect*, *épizod*, *étap* (*žizni*) which becomes *moment* (*žizni*). At the same time, *moment* takes more borrowed meaning than *tema*, *vopros*, *process*, probably filling a semantic gap in the Russian argumentative lexicon. As we could observe, *moment* serves in more analytic discourse collocations (cf. *vytjagivat' samye interesnye momenty* → *samoe interesnoe*). In any collocations with

² Cf. Julien, animé dans ce moment d'une force surhumaine, tordit un des chaînons de la chaîne qui retenait l'échelle. (Stendhal, Rouge et Noir, 1830, p. 358) (→ ‘at this moment’) | Ceci dit, ceci prouvé, il faudra agir. Ce sera le moment psychologique où Hamlet philosophe au lieu de faire. (Clémenceau, Vers réparation, 1899, p. 182) (→ ‘the psychological moment’) (TLF, *moment*).

preposition, verb or adjective, *moment*-constructions tend to figment and idiomaticity. And finally, throughout centuries, *moment*-constructions gained in usage in non-literary texts with a bigger predominance in the 21st century (82% vs. 18% for literary texts) compared to the 18th century data (72% vs. 28% respectfully).

To sum up, despite the fact that modern dictionaries still define the first meaning of *moment* as the temporal meaning of 'a very short period of time' (BTS), the corpus data certify that this meaning of *moment* is not the most used in Modern Russian. Semantic proximity of *moment* to the words *mig* and *mgnovenie* turns out to be only occasional. All these particularities of the Russian *moment*-constructions should be taken into account while dressing its 'lexicographic portrait'.

Sources

- Apresjan 2004: Apresjan, Ju. (éd). *Novyj ob''jasnitel'nyj slovar' sinonimov russkogo jazyka. 2e izd.*, Studia philologica, Moskva, Vena: Jazyki slavjanskoj kul'tury, Venskij slavističeskij al'manax, 2004, 1488 s. Open access: https://ruslang.ru/text_noss2_title?ysclid=m7rr4idcfl205043164.
- BTS 2000: Kuznecov, S. (éd). *Bol'soj tolkovyj slovar' russkogo jazyka*, 2000. 1536 c. (see <http://gramota.ru/slovari>).
- Ljaševskaja, O. & Šarov, S. (2009). *Častotnyj slovar' sovremennogo ruskojazyka (na materialax Nacional'nogo russkogo jazyka)*. Moskva: Azbukovnik. Open access: <http://dict.ruslang.ru/freq.php?>.
- NKRJa: *Nacional'nyj korpus russkogo jazyka*. Open access: <http://ruscorpora.ru>.
- Ožegov & Švedova: *Tolkovyj slovar' russkog jazyka*. Open access: <https://www.vedu.ru/expdic/>, <https://slovari.ru>.
- TLF: *Trésor de la langue française informatisé*. Open access: <http://www.atilf.fr/tlfi>.



Alexander Letuchiy

 <https://orcid.org/0000-0002-4151-2550>

Partial observation and nonstandard tense uses: Russian perfective verbs in progressive interpretation

Abstract

The article focuses on a special use of the Russian perfective past, which is virtually similar to a progressive one: it is illustrated by examples like *Smotri, samolët poletel* 'Look, a plane is flying'. I show that, in fact, this use, though characteristic of a small group of motion predicates, is a part of the whole class of uses that express the partial access of the speaker to the situation ('partial observation'): in this particular case, the speaker can only observe the middle phase of the situation, while the initial phase remains invisible. The use of perfective aspect denotes the fact that the observation period is short and represents the starting point of the observation as if it were the starting point of the situation itself. Similar phenomena are found both in the class of resultative constructions and in other construction classes: the speaker cannot observe a part of the situation and describes only the observable phase as if it was the only one. At the same time, the class of the partial observation readings is heterogeneous. First, some of them allow temporal clauses with the simultaneity meaning and other contexts, characteristic of progressive constructions, while others are incompatible with those mechanisms. Second, only some of those constructions are lexically productive, while others, like, for instance, our progressive-like constructions, characterize a small group of lexemes. I conclude that our use does not, in fact, demonstrate combinatorial properties, characteristic of progressives, thus, no nonstandard progressive past tense use is observed. The meaning of a past event is retained in the construction under analysis, and what is nonstandard is the fact that the speaker describes in the 'online' mode those events that have just taken place before the speech act (this use can be called 'immediate comment').

Keywords

perfective aspect, progressive, resultative, argument structure, evidential meanings, observation period, adverbials, nonstandard uses of tense and aspect

Introduction

Russian is known to lack a specialized progressive form: the same present tense form is used in progressive, habitual, and iterative contexts. However, some prefixed perfective verbs can have a non-canonical interpretation, typical for tense-aspect forms in richer systems. In my article, I describe a special use of the form with the prefix *po-* in the interpretation that I will label as the ‘progressive interpretation’. This use is illustrated by examples like *Smotri, samolët poletel* ‘Look, a plane is flying’, where the speaker uses a perfective verb *poletel* to describe the movement of the plane (s)he is observing.

The main aims of the study are 1) to analyze the mechanism lying behind the progressivelike interpretation of perfective past forms and 2) to show that similar mechanisms explain other nonstandard uses of grammatical forms.

The rest of the article is organized in the following way. In Section 1, I describe and analyze the progressive use of perfective verbs. Section 2 focuses on the group of resultative uses, which are similar to the progressive use under analysis in that that the situation is only partially observed by the speaker. In Section 3, other types of partial observation effects are analyzed, including evidential uses of tense forms and argument structure changes – the emergence of both phenomena result from the fact that some subevents or some participants are ignored in the description of the situation. Section 4 addresses the combinatorial properties of the progressive and similar uses: it turns out that the progressive use shows no shift in combinatorial properties in comparison to the standard use of the past tense. This leads to a conclusion that is put forward in Section 5: in the use that I called ‘progressive’, past tense forms are, in fact, used in the standard way – what is nonstandard, is the choice of the perfective aspect and the discourse function of this construction: it is used as a commentary on the event that is taking place or has just taken place. In the Conclusions section, the results of the study are briefly described.

Examples are taken from the *Russian National Corpus* (www.ruscorpora.ru) or from internet search queries. For the sake of readability, citations taken from the internet are simply marked as [WEB].

1. Progressive use

1.1. Basic description

Below I show some examples where motion verbs in the perfective past ('moved, began to move') are semantically similar to progressive constructions in some languages ('is moving'): they denote a situation that is taking place at the moment of speech.¹ The group of verbs showing this use includes *poexat'* 'begin to drive', *pojti* 'begin to go', *poplyt'* 'begin to sail', *poletet'* 'begin to fly', and so on.²

All verbs I describe contain the prefix *po-*. In the main use, they have the inceptive meaning, referring to the beginning of the situation.³ For instance, (1) shows the standard inceptive meaning of *pojti* 'begin to go', and (2) the progressive use.

- (1) *Utrom ja pošel v magazin za xlebom.*

'In the morning, I went to the shop to buy some bread.'

- (2) *Kto-to skazal, gljadja v okno: — Legavyj pošel (Legavym zvali direktora školy — Čebotarëva.*

'Someone said looking out the window: "Legavyj is going" (lit. 'went'). "Legavyj" was a nickname of Chebotarëv, the principal of the school.' [S. Dovlatov. Naši (1983)]

The context of the form *pošel* is similar to the usual context of progressive forms: the actual state of things is described, which takes place at the moment of speech: the principal Legavyj is going outside the window, and at this moment the speaker notices him. This context satisfies the definitions of a progressive, given by Chung and Timberlake (1985) (they say that progressive is a "language-particular

¹ See Guillaume (1929) and Jazez (1999) on the relation between progressive and similar meanings of the imperfective domain.

² The very phenomenon of shifted tense use have been addressed in many typological and theoretical studies: for instance, Comrie (1985, p. 16) points at the case observed in Roman letters: "... Roman society, presumably for reasons of politeness, recommended use of the recipient's deictic centre". Another case of the same sort (Comrie, 1985, p. 19) is the use of past in polite requests: "I just wanted to ask you if you could lend me a pound." Comrie (1985, p. 20) also cites the Russian use of past tense like *Ja pošel* 'lit. I came' for immediate future events and the use of past tense for surprise expression in present in Norwegian.

³ As Stojnova (2020) shows, Russian has several prefixal patterns that express the inceptive, the main being is *za-* (*zapat'* 'begin to sing', *zaigrat'* 'begin to play', *zabëgat'* 'begin to run', *zarabotat'* 'begin to work', and so on). However, only *po-* shows the use under analysis.

morphological category that signals that an event is dynamic over the event frame”) or by Bybee & Dahl (1989) (they understand progressive as “indicating the situation is in progress at reference time”) (Bybee & Dahl, 1989, p. 55). The only difference is the fact that the Russian progressive use does not have a specialized morphological marking.

The starting point of the motion is not observed by the speaker. While in non-progressive uses (1) and (6), the beginning of the motion is visible (the speaker who says *ja pošel* in (6), only begins his movement). By contrast, in examples (2) and (3)–(5), the speaker does not see the starting point of the movement of “Lexus”, of the plane or of a worm: what is relevant, is only the beginning of observation time:

- (3) *Mama znakomogo s SŠA, priexav v Biš, na každyj Leksus pal'cem pokazyvala, vo smotri, Leksus poexal.*
 ‘The mother of my friend from the USA, when she visited Bish (Bishkek), pointed by her finger to each ‘Lexus’ car, [saying]: look, a ‘Lexus’ is driving!’
 [WEB]
- (4) *Letat' xotel. Uvidit samoljët, srazu “mam, gljadi, samoljët poletel”.*
 ‘He wanted to fly. If he saw a plane, he immediately said: “Mummy, look, a plane is flying”.’ [WEB]
- (5) *Xoma govorit matuške: — Viš', do čego sgnil pux, vot kakoj červjak popolz!*
 ‘Homa says to his mother: “Look, how rotten is the fluff, what a worm is crawling.”’ [WEB]

Below I will call this feature of the progressive use ‘partial observation reading’. The partial observation reading is understood in the following way:

- (I) **The utterance has a partial observation reading if:**
- (i) **Only some phases of the situation are visible to the speaker and are described in the utterance, while other phases are invisible or irrelevant.**
 - (ii) **The speaker describes the visible phase, as if it was the only phase of the situation and does not mention other phases.**
 - (iii) **Often some elements of the utterance are re-interpreted to point only to the visible and relevant part of the situation.**

Syntactically, the progressive use is also different from the standard one: in the progressive use, the final point can be left unexpressed and even unclear from the context. In standard examples like (1), the final point is typically expressed

or clear from the context. More problematic are uses with the future meaning like (6), where the final point is often unexpressed, but is often recoverable from the context:

- (6) — *Nu, ja pošl, poka, — skazal Andrej, podxvativ svoi udočki.*

‘Now I will go, bye, Andrei said, picking up his rods.’ [E. Kolina. *Dnevnik izmeny* (2011)]

Along with intransitive verbs like *pojti*, *poletet'*, *popolzti*, some causative motion verbs also allow the progressive use (cf. examples (7) and (8)). Interestingly, they are usually used in the construction with zero 3PL subject, (8) being more natural than (9). In other words, not only the initial phase, but also the agent related to this phase are eliminated – see Section 3 about the relation between nonstandard verb form use and argument structure.

- (7) *Uvidel zaxodjaščij na posadku IL-76 i vzjal da i ljapnul: smotri, povezli trupaki UPA.*

‘He saw an IL-76 coming in for a landing, and he just blurted out: look, they are bringing UPA corpses.’

- (8) — *Gljadi, rebjata, vot soglašatelja ponesli!*

‘Look, guys, they are carrying a collaborator there!’ [Krasnyj prazdnik (1924)]

- (9) *Gljadi, rebjata, Ivanov soglašatelja ponēs.*

‘Look, guys, Ivanov is carrying a collaborator there!’

A special context is constituted by interrogative constructions with motion verbs. Of course, they can be taken to be simple uses of past tense.

- (10) — *Kuda popolz?! — ryknula tvar'!*

‘Where are you crawling, – the creature growled.’ (Kalinov most. *Trilogija*).

This type of questions has a special illocutionary force, which facilitates the progressive meaning: the speaker condemns the addressee for his / her action. Contrary to examples like (2)–(5), the initial phase can well be known. In this type of examples, we observe a pragmatic shift, rather than a semantic one: the middle phrase of the situation is central for the question, rather than the starting point of the motion.

In the following two sections, I will try to explain the two prominent features of the progressive use of perfective verbs: namely, (i) the use of the perfective aspect and (ii) the fact that it is precisely the group of motion verbs that allows this type of use.

1.2. Why perfective aspect?

The first feature of progressive uses to be explained is the choice of perfective aspect. The progressive is mostly associated with the imperfective aspect, and, in turn, the perfective aspect is taken to denote a finished situation.

The function of perfective in the progressive use under analysis is not the same as in resultative readings of examples like (11). The resultative reading of (11) ('He has fallen asleep and is sleeping now') results from the metonymy "Finished situation – result of the situation" (see also Batiukova & al., 2012): the two phases are adjacent in time, which facilitates the semantic shift.

(11) *On tol'ko čto **zasnul**.*

'He has just fallen asleep (and is sleeping now).'

This metonymy is hardly relevant for the progressive use under analysis, which does not result from the reinterpretation of more standard perfective readings – in examples like (2)–(6), the perfective reading of starting point is hardly available at all. Still, the use of perfective aspect – and, in particular, a perfective verb with the inceptive meaning – is explicable. It seems that this aspectual type is used here to express (i) the fact that the observation period (the time when the speaker can observe the situation) is short and (ii) started right before the utterance. In other words, (2) means 'I have just started to see that the director is going there.' Two other variants that could be used instead of the inceptive verb are semantically different: the standard present tense version (12) does not express any of those two components, while the variant with a non-inceptive perfective verb (13) expresses the shortness of the observation time, but does not emphasize the fact that the observation has just started.

(12) *Smotri, samolët **letit**.*

'Look a plane is flying.'

(13) *Smotri, samolët **proletel**.*

'Look, a plane flew past.'

As will be shown below, the fact that the observation period does not cover the whole situation is relevant in other cases, addressed in Section 2 below.

1.3. Why motion verbs?

The fact that the special progressive use is only or mainly compatible with motion verbs is not accidental either. It has to do with semantics of directed motion, in particular, the fact that, as noted by Dowty (1991) and Krifka (1998) the path argument of motion verbs is an incremental theme. As Rothstein puts it (2001, p. 139), the path is “used up ‘bit by bit’ as the event denoted by the verb progresses”: in every following moment the path covered by the motion is longer than in the previous one. This means that the observation of motion often (though not always) begins when the moving object becomes visible to the speaker and ends when the object is out of sight: for instance, in (2), someone said *Legavylj pošel* at the moment when he could see the director. Since the object has just started to be visible, the speaker can represent the beginning of the observation period as if the motion has just started, and this is why the inceptive prefixed verb is used.

This is not the case, for instance, with singing: the person who sings does not move to the zone of speaker’s visibility, and the moment when the speaker begins to hear the singer cannot be represented as the beginning of singing. This is why the majority of inceptives, including a large group of *za*-verbs *zapel’* ‘begin to sing’, *zaigrat’* ‘begin to play’, *zaplakat’* ‘begin to sing’, *zaprjagat’* ‘begin to jump’, *zabégat’* ‘begin to run in different directions’, and so on) do not have this reading. Examples like (14) are not used in the progressive reading:

(14) #*Petja zapel.*

Intended: ‘Petja is singing.’⁴

The perfective use under analysis are somewhat close to those addressed by Fedotov (2017, 2020) and Fedotov & Čujkova (2013), who use the notion ‘okno nabludenija’ (‘observation window’), which Plungjan (2000, 2011, 2016) introduces as a Russian analogue to Klein’s (1994) ‘topic time’. Fedotov describes it as “time interval that embraces the relevant part of situation time – the interval in which the speaker is sure” (Fedotov, 2020, p. 467). In other words, each

⁴ They are, of course, possible in the standard inceptive reading: the speaker can say (14) if (s)he perceives the moment when Petja begins to sing. This reading can be extended to the process phase (example (14) can mean ‘Petja began to sing and is singing now’). However, the same sentence is not typically used if the speaker did not hear the beginning of the song.

use of tense and aspect presupposes that only a particular interval is relevant for the description of the situation in language and behavior of various elements. However, the progressive use under analysis has some special features that are not covered by the standard notion of topic time. First, in my data, the relevant part of the situation is the only one visible to the speaker. Second, in the case under analysis, this partial observation results in a nonstandard aspect use.

Although the progressive use is only compatible with the small class of motion verbs, there are other nonstandard tense and aspect uses that can be explained using the notion of partial observation. Some of them belong to the group of resultative meanings, which I will discuss in Section 2. Others belong to another class of phenomena, where the partial observation reading adds evidential components to the sentence meaning or changes the properties of syntactic arguments. Cases of this sort will be analyzed in Section 3.

2. Other uses of tenses, related to partial observation: resultative uses and subordinate tense forms

Russian possesses a family of resultative constructions that share the partial observation reading with the progressive use. Here belong resultative uses, where a part of the situation is unavailable for observation, and some uses of tense forms in subordinate clauses, where the whole situation is observed, but only a part of the situation is focused.

2.1. Resultative uses

The main class of tense uses to show the partial observation reading are resultative uses of perfective past. In this case, neither the beginning of the situation, nor its final point can be observed. Only the resulting state of affairs ('the shop works', (15), 'the guy is sleeping', (16) and (17), 'someone thinks', (18)) is visible to the subject or observer.

- (15) *Odnadždy on uexal otdyxat' na leto, a kogda vernulsja, uvidel, čto v ego dome otkryli restoran.*

'Once he went for summer holidays, and when he returned, he saw that they had opened a restaurant at his home.' [Ė. Radzinskij. Jusupovskaja noč'].

- (16) [...] *obščij tualet s dušem, v ktorom **zasnul** kakoj-to mužik.*
 ‘Shared toilet and shower, where some bloke fell asleep.’ (and is sleeping now). [WEB]
- (17) *Nebol’šoj skverik, vokrug doma, a na skamejke **prikornul** paren’. Tëmnye volosy vzloxmačeny, budto korova liznula, rot priotkryt, kažetsja, čto sejčas zaxrapit.*
 ‘A small public garden, houses all around, and on a bench a guy is lying down. His dark hair was dishevelled, as if a cow had licked it, his mouth was open, and he looked like he was about to snore.’ [WEB]
- (18) *Ty čego **zadumalsja**? Pojdem spirtu vmažem!*
 ‘What are you thinking about? Let’s go get some alcohol...’ [M. Žukov. Tret’ja stepen’, 2012]

Resultative uses of perfective past forms in (15)–(18) differ from the progressive use under analysis (see also the discussion of example (11) above). First, in the progressive use, as I showed in Sections 1.2 and 1.3, the speaker represents the beginning of the observation as the beginning of the situation: for instance, *pošël* in (2) means ‘I began to see the motion (of the director)’, rather than ‘The director began to move’. In resultative uses (15)–(18), the speaker does not represent the beginning of the observation as the beginning of the situation: for example, in (16), *zasnul* refers to the beginning of sleeping as the inceptive phase – the thing is that the speaker cannot observe this initial phase. Second, in the resultative use of perfective past, the situation is not represented as a short one. By contrast, states and processes, illustrated in (15)–(18), are taken to be long and can be modified by adverbials like *nadolgo* ‘for a long time’. Third, in examples (15)–(18), but not in the progressive use, the resultative component constitutes a separate phase that can be distinguished from the inceptive / process one: for instance, if someone fell asleep, it is possible to distinguish the dynamic phase (falling asleep) from the statal one (sleeping). In progressive uses like (2), the dynamic phase is described, and the situation includes no resultant state. This is why resultatives (19), but not progressive uses (20) are compatible with contexts, where the resultative and the process phase are opposed: in the resultative use, the resulting state (sleeping) can be separated from the process phase (falling asleep). By contrast, in progressive uses, like (20), the process phase (driving) is inseparable from the initial phase (starting), and the sentence sounds unnatural.

- (19) *On spit, no dolgo ne mog zasnut'.*

'He is sleeping, but he could not fall asleep for a long time.'

- (20) *#On edet na rabotu, no dolgo ne mog poexat'.*

Intended: 'He is going (*lit.* 'driving') to work, but he could not leave for a long time.'

One verb, namely, *zalatit'* 'begin to say something intensely' (see example (21)), allows only the resultative use of perfective past. The past tense form *zalatil* can neither be used in the canonical inceptive meaning, nor replaced with the present tense imperfective form *ladit'*, since this verb does not have the use under analysis. As shown in (22), the resultative use is also central for intensive verbs like *razlečsja* 'lie down, sprawl out'. In those examples, the speaker emphasizes the resultant state, namely, 'the addressee says something' and 'the addressee is lying / sitting' respectively.

- (21) *Ty čego zalatil: ipostasi, ipostasi ?*

'Why do you repeat: hypostases, hypostases?'

- (22) *A poka edet po pyl'noj doroge telega s kopnoj sena, na kotoroj razleglis' troe ustalyx mal'čiček.*

'Meanwhile, a cart with a stack of hay is moving along a dusty road, on which three tired boys are sprawled out.' [WEB]

Motion verbs and semantically similar predicates are also compatible with resultative uses (23). In contrast to progressive uses, the perfective form refers to the initial phase ('the subject hid'), although the speaker only observes the resultant state ('the subject is in a secret spot').

- (23) — *Vot on gde sprjatsja, — vzvizgnul nevdaleke golos Zinki Fokinoj.*

“‘That’s where he is hiding,” Zinka Fokina’s voice shrieked closely to me.’ [Valerij Medvedev. Barankin, bud’ čelovekom! (1957)]

The resultative uses do not form a homogenous class. In particular, uses like (15) are most similar to the standard use of the past tense: although the speaker observes only the resultant state, (s)he makes an assumption about the previous, process phase – the sentence means 'The restaurant was opened and worked by the time he returned', not just 'The restaurant was worked by the time he returned' (the precise time when the shop was opened is marked in (15)).

By contrast, sentences (16) and (17) (and, most probably, (18)) mean just ‘The guy is sleeping’, ‘Why did you start thinking?’. This difference is confirmed by the fact that the temporal adverbs like *nedavno* ‘recently’ are more natural in contexts like (15) than in (16)–(18).

- (24) *On uvidel magazin, kotoryj nedavno otkryli.*

‘He saw a shop that was recently opened.’

- (25) *#On uvidel parnja, kotoryj nedavno prikornul na trave.*

Intended: ‘He saw a guy that has recently fallen asleep on the grass.’

- (26) *#On uvidel kollegu, kotoryj o čem-to nedavno zadumalsja.*

‘He saw a colleague that has recently started to think about something.’

This is why I will further refer to examples like (15) and (24) as ‘resultative contexts combining the resultative and the process phase’, while (16)–(18) and (25)–(26) will be described as simply ‘resultative uses’. Combinatorial properties of these construction types, compared to the properties of the progressive use, will be described in Section 4.

2.2. Special use of present tense in subordinate clauses

Some other uses of tense forms are also close to partial observation contexts. The difference is that one of the parts of the situation is foregrounded, and the other parts are backgrounded, though (contrary to the progressive use) all parts can be observed.

The first of those uses is the use of present tense in some subordinate clauses marked with the simultaneity marker *poka* ‘while’. Alongside the tense marking strategy, typical for adjunct clauses (28), a nonstandard strategy is possible (27).

- (27) *Reb jat, ja tut, poka sižu doma, načal rabotat' na frilanse.*

‘Guys, I started to work as a freelancer, while I am at home.’ [WEB]

- (28) *Poka ja pisal sta'ju, mne pozvonila žena.*

‘While I was writing an article, my wife called.’

In (28), both events – ‘I was writing an article’ and ‘my wife called’ – take place simultaneously and precede the speech act. As in most Russian adjunct

clauses, both verbs are marked for absolute past. By contrast, in (27), the event in the subordinate clause ('I am at home') covers a long time span that includes both the time when the speaker began to work and the speech act time. However, the speech act time tends to be more important for the speaker than any moment in the past. This is why the present tense is chosen, though this present tense marking violates the general tendency saying that if the adjunct clause denotes an event simultaneous to the main one, tense marking is the same in both clauses.

There also exists an opposite case, when the moment in past is more relevant for the speaker than the moment in present (the speech act time). This is the case with some generic contexts that describe inherent properties of the world (including real, such as 'The Sun is hot', and fantastic ones, such as 'The Sun is a glowing rock'). If this type of statement is made by a person in past, past tense marking can be used, as in (29). By the use of past tense because the speaker emphasizes the fact that the opinion (s)he describes belongs to another person in past, although Anaxagoras made his statement about a general property of the Sun.

(29) *Solnce, po ego mneniju, bylo raskalënnoj skaloj.*

'In his [Anaxagoras] opinion, the Sun was a glowing rock.' [WEB]

3. Nontemporal effects of partial observation

In the previous sections, I described the effect of partial observation reading on tense form selection. Importantly, sometimes the partial observation feature affects other features of the sentence, not reducible to tense and aspect use: for instance, it results in evidential readings or argument structure changes. These types of cases are described in the next two sections.

3.1. Past tense in memory evidence

One of the cases under analysis is what can be called 'past of memory evidence' (Aikhenvald (2004) lists memory contexts as a type of evidential contexts). The use of the past tense denotes the fact that the speaker does not remember exactly the present state of affairs but describes his / her hypothesis concerning

the present state by referring to the previous state of affairs.⁵ In fact, the state may persist at the moment of speech, but the speaker can only understand the current state through the memories about the previous one, and this is why the past tense form is used:

(30) — *Aktivirovannyj ugol' est'?* — *Vrode **byl** gde-to, — neuverenno skazala Rimma Alekseevna. — Xorošo. Rastolkite neskol'ko ložek i vysyp'te v vodu.*

‘Do you have any activated carbon? – It seems as if it was somewhere. – Well, grind a few spoonfuls and pour into the water.’ [V. Valeeva. *Skoraja pomošč'* (2002)]

(31) [...] *navigator v telephone (da i maršrut znakom), a pro compressor vrode **ležal** v багаžнике, nado proverit'!!*

‘the navigator is in the telephone, I know the path. About the compressor, it was in the boot, I have to check it!’ [WEB]

Many examples are compatible with two readings, a standard past tense reading and a memory evidence reading.

(32) — *Ty menja pomniš'?* — *Ty vrode **rabotal** na zavode v Xevrone.*

‘Do you remember me? – It seems that you worked at the factory in Hevron.’

(i) ‘Previously you worked at the factory.’

(ii) ‘It seems to me that you work at the factory.’

The past of memory evidence, as well as our ‘progressive use’, is related to the partial observation reading. The use of the past tense results from the fact that the speaker is not fully aware of the current state of affairs (as if it was invisible to the speaker) and uses the past tense form, referring to the situation in past.

3.2. Aspectual change meets argument structure change: passive participles and similar uses

Finally, the partial observation reading sometimes influences argument structure of the verb and strategies of argument expression: this can be demonstrated on

⁵ This evidential use of past does not seem to be described in any linguistic studies, contrary to some other means of expressing evidential uses in Russian (see, for instance, Xrakovskij (2005), Padučeva (2017), and many others).

some uses of passive participles. As Saj (2011) shows, the morphologically passive forms like *odet(yj)* ‘dressed’ or *otkryt(yj)* ‘opened’ are semantically resultative, but not necessarily passive: it may be the case that a person dressed him/herself, and the door opened by itself. Resultative participles share the partial observation reading with our data: only the resultative phase is visible to the speaker.

- (33) *Tam, za dver'ju, sidit' čelovek v furažke, **odetyj** v poluformennye sero-zelënye tona.*

‘Beyond the door, someone is sitting, dressed in semiformal gray-green tones.’ [Sergej Jursky. Bumažnik Goffmana (1993)]

- (34) *Dver' za spinoj moej nastež' **otkryta**, v sosednej komnate temno.*

‘The door behind my back is wide open, it is dark in the neighbouring room.’ [N. Iljina. Dorogi i sud’by (1957–1985)]

- (35) *Fortočka v kuxne **byla raspaxnuta** ot vetra, v kvartire stojal xolod.*

‘The window was opened in the wind, and it was cold in the flat.’ [WEB]

Examples (33)–(35) are similar to examples (15)–(18), where finite past tense forms have a resultative meaning. The difference is that in the participial construction, the partial observation affects the argument structure: as the agent only takes part in the dynamic phase, the partial observation semantics makes the agent / causer irrelevant for the speaker. Note that the argument structure changes in the resultative uses of participles also affect the strategy of agent / causer coding: cf. examples (36)–(38), compared to the resultative construction in (35):

- (36) *Fortočka srazu **raspaxnulas'** ot vetra, i stalo xolodno.*

‘The window immediately opened in the wind and it was cold.’

- (37) **Fortočka srazu **byla raspaxnuta** ot vetra, i stalo xolodno.*

Intended: ‘The window was immediately opened in the wind and it was cold.’

- (38) *Fortočka srazu **byla raspaxnuta** vetrom, i stalo xolodno.*

‘The window was immediately opened in the wind and it was cold.’

The construction with the preposition *ot* is only compatible with nonagentive causers. However, it mainly occurs with *-sja*-forms. Passive participles are

combined with *ot* only in the resultative meaning – in the dynamic process meaning, their combination is ungrammatical (37): *ot* is only compatible with reflexive forms, as in (36), while participial passive constructions allow only the standard instrumental agent marking, as in (38).

There are equally other cases, in which the partial observation reading results in the agent elimination. In constructions with active voice forms, the agent cannot be completely eliminated – however, the subject position can be occupied by a nonagentive and inanimate participant (a natural force or another type of inagentive cause). The canonical agent, which, of course, participated in the process phase of the situation (in (39), there was someone who covered the toilet seats with down (eider feathers)), can be left unexpressed – which is impossible in the canonical process reading (40).

(39) *Rosli tam rozy v vazax, V vazonax rozan ros, Siden'ja ž unitazov Pokryl gagačij vors.*

‘Roses grew there in vases, and the toilet seats were covered with eider feathers.’ (S. Sokolov. *Škola dlja durakov*).

(40) ??*Kogda vors pokryl sidenja unitazov, my pozvali vsex vnutr'.*

‘When the eiderdown covered the toilet seats, we called everyone inside.’

The same strategy is demonstrated by reflexive verbs, such as *smenit'sja* ‘be replaced’ in (41) or *skryt'sja* ‘be covered, hide’ in (42): only the resultative phase is observed by the speaker, while the dynamic phase (and, correspondingly, the agent that takes part in this phase) is invisible and / or irrelevant: this is why the change in the car makeup in (41) is represented as spontaneous.

(41) *Sidenja pokryla koža dvojnoj vydelki “parra”, klassičeskaja rukojatka AKPP smenilas' na elektronnyj “gribok”-kačalku...*

‘The seats are covered with double nappa leather, the classic automatic transmission knob has been replaced by an electronic ‘mushroom’ rocking arm’ [WEB]⁶

⁶ I do not consider here the stative uses of imperfective verbs. They, of course, also go back to agentive dynamic uses, but currently the stative use can be described as marking the stative phase, where the agent does not play any role. *Na nëm po-prežnemu byl tot samyj sinij gabardinovyj plašč, tol'ko teper' ego pokryvala grjaz'.* ‘He wore the same blue raincoat of gabardine, but now it was covered by dirt.’ [A. Čakovskij. *Blokada* (1968)]

- (42) *Moi zapjast'ja, šeja i pal'cy skrylis' pod urkašenijami, a golovu pokryl platok.*
 'My wrists, neck and fingers were hidden under jewellery and my head was covered by a shawl.' [WEB]

Similarly, the partial observation reading affects the meaning of another, colloquial, group of reflexive verbs. In principle, perfective *-sja*-verbs normally do not have a passive meaning. However, in the colloquial Russian, perfective *sja*-verbs allow for a resultative passive meaning, illustrated in like (43) and (44). As in the previous cases, verb forms *postavjatsja* '(they) will be put in' (43) and *otremontiruetsja* '(it) will be repaired' (44) have a resultative meaning, the role of agent in the process phase being ignored.

- (43) [...] *kogda postavjatsja porogi, štany budut večno grjaznye.*
 'When the sills are put in, the trousers will always be dirty.' [WEB]
- (44) *Kogda otremontiruetsja banja?*
 'When will the bathhouse be repaired?' [WEB]⁷

3.3. Partial observation and evidentiality

Our uses are in a sense close to another phenomenon, also associated with partial observation: namely, the inferential evidential uses. Contrary to other phenomena under analysis, the inferential meaning lacks a grammatical marking: in (45), the evidential meaning is conveyed by the parenthetical *poxože* 'likely, seemingly'.

- (45) *Prjamo ot dverej kak by vymytaja dorožka idet; poxože, vytirali tam, gde xodili, it takoe vpečatlenie, sovsem nedavno.*
 'Right from the door, there is a recently washed lane. It seems that one cleaned the place where everyone walked, and, presumably, not long ago.'
 [E. & V. Gordeevy. *Ne vse my umrëm* (2002)]

While in the reportative evidential, there is a speech act that the speaker is relying on, this is not the case with the inferential meaning like that in (45). Here the speaker observes one of phases of the situation (the resultant state 'the lane

⁷ It is very probable that a similar mechanism is also relevant for anticausative uses of reflexive verbs. Padučeva (2001) argues that the agent is not necessarily absent in the anticausative use: it can just be irrelevant for the speaker.

is clean') and can only hypothesize on the other one ('someone cleaned the lanes not long ago').

The difference is that the link between the phases is more trivial in the resultative and progressive use than in evidential constructions like (45). For instance, the fact that the lane is clean does not necessarily imply that it has recently been washed (for instance, the reason can be that nobody has lived in the house for last several months and walked along the lanes). By contrast, the resultative phase, for instance, the fact that the guy is sleeping, in examples like (17), always results from the fact that he fell asleep, and the fact that the restaurant is open (15) always implies that someone has opened it – this is why evidential markers are incompatible with contexts like (46) and (47): example (46) looks strange if used in the situation when the speaker stands in front of the shop, and (47) is incompatible with the situation when the speaker sees a sleeping guy. The logical relation between two parts of the situation is too trivial here to use an evidential marker.

(46) #*Zdes', poxože, otkryli magazin.*

'It seems that a shop was opened here.'

(47) #*Poxože, paren' zasnul.*

'It seems that the guy fell asleep.'

4. Combinatorial properties: shifts in the use of adverbs, embedded clauses and negation

As mentioned before, progressive and resultative readings presuppose a non-standard form-meaning correspondence: their interpretation is not typical for tense-aspect forms that are used. A question arises whether these nonstandard uses also demonstrate nonstandard combinatorial properties: whether temporal adjuncts, temporal embedded clauses, and other elements have a shifted interpretation in the nonstandard uses, or their interpretation does not change. In other words, if only a part of the situation is available for observation, is it true that temporal modifiers refer to the observed part of the situation? In particular, I will check combinations of nonstandard tense uses with temporal adverbs, negation and embedded clauses marking simultaneity, as well as the reference of the tense form.

Only the combinatorial properties of resultative and progressive uses of tense forms are checked below, while other uses addressed in Sections 2 and 3 are not analyzed here.

4.1. Adverbials with the meaning of present

The first diagnostic uses adverbials with the present time meaning (e.g., *segodnja* ‘today’ or *v dannyj moment* ‘at present moment’). Normally, *segodnja* can mark the state relevant for today, and *v dannyj moment* refers to the situation taking place at the moment of speech. In some nonstandard uses, the time interval that these adverbials denote shifts from the process phase to the resultant one. With **resultatives in the proper sense** like (48)–(52), those adverbials can refer to the time when the resultant state takes place (this special property is also mentioned by Padučeva (1996/2011: 271) and Fedotov (2020)).

- (48) *Novejšaja sistema otsleživaet povedenie učnikov, sčityvaet vyraženiya ix lic i opredeljaet, kto **segodnja ne podgotovilsja k uroku**.*

‘The new system traces the behavior of pupils their face expression and finds out who is not ready for the lesson today.’ [WEB]

- (49) *Prosto ja za natural'nost' i estestvennost' otnošenij, ne vižu bedy sdelat' to, čto **v dannyj moment nazrelo**, tem bolee, čto vse ravno prorvët.*

‘In fact, everything is, of course, very individual. I simply prefer relationships to be natural, and I see no trouble to do what became relevant to the moment, especially because it will burst anyway.’ [Naši deti: Podrostki (2004)]

- (50) *Oni s moim bratom segodnja oformili otnošenija. **V dannyj moment otbyli** v svadebnoe putešestvie. P'jut šampanskoe v poezde.* [N. Nesterova. Zefir v šokolade (2013)]

‘She and my brother married (lit. ‘formalized their relations’) today. At the moment, they left to their honeymoon. They are drinking champagne in the train.’

- (51) ***V dannyj moment poexal** v očerednuju komandirovku, po finansam sovsem ne spravljajus'.*

‘At the moment, I left to another business trip – I’m really struggling financially.’ [WEB]

(52) *V **dannyj moment zasnul** na rukax, no dërgaet nožkoj periodičeski ili ručkoj, budto pugaetsja čego-to.*

‘At the moment, he fell asleep on her mother’s hands, but periodically he moves his leg or arm, as if he was afraid of something.’ [WEB]

Resultative contexts **combining the resultative and the process phase** behave in a more complicated way: some present tense adverbials like *vannyj moment* ‘currently’ can refer to the resultative state (53). However, most adverbials, for instance, *segodnja* ‘today’ or *v ètom mèsjace* ‘in this month’ refer to the process state, rather than to the resultative one: (54) means that the shop was opened in this month / today, and not that in this month or today, the shop is already opened.

(53) *Tam v **dannyj moment otkryli** magazin.*

‘Currently / today a shop was opened there.’

(54) *Tam #v ètom mesjace / #segodnja otkryli magazin.*

Intended: ‘In this month / today, a shop is opened there.’

With past in the progressive interpretation, the test yields a negative result: in examples like (55) or (56), the adverbial modifiers *vannyj moment* or *sejčas* either look strange or do not have the meaning ‘the situation takes place at the moment of speech’. The adverbial modifier *tol’ko čto* ‘just’ with the precedence meaning sounds better in these contexts.

(55) *Mne pokazalos’ ili tam **tol’ko čto poplyl glaz.***

‘It seemed to me or I have just seen an eye sailing on the water.’ [WEB] (*vannyj moment* would be semantically strange).

(56) *#O, tam v **dannyj moment direktor pošël.***

Intended: ‘Look, at this moment /

4.2. Negation

Another diagnostic is the interpretation of verbal negation: if the tense use is shifted, we could expect that the observed phase is negated. However, it turns out that none of the tense uses that we have discussed allows the resultative phase to be negated. Constructions with the **progressive** (57) and **resultative** meanings (58)

are incompatible with negation – examples (57) and (58) are grammatically correct, but have the standard meaning of precedence ('The plane did not started'; 'The guy did not fall asleep'). Resultative uses **combining the resultative and the process phase** (59) behave in the same way: they can be negated, but it seems that the speaker hypothesizes on the process phase ('The addressee did not put on her hat'), rather than describes the resultative phase ('The addressee does not wear a hat').

(57) *#Samolët ne poletel.*

Intended: 'A plane is not flying.'

(58) *Paren' #ne prikornul / #ne zasnul na lavke.*

Intended: 'At the moment, the guy is not sleeping on a bench.'

(59) *#O, ty segodnja šljapu ne nadela!*

'Oh, you did not put your hat on today!'

4.3. Simultaneity clauses with *poka*

More illustrative is the test based on the use of temporal clauses marked with *poka* 'while' – a specialized temporal subordinator that denotes simultaneity. If the shift in the use of tense forms is really strong, it is expected that the construction is compatible with this type of clauses, and that *poka*-clauses state the simultaneity relation with the phase which is observed by the speaker. In fact, **progressive** uses (60) cannot be combined with *poka*-clauses, but this is possible for some resultative constructions – only for **resultatives in the proper sense**, more precisely, for those cases when the result phase is rather short and visible, as in (61) and (62). In those examples, the embedded situation is marked as simultaneous to the resultant state ('He is sleeping'; 'The door is opened'), and not to the process state ('He fell asleep'; 'Someone opened the door'). The nonstandard interpretation of *poka*-clauses is confirmed by the fact that the embedded clauses in (61) and (62) contain a perfective verb, which is not characteristic of simultaneity contexts: here it is possible, because the simultaneity relation is stated between the main situation and the result phase, which is not finished, while the process phase is irrelevant.

(60) *#Poka poletel samolët, my na nego smotreli.*

Intended: 'While a plane was flying we were looking at it.'

- (61) [...] **Poka on zasnul** na solnyske, odin mužik toporikom čast' xvosta emu ottjuknul.

'While he was sleeping on the sun, one man cut him out a part of the tale with an axe.' [WEB]

- (62) ... **poka dver' otkryli**, my bistro nyrnuli nazad, skazav, čto zamerzli.

'While the door was opened (lit. 'While they opened the door'), we fast ran in, saying that we were cold.' [WEB]

Constructions **combining the resultative and the process phase** (63)–(65) are usually incompatible with *poka*-clauses: both the dynamic state and the result are focused, and the constructions are unacceptable. Example (64) is grammatically possible, but in another meaning: 'Until she put her hat on.'

- (63) ***Poka v dome otkryli magazin / poka ona nadela pročital knigu**, on byl v otpuske.

Intended: 'While they opened a shop in the house / while he has read a book, he was on his holidays.'

- (64) #**Poka ona ne nadela šljapu**, golovu napeklo solnce.

Intended: 'While she was not wearing her hat, she has got sunstroke.'

- (65) ***Poka ona nadela pal'to**, ej bylo žarko.

Intended: 'While she was wearing her coat, she was hot.'

4.4. The use of participle forms

Although the constructions under analysis share the feature of partial observation, they differ in the degree of stability of the nonstandard tense use. To check the properties of the tense value, I use contexts where the nonstandard form is embedded under another, matrix, verb, namely, *videt'* 'see'. The nonstandard use of the finite past tense form is changed to the past participle. If this substitution is possible and does not change the meaning of the past tense form, I conclude that the nonstandard tense use is rather stable and is retained in nonfinite forms. Otherwise, the past tense form is used dynamically. It turns out that the past participle sounds strange in the **progressive use** (66), while in **constructions combining the resultative and the process phase** like *On ne nadel pal'to* 'He did not put his coat on' the participle does not express the resultative / stative

meaning (67)–(68). By contrast, **resultatives** in the proper sense (69) retain the resultative meaning when the finite form is changed to the participle:

(66) ??*Ja uvidel pošedšego direktora.*

Intended: ‘I saw the director who was going.’

(67) *Ja uvidel ženščinu, ne nadevšuju šljapu.*

(i) #‘I saw a woman who did not wear her hat.’

(ii) ‘I saw a woman who did not put on her hat (before).’

(68) *Ja uvidel otkrytyj na pervom etaže magazin.*

‘I saw a shop that was opened on the first floor.’⁸

(69) *Ja uvidel parnja, prikornuvšego na trave.*

‘I saw a guy who was sleeping / fell asleep on the grass.’

Table 1 shows combinatorial properties of progressive and other constructions. The table shows that, contrary to other nonstandard uses, the progressive use of past tense forms is incompatible with any nonstandard interpretations addressed in this section. This behavior of the progressive use will be explained in Section 5 below.

Table 1.

Combinatorial properties of progressive and resultative constructions

	Present tense adverbials	Negation	Poka-clauses	Participles
Progressive	-	-	-	-
Resultative in the proper sense	+	-	+	+
Resultative + process phase	+-	-	-	-

⁸ Example (65) differs from the rest of sentences, because passive past participles show a clear tendency to the resultative use (see 3.2 for details).

5. Progressive is not progressive: discourse characteristics

As mentioned before, Table 1 shows that no contextual diagnosis confirm the progressive nature of uses like (2)–(8). Negation, adverbials with the present time meaning and embedded clauses with the meaning of simultaneity are either incompatible with the progressive use like *Legavýj pošel* or change the meaning of relevant examples to the standard precedence reading.

In this case, a question arises whether our uses are really progressive. It turns out that a special discourse property of the constructions under analysis can give an answer to this question. The progressive use of perfective is incompatible with the indirect speech context (for instance, it is normally not used in subordinate clauses) and with a description in the narrative mode, but only compatible with the dialogue mode. Example (70) seems to be strange in the progressive reading:

(70) #*On mne skazal, čto direktor **poexal** / samolět **poletel**.*

‘He told me that the director was going.’

The other constructions addressed here do not share this property. For instance, resultative constructions are compatible with indirect speech, so they can be used in subordinate clause contexts (71):

(71) *On mne skazal, čto na trave **zadremal** paren’.*

‘He told me that a guy had fallen asleep on the grass / was sleeping on the grass.’

Provided that the perfective aspect is used to denote the dynamic nature and the short duration of the situation, I should conclude that the ‘progressive’ use is not, in fact, progressive. It rather constitutes a special type of construction that describes events that took place before the speaker’s eyes (this mode of description can thus be called ‘immediate comment’). This construction type is different from the neutral type of event description in some respects, including the word order. The immediate comment construction prefers the SOV word order with transitive verbs (72) and SV with intransitive verbs (examples (2)–(5) above), while the neutral description often uses SVO (with transitive verbs, cf. (73)) and allows VS (with intransitive verbs).

Immediate comment:

(72) *Smotri, menty xuligana povezli.*

‘Look, cops are driving a bully.’

Neutral description:

(73) *Menty povezli kuda-to xuligana / on videl, kak menty povezli kuda-to xuligana.*

‘Cops were driving a bully somewhere / He saw cops driving a bully somewhere.’

Another possible component of semantics of this use is a mirative meaning (‘the situation was not expected to occur and surprises the speaker’). For instance, in examples (3) and (4), repeated here as (74) and (75), the situation is apparently unexpected for the speaker. An unexpected event is often foregrounded in comparison to other events, which makes the use of perfective more natural, while a standard situation like ‘A car is moving there’ would rather be described with a present tense form. However, the mirative component is not obligatory, for instance, in (2), repeated as (76), it is hardly unexpected that the director appeared on the schoolyard⁹.

(74) *Mama znakomogo s SŠA, priexav v Biš, na každyj Leksus pal'cem pokazyvala, vo smotri, Leksus poexal.*

‘The mother of my friend from the USA, when she visited Bish (Bishkek), pointed by her finger to each ‘Lexus’ car, [saying]: look, there is ‘Lexus’ driving!’ [WEB]

(75) *Letat' xotel. Uvidit samolët, srazu “mam, gljadi, samolët poletel”.*

‘He wanted to fly. If he saw a plane, he immediately said: “Mummy, look, a plane is flying”’. [WEB]

(76) *Kto-to skazal, gljadja v okno: – Legavyj pošël (Legavym zvali direktora školy – Čebotarëva).*

‘Someone said looking to the window: “Legavyj is going” (lit. ‘went’). “Legavyj” was a nickname of Chebotarev, the principal of the school.’ [S. Dovlatov. Naši (1983)]

⁹ It should be noted that the semantics of immediately preceding situation is more important than the mirative meaning. As Weiss (2022) shows, some mirative constructions, such as the Russian construction *vzjal i sdelal* ‘unexpectedly did’, ‘took and did’, are compatible with the narrative mode and indirect speech.

This special discourse function of ‘immediate comment’ explains several important features of our construction. Since it marks the immediately preceding situation, the fact that it is compatible with the perfective aspect is very natural: an observer who looks at the situation and comments on them is not prepared in advance, this is why (s)he describes the situation as a sequence of short events. The fact that the construction is incompatible with modifiers and negation not only results from the fact that immediate past is described, but also from the discourse properties: the comment on the situation does not typically mention temporal relations between events or a description of events that do not take place. The immediate comment only makes sense if the speaker directly perceives the situation and immediately informs the addressee of the news. Finally, the fact that the arguments typical for motion verbs (path and final point) are often left unexpressed also results from the meanings of immediately preceding situation: the final point is unknown, while the path is clear from the context of communication: the speaker describes the motion (s)he is observing, and usually also sees the motion and understands where it is taking place.

All these facts show that, in fact, the tense use in the construction is standard (the verb is marked for past to denote an event in immediate past). The nonstandard features of the construction under analysis are (1) the discourse function of a comment on an immediately precedent event; (2) the special word order and (3) the use of the perfective aspect and the inceptive verb pattern to mark the starting point of the observation as if it was the starting point of the situation, explained in 1.2 and 1.3., rather than the use of tense.

Conclusions

In my article, I analyzed the functioning of ‘progressive use of perfective past forms’ and compared it to other nonstandard uses of tenses. It turned out that the progressive use is, in fact, different from the canonical progressive in languages like English. Although it really focuses on the situation that is in progress, it adds to the standard present tense use a dynamic component: the perfective past tense form is chosen to mark that the observation time is short and conceptualizes the beginning of the observation interval as the beginning of the situation itself (directed motion). To describe this use, I used the notion of “partial observation reading” – the reading in which the speaker observes only a particular

part of the situation and describes it as if it was the whole situation. This reading is crucial for the progressive use under analysis.

This is why the perfective aspect form is chosen for the progressive meaning, which is normally associated with imperfective verbs (the process phase is relevant and described in detail, while the final and initial states are unknown or irrelevant). This use is a result of metonymy:

Base meaning of perfective	Derived meaning of perfective
The initial stage is short (momentary)	The time of observation is short (momentary)
The situation has just begun	The observation has just begun

I showed that some other uses of Russian tense forms are also based on the partial observation, on the fact that only a part of the situation is visible to the speaker. This class includes resultative uses, marked with perfective verbs (*zasnul* 'fell asleep', *prikornul* 'fell asleep for a short time'). Many of those uses confirm their special status by the ability to combine with adverbials and adverbial constructions, associated with present tense contexts. While our progressive use is colloquial, other resultative and stative uses, addressed above, are frequent in the standard literary register.

The partial observation reading also affects phenomena outside the tense domain: for instance, it results in nonagentive uses of passive and passive-like forms and in the evidential uses of past tenses. In those cases, the irrelevance of some phases of the situation results in the elimination of participants and / or semantic components associated with those phases.

The nonstandard use of forms under analysis is to certain degree associated with nonstandard combinatorial properties: for instance, the semantic change leads to shifts in the use and semantic scope of temporal adverbs and temporal embedded clauses. However, this shift is only partial: for example, no present tense adverbs are compatible with the progressive uses. Similarly, progressive uses are incompatible with negation, while other nonstandard uses are only partially compatible with it. In other words, nonstandard tense form uses do not form a homogenous semantic class and do not behave uniformly with temporal markers.

The fact that the progressive use is incompatible with typical present tense modifiers shows that, in fact, this use is not really progressive. The past tense forms with the prefix *po-* denote an event immediately preceding the speech act and are thus used in the standard way. What is really nonstandard is that uses described here are specialized for the comment on an immediately

preceding event. Another special feature is the use of the inceptive perfective verbs to mark the starting point of the observation as if it was the starting point of the situation.

References

- Aikhenvald, A. (2004). *Evidentiality*. Cambridge University Press, Cambridge.
- Batiukova, O., Bertinetto, P.M., Lenci, A., & Zarcone, A. (2012). Semantic priming study of Russian aspect and resultativity. *Oslo Studies in Language*, 4 (1), *The Russian verb*, 177–206.
- Bybee, J., & Dahl, Ö. (1989). The creation of tense and aspect systems in the languages of the World. *Studies in Language*, 13-1, 51–103.
- Chung, S., & Timberlake, A. (1985). Aspect, tense, mood. In T. Shopen (Ed.), *Language typology and syntactic description. Vol. 3. Grammatical categories and the lexicon. (1st ed.)* (pp. 202–258). Cambridge University Press, Cambridge.
- Comrie, B. (1985). *Tense*. Cambridge University Press, Cambridge.
- Dowty, D.R. (1991). Thematic proto roles and argument selection. *Language*, 67, 547–619.
- Fedotov, M. (2020). Perfekt i okno nabljudenija (dva ekz.). In K. Semënova, A. Kibrik (Red.), *VAProsy jazykoznanija: megasbornik nanostatej* (ss. 467–475). Buki Vedi, Moskva.
- Fedotov, M. (2017). Semantika perfektiva i imperfektiva: sinxronnost' / nesinxronnost' I dinamičeskoe / statičeskoe vremja. [Handout for the talk at the conference *Russkaja grammatika: opisanie, prepodavanie, testirovanie*]. Helsinki University, 7–9 June 2017. <https://www.academia.edu/35586833/>.
- Fedotov, M., & Čujkova, O. (2013). K opredeleniju aspektual'nogo značenija limitative i k voprosu ob osobennostjax "delimitativnoj" derivacii russkogo glagola. In E. Grexova (Red.), *Iz proslogo v buduščee. Sbornik statej i vospominanij k 100-letiju professora Ju. S. Maslova* (ss. 153–203). Saint-Petersbourg University, St-Peterburg.
- Guillaume, G. (1929). *Temps et verbes. Théories des aspects, des modes et des temps*. Honoré Champion, Paris.
- Jayez, J. (1999). Imperfectivity and progressivity: The French imparfait. In T. Matthews & D. Strolovich, *SALT IX Conference proceedings* (pp. 145–162). Cornell University Press.
- Xrakovskij, V. (2005). O sootnošenii èvidencial'nosti i epistemičeskoj modal'nosti. In D. Gerasimov, S. Dmitrenko & al. (Red.), *II Konferencija po tipologii i grammatike dlja molodyx issledovatelej* (ss. 162–183). ILI RAN, Moskva.
- Klein, W. (1994). *Time in language*. Routledge, London.

- Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and Grammar* (pp. 197–235). Kluwer Academic Publishers, Dordrecht.
- Padučeva, E. (1996/2011). *Semantičeskie issledovanija*. Jazyki slavjanskoj kul'tury.
- Padučeva, E. (2001). Kauzativnyj glagol i dekauzativ v russkom jazyke. *Russkij jazyk v naučnom osveščanii*, 1, 52–79.
- Padučeva, E. (2017). *Evidential: non-direct interpretation and projection of the implied attitude holder*. [PowerPoint presentation] https://lexicograph.ruslang.ru/evidentiality_surprise_SLE.pps.
- Plungjan, V. (2000). *Obščaja morfologija*. URSS, Moskva.
- Plungjan, V. (2011). *Vvedenie v grammatičeskiju semantiku: grammatičeskie značeniya i grammatičeskie sistemy jazykov mira*. RGGU, Moskva.
- Plungjan, V. (2016). K tipologii perfekta v jazykax mira: predislovie. In T. Maisak, V. Plungjan & K. Semënova (Red.), *Acta Linguistica Petropolitana. Trudy Instituta lingvističeskix issledovanij RAN. Vypusk 7. Tipologija perfekta* (ss. 7–38). Nauka, Moskva.
- Rothstein, S. (2001). What are incremental themes? *ZAS Papers in linguistics*, 22, 139–157.
- Saj, S. [=Say S.] (2011). Pričastie. *Materials for the project of corpus-oriented description of Russian Grammar*. [Electronic resource]. www.rusgram.ru
- Stojnova, N. (2020). Načinateľnyj prefiks za-. *Materials for the project of corpus-oriented description of Russian Grammar*. [Electronic resource]. www.rusgram.ru
- Weiss, D. (2022). Mirative ‘take and do’ constructions in Russian: The impact of negation. In L. Iomdin, J. Milićević, & A. Polguère, *Lifetime linguistic inspirations: To Igor Melčuk from Colleagues and Friends for his 90th Birthday* (pp. 443–453). Peter Lang, Berlin.



Tilmann Reuther

Universität Klagenfurt
Austria

 <https://orcid.org/0000-0002-4475-9014>

Remarks on Ju. D. Apresjan's lexicographic terminology: *leksema* and *vokabula* from a multilingual perspective

Abstract

This paper discusses the Russian terms *vokabula* and *leksema*, as well as related concepts introduced to lexicography and semantics by a group of Moscow-based researchers led by Jurij D. Apresjan (commonly known as the Moscow Semantic School) and by Igor Mel'čuk (the author of the Meaning ⇔ Text Theory and, alongside Aleksander Zholkovsky, the primary proponent of the Explanatory-Combinatorial Dictionary). The Russian term *leksema* and its equivalents in other languages (Eng. *lexeme*, Fr. *lexème*, Ger. *Lexem*, Sp. *lexema*) are firmly anchored in Apresjan's and Mel'čuk's lexicographic theory and practice and have been widely accepted by the international lexicographic community. In contrast, the Russian term *vokabula* has not even been accepted by Russian lexicographers, let alone the international lexicographic community. This paper explores the reasons behind this discrepancy, which can be attributed to the nature of the lexical parallels in Russian and other languages (Eng. *vocable*, Fr. *vocable*, Ger. *Vokabel*, Sp. *vocablo*). Related issues in this paper include the widely used Russian term *slovarnaja stat'ja* (Eng. *dictionary entry*, Fr. *article de dictionnaire*, Ger. *Wörterbuchartikel*, Sp. *artículo lexicográfico*) and the Russian term *lemma* (Eng. *lemma*, Fr. *lemme*, Ger. *Lemma*, Sp. *lema*), as well as the German term *Stichwort* (Eng. *keyword*), the former in the context of computational linguistics and the latter in the context of general lexicography. As a result, we propose names for Apresjan's and Mel'čuk's term *vokabula* in German (namely, *Stichwort* or *Lemma*) and English (namely, *keyword* or *lemma*).

Keywords

linguistic terminology, lexicography, lexical parallels, Russian, English, French, German, Spanish

Introduction

This paper begins with the definitions of the Russian terms *leksema* and *vokabula*, as well as related terms. These were introduced to Russian lexicography and semantics in the 1970s by a group of Moscow-based researchers led by Jurij D. Apresjan (commonly known as the Moscow Semantic School) and by Igor Mel'čuk (the author of the Meaning $\leftrightarrow$ Text Theory and, together with Aleksander Zholkovsky, the main proponent of the Explanatory-Combinatorial Dictionary – see Mel'čuk & Reuther, 2020). The two terms were not equally accepted by the scientific community. The Russian term *leksema* and its equivalents in other languages (Eng. *lexeme*, Fr. *lexème*, Ger. *Lexem*, Sp. *lexema*) – firmly anchored in Apresjan's and Mel'čuk's lexicographic theory and practice – have been widely accepted by the international lexicographic community, albeit not always with the same clarity of definition or consistency of usage found in the work of Apresjan and Mel'čuk. They are often used as quasi-synonyms of the term *word*. Not so the Russian term *vokabula*. Anchored in Apresjan's and Mel'čuk's lexicographic terminology, this term has not even been accepted by the wider Russian lexicographic community, let alone lexicographers and linguists worldwide. We will argue that this is due to the nature of the lexical parallels¹ in Russian and other languages (English *vocable*, French *vocable*, German *Vokabel*, Spanish *vocablo*). Data from large corpora of Russian and German, as well as general dictionaries of Russian, German, French, Spanish and English, will be used to support this claim. This paper will also discuss the widely used Russian term *slovarnaja stat'ja* (Eng. *dictionary entry*, Fr. *article de dictionnaire*, Ger. *Wörterbuchartikel*, Sp. *artículo lexicográfico*) and the Russian term *lemma* (Eng. *lemma*, Fr. *lemme*, Ger. *Lemma*, Sp. *lema*), as well as the German term *Stichwort* (Eng. *keyword*), the former in the context of computational linguistics and the latter in the context of general lexicography.

The paper is organized as follows: Section 1 is documentary in nature and includes some commentary. Subsection 1.1 covers Apresjan's definitions; Subsection 1.2 deals with a recent handbook article on “Lexeme” in the authoritative ESLL (*Encyclopedia of Slavic Languages and Linguistics Online*, see Greenberg & Grenoble, 2020–); Subsection 1.3 covers Mel'čuk's definitions. Subsection 1.4 provides an intermediate summary. Section 2 is a corpus- and dictionary-based research

¹ The term *lexical parallel* is used as defined by Dubichynskyi & Reuther (2020). It refers to two or more words in different languages that are phonetically and/or graphically similar, and that have either the same, or partly different, or completely different lexical meaning. These are called full, partial, or false lexical parallels, respectively.

into the topic addressed in the title: *leksema* and *vokabula* from a multilingual perspective. Subsection 2.1 compares corpus and dictionary data on Russian *leksema* and *vokabula*, on the one hand, and on German *Lexem* und *Vokabel*, on the other hand; Subsection 2.2 contains general dictionary data on English *vocable*, French *vocable* and Spanish *vocablo*. The short Section 3 is devoted to the related terms English *lemma* and *keyword*, German *Lemma* and *Stichwort*, which are widely used in general, computational and corpus linguistics. This leads to Section 4, in which we propose names for Apresjan's and Mel'čuk's term *vokabula* in German (*Stichwort* or *Lemma*) and English (*keyword* or *lemma*). The paper ends with the Conclusion and the Bibliography.

Section 1: Russian, French and English terminology

1.1 Apresjan (2009; 2014): the Russian terms *leksema* and *vokabula*

Let us begin with Apresjan's definition of the term *leksema* as the core concept of his "Conceptual apparatus of systemic lexicography" (Apresjan, 2009, p. 509; translation mine – T.R.²):

"A *leksema* is a word in one of its senses, together with all its inherent properties in this sense, essential properties being those to which the rules of the given language refer. [...] It is the *leksema*, not the word, that is the real lexical unit of the language, since normally every polysemic word is used in an utterance in one of its senses. Unlike the meaning, the *lexeme* is a multifaceted unit of language: it has a signifier (phonetic shell), a signified (meaning) and a syntactics (features of combinability)."³

² For the sake of accuracy and transparency, the relevant Russian terms appear in italics and transliteration, rather than in translation from the original Russian.

³ The original Russian text reads as follows: «Лексема – слово в одном из его значений, но во всей совокупности присущих ему в этом значении свойств, причем существенными свойствами признаются те, к которым обращаются правила данного языка. [...] Именно лексема, а не слово, является реальной лексической единицей языка, потому что в норме каждое многозначное слово используется в высказывании в каком-то одном из своих значений. В отличие от значения, лексема – многосторонняя единица языка: у нее есть означающее (фонетическая оболочка), означаемое (значение) и синтактика (особенности сочетаемости)».

Apresjan's definition of the *leksema* highlights the traditional structuralist concept of the dual nature of the linguistic sign: the signifier (Rus. *označajuščee*) and the signified (Rus. *označaeomoe*). This concept was enriched by Melčuk's syntactics (Rus. *sintaktika*) as the third component of the linguistic sign in Meaning $\Leftrightarrow$ Text Theory (Melčuk, 1974).

This definition of *leksema* relies on the concepts of *word* (Rus. *slovo*) and *meaning* (Rus. *smysl*), as well as the ability to identify the (signified) meaning of one (signifying) element within the (signified) meaning of a (signifying) utterance. Further on, the concept of *polysemy* (Rus. *mnogoznačnost'*) of words is crucial for linguists when compiling dictionaries. Apresjan also mentions his approach to linguistic modelling – the concept of interaction between the lexemes of the lexicon and the rules of grammar. However, the term *vokabula* was not included in the “Conceptual apparatus” of (Apresjan, 2009). It can be found later in (Apresjan, 2014, pp. 7–8):

“*Vokabula* and *leksema*”

The basic unit of description in the AS [= *Aktivnyj slovar' russkogo jazyka* / *Active dictionary of Russian* – T.R.], as well as in any other explanatory dictionary, is a single word, or a *vokabula*. [...] ⁴.

The second very important unit of the dictionary is the word in one of its inherent senses, or the *leksema*. More precisely, a *leksema* in the AS is understood as a word taken in one of its senses and with the totality of the linguistically essential properties inherent to the word in this sense. Linguistically essential properties are properties on which some rules of a given language are based. [...]. They include the stylistic properties of a *leksema*, its phonetic properties, the specific prosody of the *leksema*, its grammatical forms, if they differ from the grammatical paradigm of the whole word, the possible meanings of grammatical forms, the *dictionary meaning* (*tolkovanie*) of the *leksema*, its regular modifications in certain contextual conditions, some information about the realities behind the given *leksema*, the government properties of the *leksema*, [...], non valency-controlled syntactic constructions characteristic of the *leksema*, its lexical-semantic combinability and,

⁴ Speaking about the *vokabula* as the basic (!) unit of description in the *Active Dictionary of the Russian language*, Apresjan contradicts his own words („It is the *leksema*, not the word, that is the real lexical unit of the language”, see above) and Melčuk's approach (see below). For more details see also Footnote 12.

finally, its lexical world, i.e., synonyms, analogues, conversions, antonyms and derivatives.”⁵

As can be seen, Apresjan provides a lot of information about the concept of *leksema*. In contrast, the term *vokabula* remains rather underdefined, as does the related term *word* (Rus. *slovo*).⁶

1.2. Spencer & Wiemer (2020): *lexeme* and “*vocable*” in ESSL

For the sake of comparison, and to demonstrate the international recognition of the approach of Apresjan and Melčuk (for the latter, see Section 1.3), the following two paragraphs are taken from Spencer & Wiemer's (2020) recent encyclopedic article on “Lexeme” in the *ESLL* (Greenberg & Grenoble, 2020–):

“Like Zaliznjak (1967), Apresjan and his collaborators conceive of lexemes as abstract units represented by word forms or phrases (i.e., analytical/periphrastic expressions), called “lexes”, organized in paradigms (Melčuk, 2006, p. 21). Lexemes are opposed to “vocables” (Rus. *vokabuly*). The latter in practice correspond to dictionary entries (as defined, e.g., in Katamba 1993, pp. 17f.): a polysemic vocable is a *signifiant* that comprises as many lexemes as the number of meanings

⁵ The original Russian text reads as follows: «Вокабула и лексема. Основной единицей описания в АС, как и во всяком другом толковом словаре, является отдельное слово, или вокабула. [...] Второй важнейшей единицей словаря считается слово в одном из присущих ему значений, или лексема. Более точно, под лексемой в АС понимается слово, взятое в одном из имеющих у него значений, но во всей совокупности присущих ему в этом значении лингвистически существенных свойств. Лингвистически существенными считаются свойства, на которые опираются какие-то правила данного языка. [...] К ним относятся стилистические свойства лексемы, ее фонетические свойства, специфическая просодия лексемы, ее грамматические формы, если они отличаются от грамматической парадигмы всего слова, возможные значения грамматических форм, словарное значение (толкование) лексемы, его закономерные модификации в определенных контекстуальных условиях, некоторые сведения о стоящих за данной лексемой реалиях, управляющие свойства лексемы, [...], характерные для лексемы невалентные синтаксические конструкции, ее лексико-семантическая сочетаемость и, наконец, ее лексический мир, а именно синонимы, аналоги, конверсивы, антонимы и дериваты».

⁶ In Meaning ⇔ Text Theory. “word” is not a term, or at least should not be used as such, due to its ambiguity: 1. polysemous word (= vocable in MTT terminology) vs. 2. lexeme (Fr. *acception d'un vocable*; Eng. *wordsense* in MTT terminology) vs. 3. wordform (element of a lexeme, one of its lexes in MTT terminology). The term “vocable” has been put forward precisely in order to avoid this ambiguity. For a recent discussion of the term *word* in linguistic typology, see Haspelmath, 2023.

[read: senses!⁷ – T.R.] defined by the metalinguistic paraphrases; the paraphrases have to reveal how the meanings are connected (Melčuk, 1995c, p. 500 n3; 1997, p. 346; see Polysemy). From this it follows that homonymous units belong to different vocables. This is summarized in the following definition:

A *lexeme* is a minimal unit of lexicographic description; it corresponds to a dictionary entry. A lexeme is a word in a specific sense, with all its syntactic and morphological peculiarities (characteristic precisely of this sense). A vocable is a family of lexemes such that: (i) their signifiers are identical; (ii) their signifiers [read: signifieds! – T.R.] contain a non-trivial common component. Two lexemes belonging to the same vocable are in relation of *polysemy*; two lexemes having identical signifiers but belonging to two different vocables are in relation of *homonymy*. (Melčuk, 1995a, p. 132 n1; see also Melčuk, 1974, p. 141 n1; 1995b, pp. 415f.; 1997, pp. 102f., 329; 2006, p. 21; Melčuk & Zholkovsky, 1995, p. 51 n2)."

Obviously, Spencer & Wiemer (2020) use the English term *lexeme* in a natural way as the equivalent of the Russian term *leksema*, while putting "vocable" in quotation marks before using it – in the second paragraph – as the plain English equivalent of the Russian term *vokabula*. The authors also link the terms *lexeme* and *vocable* to the notion of *dictionary entry*: a *vocable* is said to correspond "in practice" to a dictionary entry, while a *lexeme* is said to be "the minimal unit of lexicographic description".⁸

⁷ In Meaning ⇔ Text Theory, as generally in linguistics, different meanings of a polysemic word are called (*word*) *senses*. On behalf of "lexes" see Footnote 13 below.

⁸ It is worth noting that Spencer & Wiemer use the three terms (*signifié* – *meaning* – *sense*) synonymously, and that they fail to mention a third component of the linguistic sign, the so-called *syntactics* (of the lexeme). Further on, Spencer & Wiemer (2020) discuss details that are not directly relevant to our exploration of this paper's main topic. However, I believe they are interesting in their own right and worthy of a mention as a side note:

"Since a lexeme is a word (vocable) with a specific meaning, and paradigms are built on lexemes (not on vocables; [...]), we are left with the question how lexical meanings are identified and formulated. This is where most theories, and thus approaches to the lexeme, differ. According to the earlier tradition (cf. Reformatskij, 1967), the meaning of a lexeme is effectively just what is left when we eliminate the grammatical meanings from the meanings of wordforms. This presupposes that the grammatical meanings of a given language are known and that a distributional analysis of all wordforms in the language (including their morphemic structure and their syntactic distribution) is possible, allowing us to isolate sets of forms with an identical lexical meaning (e.g., Zaliznjak, 1967; Kempgen, 1981 on Russian; Laskowski, 1984, pp. 22–26; 1987; 1998, pp. 42–52 on Polish). What all these approaches have in common is that they start from inflection and paradigmatically organized wordforms (however, Laskowski, 1984; 1998 explicitly accounts for noninflected word-form classes) but do not appeal to lexical semantic representations."

1.3. Mel'čuk: *lexie*, *article de dictionnaire*, *vocabulaire* in DEC and ESLL

Along with Russian, French is the best represented language in terms of lexicographic theory and practice of the Meaning $\Leftrightarrow$ Text Theory. See, for example, the four volumes of DEC, the *Dictionnaire explicatif et combinatoire du français contemporain*, i.e., the French version of the *Explanatory-Combinatorial Dictionary* (Mel'čuk et al. (1984, 1988, 1992, 1999)). The theoretical foundations were set out in Mel'čuk, Clas & Polguère (1995). As can be seen below, Mel'čuk, Clas & Polguère (1995, pp. 15–16) proposed the French term (*la*) *lexie* as a hyperonym of the Russian term *leksema*:⁹

“La *lexie* est l'unité de base de la lexicologie — en fait, son objet central et même, en schématisant un peu, son seul et unique objet. [...] Le concept de *lexie* est une formalisation et, simultanément, une généralisation de la notion de MOT. Il n'est pas dans notre intention d'entreprendre ici une analyse poussée de cette notion, même si l'on sait bien que le mot est une unité centrale de la langue: [...] on trouvera une étude détaillée de la question et des propositions de définitions dans Mel'čuk, 1993, Partie I, p. 97 *ssq.*”

Again, as before with Apresjan and Spencer & Wiemer, the understanding of the term *word* is taken to be intuitively clear and prior to the definition of the term *lexie*. Further on, *lexie* is revealed to be synonymous with *unité lexicale*, and hyperonymous to *lexème* and *phrasème*:

“Pour le moment, il nous suffit de dire qu'une *lexie* (n1: Nous utilisons le terme *lexie* tel que propose et défini par B. Pottier (1991)) ou *unité lexicale* est soit un mot pris dans une acception bien spécifique (= *lexème*), soit encore une locution, elle aussi prise dans une acception bien spécifique (= *phrasème*).”¹⁰

Later in their book, in Chapter IV on the macrostructure of the dictionary, Mel'čuk, Clas & Polguère (1995, pp. 155–156), explain the connection between the terms *lexie* and *article de dictionnaire*.

“L'unité de base du DEC, comme nous l'avons déjà dit, est la *lexie*; chaque *lexie* est décrite dans un article de dictionnaire spécifique. Une *lexie* ne correspond qu'à

⁹ As we will see below, Fr. *lexie* (= Fr. *unité lexicale*) is a hyperonym of Fr. *lexème* (Rus. *leksema*) and of Fr. *phrasème* (Rus. *frazema*).

¹⁰ A later term that was proposed instead of the too general *phrasème* is *locution* = *idiom*.

un seul article de dictionnaire, et *vice versa*, un article de dictionnaire correspond toujours à une seule lexie.”

In a further step, the French equivalent of the Russian term *vokabula* is established: the authors refer to it as *(le) vocable*, and to its corresponding lexicographic description as *le superarticle de dictionnaire*:

“On constate, cependant, que, dans de nombreux cas, des lexies ont un même signifiant et, en plus, manifestent entre elles des liens sémantiques assez évidents. Par exemple, on a en français au moins trois lexies ayant le signifiant /lapɛ̃/ qui sont sémantiquement liées: LAPIN (animal) (*Le lapin mange des carottes*), LAPIN (viande de cet animal) (*une excellente terrine de lapin*) et LAPIN (fourrure de cet animal) (*un manteau en lapin*). C’est le phénomène bien connu de la *polysémie* [...] Le DEC, [...] aura au moins trois articles différents pour les trois lexies LAPIN illustrées ci-dessus — trois articles différents réunis cependant sous le même vocable LAPIN. Dans le DEC, donc, un vocable n’est pas un article de dictionnaire mais un ensemble d’articles apparentés par leur sémantisme et leur forme — un *superarticle*.”

More recently, in their ESSL, article on the *Explanatory-Combinatorial Dictionary* Melčuk & Reuther (2020) used the term “vocale” in quotation marks. Building on the well-established term *lexical unit* (abbreviated LU), they introduced the term (*lexicographic*) *superentry* to refer to the collection of dictionary entries describing the *vocable*:

“The target of a lexical entry in the ECD is, as stated above, a “monosemous LU” – a lexeme or an idiom. Each LU is described by one dictionary entry, and each entry describes one LU. But what about “polysemous words,” which are so numerous in any language? This brings us to the notion of “vocale,” i.e., a lexicographic superentry: if a lexeme is analogous to a “word sense,” then a vocable corresponds to a “full word.”

A “vocale” is the set of all LUs such that

- (i) all of them have an identical signifier, and
- (ii) any one of them is semantically linked to all the others.”

The background for the quotation marks will be the subject of discussion in Section 2 below.

1.4. Intermediate summary

Let us summarize what has been discovered so far.

- The term *leksema* as defined by Apresjan and Melčuk is a key concept in lexicology¹¹ and is directly linked to the corresponding concept of dictionary entry (Rus. *slovarnaja stat'ja*) in lexicography.
- In Apresjan's approach, "the basic unit of description in the AS is the *vokabula*," whereas Melčuk holds that "l'unité de base du DEC est la *lexie*."¹²
- The Russian term *leksema* has equivalents in other languages, such as Eng. *lexeme*, Fr. *lexème*, while the Russian term *vokabula* (cf. Latin *vocabulum*) has evolved as a technical term in Russian linguistics and has been assigned terminological equivalents in French (*le*) *vocable* and in English (*the*) *vocable*.
- The French term (*la*) *lexie* has no generally accepted equivalent in other languages. Spencer & Wiemer (2020, see above) put "lexes" in quotation marks.¹³
- The French term *lexie* is said to be synonymous with Fr. *unité lexicale*; its equivalent in English is *lexical unit* (Rus. *leksičeskaja edinica*).
- The Russian term *slovarnaja stat'ja* has the commonly known equivalents *article de dictionnaire* in French and *dictionary entry* in English.
- The innovative French term (*le*) *superarticle* (*de dictionnaire*) has its equivalent (*lexicographic*) *superentry* in English, whereas there is no such established equivalent in Russian.¹⁴

¹¹ However, note that the older Russian term *leksiko-semantičeskij variant slova*, which was abandoned by the Moscow Semantic School in favor of *leksema*, can still be found in the works of influential Russian lexicographers.

¹² The difference in wording between Apresjan and Melčuk must not be overestimated. It stems from a difference in the perspective of linguistic modelling, which is evident in almost all of their work. Apresjan often "looks in the direction" of linguistic analysis from Text to Meaning. In the case of the dictionary, he starts with a word and all its possible usages, moving towards defining separate meanings, while keeping them under the "umbrella" of one *vokabula* as long as possible, i.e., for as long as polysemy can be argued for (homonymy being the alternative case for establishing two *vokabuly*). Melčuk's modelling always (!) proceeds in the direction of synthesis from Meaning to Text. This means moving from a semantic representation (of what the speaker wants to say) to identifying lexical units and grammatical relations that express this Meaning as a Text. In this approach, the dictionary contains all the necessary information about *lexical units*, which, under the condition of polysemy, come as a set called a *vocable*. However, both Apresjan and Melčuk share the concept of two major inherently connected components of linguistic description: the lexicon and the grammar.

¹³ In Meaning $\leftrightarrow$ Text theory the term *lex* is defined as an 'element of a lexeme', and not as the English equivalent of *lexie*, which is called a *lexical unit* in English.

¹⁴ One could think of *slovarnaja sverxstat'ja* or *slovarnaja superstat'ja*.

Section 2: The multilingual perspective

In this section, we will examine the favorable and unfavorable conditions for the existence and/or emergence of equivalents of *leksema* and *vokabula* in four major European languages (English, French, German, Spanish).

2.1 Data from corpora and general dictionaries on Russian *vokabula* / *leksema* and German *Vokabel* / *Lexem*

All four words are attested in corpora of Russian and German, respectively, as well as in general dictionaries of the two languages.

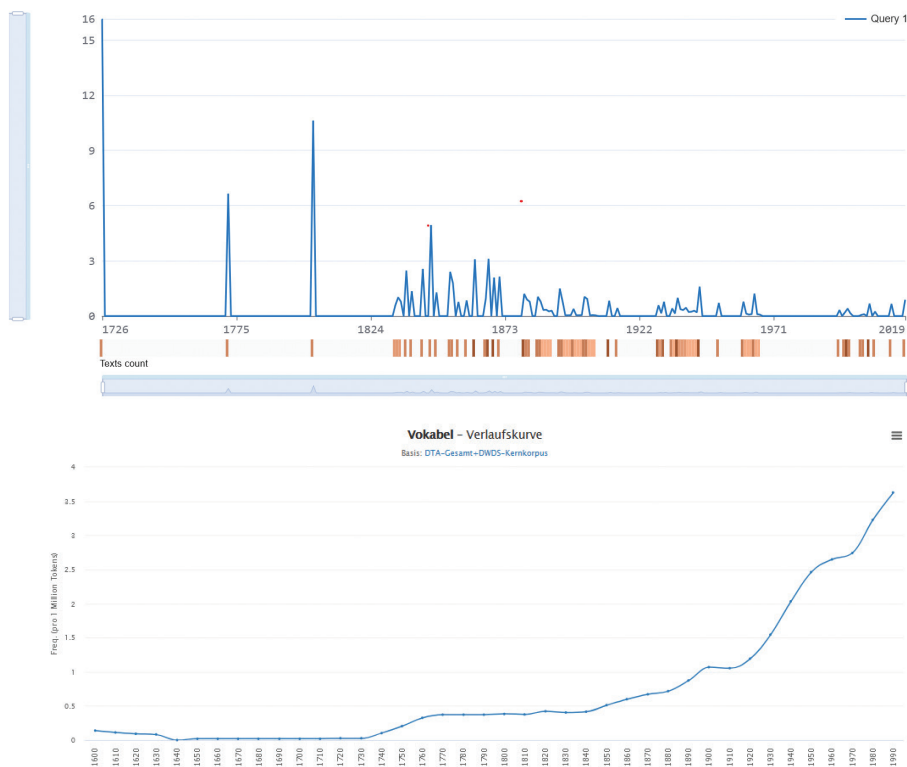
In Russian, the item *vokabula* is attested in the *National Corpus of the Russian Language* (ruscorpora.ru) and in general dictionaries, where it is found with one archaic sense (see below). As will become clear, this sense is different to that introduced by Apresjan and Mel’čuk for the lexicographic term.

In German, the item (*die*) *Vokabel* is equally attested in corpora — we will refer to Cosmas II corpus (cosmas2.ids-mannheim.de) — and is found in general dictionaries in two senses. In one of these senses, Ger. *Vokabel* is widely used in the teaching and learning of foreign languages, referring to a word that the student should learn, always in pair with an equivalent in the student’s mother tongue (and usually written down for memorization and repetition in a special notebook called *Vokabelheft*). The question is whether there is enough “favorable space” for another language-related sense of *Vokabel*? Referring to corpus data and dictionary data, we will argue that there is not.

The following quantitative data were taken from the subcorpora specified as “Corpus name: main” in the *Russian National Corpus*, and *W Archiv der Geschriebenen Sprache; Korpus W-öffentlich - alle öffentlichen Korpora des Archivs W (mit Neuakquisitionen)*; *Archiv-Release: Deutsches Referenzkorpus DeReKo-2025-I* in the *Cosmas II corpus*:

Russian <i>vokabula</i>	Russian <i>leksema</i>	German <i>Vokabel</i>	German <i>Lexem</i>
Request 28.02.2025 Documents found: 61 Examples found 81	Request 01.03.2025 Documents found: 84 Examples found: 678	Request 01.03.2025 Examples found: 26.434	Request 01.03.2025 Examples found: 356

While *leksema* (678 examples found) and *Lexem* (356 examples found) appear to be closely related in Russian and German, the data for *vokabula* (81 examples found) and *Vokabel* (26434 examples found) are strikingly different in the two languages. The timelines of the Russian and German items provide more details: *vokabula* was popular in the 19th century, specifically from the 1820s to the 1870s, and is rarely used nowadays, whereas *Vokabel* has steadily increased in popularity from the 18th century to the present day.



Further on, in the dictionaries, the Russian word *vokabula* is monosemic and marked as archaic. The German word *Vokabel* is polysemic with one sense as provided above that is roughly equivalent to the archaic sense of the Russian word *vokabula*. Detailed lexicographic definitions¹⁵ and examples from texts are given below:

¹⁵ The original definition of Russian *vokabula*, taken from the four-volume *Small Academic Dictionary of Russian*, reads as follows: “Устар. Иностранное слово с переводом на родной язык как предмет заучивания при обучении какому-л. языку.” (<http://feb-web.ru/feb/mas/mas-abc/default.asp>). The German definition of *Vokabel* is taken from the online resource *Der deutsche Wortschatz von 1600 bis heute* (<https://www.dwds.de/wb/Vokabel>).

	<i>vokabula</i>	<i>Vokabel</i>
definition	<i>Archaic.</i> A foreign word with a translation into a person's mother tongue as an object of memorisation in foreign language teaching and learning	1. Einzelnes Wort einer fremden Sprache 2. Bezeichnung, Ausdruck
examples	<i>Učen'e vokabul, grammatiki i razgovorov šlo samo po sebe v klasse; no doma ja ničego ne učil naizust' 'The study of vocabulary, grammar, and conversations went on by itself in class; but at home I learned nothing by heart' [S. Aksakov, 1856]; Čto èto ty francuzskie vokabuly učiš'? — kivnul Ivan na tetradku, ležavšuju na stole 'So you're learning French word lists? — Ivan nodded toward the notebook lying on the table' [F. Dostoevskij, 1880].</i>	1. <i>Bei solch' tollen Temperaturen Vokabeln pauken?</i> [Nürnberger Nachrichten]; <i>Eine neue Vokabel aus der Wirtschaftswelt ist in Rußland in aller Munde, obwohl noch die wenigsten etwas mit ihr anzufangen wissen: der Voucher, auch Privatisierungsscheck genannt.</i> [Süddeutsche Zeitung]; <i>Dabei wird neben Grammatik und Vokabeln auch der kulturelle Hintergrund unterrichtet.</i> [News]. 2. <i>"Katastrophal" – das war die meistgebrauchte Vokabel unter den Aktienanalysten, als sie die Gewinnprognosen des Siemens-Konzerns für das gerade erst begonnene Geschäftsjahr 1996/97 vernahmen.</i> [Nürnberger Nachrichten]; <i>Nicht eben das, was man sich unter der Vokabel "Sofortvollzug" vorstellt, den die Stadtentwässerung fordert.</i> [Süddeutsche Zeitung]; <i>Das Wort "löschen" ist die meistgebrauchte Vokabel der theatralischen gesellschaftlichen Bestandsaufnahme "Letzter Aufruf".</i> [News].

As a result, general dictionaries report the archaic Russian monosemic item *vokabula* and its polysemic German phonic equivalent *Vokabel*. The latter has two senses: *Vokabel 1*, which is related to language teaching and learning and is similar in meaning to *vokabula*, and *Vokabel 2*, which means ‘denomination, expression’.

Adding the Russian linguistic term, as defined above by Apresjan and Melčuk (A&M), as a second lexeme *vokabula 2* would result in a polysemic Russian word *vokabula* with both archaic and modern senses. The “semantic bridge” between these two senses is ‘word of a language’ (a proof of their polysemic nature). The rest of the two senses, however, is to a certain degree antonymous, cf. ‘one word in one language with its translation into another language’

for *vokabula 1* versus 'set of formally identical and semantically closely related lexemes of one language' for *vokabula 2*.

For the German word *Vokabel*, A&M's meaning is not attested, and I believe that there is an unfavorable condition to establish, along with *Vokabel 1* and *Vokabel 2* as defined above, a third lexeme *Vokabel 3* with A&M's terminological meaning. *Vokabel 1* is so well known and frequently used that another language-related, and to a certain degree antonymous (see above for *vokabula 2*) lexical unit *Vokabel 3* would confuse many general language users, as well as specialists. However, the remaining problem is how to express the meaning of A&M's term *vokabula 2* in German.

After turning to French and Spanish, as opposed to English, an alternative for both German and English will be given in Section 4 below.

2.2 French *vocable*, Spanish *vocablo*, English *vocable*

In this subsection we will demonstrate how French and Spanish differ from English in terms of the acceptability of the phonetic equivalents of *vokabula 2*.

2.2.1 French *vocable*

Checking authoritative online dictionaries of French, the *Larousse*, the *Robert*, and the *Oxford French Dictionary* reveals the following:

Larousse en ligne <https://www.larousse.fr/dictionnaires/francais/vocable/82341>)

1. Mot, terme d'une langue considéré surtout sous le rapport de sa signification, de son individualité lexicale (*petit, petite, petits* sont trois formes d'un même vocable).
2. Nom du saint ou du mystère sous le patronage duquel une église est placée.

Robert en ligne <https://dictionnaire.lerobert.com/definition/vocable>

Mot d'une langue, considéré dans sa signification, sa valeur expressive.

Oxford French Dictionary online

1. (mot) term;
2. *Religion* chapelle sous le vocable de chapel dedicated to.

As can be seen, all definitions refer to the hyperonym *mot* (word), and the meaning of *vocable 1* in *Larousse* is the closest to A&M's lexicographic term *leksema*. This is a favorable precondition for the integration of A&M's term as a sense of *vocable* into the French language.

2.2.2 Spanish *vocablo*

For Spanish, the reference sources are three dictionaries, the *Spanish Wiktionary online*, the *Diccionario del estudiante Real Academia Española online*, and the *Oxford Spanish Dictionary online*:

Spanish Wiktionary online <https://es.wiktionary.org/wiki/vocablo>

Unidad mínima del lenguaje escrito o hablado que puede darse separadamente

Diccionario del estudiante Real Academia Española online <https://www.rae.es/diccionario-estudiante/vocablo>

Palabra (sonido o conjunto de sonidos dotados de significado que constituyen una unidad indivisible del discurso, y su representación gráfica). *El vocablo “fútbol” proviene del inglés “football”. ¿Cuál es el significado de este vocablo?*

Oxford Spanish Dictionary online

(formal) word vocablos extranjeros

As can be seen, the Spanish item *vocablo* is a synonym of *palabra* (word). In the *Diccionario del estudiante Real Academia Española online*, there is a quasi-linguistic definition of what a *word* (*palabra*) is: “Word (sound or set of sounds with a meaning that constitute an indivisible unit of speech, and its graphic representation)”. This fact is a favorable condition for the integration of A&M’s term as a sense of *vocablo* into the Spanish language.¹⁶

2.2.3 English *vocable*

The *Oxford dictionary* (for British English) and the *Merriam Webster Dictionary* (for American English) both refer to the hyperonym *word*:

Oxford dictionary of English online

a word, especially with reference to form rather than meaning.

¹⁶ Indeed, Alonso Ramos (2004), in her book related to the framework of the *Explanatory-Combinatorial Dictionary*, defines *vocablo* as a lexicographic term in line with A&M’s definition (“vocablo or polysemic word”): “Así, un vocablo o palabra polisémica como *dar* agrupará lexemas que tienen un comportamiento libre, es decir, que son escogidos según su sentido, y otros lexemas que son restringidos, es decir, que son controlados por otros ya escogidos previamente en el discurso. En otras palabras, un vocablo puede estar constituido por lexemas autónomos o libres y por lexemas colocativos.”

Merriam-Webster online <https://www.merriam-webster.com/dictionary/vocable#dictionary-entry-1>

term, specifically: a word composed of various sounds or letters without regard to its meaning. *The birds' phrasings are both melodic and mechanical, cyclical and spontaneous, like the wordless vocables of scat singers* [New York Times, 11 Feb. 2020].

As can be seen, both dictionaries offer only one meaning, which – although referring to *word* – is much narrower than the meaning of A&M's lexicographic term. In English, *vocable* is exclusively related to the form of a word, i.e., to its signifier, and not to the signified. This is an unfavorable condition for the integration of the word *vocable* in A&M's sense into the English language. Nevertheless, it is found as a term in quotation marks (Melčuk & Reuther, 2020, see above). An alternative name will be presented below.

To conclude this section, it became clear that the French *vocable* and the Spanish *vocablo*, both of which refer to *word/mot/palabra* as hyperonyms, can easily be used as lexicographical terms: *word/mot/palabra* refer to form and meaning, the traditional dual nature of lexical items. This creates favorable conditions for the integration of A&M's term into French and Spanish. On the contrary, the definition of *vocable* in English does not include the signified component of *word* (cf. 'a word composed of various sounds or letters without regard to its meaning'), resulting in unfavorable conditions for the integration of *vocable* in A&M's sense into the English language.

Section 3.

A short exploration of the terminology of traditional, computational and corpus linguistics: *lemma*, *keyword*

In order to solve the problem of how the meaning of A&M's term *vokabula* can be expressed in English and German, let us briefly consider two examples of terminology from computational and corpus linguistics.

Firstly, the term *keyword* has been used in lexicography for centuries, as has its German equivalent *Stichwort*. However, until now, neither *keyword* nor *Stichwort* have been used as equivalents of Russian *vokabula*.

Secondly, the recently updated version of the Russian National Corpus uses the term Rus. *lemma* (written in Cyrillic as *лемма*) for the top search box in

the Russian user interface in parallel with Eng. *lemma* for the top search box in the English user interface. This demonstrates the complete integration of the term *lemma* into both Russian and English. However, *lemma* has not yet been suggested as a potential equivalent to *vokabula* in English.

In the next section, we will suggest German and English terminological equivalents of the lexicographic term Rus. *vokabula* (Fr. *vocable*, Sp. *vocablo*) in A&M's sense which make use of Ger. *Stichwort* (alternatively, *Lemma*) and Eng. *keyword* (alternatively, *lemma*).

Section 4: A traditional and a corpora-oriented proposal for equivalents of Apresjan & Mel'čuk's term *vokabula/vocable* in German and English

Obviously, the term *Stichwort* can be used to refer either to one lexeme, or to the set of lexemes that fall "under the umbrella" of a *vokabula* in the sense of A&M's term, cf. the following example: "Zum <Unter dem> Stichwort HASE findet man im Wörterbuch drei Bedeutungen." (Rus. "Vokabula HASE imeet v slovare tri znachenija." – for English, see below). Consequently, when it comes to the term *dictionary entry*, the term *Stichworteintrag* can be used referring to the "name of the umbrella" where all information about one or more lexemes that fulfil the definition of A&M's term *vokabula* is stored. Cf. "Der Stichworteintrag HASE enthält drei Unterartikel zu drei verschiedenen Bedeutungen des Wortes Hase." In other words, a *Stichworteintrag* is a dictionary entry for either a monosemic or for a polysemic word.

What remains to propose is a more favorable name for A&M's notion of *vokabula* in English. In my opinion, this could be either the traditional term *keyword* or the international term of Greek origin Eng. *lemma* (Rus. *lemma*, Fr. *lemme*, Ger. *Lemma*, Sp. *lema*, etc.).

If we agree to this solution, the quotation from Mel'čuk & Reuther (2020), cited in Section 1, Subsection 1.3, would read as follows:

"The target of a lexical entry in the ECD is, as stated above, a "monosemous LU" – a lexeme or an idiom. Each LU is described by one dictionary entry, and each entry describes one LU. But what about "polysemous words," which are so numerous in any language? This brings us to the notion of keyword (or lemma), i.e.,

a lexicographic superentry: if a lexeme is analogous to a “word sense,” then a keyword (or lemma) corresponds to a “full word.”

A keyword (or lemma) is the set of all LUs such that

- (i) all of them have an identical signifier, and
- (ii) any one of them is semantically linked to all the others.”

An alternative approach is to do without *keyword* and *lemma* and stick to (*lexicographic superentry*), cf.: “... if a lexeme is analogous to a “word sense,” then a superentry corresponds to a “full word.”

A superentry is the set of all LUs such that

- (i) all of them have an identical signifier, and
- (ii) any one of them is semantically linked to all the others.”

Consequently, the translations of the German sentences above are the following: “*In the dictionary, the keyword <lemma, superentry> HASE has three meanings.*”; “*The superentry HASE contains three subentries for three different meanings of the word Hase.*”

That concludes my contribution as a linguist and a native speaker of German. It is for linguists and native speakers of English to decide on the solution for English.

Our findings presented in Sections 2 to 4 are as follows:

1. Firmly anchored in Apresjan's and Mel'čuk's lexicographic theory and practice, the Russian term *leksema* and its lexical parallels (Eng. *lexeme*, Fr. *lexème*, Ger. *Lexem*, Sp. *lexema*) have made their way into international lexicographic terminology.
2. The Russian term *vokabula* can easily be integrated into the French and Spanish linguistic terminology but meets resistance in German and English. Luckily enough, Ger. *Stichwort* and Eng. *Lemma*, in our opinion, fulfill favorable conditions for their integration as equivalents of A&M's Russian term *vokabula* into German and English. In combination with the terms Ger. (Wörterbuch) *eintrag* and Eng. (dictionary) *entry*, we consider Ger. *Stichworteintrag* and Eng. *superentry* as favorable terms for the complete information for all lexemes that fall “under the umbrella” of a Rus. *vokabula*, Ger. *Stichwort*, Eng. *lemma*.

Conclusion

Given the data on Russian, English, French, German and Spanish, the terms Rus. *leksema* – Eng. *lexeme* – Fr. *lexème* – Ger. *Lexem* – Sp. *lexema* are full terminological lexical parallels. On the contrary, the words Rus. *vokabula* – Eng. *vocabulary* – Fr. *vocabulaire* – Ger. *Vokabel* – Sp. *vocablo* have a historical background, dating back centuries before their terminological sense was first introduced by Apresjan and Melčuk for Russian. The lexical parallels of Rus. *vokabula* – Fr. *vocabulary* – Sp. *vocablo* proliferated easily into French and Spanish terminology, leaving the English equivalent “*vocabulary*” halfway in quotation marks, and the German *Vokabel* undocumented.

The situation in English and German is the following: In modern English, *vocabulary* is defined as ‘a word, especially with reference to form rather than meaning’. In modern German, *Vokabel* is defined as a one-to-one correspondence between two words, one in a foreign language, and one in a learner’s mother tongue. Both definitions contradict the idea of *vocabulary* as a set of lexemes in one and only one language.

In order to integrate the concept of *vocabulary* into German and English terminology, the following is proposed: The terms Ger. **Stichwort** (alternatively, **Lemma**), and Eng. **keyword** (alternatively, **lemma**) can refer either to one lexeme, or to the set of lexemes that fall “under the umbrella” of a *vocabulary* in the sense of Apresjan & Melčuk’s terminology; consequently, Ger. *Stichwort* or *Lemma* and Eng. *keyword* or *lemma*, should be used as equivalents of the terms Rus. *vocabulary* – Fr. *vocabulary* – Sp. *vocablo*. Further on, Ger. **Stichworteintrag** (alternatively, **Lemmaeintrag**) and Eng. **superentry** (as proposed by Melčuk) can be used, referring to the “name of the umbrella” where all information about one or more lexemes that fulfil the definition of A&M’s term *vocabulary* is stored.¹⁷

References

- Alonso Ramos, M. (2004). *Las construcciones con verbo de apoyo*. Visor Libros, Madrid.
 Apresjan, Ju. (1986). Integral’noe opisanie jazyka i tolkovyj slovar’. *Voprosy jazykoznaniya*, 2, 57–70.

¹⁷ Many thanks to the reviewers for their comments on the first version of this paper, and to Igor Melčuk for his feedback and clarification of his position on several drafts of this paper. The remaining views on terminology are my own.

- Apresjan, Ju. (2009). Ponjatijnyj apparat sistemnoj leksikografii, In Ju. Apresjan, *Issledovanija po semantike i leksikografii. Tom 1. Paradigmatika* (p. 486–552). Jazyki slavjanskix kul'tur, Moskva.
- Apresjan, Ju. (2014). Ob aktivnom slovarе russkogo jazyka. In Ju. Apresjan (Ed.), *Aktivnyj slovar' russkogo jazyka. Tom 1. A-B* (pp. 5–36). Jazyki slavjanskoj kul'tury, Moskva.
- Dubichynskij, V., & Reuther, T. (2020). *Leksygrafuvannja leksyčnyx parallelej/Lexikografie lexikalischer Parallelen. Teoretyčni poležennja ta ukrajins'ko-nimec'kyj slovnyk. Monografija / Theoretische Grundlagen und ukrainisch-deutsches Wörterbuch. Monografie*. Peter Lang, Berlin.
- Greenberg, M. L., Grenoble, L. A., & others (Eds.). (2020–). Encyclopedia of Slavic Languages and Linguistics Online. Leiden & Boston: Brill. <https://doi.org/10.1163/2589-6229>
- Haspelmath, M. (2023). Defining the word. *Word*, Vol. 69, No. 3, 283–297, <https://doi.org/10.1080/00437956.2023.2237272>
- Katamba, F. (1993). *Morphology*. Macmillan, London.
- Kempgen, S. (1981). "Wortarten" als klassifikatorisches Problem der deskriptiven Grammatik. *Historische und systematische Untersuchung am Beispiel des Russischen*. Otto Sagner, München.
- Laskowski, R. (1984). Podstawowe pojęcia morfologii. In R. Grzegorzczkova et al., *Gramatyka współczesnego języka polskiego, vol. I: Morfologia* (pp. 9–57). Wydawnictwo Naukowe PWN, Warszawa.
- Laskowski, R. (1987). On the concept of the lexeme. *Scando-Slavica*, 33, 169–178.
- Laskowski, R. (1998). Zagadnienia ogólne morfologii. In R. Grzegorzczkova et al., *Gramatyka współczesnego języka polskiego, vol. I: Morfologia*. (pp. 27–86). Wydawnictwo Naukowe PWN, Warszawa.
- Melčuk, I. (1974). *Opyt teorii lingvističeskix modelej "Smysl ⇔ Tekst"*. Nauka, Moskva.
- Melčuk, I. (1993). *Cours de morphologie générale (théorique et descriptive)*. Vol. 1. *Introduction et Première partie: Le mot*. Les Presses de l'Université de Montréal – CNRS, Montréal.
- Melčuk, I. (1995a). Countability vs. non-countability of nouns in Russian and their lexicographic description. In I. Melčuk, *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (pp. 125–133). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach.
- Melčuk, I. (1995b). La combinatoire des numéraux du type DVOE en russe contemporain. In I. Melčuk, *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (pp. 383–422). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach, Moskva – Wien.
- Melčuk, I. (1995c). Slovoobrazovanie v lingvističeskix modeljax tipa Smysl-Tekst (predvaritel'nye zamečanija). In I. Melčuk. *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (p. 475–504). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach, Moskva – Wien.

- Melčuk, I. (1997). *Kurs obščej morfologii. Tom I. Čast' 1: Vvedenie. Slovo*. Jazyki russkoj kul'tury, Wiener Slawistischer Almanach, Moskva – Wien.
- Melčuk, I. (2006). *Aspects of the theory of morphology*. Walter de Gruyter, Berlin–New York.
- Melčuk, I., et al. (1984, 1988, 1992, 1999). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques (I–IV)*. Les Presses de l'Université de Montréal, Montréal.
- Melčuk, I., Clas, A., & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Éditions Duculot, Louvain-la-Neuve.
- Melčuk, I., & Reuther, T. (2020). *Explanatory Combinatorial Dictionary*. In M. L. Greenberg, L. A. Grenoble et al. (Eds.), *Encyclopedia of Slavic Languages and Linguistics Online*. Brill <https://doi.org/10.1163/2589-6229>
- Melčuk, I., & Zholkovsky, A. (1995). *An explanatory combinatorial dictionary of modern Russian*. In I. Melčuk, *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (pp. 17–54). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach, Moskva–Wien.
- Plungjan, V. (2000). *Obščaja morfologija: Vvedenie v problematiku*. URSS, Moskva.
- Pottier, B. (1991). La lexie: une mise au point nécessaire. In A. Zampolli (Ed.), *Computational lexicology and lexicography. Special issue dedicated to Bernard Quemada*. Giardini, Pisa.
- Reformatskij, A. (1967). *Vvedenie z jazykovedenie*. 4-e izdanie. Prosveščenie, Moskva.
- Spencer, A., & Wiemer, B. (2020) *Lexeme*. In M. L. Greenberg, L. A. Grenoble et al. (Eds.), *Encyclopedia of Slavic Languages and Linguistics Online*. Brill, Leiden, Boston.
- Zaliznjak, A. (1967). *Russkoe imennoe slovoizmenenie*. Nauka, Moskva.



Rémi Camus

SeDyL, INaLCO
France

 <https://orcid.org/0000-0002-0806-5939>

***Poskakali!* « Filons ! » / « Assez couru ! » : construire l’instable (les verbes *nesti*, *nosit'*, *skakat'* et le préfixe *po-*)**

***Poskakali!* “Let’s Flee!” / “Enough Running!”: Constructing Instability (The verbs *nesti*, *nosit'*, *skakat'* and the prefix *po-*)**

Abstract

Russian verbs capable of expressing spatial displacement (among other meanings) allow the prefix *po-* to express two contradictory meanings: inchoation (onset of action: ‘Let’s go!’) and delimitation (temporal bounding: ‘Enough running!’). This article explores the conditions triggering this duality, focusing on the grammatical category of so-called “verbs of motion”. Within this small set, which contrasts determinate and indeterminate verbs, the former tendentially select the inchoative meaning, while the latter select the delimitative meaning. The analysis is carried out using the example of the determinate verb (*po*)*nesti* and the indeterminate (*po*)*nosit'*, with extension to non-motion uses.

The core argument is that determinate verbs – like other motion-capable verbs – rely on an unstable term *a*, whereas indeterminate verbs, as products of a morphological derivation process, operate on the feature that distinguishes verbs within this category: the mechanism creating this instability. In the case of *nesti/nosit'*, this involves a contingently defined locator (e.g., a “carrier”). Consequently, *po-* prefixed to determinate verbs targets the existence of *a*, while with indeterminate verbs, it modifies the properties of the “carrier”. For verbs outside the determinate/indeterminate opposition, the ambiguity stems directly from the verbal base’s interpretation, generating specific semantic effects.

Keywords

Linguistics, Russian verb motion, determinate, indeterminate, prefixation, preverb, morphology, derivation invariant, variation

Initialement exposée lors du colloque « Le russe du XX^e siècle » qui s'est déroulé à Aix-en-Provence en mai 2000 avec la participation remarquée de Jurij Derenikovič Apresjan, la présente étude se veut un humble hommage à la mémoire de ce grand linguiste. Plusieurs des questions soulevées dans ce travail resté inédit avaient fait écho aux préoccupations de Jurij Derenikovič, et stimulé une discussion de couloir dont je garde un souvenir reconnaissant : elle fut un encouragement à élargir la recherche d'enjeux scientifiques partageables entre théories concurrentes. La discussion avait débouché sur le thème des exemples qu'on porte en soi pour s'y référer toute sa vie durant. Quiconque connaît cette inquiétude pour des faits singuliers qui ne laissent de titiller et insister en tapinois, peut, même si ces faits lui résistent, se dire fièrement un disciple d'Apresjan.*

0. Les termes du problème

Les deux traductions opposées de l'exclamation *Poskakali!* – ‘Filons!’ et ‘Assez couru!’ – font écho à deux interprétations bien répertoriées du préfixe (préverbe) russe *po-* combiné à certains verbes susceptibles de décrire un déplacement. Pour une partie de ces verbes, cette différence d'interprétation est généralement associée à celle qui oppose les deux séries de verbes constitutives de la catégorie grammaticale connue des slavistes sous les termes « verbes de mouvement » ou encore « sous-aspect déterminé / indéterminé » (Garde, 1998, §606) : *ponesti* – *ponosit'* ‘porter, subir etc.’ Au paradoxe énantiosémique initial succède alors une énigme plus générale : comment rendre compte de la répartition des deux interprétations de *po-* dans ces verbes ? Le présent article prend le parti d'y répondre par la combinaison des verbes et du préfixe figurant dans le titre : les verbes *nesti-ponesti* (§1–2), *nosit'-ponosit'* (§3–4), *skakat'-poskakat'* (§0.1 et §5), et en mobilisant une théorie du sous-aspect compatible avec notre approche construisant la signification à partir des formes.

Cette feuille de route exigeante impose quelques éclaircissements au préalable.

* Colloque « Le russe du XX^e siècle » (Aix-en-Provence, 5–7 mai 2000), cf. le recueil *Russkij jazyk peresekaja granicy*. Dubna, 2001. Des fragments ou versions antérieures du texte qu'on va lire ont bénéficié, à divers stades, des observations de Chr. Bonnot, S. De Vogüé, Sv. Krylosova, D. Pailard, A. Tokmakov, qui ne sauraient être tenus responsables de mes oublis, erreurs ou désaccords. Les mot à mot des exemples ne sont détaillés qu'en tant que de besoin et ne mentionnent pas de morphèmes zéros (« Ø ») ; les abréviations sont regroupées en fin d'article.

0.1. L'exemple du mot *poskakali*

Le verbe cité est un verbe perfectif au passé. Pour fixer une terminologie de convention, indiquons les grandes valeurs du perfectif passé auxquelles il sera fait référence : la valeur d'aoriste (événement ponctuel dans un récit : *proskakal...* 'il galopa [...2 km]'), de parfait (*uskakal* 'il a filé (il n'est pas là)'), de résultat visé et atteint (*doskakal!* 'il est arrivé au bout ! (en galopant)').

Du préfixe *po-* qui perfective *skakat'*, disons pour l'instant qu'il se caractérise par son apparente « légèreté » sémantique ; ses emplois sont le plus souvent associés aux « modalités d'action », dont justement les modalités inchoative (entamer l'action) et délimitative (agir un temps limité).

Quant au radical alternant *skak-/skok-*, il est utilisé comme une onomatopée dans *skok v vodu!* 'Hop dans l'eau !', ou sous forme verbale décrivant le galop du cheval et, plus généralement, un déplacement rapide :

- (1) *Po-skaka-l-i!*
po-galoper-PASS-PLUR (perfectif)
 'Filons !'
 'Assez couru !'

La traduction 'Filons !' répond à l'usage injonctif du parfait. Il y a inchoation dans la mesure où il s'agit de quitter le lieu où l'on se trouve, et donc de se mettre en route, qu'il y ait ou non prise en compte d'une destination :

- (1a) *Po-skaka-l-i* *ot-sjuda!*
po-bondir-PASS-PLUR de-ici
 'Filons d'ici !'

Suivant la seconde traduction, *Poskakali!* relève également du parfait, mais cette fois-ci sous forme de bilan performatif : la profération de l'énoncé clôt le procès, ce qui correspond à la valeur délimitative. Le contexte de (1b) ne pourra être rétabli qu'en fin d'article :

- (1b) *Po-skaka-l-i,* *xvat-it,* *pora* *umirat'.*
po-bondir-PASS-PL suffire-3.SG, il_est_temps mourir.
 'Fin de galoper, ça suffit / Assez galopé. Il est temps de mourir.'

Les lignes qui suivent détaillent ces valeurs étiquetées respectivement PO-INCH et PO-DEL.

0.2. *Po-*, modalités d'action et types de procès

PO-INCH et PO-DEL ont suscité des interrogations de nature différente dans la littérature.

Le problème généralement soulevé pour PO-INCH est l'existence, à côté de *po-*, d'autres préverbes participant à la construction de verbes inchoatifs, notamment *za-*, *voz-*, *raz*. On s'accorde sinon sur l'inventaire, au moins sur le nombre réduit des unités où c'est *po-* qui est choisi ; LeBlanc (2010 : 44) les énumère sous trois rubriques citées ici sans commentaires :

- verbes dits « de mouvement », *idti* 'aller', *nesti* 'porter' (cf. *infra*) ;
- verbes exprimant un déplacement (*mčat'sja* 'filer', *dut'* 'souffler') ;
- verbes de « perception ou d'état mental » (*ljubit'* 'aimer', *slyšat'sja* 'se faire entendre', *čuvstvovat'* 'ressentir').

Cet emploi de *po-* est généralement distingué de l'expression d'un pur commencement. Pour Tauscher, Kirschbaum (1989 : 275), le préfixe représente ici souvent un « simple marqueur aspectuel » ; selon Zaliznjak, Šmel'ev (2000 : 109–110), il en vient à désigner « l'action en elle-même » : « le commencement d'une situation donnée s'avère équivalent à sa réalisation ».

À l'inverse de leur traitement de PO-INCH, les manuels n'indiquent pas de préfixation concurrente à PO-DEL. Selon Zaliznjak, Šmel'ev (2000 : 111), ces préverbes en *po-* « décrivent une 'portion' d'action jugée petite, et limitée par le temps durant lequel cette action se déroule ». Les verbes susceptibles de recevoir PO-DEL constituent une liste ouverte, de sorte que très tôt l'attention des chercheurs s'est portée moins sur l'inventaire des unités concernées que sur le type de procès désigné. Selon Mehlig (1985) adaptant au russe la classification de Z. Vendler, la valeur délimitative apparaît lorsque *po-* s'applique aux procès baptisés *activities* : *ležat'* 'être couché', *igrat' v tennis* 'jouer au tennis', etc. Sémon (1986) a redéfini cette classe de verbes pour y intégrer certains procès statiques (*byt' u kogo* 'être chez quelqu'un') : tous seraient des procès pour lesquels « sont naturelles aussi bien l'interruption que la réitération ». Dont le galop (*skakat'*), ou encore la seconde série des verbes dits « de mouvement » (cf. *infra*).

Sans rappeler ici les débats célèbres alimentés par ces définitions et classifications, notons que les notions d'inchoation et de délimitation diffèrent en fonction des unités et de leurs contextes. *Poskakali* est justement représentatif d'un ensemble de verbes compatibles avec les deux interprétations, alors que celles-ci se répartissent tendanciellement sur chacune des deux séries constitutives de la catégorie du sous-aspect.

0.3. *Po-* et le sous-aspect

Du sous-aspect relève une petite quinzaine de paires (cf. §3) réparties en deux séries parallèles : la série « déterminée » et la série « indéterminée », respectivement *nesti* et *nosit'*. Le sous-aspect occupe une place de choix dans les descriptions de *po-* qui consignent une corrélation « PO-INCH + déterminés / PO-DEL + indéterminés » qui est massive (sinon exclusive, cf. §4) :

- (2) *Vzja-l menja na ruk-i i kuda-to po-nesti^{DETERMINÉ}.PASS*
 prendre-PASS moi.ACC dans bras-PL et quelque part
 'Il m'a pris dans ses bras et m'a **emportée** quelque part.' (PO-INCH)

- (3) (au sujet d'un bracelet en cuir)
Muž nemnogo ponosi-l i peresta-l.
 mari un_peu *po-nosit'^{INDETERMINÉ}.PASS* et cesser-PASS.
 'Mon mari l'a **porté** un peu, et a renoncé.' (PO-DEL)

Notre fil rouge pour la présentation du sous-aspect sera l'article Iordanskaja, Krylosova, Mel'čuk, Mixel' (2020) qui reprend la notion d'« orientation » introduite par A. V. Isačenko. L'examen détaillé de la paire *letet'^{DETERMINÉ} / letat'^{INDETERMINÉ}* 'voler, tomber etc.' conduit les auteurs à opposer les deux séries en ces termes : « **verbes déterminés** : 'vers L1 en provenance de L2 et à un moment donné' ; **verbes indéterminés** : 'soit (a) vers L1 en provenance de L2 mais pas à un moment donné, soit (b) {« sans L1 ni L2 »}' » (Iordanskaja *et al.*, 2020 : 51). La séquence « {« sans L1 ni L2 »} » est une instruction d'effacement sémantique : « supprimer du sens du verbe les points de départ et d'arrivée du déplacement ». Voici un des exemples d'effacement donnés dans l'article, caractérisé par l'absence de direction assignée au vol de la feuille :

- (4) *Zolot-oj list dolgo leta-l nad luž-ej.*
 doré-MASC feuille longtemps *letat'^{INDETERMINÉ}.PASS* au-dessus
 flaque-INSTR
 'La feuille dorée voleta/resta longtemps au-dessus de la flaque.' (p. 41)

Selon Iordanskaja *et al.* (2020), c'est l'ajout d'un « moment donné » aux points de départ et d'arrivée qui conditionne l'emploi de la série déterminée. L'inconvénient de cette présentation réside dans les points de départ et d'arrivée qu'elle attribue à tous les verbes de cette catégorie. Cette décision entraîne

l'ajout d'une règle d'effacement *ad hoc* activée lorsque aucune direction n'est spécifiée.¹

Pour éviter la conclusion paradoxale que les verbes traditionnellement identifiés comme « verbes de mouvement » se trouveraient privés de mouvement, l'article introduit un composant interprétatif ineffaçable : toute description d'un verbe de cette catégorie mobiliserait *a priori* le verbe *dvigat'sja* 'bouger, se mouvoir' désignant « un mouvement chaotique sans points de départ et d'arrivée distingués ». Le verbe *dvigat'sja* se voit glosé au moyen de l'auxiliaire *načínat'* « commencer à » : « X commence à occuper successivement les points L1, L2, disposés sur une ligne (...) » : le contraste suggéré entre ce mouvement chaotique et une stabilité initiale semble en contradiction avec l'absence de « point de départ et d'arrivée distingués ». Cette présentation qui contient deux artifices théoriques (règle effaçable, redoublement du « mouvement ») serait-elle, de surcroît, incohérente ?

Au total, l'examen des données effectué par Iordanskaja *et al.* (2020) confirme que les notions d'orientation et d'actualisation sont primordiales pour la description des verbes déterminés et indéterminés. On peut légitimement en déduire l'hypothèse qu'elles concernent aussi la répartition de PO-DEL et PO-INCH dans la catégorie du sous-aspect et au-delà (*poskakat'*). Reste la notion problématique de « mouvement », que nous avons d'autres raisons de ne pas retenir.

¹ Il s'agit d'un écho de la présentation structuraliste empruntée à R. Jakobson par A. V. Isačenko. Citons cette thèse : « Les verbes du type *xodit'* sont le membre faible de la corrélation de mode de procès. (...) À la différence des verbes de type *idti*, les verbes de la série *xodit'* se bornent à ne pas stipuler de direction univoque du mouvement. Cela signifie que dans leurs emplois concrets, les verbes de la série *xodit'* peuvent signifier aussi bien un mouvement effectué dans une direction (*On xódit v škólu čerez park* 'Il se rend à l'école en passant par le parc'), qu'un mouvement multidirectionnel (*On xódit pó lesu* 'Il marche dans la forêt'). (Isačenko, 1960 : 311–312). Cette affirmation peut sembler en contradiction avec les statistiques montrant que les déterminés sont en moyenne les plus fréquents, cf. le tableau des correspondances statistiques extraites de la base ruscorpora, sous-corpus 1950–2014, par Larsen (2014 : 28–29). La correspondance ne s'inverse nettement au profit de l'indéterminé que pour les verbes *bresti* (3416) et *brodit'* (9646), exclus de la catégorie par les manuels sur la foi de leur différence de signification (respectivement 'aller lentement' et 'errer'). Elle s'équilibre avec un léger avantage au profit de l'indéterminé dans le cas des verbes *nes-ti* (25586) – *nosit'* (29424).

0.4. Postulats

Deux thèses qui soutiennent notre description sont en rupture avec les analyses récentes du sous-aspect :

1) Primat de la sémasiologie

Comme les publications antérieures auxquelles il se réfère (notamment Majsak & Raxilina, 1999, cf. aussi 2007), ou encore Kagan (2010), Bernitskaïa (2019), l'article Iordanskaja *et al.* (2020) part d'un postulat qui occulte les conditions de mise en place des points d'arrivée et d'origine : sont exclus de l'examen tous les emplois qui ne décrivent pas un trajet balisable. Sont qualifiés d'homonymes et laissés de côté les emplois du verbe *letet'* traduits 'tomber' en français : 'Les verres tombent de l'étagère', 'Tous les boutons tombent (ou : s'en vont)', 'Les disques durs tombent constamment en rade (ou : sautent constamment)'. Ne sont retenus que les emplois où l'unité est « un verbe de déplacement ou ses 'dérivés' métaphoriques ('voler pour ainsi dire') ». Or, aucun critère explicite ne permet de départager les emplois métaphoriques des homonymes. Pour éviter ce biais de découpage arbitraire, l'analyse partira de la distribution et des interprétations des verbes, et non d'une notion sémantique liée à des interprétations locales. C'est pour insister sur ce point que furent choisis ici *nesti* / *nosit'* appartenant aux couples transitifs du sous-aspect : la variété de ses constructions syntaxiques en fait un point d'observation privilégié des modalités de construction de l'instable.

2) Pertinence de la morphologie

Les verbes indéterminés non préfixés seront présentés comme le produit d'un processus morphologique général par opposition aux verbes déterminés primaires (cf. §3.1). Cet argument morphologique fort invite à ne pas interpréter le préfixe *in-* dans le terme « indéterminé » comme renvoyant à un manque et les verbes de cette série comme les membres « par défaut » (Kagan, 2010) du sous-aspect. Ils présentent au contraire une articulation supplémentaire dont on verra comment les verbes comme *skakat'* peuvent faire l'économie. L'analyse des déterminés *nesti-ponesti* qui entame l'étude peut à ce titre être vue comme l'analyse d'emplois verbaux d'une racine {N(')OS}² commune avec celle de *nosit'* dont les emplois seront circonscrits dans un second temps.

² Cette transcription rappelle que phonologiquement, les deux avatars de la racine ne diffèrent que par le caractère « dur » (vélarisé) ou « mou » (palatalisé) de l'initiale : /nos/ ou /nòs/. Le /s'/ mou de *nosit'* /nos'it'/ est une conséquence de l'opération de dérivation et n'appartient pas au morphème radical en propre. La translittération du cyrillique note respectivement *nos-* pour le premier (нос), et *nēs-* en syllabe accentuée, *nes-* ailleurs pour le second (нēs, нес).

1. Invariance de la racine et *nesti*

1.1. *Instabilité*

À la notion de déplacement, les analyses qui suivent substituent celle d'instabilité³ d'une entité *a*, envisagée comme la caractéristique centrale des unités relevant du sous-aspect et d'autres verbes possédant des emplois associés à l'expression de déplacements.

En l'occurrence, {N(')OS} signifie que l'entité *a* est instable car privée de site (localisateur stable de référence) *L*. On notera *L'* le localisateur contingent qui y supplée, ou son absence.

Pour prévenir tout malentendu, précisons ici-même que le terme « site » n'est pas utilisé en référence au couple « site » / « cible » introduit par Vandeloise (1986) dans son étude sur les prépositions spatiales du français. Notre relation de localisation (localisateur / localisé) ne concerne pas la conceptualisation des relations spatiales et fait référence à la théorie du repérage d'A. Culioli⁴, appliquée à des entités de langue au statut variable : termes, procès, propriétés.

Deux cas seront distingués en fonction des constructions syntaxiques :

(a) *L'* est distinct du site *L*

L' est représenté par le sujet syntaxique du verbe *nesti*. La représentation d'un « porteur » ou « support » provient de ce que *L'* n'est pas le localisateur *L* de *a* naturel, exclusif, originel, visé, légitime, etc. *L'* substitué au site *L* devient, suivant le cas, localisateur circonstanciel, auxiliaire, fortuit, temporaire, illégitime, etc. C'est ce statut secondaire de son localisateur qui rend *a* instable. Un cas particulier d'instabilité provient de la caractérisation de *L'* comme mobile. On dit alors que *a* est *porté* (*emporté, transporté, apporté*), *charrié* etc. par *L'* dans une certaine direction : *On nesët / nosit pis ma v 5 kabinet* 'Il porte les lettres en salle 5'.

(b) *L'* correspond à l'absence de site *L*

Cette interprétation est celle des structures dites « impersonnelles », sans sujet syntaxique : *nes-ët* ou *nos-it*, *litt.* : « (ça) porte ». Le verbe est par défaut à la

³ Par opposition au premier sens de « stable » identifié dans le *Petit Larousse* (édition en ligne) : « Qui repose solidement sur sa base et dans une position d'équilibre bien assurée. »

⁴ « Dans les constructions que l'on effectue, dans les relations que l'on établit, tout terme entrant dans la relation doit être nécessairement situé, c'est-à-dire être stabilisé dans un schéma par rapport à un autre terme. Être situé, trouver une stabilité, signifie avoir une assiette, ou, dans un style plus moderne, avoir un site. » (Culioli, 1990 [1987] : 119).

3^{ème} pers. du singulier dans les formes conjuguées, au genre neutre dans les formes accordées du passé. Ce schéma syntaxique ne ménage pas de place pour **L'** qui échoue à se constituer en site **L**. Le terme **a** est alors privé d'attaches, abandonné à un flux :

- (5) *Lodk-u* *nesët* *vdal'*.
 Barque-ACC *nesti*-3SG au.loin
 'La barque dérive vers le large.'

Il arrive qu'il soit impossible d'identifier **a**. (6) est un emploi impersonnel intransitif de *nesti* ; **a** correspond à un « je ne sais quoi » invisible, impalpable et diffus, flottant loin de sa source **L** :

- (6) *Ot* *n-ego* *nes-ët* [*vodk-oj*].
 de lui-GEN *nesti*-3.SG [vodka-INSTR]
 'Il sent (empeste) [la vodka].'

Le **a** en question n'est pas de la vodka, mais une odeur, émanation désincarnée qui peut être déterminée au moyen du complément à l'instrumental. Cette construction impersonnelle renvoyant à l'olfaction se retrouve avec des verbes qui pointent tantôt la source (*otdavat'* avec le préfixe *ot-* marquant le détachement), tantôt l'effet produit sur un sujet de perception (*razit'*, qui signifie également 'frapper' et s'associe à une odeur forte), tantôt ce qui charrie l'odeur (*vejat'*, cf. *veter* 'vent'), tantôt encore – si l'on suit les conjectures des dictionnaires étymologiques – des effluves ou des émanations (*vonjat'* 'puer' et *von'* 'puanteur' cf. lat. *anima* ? ; *pax-nut'* 'sentir' cf. racine *pax-/pyx-* onomatopéique ?). *Nesët* décrit quant à lui l'instabilité de **a** séparé de son site **L**. Quant au sujet de perception, il atteste la présence de l'odeur sans pouvoir la fixer.

En (a) comme en (b), l'instabilité implique l'existence du pôle de stabilité **L**. Néanmoins, c'est la prise en compte de **L'** qui prévaut, et non celle de **L** : ce que prédisent les verbes en {**N**(')**OS**} se réduit à l'instabilité de **a**⁵. C'est pour cette raison que notre définition qui rejette les points L1 et L2 de Iordanskaja *et al.* (2020), écarte aussi les notions de déplacement et de causation centrales dans les

⁵ L'instabilité est au cœur des débats sur la reconstruction des étymons de *nesti*, dont l'avatar grec *enenkein* est en relation de suppléation avec *ferô* – cf. fr. *porter*, angl. *bear* etc. Les réflexes indo-européens identifiés se partagent entre deux valeurs opposées : l'une est traduite 'porter, acheminer, soutenir' ; l'autre, 'atteindre'. Ramón (1999) suppose l'existence de deux racines indo-européennes, avec deux laryngales différentes, dont les destins se seraient tantôt croisés, tantôt auraient divergé. En synchronie, cette question sépare *nesti* / *nosit'* du français *porter* (cf. Franckel, 1992).

descriptions des verbes *nesti* et *nosit'* (cf. notamment Apresjan, 1995 : 136, 254 ; Bulynina, 2004). L'argumentation, qu'il faudra étayer, est la suivante :

- 1) Ces verbes n'impliquent pas en soi l'existence d'un causateur. Dans l'exemple ci-dessus où il n'y a pas de sujet syntaxique, s'il faut imputer une cause, celle-ci est à tout prendre l'attache originelle *L* (source de l'odeur) qui n'est pas celle que prédique *nesti*.
- 2) Le mouvement n'est qu'une manifestation parmi d'autres de l'instabilité. La comparaison de (7)–(10) corrige sur ce point l'opposition des manuels entre *nesti* jugé « dynamique », et *deržat'* '(sous)tenir' jugé « statique » :

(7) peut décrire des colonnes de manifestants en mouvement. Même en arrêt devant un barrage des forces de l'ordre, ces colonnes sont « en marche » :

- (7) 4 kolonn-y *nes-ut* *transparent-y*.
 4 colonnes-GEN *nesti*-3.PL banderoles-PL.
 'Quatre colonnes **portent** des banderoles.' (cf. la vignette du manuel Mouraviova, 1985 : 28)

Dans (8), les colonnes sont des piliers immobiles :

- (8) (La description de l'autel d'une cathédrale s'arrête sur la construction qui surplombe l'autel, un *ciborium*)
 4 malaxitovyx kolonny *nes-ut* 5-glavyj zoločennyj baladaxin⁶.
 4 colonnes de malachite *nesti*-3.PL baldaquin doré à 5 sommets
 'Quatre colonnes de malachite **supportent** un baldaquin doré à 5 sommets.'

Les colonnes, en tant qu'éléments constitutifs du *ciborium*, représentent l'assise naturelle du baldaquin. C'est *nesti* qui leur confère le statut de *L'* dans le cadre d'une description qui détaille l'édifice en un porteur *L'* (la colonnade) et un porté *a* (le baldaquin). Les colonnes n'apparaissent plus inhérentes à l'édifice, mais comme support du baldaquin ; celui-ci, en retour, figure un élément surélevé qui requiert un étayage. On dit de même *nesuščaja stena* 'mur porteur'.

Dans (9), cet emploi statique de *nesti* marque l'incommensurabilité entre un fardeau et son frêle porteur circonstanciel, par contraste avec le verbe *deržat'* '(sou)tenir' dont le sujet représente le site de stabilisation :

⁶ L'orthographe de l'original en ligne (<https://www.hometravel.ru/volga/pages/diveevo.shtml>) est conservée.

- (9) *Atlant nesët (na svoix plečax) vs-ju tjažest' vseleonn-oj.*
 Atlas **nesti**-3SG (sur ses épaules) tout-ACC poids univers-GEN
 'Atlas supporte (sur ses épaules) tout le poids de l'univers.'

- (10) *Atlant derž-it vseleonn-uju*
 Atlas tenir-3SG univers-ACC
 'Atlas soutient l'univers.'

- 3) {N(')OS} préserve la solidarité de la relation **L'-a**, condition *sine qua non* de l'instabilité qu'il décrit. Quelle que soit l'interprétation de *On nesët / nosit pis'ma v 5 kabinet* 'Il porte les lettres en salle 5', le dépôt des lettres en salle 5 reste hors cadre.

En somme, la stabilité sur laquelle fait fond le verbe *nesti* ressortit à des déterminations extérieures, qu'il convient de séparer de l'apport du verbe.

1.2. De sujet agent à sujet patient

La valeur de transit découle d'un dédoublement du site **L** en un localisateur d'origine **L_A** et un second localisateur (lieu et/ou bénéficiaire) **L_B** visé par le porteur **L'** :

- (11) *Saš-a nes-ët suxarik-i utk-am v park.*
 Sacha-NOM **nesti**-3.SG croûton-PL(.ACC) canard-DAT.PL dans
 parc(.ACC)
 'Sacha apporte du pain sec aux canards du parc.'

- (12) illustre un cas de délocalisation où seul **L_A** est identifiable :

- (12) *Saš-a nes-ët baton-y [s (xlebo)zavod-a].*
 Sacha-NOM **nesti**-3.SG baguette-PL [de usine_à_pain-GEN]
 'Sacha chaparde des baguettes (à la boulangerie industrielle).'

L'instabilité conférée aux baguettes constitue un délit, et qualifie par retour d'il-légitime le sujet **L'**, décrit par le dérivé nominal péjoratif *nesun* 'employé qui chaparde sur son lieu de travail' hérité des dernières décennies de l'URSS⁷.

⁷ En revanche, le terme commun à valeur agentive désignant tout « porteur » partage son radical avec le verbe indéterminé *nosit'* : le dérivé régulier *nositel'*, cf. §3.

Dans ces emplois, le sujet du verbe *nesti* cumule deux fonctions : il est support L' dans le cadre du fonctionnement de *nesti*, et agent engagé dans une trajectoire L_A / L_B. Dans les emplois non directionnels, seule la première fonction demeure : les sujets, investis d'une agentivité dégradée ou déniée, sont les supports circonstanciels du nom à sémantique prédicative désigné par *a*. Avec *otvetstvennost'* 'responsabilité' se construit la déclaration *Nesu otvetstvennost' [za nego]* 'je me porte garant (de lui)' qui fait du sujet un contractant formel, et non celui qui prend fait et cause et 'répond de (quelqu'un)' (*otvečat' za*). *Vaxtu* 'garde' dans *nesu vaxtu* 'je suis de garde' désigne une fonction mise en partage entre le sujet considéré et d'autres : il n'en est pas le détenteur exclusif et stable. *Gluposti* 'des bêtises' dans *nesěš' gluposti* 'tu dis des bêtises' déconsidère les paroles *a* de l'interlocuteur en les réduisant à leur évaluation négative. *Nesěm ubytki* 'nous subissons des pertes' (et non **pribyl'* 'du profit') déprend le sujet du statut visé de bénéficiaire. Responsabilité de pure forme, fonction en partage, paroles sans valeur et préjudice empêchent le sujet syntaxique de se constituer en site L (responsable, gardien, interlocuteur, bénéficiaire).

En résumé, dans les séquences commentées, aucun constituant ne remet en question l'instabilité fondamentale de *a* que prédique *nesti*. Il reste à envisager une ultime source de stabilisation, énonciative.

1.3. Déplacements et dynamique énonciative

Il faut faire la part d'un effet de stabilisation qui se manifeste dans la contextualisation des énoncés et que nous attribuons à la catégorie baptisée « aoristique »⁸. Il sera présenté au travers de deux séries d'emplois de *nesti* : un emploi personnel (I.), puis impersonnel (II.) :

⁸ Cf. Culioli (1978), De Vogüé (1995). La catégorie que désigne cet adjectif substantivé est plus large que l'emploi traditionnel du nom « aoriste » (thème en grec, une des valeurs du passé PF dans cet article, etc.). Elle se manifeste de diverses manières sur l'ensemble des unités constitutives d'un énoncé. En russe, outre les interprétations bien identifiées des verbes telles que le passé à valeur d'aoriste, le présent de narration ou les emplois performatifs, l'aoristique a aussi trait à l'interprétation des syntagmes nominaux et leur relation au verbe. Cf. dans (15) la représentation de *topor* 'hache' comme ustensile efficace, et non plus seulement utile, ni *a fortiori* menaçant.

I. *Kostja-a nes-ët topor* [Kostia-NOM *nesti*-3SG hache]

- état qualifiant le sujet :

(13) *Kostja Lebedev (...) ostalsja na ulice i styditsja vojti, potomu čto nesët topor.*

‘Kostia Lebedev (...) est resté dehors et n’ose pas entrer, parce qu’il **porte** une hache.’ [F. Dostoevskij, *Idiot*, IV, 5 (1874)]

Le port d’une hache rend Kostia Lebedev menaçant (motif justifiant son souhait de rester dehors) ; la hache s’interprète alors comme une arme potentielle.⁹

- processus en cours :

(14) – *Ura, Kostja nesët topor !*

‘– Hourra, Kostia apporte une hache !’

L’énoncé décrit Kostia en marche, pourvoyeur de la hache envisagée comme un instrument utile pour le locuteur (**L_B**).

- événement ponctuel :

(15) *Dostraivaem izbu na opuške lesa. Instrumentov ne xvataet. Nakonec-to, Kostja nesët topor. Srubaem paru lišnix sosen.*

‘Nous achevons la construction d’une cabane en lisière de forêt. Il nous manque des outils. Finalement survient Kostia qui **apporte** une hache. Nous coupons quelques pins supplémentaires.’

Le contexte droit implique la fin de l’instabilité : il y a localisation en **L_B** de la hache qui d’objet transporté devient ustensile efficace. Entre le transport de la hache et son utilisation, aucun lien causal n’est explicité. On retrouve la description par É. Benveniste de la catégorie de l’histoire : « Les événements sont posés comme il se sont produits, à mesure qu’ils apparaissent à l’horizon de l’histoire. Personne ne parle ici ; les événements semblent se raconter eux-mêmes » (Benveniste, 1959 [1966 : 241]). L’instabilité prédiquée par *nesti* de **a** n’est en soi pas remise en cause.

⁹ On reconnaît, jusque dans le détail de l’homme en arme, la catégorie qu’Aristote baptisait *ekhein* ‘être en état’ (cf. les commentaires de Ildefonse et Lallot dans leur édition du premier volume de l’*Organon* (Aristote, 2002 : 139–144).

II. **a** *nes-ët / nes-l-o* [SN.ACC *nesti*-3SG / *nesti*-PASS-N]

En l'absence de sujet syntaxique, l'instabilité de **a** tient à l'impossibilité de lui attribuer un site L. Le SN à l'accusatif nomme un **a** affecté par un flux qu'il ne contrôle pas.

– état : diarrhée

- (16) *Rebënk-a* **nesët.**
 L'enfant-ACC **nesti**-3SG
 'L'enfant a la diarrhée.'

Le verbe identifie un état d'incontinence. À comparer avec le dérivé nominal correspondant, *ponos*, qui est ambivalent : terme non comptable identifiant une pathologie (nom savant *diareja*), ou terme comptable correspondant à un accès de diarrhée.

– processus en cours : logorrhée

- (17) *Ostap-a* *neslo.*
 Ostap-ACC **nesti**-PASS-N
 'Ostap était déchaîné.' [I. Il'f, E. Petrov, *Dvenadcat' stul'ev*, I/16 (1928),
 trad. A. Préchac)

(17) décrit un comportement *hic et nunc*, un flot irrepressible de paroles.

– processus orienté : dérive

- (18=5) *Lodk-u* **nesët** *vdal'.*
 Barque-ACC **nesti**-3SG au.loin
 'La barque dérive vers le large.'

Le complément de direction dispose (18) à une insertion dans un enchaînement narratif dont la suite entérinerait un état stabilisé : la disparition de la barque.

2. *Po- et nesti*

Ponesti est plus fréquent dans les textes que *ponosit*¹⁰. Son analyse requiert une hypothèse de travail sur le préfixe.

2.1. Le paradoxe de *po-*

Nous résumerons l'invariant proposé dans Camus (1998) de la manière suivante : la préverbalisation en *po-* signifie que la signification de la base verbale n'est envisagée qu'à partir de la portion d'espace-temps où elle s'inscrit. Les théories du perfectif russe s'accordent généralement pour dire que la perfectivation introduit une limitation du procès indépendamment de sa manifestation spatio-temporelle¹¹. Le paradoxe de *po-* est de restreindre cette limitation à l'espace et au temps : la mesure de ce qui advient n'est autre que son advenue même.

Ainsi, à partir de *znat'* 'savoir ; connaître' qui prédique une propriété, est construit le préfixé *poznat'* qui transforme la connaissance en expérience. *Poznat' istinu* 'accéder à la vérité' ou *poznat' plen* 'connaître la captivité' signifient une connaissance réduite au vécu d'un sujet, réduction impossible lorsque l'objet impose son format à la connaissance : 'connaître (**poznat'*) le numéro de téléphone' (cf. Camus & Dennes, 2004).

2.2. *po-nesti*

Dans *ponesti*, l'instabilité de *a* est circonscrite par son actualisation dans le temps et l'espace et ne débouche pas sur une stabilisation. Ce point est manifeste en emplois non directionnels. L'impersonnelle *Poneslo!* (*po-nesti*-PASS-N), décrit un flot de paroles, mais pas l'épisode ponctuel d'un accès de diarrhée qui se dit *proneslo*. Le procès *ponesti ubytki* 'subir, essayer des pertes' est coextensif aux pertes constatées à un moment donné, et non rapporté à un préjudice de référence qui serait fixé de manière indépendante.

¹⁰ L'application en ligne Ngram viewer (Google) fait état, entre les formes de *ponosit'* (en cumulant les homographes *ponosit'* « porter » et *ponosit'* « réprimander ») et *ponesti*, d'un rapport de 1 à 14 pour la première personne du présent, 1 à 8 pour les formes de passé masculin, et du simple au double pour les infinitifs.

¹¹ Cf. les notions de *predel* « limite » (Vinogradov, 1947 : 493–498, Veyrenc, 1973 : 70), de totalité (Isačenko, 1960 : 130–136), ou encore de changement de situation (cf. Zaloznjak & Šmel'ev, 2000 : 134–135, en référence aux travaux de E. Padučeva).

En contextes directionnels, la particularité de *ponesti* apparaîtra bien en comparaison avec le préfixé en *ot-* marquant une séparation¹² :

- (19) *Mus-ja* *ot-nes-l-a* *kotjat* *v spal'n-ju*.
 Moussia-NOM *ot-nesti*-PASS-F chaton.ACC.PL dans chambre-ACC
 ‘Moussia a déposé les chatons dans la chambre.’

- (20) *Mus-ja* *po-nes-l-a* *kotjat* (*v spal'n-ju*) / (**v spal'nju*).
 Moussia-NOM *po-nesti*-PASS-F chaton.ACC.PL(+/*dans chambre-ACC)
 => complément de direction possible : ‘Moussia porta/porte les chatons
 (+ dans la chambre).’
 => complément de direction impossible : ‘Moussia attend/porte des
 chatons (*dans la chambre).’

Dans (19), *ot-* signifie que Moussia *L'* ne porte plus les chatons *a*, désormais localisés par le site *L_B* : *v spal'nju* « dans la chambre » ne peut pas être supprimé. Cette contrainte forte sur le complément directionnel n'existe pas avec d'autres préfixes (par ex. *u-* ou *vy-*) qui n'impliquent pas la séparation *L'-a* : l'essentiel est que ces préverbes stabilisent *a* en *L*, lequel constitue un critère distinguant la cessation du procès de la pure et simple interruption.

À l'inverse, aucune des deux interprétations de *ponesti* dans (20) ne met en jeu de stabilisation :

- Gestation : dans cet emploi sans complément de direction, le sujet reçoit une interprétation non agentive. Moussia accède au statut de gestatrice *L'*¹³, et *a* tend à ne pas être explicité par un complément : *Musja ponesla* ‘Moussia attend des petits’. La stabilisation de *a* correspondant à la mise à bas demeure au-delà de la relation constituée (au contraire de l'interprétation commune du français *portée* qui désigne les enfants mis au monde).¹⁴

¹² Cf. Dobrušina *et al.* (2001). Suivant la théorie de la préverbalisation exposée dans ce livre, *ot-*, *u-* et *vy-* cités ici font partie des préverbes qui représentent l'événement central du mot verbal en marginalisant la base préfixée.

¹³ On pourrait détailler les différences de comportement de *nesti* suivant qu'il s'agit d'ovipares (*nesti jajca* ‘pondre des œufs’), de mammifères quelconques ou d'humains. En particulier, le statut de mère accordé à un sujet humain est faiblement compatible avec le rôle secondaire de *L'* par rapport à *a*. D'où la tonalité nonchalante d'énoncés comme le suivant, pourtant attestés dans l'ancienne langue : *Nu vot, ponesla, i uže na pjatom mesjace* ‘Voilà, elle est engrossée, et déjà au cinquième mois’ (énoncé attesté, communiqué par Sv. Krylosova). L'indéterminé *nosit'* non préverbe lève cette contrainte (cf. §3).

¹⁴ De rares énoncés accordent à *ponesti* le sens ‘mettre bas’ et exigent la présence d'un complément. Jugés très familiers, ces exemples paraissent relever d'interférences linguistiques. Dans

- Transport : comme pour *nesti*, cette valeur de *ponesti* implique la visée par l'agent du site de *a* représenté par le complément directionnel *v spal'nju* 'dans la chambre'. *Po-* bloque la stabilisation de *a* et la cessation de *nesti*. En réponse aux questions « Où sont les chatons ? », « Où est Moussia ? », une réponse avec le verbe *ponesla* (valeur de parfait) se traduirait 'Moussia les porte / porte les chatons dans la chambre'. On retrouve l'observation déjà citée de Zaliznjak & Šmel'ev (2000) suivant laquelle PO-INCH en vient à « désigner l'action elle-même », en l'occurrence : l'instabilité de *a*.

2.3. Remarque : solidarité L'-a

Ponesti préserve la relation établie entre *L'* et *a*, c'est au contexte qu'incombe leur éventuelle désolidarisation *via* la stabilisation de *a* :

- (21) *Ona broсила свои цветы в канаву. Rasterjavšis' nemnogo, ja vsë-taki pod-njal ix i podal ej, no ona, usmexnuvšis', ottolknuła cvety i ja ponës ix v rukax. Tak šli molča nekotoroje vremja, poka ona ne vynula u menja iz ruk cvety, ne broсила ix na mostovuju...*

'Elle jeta ses fleurs dans le caniveau. Un peu décontenancé, je les ramassai malgré tout et lui tendis, mais elle les repoussa avec un ricanement, et je me retrouvai à les porter (*po-nesti*-PASS) moi-même. **Nous marchâmes ainsi en silence un moment, avant qu'elle ne m'arrache les fleurs des mains pour les jeter sur le pavé ...**' [M. Bulgakov, *Master i Margarita*, I/13 (1984, 8^e éd.)]

Les exemples suivants représentent trois types de contextes forçant la stabilisation de *a* :

1. Valeur distributive. Dans (22), le sujet au pluriel permet d'envisager des occurrences singulières de stabilisation sans signifier la cessation du procès *nesti* lui-même :

l'exemple suivant, une internautes ukrainienne réagit de manière ironique à un message qui s'indignait du sort accordé à une portée de chiots : *Kakoj užas, u ljudej sobaka ponesla ščenkov, i oni pytalis' ix pristroit'.* *Prosto košmar, dejstvitel'no!* 'Quelle horreur, le chien de ces gens a eu (*po-nesti*-PASS-F) des chiots, et eux, ils ont essayé de leur trouver des maîtres. C'est affreux, en effet !'. *Po-* est ici substitué au préfixe *pry-* utilisé en ukrainien et biélorussien dans ce sens (*prynesla, prynjasila*).

- (22) *Graždane Rossii masovo ponesli svoi sbereženija v banki. Po dannym CB, tol'ko za ijun' ob''ëm vkladov vyros na 500 milliardov rublej (...).* (www.m24.ru/articles/banki/22072015/79885)

‘Les citoyens russes ont massivement apporté (*po-nesti*-PASS.PL) leur épargne dans les banques. Selon les données de la Banque centrale, durant le seul mois de juin, le volume des dépôts a augmenté de 500 milliards de roubles (...).’

2. Succession narrative factuelle. Dans (23), la désolidarisation de *Ľ-a* relève de l’effet aoristique décrit plus haut :

- (23) *Snjal ja katafoty (...) i prinës domoj. Otmyl ix i ponës na kuxnju. Na kuxne v bol'soj kastrjule sdelał kipjatok, (...) I iskupal v kipjatočke katafoty.* (www.drive2.ru/l/452887191287234794/)

‘Alors j’ai enlevé les réflecteurs (...) et les ai apportés chez moi. Je les ai décrassés, puis je les ai portés (*po-nesti*-PASS) à la cuisine. Dans la cuisine, j’ai fait bouillir de l’eau dans une grande casserole, et (...) j’ai y ai fait tremper les réflecteurs.’

3. Cliché. L’exemple (24), critiqué par certains locuteurs mais largement attesté, décrit une des étapes symboliques d’un rituel intégrant la stabilisation de *a* :

- (24) *Vot prezident ponës cvety k neizvestnomu soldatu – èto konečno deševle, čem okazat' real'nuju pomošč' eščë živym starikam.* (www.kharkovforum.com/threads/etozh-kak-o-uet-nado-zla-nekhvataet.80615/)

‘Tenez, le président a porté (*po-nesti*-PASS) des fleurs sur la tombe du soldat inconnu (*litt.*: ‘au soldat inconnu’) – c’est assurément moins coûteux que d’offrir un soutien substantiel aux personnes âgées encore en vie.’

Ponesti décrit ici un geste ostentatoire (*cf.* la différence entre *dat' nož* ‘donner un couteau’ et *podat' nož* ‘tendre un couteau’) appartenant au cérémonial de dépôt de gerbe qui inclut le destinataire ‘soldat inconnu’ comme une spécification adverbiale (*cf. infra*).

2.4. Cas particulier : l'impératif injonctif

La forme d'impératif perfectif *ponesi*, lorsqu'elle a une valeur d'injonction ('porte!'), est incompatible avec un complément de direction. Cela distingue ce préfixé de *unesil* 'emporte!', *otnesil* 'dépose!', *prinesi* 'apporte!', mais aussi du simple *nesil* 'porte!' :

- (25) *Po-nes-i* *kotjat!* (**v* *spal'n-ju!*)
po-nesti-IMP chatons.ACC (*dans chambre-ACC)
 'Porte les chatons!' (*dans la chambre)

Cette particularité du perfectif *ponesi!* tient à sa composante exclusivement inter-subjective. Suivant les analyses de Culioli et Paillard (1998), l'agent est ici scindé en deux instances subjectives qui définissent deux positions relativement au procès **P** :

- validation de **P** en suspens : l'instabilité de *a* prédiquée par l'imperfectif *nesti* est le motif d'une discordance entre les sujets (insistance, attermolements, permission etc.) ;
- visée de **P** : elle répond à la stabilisation de *a* opérée par les perfectifs préfixés *unesti*, *otnesti*.

Or, *po-* conserve l'instabilité de *a* de sorte que la visée ne peut concerner que le choix du porteur **L**. À moins d'intégrer la destination comme modalité de réalisation du procès à l'instar de (24). C'est le « coup de force » qu'effectue Nabokov au moyen de l'adverbe directionnel *naverx* « en haut » dans sa propre traduction russe de son roman *Lolita* :

- (26) *Lolita (...) pripodnjav koleno, protjanula ko mne ogolënnye ruki:*
 — *Ponesi menja naverx, požalujsta. Ja čto-to v romantičeskom nastroenii.*
 'Lolita, soulevant légèrement le genou, tendit vers moi ses bras dénudés : « Porte-moi en haut, s'il te plaît. Je me sens toute romantique, ce soir »'. *Original anglais* : « Carry me upstairs, please. I feel sort of romantic tonight. » [Vl. Nabokov, *Lolita*, (en anglais 1955, en russe 1967), trad. française de E. H. Kahane]

L'effet produit par cet énoncé (tantôt rejeté, tantôt jugé « poétique » par les locuteurs consultés) pourrait être rendu en marginalisant la direction : « Porte-moi en haut *dans tes bras* » met un cliché « romantique » dans la bouche d'un enfant sous la forme d'une tournure archaïsante issue des contes merveilleux (« Tapis volant, emporte-moi [*ponesi menja*] au royaume d'Arbazam ! »).

2.5. Deux types d'inchoation

L'interprétation inchoative provient de ce que *po-* projette l'instabilité de *a* dans l'espace et le temps. Lorsque *L* n'est pas une coordonnée spatio-temporelle assignable, la manifestation de *a* équivaut à son entrée dans un état instable (préjudice, gestation, état incontrôlable : *Poneslo!* 'Le voilà reparti !'). Lorsque le site *L* représente une visée elle-même spatio-temporelle, il y a saisie partielle d'un trajet et donc commencement.¹⁵ Suivant que le contexte privilégie *L* en perspective ou *L'*, *Saša ponēs Musju na kuxnju* 'Saša porte / porta / a porté Moussia à la cuisine' signifie le lancement d'un déplacement orienté, ou constate le début d'une absence (cf. 20).

PO-DEL nécessitera la prise en compte de la série indéterminée.

3. *Nosit'* et la dérivation indéterminée

3.1. Verbes primaires vs dérivés

Les verbes déterminés appartiennent à des modèles de flexion hétéroclites représentés par des ensembles clos de verbes (*nesti*, *nesu* cf. *gresti*, *grebu* 'ramer' ; *letet'* cf. *vertet'*, *verču* 'faire tourner' etc.). Certains de ces verbes sont totalement irréguliers : *bežat'* 'courir' est un mixte unique des deux conjugaisons régulières (conj. I : *begut* ; conj. II : *bežit*) ; *exat'* 'aller, rouler' ; *idti* 'aller, marcher' et *gnat'* 'chasser' ont des radicaux alternants. À deux exceptions près (cf. *infra* *katit'*, *taščit'*), les modèles de flexion de ces verbes sont non productifs. Tous peuvent être décrits comme des verbes primaires¹⁶.

À l'inverse, chacun des verbes indéterminés suit le modèle de flexion le plus productif de l'une des deux conjugaisons régulières I et II :

- les verbes de conjugaison I suivent le modèle productif à thème en *-a(j)-* ; par ex. *letat'* 'voler, tomber', *begat'* 'courir, filer' ;
- les verbes de conjugaison II suivent quant à eux le modèle productif à thème en *-i/Ø-* ; par ex. *nosit'* 'porter'.

¹⁵ Sur le *commencement* comme existence partielle (fr. *commencement de preuve*), cf. Camus (2004 : 86–87).

¹⁶ Nous reprenons la notion de « verbe primaire » utilisée dans son acception diachronique par Regnéll (1944) qui traçait un parallèle entre la genèse de l'opposition perfectif / imperfectif et les verbes déterminés / indéterminés.

Régularité et productivité sont des caractéristiques typiques des verbes dérivés. De plus, entre modèles productifs existe une hiérarchie de conjugaisons ici strictement respectée : I>II. Les seuls déterminés relevant d'une conjugaison productive sont de conjugaison II (*katit'*, *taščit'*) ; les dérivés indéterminés associés suivent par conséquent la conjugaison I. Le tableau ci-dessous récapitule toutes les formes ; aux colonnes 1–3 des présentations familières aux slavistes sont ajoutées les colonnes 4–5 commentées plus bas :

Tableau 1.

Les radicaux de la catégorie du sous-aspect

1. DETERMINES	2. INDETERMINES	3. traductions indicatives	4. noms (thèmes non suffixés)	5. Radicaux C(x)V(x)C
Verbes primaires	Verbes dérivés en /-i-/, conj. II régulière, productive			
<i>idítí</i> / <i>šed-</i> irr. non pr., I	<i>xodít'</i>	'aller, marcher'	<i>xod</i> 'allure, marche'	XOD
<i>éxat'</i> irr. non prod., I	<i>ézdít'</i>	'aller, rouler'	<i>ezd-a</i> 'route ; équitation'	jEZD
<i>nestí</i> non prod., I transitif	<i>nosít'</i>	'porter'	<i>po-nós</i> 'diarrhée'	NOS
<i>vestí</i> non prod., I transitif	<i>vodít'</i>	'mener'	<i>za-vód</i> 'usine ; remontoir'	VOD
<i>vezítí</i> non prod., I transitif	<i>vozít'</i>	'conduire, transporter'	<i>voz</i> 'charriot'	VOZ
<i>brestí</i> non prod., I	<i>brodít'</i> (cf. note 1)	« se traîner » // « errer »	<i>brod</i> 'gué'	BROD
	Verbes dérivés en /-a-/ conj. I régulière, productive			
<i>letét'</i> non prod., II	<i>letát'</i>	'voler'	<i>po-lét</i> 'vol'	LËT
<i>bežát'</i> irr., conj. isolée.	<i>bégat'</i>	'courir, aller vite'	<i>beg</i> 'course'	BEG
<i>polztí</i> non prod., I	<i>pólzat'</i>	'ramper'	<i>za-polz</i> ¹⁷	POLZ

¹⁷ Occasionalisme. Course de bébés à Arkhangelsk : « Les vainqueurs de la compétition furent déclarés à l'issue de quatre épreuves de course (de rampement) [*četyry zabega (zapolza)*] ».

Tableau 1 (Continuation)

1. DETERMINES	2. INDETERMINES	3. traductions indicatives	4. noms (thèmes non suffixés)	5. Radicaux C(x)V(x)C
<i>lezt'</i> non prod., I	<i>lázat'</i> (ou <i>lázit'</i> II)	'grimper, passer, aller'	<i>laz</i> 'orifice, chatière'	LAZ
<i>plyt'</i> non prod., I	<i>plávat'</i>	'nager, naviguer, flotter'	<i>plav</i> / <i>-plyv</i>	PLAV
<i>gnat'</i> irr., II transitif	<i>gonját'</i>	'faire avancer'	<i>gon</i> 'rut'	GON
<i>katít'</i> prod. , II transitif	<i>katát'</i>	'faire rouler'	<i>za-kát</i> 'coucher de soleil'	KAT
<i>taščít'</i> prod. , II transitif	<i>taskát'</i>	'porter, tirer'	<i>tásk-a</i> 'punition'	TASK

L'examen des colonnes 1 et 2 suffit à se convaincre qu'aucune règle unifiée ne saurait déduire les indéterminés des déterminés. La dérivation s'effectue à partir d'une liste de radicaux normalisés qui sont tous de forme canonique C(x)V(x)C où x représente une éventuelle consonne de moindre sonorité, par ex. : PLAV, JEZD¹⁸ (colonne 4). Ces séquences sont disponibles pour toute opération de dérivation, y compris en l'absence de suffixe autre que les flexions nominales (colonne 3). Le radical NOS- du verbe indéterminé NOS-*i-t'* ne figure pas sous sa forme nue : le substantif *nos* 'nez' est étymologiquement relié à fr. *nez*, all. *Nase*, angl. *nose*. Néanmoins, le radical NOS- se déduit d'une riche famille de composés et dérivés de diverses catégories où les substantifs dominent. En voici un bref échantillon :

<i>nos-k-a</i>	NOS-SUFF-FLEXION 'usage, port'
<i>nos-k-ij</i>	NOS-SUFF-FLEXION (adjectif) 'inusable'
<i>noš-a</i>	NOS+ alternance s/š-FLEXION 'charge, fardeau'
<i>noš-en-i-e</i>	NOS+ alternance s/š-SUFF-SUFF-FLEXION 'charge, fardeau'
<i>nos-i-tel-'</i>	NOS-voyelle thématique-SUFF-FLEXION 'vecteur, porteur, locuteur'

¹⁸ Ce fonctionnement morphologique n'est pas isolé en russe. Un procédé analogue de façonnage de radical monosyllabique fournit des prénoms diminutifs susceptibles à leur tour d'être suffixés. À partir d'*Aleksandr* et *Aleksandra* voire d'autres prénoms plus rares, est construit *Saš-*, base de *Saša* et de ses multiples correspondants suffixés *Saška*, *Sašen'ka*, *Sašulja*, *Sašura* etc. La différence avec *nos-* est que les radicaux de ces diminutifs sont proliférants : à côté de *Saš-* existent *San-* ou *Saj-* (*Sanja*, *Saja*), ou encore *Šur-* (*Šur* < *Sašura*).

<i>nos-i-l-k-i</i>	NOS-voyelle thématique-SUFF-SUFF-FLEXION ‘brancard’
<i>za-nos-Ø</i>	PREF-NOS-FLEXION 1. ‘dérapage’ ; 2. ‘usure’
<i>po-nos-Ø</i>	PREF-NOS-FLEXION ‘diarrhée ; accès de diarrhée’
<i>po-nos-k-a</i>	PREF-NOS-SUFF-FLEXION ‘(objet de) rapport’ (terme de dressage canin)

L’adaptation aux modèles verbaux productifs n’est pas l’apanage du sous-aspect : il est une caractéristique de tous les verbes dérivés, en particulier :

- fréquentatifs (et assimilés) en *-a* :
byt’ ‘être’ (irrégulier) – *byvát’* ; *videt’* ‘voir’ (non productif) – *vidát’*
- dérivés aspectuels (perfectif => imperfectif) en *-a* (avec ou sans suffixe *-iv-*) :
lišit’^{IPF} ‘priver de’ conj. II – *lišat’*^{IPF} conj. I

Enfin, les deux séries du sous-aspect se prêtent à la dérivation aspectuelle¹⁹. Dans l’exemple suivant, le radical *NOS-* est sollicité dès qu’intervient un processus dérivationnel, quelle que soit la série concernée ; son allomorphe *NAŠ-* est prévisible à partir des modalités d’application de la dérivation suffixale en jeu :

Tableau 2.
NOS- en dérivation aspectuelle et sous-aspectuelle

	Imperfectif	Perfectif préverbé	Imperfectifs dérivés
série DÉT. :	<i>nesti</i> ^{IPF}	<i>za-nesti</i> ^{PF} « porter, déposer... »	<i>za-NOS-it</i> ^{IPF}
série INDÉT. :	dérivé <i>NOS-it</i> ^{IPF}	<i>za-NOS-it</i> ^{IPF} « user (en portant) »	<i>za-NAŠ-iv-at</i> ^{IPF}

Il s’ensuit que la distinction entre déterminés et indéterminés n’est pas marquée dans les dérivés non verbaux, cf. *(po)nesti* – *poNOS* ‘diarrhée’. Le composé *zaNOS* est susceptible de renvoyer à deux perfectifs préfixés : le déterminé *zanes-ti* (et son dérivé imperfectif *zaNOSit’*) pour signifier par exemple l’« embardée » d’une voiture ; l’indéterminé *zaNOSit’* (et son dérivé imperfectif *zaNAŠivat’*) au sens d’ ‘usure’.

¹⁹ Il est établi que les verbes déterminés et indéterminés ne dérogent pas aux règles communes de perfectivation et imperfectivation. Cf. par exemple Isačenko (1960 : 325–344), Veyrenc (1973 : 98–100 ; 1980 : 198–199), Garde (1980 : 383–390), Anna Zaliznjak (2024). L’absence de dérivés imperfectifs pour certains indéterminés préfixés, qui tient au fonctionnement de la préfixation verbale (« prévervation »), n’affecte pas ce principe.

Retenons que les radicaux des verbes indéterminés sont structurellement aptes à former des noms sans suffixe dérivationnel supplémentaire : les indéterminés peuvent être décrits comme dénominatifs²⁰. *A contrario*, les radicaux des verbes déterminés présentent des potentialités dérivationnelles réduites qui témoignent de leur affinité pour la prédication verbale. Les dérivés *nesun* ‘chapeur’ ou *nesenie* ‘le fait de porter’ sont des déverbatifs et renvoient exclusivement aux valeurs du déterminé *nesti*.

3.2. Dérivation indéterminée et double instabilité

Ce n’est donc pas à son radical qu’un dérivé indéterminé doit son caractère verbal, mais bien à l’opération de dérivation dont il procède. Son radical, quant à lui, permet de nommer une notion. À l’écart entre ce dont on parle et ce qu’un prédicat verbal en dit, la liste des radicaux indéterminés ajoute celui résidant entre, d’une part, la singularité d’un dire et, d’autre part, la généricité des notions qu’ils nomment.²¹

C’est cet écart entre notion et occurrence singulière qu’exploite Fontaine (1983). Après une analyse des verbes déterminés en terme d’actualité (le « moment donné » de Iordanskaja *et al.*, 2020 : 96, 106), la série indéterminée est caractérisée par la désactualisation qu’elle opère : « Sa fonction est d’abstraire l’idée de l’action (...), de signifier, pour un procès donné, l’acte générique correspondant lexicalement à ce procès (...) ». Sont ensuite présentés trois « degrés de désactualisation » du plus au moins général, repris ci-dessous pour ordonner nos observations sur *nosit’* et quelques renvois à des débats pertinents pour notre propos.

1) L’« aptitude fonctionnelle » (p. 97) correspond à une qualification du sujet et a suggéré la dénomination générale de « mouvement de fonction » pour les indéterminés (Veyrenc, 1980). *Saša nosit pis’ma v 5 kabinet* ‘Sacha porte les lettres en salle 5’ peut désigner l’emploi (subalterne) de Sacha. En dehors des contextes directionnels sur lesquels se concentre J. Fontaine, l’instabilité de *a* en fait un attri-

²⁰ L’origine dénominale en diachronie est promue notamment par Nichols (2010). Son raisonnement, fondé sur les attestations de lexèmes nominaux dans les anciens textes, n’est pas assuré pour la totalité des dérivés. Cela n’affecte pas la haute probabilité de ce traitement, ni surtout son profit pour une approche synchronique : une relation de dérivation n’opère pas toujours à partir de lexèmes attestés.

²¹ Cf. De Vogüé (2004) sur le français *filer* / *fil*. En l’absence de différentiel morphologique analogue aux dérivés indéterminés russes, cet article n’aborde pas l’ambivalence des verbes dérivés. Cependant, de nombreux verbes des 1^{er} et 2^{ème} groupes des manuels de français semblent partager cette ambivalence. La différence posée ici entre les indéterminés et les déterminés n’est pas sans rappeler, *mutatis mutandis*, ce qui distingue *solutionner* (< solution) de *résoudre*.

but distinctif : *habitus* de parure dans *on nosil borodu* 'il portait la barbe' (mais pas **golubye glaza* 'yeux bleus' : la contingence exclut la possession inaliénable) ; signe caractéristique : *Oleni nosjat roga* 'Les rennes portent des bois' ; marque singularisant son support : *Mebel' nosit otpečatok stariny* 'Ces meuble portent le cachet de l'ancien'²². *Nosit'* désigne également la maternité : *Saša nosila rebënka ot milliardera* 'Sacha portait un enfant d'un milliardaire' ; elle s'accompagne alors de la mention du père qui relègue la mère au statut secondaire de **L'**, et rejaillit sur la qualification de l'enfant : enfant de milliardaire, de maniaque, du voisin, etc.

2) Par « avatar de l'acte générique » (p. 98), Fontaine entend un « degré inférieur de la désactualisation » où il s'agit « non pas de donner à voir, mais seulement de mentionner l'activité ». Cette rubrique contient notamment les emplois négatifs où aucun déplacement n'intervient. (27) illustrera la notion plus discutée d'aller-retour utilisée à des fins pratiques dans les manuels (Mouraviova, 1985 [1974] : 50 ; Roudet, 2016 : 156) :

- (27) *Kogda menja posadili, on sotrjasën byl, pisal mne boevuju xarakteristiku polučše, nosil komdivu na podpis'. Demobilizovavšis', on eščë iskal čerez rodnyx – kak by mne pomoč'.*

'Lorsqu'on m'arrêta, il fut profondément secoué, il rédigea sur moi le meilleur rapport qu'il put, alla (*nosit'*-PASS) le faire signer par le commandant de division. Une fois démobilisé, il chercha encore, par l'intermédiaire de parents, de quelle manière il pouvait m'aider'. [A. Solženicyn, *ARXIPELAG Gulag*, I/4 (1990), trad. J. Johannet, J. Lafond, légèrement modifiée].

Padučeva (1996), Bernitskaïa (2007) et Kagan (2010) démontrent l'inadéquation de la notion d'aller-retour. L'idée de retour représente une inférence due au point de vue rétrospectif sur une action réversible (Cf. « Quelqu'un a allumé la lampe hier » => fut-elle éteinte ?). L'indéterminé est ici employé dans la valeur de mention d'un fait (« *obščefaktičeskoe* ») de l'imperfectif passé. L'enjeu de (27) n'est pas de localiser la lettre, mais de décrire l'action du chef : loin de rester inactif, celui-ci a porté en personne le rapport pour signature. Le choix de l'indéterminé *nosit'* subordonne le trajet à son motif : recueillir en mains propres la signature du commandant.

3) Le « degré de désactualisation le plus faible » (p. 104) est analysé comme une présentation globale d'un procès attesté sous forme d'occurrences singulières :

²² Ce dernier emploi possède un correspondant avec le déterminé *nesti*.

(28) *Eščë so vremën sovxoza*

vod-u **nos-jat** sjuda iz kolodc-ev.

eau-ACC **nosit'**-3.PL ici (direction) en_provenance_de puits-GEN.PL

‘Depuis l’époque du sovkhoze, rien n’a changé ici : l’eau des puits est **ache-minée** à bras d’hommes.’

L’itération n’est pas un critère suffisant pour le choix de l’indéterminé (cf. Garde, 1965 : 181–182 ; puis Veyrenc, 1980 : 181–184). Même répété, coutumier ou stigmatisant, le chapardage sur son lieu de travail de l’exemple (12) impose la série déterminée. L’indéterminé neutralise la singularité des occurrences spatio-temporelles (les actes de rapine qui font un chapardage) au profit de la sémantique du radical : (28) signifie que l’eau n’est acheminée ni par tuyaux, ni par citernes, mais portée à bras nus.

Comparons la manière dont le déterminé *nesti* et l’indéterminé *nosit'* mettent en scène l’instabilité de **a** :

- avec *nesti*, l’instabilité de **a** est circonstancielle et ne débouche pas sur une qualification du sujet **L'** indépendante de **a**. En témoignent les effets de désagentivisation décrits pour la structure Sujet-Complément : un garant (*nesti + otvetstvennost'* ‘responsabilité’) endosse une responsabilité de circonstance qui, serait-elle même contractuelle, n’est ni une vocation, ni un engagement en conscience ; un mur est qualifié de porteur (*nesuščaja stena*) dans une configuration particulière qui n’altère pas sa fonctionnalité intrinsèque de cloison. La prise en compte de **L** dans les emplois directionnels (syntagmes prépositionnels ou adverbiaux) fait du sujet la source d’une visée indépendante, mais *nesti* le décrit uniquement comme localisateur intermédiaire et temporaire.
- le dérivé *nosit'*, en plus d’opérer la mise en scène de l’instabilité de **a** dans le cadre de la proposition, nomme le scénario en jeu sous l’espèce de la notion NOS-, et redistribue les rôles sur les constituants de la proposition. En l’absence de **L** assignable dans le temps et l’espace, le sujet **L'** devient porteur « patenté » – de barbe, armes ou cornes –, ce qui bloque les emplois où prime la circonstancialité (cf. les noms à sémantique prédicative : **nosit' otvetstvennost'*) ; dans le cas contraire, **L'** est investi d’une fonction ou d’une tâche définies par la visée de **L**. Enfin, le verbe lui-même décline de diverses manières ce qui distingue son radical de ceux des autres verbes transitifs du tableau I : porter dans ses bras, à mains nues etc.

On comprend mieux l’insistance avec laquelle les manuels font de l’itération une valeur exclusive des indéterminés malgré l’inexactitude notoire de cette description. Il y a bien itération, mais celle-ci n’est pas dans la valeur référentielle des énoncés. Elle réside dans la manière de construire la prédication, qui mobi-

lise doublement l'instabilité : d'une part pour construire la proposition verbale, d'autre part en nommant cette mise en relation.²³

Les indéterminés sont généralement associés à PO-DEL, qui peut à présent être examiné.

4. PO-DEL

Les locuteurs rejettent les tentatives suivantes de traduire 'Il a un peu porté la hache, et puis...', ou 'Finis de porter...' au moyen du verbe déterminé *ponesti* :

- (29) ?*Vsë, ponës (topor), i xvatit!*
 Tout, *po-nesti*-PASS (hache), et ça suffit

- (30) ?*On ponës(-ponës) (nemnogo) topor, i otdal ego Bore.*
 On *po-nesti*-PASS(--REDUPL) (un peu) hache, et l'a passée/
 rendue à Boria.

Comme le suggérait (1b) *Poskakali – xvatit!* « Assez galopé, ça suffit ! », PO-DEL marque une cessation étayée par un jugement. Suivant les termes de Sémon (1986), « l'acte finit par être congru à un désir, un modèle, une durée ou une intensité nécessaires au surgissement de quelque conséquence ». Mais plutôt que la « conséquence » à proprement parler, l'enjeu est l'absence de **P**, en l'occurrence : son interruption. Soit parce que **P** a duré plus qu'il ne le devait ou méritait (1b : « c'est assez ! »), soit parce qu'au contraire **P** aurait dû se poursuivre : *Vot i vsë. Poskakali i ušli...* 'C'est tout. Deux petits tours de pistes, et les voilà partis'.

Cette opération est *a priori* incompatible avec le déterminé *ponesti* qui se borne à prédiquer l'existence instable de **a**. L'incompatibilité est levée lorsque l'absence de **P** représente non pas son interruption, mais figure une alternative. Ainsi, c'est bien PO-DEL qui apparaît dans l'injonction suivante :

- (31) *Daj ponesti topor!*
 Laisse *po-nesti* hache !
 'Laisse-moi **porter** la hache !'

²³ Sur cette mise en abyme de la prédication comme potentiel de la *lexis* culiolienne (matrice de gloses), et l'hypothèse de sa mise en œuvre dans la dérivation, cf. De Vogüé (2011), ch. 2.5.

Il s'agit alors d'un procès qui, ne durant que le temps qu'il dure, ne portera justement pas à conséquence.

La série indéterminée, quant à elle, confère en soi au procès une plénitude notionnelle en plus de son advenue dans le temps. *Po-* assigne un *quantum* d'espace-temps à une notion. C'est ce contraste qui engendre l'effet de restriction souhaitée ('c'est assez!') ou regrettée ('c'est tout?'). Les degrés de désactualisation de J. Fontaine ne sont donc plus valides, puisque *po-* nécessite une actualisation. L'*habitus* de parure devient un caprice temporaire :

(32=3) (au sujet d'un bracelet en cuir)

<i>Muž</i>	<i>nemnogo</i>	<i>ponosi-l</i>	<i>i</i>	<i>peresta-l.</i>
mari	un_peu	<i>po-nosit'</i> -PASS	et	cesser-PASS.

'Mon mari l'a **porté** un peu, et a arrêté.'

L'emploi en fonction de mention d'un fait est incompatible avec la représentation d'une durée, alors que la valeur globale oppose bien la notion à une pluralité d'occurrences. Par conséquent, (33) est impossible alors que (34) est proche de (32=3), *modulo* la classe des occurrences et la pluralisation du complément :

(33) *On ***po-nosi-l*** *xarakteristik-u komdiv-u* *na*
podpis'.
 Il ***po-nosit'***-PASS rapport-ACC.SG commandant-DAT.SG pour
 signature

(34) *Mesjac-dva on* ***ponosi-l*** *xarakteristik-i komdiv-u*
na *podpis', potom brosil, rešil podpisivat' sam.*
 mois-deux il ***po-nosit'***-PASS rapport-PL commandant-DAT
 pour signature (...)
 'Un ou deux mois il **porta** des rapports à signer au commandant de division, puis abandonna, décidant de les signer lui-même.'

Au balancement syntaxique de (32) et (34) répond le schème lexico-syntaxique *Po-V (i) xvatit* litt. 'Po-V (et) ça suffit', rappel menaçant de la fugacité et de la vanité des choses ici-bas. Il entre en écho avec l'instabilité prédiquée par *nosit'* dans cet aphorisme imaginé au pied d'une guillotine qui transforme l'inaliénable en aliénable : *Ponosil i xvatit!* 'Tu l'as assez portée !' (V. Afončenko), ou par *skakat'* dans notre titre.

5. Poskakali i xvatit!

Skakat' représente la série ouverte de verbes qui prédisent l'instabilité d'un terme *a* tout en échappant à l'opposition de sous-aspect. Parmi ceux compatibles avec PO-INCH et PO-DEL, citons aussi *pošagat'* 'faire des pas', *posemenit'* 'trotter', *pokovyljat'* 'marcher en chancelant', *poskol'zit'* 'glisser', *potrusit'* 'trotter', *poprygat'* 'sauter', etc. Comparer (35) et (36) :

- (35) On *potrusi-l* *sled-om* *za* *n-ej*.
 Il *po*^{INCH}-trotter-PASS suite-INSTR derrière elle-INSTR
 « Il la suivit en trottant. »

- (36) On *potrusi-l* *tri* *minut-y*
 Il *po*^{DEL}-trotter-PASS trois minute-GEN
 'Il trotta trois minutes.'

Ces verbes identifient certes plusieurs « manières de se déplacer » (Beliakov & Stosic, 2020), mais ont aussi en commun leur pittoresque. Ce qui les rapproche d'autres verbes évoquant des impressions sensorielles produites par un déplacement : *šarkat'* 'faire du bruit avec ses pieds' ou *šuršat'* suscitant des impressions auditives et tactiles (tissu ou papier froissé, murmure ou chuchotement, grésillement ou grincement léger) :

- (37) *Pružina v časax pošurša-l-a-pošurša-l-a i*
sta-l-a.
 Ressort dans montre *po*^{DEL}-šuršat'-PASS-F~ REDUPL et
 stopper-PASS-F
 'Le ressort de la montre chuchota un temps, puis s'arrêta.' [D. Mamin-Sibirjak, *Dikoe sčast'e*, XV (1884)]

- (38) *S peril verand-y pauk pošuršal*
čerez stol.
 de rambarde.GEN veranda-GEN araignée *po*^{INCH}-šuršat'-PASS
 à-travers table
 'Quittant la rambarde de la véranda, l'araignée s'engagea sur la table dans un froufroutement.'

Leur force expressive dispose ces verbes au fonctionnement de PO-DEL qui confronte une représentation notionnelle à ce qui advient dans l'espace et le temps, représentation susceptible de réélaborations :

(39=1b) *Poskaka-l-i* – *xvat-it.* *Pora* *umirat'*
Po-skakat'-PASS-PL – suffire-3SG *il_est_temps* mourir
 'Assez galopé, fini ! C'est le moment d'aller mourir.'

Cet énoncé est un message déposé sur un forum russophone en avril 2024, commentaire sarcastique d'une vidéo montrant l'enlèvement d'un homme dans la rue par des militaires dans la région de Dnipro, probablement dans le but de l'envoyer au front.

L'emploi délimitatif active la valeur notionnelle de *skakat'*. Les paramètres abstraits attribués à *skakat'* par l'analyse en traits lexicaux de Beliaikov et Stosic (2020) – « [vitesse] », « [schéma moteur] », « [moyen de locomotion] » – trouvent dans ce cas à s'incarner. Le verbe évoque alors le substantif suffixé *skakun*, litt. 'étalon, cheval de course' où, le schème lexico-sémantique *Po-V (i) xvatit* 'Po-V (et) ça suffit' aidant, on peut entendre la valeur péjorative qu'il prenait dans *nesun* 'chapeardeur'. Mais c'est surtout une référence à un célèbre soldat de l'armée ukrainienne, Vitalii Skakoune (1996–2022), dynamitant un pont au prix de sa vie pour retarder l'offensive ennemie. Il symbolise la bravoure et la résilience des Cosaques ici rapportées à la fugacité du monde, d'où la traduction : 'Fin de faire/jouer les Skakoune'.

6. Conclusion et envoi

Nesti, *nosit'* et *skakat'* admettent des compléments de direction non parce qu'ils seraient des « verbes de déplacements » (*nesti* et *nosit'* outrepassent cette valeur), mais parce que leur propriété commune est de construire l'instabilité d'un terme *a*. De son côté, le préfixe *po-* inscrit *a* instable dans l'espace et le temps, ce qui induit également une idée d'impermanence, et se solde par une interprétation « inchoative » lorsque le pôle de stabilisation est spatio-temporel (*quantum* de procès), ou « délimitative » lorsqu'intervient un jugement sur une représentation notionnelle du procès (point de vue qualitatif).

Cette seconde valeur de *po-* sollicite le fonctionnement double de la série indéterminée des verbes du sous-aspect, qui sont des dérivés adnominaux :

ils installent le scénario d'instabilité dans la syntaxe de la proposition verbale, et ils en nomment la notion générique qui rejaillit sur tout l'énoncé. Un autre effet d'« écho » interprétatif se fait jour avec les verbes étrangers au sous-aspect (*skakat'*, *trusit'*, *šuršat'*, etc.) : leur force expressive les rend accueillants au bruit du monde, qu'ils ne nomment pas mais imitent ou reproduisent. À preuve, la manière dont *Poskakali!* parvient, à partir d'un simple radical onomatopéique, *via* le balancement d'un schème lexico-syntaxique, à faire surgir la figure singulière de Vitalii Skakoune.

Nommer et reproduire le monde sont assurément des entreprises différentes, mais peut-être peuvent-elles se rejoindre. Le radical de dérivation NOS-, outre sa valeur structurale définie au sein d'un petit paradigme clos, possède son identité propre, qui n'est pas cantonnée aux verbes indéterminés mais ouverte à toutes les chicanes. Il est alors tentant de retrouver la piste des fausses étymologies associant le radical NOS- au substantif *nos* 'nez', et d'ajouter l'exemple suivant à la « nosologie » (V. V. Vinogradov) suscitée par les lectures de la nouvelle éponyme de N. Gogol :

- (40) *No skol'ko raz ni **podnosil** on ego na ego že sobstvennoe mesto, staranie bylo po prežnemu neuspešno.*

'Mais il avait beau réitérer les tentatives de le [son nez] **remettre** (*podnosi-l* préfixe-NOSit'-PASS) à sa place légitime, ses efforts restaient vains.'
[N. Gogol, *Nos*, II (1835)]

Podnosil est verbe imperfectif obtenu par dérivation à partir d'un verbe déterminé (*podnesti*). Le coq-à-l'âne lexical suggéré revient en somme à donner chair au rapprochement morphologique constaté entre deux catégories grammaticales : l'aspect et le sous-aspect. Il rejoue deux fois l'instabilité prédiquée par NOS- : au titre du radical du verbe *podnosit'* « porter à », mais aussi du nez (*nos*) qui a quitté l'emplacement stable qui était le sien au milieu du visage de l'assesseur de collègue Kovaliov.

Liste des abréviations

ACC – accusatif

conj. – conjugaison

DAT – datif

F – féminin

GEN – génitif

IMP – impératif
INSTR – instrumental
irr. – irrégulier
litt. – littéralement
M – masculin
N – neutre
NOM – nominatif
PL – pluriel
PART – particule
PASS – passé
PREP – prépositif
REDUPL – reduplication
prod. – productif
SG – singulier
SN – syntagme nominal

Références citées

Sauf mention d'auteur, les exemples cités sont des énoncés attestés oralement ou en ligne. Tous ont été soumis à plusieurs locuteurs natifs.

- Apresjan, Ju. (1995 [1970]). *Izbrannye trudy, t.1. Leksičeskaja semantika*, 2^{ème} édition. Jazyki russkoj kul'tury.
- Aristote (2002). *Catégories*. Présentation et traduction de Fr. Ildefonse & J. Lallot. Seuil.
- Beliakov, Vl. & Stosic, D. (2018). Les verbes exprimant la manière de se déplacer en russe, *Revue des études slaves* 89(1–2), 55–71
- Benveniste, É. (1959). Les relations de temps dans le verbe français. *Bulletin de la société de linguistique* 54(1), 69–82. Cité d'après la réédition *Problèmes de linguistique générale I*, Gallimard, 1966, 237–250.
- Bernitskaïa, N. (2018). *Vinni Pux xodil v gosti k Pjatačku. Vernulsja li on domoj?* Problème de l'aller-retour. In Vl. Beliaikov & Chr. Bracquenier (éds), *Contribution aux études morphologiques, syntaxiques et sémantiques en russe* (43–56). Presses universitaires du Midi.
- Bernitskaïa, N. (2019). O grammatičeskoj opozicii glagolov dviženija tipa 'iditi / xoditi' v russkom jazyke. *Voprosy jazykoznanija* 1, 75–92.
- Bulynina, M. (2004). *Glagol'naja kauzacija dinamiki sintaksičeskogo koncepta (na materiale russkoj i anglijskoj leksiko-semantičeskix grupp glagolov peremeščenija ob'ekta)*. Résumé de thèse pour le grade de docteur ès lettres. Voronež.

- Camus, R. (1998). Quelques considérations sur le préfixe *po-* en russe contemporain. *Revue des études slaves* 70(1), 101–112.
- Camus, R. (2004). Quelques aspects de *commencer*. In R. Camus & S. De Vogüé (éds), Numéro thématique « Variation sémantique et syntaxique des unités lexicales : étude de six verbes français » (81–101). *Linx*, 50.
- Camus, R. & Dennes, M. (2004). *Poznat', poznanie* : quand la connaissance devient une expérience. In B. Cassin (éd.), *Vocabulaire européen des philosophies* (136–138). Seuil/Le Robert.
- Culioli, A. (1978). Valeurs aspectuelles et opérations énonciatives : l'aoristique ». In J. David & R. Martin (éds), *Actes du colloque sur la notion d'aspect* (181–194). Centre d'analyse syntaxique de l'Université de Metz (repris dans A. Culioli, *Pour une linguistique de l'énonciation*, T. 2. Ophrys, 1999, 127–143).
- Culioli, A. (1986). Formes schématiques et domaine. *BULAG* 13, 7–16 (cité d'après la réédition dans *Pour une linguistique de l'énonciation*, T. 1. Ophrys, 1990, 115–126).
- Culioli, A. & Paillard, D. (1987). À propos de l'alternance imperfectif/perfectif dans les énoncés impératifs. *Revue des études slaves* 59(3), 527–534.
- De Vogüé, S. (1995). L'effet aoristique. In J. Bouscaren, J.-J. Franckel, St. Robert (éds), *Langues et langage. Problèmes et raisonnement en linguistique. Mélanges offerts à Antoine Culioli* (247–259). Presses universitaires de France.
- De Vogüé, S. (2004). Syntaxe, référence et identité du verbe *filer*. In R. Camus & S. de Vogüé (éds), Numéro thématique « Variation sémantique et syntaxique des unités lexicales : étude de six verbes français » (135–167). *Linx*, 50.
- De Vogüé, S. (2011). Les principes organisateurs de la variété des constructions verbales. *Revista virtual de estudos da linguagem* 9(16), 276–315.
- Dobrušina, E., Mellina, E. & Pajjar, D. [= Paillard, D.] (2001). *Russkie prístavki: mnogoznačnosť i semantičeskoe edinstvo* (sbornik). Russkie slovari.
- Fontaine, J. (1983). *Grammaire du texte et aspect du verbe en russe contemporain*. Institut d'études slaves.
- Franckel, J.-J. (1992). Le sens hors de portée. *Cahiers de lexicologie* 61, 18–39.
- Garde, P. (1965). Verbes indéterminés et verbes itératifs dans les langues slaves. *Slavia* 34(2), 177–189.
- Garde, P. (1998 [1980]). *Grammaire russe. Phonologie et morphologie*. 2^{ème} éd. Institut d'études slaves.
- Išačenko, A. (1960). *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim*, Morfologija, č. 2. Izdatel'stvo slovackoj akademii nauk.
- Iordanskaja, L., Krylova, Sv., Melčuk, I. & Mixel', P. (2020). Kategorija napravlenosti u russkix parnyx glagolov peremeščeniya (na primere glagolov LETET' / LETAT'). *Voprosy jazykoznaniya* 1, 27–64.

- Kagan, O. (2010). Aspects of motion. On the semantics and pragmatics of indeterminate aspect. In V. Hasko & R. Perelmutter (éds), *New approaches to slavic verbs of motion* (47–65). John Benjamins.
- Larsen, A. (2014). *Bespristavočnye glagoly russkogo jazyka v prjamom i perenosnom značenii v aspekte prepodavanja norvežskim učaščimsja*. Universitetet i Tromsø.
- LeBlanc, N. L. (2010). *The polysemy of an “empty” prefix: A corpus-based cognitive semantic analysis of the Russian verbal prefix PO*. Thèse de doctorat. University of North Carolina.
- Majsak, T. & Raxilina, E. (1999). Semantika i statistika: glagol *idti* na fone drugix glagolov dviženija. In N. Arutjunova & I. Šatunovskij (éds), *Logičeskij analiz jazyka: jazyki dinamičeskogo mira* (53–67). Meždunarodnyj universitet «Dubna».
- Majsak, T. & Raxilina, E. (éds) (2007). *Aquamotion – Glagoly dviženija v vode: leksičeskaja tipologija*. Indrik.
- Mouraviova, L. (1985) [1974]. *Verbes de mouvement en russe*. Langue russe.
- Nichols, J. (2010). Indeterminate motion verbs are denominal. In V. Hasko & R. Perelmutter (éds), *New approaches to slavic verbs of motion* (47–65). John Benjamin.
- Paillard, D. (2004). À propos des verbes préfixés. *Slovo* 31, 13–44.
- Ramón, J.L.G. (1999). Zur Bedeutung indogermanischer Verbalwurzeln: $*h^2nek^{\wedge}$ “erreichen, reichen bis”, $*h^1nek^{\wedge}$ “erhalten, (weg)nehmen”. In J. Habisreiter et al. (éds), *Gering und doch von Herzen. 25 indogermanische Beiträge Bernhard Forssman zum 65. Geburtstag* (47–80). Reichert Verlag.
- Regnéll, C. G. (1944). *Über den Ursprung des slavischen Verbalaspektes*. Verlag C. W. K. Gleerup.
- Roudet, R. (2016). *Grammaire russe*. Institut d'études slaves.
- Sémon, J.-P. (1986). *Postojat'* ou la perfectivité de congruence: définition et valeurs textuelles. *Revue des études slaves* 58(4), 609–635.
- Tauscher, E. & Kirschbaum, E.-G. (1989). *Grammatik der russischen Sprache*. Brücken-Verlag.
- Vandeloise, C. (1986). *L'espace en français: sémantique des prépositions spatiales*. Le Seuil.
- Veyrenc, Ch.-J. (1973). *Grammaire du russe*. Presses universitaires de France.
- Veyrenc, J. (1980) [1966]. Russe *idti* et *xodit'*: mouvement de déplacement et mouvement de fonction. In *Études sur le verbe russe*. Institut d'études slaves.
- Zaliznjak, A. [= Zalizniak, Anna] & Šmelëv, A. (2000). *Vvedenie v russkiju aspektologiju*. Jazyki russkoj kul'tury.
- Zaliznjak, A. (2001). Semantičeskaja derivacija v značenii russkoj pristavki *u*. *Moskovskij lingvističeskij žurnal* 5(1), 69–84.
- Zaliznjak, A. (2017). Russian Prefixed Verbs of Motion Revisited. *Zeitschrift für Slawistik* 62(1), 1–28.



Svetlana Krylosova

INaLCO, CREE

France

 <https://orcid.org/0000-0002-1324-1941>

Définir pour apprendre et apprendre à définir : repenser la lexicographie pour l'enseignement des langues

Defining to Learn and Learning to Define: Rethinking Lexicography for Language Teaching

Abstract

This article examines the pedagogical potential of the *Active Dictionary of the Russian Language*, edited by Ju. D. Apresjan and developed within Melčuk's framework of Explanatory and Combinatorial Lexicography (ECL) and Apresjan's Active Lexicography (AL). Rather than treating the *Active Dictionary* solely as a reference work, the article approaches it as a methodological tool for integrating lexis and grammar in the teaching of Russian as a foreign language. Following a concise overview of the theoretical principles underpinning ECL–AL, it shows how dictionary definitions can be adapted for pedagogical purposes. Each example highlights the controlled simplification of definitions and demonstrates how they can foster paraphrastic competence and support learners in justifying lexical choices in the target language. The approach proposed here argues for a more systematic incorporation of lexicographic methodology into language instruction.

Keywords

Explanatory and Combinatorial Lexicology (ECL), Active Lexicography (AL), *Active Dictionary of the Russian Language*, pedagogical definition, Russian as a foreign language, lexicon and grammar, Ju. D. Apresjan

À la mémoire de Ju. D. Apresjan

Introduction

Ouvrage sans équivalent dans la tradition lexicographique, le *Dictionnaire actif de la langue russe*, dirigé par Ju. Apresjan (Apresjan, 2014a, 2014b, 2017, 2023, 2024), s'appuie à la fois sur les acquis majeurs de la lexicographie russe et européenne et sur les résultats récents de la linguistique théorique. Son objectif est clair : fournir, pour chaque unité lexicale, toutes les informations nécessaires non seulement à sa compréhension en contexte, mais aussi à son emploi correct. Formes grammaticales, prosodie spécifique, définition analytique, variations contextuelles du sens, constructions syntaxiques, schémas actantiels, combinatoire lexicale, conditions pragmatiques, réseau de relations lexicales – autant d'éléments réunis dans chaque article, illustré par de nombreux exemples issus du *Corpus national de la langue russe* (ruscorpora.ru). L'ouvrage s'adresse ainsi aussi bien aux spécialistes (enseignants, traducteurs, éditeurs) qu'aux locuteurs désireux de perfectionner leur maîtrise de la langue.

Par l'ampleur de la description et l'intégration des avancées théoriques, le *Dictionnaire actif* est considéré comme une œuvre « colossale par son volume, sa précision et son importance scientifique et culturelle » (Melčuk, 2024). Pourtant, dans la littérature spécialisée consacrée à l'enseignement du russe langue étrangère (RLE), les réactions restent rares. L'une des exceptions notables est la prise de position d'Olga Glazunova, spécialiste du RLE connue pour ses ouvrages *La grammaire du russe en exercices et commentaires*.

En 2016, elle publie dans la revue *Vestnik SPbGU* un article intitulé « Sur le *Dictionnaire actif de la langue russe* et la description systémique du lexique du point de vue de la théorie et de la méthodologie de l'enseignement du russe langue étrangère »¹. Glazunova salue la pertinence de l'ambition : « L'idée d'une description systémique du lexique qui sous-tend ce dictionnaire est d'une actualité extrême » (Glazunova, 2016 : 121), tout en déplorant l'absence de dialogue entre deux traditions pourtant proches par leurs objectifs : la lexicographie et la didactique du RLE. Certaines questions traitées dans le *Dictionnaire actif* trouvent déjà des réponses dans les travaux des spécialistes du RLE, mais ces apports n'ont pas été intégrés, notamment en raison de l'isolement institutionnel des deux domaines (Glazunova, 2016 : 130). La reconnaissance initiale laisse donc rapidement place à une analyse plus critique : selon Glazunova, la richesse de l'ouvrage ne suffit pas à en faire un outil directement exploitable en classe. Sans médiation pédagogique, la présentation lexicographique ne prévient pas les

¹ Nous tenons à remercier Lidia Kolzun de nous avoir signalé cette publication.

erreurs typiques et ne guide pas efficacement le lecteur dans ses choix lexicaux. Pour devenir pleinement opérationnel, le matériel devrait être accompagné de commentaires théoriques et d'exercices spécifiquement ciblés sur les phénomènes combinatoires, présentés sous une forme « généralisée, logiquement structurée et accessible » (Glazunova, 2016 : 129).

Dans notre article, nous proposons de changer de perspective : plutôt que de considérer le *Dictionnaire actif* uniquement comme un ouvrage de référence au sens classique, nous avançons qu'il constitue, par la méthode de travail sur le lexique qu'il propose, un outil inédit et fécond pour l'enseignement du russe langue étrangère. Comme le rappelle Olga Glazunova, ce dictionnaire s'appuie sur la Théorie Sens-Texte², et plus précisément sur sa composante lexicale, la Lexicologie Explicative et Combinatoire (LEC) de Melčuk (Melčuk *et al.*, 1995 ; Melčuk & Polguère, 2007), mise en œuvre notamment dans les *Dictionnaires explicatifs et combinatoires*³ du russe et du français (Melčuk & Žolkovskij, 1984 ; Melčuk *et al.*, 1984–1988–1992–1999). Si Glazunova estime que ce modèle peut limiter la capacité à restituer toute la complexité du lexique, nous considérons au contraire que cette formalisation, par sa clarté et sa cohérence interne, constitue non pas une contrainte, mais un levier, ouvrant de nouvelles possibilités d'exploitation en didactique.

Les définitions élaborées selon les principes de la LEC – et donc selon ceux de la Lexicographie Active (LA) d'Apresjan⁴ – peuvent devenir un instrument métho-

² Glazunova formule ici encore la réserve suivante : « Un aspect important à souligner est que le langage du modèle *Sens-Texte* a été développé sous l'influence des recherches en traduction automatique, en analyse et en synthèse de textes, selon les principes de la description sémantique formelle, utilisant une symbolique univoque et logiquement cohérente. Le critère obligatoire ayant présidé à la création de ce code linguistique symbolique était sa possible mise en œuvre dans des applications informatiques [...]. C'est précisément là qu'apparaissent certaines limites : comme tout système, le modèle proposé simplifie la réalité et ne permet pas de décrire de manière satisfaisante la complexité structurelle et sémantique des constructions linguistiques » (Glazunova, 2016 : 126).

³ Le *Dictionnaire explicatif et combinatoire du russe* (Melčuk & Žolkovskij, 1984) a été élaboré par une équipe sous la direction d'Igor A. Melčuk et comprend 286 entrées. Le *Dictionnaire explicatif et combinatoire du français contemporain* (Melčuk *et al.*, 1984, 1988, 1992, 1999) propose au total la description formelle de 510 vocables du français. En parlant de la LEC, on ne saurait passer sous silence l'ouvrage sans précédent Iordanskaja & Paperno (1996), qui recense 140 noms des parties du corps en russe et en anglais, ainsi que *l'Introduction à la lexicographie explicative et combinatoire* (Melčuk *et al.*, 1995).

⁴ Nous utilisons ici le terme *Lexicographie active d'Apresjan* pour désigner la composante lexicographique de ce qu'Apresjan lui-même appelle *École de Moscou de sémantique, de description linguistique et de lexicographie systémique* (Apresjan, 2006 : 25 ; Melčuk & Polguère, 2007). Sur l'histoire du développement de la lexicographie active depuis le dictionnaire de Reum (1953), voir Apresjan (2010 : 19–30).

dologique puissant, servant de base à l'élaboration de définitions pédagogiques adaptées à des publics variés, y compris débutants. Le *Dictionnaire actif*, tout comme les *Dictionnaires explicatifs et combinatoires*, ne doivent pas être envisagés uniquement comme d'excellents outils prêts à l'emploi, mais également comme des méthodes à adapter. Dans cette optique, leur rigueur descriptive, loin de constituer un obstacle, devient un atout : intégrée à une démarche didactique réfléchie, elle fournit à l'enseignant un cadre solide pour construire ses propres définitions, concevoir des exercices et guider l'utilisateur dans sa progression linguistique.

Cette approche unifiée – la LEC de Mel'čuk et la LA d'Apresjan⁵ – fournit ainsi un cadre solide pour développer une méthode d'enseignement du vocabulaire orientée non seulement vers la compréhension, mais aussi vers « la production de textes » (terme d'Apresjan dans Apresjan, 2014a : 6). Nous nous appuyons ici sur notre expérience d'enseignement du russe langue étrangère à l'Inalco (Paris)⁶. Notre propos ne portera pas sur l'utilisation directe, en classe, d'articles de dictionnaire déjà rédigés⁷, mais sur la manière dont nous transposons la méthodologie de construction d'une définition lexicographique à des objectifs concrets : stimuler la prise de parole en expression orale, clarifier certains points délicats en cours de grammaire, ou encore expliquer les blocages rencontrés en version et en thème.

L'article est structuré comme suit : la première section pose le cadre théorique en rappelant les six principes fondamentaux qui régissent la définition dans la LEC-LA et en précisant leur portée pour la description lexicographique ; la deuxième section propose des modèles de définition pédagogique directement inspirés de ces principes, conçus pour répondre aux besoins de publics apprenant une langue étrangère et illustre la mise en œuvre de ces modèles à travers différents types d'activités didactiques.

⁵ Lorsque nous parlons de la LEC de Mel'čuk ou de la LA d'Apresjan, nous gardons à l'esprit que, derrière ces deux grandes figures emblématiques, se trouvent des collaborateurs, des disciples et, plus largement, une véritable école linguistique, dont les travaux ont contribué à élaborer et à enrichir ces approches.

⁶ Institut national des langues et civilisations orientales. Fondé en 1669 (date de création de son prédécesseur, l'École des langues pour la jeunesse), l'Inalco accueille aujourd'hui plus de 9 000 étudiants, qui y apprennent plus de cent langues. L'enseignement du russe y débute en 1817 et le département de russe est officiellement créé en 1876. Il est actuellement le plus important de son genre en France et en Europe, avec environ 700 étudiants et 30 enseignants (Inalco, 2024).

⁷ Même si, avec les étudiants de troisième année de Licence et, plus encore, de Master, nous exploitons le *Dictionnaire actif* et constatons qu'il suscite souvent, chez eux, le plaisir de lire un dictionnaire – parfois pour la première fois, et cela en langue étrangère. Ce fait est d'autant plus remarquable que, même dans leur langue maternelle, ils consultent rarement des dictionnaires papier, les outils numériques ayant largement modifié leurs habitudes.

1. Définition lexicographique selon la LEC et la LA

Dans cette section, nous examinerons, d'une part, la définition lexicographique telle qu'elle est conçue dans la Lexicologie explicative et combinatoire (1.1), et, d'autre part, la manière dont ces principes sont mis en œuvre et adaptés dans la Lexicographe active d'Apresjan (1.2).

1.1. Définir selon la LEC

Comme nous l'avons indiqué plus haut, la Lexicologie Explicative et Combinatoire constitue la composante lexicologique – à la fois théorique et descriptive – de la théorie linguistique Sens-Texte. Dans cette approche, la définition lexicographique d'une unité lexicale est conçue comme une paraphrase analytique rigoureuse de cette lexie, étroitement liée à la modélisation de l'ensemble de ses propriétés linguistiques, y compris de ses comportements combinatoires⁸. Les exigences formelles relatives à la structure des définitions ont commencé à se formuler dans les années 1960 (Žolkovskij *et al.*, 1961 ; *Mašinnyj perevod i prikladnaja lingvisika*, 1964). Elles ont ensuite été exposées et précisées dans plusieurs travaux fondateurs, notamment chez Apresjan (1974 : 95 et suiv.), Melčuk (Melčuk, 1974 : 111 et suiv. ; Melčuk *et al.*, 1988 : 30-37 ; Melčuk, 2013 : 279 et suiv.), ainsi que dans l'*Introduction à la lexicologie explicative et combinatoire* (Melčuk *et al.* 1995 : 72 et suiv.).

L'élaboration des définitions, dans la LEC, repose sur six principes fondamentaux : les trois premiers concernent le contenu informationnel des définitions, et les trois autres, leur structuration (Melčuk *et al.*, 1995 ; Melčuk & Polguère, 2016 : 62–78). En nous appuyant sur l'article Melčuk & Polguère, 2016, nous proposons une présentation de ces principes, afin d'en dégager l'essentiel. L'objectif n'est pas ici d'en donner une exposition exhaustive, mais de souligner qu'ils sont exploitables dans une perspective didactique. Il convient de préciser que, pour des raisons d'économie d'espace et de lisibilité, nous avons choisi de donner, dans le corps du texte, la traduction directe des définitions initialement rédigées en russe, d'où la maladresse de certaines formulations. Il va de soi que les unités lexicales d'une langue L doivent être définies dans cette même langue L. Voici, dans leur formulation nécessairement simplifiée, les principes en question.

⁸ La mention de la notion de fonction lexicale s'impose dans le présent exposé, et nous renvoyons notre lecteur aux nombreuses publications qui l'introduisent, notamment Melčuk *et al.* (1995), section 3.5, Melčuk (1996).

1. Principe de paraphrasage (ou d'adéquation) – la définition doit établir une équivalence sémantique stricte entre deux entités linguistiques :

| *Le défini* $\equiv$ *Le définissant*.

Chaque composante du définissant doit être nécessaire, et l'ensemble des composantes suffisant pour rendre compte de tous les emplois de la lexie définie. La définition doit pouvoir, en principe, se substituer au défini dans tout contexte approprié, sans perte ni ajout de sens, comme l'illustre ci-dessous la définition de **xolodI.1**⁹ 'froid', lexie d'un vocable correspondant¹⁰.

xolodI.1¹¹ sg. 'froid'

xolod X v Y-e_{PRÉP}, oščuščajemyj Z-om_{INSTR} :
'froid X dans Y ressenti par Z'

température X de l'air dans un lieu Y ressentie
par Z

- qui est en-dessous de la température normale de confort pour Z

NB ! ASém 'X' ne peut pas être exprimé¹².

U nas v auditorii stoit xolod.

litt. 'Chez nous dans salle de cours, il y a froid'

On peut valider la définition ci-dessus au moyen de l'équivalence paraphrastique suivante :

⁹ Les définitions de **xolodI.1** et **morozI.1** ont été élaborées lors d'une séance de travail du projet LEC-ru (*Lexicologie explicative et combinatoire du russe*, Inalco Paris, ATILF CNRS Nancy et Université de Montréal) par Lidia Iordanskaja, Lidia Kolzun, Svetlana Krylosova, Igor Melčuk et Alain Polguère.

¹⁰ Un *vocable* est un ensemble de lexies aux signifiants identiques, sémantiquement distinctes, mais dont les sens possèdent une intersection significative (« pont sémantique »). Le vocable **xolod** est polysémique : il contient plus d'un lexème. **xolodI.1** est un lexème de base de ce vocable, et il possède plusieurs co-polysèmes : *froid* – température d'un objet (*xolod ètogo kamnja* 'le froid de cette pierre') ; *froid* – sensation (*drožat' ot xoloda* 'trembler de froid') ; *froid* – endroit (*vynešti produkty na xolod* 'sortir les provisions dans un endroit froid') ; *froid* (**xolod** au pl.) – période (*uspet' do xolodov* 'réussir à faire quelque chose avant l'arrivée du froid') ; *froid* – absence d'empathie (*xolod eë vzgljada* 'le froid de son regard') ; *froid* – sentiment (*xolod v duše* 'froid dans l'âme'). La polysémie de **xolod** est étudiée, dans le cadre du projet LEC-ru, par Lidia Kolzun.

¹¹ **холодI.1** ед.: *холод X в Y-е, оощуцаемый Z-ом* 'температура воздуха X в пространстве Y, оощуцаемая Z-ом • ниже нормальной комфортной температуры для Z'.

¹² Cet actant n'est pas exprimable directement : on ne peut pas dire en russe *desjatigradusnyj xolod* ou *xolod v desjat' gradusov* 'le froid (xolod) de dix degrés'.

U nas v auditorii xolod.

‘Il fait froid dans notre salle de cours.’

≡

U nas v auditorii temperatura vozduxa niže normal'noj, komfortnoj dlja čeloveka.

‘Dans notre salle de cours, la température de l’air est inférieure à la normale, c’est-à-dire à celle qui est confortable pour l’être humain.’

Voici, pour comparaison, la définition du quasi-synonyme de **XOLOD I.1**, **MOROZ I.1**, qui dénote également le froid. Les deux lexèmes partagent un élément sémantique important (*température de l’air basse*), mais ils ne sont pas interchangeables dans tous les contextes, ce qui en fait une difficulté particulière pour les apprenants du russe (Krylosova, 2022). Nous avons surligné en gris les éléments absents de la définition de **XOLOD I.1**.

MOROZ I.1 sg. ‘froid’

X-yj moroz v Y-č_{PRÉP}, oščuščаемyj Z-om_{INSTR} :
‘froid de X [degrès] dans Y ressenti par Z’

température **X** de l’air dans un lieu Y ressentie par Z

- **en tant que condition climatique** dans le lieu Y
- tel que X est basse **au point que l’eau gèle**¹³

V fevrale v Monreale stojal 20-gradusnyj moroz.

litt. ‘En février, à Montréal, il y avait un froid (*moroz*) de 20 degrés’

Les éléments surlignés dans la définition de **MOROZ I.1** ne constituent pas une surcharge arbitraire : ils correspondent à une différence réelle de sens, observée dans des contextes où **XOLOD I.1** et **MOROZ I.1** ne peuvent pas se substituer l’un à l’autre. Ainsi, les deux phrases suivantes sont incorrectes en russe :

**V fevrale v Monreale stojal 20-gradusnyj xolod.*

litt. ‘En février, à Montréal, il y avait un froid (*xolod*) de 20 degrés.’

¹³ мороз I.1 ед.: X-ый мороз в Y-е, ощущаемый Z-ом ‘X-вая температура воздуха в пространстве Y, ощущаемая Z-ом’ • являющаяся одной из характеристик погоды в пространстве Y • при которой X такой низкий, что вода замерзает’.

**Segodnja na ulice stoit moroz, da k tomu že dožd' s utra.*

litt. 'Aujourd'hui, dehors, il y a un froid (*moroz*), et en plus il pleut depuis ce matin.'

Ces deux définitions montrent aussi un aspect important de la LEC : on n'y définit pas seulement des lexies isolées comme *ХОЛОД I.1* ou *МОРОЗ I.1*, mais leur *forme propositionnelle* – c'est-à-dire le sémantème concerné et les positions actantielles qu'il contrôle –, respectivement *froid X dans Y ressenti par Z* et *froid de X [degrés] dans Y ressenti par Z*.

2. Principe de décomposition sémantique – définir une lexie, c'est en analyser le sens. Le définissant doit être composé exclusivement de sémantèmes plus simples que le sémantème défini. Deux conséquences principales découlent de cette exigence :

- absence de définition par synonymie simple ;
- prévention des cercles vicieux dans le système définitionnel.

Voici comment la définition de *NENAVIST' I.1* ('haine', *nenavist' k vragu* – 'haine pour l'ennemi') est formulée dans le cadre de la LEC (Krylosova, 2024) :

NENAVIST' I¹⁴

nenavist' X-a_{GEN} k Y-y_{DAT} iz-za Z(Y)_{GEN} :
'la haine de X pour Y à cause de Z (Y)'

sentiment intense très négatif d'une personne X envers une personne Y causé¹ par l'opinion de X à propos de Z(Y)

- X considère Z(Y) comme extrêmement négatif
- et qui [= le sentiment] cause¹ que X veut faire du mal à Y jusqu'à faire disparaître Y

Esli vy ispytyvaete sil'nuju nenavist' k sebe, rekomenduetsja obratit'sja k vraču.

'Si vous éprouvez une *haine* intense envers vous-même, il est recommandé de consulter un médecin.'

¹⁴ ненависть I: *ненависть X-a_{GEN} к Y-y_{DAT} из-за Z(Y)_{GEN}* 'очень сильное неприятное чувство, которое испытывает лицо X по отношению к лицу Y, казуированное¹ мнением X-а о Z(Y)-е • мнение X-а об Z(Y)-е крайне отрицательное • и которое [= чувство] казуирует¹, что X хочет причинить вред Y-у и сделать так, чтобы Y-а не стало.'

REM 1. Le sémantème '*personne*' peut être interprété de façon plus large, englobant également les collectivités, pays, organisations, etc.

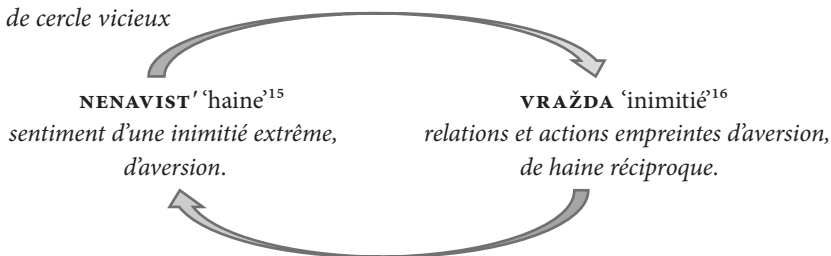
REM 2. «Z(Y)» signifie 'action, activités ou caractéristiques Z de Y'

REM 3. Le sémantème '*causer1*' correspond à la causalité non agentive, involontaire – 'X est la cause de Y'; il s'oppose à '*causer2*' – la causalité agentive et volontaire, où 'X est le causateur de Y' (Melčuk, 2012, ch. 5).

À titre de comparaison, le *Petit dictionnaire académique de la langue russe* définit NENAVIST' '*haine*' comme un *sentiment d'une inimitié extrême, d'aversion* (Evgen'eva, 1981–1984). Ce n'est pas, à proprement parler, une définition, mais une simple énumération de synonymes approximatifs, une sorte de pseudo-définition (Melčuk & Polgère, 2016 : 66). Plus problématique encore, cette définition renvoie à la lexie VRAŽDA '*inimitié*', elle-même définie comme *relations et actions empreintes d'aversion, de haine réciproque* (Evgen'eva, 1981–1984). Autrement dit, dans la définition de NENAVIST', on utilise VRAŽDA, et réciproquement – un cercle vicieux :

Figure 1.

Exemple de cercle vicieux



Cet exemple illustre, *a contrario*, l'intérêt du principe de décomposition sémantique : ramener le sens défini à une configuration de composants plus simples – '*sentiment*', '*négatif*', '*intense*', etc. – tout en évitant la synonymie simple et les renvois circulaires, afin de formuler des définitions véritablement explicatives.

3. Principe d'univocité – chaque mot du définissant doit correspondre à un sémantème donné, et chaque sémantème doit être exprimé de manière unique

¹⁵ ненависть : чувство сильнейшей вражды, неприязни.

¹⁶ вражда : отношения и действия, проникнутые неприязнью, взаимной ненавистью.

au sein des définissants dans l'ensemble du dictionnaire. Concrètement, cela implique deux exigences :

- Idéalement, dans une définition, chaque sémantème utilisé doit être muni d'un numéro lexicographique : on ne parle pas du sémantème LJUBOV' 'amour', mais de LJUBOV'**I.1** (*ljubov' k svoim detjam* 'amour envers ses enfant'), LJUBOV'**I.2** (*ljubov' k ženščine* 'amour pour une femme') ou de LJUBOV'**II** (*ljubov' k tomatnomu soku* 'amour pour le jus de tomate').
- Les patrons de définition doivent être harmonisés pour les lexies sémantiquement très proches. Ainsi le sémantème 'température', comme nous l'avons vu ci-dessous est présent dans les définitions des quasi-synonymes XOLOD**I.1** et MOROZ**I.1** 'froid', et il devrait également l'être dans celles de ŽARA, ZNOJ 'chaleur', PROXLADA 'fraîcheur', etc. (voir *infra*, sous-section 2.2). Quant au champ sémantique des sentiments¹⁷ auquel appartient le lexème NENAVIST'**I** défini plus haut, toutes les lexies qui en relèvent doivent comporter dans leur définition le sémantème 'sentiment' en position stratégique. Voici quelques ébauches¹⁸ de définitions de certaines unités lexicales appartenant à ce champ.

Tableau 1.

Ébauches de définitions d'unités lexicales du champ des sentiments

GOREČ' II 'amertume' <i>goreč' X-a_{GEN} / Y-a_{GEN} :</i>	'sentiment négatif durable et profond de la personne X cause 1 par le fait que X perçoit comme une injustice le fait Y qui lui est arrivé – 'comme si' X éprouvait GOREČ' I.1 ¹⁹ <i>goreč' poraženija</i> 'amertume de la défaite'
LJUBOV' II 'amour' <i>ljubov' X-a_{GEN} k Y-y_{DAT} :</i>	'caractéristique d'un être vivant X d'éprouver un sentiment positif assez intense envers une entité ou un fait Y. Ce sentiment est cause 1 par le contact entre X et Y, il cause 1 que X veut être en contact avec Y' <i>ljubov' k tomatnomu soku</i> 'amour pour le jus de tomate' REM 1. Dans le cas où Y est une situation, 'se trouver en contact avec Y' signifie 'être participant de cette situation'

¹⁷ À propos des champs sémantiques : cf. Polguère, 2013.

¹⁸ Nous avons pleinement conscience que les définitions présentées ici ne sont pas exhaustives et gagneraient à être précisées et affinées.

¹⁹ GOREČ'**I.1** (amertume de X) pourrait, à son tour, être défini comme 'gout amer de l'entité X'. Exemple : *He чувствую горечи, он выпил лекарство*. 'Il a avalé le médicament sans en sentir l'amertume'. Voir également l'analyse des lexèmes du vocable AMERTUME en français dans Polguère, 2013.

Tableau 1 (Suite)

NADEŽDA ‘espoir’ <i>nadežda</i> X-a _{GEN} na Y _{ACC} :	‘ sentiment positif assez intense causé ¹ par la pensée qu’il peut lui arriver quelque chose de bien Y’ <i>nadežda na lučšee</i> ‘espoir du meilleur’
---	--

Toujours à titre de comparaison, voici les définitions correspondantes tirées du *Petit dictionnaire académique* (Evgeněva, 1981–1984) :

GOREČ'3 (*goreč' poraženija* ‘amertume de la défaite’) : ‘**sentiment** pénible provoqué par un malheur, un échec, une offense, etc.’²⁰

LJUBOV'3 (*ljubov' k tomatnomu soku* ‘amour pour le jus de tomate’) : ‘**penchant** pour quelque chose, préférence pour quelque chose’²¹ ;

NADEŽDA ‘espoir’ (*nadežda na konec vojny* ‘espoir de la fin de la guerre’) : ‘**attente** de quelque chose de souhaité, accompagnée de la certitude de la possibilité de sa réalisation’²².

Comme on peut le constater, seul le premier article, GOREČ'3, emploie le sémantème ‘*sentiment*’, comme si ces trois lexies n’avaient rien de commun.

En LEC, pour les lexies sémantiquement proches (noms d’animaux, de sensations, d’actes de communication verbale, etc.), on élabore des schémas définitionnels communs. Ainsi, pour définir la lexie NENAVIST'1 (Krylosova, 2024), nous avons suivi le patron établi pour un sous-ensemble de noms appartenant au champ « *sentiments* » (Iordanskaja & Melčuk, 2022 ; Apresjan & Apresjan, 1993 ; Milićević, 2016 ; cf. également Wierzbicka, 1999 relevant du cadre du *Natural Semantic Metalanguage*) :

²⁰ горечь3: тяжелое чувство, вызываемое бедой, несчастьем, неудачей, обидой и т. п.
²¹ любовь3: внутреннее стремление, влечение, склонность, тяготение к чему-л.
²² надежда : ожидание чего-л. желаемого, соединенное с уверенностью в возможности осуществления.

Tableau 2.
Schéma définitionnel pour un sous-ensemble de noms du champ « sentiments »

DÉFINI	<i>nemavist' X-a_{GEN} k Y-y_{DAT} iz-za Z(Y)_{GEN}</i> 'la haine de X pour Y à cause de Z'	
composante centrale	'sentiment intense très négatif, d'une personne X envers une personne Y'	COMMENTAIRE : Cette paraphrase minimale contient le sémantème <i>čuvstvo</i> 'sentiment' que nous caractérisons à travers les deux paramètres suivants : 1. L'intensité (sur l'échelle 'intense' ~ 'modéré') 2. La polarité (sur l'échelle 'ressentir quelque chose de bien' ~ 'ressentir quelque chose de mal')
composantes périphériques • composante de stimulus • composante de l'effet	'causé1 par l'opinion de X à propos de Z(Y) • X considère Z(Y) comme extrêmement négatif • et qui [= le sentiment] cause1 que X veut faire du mal à Y jusqu'à faire disparaître Y'	COMMENTAIRE : Le schéma de définition pour les noms de sentiments peut comporter deux composantes périphériques principales : le <i>stimulus</i> (la cause du sentiment) et l' <i>effet</i> (la description des types de comportement de X provoqué par le sentiment). Bien que la composante de l'effet ne figure pas dans toutes les définitions des noms des sentiments, elle est caractéristique des noms des sentiments intenses, tels que haine.

Comme les lexies sémantiquement apparentées tendent à avoir la même structure de définition, cette manière de description par champ lexical, avec des blocs définitoires standard, permet de respecter le principe d'univocité. Et, comme nous le verrons plus bas, elle s'avère très efficace, par exemple, pour faire parler les apprenants, d'une langue étrangère – même débutants – en cours d'expression orale (voir *infra*, section 2).

4. Principe de décomposition minimale (ou de bloc maximal) – tout sémantème non primitif est décomposable, mais l'analyse doit s'arrêter au niveau le plus

élevé possible afin de préserver la lisibilité, la fonction paraphrastique et une profondeur uniforme de décomposition.

Reprenons la définition de NENAVIST'1. Sa composante périphérique de stimulus contient le sémantème 'mnenie' ('opinion') : *litt.* 'opinion de X sur Z(Y) est extrêmement négative'. Puisque 'mnenie' ('opinion') n'est pas un sémantème primitif, il est décomposable :

MNENIE²³ 'opinion'

<i>mnenie</i> X-a, čto Y – P :	information P(Y) présente dans l'esprit de X • que X a décidé de considérer comme vraie à la suite du fait que X a réfléchi à des données relatives à Y ou à P
--------------------------------	--

Nous aurions donc pu poursuivre la décomposition du définissant de NENAVIST'1 et substituer à *mnenie* sa paraphrase ('information P(Y) présente dans l'esprit de X...'). Nous nous en abstenons toutefois, car une telle expansion compromettrait la lisibilité de la définition et en limiterait la maniabilité²⁴. C'est pour cette raison que la LEC impose de s'arrêter à la **décomposition sémantique minimale**²⁵.

5. Principe de structuration hiérarchique. Comme nous l'avons vu plus haut avec la définition du lexème NENAVIST'1, son définissant comprend des composantes, c'est-à-dire des configurations de sémantèmes qui forment un module définitionnel. Ces composantes sont organisées de façon structurée et hiérarchique et n'ont pas le même statut : on distingue une composante centrale (*genre prochain*), unique et obligatoire, et des composantes périphériques (*différences spécifiques*). Dans le cas des noms de sentiments, il s'agit notamment d'une composante périphérique de stimulus (la cause du sentiment) et, éventuellement, d'une composante périphérique de l'effet (réaction au sentiment).

²³ мнение1: мнение X-a_{GEN}, čto Y – P 'информация P(Y), имеющаяся в уме X-a • которую X решил считать истинной в результате того, что X обдумал данные, относящиеся к Y или к P'. Il s'agit de la définition proposée par le *Dictionnaire explicatif et combinatoire du russe*, modifiée à la lumière des remarques critiques formulées par Juri Apresjan (Apresjan, 2010 : 27).

²⁴ Cf. cette remarque d'Apresjan : « Ce sont les définitions construites à partir de blocs relativement importants qui sont psychologiquement les plus acceptables, car elles conservent la propriété de transparence. Si une signification lexicale assez complexe est immédiatement ramenée à des primitifs, elle risque, malgré toute sa précision, de devenir pratiquement méconnaissable » Apresjan (1994).

²⁵ Dans d'autres approches, par exemple dans le cadre des recherches transculturelles, la décomposition maximale peut être justifiée. C'est, par exemple, le cas du *Natural Semantic Metalanguage* (Wierzbicka 1972, 1985 et travaux ultérieurs).

La représentation formelle d’une définition dans la LEC encode explicitement le statut de chacune de ses composantes, ce qui rend visibles les relations de dépendance.

Cette structuration peut être illustrée avec la définition du lexème LAVINAL.1 ‘avalanche’ :

LAVINAL.1 ²⁶ ‘avalanche’	
<i>lavina, iduščaja na Y</i> : ‘avalanche qui arrive sur Y’	‘masse de neige • qui est énorme • qui se déplace accidentellement rapidement vers le bas de la pente d’une montagne sur Y • (et cause des dommages à Y)’

Nous voyons ici que LAVINAL.1 possède la structure hiérarchique où chaque composante périphérique est subordonnée, du point de vue de l’information, à la composante centrale²⁷ ‘masse de neige’ :

Tableau 3.
Structure hiérarchique de la définition de LAVINAL.1

‘masse de neige’			
‘qui est énorme’	‘qui se déplace rapidement’	‘qui se déplace le long de la pente d’une montagne sur Y’	(‘qui cause des dommages à Y’)

La composante (‘qui cause des dommages à Y’) est mise entre parenthèses, car elle n’est pas nécessairement activée lorsque le locuteur utilise LAVINAL.1 dans un énoncé. Ce type de composante, dans une définition, est qualifié de *composante sémantique faible* (Melčuk & Polguère, 2016 : 73).

Cette organisation en composantes offre, entre autres, un cadre clair pour présenter les définitions aux apprenants, en mettant en relief les éléments essentiels du sens et leurs interrelations.

²⁶ лавинаI.1: лавина, идущая на Y_{acc} ‘огромная масса снега, • которая стремительно перемещается вниз по склону горы на Y • (и наносит ущерб Y-y)’. Les définitions de LAVINAL.1 et de ses co-polysèmes ont été formulées par L. Kolzun, S. Krylosova, A. Polguère dans le cadre du projet LEC-ru (voir Kolzun *et al.*, 2018).

²⁷ Une composante périphérique peut également être subordonnée à une autre composante périphérique.

6. Principe de formalisation en réseau sémantique. Dans la Théorie Sens-Texte, une représentation sémantique s'écrit nécessairement sous la forme d'un réseau. Celui-ci est constitué de nœuds, reliés entre eux par des flèches : les nœuds portent les étiquettes des éléments sémantiques, tandis que les flèches sont assorties de numéros indiquant les relations de type « prédicat ~ argument » (voir Melčuk 1988 : 52 et suiv. ; Melčuk *et al.* 1995 : 75). À l'instar des auteurs de Melčuk *et al.* (1995), nous n'avons pas recours à ce formalisme des réseaux sémantiques pour décrire le sens des lexies. Nous privilégions plutôt des définitions rédigées en russe, proches dans leur forme de celles des dictionnaires usuels. Autrement dit, dans notre pratique, nous donnons la priorité à la linéarité sur la multidimensionnalité (voir les exemples de définitions présentés plus haut ; cf. Melčuk *et al.* 1995 : 73–75).

1.2. Définir selon la LA

C'est en s'appuyant sur ces principes que l'équipe dirigée par Ju. Apresjan élabore le *Dictionnaire actif de la langue russe*. Selon lui, cet ouvrage « ne sera pas une simple copie, même des meilleurs dictionnaires étrangers » (Apresjan, 2010), mais visera à établir un standard original du type *dictionnaire actif*, fondé sur les principes de la LEC. L'ambition affichée n'est donc pas de transposer tel quel, sur le matériau russe, le modèle élaboré dans la tradition européenne, mais bien de le repenser en profondeur. Tout en préservant le meilleur de cette tradition, il s'agit de la moderniser radicalement, en s'appuyant sur les acquis théoriques les plus avancés de la linguistique d'aujourd'hui.

Le projet définit deux caractéristiques essentielles :

1. être conçu selon des principes lexicographiques modernes, recourir aux technologies linguistiques contemporaines et intégrer les résultats les plus récents de la théorie linguistique ;
2. s'adresser à un large public de locuteurs de la langue ne possédant pas de formation linguistique spécialisée – c'est-à-dire ayant, par exemple, un niveau d'études équivalent au baccalauréat en France –, et présenter les acquis les plus récents de la théorie linguistique dans une langue claire et accessible au non-spécialiste.

Un tel dictionnaire est pensé pour remplir simultanément deux fonctions :

1. servir de source d'orientation théorique et de matériaux de base pour la description lexicographique scientifique du russe ;
2. constituer un outil de type actif destiné à la consultation pratique.

Dans sa première fonction, il pourra devenir un point d'appui majeur pour des recherches théoriques variées ; dans la seconde, il contribuera à une maîtrise pratique complète de la langue russe (Apresjan, 2010).

Au cœur de la démarche lexicographique adoptée dans le *Dictionnaire actif* se trouvent les définitions, conçues conformément aux principes bien établis de la LEC²⁸. Dans cette perspective, une bonne définition doit, idéalement, répondre à trois exigences fondamentales : **a)** être complète, **b)** éviter toute redondance et **c)** ne pas être tautologique²⁹ (Apresjan, 2010). Pour illustrer ces principes, nous proposons ci-dessous la comparaison entre deux définitions du verbe *VOSXIŠČIAT'SJA* 'admirer' : la première, extraite du *Dictionnaire explicatif et combinatoire* du russe et légèrement remaniée, et la seconde, issue du *Dictionnaire actif* de la langue russe.

Tableau 4.
Définitions du verbe VOSXIŠČIAT'SJA 'admirer' en LEC et en LA

VOSXIŠČIAT'SJA ³⁰ 'admirer' (LEC)	VOSXIŠČIAT'SJA ³¹ 'admirer' (LA)
<i>X admire Y</i> : 'X éprouve un sentiment positif assez intense, causé ¹ par le fait que X considère ² que Y qu'il perçoit, est très bon (, et X prononce une énonciation exprimant ce sentiment); ce sentiment est tel qu'il est habituellement dans la situation indiquée.'	<i>A1 admire A2</i> : 'La personne A1 éprouve et, éventuellement, exprime par des paroles A3 un sentiment fort et agréable, suscité par l'objet ou la situation A2, qu'elle évalue très positivement.'

²⁸ Ces principes qui ont été ajustés dans le cadre des travaux sur le *Nouveau dictionnaire explicatif de synonymes de la langue russe* (voir Iordanskaja, ce volume), témoignant ainsi d'une réflexion continue sur cet aspect central de la lexicographie (Apresjan, 1994).

²⁹ Néanmoins, Apresjan souligne que, comme le montre l'expérience accumulée lors du traitement lexicographique de vastes corpus, cet idéal est loin d'être toujours atteignable. Cela n'enlève rien à sa valeur, mais il faut clairement comprendre que, dans certains cas, le respect strict de ces exigences devient impossible et qu'elles doivent être assouplies pour les besoins du travail lexicographique concret (Apresjan, 2010). Ces observations, particulièrement pertinentes, mériteraient à elles seules de faire l'objet d'un article distinct.

³⁰ *X восхищается Y-ом* 'X испытывает достаточно интенсивное положительное чувство, каузированное¹ тем, что X считает² воспринимаемый им Y очень хорошим (, и X произносит высказывание, выражающее это чувство); это чувство таково, какое обычно бывает в указанной ситуации'. Nous avons apporté quelques ajustements à la définition du DEC du russe. À l'époque, les auteurs du dictionnaire avaient fait le choix méthodologique de ne pas recourir au sémantème 'čuvstvo' sentiment.

³¹ *A1 восхищается A2* 'Человек A1 испытывает и, возможно, выражает словами A3 сильное и приятное чувство, вызванное объектом или ситуацией A2, которые он очень высоко оценивает'. В форме СОВ часто указывает на словесное проявление эмоции: *Ты без пальто? – восхитился^{PERF} он.*

Tableau 4 (Suite)

	À la forme perfective, indique souvent une manifestation verbale de l'émotion : <i>Ty bez pal'to? – vosxilitilsja^{PERF} on.</i> <i>'Tu es sans manteau? – s'émerveilla-t-il.'</i>
--	--

La comparaison de ces deux définitions montre que les auteurs du *Dictionnaire actif* mettent en œuvre les principes de la LEC présentés plus haut (paraphrasage, décomposition sémantique, etc.). L'examen attentif révèle toutefois une inflexion propre à leur démarche : à ces principes s'ajoutent deux orientations spécifiques, identifiées par Apresjan. Les voici :

1. Le principe de convivialité des définitions : elles doivent être formulées de manière à ce qu'elles soient généralement acceptées et facilement reconnaissables. En cas de conflit entre, d'une part, l'exigence de vérité scientifique – qui peut conduire à des formulations excessivement complexes – et, d'autre part, la nécessité de rester compréhensible pour un lecteur non spécialiste, le *Dictionnaire actif* privilégie la clarté et l'accessibilité (Apresjan, 2010).

Nous voyons, en comparant les deux définitions du lexème *VOSXIŠČIAT'SJA* 'admirer', que dans la seconde, le langage définitoire est allégé : la composante périphérique faible – '*exprime (ce sentiment) par des paroles*' – est intégrée dans le corps du texte, sans être isolée entre parenthèses. Par ailleurs, le sématème '*causé1*' (*kauzirovať1*), terme artificiel du métalangage lexicographique – inexistant en russe, mais indispensable dans une définition à visée théorique – est ici remplacé par une formulation plus simple, plus immédiatement compréhensible pour un lecteur non spécialiste '*suscité par l'objet ou la situation A2*'.

2. Le principe d'intégralité de la description linguistique : formulé dès les travaux d'Apresjan (1980, 1995), il repose sur l'idée que les deux composantes principales d'une description linguistique – le dictionnaire et la grammaire – doivent être conçues comme un ensemble intégré et cohérent³². Comme le note Apresjan, le lexicographe, en décrivant une nouvelle lexie, doit travailler sur l'ensemble des règles linguistiques et attribuer à chaque lexie toutes les propriétés susceptibles d'être mobilisées par ces règles (réglage du dictionnaire sur la grammaire).

³² On retrouve cette idée également dans les travaux de Mel'čuk et de ses disciples. Le principe d'une description intégrale de la langue dans le dictionnaire – sans toutefois qu'elle soit fixée terminologiquement – a été étendu à la syntaxe à partir des années 1970. Voir Apresjan, 2006 : 42.

Cela exige souvent d'intégrer certaines règles directement dans le dictionnaire, notamment lorsqu'elles ne concernent qu'une lexie donnée ou un petit groupe de lexies, auquel cas leur place la plus naturelle est précisément dans leurs articles de dictionnaire³³ (Apresjan, 2006 : 43).

Dans le cas du verbe *vosxiščat'sja* 'admirer', on voit que l'article de dictionnaire inclut non seulement l'information selon laquelle ce verbe possède une forme perfective, mais aussi la précision qu'à la forme perfective, il indique souvent une manifestation verbale de l'émotion. Autrement dit, un composant faible de la définition cesse de l'être lorsque le verbe est employé au perfectif.

Ajoutons que c'est précisément cette approche (réglage du dictionnaire sur la grammaire) qui fait du *Dictionnaire actif* un outil inestimable pour l'enseignant de russe langue étrangère.

De manière générale, la zone des commentaires qui accompagne les définitions dans le *Dictionnaire actif* d'Apresjan possède un potentiel considérable pour la préparation de cours de RLE. Faute de pouvoir illustrer, dans le cadre de cet article, tous les types de commentaires proposés (informations pragmatiques, règles sémantiques, connotations, etc.), nous nous arrêterons sur l'un d'entre eux : la comparaison succincte d'une lexie avec son plus proche synonyme. Tout enseignant de langue étrangère sait combien il peut être difficile, parfois, de répondre à la question « Quelle est la différence entre ces deux mots ? » et combien les apprenants se montrent insatisfaits d'une réponse du type « On peut utiliser l'un ou l'autre »³⁴ d'autant plus que les dictionnaires traditionnels, dans ce genre de cas, apportent rarement une aide réelle.

Prenons l'exemple du commentaire associé au lexème *istina* 'vérité'. En russe, pour la 'vérité' on utilise deux termes – *pravda* et *istina* –, dont la distinction a été abondamment étudiée dans la littérature linguistique (voir, entre autres, Arutjunova, 1995, 1998 ; Levontina, 1995 ; Bulygina & Šmel'ev, 1997 ; Vežbickaja, 2002). La question de la différence entre ces deux synonymes apparaît fréquemment chez les apprenants, parfois dès les premières étapes de l'apprentissage³⁵, et nécessite de la part de l'enseignant une réponse précise, au-delà d'une

³³ Apresjan souligne également que l'importance accordée, dans ses travaux, aux règles intégrées dans un dictionnaire répond aussi à une motivation d'ordre psychologique : « il semble difficile au lexicographe d'accepter l'idée que le dictionnaire doive devenir un dépôt pour de nombreuses règles » (Apresjan, 2006 : 23, note 15).

³⁴ Dans notre pratique, nous incitons d'ailleurs les étudiants à ne jamais se contenter d'une telle explication.

³⁵ Avec son aimable autorisation, nous citerons ici un message de l'une de nos étudiantes (grande débutante, deux mois d'apprentissage du russe, nous traduisons du russe) : « Madame, voici les exercices de phonétique. Pourriez-vous me dire si j'ai fait des erreurs ? De plus, j'ai deux questions : existe-t-il une différence entre *pravda* et *istina* 'vérité' et entre *bazar* et *rynok*

simple équivalence lexicale. Voici la définition, ainsi que le commentaire qui l'accompagne, proposée par le *Dictionnaire actif* pour ISTINA :

ISTINA¹³⁶, nom ; féminin ; pas de pluriel ; *philos. ou vieilli.*

‘Connaissance complète de la structure du monde, constituant la plus haute valeur spirituelle, qui donne aux hommes la réponse à toutes les questions importantes de l’existence et dont l’acquisition est pour eux un objectif très important et difficile à atteindre’

COMMENTAIRE

Le mot ISTINA se rapproche du nom synonyme PRAVDA^{3.2} par le trait qui les réunit : *istina* comme *pravda* est difficile à atteindre et hautement valorisée par les êtres humains. PRAVDA^{3.2} désigne une représentation morale correcte du monde, propre aux individus, ainsi que le comportement qui lui correspond : *vivre selon la pravda*. ISTINA, en revanche, désigne un savoir suprême, impersonnel, dont l’acquisition permet aux êtres humains de comprendre le sens de la vie et de répondre à toutes les grandes questions de l’existence : *quête éternelle de l’istina* [...]. Contrairement à *pravda*, *istina* appartient à une catégorie philosophique et religieuse, et peut être employée dans un sens terminologique.

Ce type d’analyse fine, absent des dictionnaires classiques, constitue l’une des spécificités majeures du *Dictionnaire actif*.

Cette section a permis de mettre en évidence – autant que l’autorisait l’espace disponible – le potentiel pédagogique exceptionnel et, à bien des égards, inédit dans la tradition lexicographique, des définitions élaborées dans le cadre de la LEC-LA. Nous allons à présent examiner comment la méthodologie développée par les équipes de Melčuk et d’Apresjan peut être adaptée aux besoins spécifiques de l’enseignement des langues étrangères.

‘marché (endroit)’. Merci beaucoup de vos réponses. Jeanne.» (correspondance personnelle, le 4 novembre 2021).

³⁶ ‘Являющееся высшей духовной ценностью полное знание о том, как устроен мир, которое даёт людям ответ на все важные вопросы бытия и обретение которого является для них очень важной и труднодостижимой целью’.

2. Définitions pédagogiques selon la LEC-LA

2.1. Qu'est-ce qu'une définition pédagogique ?

Dans le cadre de cet article, nous entendons par *définition pédagogique ou définition didactisée* une définition élaborée selon les principes de la LEC-LA, mais spécifiquement adaptée aux besoins de l'apprentissage d'une langue étrangère. Ce type d'approche, dont Jasmina Milićević (2008, 2016) a été la pionnière pour le français LE³⁷, a montré qu'une définition construite selon les principes de la LEC-LA peut constituer non seulement un outil de description lexicographique rigoureux, mais aussi un support efficace pour développer la précision, la cohérence et l'autonomie dans l'usage lexical en contexte d'apprentissage. Dans les travaux de J. Milićević et de ses collègues, cette démarche a notamment été exploitée pour affiner la distinction entre quasi-synonymes (*mécontentement* ~ *insatisfaction* ~ *déplaisir*) et entre lexies de sens proches (*mécontentement* ~ *déception* ~ *frustration*).

La définition pédagogique conserve de la LEC-LA le même souci de précision et de structuration, la finesse méthodologique et l'orientation vers la production de la parole. Il n'y a pas de renoncement à l'emploi d'un certain formalisme – d'autant que nous travaillons avec des étudiants qui maîtrisent déjà, dans une certaine mesure, les conventions lexicographiques – mais sa formulation peut être allégée, de manière à rester simple sans être simpliste (Milićević, 2016 : 101–102).

Illustrons les modifications possibles par rapport aux définitions classiques de la LEC-LA à partir d'un couple de quasi-synonymes dénotant le froid – XOLODI.1 ('température de l'air basse') et MOROZI.1 ('température de l'air basse au-dessous de zéro, à l'extérieur') – que nous avons déjà évoqué plus haut. Dans la colonne de droite figurent les définitions au format didactisé que nous proposons dans le manuel d'expression orale *Mel'nica* (Krylosova, 2025b), destiné aux grands débutants ; dans le manuel, elles sont, bien entendu, données en russe.

³⁷ Parallèlement, des travaux importants sont menés par nos collègues de l'Université de Lorraine (Nancy) – en particulier Veronika Lux-Pogodalla et Alain Polguère – sur la conception de définitions didactisées, destinées cette fois à l'enseignement du français langue maternelle, et dont de nombreuses orientations méthodologiques rejoignent nos propres préoccupations (Lux-Pogodalla & Polguère, 2011 ; Lux-Pogodalla & Polguère, 2021).

Tableau 5.
Définitions classiques et didactisées

Définitions lexicographiques classiques	Définitions lexicographiques didactisées
XOLODI.1 sg. <i>xolod X v Y-e, oščuščaemyj Z-om</i> température de l'air X dans l'espace Y ressentie par Z • en-dessous de la température normale de confort ASém 'X' ne peut pas être exprimé.	XOLODI sg. – TEMPERATURE DE L'AIR à l'intérieur ou dehors • Cette température est basse, et elle n'est pas confortable pour un être humain.
MOROZI.1 sg. <i>X-yj moroz v Y-e, oščuščaemyj Z-om</i> température X dans un lieu Y ressentie par Z • en tant que condition climatique dans le lieu Y • tel que X est bas au point que l'eau gèle	MOROZI sg. <i>moroz v X gradusov</i> TEMPÉRATURE DE L'AIR de X degrés dehors • Température de X degrés est une condition météorologique (<i>litt.</i> une caractéristique de la météo). • Température de X degrés est si basse que l'eau gèle.

Voici les principales modifications pédagogiques que nous avons apportées au format classique des définitions (voir à ce propos Krylosova, 2022).

1. Afin de permettre à l'enseignant d'évoquer, si nécessaire, la polysémie des vocables XOLOD et MOROZ (« Ce n'est pas de ce froid-là que nous parlons. »), nous avons conservé les numéros d'acception, en les simplifiant toutefois.

2. Pour XOLOD**I**, nous avons renoncé à la forme propositionnelle (en italique dans la colonne de gauche) – que nous avons jugée lexicalement et grammaticalement difficile à appréhender à ce stade de l'apprentissage du russe.

3. Pour faciliter la lecture des définitions, nous avons supprimé ou lexicalisé certaines variables. Par exemple, pour XOLOD**I**, 'dans l'espace Y' a été remplacé par 'à l'intérieur ou à l'extérieur (*litt.* à la maison ou dans la rue)', et la variable Z par 'être humain'. Pour MOROZ**I**, nous n'avons conservé qu'un seul actant, X, qui peut s'exprimer comme dépendant direct (*moroz de X degrés*). La différence

avec le lexème **XOŁOD1**, pour lequel cet actant n'est pas exprimable directement, devient ainsi plus manifeste.

4. Nous avons simplifié le vocabulaire et la syntaxe du langage définitoire. Au stade où ces définitions apparaissent dans le manuel, les étudiants connaissent deux cas obliques du russe (et cela uniquement pour les substantifs ainsi que les pronoms personnels et interrogatifs) : nous introduisons le Prépositif et le Génitif dans nos définitions, en évitant toute autre forme fléchie. De même, nous avons écarté les constructions complexes de la définition classique (proposition participiales et relatives) en les remplaçant par des phrases complètes.

De manière générale, dans notre travail sur les définitions pédagogiques, nous suivons les recommandations formulées par J. Milićević (2016 : 108) :

1. Lorsque nous envisageons d'alléger une définition par l'omission ou le remplacement de certaines composantes, nous veillons à ce que cette modification ne contrevienne pas au principe d'adéquation de la LEC-LA – rappelons que, selon ce principe, la définition doit pouvoir se substituer au défini dans tout contexte approprié.

Malgré les modifications apportées à la définition classique de **MOROZ1.1**, la définition pédagogique peut être validée par le remplacement du lexème en question par la paraphrase de son définissant :

Na ulice segodnja desjatigradusnyj moroz.

litt. 'Dans la rue, aujourd'hui il y a un froid (*moroz*) de dix degrés.'

≡

Na ulice segodnja nizkaja tempretura vozduxa v desjat' gradusov. Èto xarakteristika pogody. Èta temperatura takaja nizkaja, èto voda zamerzaet.
'Dans la rue aujourd'hui il y a une température de dix degrés [en dessous de zéro]. C'est une caractéristique de la météo. Cette température est si basse que l'eau gèle'.

2. Nous nous autorisons à déroger au principe de la décomposition minimale (ou du bloc maximal) lorsque cette adaptation s'avère nécessaire pour atteindre nos objectifs pédagogiques. Ainsi, dans la définition de **XOŁOD1**, le sémantème '*espace (prostranstvo)*' a été jugé trop complexe pour des étudiants en année d'initiation. Afin d'éviter une difficulté lexicale supplémentaire et de leur permettre de se concentrer sur l'essentiel, il a été remplacé par l'élément '*à la maison ou dans la rue*'³⁸.

³⁸ De même, pour le verbe **VOSXIŠČAT'SJA** 'admirer' que nous avons vu plus haut, le respect strict du principe aurait conduit à une définition du type : '*éprouver VOSXIŠČENIE (admiration)*'.

3. Nous cherchons à mettre en évidence les blocs définitionnels et le rôle des différentes composantes de la définition et nous n'hésitons pas à élargir la zone des commentaires lorsque cela peut aider les étudiants. Citons ici un exemple du verbe POPREKAT'2 '≈ reprocher de ne pas 'renvoyer l'ascenseur' après un service rendu'. Il s'agit d'un lexème lacunaire, « intraduisible » en français ; nous en donnerons donc d'abord une explication.

Ce verbe de communication évaluative³⁹ décrit la situation où une personne, après avoir rendu un service ou accompli une bonne action envers un proche, estime avoir acquis un droit à des avantages réciproques, à l'obéissance ou, à tout le moins, à une gratitude constante. Elle évoque alors ses sacrifices passés pour maintenir le bénéficiaire dans une position de dépendance, exerçant ainsi une pression morale. *Poprekat'* désigne un comportement socialement réprouvé⁴⁰. Paradoxalement, le bienfaiteur devient vulnérable : un mot susceptible d'être interprété dans ce sens peut, parfois pour toute la vie, entraîner l'accusation de *poprekat'* (Zaliznjak *et al.*, 2005 ; Krylosova, 2022).

Voici la définition didactisée de POPREKAT'2 que nous proposons aux apprenants d'un niveau intermédiaire (fin de la première année de Licence ; la définition est donnée en russe).

Tableau 6.
Définition didactisée de POPREKAT'2

structure de définition	POPREKAT'2, pf. POPREKNUT'2 (traduction non donnée aux étudiants)
Forme propositionnelle	Človek X poprekaet človeka Y faktom Z X : Nom. Y : Acc. Z : Instr. ČTO

Toutefois, dans certains cas, la clarté prime sur l'application littérale de ce principe. C'est pourquoi, dans le cadre de cet article, et en pensant à notre lecteur, nous avons remplacé VOSXIŠČENIE par sa décomposition explicite : 'un sentiment positif assez intense [...]']

³⁹ Un schéma définitionnel applicable à ce type de verbes a été proposé, entre autres, dans Dostie *et al.* (1999), Alonso Ramos (2003) et Iordanskaja (2007). En ce qui concerne la didactisation des définitions de cette catégorie verbale selon les principes de la LEC, on pourra se reporter à l'article particulièrement éclairant de Milićević (2008). Voir également l'article UPREKAT' 'reprocher' dans Apresjan (2004).

⁴⁰ À la première personne, il s'emploie presque exclusivement à la forme négative, pour se défendre : *Ja tebja ne poprekaju!* 'Je ne te dis pas cela pour étaler mes bienfaits!' (Zaliznjak *et al.*, 2005).

Tableau 6 (Suite)

structure de définition	POPREKAT'2, pf. POPREKNUT'2 (traduction non donnée aux étudiants)
<i>Présumé</i>	X a fait Z pour Y X pense que Z est BON pour Y X pense que Y ne se rend pas compte en quoi il est redevable à X
<i>Assertion</i>	X PARLE à Y du Z • pour rappeler à Y qu'il a une dette morale envers X • parce que X a fait Z pour Y • pour faire par cet acte que Y se sente mal
<i>Commentaire</i>	La personne qui emploie le verbe POPREKAT'2, pense que POPREKAT'2 n'est pas bien.
<i>Exemples</i>	<i>On vzjal eë замуž s reběnkom, i vsju žizn' ètim poprekaet.</i> 'Il l'a épousée alors qu'elle avait déjà un enfant et, toute sa vie, il lui rappelle qu'elle doit lui en être reconnaissante' <i>Ne xoču ja tvoix deneg! Ty potom budeš' menja imi poprekat'.</i> 'Je ne veux pas de ton argent ! Plus tard, tu me rappelleras régulièrement que je t'en suis redevable'

En présentant la définition de cette manière, nous offrons aux étudiants une sorte de scénario de pièce de théâtre – le fait Z, la personne Y à l'origine de ce fait, et une personne X qui l'évalue et, éventuellement, communique cette évaluation à Y ou à une tierce personne –, scénario que l'on retrouve dans tous les autres verbes de communication évaluative (*kritikovat'* 'critiquer', *obvinjat'* 'accuser', *osuždat'* 'blâmer', etc.) et que les étudiants peuvent ainsi suivre et reproduire plus facilement.

Ces exemples d'aménagements des définitions LEC-LA conduisent naturellement à s'interroger sur la fonction de la définition pédagogique en cours de langue.

2.2. Définition pédagogique : pourquoi et pour quoi ?

À l'Inalco, nous introduisons la définition pédagogique LEC-LA dans nos cours depuis près de dix ans. Il n'est pas possible d'en dresser ici un bilan complet ; nous nous limiterons donc à quelques exemples représentatifs.

L'usage de la définition pédagogique en classe peut répondre à deux objectifs principaux :

1. Introduire ponctuellement une définition, par exemple, pour expliciter le sens d'un lexème lacunaire en français (comme POPREKAT'2, cf. *supra*), pour distinguer deux quasi-synonymes (XOLOD1 et MOROZ1 'froid'), clarifier la différence entre des lexies de sens proches (VOSXIŠČAT'SJA et LJUBOVAT'SJA 'admirer') ou pour expliquer pourquoi, dans des cas pourtant transparents au premier coup d'œil, et qui ne devraient poser aucun problème, les étudiants se heurtent néanmoins à des difficultés de traduction.

Prenons une phrase en français *Dès que l'avalanche commence, tentez de vous frayer un passage vers la surface en agitant les bras*. Sa traduction en russe – **Kak tol'ko načinaetsja lavina, pytajtes' probit'sja k poverxnosti, aktivno rabotaja rukami* – peut sembler, au premier coup d'œil, satisfaisante. La phrase n'est pas facile à traduire, mais l'on sait au moins que *avalanche* se traduit par *lavina*, comme le confirment tous les dictionnaires, et même les systèmes d'intelligence artificielle ne détectent aucun problème dans ce choix. Pourtant, un piège inattendu se cache ici. En russe, *lavina* ('avalanche') ne peut pas *commencer, durer, se terminer* ou *se produire*. À la différence du français, *lavina* n'est pas un nom d'événement (la chute d'une masse de neige), mais désigne la masse de neige elle-même (Kolzun et al., 2018).

C'est précisément parce que les composantes sémantiques centrales de LAVINA1 ('masse de neige [qui descend]') et d'AVALANCHE1 ('descente [d'une masse de neige]') ne coïncident pas que leur combinatoire diffère en russe et en français, ce qui peut être à l'origine de difficultés de traduction. Pour rendre cette différence visible aux étudiants dans ce cas concret, et pour les amener à développer une démarche réflexive sur le lexique, l'enseignant pourrait proposer un travail d'analyse comparée sur ces deux définitions :

Tableau 7.
Définitions pour une analyse comparée

LAVINA1 ⁴¹ 'avalanche'	AVALANCHE1
<i>lavina, kotoraja idët na Y</i>	<i>avalanche qui arrive sur Y</i>
'énorme masse de neige'	'glissement accidentel d'une énorme masse de neige'
• qui se déplace accidentellement très rapidement vers le bas d'un versant de montagne jusqu'à Y	• qui se détache d'une montagne ou d'un élément du relief similaire

⁴¹ лавина1 'огромная масса снега, • которая очень быстро перемещается вниз по склону горы на Y • (и наносит ущерб Y-y)' *Со склона соседней горы сорвалась лавина*. Les définitions de LAVINA1.1 et AVALANCHE1.1 ont été élaborées, dans le cadre du projet LEC-ru, par L. Kolzun, S. Krylosova et A. Polguère.

Tableau 7 (Suite)

LAVINAL 'avalanche'	AVALANCHE1
• (et cause des dommages à Y) <i>Du versant d'une montagne voisine, une avalanche s'est détachée.</i>	• qui se déplace rapidement sur cet élément du relief • qui atteint Y (et lui cause des dommages) <i>Une avalanche s'est produite hier dans le massif alpin en provoquant la mort d'une personne.</i>

Voici un autre exemple de difficulté lexicale, cette fois plus prévisible, qui peut se présenter aussi bien en cours de traduction que lors d'activités de grammaire ou d'expression. Tout enseignant de russe langue étrangère travaillant avec des étudiants francophones sait que la paire *vosxiščat'sja* et *ljubovat'sja* 'admirer' pose un réel problème. Ces deux verbes sont souvent introduits simultanément, au moment de l'apprentissage du cas instrumental, c'est-à-dire dès la première année. La difficulté tient au fait qu'ils se traduisent tous deux par 'admirer' en français. Il est tentant, pour l'enseignant, de recourir à la langue maternelle des étudiants et de proposer une distinction simplifiée : *vosxiščat'sja* correspondrait à 'admirer mentalement', et *ljubovat'sja*, à 'admirer visuellement'. Une telle explication, bien que pouvant sembler efficace sur le moment, ne permet pas de développer la capacité des étudiants à justifier leurs choix lexicaux en russe. En revanche, en s'appuyant sur le modèle LEC-LA, il est possible de proposer pour ces deux verbes des définitions précises et de construire, à partir de celles-ci, une série d'exercices ciblés de choix lexical.

Tableau 8.
Ébauche des définitions didactisées des verbes *vosxiščat'sja* et *ljubovat'sja*

VOSXIŠČAT'SJA	LJUBOVAT'SJA
<i>X vosxiščatsja Y_{instr}</i> 'La personne X éprouve et, éventuellement, exprime par des paroles Z un sentiment fort et agréable, suscité par l'objet ou la situation Y, qu'elle évalue très positivement.'	<i>X ljubuet'sja Y_{instr}</i> 'La personne X regarde Y et trouve Y très beau, ce qui fait que X ressent un plaisir intense.'

L'objectif n'est pas seulement que l'étudiant identifie le verbe approprié, mais qu'il soit en mesure, en russe, de justifier son choix par une explication claire – parfois longue, mais toujours accessible et transparente –, en mobilisant le vocabulaire et les structures syntaxiques déjà fournis dans les définitions. Voici un modèle de raisonnement :

Dans la phrase *Mark ... smelost'ju Žanny d'Ark* 'Marc... le courage de Jeanne d'Arc', je choisis le verbe *vosxiščat'sja* parce que Marc éprouve un sentiment intense et agréable, et qu'il évalue très positivement le courage de Jeanne d'Arc. Je ne peux pas dire ici **Mark ljubuetsja smelost'ju Žanny d'Ark* car il n'est pas possible de contempler courage comme tel.

Ainsi nous abordons un second usage de la définition pédagogique dans l'enseignement d'une langue étrangère : un usage plus ambitieux, qui consiste à concevoir un cours – voire une série de cours – structuré(s) autour d'une définition donnée.

2.3. Construire un cours autour d'une définition

Nous névoquerons ici que deux exemples de cours : l'un destiné aux grands débutants et l'autre aux étudiants d'un niveau intermédiaire.

Faire parler les étudiants en cours d'expression orale

Il s'agit d'un cours de 1h30 et conçu autour des définitions pédagogiques de *xolod1* et *moroz1* présentées plus haut, dans un groupe d'initiation (grands débutants, quinzième semaine d'apprentissage du russe). L'objectif est double : amener les étudiants à s'exprimer en russe tout en travaillant sur le sens lexical et sur les liens paradigmatiques (*xolodnyj* 'froid (adj.)') et syntagmatiques (*sobačij xolod* 'froid de chien'). Les définitions sont fournies en russe, et l'ensemble des activités se déroule exclusivement dans cette langue : les enseignants chargés de ce cours sont invités à ne donner aucune explication en français. Nous présentons ici, de manière synthétique, les modules de travail élaborés à partir de ces définitions⁴².

- Travail de paraphrasage (apprendre à dire autrement). Pour aider les étudiants à s'approprier rapidement les définitions lexicographiques des deux lexèmes, nous leur proposons de reformuler des phrases telles que *Na ulice moroz* 'litt. Dehors, il y a le froid (*moroz*)', etc.

⁴² Nous nous appuyons ici sur les analyses présentées dans Krylosova (2022). Le cours, extrait du manuel d'expression orale Krylosova (2025b), auquel il est fait référence, constitue un document de travail non publié, qui peut être communiqué sur demande aux collègues intéressés.

- Travail sur le champ sémantique. Nous demandons aux étudiants : a) soit de formuler les définitions des lexèmes comme *ŽARA* ‘chaleur’, *PROXLADA* ‘fraîcheur’, *TEPLO* ‘douceur’ ; b) soit, à partir d’une définition donnée par l’enseignant d’identifier le lexème défini. Par exemple : ‘*Žara* est une température de l’air. C’est une température élevée, et elle n’est pas confortable pour un être humain⁴³’.

Travail sur le potentiel dérivationnel des lexèmes. Nous nous intéressons aux adjectifs dérivés de *XOLOD1* et *MOROZ1* et, plus précisément aux syntagmes *xolodnyj den'* et *moroznyj den'* ‘journée froide’. Il s’agit dans ce cas d’une dérivation structurale des lexèmes étudiés, autrement dit d’un changement de partie de discours sans ajout de sens. Les exercices avec les adjectifs reposent donc sur la connaissance du sens des substantifs correspondants. Par exemple, les étudiants doivent déterminer s’il s’agit de *xolodnyj den'* ou *moroznyj den'* ‘journée froide’ dans la situation suivante : ‘Pendant une telle journée, il peut pleuvoir’, etc.

- Travail sur la combinatoire lexicale restreinte : Nous proposons aux étudiants une série de tableaux et d’exercices leur permettant d’observer la combinatoire de *XOLOD* *XOLOD1* et *MOROZ1*, et d’établir des listes de lexies pouvant se combiner avec l’un ou l’autre de ces lexèmes, ou avec les deux (*slabyj* ‘faible’ uniquement avec *MOROZ1* ; *sobačij* ‘de chien’ uniquement avec *XOLOD1* ; *sil'nyj* ‘fort’ – avec les deux, etc.).
- Travail avec les actants et les verbes supports (partie grammaticale du cours). Les étudiants travaillent sur les trois façons d’exprimer l’actant X de *MOROZ1* (*moroz de X degrés*) et sur deux verbes supports, sémantiquement vides en contexte, qui sont communs pour *XOLOD1* et *MOROZ1* : *byt'* et *stojat'*⁴⁴.
- Travail de distinction des quasi-synonymes. Nous demandons aux étudiants de compléter une dizaine de phrases avec *xolod* ou *moroz* – ou d’indiquer que les deux lexèmes sont possibles – selon le contexte. Cet exercice permet de vérifier à la fois la maîtrise du sens des deux quasi-synonymes et celle de leur combinatoire. Pour chaque phrase, les étudiants doivent argumenter en russe leur choix. Par exemple, *Dans la phrase numéro cinq, je choisis moroz, parce qu’ici, il y a l’épithète slabyj ‘faible’*⁴⁵.

⁴³ *Жара – это температура воздуха. Это высокая температура, и она некомфортна для человека.*

⁴⁴ Les étudiants révisent certains points de grammaire importants : 1) l’emploi des substantifs après les numéraux ; 2) le passé et le futur du verbe *byt'* ‘être’ (au présent, il est à la forme zéro) ; 3) la conjugaison du verbe *stojat'*.

⁴⁵ *Во фразе номер пять я выбираю слово мороз, потому что здесь есть эпитет слабый.*

- Production libre : décrire une image ou improviser un dialogue « Alexa (Alissa), quel temps fait-il aujourd'hui ? » en mobilisant tout le lexique acquis.

Exigeant même pour des locuteurs natifs, ce travail sur le lexique valorise les étudiants : après quelques mois seulement, ils peuvent discuter, en russe, des mots de la langue qu'ils étudient. Ce type de séance, centrée sur un couple de lexèmes, montre que comprendre une unité lexicale ne se réduit pas à la traduire, mais suppose de l'inscrire dans un réseau lexical, grammatical et culturel.

Expliquer un point de grammaire délicat

Nous présentons ici une stratégie d'enseignement portant sur un petit groupe de verbes (18 paires) dits *verbes de déplacement appariés (de couple)* non préfixés, tels que *idti / xodit' 'aller'*, dans le cadre d'un cours de grammaire théorique destiné aux étudiants de première année de Licence. Ce groupe constitue l'un des points les plus complexes de la grammaire russe.

Ces verbes se distinguent des autres verbes de déplacement – et, plus largement, de l'ensemble des verbes du russe – par le fait qu'ils sont les seuls à posséder deux formes imperfectives sémantiquement non équivalentes : l'une déterminée (*monorientée*) et l'autre indéterminée (*non orientée*). En raison de leur grande fréquence d'usage, ces verbes doivent être introduits dès les premiers stades de l'apprentissage du RLE. Pourtant, leur emploi soulève d'importantes difficultés pour l'apprenant : dès qu'il veut construire la phrase la plus simple avec l'un de ces verbes, il se heurte à un problème lié au choix du bon membre du couple verbale (Iordanskaja et al., 2020, Krylosova, 2025a).

Avant d'aborder ce thème en grammaire théorique, les étudiants ont déjà rencontré et utilisé ces verbes en pratique – leur fréquence rend impossible d'attendre la maîtrise de bases théoriques solides pour les introduire à l'oral. En revanche, c'est dans le cadre du cours théorique que nous les étudions pour la première fois de manière systématique. Notre démarche repose sur un examen comparatif des définitions d'une paire de ces verbes. La discussion s'ouvre par des questions fondatrices : « Que signifie *bežat'*^{DÉT} 'courir' (*letet'*^{DÉT} 'voler' ou *plyt'*^{DÉT} 'nager') ? » et, corrélativement, « Que signifie *begat'*^{INDÉT} 'courir' (*letat'*^{INDÉT} 'voler' ou *plavat'*^{INDÉT} 'nager') ? ».

Voici, en guise d'exemple, les définitions pédagogiques des verbes EXAT'1 et EZDIT'1 'se déplacer en véhicule' (les étudiants les reçoivent en russe) :

Tableau 9.

Définitions pédagogiques des verbes EXAT'1 et EZDIT'1 'se déplacer en véhicule'

EXAT'1^{DÉT46} (ja edu, ty edeš', oni edut) imperfectif <i>X edet na Y po Z v točku A iz točki B</i> 'La personne X se déplace en moyen de transport Y par une surface ou dans un espace Z' La personne X • se déplace ◦ au point A depuis le point B, et ce déplacement peut être observé à un moment donné. X = Nom. Y = na + Prép. (<i>na avtobuse</i>) Instr. (<i>avtobusom</i>) Z = po + Dat. (<i>po doroge</i>) point A = v + Prép., na + Acc., k + Dat. ... point B = iz + Gen., s + Gen., ot + Gen. ... EXEMPLES: 1. Ty edeš' sliškom bystro ! 'Tu va trop vite' [Procès en cours au moment de l'énoncia- tion]	EZDIT'1^{INDÉT47} (ja ezžu, tu ezdiš', oni ezdjat) imperfectif <i>X ezdit na Y po Z (v točku A iz točki B)</i> 'La personne X se déplace en moyen de transport Y par une surface ou dans un espace Z' La personne X • se déplace ◦ soit (a) vers le point A depuis le point B, mais ce déplacement est impossible à observer à un moment donné ; ◦ soit (b) « sans point A ni point B ». (a) X = Nom. Y = na + Prép. (<i>na avtobuse</i>) Instr. (<i>avtobusom</i>) Z = po + Dat. (<i>po doroge</i>) point A = v + Prép., na + Acc., k + Dat. ... point B = iz + Gen., s + Gen., ot + Gen. ... EXEMPLES : 8. Veclav často ezdit v Pariž na konferencii 'Wiesław se rend souvent à Paris pour des colloques.' [Déplacement fonctionnel habituel]
--	---

⁴⁶ ехать: X едет на Y по Z в точку A из точки B 'Человек X движется на транспортном средстве Y по поверхности или в пространстве Z'. При этом человек X • движется в точку A из точки B, и это движение можно наблюдать в какой-то момент'.

⁴⁷ ездить: X едет на Y по Z в точку A из точки B 'Человек X движется на транспортном средстве Y по поверхности или в пространстве Z'. При этом человек X • движется либо (a) в точку A из точки B, но это движение невозможно наблюдать в какой-то момент; либо (b) «без точки A и точки B».

Tableau 9 (Suite)

2. <i>Kogda my exali na poezde na jug, my igra-li v karty</i> 'Pendant que nous allions en train vers le sud, nous jouions aux cartes.' [Narration descriptive] 3. <i>Korga budeš' exat' mimo Ėjfelevoj bašni, sfotografiruj eě dlja menja</i> 'Quand tu passeras près de la tour Eiffel, prends-la en photo pour moi.' [Toile de fond pour un autre procès] 4. <i>Každyj god, kak tol'ko prixodit leto, Katia edet v derevnju</i> 'Chaque année, dès que l'été arrive, Katia va à la campagne.' [Procès pris dans une succession; événement déclencheur] 5. <i>Skol'ko vremeni vy exali do Berlina?</i> 'Combien de temps avez-vous mis pour aller à Berlin ?' [Durée d'un trajet] 6. <i>Obyčno my edem v Kanny čerez Pariž</i> 'D'habitude, nous nous rendons à Cannes en passant par Paris.' [Chemin suivi] 7. <i>Uže zavtra my edem v Varšavu!</i> 'Demain, nous partons pour Varsovie' [Anticipation sur une action à venir]	9. <i>Veclav eszdil v Pariž i privěz novyj roman Ėmanuelja Karrera</i> 'Wiesław est allé à Paris et a rapporté le nouveau roman d'Emmanuel Carrère.' [Déplacement fonctionnel dans une situation passée unique] (b) X = Nom. Y = na + Prép. (<i>na avtobuse</i>) Instr. (<i>avtobusom</i>) Z = po + Dat. (<i>po doroge</i>) (sans point A et point B) EXEMPLES : 10. <i>Devočke vsego tri goda, a ona uže ezdit na rolikax!</i> 'La petite fille n'a que trois ans et elle fait déjà du roller' [Acte générique: aptitude] 11. <i>Včera šěl sil'nyj sneg, i ja nikuda ne ezdit</i> 'Hier, il a neigé fort et je ne suis allé nulle part.' [Énoncé négatif au passé: difficulté de visualiser un mouvement qui n'a pas lieu] 12. <i>Uže celyj čas pered našim domom ezdit kakoj-to mal'čik na samokate</i> 'Voilà déjà une heure qu'un garçon circule en trottinette devant notre maison.' [Mouvement de va-et-vient] 13. <i>Mal'čik, ne ezdi zdes'!</i> 'Petit, ne roule pas ici !' [Énoncé négatif à l'impératif]
---	--

L'analyse minutieuse des éléments présents dans l'une des définitions et absents dans l'autre permet de mettre en évidence la non-substituabilité de ces verbes dans la plupart des contextes, et de répondre à la question : qu'y a-t-il dans le sens de ces verbes – souvent traduits de la même manière en français (*nager, courir,*

voler, etc.) – qui les empêche de s’interchanger dans un contexte donné? C’est par ce biais que nous introduisons la catégorie grammaticale de directionnalité (Iordanskaja *et al.*, 2020) :

Tableau 10.
Catégorie grammaticale de directionnalité d’après Iordanskaja et al., 2020

	point final : ‘au point A’	actualité: ‘au moment donnée’	exemple
monodirectionnalité (<i>exat’</i>)	+	+	<i>Smotri, kak bystro on edet na samokate prjamo na nas!</i> ‘Regarde comme il arrive vite en trottinette, droit sur nous!’
nondirectionnalité (<i>ezdit’</i>)	+	–	<i>Nina často ezdit k deduške v Monreal’</i> ‘Nina va souvent chez son grand-père à Montréal.’
	–	+	<i>Ne mogu usnut’ : mal’čiški ezdjat na velosipedax po dvoru u kričat</i> ‘Je n’arrive pas à m’endormir : des garçons font du vélo dans la cour et crient.’
	–	–	<i>Neuželi etot malyš uže ezdit na skejte?</i> ‘Est-il possible que ce petit fasse déjà du skate?’

Dans notre pratique, nous consacrons 4 h 30 à l’analyse de ces verbes en grammaire théorique⁴⁸, suivies d’un volume horaire équivalent en grammaire pratique. Il n’est donc pas possible, dans le cadre de cet article, de détailler l’ensemble de la progression pédagogique ; celle-ci est présentée pas à pas dans Krylosova (2025a).

⁴⁸ Depuis six ans, nous avons introduit dans nos cours un travail inédit sur les définitions, qui permet aux étudiants d’aborder les verbes de déplacement de manière active (Krylosova, 2025a). À la suite de cet exercice approfondi, le cours théorique « classique » consacré aux verbes de déplacement reprend l’analyse en examinant les critères de présence d’un point final et d’observabilité (ou actualisation) du déplacement, considérés comme déterminants dans l’opposition entre verbes déterminés et indéterminés. Nous nous appuyons ici sur la version la plus récente de ce cours (Bonnot & Krylosova, 2025), actualisée chaque année par l’intégration d’exemples contemporains. La version initiale, rédigée en 1991 par Christine Bonnot, s’inspirait notamment de Fontaine (1983), qui, dans le sillage de Karcevski (1927/2024) et de Foot (1967), avait introduit le critère d’« actualisation » comme élément central de la distinction entre les membres des paires de verbes de déplacement.

Ici, nous avons simplement souhaité souligner la possibilité d'exploiter des définitions lexicographiques même dans le cadre d'un cours de grammaire, et mettre en évidence l'intérêt qu'il y a à montrer aux étudiants l'interaction constante entre le lexique et la grammaire, telle qu'elle est mise en œuvre dans la méthode LEC-LA.

Conclusion

Loin d'être uniquement des dictionnaires au sens classique, les dictionnaires conçus dans le cadre de la LEC-LA constituent avant tout un véritable cadre méthodologique, apte à inspirer des dispositifs d'enseignement où lexique et grammaire se construisent conjointement. En s'appuyant sur les principes de la lexicographie active, l'enseignant peut créer ses propres définitions, même à partir de zéro, sélectionner les combinaisons les plus pertinentes, structurer les contenus en fonction des besoins concrets des apprenants et ancrer les acquisitions dans une compréhension fine des régularités du lexique. Utilisé comme matrice, ce type de dictionnaire permet de transformer l'analyse d'un mot en un parcours d'apprentissage complet, conduisant l'apprenant à observer, comparer, justifier et produire – dans la langue étudiée. On l'amène ainsi à se poser les bonnes questions, à envisager les mots sous l'angle de l'analyse du sens et de l'interconnexion des sens, démarche qui conduit à une véritable maîtrise de la langue.

Remerciements

Je voudrais exprimer toute ma gratitude à mes collègues du Département d'études russes de l'Inalco, qui ont accueilli avec bienveillance l'idée d'intégrer le travail lexicographique dans leur programme et qui m'ont offert un retour attentif, riche et stimulant. Leur ouverture a rendu possible cette expérience, menée chaque année avec une centaine d'étudiants. Merci du fond du cœur à ces étudiants, qui se sont prêtés au jeu avec curiosité et enthousiasme, apprenant à définir pour apprendre et apprenant en définissant. Ma reconnaissance va également à Alain Polguère, à tous les membres de l'équipe LEC-ru (CREE Inalco & ATILF CNRS), et aux évaluateurs de ce texte, dont les remarques constructives m'ont été précieuses. Merci à Paweł Golda pour sa patience et pour le travail accompli sur ce texte, particulièrement

exigeant sur le plan typographique. Enfin, je ne remercierai jamais assez Igor Melčuk. Avoir eu la chance de collaborer et d'échanger avec lui pendant plusieurs années a été un bonheur, un privilège et un apprentissage constant. Il a lu la toute première version de cet article et m'a confié des commentaires d'une finesse inestimable, qui ont nourri ma réflexion et suscité de riches discussions, même si j'ai parfois choisi de conserver ma propre formulation. Peut-être parce qu'une définition unique et parfaite n'existe pas ? Les imperfections qui demeurent sont, bien sûr, entièrement les miennes.

Références citées

- Alonso Ramos, M. (2003). Éléments du frame vs. Actants de l'unité lexicale. *Proceedings of the First International Conference on Meaning-Text Theory* (77–89). École Normale Supérieure.
- Apresjan, Ju. (1974). *Leksičeskaja semantika. Sinonimičeskie sredstva jazyka*. Nauka.
- Apresjan, Ju. (1980). *Tipy informacii dlja poverxostno-semantičeskogo komponenta modeli Smysl-Tekst*. Wiener Slawistischer Almanach.
- Apresjan, Ju. (1994). O jazyke tolkovaniij i semantičeskix primitivax. *Voprosy jazykoznanija* 4, 3–16.
- Apresjan, Ju. (1995). *Izbrannye trudy. Tom II. Integral'noe opisanie jazyka i sistemnaja leksikografija*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2004). *Novyj obščanitel'nyj slovar' sinonimov russkogo jazyka*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2006). *Jazykovaja kartina mira i sistemnaja leksikografija*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2010). *Prospekt aktivnogo slovarja russkogo jazyka*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2014a). *Aktivnyj slovar' russkogo jazyka. T. 1*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2014b). *Aktivnyj slovar' russkogo jazyka. T. 2*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2017). *Aktivnyj slovar' russkogo jazyka. T. 3*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2023). *Aktivnyj slovar' russkogo jazyka. T. 4 (1)*. Jazyki slavjanskoj kul'tury.
- Apresjan, Ju. (dir.) (2024). *Aktivnyj slovar' russkogo jazyka. T. 4 (2)*. Jazyki slavjanskoj kul'tury.

- Apresjan, Ju. & Apresjan, V. (1993). Metafora v semantičeskom predstavlennii èmocij. *Voprosy jazykoznaniija* 3, 27–35.
- Arutjunova, N. (1995). Istina i ètika. In N. Arutjunova & N. Rjabceva (éds), *Logičeskij analiz jazyka. Istina i istinnost' v kul'ture i jazyke* (7–23). Nauka.
- Arutjunova, N. (1998). *Jazyk i mir čeloveka*. Jazyki russkoj kul'tury.
- Bonnot, C. & Krylosova, S. (2025). *Les verbes de mouvement simples (déterminés et indéterminés)*. *Cours de grammaire théorique du russe de Chr. Bonnot complété par S. Krylosova*. Licence 1 [document de travail non publié]. Inalco.
- Bulygina, T. & Šmelëv, A. (1997). *Jazykovaja konceptualizacija mira (na materiale russkoj grammatiki)*. Jazyki russkoj kul'tury.
- Dostie, G., Melčuk, I. & Polguère, A. (1999). Méthodologie d'élaboration des articles du *Dictionnaire explicatif et combinatoire du français contemporain*. In I. Melčuk et al. (éds), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I–IV* (11–27). Presses de l'Université de Montréal.
- Evgenëva, A. (dir.) (1981–1984). *Malyj akademičeskij slovar' russkogo jazyka*, Russkij jazyk.
- Fontaine, J. (1983). *Grammaire du texte et aspect du verbe russe contemporain*. Institut d'études slaves.
- Footo, I. P. (1967). Verbs of motion. *Studies in the Modern Russian Language* 1, 4–33.
- Forsyth, J. (1963). The Russian verbs of motion. *Modern Languages* 44, 147–152.
- Glazunova, O. (2016). Ob Aktivnom slovare russkogo jazyka i sistemnom opisanii leksiki s točki zrenija teorii i metodiki prepodavanija russkogo jazyka kak inostrannogo. *Vestnik SPbGU* 9(2), 121–133.
- Inalco (2024). *Études russes*. <https://www.inalco.fr/etudes-russes>, consulté le 5.12.2024.
- Iordanskaja, L. (2007). Lexicographic Definition and Lexical Co-occurrence: Presuppositions as a 'No-go' Zone for the Meaning of Modifiers. In K. Gerdes, T. Reuther & L. Wanner (éds), *Proceedings of the Third International Conference on the Meaning-Text* (209–218). Wiener Slawistischer Almanach.
- Iordanskaja, L. & Melčuk, I. (2022). Names of Feelings in the Dictionary. *International Journal of Lexicography* 35(1), 20–52.
- Iordanskaja, L. & Paperno, S. (1996). *A Russian-English Collocational Dictionary of the Human Body*. Slavica Pub.
- Iordanskaja, L., Krylosova, S., Melčuk, I. & Mixel', P. (2020). Kategorija napravlennosti u russkix parnyx glagolov peremeščenija (na primere glagolov LETET' / LETAT'). *Voprosy jazykoznaniija* 1, 27–64.
- Karcevski, S. (1927/2024). *Système du verbe russe: essai de linguistique synchronique*. Institut d'études slaves.
- Kolzun, L., Krylosova, S. & Polguère, A. (2018). Idut laviny odna za odnoj : k leksikografičeskomu opisaniju russkogo i francuzskogo suščestvitel'nyx so značeniem 'lavina'.

- Proceedings of the International Conference on Russian Studies at the University of Barcelona (MKP-Barcelona 2018)*, 1204–1217.
- Krylosova, S. (2022). Enseigner le lexique « culturellement spécifique » : et si la solution était dans la définition lexicographique ? In L. Iomdin, Ja. Miličević, & A. Polguère (éds), *Lifetime linguistic inspirations: To Igor Mel'čuk from colleagues and friends for his 90th birthday* (253–266). Wiener Slawistischer Almanach, Peter Lang Verlag.
- Krylosova, S. (2024). Le temps des sentiments monochromes : pour une description lexicographique des lexèmes *nenavist'* 'haine' en russe contemporain. *Slovo*, 39–77.
- Krylosova, S. (2025a). O vzaimodejstvii grammatiki i slovarja na zanjatijax po RKI: glagoly peremeščeniya i ix leksikografičeskie tolkovaniya. *Russian Language Journal*, <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1499&context=rlj>, consulté le 25.09.2025.
- Krylosova, S. (2025b [2009 pour la première version]). *Mel'nica : cours d'expression orale pour francophones, partie 2*. Inalco, Paris. [Document de travail non publié].
- Levontina, I. (1995). Zvezdnoe nebo nad golovoj. In N. Arutjunova & N. Rjabceva (éds), *Logičeskij analiz jazyka. Istina i istinnost' v kul'ture i jazyke* (31–35). Nauka.
- Lux-Pogodalla, V. & Polguère, A. (2021). *Méthodes pour l'étude du lexique*. Université de Lorraine & ATILF-CNRS. [Document non publié], hal-03443313, consulté le 25.09.2024.
- Lux-Pogodalla, V. & Polguère, A. (2011). Construction of a French Lexical Network. Methodological Issues, *Proceedings of the First International Workshop on Lexical Resources*. Ljubljana, 54–61.
- Mašinnyj perevod i prikladnaja lingvistika*, 8 (1964). MGPIIJA.
- Mel'čuk, I. & Polguère A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. De Boeck & Larcier.
- Mel'čuk, I., Clas A. & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Duculot.
- Mel'čuk, I. (1974). *Opyt teorii lingvističeskix modelej «Smysl–Tekst»*. Nauka.
- Mel'čuk, I. (1988). *Dependency Syntax: Theory and Practice*. State University of New York Press.
- Mel'čuk, I. (1996). Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In L. Wanner (éd.), *Lexical Functions in Lexicography and Natural Language Processing* (37–102). Benjamins.
- Mel'čuk, I. (2012). *Semantics: From Meaning to Text. Vol. 1*. John Benjamins.
- Mel'čuk, I. (2013). *Semantics: From Meaning to Text. Vol. 2*. John Benjamins.
- Mel'čuk, I. (2024). An epoch has ended: to Apresjan's dear memory. *Russian Journal of Linguistics* 28(2), 480–483.

- Melčuk, I., Arbatchesky-Jumarie, N., Elnitsky, L., Iordanskaja, L. & Lessard, A. (1984). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques*. Vol. I. Presses de l'Université de Montréal.
- Melčuk, I., Arbatchesky-Jumarie, N., Dagenais, L., Elnitsky, L., Iordanskaja, L., Lefebvre, M.-N. & Mantha, S. (1988). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques*. Vol. II. Presses de l'Université de Montréal.
- Melčuk, I., Arbatchesky-Jumarie, N., Iordanskaja, L. & Mantha, S. (1992). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques*. Vol. III. Presses de l'Université de Montréal.
- Melčuk, I., Arbatchesky-Jumarie, N., Iordanskaja, L., Mantha, S. & Polguère, A. (1999). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques*. Vol. IV. Presses de l'Université de Montréal.
- Melčuk, I. & Polguère, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. De Boeck & Larrier.
- Melčuk, I. & Polguère, A. (2016). La définition lexicographique selon la Lexicologie explicative et combinatoire. *Cahiers de lexicologie* 109(2), 61–91.
- Melčuk, I. & Žolkovskij, A. (1984). *Tolkovo-kombinatornyj slovar' sovremennogo russkogo jazyka: opyty semantiko-sintaksičeskogo opisanija ruskoj leksiki*. Wiener Slawistischer Almanach.
- Miličević, J. (2008). Structure de la définition lexicographique dans un dictionnaire d'apprentissage explicatif et combinatoire. In E. Bernal & J. DeCesaris (éds), *Proceedings of the XIII EURALEX* (551–561). University Institute for Applied Linguistics, Pompeu Fabra University.
- Miličević, J. (2008). Structure de la définition lexicographique dans un dictionnaire d'apprentissage explicatif et combinatoire. In E. Bernal & J. DeCesaris (éds), *Proceedings of the XIII EURALEX* (551–561). University Institute for Applied Linguistics, Pompeu Fabra University.
- Miličević, J. (2016). La définition lexicographique pédagogique: enjeux et difficultés. *Cahiers de lexicologie* 109(2), 93–115.
- Polguère, A. (2013). Les petits soucis ne poussent plus dans le champ lexical des sentiments. In F. Baider & G. Cislaru (éds), *Cartographie des émotions. Propositions linguistiques et sociolinguistiques* (21–42). Presses Sorbonne Nouvelle.
- Reum, A. (1953), *Petit dictionnaire de style à l'usage des Allemands*. Édition remaniée par Henrik Becker.
- Vežbickaja, A. [= Wierzbicka, A.] (2002). Russkie kul'turnye skripty i ix otryaženie v jazyke. *Russkij jazyk v naučnom osvveščeni* 2(4), 6–34.
- Wierzbicka, A. (1972). *Semantic Primitives*. Athenäum.

- Wierzbicka, A. (1985). *Lexicography and Conceptual Analysis*. Ann Arbor MI.
- Wierzbicka, A. (1999). Emotional Universals. *Language Design: Journal of Theoretical and Experimental Linguistics* 2, 23–69.
- Wierzbicka, A. (1999). Emotional Universals. *Language Design: Journal of Theoretical and Experimental Linguistics* 2, 23–69.
- Zaliznjak, A., Levontina, I. & Šmel'ev, A. (2005). *Ključevye idei russkoj jazykovoj kartiny mira*. Jazyki slavjanskoj kul'tury. Nauka.
- Žolkovskij, A., Leont'eva, N. & Martem'janov, Ju. (1961). O principial'nom ispol'zovanii smysla pri mašinnom perevode. *Mašinnij perevod* 2, 17–46.

Rédaction / Copy editing

Paweł GOLDA (textes français / French texts)

Marzena WYSOCKA-NAREWSKA (textes anglais / English texts)

Projet de la couverture et de la page de titre / Cover and front page design

Tomasz JURA

Préparation de la couverture à l'impression / Cover preparation for printing

Paulina DUBIEL

Composition / Typesetting

Paulina DUBIEL

ISSN 2353-088X

Attribution – Partage dans les mêmes conditions 4.0 International (CC BY-SA 4.0) /

Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)



Auparavant, la revue paraissait sous forme imprimée suivie de l'identificateur ISSN 0208-5550 /

This journal was formerly published in print with the following identifier ISSN 0208-5550

La version primaire référentielle de la revue est sa version électronique (en ligne) /

The primary referential version of the journal is its electronic (online) version

La revue est distribuée gratuitement /

The journal is distributed free of charge

Maison d'édition / Publisher

Wydawnictwo Uniwersytetu Śląskiego / University of Silesia Press

ul. Jordana 18, 40-043 Katowice

<https://wydawnictwo.us.edu.pl>

e-mail: wydawnictwo@us.edu.pl

Printed sheets: 25.75. Publishing sheets: 27.5.

Free copy

ISSN 2353-088X

9 772353 088509

5 5

For more details

