

## ФОРМУВАННЯ СЕМАНТИЧНОЇ СТРУКТУРИ СЛОВА «КОРПУС» У ФРАНЦУЗЬКІЙ ЛІНГВІСТИЧНІЙ ТРАДИЦІЇ

Страшко Ірина Володимирівна

кандидатка філософських наук,

докторантка кафедри прикладної лінгвістики, порівняльного мовознавства та перекладу  
Національного педагогічного університету імені М. П. Драгоманова  
вул. Тургенєвська, 8/14, Київ, Україна

Аналізуються генеза і семантичний розвиток слова «корпус» у французькій мові. Залучення порівняльно-історичного, описового, контекстного і дефініційного методів дало можливість встановити смислові відмінності в різні періоди його вживання, експлікувати його семантичну структуру, уточнити денотативні, сигніфікативні і конотативні значення, що її реалізують, виявити появу нових семем, провести спостереження над його трактуванням у словниках французької мови, враховуючи й доступні на лексичному порталі платформи CNRTL. Проведений історико-етимологічний аналіз дає підстави говорити про розгортання семантичної структури слова: від релігійного поняття до мовного терміна. Його інтерпретація в лінгвістичному ключі зустрічається у лексикографічних джерелах першої половини ХХ ст., а з шістдесятих років минулого століття воно набуває нового значення, пов'язаного з корпусною лінгвістикою. Становлення терміна «корпус» у французькому мовознавстві супроводжувалося пошуком його релевантних характеристик і спробами дати йому класифікацію. Розширення самої типології корпусів зумовило виникнення й відповідних номінацій. Останні утворені синтаксичним способом і характеризуються різною структурою. Виділення значимих характеристик корпусу реалізується атрибутивними компонентами, які конкретизують вихідне поняття, утворюють його видові кореляти. При цьому базовою моделлю такого термінотворення є атрибутивна словосполучка з ад'єктивною постпозицією; кількісно домінуюча модель – двокомпонентний термін безприменникової структури з атрибутивним зв'язком; продуктивною є модель N+ PART; менш частотною – модель субстантивного поєднання з включеним прийменником. Зміна частоти вживання аналізованого слова показана за допомогою мовних сервісів Google Books Ngram Viewer та Usage à travers le temps. Попри те, що зазначені пошукові інструменти ґрунтуються на різних підходах і оперують різними масивами даних, отримані результати дозволяють констатувати схожі тенденції.

**Ключові слова:** французька мова, корпус, етимологія, семантика, корпусна лінгвістика.

Використання писемних джерел, що відповідають елементам мови, і вважаються гідними передавання майбутнім поколінням (релігійні, епіграфічні, літературні тексти, написи тощо), експлуатація яких зумовлена проблемами збереження, має довгу традицію. Якою б сучасною не здавалася ідея про те, що мовні дослідження можуть ґрунтуватися виключно на зібраннях даних, і обґрунтування гіпотез можливе з посиланням на випадки зустрічальності, вона не нова [8], оскільки її існування є нерозривно пов'язаним з поняттям корпусу. Поряд зі стійким дослідницьким інтересом до корпусних досліджень загалом, пошуком точного визначення корпусу, який як емпіричний об'єкт, зазнає стількох пояснювальних варіацій, рефлексією над корелюючими поняттями тексту, архіву, дискурсу, у французькій лінгвістичній традиції значимим є аналіз семантичної історії слова корпус. Водночас французька академічна думка, утім як і французька мова, в останні десятиліття лишуються на периферії наукових пошуків українських дослідників. **Актуальність** цієї розвідки

зумовлена відсутністю діахронічного аналізу становлення лексеми *корпус* у французькій мові в роботах українських дослідників.

Теоретичним підґрунтям дослідження стали розвідки українських і зарубіжних науковців в галузі корпусної лінгвістики (О. Демська-Кульчицька, В. Жуковська, В. Захаров, Є. Карпіловська, В. Плунгян, В. Риков, О. Baude, J. P. Dalbera). Серед **останніх** слід виділити роботи G. Bergounioux, I. Eshkol-Taravella та A. Lefeuvre-Halftermeyer, P. Charaudeau, A. Herschberg Pierrot, S. Mellet, в яких розглядаються теоретичні і методологічні передумови створення корпусів для французької мови, формулюються епістемологічні проблеми їх конституювання і експлуатації, вивчається роль корпусних досліджень. Окремо варто відзначити дослідження J. Pruvost щодо історичного, культурного й лінгвістичного розвитку поняття *корпус* у французькій мові. Методологічною основою роботи є аналіз і узагальнення наявного наукового досвіду з можливими підходами до дефініції досліджуваної лексичної одиниці.

**Метою статті** є історико-етимологічний аналіз трансформацій значення лексеми *корпус* у французькій мові. Зазначена мета досягається вирішенням взаємопов'язаних **завдань**:

– розглянути етапи історичного розвитку і основні тенденції використання слова *корпус* у французькій мові;

– показати динаміку частотності вживання аналізованої лексичної одиниці за обраний діапазон років.

Розвиток аналізованої лексичної одиниці простежується за допомогою тлумачних і етимологічних словників, тезаурусів сучасної французької мови, (онлайн) енциклопедій [15-29] та набору комп'ютеризованих мовних ресурсів і засобів обробки мови, розміщених на лексичному порталі платформи Національного центру текстів та лексичних ресурсів (<https://www.cnrtl.fr/>).

На загальну думку, у французькій мові слово *corpus* сходиться до латинських коренів. Вважається, що витoki цієї лексичної одиниці лежать у релігійній площині: його пов'язують із францизованим латинським *corpus Deu – тіло Христа*, що позначає Євхаристію (фр. *le corps du Christ*) [14, с. 389; 21]. Зустрічається вказівка на запозичення з християнізованої латини *corpus Domini, corpus Christi* [21], яке, до речі, не фіксується словниками французької мови XVII-XVIII століть. Версія релігійних коренів слова виходить зі значної ролі церкви в житті тогочасного французького суспільства, що визначило його місце у словниковому складі французької мови і надовго усталило його сприйняття у повсякденній свідомості французів як церковного терміна. Приміром, у 8-му виданні *Dictionnaire de l'Académie française* (1932-1935) зазначається, що *corpus* походить з класичної латини у складі виразу *corpus juris* (укр. *збірка законів*) [17]. Інше пояснення говорить про його німецьке походження: запозичення з нім. *corpus / korpus*, засвідчене з 1787 р. для позначення набору даних певного типу, зібраних для наукового вивчення [21]. Серед версій проникнення аналізованого слова у лексичний склад французької мови є й та, яка пов'язана з діяльністю кореспондента Інституту Франції, професора французької літератури Геттінгенського університету Шарля Віллера, в обов'язки якого входило складання звіту про стан літератури в Німеччині. Припускають, що саме він ввів у вжиток термін *корпус* [21]. Однак, в цілому, шляхи проникнення цього слова у французьку мову (вперше воно зафіксоване у старофранцузькій мові кінця XII ст.) зрозумілі: домінуюча версія латинського походження слова *corpus* ґрунтується

перш за все на релігійній суті позначуваного явища, чим воно, безумовно, не обмежується. На думку Ж. Прюво релігійні конотації все ще залишаються помітними й у XX столітті. Посилаючись на приклади зі словника *Trésor de la langue française*, 1957 р., він пояснює це тим, що в рамках християнської традиції слово *корпус* латиною відтворювалось під час кожної церковної служби, зокрема й у складі священницьких т.зв. «слів посвяти» *Hoc est enim corpus meum* (укр. *Це тіло моє*) і тому неусвідомлено є добре вкоріненим у колективній пам'яті французького народу і його узусі. У цьому вузькому, релігійному значенні, яке збереглося і сьогодні, зустрічається, приміром, у Стендаля: "*C'est demain la fête du Corpus Domini – la Fête-Dieu*" [14, с. 389]. До речі, назва *la Fête-Dieu* існує тільки у французькій мові. Більш відома офіційна назва цього релігійного свята Католицької Церкви – *Fête du Saint-Sacrement* або *fête du Corpus / Corpus Christi* – сходиться коренями до латино – католицької назви – *ceremonitas Sanctissimi Corporis* та *Sanguinis Christi* [16].

Зважаючи на давнє існування цього поняття як явища народного життя, воно досить рано з'являється в писемних джерелах – згадка про корпус засвідчена в епічній поемі кінця XII сторіччя "*La mort de Garin le Loherain*", авторство якої приписується Жану де Флажі (Jean de Flagy). Згодом лексема *корпус* реєструється у Біблії Гіо у 1206 у складі латинської назви *corpus Domini* – «Свято Тіла і Крові Христових» [21]. Ця знаменита Біблія є твором Гіо де Провена, відомого завдяки двом сатиричним поемам, в яких піддавались нещадній критиці як вади простого люду, так і аморальність дворян. Цікаво, що в середньовіччі слово «біблія» слугувало синонімом «сатири» [14, с. 389] й означало сатиричний твір, в якому осміювались основні суспільні стани (<https://www.cnrtl.fr/definition/bible>). Згодом вживання слова *корпус* реєструється лише у двох джерелах, датованих відповідно кінцем XVI ст. і першою половиною XVII ст. [21]. Поступово воно витісняє інше – *corps* (теж латинського кореня і в тому ж значенні), поодинокі свідчення якого траплялись ще в кінці XIII століття. Останнє не відмічається джерелами протягом більш, ніж двох століть і знову з'являється тільки починаючи з 1671 року (зокрема *Trésor de la Langue Française* дає його у трактуванні «*застарілий синонім до слова корпус*») [21]. В середні віки і до XVIII століття аналізоване слово у значенні «просфора» метафорично позначає «тіло Христа» [12], що є результатом його співвіднесення з позамовною дійсністю. Ця «органічна метафора», за Анн

Хершберг П'єро, імпліцитно існує й сьогодні, оскільки корпус, як сукупність висловлювань, «дійсно є тілом теорії, яку він ілюструє, створенню якої він слугує». [Там само].

Лексема *корпус* не реєструється одномовними словниками XVII-XVIII століть. У XIX ст. вона зафіксована у словниках 1845 року (*Grand dictionnaire classique de la langue française*), 1863 року (*Dictionnaire de la langue française*) і 1869 р. (*Grand Dictionnaire universel du XIXe siècle*), словникові дефініції яких мало різняться: корпус позначає збірник римського права – *corpus juris*. Жан Прюво дотримується думки, що повна латинська назва (лат. *corpus juris civilis*), яка іноді вживалась й у зредукваному варіанті *juris civilis*, певним чином передбачає значення, яке шукане слово набуде у другій половині XIX століття, позначаючи колекцію документів, що стосуються однієї окремої дисципліни. До цих пір, за Прюво, «будучи дещо недооціненим лексикографами», *corpus* залишається добре відомим юридичним терміном і не реєструється лінгвістичними термінологічними словниками [14, с. 391].

Відмітно, що у другому томі другого видання словника *Le Littré*, датованого 1873 роком, вживання лексеми *corpus* фіксується у складі запозиченої терміносполуки зі сфери англійського права – *habeas corpus*. Наводиться також історична довідка з детальним поясненням походження і змісту цього юридичного терміна. Зазначається, що “*Habeas Corpus*” – закон, прийнятий у травні 1679 р. під час правління Карла II, що закріпив особисту свободу всіх громадян. Згідно з ним, будь-яка ув'язнена особа (або інша людина від її імені) могла звернутися до лорд-канцлера або одного з дванадцяти великих суддів Англії з проханням отримати судовий напис або ухвалу *habeas corpus*. Тюремник був зобов'язаний доставити в'язня до судді та засвідчити, ким і за що той був ув'язнений. Суддя міг відпустити ув'язненого на волю або допустити його до внесення застави, за умови, що йдеться не про державну зраду чи скоєння одного з тяжких злочинів, перелік яких вказаний у ордері на арешт [20]. На платформі CNRTL перше свідчення виразу *habeas corpus* у французькій мові датується 1672 роком. Там же пропонуються трактування, які змістовно включають уявлення про закон, що дозволяє уникнути юридичного свавілля та відповідну постанову, яка позначає права і роль громадянина у процесах функціонування карної системи: *A. Loi votée au XVII<sup>e</sup> siècle par le Parlement anglais et garantissant la liberté individuelle, qui évite l'arbitraire de la détention par une*

*justification judiciaire de celle-ci en donnant le droit au détenu de comparaître immédiatement. B. (Ordonnance d')Habeas Corpus. Acte délivré à la requête d'un détenu en vertu duquel ce dernier doit être amené immédiatement devant le juge qui doit vérifier les motifs de la détention et prononcer éventuellement sa mise en liberté définitive ou sous caution.* (<https://www.cnrtl.fr/definition/habeas%20corpus>). У світлі викладеного варто пояснити, що *habeas corpus* є частиною латинської фрази *habeas corpus ad subjiciendum*, буквально: «представ арештованого особисто до суду» [22]. Отже, вживання лексеми *corpus* як правового терміна, що виник на основі рецепції понятійно-термінологічного апарату римського права, можна вважати наступним етапом її входження у французьку мову XVII сторіччя.

П'єр Ларусс є одним із перших лексикографів, який запропонував визначення, що не є ні релігійним, ані юридичним. Зрештою, наведена ним словникова дефініція близька за змістом в тому сенсі, чим може бути корпус у сучасному розумінні. Так, у четвертому томі *Grand Dictionnaire universel du XIXe siècle*, опублікованому у 1869 році, у колонці зі словом *corpus* вказується, що це латинське слово було взяте за назву для позначення повної колекції літературних творів або найважливіших документів: “*Corpus : mot latin qui signifie corps et que plusieurs savants ont adopté comme titre pour désigner un recueil complet de pièces littéraires d'un genre donné ou de documents essentiels.*” [Цит. за: 14, с. 390]. У цьому визначенні поняття *корпус* в першу чергу стосується збірника текстів, що становлять певну єдність жанру чи періоду. У цьому ж виданні, для ілюстрації, згадується збірка творів латинських поетів, наводиться назва релігійного свята (*Corpus Christi*), дається короткий словникова стаття, присвячена прихильникам корпускулярної філософії з відповідним терміном – *corpusculiste* – та стаття з аналізом поняття *corpus delicti*.

Доречно буде також згадати, що у 1815 р. термін *corpus* був обраний Берлінською академією для назви збірки графічних свідчень мов, які вийшли з ужитку “*Corpus inscriptiones graecorum*”, перший том якої вийшов у 1828 р. і передував появі «*Corpus Inscriptionum Latinarum*», як об'єднаний в єдиному томі колекції усіх творів латинських поетів, створеної Теодором Моммзенем у 1847 році [9]. Згадуючи “*Corpus inscriptionum graecarum*”, Ларусс дає інформацію про те, що корпус «був написаний під керівництвом Бека, за рахунок Берлінської академії» і «містить усі епіграфічні документи, які збереглися грецькою мовою».

[Цит. за: 14, с. 392]. У цьому ж контексті слід назвати й "Corpus poetarum latinorum" – збірку творів латинських поетів, укладену у Франкфурті в 1833 році та "Corpus scriptorum historiae Byzantinae" – збірник грецьких істориків Візантійської імперії, з перекладом на латину. [Там само, с. 391]. «Корпус навіть стає власне французьким словом, як тільки воно більше не вкладається в латинську фразу, і ми говоримо про Corpus des Troubadours або про Corpus des poètes de la Renaissance». [9, с. 1-2]. На думку Дальбера, такий тип корпусу – а він може складатися не тільки з текстів, а й зі слів, записаних або транскрибованих усних свідчень – не є «прерогативою лінгвіста», адже з аналогічними об'єктами працює і історик, і філолог, і юрист, навіть якщо «евристична перспектива» кожного з них буде помітно іншою, враховуючи різноманітність самих матеріалів, диверсифікованість підходів і різноманітність концепцій, пов'язаних з галузями дослідження [Там само].

Отже у XIX столітті слово *corpus* секуляризується і починає вживатися на позначення збірки творів та документів, що створена за аналогією наукового довідника й стосується однієї дисципліни [12].

Інтерпретація слова *корпус* в лінгвістичному ключі зустрічається у лексикографічних джерелах першої половини XX сторіччя. Приміром, у 9-му виданні Dictionnaire de l'Académie française (1935) воно трактується як сукупність документів, що використовуються для лінгвістичного аналізу [17]. Разом з тим, у словниках того періоду, зокрема у 8-му виданні Dictionnaire de l'Académie française (1932-1935), зберігається й найбільш загальне визначення корпусу, як збірки творів, документів, яке тією чи іншою мірою повторюється в лексикографічних джерелах і далі: "Recueil de pièces, de documents concernant une même matière" [18].

Спираючись на цитату із твору Е. Анріо "Les Romantiques: Courrier littéraire du XIXe siècle", опублікованого у 1952 р. про «дорогоцінний корпус неопублікованих листів і документів», Ж. Прюво звертає увагу на те, що слово *corpus* «вже добре інтегрувалось поза релігійним контекстом». «Відтепер, літературні тексти утворюють безперечний «корпус»», – стверджує дослідник і робить висновок про те, що появи власне лінгвістичного корпусу як мовного феномена передували вживання цього слова для опису зібрання релігійних, епіграфічних, музичних, літературних творів. Він обстоює думку про те, що у 1972 році, у день публікації Grand Larousse de la langue française, за редакцією Луї Гілберт, Рене Лаган і Жорж Ніобей,

лінгвістичний корпус народився як «формула і експлуатаційний концепт». Ілюстрацією зазначеної точки зору слугують дані словникової статті, на яку посилається науковець. Констатуючи її обмежений обсяг, однією з причин він називає брак досліджень для створення оглядової статті, присвяченої корпусам. Як зауважує Прюво, автори словника наводять корпус як скорочене від латинського *corpus juris* і визначають його спочатку як «збірник текстів римського права», у другому, дещо розширеному розумінні, – як «збірник документів тієї самої дисципліни. І тільки у третій дефініції корпус – «сукупність документів, що слугують основою для опису чи вивчення лінгвістичного явища». «Ігнорувати цей особливий сенс вже неможливо», – наголошує він [14, с. 393]. Перший слід слова *корпус*, «наділеного значенням, характерним для мовознавства» фіксується у словнику Le Robert 1961 року [Там само, с. 392]. Згодом, дуже точна дефініція корпусу як сукупності текстів, створених для лінгвістичного вивчення дається у шостому томі Trésor de la Langue Française, опублікованому у 1978 році. Хоч у такому трактуванні все ще не йдеться про справжній «мовний корпус», ілюстративні приклади, що його супроводжують (*le corpus des textes, d'une revue; un corpus littéraire; le corpus du vocabulaire français*), дають підстави говорити про виникнення самого поняття «мовний корпус». Проте вже у другому виданні Grand Larousse Universel 1995 року наводиться більш деталізоване формулювання: «Сукупність письмових чи записаних тверджень, констатуваних для їх лінгвістичного аналізу» [14], хоч і воно стосується лише текстових корпусів. Визначення, подане у передмові до Trésor de la langue française, с. 20: «Словникова номенклатура та лексикографічна інформація, яка приєднується до лексичних одиниць, що її складають [...] міститься не лише у встановлених хронологічних межах; яка також є функцією використаної документації того, що більш академічною мовою [...] називають корпусом» [Там само] є, мабуть, найбільш вдалим, зокрема тому, що може бути застосованим не лише до процесу, але й до результату, тобто до конкретних фактів укладання лінгвального матеріалу, який може бути названий «корпусом». У словниковій статті Petit Larousse illustré 1990 р. лексема *corpus* з маркуванням «дидакт.» характеризується змістом, що дозволяє констатувати збереження семантичного ядра аналізованої лексеми: «Набір текстів, документів, наданих за традицією або зібраних для вивчення, зокрема для лінгвістичного дослідження»

[15, с. 264]. Прикметно, що ця позначка зберігається й у сучасному електронному словнику Le Robert. Проте значення є вузькоспрямованим і передбачає конкретне використання: “*Ensemble fini de textes choisi comme base d’une étude*” [28].

У комп’ютеризованій версії словника XIX і XX століть – Le Trésor de la Langue Française Le Trésor de la Langue Française Informatisé – групуються усі попередні визначення лексеми, з відповідними історично обґрунтованими позначками («філол., гум. і суспільні науки», «лінгв.», «спеціальне»), тим самим підтверджуючи збереження основних компонентів значення, й додається інше, яке стосується сучасних технологій комп’ютерної обробки даних: «*Набір даних, які можна використовувати в аналізі або в автоматичному пошуку інформації*» [29], що свідчить про розширення її семантичного обсягу. Le grand dictionnaire terminologique (GDT), створений науковцями Квебекського офісу французької мови [27] демонструє розмежування трьох семем: 1) «*Сукупність усних та письмових джерел, що стосуються досліджуваної галузі та які використовуються у термінологічній роботі*». З позначками «філософія» і «право» наводяться й два інші варіанти значення: 2) «*Обмежений набір елементів (висловлювань), на яких ґрунтується дослідження явища*»; 3) «*Корпус (тіло) становить матеріальний елемент володіння, позначаючи фактичну владу, яка здійснюється над річчю*». Порівняно зі словниками попереднього етапу, автори від початку відштовхуються від трактування, що ґрунтується на сфері використання терміна. Воно й не дивно, оскільки сама ідея створення такого словника виникла із потреб, пов’язаних з введенням у Квебеці французької мови як робочої. З 1974 року сотні тисяч файлів, отриманих в результаті термінологічних досліджень, були депоновані в Банку термінології Квебека (BTQ), який в 1997 році став Le grand Dictionnaire terminologique [Там само].

Принципових відмінностей у визначеннях поняття *corpus* не спостерігається й у інших джерелах [24; 26]: як правило, це сукупність документів, що стосуються тієї ж теми, дисципліни/дослідження, письмових або записаних висловлювань, зібраних для збереження або (лінгвістичного) аналізу. Але таке широке та розмите значення означає конкретне використання та є пов’язаним з певними галузями знань. Приміром, у галузі права термін відноситься до *corpus juris*, тоді як у фізиці, за Дальбера [9], він майже не використовується, а лексема *corpuscule* набуває значення *particule* «частинка», дискретна складова речовини і не є демінутивом.

Як зазначає С. Меле, хоч поняття корпусу й здається досить простим і міцно закріпленим у певних традиціях гуманітарних і соціальних наук – корпус складається з набору даних, відібраних і зібраних для конкретної дисципліни, – з огляду на різноманітність практик, цілей створення і експлуатації корпусів, воно дещо ускладнилося за останні десятиліття [13, с. 1]. Не в останню чергу це пов’язано з тим, що становлення терміна «corpus» у лінгвістиці супроводжувалося пошуком релевантних характеристик цього способу досліджень мовних явищ і спробами дати йому класифікацію. Саме складність цього мовного феномена, проблеми, спричинені розумінням корпусу, призводять до розмаїття думок і наявності безлічі визначень терміну «корпус» у роботах франкомовних авторів (від широких до більш конкретних, вузькоспрямованих): «регульована принципами сукупність емпіричних даних, текстів (або фрагментів текстів), які є відповідними для певного дискурсу, і наділені відповідно репрезентативною цінністю» [8]; «сукупність елементів, на яких базується вивчення мовного явища» [9]; «упорядковані колекції записів усних та мультимодальних творів» [7]. Ми не вважаємо за потрібне зупинятися на аналізі трактувань терміну *corpus* в україно- і англійськомовних працях, оскільки вони вже досить детально прописані (див. напр. [1; 2; 3; 4, с. 43-45]). Однак, відсутність «авторитетного визначення, на яке можна посилається, коли необхідно обґрунтувати корпус», можливо, «[...] є симптомом того факту, що корпус не існує сам по собі, а [...] залежить від теоретичного позиціонування, на основі якого він розглядається» [11, с. 37], що, як і раніше, залишає місце для неоднозначності.

Як свідчить мовний матеріал, отриманий у результаті вибірки зі спеціальної літератури (як сфери функціонування), *corpus*, як один із базових термінів в корпусній лінгвістиці, є стрижневим в атрибутивних конструкціях, крім, власне, назви предметної галузі – *linguistique de corpus*. Ці номінації утворені синтаксичним способом, з різним планом вираження (структурою), як-то: N + A, N+ PART, N<sub>1</sub> + PREP + N<sub>2</sub>+A, N<sub>1</sub> + Art + N<sub>2</sub>, N<sub>1</sub> + PREP + N<sub>2</sub>, N<sub>1</sub> + PREP + N<sub>2</sub>+ PART. Виділення значимих характеристик корпусу реалізується атрибутивними компонентами, вираженими здебільшого прикметниками, дієприкметниками минулого часу різної структури; менш продуктивною є модель субстантивного поєднання з прийменником *de*. Вони конкретизують вихідне поняття, утворюють його видові кореляти, що є підставою для

типологізації корпусів. Отже, залежно від виду мовного матеріалу розрізняють: *corpus de textes*, *corpus d'écrits*, *corpus oral / corpus d'oralité*, *corpus de mots*, *corpus d'expressions*, *corpus d'unités de langue*. Інший поділ проводиться між корпусами, реалізованими як репрезентативні зразки фактів мови (*corpus représentatif*, *corpus de référence*) і тими, які є вичерпними у певній галузі (*corpus exhaustif*, *corpus partiel*, *corpus échantillonné / corpus d'échantillons*). За характером даних, що становлять корпус (*corpus ouvert*, *corpus clos*, *corpus non clos*, *corpus de suivi*), обсягом (*corpus de grande taille*) критеріями відбору матеріалу (*corpus construit – corpus non construit*), способом і ступенем обробки мовного матеріалу (*corpus aligné*, *corpus brut*, *corpus nu*, *corpus ordonné*, *corpus annoté*, *corpus étiqueté*), хронологічними рамками (*corpus chronologique*), спрямованості на певні типи досліджень (*corpus du français*, *corpus de français parlé*, *corpus de langue parlée*, *corpus des régionalismes*, *corpus du vocabulaire*, *corpus discursif*), типом джерельної бази

(*corpus littéraire*, *corpus linguistique*, *corpus lexicographique*, *corpus de presse*, *corpus en lettres et sciences sociales*), принципами відбору текстів (*corpus parallèle*, *corpus comparable*). В кінцевому підсумку, множинні комбінації цих варіантів утворюють досить багату палітру корпусів і «так чи інакше фіксують основні компоненти цього поняття» (Захаров).

Динаміка частотності слова *corpus* у писемній французькій мові наглядно простежується за допомогою мовного сервісу Google Books Ngram Viewer. На рис. 1 наведено криву відносної частоти його появи за кожен рік, розраховану при т. зв. «нульовому згладжуванні» (smoothing of 0) з часовим охопленням з 1600 по 2019 р.

З огляду на те, що при відсутності згладжування сервіс надає тільки необроблені дані, для отримання більш виразних тенденцій зустрічальності аналізованого слова, коефіцієнт згладжування був збільшений до 10 (рис. 2).

За визначений діапазон років (1600-2019 рр.) побудовані графіки в цілому не показують суттєвих відмінностей, що відповідає вищенаведеним

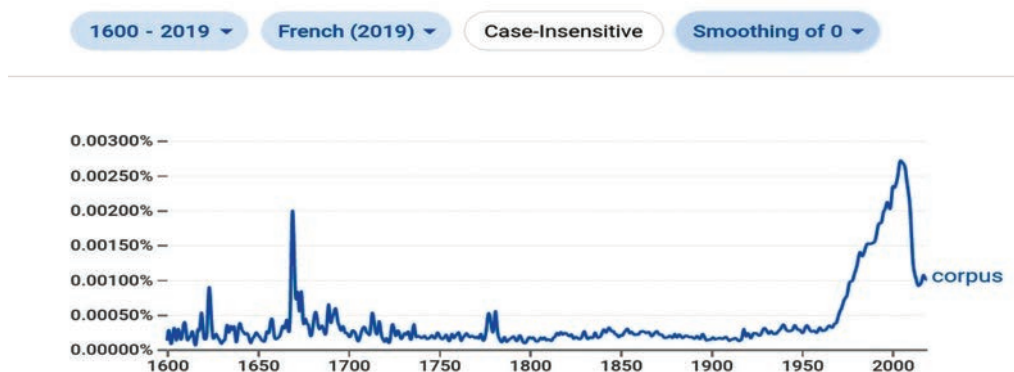


Рис. 1. Графік частоти згадування слова *corpus* у книгах французькою мовою, розрахована при нульовому згладжуванні у сервісі Google Books Ngram Viewer

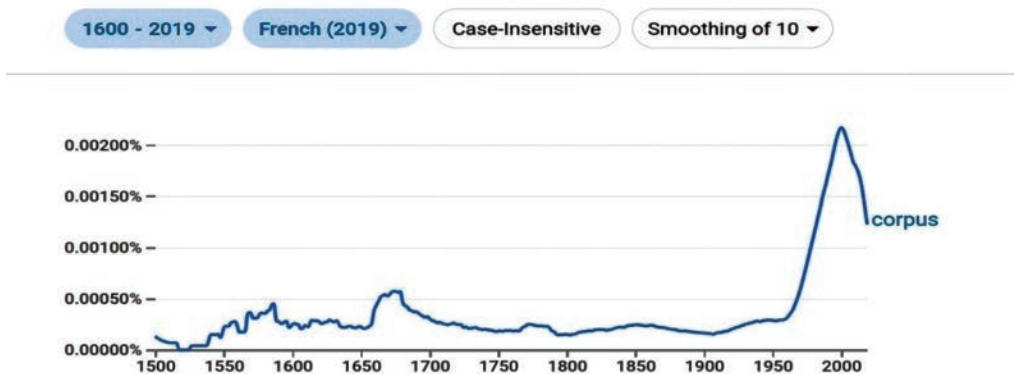


Рис. 2. Графік частоти згадування слова *corpus* у книгах французькою мовою, розрахована при коефіцієнті згладжування 10 у сервісі Google Books Ngram Viewer

уявленням про історичний розвиток аналізованої лексичної одиниці. Приміром, наведені фрагменти текстів XVI-XVII століття виключно на релігійну, церковну тематику; знайдені за запитом посилання на тексти початку XVII ст. і до кінця восьмидесятих років XX ст. – це, переважно, уривки із філософських, епіграфічних, правничих і, пізніше, літературних творів; наступний період включає фрагменти літературних творів; з 1999-2006 рр., крім власне літературних, наводяться й фрагменти мовознавчих робіт; з 2007-2019 р. домінують посилання на праці лінгвістичного характеру, в яких аналізується створення і експлуатація різних корпусів. Цікаво, що крива, яка відображає вживання слова *corpus*, починає збільшуватися на початку 1960-х років (що, з огляду на його вживаність у лексикографічних джерелах того часу, є цілком обґрунтованим) і сягає максимальних значень у 2000-х роках, що, вірогідно, викликане як розвитком корпусної лінгвістики у Франції і франкомовних країнах, так зростанням кількості корпусних ресурсів в цілому. Після чого крива опускається, а з 2016 р. фіксується незначне зростання кількості згадувань (рис. 1).

Для порівняння розглянемо графік зустрічальності слова *corpus* з мовного сервісу Usage à travers le temps в рамках проекту онлайн-словника Le Dictionnaire Vivant de la Langue Française [<https://dvlf.uchicago.edu/mot/corpus>]:

Відображена в усіх графіках динаміка використання слова *corpus* в цілому схожа протягом більш ніж трьох сторіч, проте відсутність даних після 1950-х років у другому інструменті суттєво утруднює достовірне оцінювання й інтерпретацію. Варто також зауважити, що при побудові графіків у Google Books Ngram Viewer,

при нульовому згладжуванні, враховується відносна частота його появи за кожен рік, тоді як крива, отримана на основі аналізу матеріалів бази даних ARTFL-Frantext, показує частотність аналізованої лексичної одиниці на мільйон слововживань.

Таким чином, історико-етимологічний аналіз формування семантичної структури слова *corpus* у французькій мові, проведений на основі лексикографічних, корпусних даних і онлайн ресурсів, дає підстави говорити про розгортання семантичної структури слова: від релігійного поняття до мовного терміна. Інтерпретація аналізованої лексики у лінгвістичному ключі зустрічається у лексикографічних джерелах першої половини XX сторіччя, а розуміння корпусу як лінгвістичного терміна з'являється у шістдесятих роках XX ст. Становлення терміна *corpus* у французькому мовознавстві супроводжувалося пошуком його релевантних характеристик і спробами дати йому класифікацію. Виділення значимих характеристик корпусу у французькій мові реалізується атрибутивними компонентами, вираженими здебільшого прикметниками, дієприкметниками минулого часу різної структури; менш продуктивною є модель субстантивного поєднання з прийменником *de*. Вони конкретизують вихідне поняття, утворюють його видові кореляти, що є підставою для типологізації корпусів. Діахронічний аналіз динаміки частоти вживання слова *corpus* за допомогою мовних сервісів Google Books Ngram Viewer і Usage à travers le temps дає можливість констатувати схожі тенденції, попри те, що зазначені сервіси ґрунтуються на різних підходах і оперують різними масивами даних.

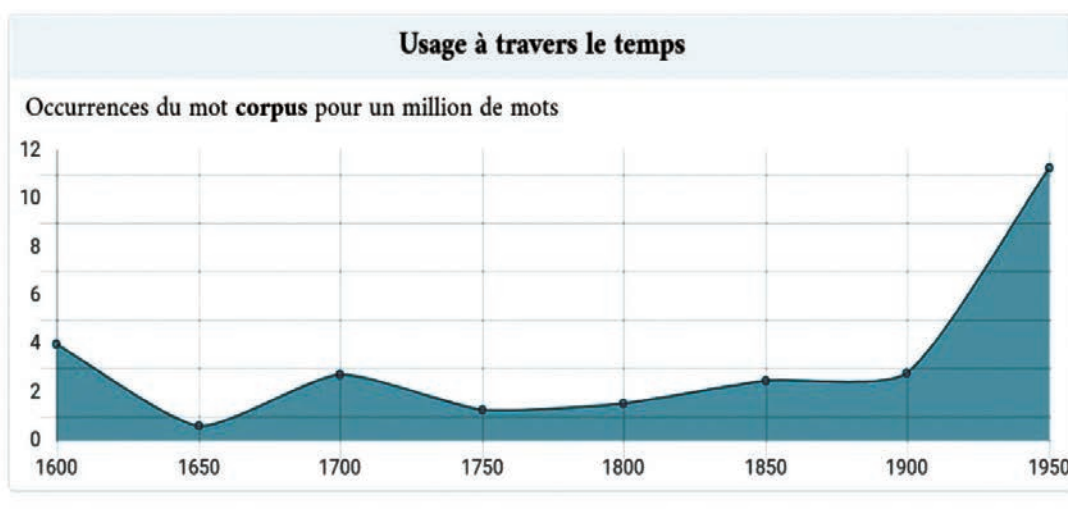


Рис. 3. Графік частоти згадування слова *corpus* з ресурсу «Usage à travers le temps»

## ЛІТЕРАТУРА

1. Демська-Кульчицька О. Дещо про класифікацію текстових корпусів. *Наукові записки. Серія: Мовознавство*. 2004. Т. 1. №. 11. С. 153-157.
2. Жуковська В.В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ імені Івана Франка, 2013. 142 с. 3. Захаров В.П. Корпусная лингвистика: Учебно-метод. пособие. СПб., 2005. 48 с.
4. Плахотнікова О.Ю. Корпус українського усного мовлення: теоретичні засади побудови й основи практичного втілення: дис. ... канд. філол. наук : 10.02.01; Київ. нац. ун-т ім. Тараса Шевченка. Київ, 2017. 280 с.
5. Плунгян В.А. Корпус как инструмент и как идеология: О некоторых уроках современной корпусной лингвистики. *Русский язык в научном освещении*. 2008. № 2 (16). С. 7-20.
6. Рыков В. В. Корпус текстов как реализация объектно-ориентированной парадигмы. Компьютерная лингвистика и интеллектуальные технологии: Тр. междунар. семинара "Диалог – 2002": В 2 т. Т. 1. Москва, 2002. С. 390-393.
7. Baude O., Blanche-Benveniste C., Calas M.-F., Carpeau P., Cordereix P., et al. (2006). *Corpus oraux, guide des bonnes pratiques*. CNRS Editions, Presses Universitaires Orléans.
8. Bergounioux, G. (2016). La linguistique de corpus et la partition des structuralismes. *Histoire Epistémologie Langage*, 38(2), 41-54.
9. Dalbera, J. P. (2002). Le corpus entre données, analyse et théorie. *Corpus*, (1).
10. Eshkol-Taravella, I., & Lefevre-Halftermeyer, A. (2017). Linguistique de corpus: vues sur la constitution, l'analyse et l'outillage. *Corela. Cognition, représentation, langage, (HS-21)*.
11. Charaudeau, Patrick. "Dis-moi quel est ton corpus, je te dirai quelle est ta problématique." *Corpus 8* (2009): 37-66.
12. Herschberg Pierrot, A. (2006). Style, corpus et genèse. *Corpus*, (5).
13. Mellet, S. (2002). Corpus et recherches linguistiques. Introduction. *Corpus*, (1).
14. Pruvost, J. (2017). Du Corpus Christi au Corpus linguistique en passant par le Corpus juris civilis et quelques autres. *Ela. Etudes de linguistique appliquée*, (4), 389-395.
15. Petit Larousse illustré. 1990, Librairie Larousse. 1680 p.
16. «FÊTE-DIEU ou FÊTE DU SAINT-SACREMENT», Encyclopædia Universalis [en ligne]. URL: <https://www.universalis.fr/encyclopedie/fete-dieu-fete-du-saint-sacrement/> (consulté le 7 septembre 2021)
17. Dictionnaire de l'Académie française (1932-1935) [en ligne]. URL: <https://www.cnrtl.fr/definition/academie8/corpus> (consulté le 12 juin 2021)
18. Dictionnaire de l'Académie française 8-e édition [en ligne] URL: <https://www.cnrtl.fr/definition/academie8/corpus>
19. Dictionnaire de l'Académie française 9-e édition [en ligne] URL: <https://www.academie-francaise.fr/le-dictionnaire/la-9e-edition>
20. Dictionnaire de la langue française (Litttré). Tome 1. Tome 2 [en ligne]. URL: <https://artflsrv03.uchicago.edu/philologic4/publicdicos/query?report=bibliography&head=corpus>
21. Definition corpus URL: <https://www.cnrtl.fr/definition/corpus>
22. Encyclopædia Universalis [en ligne]. URL : <https://www.universalis.fr/encyclopedie/habeas-corpus-act/>
23. Grand Dictionnaire universel du XIXe siècle [en ligne] URL: <https://gallica.bnf.fr/ark:/12148/bpt6k205356p.image.f948.langEN>
24. Larousse. [en ligne]. URL: <https://www.larousse.fr/dictionnaires/francais-monolingue/>
25. Le dictionnaire de la langue française [en ligne] URL: <https://www.lalanguefrancaise.com/dictionnaire/definition/corpus#3>
26. Le dictionnaire vivant de la langue française (DVLF) [en ligne] URL: <https://dvlf.uchicago.edu/>
27. Le grand dictionnaire terminologique. URL: <http://gdt.oqlf.gouv.qc.ca/index.aspx>
28. Le Robert[en ligne] URL: <https://dictionnaire.lerobert.com/definition>
29. TLFi : Trésor de la langue Française informatisé, URL: <http://www.atilf.fr/tlfi>, ATILF – CNRS & Université de Lorraine.
30. Google books Ngram viewer. URL: <https://books.google.com/ngrams> (reached 10.09.2021).

## REFERENCES

1. Dems'ka-Kul'chyts'ka, O. (2004). *Deshcho pro klasyfikatsiyu tekstovykh korpusiv*. [Something about the classification of text corpora]. *Naukovi zapysky. Seriya: Movoznavstvo*. Vol. 1. №. 11. P. 153-157. [in Ukrainian].
2. Zhukovs'ka, V.V. *Vstup do korpusnoyi linhvistyky: navchal'nyy posibnyk*. [Introduction to corpus linguistics: textbook]. Zhytomyr: Vyd-vo ZHDU imeni Ivana Franka, 2013. 142 s. [in Ukrainian].
3. Zakharov, V.P. (2005). *Korpusna linhvistyka: Navchal'no-metod. posibnyk*. [Corpus linguistics: Educational and methodical manual]. SPb. 48 s. [in Russian].
4. Plakhotnikova, O.Yu. (2017). *Korpus ukrayins'koho usnoho movlennya: teoretychni zasady pobudovy y osnovy praktychnoho vtilennya*. [Corpus of Ukrainian oral speech: theoretical principles of construction and basics of practical implementation]: dys. ... kand. filol. nauk : 10.02.01; Kyiv. nats. un-t im. Tarasa Shevchenka. [in Ukrainian].
5. Plunhyan, V.A. (2008). *Korpus yak instrument ta yak ideolohiya: Pro deyaki uroky suchasnoyi korpusnoyi linhvistyky*. [Corpus as a tool and as an ideology: About some lessons of modern corpus linguistics]. *Russkyy yazyk v nauchnom osveshcheniyi*. № 2 (16). S. 7-20. [in Russian].



6. Rykov, V. V. (2002). *Korpus tekstiv kak realizatsiya ob'yektno-oriyentovanoi paradyhmy*. [Corpus of texts as an implementation of object-oriented paradigm]. Komp'yuterna linhvistyka ta intelektual'ni tekhnolohiyi: Tr. mizhnar. seminaru "Dialoh – 2002": V 2 t. T. 1. Moskva, 2002. S. 390-393. [in Russian].
7. Baude, O. et al. (2006). *Corpus oraux, guide des bonnes pratiques*. CNRS Editions, Presses Universitaires Orléans. [in French].
8. Bergounioux, G. (2016). La linguistique de corpus et la partition des structuralismes. *Histoire Epistémologie Langage*, 38(2), 41-54. [in French].
9. Dalbera, J. P. (2002). Le corpus entre données, analyse et théorie. *Corpus*, (1). [in French].
10. Eshkol-Taravella, I., & Lefevre-Halftermeyer, A. (2017). Linguistique de corpus: vues sur la constitution, l'analyse et l'outillage. Corela. *Cognition, représentation, langage*, (HS-21). [in French].
11. Charaudeau, Patrick. "Dis-moi quel est ton corpus, je te dirai quelle est ta problématique." *Corpus* 8 (2009): 37-66. [in French].
12. Herschberg Pierrot, A. (2006). Style, corpus et genèse. *Corpus*, (5). [in French].
13. Mellet, S. (2002). Corpus et recherches linguistiques. Introduction. *Corpus*, (1). [in French].
14. Pruvost, J. (2017). Du Corpus Christi au Corpus linguistique en passant par le Corpus juris civilis et quelques autres. *Ela. Etudes de linguistique appliquée*, (4), 389-395. [in French].
15. Petit Larousse illustré. 1990, Librairie Larousse. 1680 p. [in French].
16. « FÊTE-DIEU ou FÊTE DU SAINT-SACREMENT », Encyclopædia Universalis [en ligne]. URL: <https://www.universalis.fr/encyclopedie/fete-dieu-fete-du-saint-sacrement/> (consulté le 7 septembre 2021) [in French].
17. Dictionnaire de l'Académie française (1932-1935) [en ligne]. URL: <https://www.cnrtl.fr/definition/academie8/corpus> (consulté le 12 juin 2021) [in French].
18. Dictionnaire de l'Académie française 8-e édition [en ligne] URL: <https://www.cnrtl.fr/definition/academie8/corpus> [in French].
19. Dictionnaire de l'Académie française 9-e édition [en ligne] URL: <https://www.academie-francaise.fr/le-dictionnaire/la-9e-edition> [in French].
20. Dictionnaire de la langue française (Littré). Tome 1 [en ligne]. URL: <https://artflsrv03.uchicago.edu/philologic4/publicdicos/query?report=bibliography&head=corpus> [in French].
21. Dictionnaire de la langue française (Littré). Tome 2 [en ligne]. URL: <https://artflsrv03.uchicago.edu/philologic4/publicdicos/query?report=bibliography&head=corpus> [in French].
22. Encyclopædia Universalis [en ligne]. URL : <https://www.universalis.fr/encyclopedie/habeas-corpus-act/> [in French].
23. Grand Dictionnaire universel du XIXe siècle [en ligne] URL: <https://gallica.bnf.fr/ark:/12148/bpt6k205356p.image.f948.langEN> [in French].
24. Larousse. [en ligne]. URL: <https://www.larousse.fr/dictionnaires/francais-monolingue/> [in French].
25. Le dictionnaire de la langue française [en ligne] URL: <https://www.lalanguefrancaise.com/dictionnaire/definition/corpus#3> [in French].
26. Le dictionnaire vivant de la langue française (DVLF) [en ligne] URL: <https://dvlf.uchicago.edu/> [in French].
27. Le grand dictionnaire terminologique. URL: <http://gdt.oqlf.gouv.qc.ca/index.aspx>
28. Le Robert [en ligne] URL: <https://dictionnaire.lerobert.com/definition> [in French].
29. TLFi : Trésor de la langue Française informatisé, URL: <http://www.atilf.fr/tlfi>, ATILF – CNRS & Université de Lorraine. [in French].
30. Google books Ngram viewer. URL: <https://books.google.com/ngrams> (reached 10.09.2021). [in French].

## FORMATION OF THE SEMANTIC STRUCTURE OF THE WORD CORPUS IN THE FRENCH TRADITION

**Strashko Iryna Volodymyrivna**

*Candidate of Philosophical Sciences,*

*doctoral candidate at the Department of Applied Linguistics, Comparative Linguistics and Translation*

*National Pedagogical Dragomanov University*

*8/14, Turgenevska str., Kyiv, Ukraine*

*The paper presents the genesis and the semantic development of the word corpus in French. The use of comparative, historical, descriptive, contextual and definitional methods allowed to establish semantic differences in different periods of its use, to explicate its semantic structure, to clarify denotative, significative and conotative meanings, to reveal the appearance of new sememes, to observe its interpretation in French dictionaries, including those available on the lexical portal of the*

*CNRTL platform. Historical and etymological analysis allows to speak about the development of the semantic structure of a word from a religious concept to a linguistic term. Its interpretation in a linguistic way is found in lexicographic sources of the first half of the twentieth century, and since the sixties of the last century it has acquired a new meaning associated with corpus linguistics. The formation of the term corpus in French linguistics was accompanied by a search for its relevant characteristics and the attempts to classify it. The expansion of corpora typology led to the emergence of the appropriate nominations. They are formed in a syntactic way and characterized by a different structure. The significant characteristics of the corpus are expressed by attributive components, which concretize the original concept and form its correlates. The basic model of a term formation is an attributive component with adjective postposition; the quantitatively dominant model is a two-component term of the non-propositional structure; model N + PART is productive; less frequent is a noun combination with an included preposition. The change in the frequency of this lexical unit use is shown using the language services Google Books Ngram Viewer and Usage à travers le temps. Despite the fact that these services are based on different approaches and operate on different data sets, the analysis results allow us to state similar trends.*

**Key words:** French language, corpus, etymology, semantics, corpus linguistics.