

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ  
ІНСТИТУТ СХОДОЗНАВСТВА ім. А. Ю. КРИМСЬКОГО

**КРИКОНЮК Катерина Миколаївна**



УДК 811.222.1'06'116'366'373.611

**ФОРМАЛЬНА МОРФОЛОГІЯ  
ПИТОМИХ ЗАСОБІВ СЛОВОТВОРУ  
СУЧАСНОЇ ПЕРСЬКОЇ МОВИ**

10.02.13 – мови народів Азії, Африки,  
аборигенних народів Америки та Австралії

**АВТОРЕФЕРАТ**  
дисертації на здобуття наукового ступеня  
кандидата філологічних наук

Київ – 2014

Дисертацією є рукопис.

Роботу виконано в Інституті сходознавства імені А.Ю. Кримського НАН України

**Науковий керівник:**

доктор філологічних наук, професор  
**Тищенко Костянтин Миколайович,**  
Київський національний університет  
імені Тараса Шевченка,  
професор кафедри Близького Сходу

**Офіційні опоненти:**

доктор філологічних наук, доцент  
**Кшановський Олег Чеславович,**  
Київський національний  
лінгвістичний університет,  
професор кафедри східної філології

кандидат філологічних наук  
**Максимів Ольга Йосипівна,**  
Львівський національний університет  
імені Івана Франка,  
доцент кафедри сходознавства  
імені професора Ярослава Дашкевича

Захист відбудеться 27 листопада 2014 р. о 10 годині на засіданні спеціалізованої вченої ради Д 26.174.02 в Інституті сходознавства ім. А. Ю. Кримського НАН України за адресою: м. Київ – 01001, вул. Грушевського, 4.

З дисертацією можна ознайомитись у бібліотеці Інституту сходознавства ім. А. Ю. Кримського НАН України за адресою: м. Київ – 01001, вул. Грушевського, 4, к. 207.

Автореферат розісланий 24 жовтня 2014 р.

Вчений секретар  
спеціалізованої вченої ради



В.А. Мусійчук

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Сучасна словотвірна система перської мови є наслідком складного історичного розвитку перської лексики. Найбільш усталений тип словотвору питомих перських слів (суфіксально-префіксальний) властивий усім індоєвропейським мовам, – але в сучасній перській мові суфіксальний спосіб переважає. Крім успадкованої питомої лексики іранського походження словник містить пласти арабських, тюркських та французьких запозичень. Різною мірою вони також вплинули на словотвірну систему сучасної перської мови. Співвідношення питомих і запозичених засобів перського словотвору звичайно оцінюється інтуїтивно і ближче не досліджене.

Вивчення словотвору є необхідним кроком при відтворенні цілісного уявлення про структуру мови. Складність словотвірної системи зумовлена її нерозривними зв'язками з іншими рівнями мови, рухливістю системи і неможливістю проведення чіткої грані між її потенцією та фактичною реалізацією, між словотвором та словозміною, а також різноманітністю та комунікативною значимістю самих функцій словотвору

Свого часу проблемами перського словотвору займалися Є. Е. Бертельс, Ю. А. Рубінчик, Л. С. Пейсіков, І. К. Овчиннікова, А. К. S. Lambton. Серед іранських науковців цієї теми торкалися Х. Резаї Багбілі, Х. Фаршідфард та П. Н. Ханлері. Окремі аспекти морфології перської мови відображені у дослідженнях з комп'ютерної лінгвістики К. Мігердуміян, М. Арабсорхі та М. Шамсфард, В. Sagot та G. Walter. Незважаючи на багату традицію граматичних описів перської мови, її словотвірна система як реальне багатоаспектне і розгалужене ціле ще не була предметом спеціального вивчення.

**Актуальність** теми дослідження зумовлена відсутністю системного викладу словотвірних моделей і правил сучасної перської мови. Аналіз наявної літератури з питань перського словотвору переконує, що ця тема потребує подальшої розробки та удосконалення. Запропоновані принципи та критерії формально-морфологічного підходу виглядають достатніми, аби ефективно дослідити словотвірну систему перської мови і в кінцевому результаті адекватно відтворити картину словотвірної структури перського лексикону.

**Зв'язок роботи з науковими планами, темами.** Ця тема дотична до загальної тематики «Релігійно-доктринальні тексти Сходу: семантика і прагматика, рецепція та інтерпретація» (№ державної реєстрації 0112U001950).

**Мета** дисертаційного дослідження полягає у створенні опису словотвірної системи сучасної перської мови. Для досягнення поставленої мети необхідно виконати такі завдання:

- описати принципи та методологічні засади формально-морфологічного підходу до вивчення перського словотвору;
- створити інвентарій словотвірних формул сучасної перської мови;
- з'ясувати кількісне співвідношення морфологічних класів перської мови та прямих і опосередкованих запозичень з інших мов;
- виявити продуктивні моделі словотвірної системи та найвживаніші морфеми;
- дати вичерпну і водночас оглядну кількісну характеристику морфологічних класів;
- виявити основні закономірності структури питомих засобів перського словотвору;
- дослідити види дистрибуцій, які мають місце у системі перського словотвору;
- окреслити основне коло загальнолінгвістичних проблем структурної лінгвістики, які постають на ґрунті матеріалу, отриманого шляхом здійснення морфологічного аналізу.

**Об'єкт дослідження** – словотвірна система сучасної перської лексики, відображеної у тлумачному словнику М. Каріба. (Проаналізовано понад 29 тис. слів. Аргументація вибору наведена нижче.)

**Предмет дослідження** – інвентарій, структура та регламент питомих засобів словотвору сучасної перської мови; причому інвентарій трактовано як вичерпний перелік або реєстр елементарних словотвірних моделей, структура як внутрішні морфологічні зв'язки компонентів моделі, а регламент як зовнішні зв'язки самих моделей (інакше – інтегранти) у складі віртуальної системи перського словотвору.

Дисертація містить дві теоретичні частини: у першій досліджено морфемічний аспект морфології (тобто, який стосується морфеміки), а саме – формально-морфологічні засоби до вивчення словотвору (з акцентом на морфологічному аналізі) та методологічні засади формально-морфологічного підходу; другий теоретичний блок є наслідком осмислення локалізації побудованого словотвірного опису сучасної перської мови в рамках структурної лінгвістики.

**Фактологічна база роботи.** Загальновідомо, що кількість і якість словників загальномовної лексики є своєрідним відображенням історичного стану суспільства. Внаслідок стрімкого розвитку Ірану у ХХ ст. і як результат – потужного напливу неологізмів й стихійного терміноутворення, для перської мови постала нагальна потреба впорядкування лексичого фонду.

Першу серйозну спробу укладання тлумачного словника перської мови (видання розпочате 1931 року) зробив іранський вчений Алі-Акбар Дегхода. Робота над цим словником триває дотепер під егідою Інституту Дегходи, підрозділу Тегеранського національного університету, і останнє видання склало 16 томів (більше 27000 сторінок).

Іншими авторитетним лексикографічними джерелами середини ХХ ст. вважають словник Мугаммада Моїна в шести томах (1963-1973рр.), тлумачний словник Гасана Аміда в двох томах (перше видання – 1973р.) та однотомний словник Мугаммада Каріба (1977).

При незаперечно високому рівні цих словників, їхньою культурно-історичною особливістю є те, що у виборі першоджерел інформації вони надають перевагу класичній поезії та літературі, наслідком цього є численні застарілі слова та діалектизми у реєстрі водночас із браком слів, які відображають сучасні їм реалії життя іранців.

Серед сучасних тлумачних словників на особливу увагу заслуговують однотомний Словник перської мови Магшід Мошїрі (2004; побудований за принципом *dictionnaire alphabétique et analogique* «алфавітного та аналітичного», або, інакше, «асоціативного» словника французького лексикографа Поля Робера), Словник сучасної перської мови Кулямгусейна Садрі Афшара, Насрін та Настаран Гакамі (два видання – 1995-1996 роки; з акцентом на лексику, вживану в сучасних медіаресурсах), а також восьми томний Великий словник «Сохан» Гасана Анварі (2004).

Головним критерієм у виборі тлумачного словника Каріба як основного джерела дослідження став помірний лексичний обсяг його реєстру – 29162 слова.

Крім того, як зазначено у передмові словника, до реєстру також заведена лексика підручників і навчальних посібників для гімназій і університетів. Словник Каріба дотепер відображує норму сучасної літературної перської мови.

Для ролі фактологічної бази дисертаційної роботи масив реєстрових слів тлумачний словник «*واژه نامه نوین*» Мугаммада Каріба був придатний ще й тому, що відображує загальноповивану лексику сучасної перської мови без спеціальних слів, діалектизмів та історизмів. У разі потреби з метою точнішого з'ясування походження та функціональної ролі тієї чи іншої морфемі як допоміжне джерело залучено тлумачний словник Дегходи.

**Методи дослідження.** Дослідження спирається на: 1) метод морфологічного аналізу (для встановлення інформації про лексему); 2) метод формалізації, що дав змогу у формальному вигляді зафіксувати сутнісні властивості предмета дослідження; 3) описовий метод (для характеристики отриманих результатів), 3) матричний метод (для

впорядкування результатів дослідження та діаграмного унаочнення матеріалу); 4) метод перебору, метою якого є максимально оптимізувати структуру матриць і таблиць.

**Наукова новизна** дослідження полягає в тому, що у ньому вперше застосовано процедури формально-морфологічного аналізу до словотвірної системи сучасної перської мови як цілого.

**Практичне значення.** Одержані результати можуть знайти дальше застосування у мовно-типологічних студіях, а також для вдосконалення змісту навчання перської мови з метою полегшити засвоєння словотвору перської лексики.

**Теоретичне значення.** У теоретичному плані здійснене дослідження може бути корисним для підтвердження або спростування певних морфологічних механізмів чи словотвірних правил у загальнолінгвістичних моделях мови.

**Апробація.** Основні положення дисертації апробовані автором особисто на кількох науково-практичних конференціях у Львові та Києві. Назви і дати конференцій: 1) Міжнародна науково-практична конференція «Орієнталістика в Україні: традиції та актуальні проблеми», присвячена 75-літтю від дня народження Я. Є. Полотнюка в рамках святкування 350-річчя заснування Львівського національного університету імені Івана Франка (Львів, 14-15 жовтня 2010 року); 2) Сходознавчі читання А. Ю. Кримського, присвячені 90-літтю Омеляна Прицака (Київ, 2011 року); 3) Міжнародна наукова конференція «Топонімія і контекстуальний метод К. М. Тищенка» (Київ, 20 квітня 2012 року); 4) «Міжнародна наукова конференція «Мова і культура» імені Сергія Бураго (Київ, 23-26 червня 2014 року).

**Публікації.** Основні результати дослідження викладені у 6 публікаціях, з них у фахових виданнях, затверджених МОН України, – 4, та 2 – закордонні публікації.

**Особистий внесок здобувача** полягає в застосуванні методу морфологічного аналізу до цілого масиву лексем словника, у розчленуванні його засобами комп'ютерного програмування (написання макросів для класифікації матеріалу), у створенні оптимізованих таблиць перського словотвору, у визначенні місця здійсненого дослідження у сучасній структурній лінгвістиці, у доборі літератури, а також у виявленні морфологічних закономірностей, що діють у словотвірній системі перської мови. Робота є результатом самостійної праці дисертанта. Наукових праць, написаних у співавторстві, немає.

**Структура роботи.** Дисертаційне дослідження складається зі вступу, чотирьох розділів, висновків, списку використаних джерел і шести додатків. Обсяг основного тексту роботи становить 179 сторінок і включає 8 таблиць та 42 матриці. Список використаних джерел містить 156 найменувань (з них 50 – іноземними мовами). Робота має 6 додатків на 24 сторінках.

## **ОСНОВНИЙ ЗМІСТ РОБОТИ**

У **вступі** мотивовано вибір теми дослідження та її актуальність, розкрито стан її вивчення, перелічено теоретичні та методологічні засади і наукову новизну, визначено мету, об'єкт та основні завдання дисертаційної роботи.

У **першому розділі «Короткий огляд історії вивчення морфології та словотвору»** наведено огляд вивчення морфології у лінгвістиці. Упродовж усієї історії мовознавства об'єктом дослідження вчених були морфеми та морфологічні ланцюжки, лексичні, семантичні та граматичні категорії, морфологічні та морфонологічні правила, сублексичні фонологічні компоненти і взаємозв'язки між ними, при чому увага лінгвістів перемістилася з площини граматичних категорій до вивчення формальних елементів. Сьогодні більшість лінгвістів сходяться на думці, що об'єктом морфології є мінімальні двосторонні одиниці мови (плану вираження і плану змісту) – морфеми та комплекси цих одиниць, яким притаманна низка характерних особливостей (такі

комплекси називають словоформами або просто словами). Таким чином, можна сказати, що вся морфологічна проблематика вміщується у просторі між морфемою та словоформою.

Кожна лінгвістична школа зробила свій вагомий внесок у вивчення морфології. Тенденція до класичного перегляду лінгвістичних понять і методів дослідження мови виникла у межах класичного мовознавства; у ході цього перегляду вироблялися елементи структурного підходу до мови. Засади структурної лінгвістики були закладені Ф. Ф. Фортунатовим, І. А. Будуеном де Куртене та Ф. де Сосюром. Історично структурну лінгвістику розвивали три лінгвістичні школи: *празька* на чолі з Н. С. Трубецьким та Р. О. Якобсоном (*функціональна лінгвістика*), *копенгагенська* або *данська* на чолі з Л. Єльмслевим (*гloseматика*) і *американська* на чолі з Л. Блумфілдом та З. Херрісом (*дескриптивна* або *дистрибутивна лінгвістика*).

Американська дистрибутивна лінгвістика не є теорією мови у загальноприйнятому значенні цього слова. Її можна розглядати як схему процесів, що уможливають розкриття граматики певної мови, або як експериментальну техніку збирання та первинної обробки сирого мовного матеріалу. Розуміння морфології американськими структуралістами зводиться до таких основних принципів (за С. Андерсоном): 1) морфема є однорідними неподільними мовними одиницями плану вираження, що з'єднують деякі компоненти плану змісту з групою взаємовиключних аломорфів, які їх виражають; 2) кожна морфема має певний зміст, а кожен аломорф – певне фонологічне вираження; 3) морфема є найменшими значенневими одиницями, з яких утворюються слова; 4) кожна морфема у слові представлена лише одним аломорфом, і, навпаки, кожен аломорф представляє лише одну морфему. Американські дескриптивісти також докладно опрацювали методику *дистрибутивного аналізу*, суть якого полягає в отриманні компактного опису структурних властивостей мови в плані вираження на основі розподілу окремих одиниць (у тому числі і морфем) у тексті.

На відміну від дистрибутивної лінгвістики гloseматика була задумана не як експериментальна техніка обробки текстів, а як *універсальна лінгвістична теорія*. Вчені копенгагенської школи вважали, що за структурного підходу до мови знімається відмінність між усіма одиницями мови – реченнями, словами, морфемами, складами, фонемами – усі вони характеризуються одотипними функціями і можуть бути розглянуті в одному процесі аналізу. Цим самим розвивається відмінність між традиційними розділами мовознавства – лексикологією і граматиною, граматиною і фонетикою і т.д. Структурний підхід ніби дозволяє уникнути «старого недосконалого поділу лінгвістики на фонетику, морфологію, синтаксис, лексикографію і семантику». Натомість у гloseматиці вводиться членування мови на чотири пласти або шари (*strata*), виходячи з двох дихотомій: *зміст – вираження*, *субстанція – форма*. Через загальнолінгвістичну природу гloseматики у ній приділено небагато уваги питанням морфології та словотвору. Єльмслев розглядає морфему як флективний елемент плану змісту.

Празька функціональна лінгвістика – це теорія реальних явищ і процесів природної мови. Основною одиницею морфологічної системи мови вважається морфема; будь-яку морфему можна представити як пасмо елементарних морфологічних опозицій, наприклад опозицій дієслівних форм часу і т.п. «Репертуар» семантичних ознак, які слугують основою для цих опозицій, виділяється способом, цілком аналогічним тому, за допомогою якого виділяються фонологічні ознаки у фонології. Морфологічні опозиції розцінюються як бінарні та асиметричні.

З появою комп'ютерних технологій почався розвиток теорії «породжувальної граматики» Ноама Хомського. Генеративний напрямок представлений широким спектром дедуктивних теорій – від радикальних лексикалістських, деякі з яких взагалі відкидають поняття «морфема» чи дихотомію словотвір/ словозміна (Р. Лібер,

С. Лапойнт, А. М. ді Шулло та Е. Вільямз, А. Марантца, Д. Песецький, Р. Спроат, М. Аронов), до поміркованіших, котрі намагаються пояснити морфологічні правила фонологічними (неконкатенативна модель «автосегментної фонології» Дж. МакКарті та А. Прінса (1986)), морфонологічними (теорія «лексичної фонології» П. Кіпарського (1982), М. Галле і Р. Моганана (1985), Г. Буйджі й Р. Б. Рубака (1987), Ш. Інкеласа (1993)) та синтаксичними («теорії управління та зв'язування» М. Бейкера (1988), концепція «наносинтаксису» М. Старке (2010)) механізмами певної універсальної моделі мови.

Ефективну методику вивчення морфологічного аспекту мови розробили лінгвісти школи формальної морфології. Формальна морфологія виникла у рамках структурної лінгвістики у 1960-их рр. Одним з її методів є формально-морфологічний аналіз, свого часу опрацьований І. Братчиковим, С. Фітіаловим і Г. Цейтніним для потреб машинного перекладу у площині словозміни. За запропонованою ними методикою формалізовані описи морфології словозміни різних мов здійснив у 1963-2003 рр. К. М. Тищенко. Це формальні морфології дієслівних підсистем французької, італійської, португальської, іспанської, перської, фінської мов; опис акцентуаційних моделей іменного словотвору російської мови, а також системи словозмінних моделей частотної лексики української мови. З окремих питань формальної морфології під науковим керівництвом К. М. Тищенка написано низку дипломних робіт (студенти В. Турчанінова, О. Тимошук, О.Бедов, А. Півторак та ін.).

Ще одним сучасним підходом до вивчення морфології є ономасіологічний. Ця лінгвістична концепція вперше запропонована М. Докулілою і детально розроблена у працях Є. С. Кубрякової, В. Н. Телії, П. Штекауера, Л. Бауера та інших дослідників. Ономасіологія розглядає словотвір у світлі теорії номінації і використовує для його інтерпретації поняття *ономасіологічного базису*, *ономасіологічної ознаки*, *зв'язки* та *ономасіологічної категорії*. В основі ономасіологічної структури похідного слова лежить з'єднання ономасіологічного базису і ономасіологічної ознаки за допомогою ономасіологічної зв'язки. Ономасіологічний базис, який визначає категоріальну приналежність найменування, завжди характеризується простою структурою, у той час, як структура ономасіологічної ознаки, покладеної в основу номінації, може мати різний ступінь складності. Визначення морфеми як білатеральної мовної одиниці у межах ономасіології дає підстави робити висновок про ієрархічний характер структури мовних рівнів. Словотвір в ономасіологічному підході розцінюється як основний (непроміжний) мовний рівень. Словотвірні правила, які в контексті цього підходу мають термінологічну дефініцію *словотвірні типи*, вважаються систематичними і передбачуваними. Головними поняттями ономасіологічного аналізу є *семантичні категорії*, що відображають комплексний мовленнєво-мисленнєвий процес людини як соціальної істоти.

Огляд постулатів різних морфологічних шкіл привів нас до висновку, що завдання сучасної морфології повинно полягати у встановленні балансу між дослідженням змісту і форми морфеми, а також механізмів її взаємодії з фонологічною, морфонологічною, синтаксичною та семантичною площинами у світлі універсальніших морфологічних теорій.

У другому розділі «Теоретичні настанови формально-морфологічного підходу» детально розглянуто теоретичні засади формально-морфологічного підходу, за допомогою якого у роботі здійснений опис словотвірної системи сучасної перської мови. Ключовим методом формально-морфологічного підходу є морфологічний аналіз. Морфологічний аналіз передбачає встановлення інформації про словоформу шляхом чіткого виділення елементів лексеми та одночасну їх класифікацію. Можливості сполучуваності морфем завжди обмежені (у синхронічному зрізі); через те, можна

очікувати, що процедура морфологічного аналізу складатиметься з кінцевої кількості кроків. Крім того, формальний аспект підходу дає змогу стисло, у скороченому вигляді зафіксувати і передати відношення між слотовірними формантами, оцінити їх характер та виявити логічну структуру цілої системи.

При побудові морфологічного опису певної мови прийнято спиратися на низку вихідних методичних настанов: повноту, простоту, однорідність та формалізацію. *Повноту* опису досягають вичерпним залученням на вихідному етапі усіх без винятку мовних явищ досліджуваної галузі, зафіксованих певним авторитетним джерелом або групою джерел. *Простоту* опису забезпечують відбором економної метамови дослідження, прагненням до прозорості процедури аналізу, досяжної для повторення або перевірки, оглядної і несуперечливої. При цьому не слід забувати, що метамова опису не мусить бути складнішою від самого описуваного мовного об'єкта. *Формалізацію* дослідження забезпечують використанням лаконічної системи умовних позначень морфологічних категорій. Принцип *однорідності* спрямований на уникнення еkleктичного або механічного поєднання для потреб слотовірного аналізу часом взаємно несумірних засобів дослідження, опрацьованих на інших методологічних засадах. Принцип однорідності також стосується і вибору предмета дослідження: у центрі уваги лежать не індивідуальні явища, а класи явищ, частота застосування яких рівномірніша. Належність слотовірної формули до того чи іншого класу визначається не відносно індивідуальних формул, а відносно класів слотовірних формул.

Формально-морфологічний підхід передбачає опис процесу членування слів у формалізованому вигляді з метою виявлення загальної картини системи слотовірних моделей перської лексики. Слова кодувалися за такою формулою:

$$R'' = mR'm_1mR'_2\dots,$$

де R – радикал, або нерозкладне слово, що відноситься до певної частини мови і позначається її символом (N, Ad, Verb, Aj і т.д.); m – граматична або лексико-граматична морфема (суфікс або префікс), що передається без змін засобами латинської транскрипції. Символ «'» (два штрихи) означає, що слово складне, тобто містить радикал (кілька радикалів) та одну (або декілька) граматичну чи лексико-граматичну морфему.

Метаапарат формально-морфологічного підходу утворюють поняття метамови, інвентарію слотовірних формул, морфологічного класу, продуктивності та валентності, ініціалі, медіалі, фіналі, гапаксу, унікального та рекурентного корпусів системи словотвору. *Метамовою* ми вважали універсальний формалізований семіотичний код (або «мову другого рівня») морфологічного опису сучасної перської мови («мови першого рівня»), який уможлиблює подальшу класифікацію отриманого матеріалу, встановлення різного роду відношень між його складовими та здійснення розрахунків з метою виявлення лінгвостатичних закономірностей у словотворі. Сукупність усіх слотовірних формул словникового реєстру, отриманих методом морфологічного аналізу, складає *словотвірний інвентарій*. Робота з ним дала змогу виявити низку лінгвістичних закономірностей, що діють у перському словотворі. Словотвірний інвентарій є своєрідною моделлю кінцевих станів (морфологічна граматики словотвору), яка уможлиблює створення морфологічного опису першого рівня. Для позначення слотовірних моделей, які утворюють лише одну лексему, обрано термін *гапакс* (від грецьк. ἄπαξ λεγόμενον – «тільки один раз назване»). Гапакси не потребують морфологічного аналізу – їх можна розглядати як словникові (а не граматичні) одиниці. Група слотовірних одиниць-гапаксів складає *корпус унікальних моделей* інвентарію. До



корпусу *рекурентних моделей* інвентарію відійшли такі, що утворюють два похідні слова і більше.

З метою опрацювання інвентарію словотвірних формул було запроваджено терміни *ініціаль*, *медіаль* та *фіналь*. Ініціаллю називалася перша морфема зліва (у латинській транслітерації), фіналлю – остання права морфема словотвірної формули, а медіаллю – одна або кілька морфем між ініціаллю та фіналлю.

Під *продуктивністю* словотвірної формули ми розуміли її максимальну здатність утворювати лексеми. За іншу важливу рису словотвірної системи було обрано *характеристику валентності*, виражену її індексом – кількісним показником сполучуваності певної морфеми.

Частини мови позначалися терміном *морфологічний клас*, оскільки за основний критерій належності слова до тієї чи іншої групи у дослідженні правив морфологічний. Як допоміжний, до уваги брався також семантично-синтаксичний критерій.

**Третій розділ «Опис питомої словотвірної підсистеми сучасної перської мови»** становить ядро нашого дослідження. У ньому подано реальний опис словотвірної системи сучасної перської мови засобами формальної морфології. Найвагоміші висновки третього розділу подаються далі.

Передусім було з'ясовано кількісні співвідношення морфологічних класів перської мови та прямих і опосередкованих запозичень з інших мов. За даними виконаного дослідження морфологічні класи лексем розподіляються таким чином: 55.44% від загальної кількості слів припадає на іменники і 23.6% – на прикметники; дієслова, прислівники, іменники-прикметники та прикметники-прислівники становлять відповідно 12.91%, 3.72%, 2.26% та 1.3% від усього словникового реєстру. Окрім того, окремо було отримано квантитативні характеристики кореляцій морфологічних класів простих та складних слів. Для кращого розуміння одержаних кількісних показників ми використали дані співвідношень частин інших мов, зіставлення яких виявило потужну роль іменників у перському словотворі.

Загальна цифра усіх запозичень (прямих та опосередкованих) складає 52.4% словника. Найбільша частка лексичних запозичень припадає на арабську мову – 46.4%, запозичення з французької становить 2.8%, з турецької – 1.04%, з англійської – 0.2% від загальної кількості словника. Найбільше опосередкованих запозичень у перській мові – з грецької через арабську (0.23%).

До словотвірного інвентарію перської мови увійшло сім повнозначних морфологічних класів (іменник, прикметник, дієслово, прислівник, числівник, займенник, вербоїд) та шість службових класів (прийменник, артикль, сполучник, вигук, питальне слово, частка). Окремі морфологічні класи склали слова, що не піддаються атрибуції й за певних умов переходять з однієї частини мови в іншу; усього таких конверсійних класів нараховано сімнадцять (здійснений аналіз показав, що це здебільшого слова арабського походження). Морфологічний клас іменників утворений 407 словотвірними формулами, прикметників – 241 словотвірною формулою, прислівників – 71, дієслів – 15 словотвірними формулами. Усього словотвірних формул у конверсійних морфологічних класах – 80. У відсотках склад словотвірних одиниць інвентарію розподіляється таким чином: 50% інвентарію належать іменникам, 29.61% – прикметникам, 8.73% – прислівникам, 1.84% – дієсловам, 9.82% – конверсійним морфологічним класам.

У ході дослідження словотвірну систему питомих засобів перської мови було осмислено як динамічну аналітичну модель, яка має багатогранну структуру з чітко вираженими внутрішніми закономірностями та взаємозв'язками. При чому складові частини цієї моделі організуються на основі характеру ініціалей та фіналей – чи то їх утворюють афікси, чи морфологічні класи (поділ системи на чотири морфологічні

підсистеми), продуктивності чи непродуктивності словотвірних формул (поділ інвентарію на рекурентний та унікальний корпуси; розмежування трьох зон Бредфорда у рекурентному корпусі), а також належності словотвірної моделі до певної частини мови (розподіл словотвірних формул за морфологічними класами). Виявлені тенденції перського словотвору стосуються як внутрішньої будови словотвірних формул (закономірності ініціалей на медіалі), так і кількісно-рангового, морфемно-валентного та кількісно-морфемного розподілу словотвірних формул у морфологічних класах (залежності Есту-Зіпфа 1, 2, 3).

Впорядкування формул словотвірного інвентарію за допомогою матриць, а також на основі принципів їх продуктивності та кількості морфем у слові, дало змогу виявити низку лінгвістичних закономірностей, за якими відбувається організація морфологічних складових у лексемі. При чому остаточний варіант матриць став можливим завдяки їх оптимізації. Внаслідок оптимізації матриць іменникового словотвору було розрізнено 6 множин ( $\{a-\emptyset-C\}$ ,  $\{a-C-a\}$ ,  $\{C-\emptyset-a\}$ ,  $\{C-C-a\}$ ,  $\{C-a-a\}$  та  $\{C-aC/Ca/aaC-a\}$ , де  $a$  – афікс,  $C$  – морфологічний клас,  $\emptyset$  – нульова морфема), кожна з яких містить певну кількість структурних типів (вони можуть бути визначені як за критерієм ініціалі, так і фіналі).

Для прикладу наведемо одну матрицю продуктивних словотвірних моделей, яка містить закономірність утворення іменників, початковою морфемою яких є ініціаль-префікс (див. таблицю 1). Цю закономірність можна сформулювати так: *в разі, коли ініціаллю словотвірної формули виступає префікс, а фіналлю – морфологічний клас, то медіаль утворює нульова морфема (напр., dar- + N(درخانه, دربار), dar- + V(دریند)). Виняток з цього правила становить словотвірна формула N''=be+V+N – بیستانکار*. Досконала візуалізація матеріалу у цій матриці уможливило виділення 7 структурних типів у межах множини  $\{a-\emptyset-C\}$  словотвірних моделей перської мови:

**Тип 1. (-N, -W, -V, -Aj): ham-** : همآورد، همجوش، همجوار؛ همدرست؛ همراه، همسال، همکار، همسر، همدم، همچشم، هم ماهنگ، هم چشم، همزاد؛

**Тип 2. (-N, -W, -V): dar-**: دربار، درخانه، دروا؛ دریند؛ درگذشت، درآمد، درخواست

var-: ورتاج، ورسنگ؛ ورنانداز، ورسکست

bar-: برچسب، برمال، بریاره، برنامه؛ برانداز، برگذار؛ برآورد، برخورد، برداشت، برگشت

bāz- : بازدم، بازدار، بازرس؛ بازخواست، بازدید، بازیافت

**Тип 3. (-N, -W): bi-**: بی سیم، بیبده، بی همال، بی سکه؛ بیداد

foroo-: فرودست، فرومایه؛ فروداشت

**Тип 4. (-W, -V): vā-** : وایند، واریز؛ واخواست، وانمود

**Тип 5. (-N, -Pron): doš-** : دشنام، دشمن

У дослідженні шляхом метризації та оптимізації виявлено три закономірності формальної організації морфем у словотвірних формулах іменників.

Упорядкування формул словотвірного інвентарію за критеріями продуктивності та кількості морфем дало змогу виявити дві лінгвостатистичні закономірності Есту-Зіпфа, що діють у системі словотвору перської мови: 1) *якщо впорядкувати словотвірні формули від найпродуктивнішої до гапаксів (вісь f), то залежність між рангом тієї чи іншої словотвірної моделі (вісь r) та кількістю слів, яку вона утворює, виражатиме гіперболічна крива Есту-Зіпфа; іншими словами, чим менший ранг словотвірної формули, тим більше слів вона утворює (закономірність Есту-Зіпфа 1); 2) що менше морфем у словотвірній формулі, то більше таких формул (закономірність Есту-Зіпфа 2).*

## Матриця множини {a-Ø-C} продуктивних моделей складних іменників

Морфологічний клас	Finale →	-N	-W	-V	-Aj	-Pron
	Initiale ↓					
N	ham-	-Ø- ( <i>ham-sar</i> )	-Ø- ( <i>ham-āvard</i> )	-Ø- ( <i>ham-joosh</i> )	-Ø- ( <i>ham-javār</i> )	
	dar-	-Ø- ( <i>dar-bār</i> )	-Ø- ( <i>dar-āmad</i> )	-Ø- ( <i>dar-band</i> )		
	var-	-Ø- ( <i>var-sang</i> )	-Ø- ( <i>var-šekast</i> )	-Ø- ( <i>var-andāz</i> )		
	bar-	-Ø- ( <i>bar-nāmeḥ</i> )	-Ø- ( <i>bar-dāšt</i> )	-Ø- ( <i>bar-andāz</i> )		
	bāz-	-Ø- ( <i>bāz-dam</i> )	-Ø- ( <i>bāz-did</i> )	-Ø- ( <i>bāz-res</i> )		
	bi-	-Ø- ( <i>bi-sim</i> )	-Ø- ( <i>bi-dād</i> )			
	foroo-	-Ø- ( <i>foroo-dast</i> )	-Ø- ( <i>foroo-dāšt</i> )			
	vā-		-Ø- ( <i>vā-khāst</i> )	-Ø- ( <i>vā-riz</i> )		
	doš-	-Ø- ( <i>doš-nām</i> )				-Ø- ( <i>doš-man</i> )
	bād-	-Ø- ( <i>bād-zahre</i> )				
	be-	-V- ( <i>be-setān-kār</i> )				
	andar-	-Ø- ( <i>andar-gāh</i> )				
	khar-	-Ø- ( <i>khar-gooš</i> )				
	deš-	-Ø- ( <i>deš-pil</i> )				
	farā-	-Ø- ( <i>farā-bar</i> )				
	pā-	-Ø- ( <i>pā-zahr</i> )				
	pād-	-Ø- ( <i>pād-afrah</i> )				
na-		-Ø- ( <i>na-bood</i> )				

Систематизація фіналей словотвірних формул у матрицях за трьома зонами Бредфорда уможливила формулювання ще однієї словотвірної закономірності: *якщо словотвірні фінали, розташовані у спадному порядку валентності, поділити на три відносно рівні частини, то кількість слів, утворених за словотвірними моделями в межах зони, зменшуватиметься у кожній наступній зоні.*

У цьому розділі також було визначено кількісні характеристики корпусів унікальних та рекурентних формул словотвірного інвентарію, які є важливими показниками співвідношення словотвірних правил та винятків в описах лінгвістичних моделей. Дослідження показало, що сектор унікальних формул складає 56.7% від загальної кількості словотвірних одиниць інвентарію, тоді як на сектор рекурентних формул припадає 43.3%. За унікальними формулами утворено 8.325% складних слів реєстру, за рекурентними – 91.675%.

За допомогою спеціально створеного макроса також вдалося виявити найпродуктивніші моделі словотвірної системи сучасної перської мови:  $N''=N+N$ ,  $N''=N+V$ ,  $N''=N+e$ ,  $N''=N+i$ ,  $N''=Aj+i$ ,  $N''=Aj+N$ ,  $N''=Num+N$ ,  $N''=N+ak$ ,  $N''=N+gāh$ ,  $N''=V+eš$ ,  $Aj''=N+i$ ,  $Aj''=Aj+N$ ,  $Aj''=N+V$ ,  $Aj''=W+e$ ,  $Aj''=N+N$ ,  $Aj''=N+Aj$ ,  $Aj''=N+in$ ,  $Aj''=Aj+V$ ,  $Aj''=N+mand$ ,  $Aj''=V+ā$ ,  $Ad''=N+an$ ,  $Ad''=Aj+an$ ,  $Ad''=N+N$ ,  $Ad''=Ad+N$ ,  $Ad''=N+i$ ,  $Ad''=N+be+N$ ,  $Ad''=Pron+N$ ,  $Ad''=ham+Ad$ ,  $Ad''=Aj+Aj$ ,  $Ad''=Ad+ā+Ad$ ,  $Verb''=W+an$ . Ці моделі складаються з таких найуживаніших морфем:  $N$ ,  $V$ ,  $W$ ,  $Aj$ ,  $Num$ ,  $-e$ ,  $-i$ ,  $-ak$ ,  $-gāh$ ,  $-eš$ ,  $-ā$ ,  $-mand$ ,  $be-$ ,  $ham-$ .

У четвертому розділі «Локалізація результатів морфологічного аналізу у системі основних понять структурної лінгвістики» зроблена спроба з'ясувати, яким чином результати здійсненого морфологічного аналізу перського словника співвідносяться з формальними моделями структурної лінгвістики.

Було досліджено види дистрибуцій, які мають місце у системі перського словотвору. Як відомо, під дистрибуцією будь-якої морфеми розуміють суму всіх контекстів, в яких вона може зустрічатися, на відміну від тих контекстів, в яких вона зустрічатися не може. Розрізняють чотири основні типи дистрибуції: додаткову, вільно варіативну, контрастну і дистрибуцію часткової еквівалентності. У ході аналізу словотвірного інвентарію вдалося встановити додаткову та вільно варіативну дистрибуції. Додаткова дистрибуція, зумовлена оточенням голосна/приголосна попередньої морфеми, встановлена для двох груп ініціалей-префіксів та двадцяти груп фіналей-суфіксів (див. табл. 2), причому морфеми у цій дистрибуції пов'язані між собою привативними опозиційними відношеннями. Водночас усі морфеми протиставлені за ознакою довжини морфем (кожна наступна морфема на один звук довша, ніж попередня (на фонеми /y/, /e/, /i/, /t/, /g/, /qh/, /ā/)).

Таблиця 2.

### Опозиційні групи ініціалей-префіксів та фіналей-суфіксів у додатковій дистрибуції

pā-/pād-	-eš/-yeš	-āt/-yāt	-ande/-yande
dar-/andar-	-gān/-egān/-yegān	-j/-ej/-ij	-vand/-āvand
-e/ye/īye/-te	-gi/-egi	-ā/-yā	-če/-iče
-varanjan/-āvaranjan	-bān/-ebān	-gin/āgin	-qhand/-āqhand
-ān/-yān/-iyān	-mān/-emān/-yemān	-gen/-āgen	-āle/qhāle
	-gār/-egār	-ak/-yak/-gak	

До групи вільно варіативної дистрибуції входять три групи ініціалей-префіксів та двадцять груп фіналей-суфіксів (див. табл. 3).

Таблиця 3.

**Опозиційні групи ініціалей-префіксів  
та фіналей-суфіксів у вільно варіативній дистрибуції**

<b>deš-/dež-/doš-</b>	<b>-vār/-vāre</b>	<b>-an/-ane</b>	<b>-ji/-či</b>
<b>bād-/pād-/ped-</b>	<b>-ān/-āne</b>	<b>-and/ande</b>	<b>-egi/-yeki</b>
<b>be-/pe-</b>	<b>-yān/-yāne</b>	<b>-gān/-gāne</b>	<b>-iyyat/-iyye</b>
<b>-vān/-vāne</b>	<b>-āl/-āle/-al</b>	<b>-eč/-ej</b>	<b>-vā/-bā</b>
<b>-in/-ine</b>	<b>-baranjan/-varanjan</b>	<b>-eš/- ešt</b>	<b>-vaš/-faš</b>
<b>-gāh/-gah</b>	<b>-āvi/-ovi</b>	<b>-āgand/-āqhand</b>	

З дистрибуцією тісно пов'язане поняття потужності морфологічних груп, теоретичний аспект якого розглянуто у праці Ю. С. Степанова в контексті теоретико-множинного методу, розробленого на основі категорій алгебри та логіки. Потужність морфологічних груп представлена як лінгвістичний аналог ступеню (або квантифікації) логічних відношень дистрибуції – повноти, симетрії, рефлексивності та транзитивності.

Під терміном «потужності морфологічних груп» розуміють кількість одиниць, що поєднуються з даною одиницею, однак це кількість не в абсолютному підрахуванні, а лише у відносному, що визначається словами «дорівнює», «більше за ...», «менше за...», «значно більше за...», «значно менше за...» по відношенню до інших дистрибутивних груп.

Виявлення потужності морфологічних груп розкриває процеси граматикалізації, які мають місце у морфологічній площині мови. Граматикалізацією – у загальному визначенні – називають суму поступових процесів зміни одиниць різного рівня, що спричиняє втрату семантичної складності, прагматичної значущості, синтаксичної свободи і фонетичної субстанції. Граматикалізація відбувається не довільно, а тому, знаючи тільки синхронічний стан будь-якої граматичної системи, дослідник може висувати правдоподібні гіпотези як про походження тих чи інших елементів, так і про їхні можливі еволюції.

Огляд потужності основи теперішнього часу дієслова у перському словотворі (в позиції фіналі) дав змогу виявити особливості процесу граматикалізації цієї морфологічної групи. Ми проаналізували усі іменники, прикметники та іменники/прикметники словникового реєстру, другою морфемою яких виступає основа теперішнього часу дієслова.

У традиційній перській граматиці не існує однозначної оцінки морфологічної ролі кореня теперішнього часу дієслова. Морфемами, які його утворюють – такі, як-от *-dār*, *-keš*, *-āvar*, *-gir* і т.д. – вважаються напівсуфіксами (напівафіксами) або ж морфологічними частками. З'ясування потужності цієї морфологічної групи може допомогти пояснити його функцію у перському словотворі: група морфем з індексом вищим за 6-7 (тобто морфемами, до яких зліва приєднуються більше за шість- сім одиниць) доцільно вважати віддієслівними аналітичними афіксами. Група з індексом 4-6 виглядає як перехідна до аналітичних афіксів.

Проведений аналіз також показав, що валентність та продуктивність морфем (чи морфологічного відрізка) характеризують дві структурні риси словотвірної формули та її складових: внутрішню, яка є показником частотної сполучуваності морфем з

іншими словотвірними формантами всередині словотвірної формули, та зовнішню, що стосується її «породжувальної здатності», тобто участі у творенні нових лексем. Було виявлено валентність словотвірних морфем у різних морфологічних класах. Найчастотнішу сполучуваність у межах класу іменників демонструють суфікси *-e* (36), *-i* (31), *-ak* (16), *-ān* (15) та префікси *bar-* (8), *ham-* (6), *bā-* (5), *bi-* (5), *doš-* (5), *vā-* (5), *var-* (5), причому перші є ще і найпродуктивнішими у масштабі усієї системи словотвору сучасної перської мови, тобто за їх участю утворено найбільшу кількість лексем. Найвищу валентність у морфологічних класах прикметників, іменників-прикметників, прикметників-прислівників мають відповідно такі суфікси та префікси: *-e* (23), *-i* (20), *-ān* (7), *-oo* (7), *-in* (7), *-vār* (6) та *nā-* (18), *bi-* (6), *farā-* (5), *dar-* (4); *-e* (9), *-i* (2), *-ā* (2), *-gār* (2) та *ham-* (5), *nā-* (3), *dar-* (2); і врешті суфікси *-ān* (5), *-āne* (2), *-in* (2). До найвалентніших суфіксів та префіксів у морфологічному класі прислівників належать: *-i* (9), *-an* (5), *-ān* (5), *-e* (5) та *bā-* (3), *be-* (3), *ham-* (3).

Завдяки валентному аналізу було встановлено ще одну лінгвостатистичну закономірність (закономірність Есту-Зіпфа 3): *що більше валентностей має суфікс, тим менший його ранг* (див. рис. 1).

Отриманий внаслідок морфологічного аналізу матеріал став корисним у типологічних дослідженнях мов. Зокрема, реєстер словника містить інформацію про ступені суфіксації, деривації та префіксації – показники, що використовуються у встановленні морфологічного типу тієї чи іншої мови. Вони разом з іншими критеріями були свого часу виділені Дж. Грінбергом у модифікованій схемі, створеній на базі типологічної класифікації Сепіра (її основу склали ступені синтетичності, аглютинативності, компактності, деривативності, флективності, префіксації, суфіксації). Ця схема дає змогу використовувати кількісні показники під час порівняння текстів на різних мовах.

У дослідженні було з'ясовано, що ступінь префіксації усієї системи словотвору перської мови становить **0.001**. Значно вищим є показник ступеню суфіксації – **0.005**.

У процесі зіставлення ступенів суфіксації та деривації між морфологічними класами у дослідженому словниковому реєстрі була виявлена така особливість: суфікси та префікси беруть найактивнішу участь у творенні прикметників. Інакше кажучи, *у порівнянні з іншими морфологічними класами прикметники мають найвищі ступені суфіксації та префіксації*.

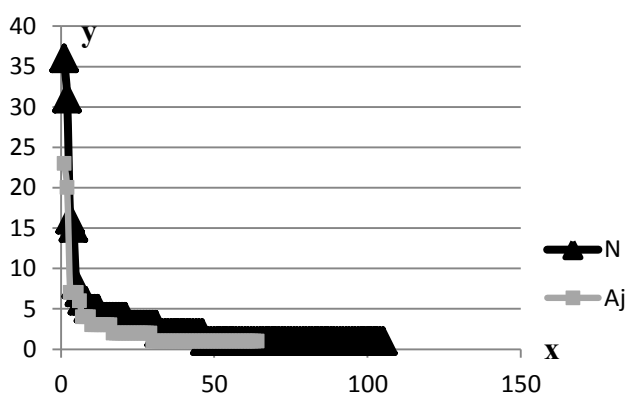


Рис. 1. Залежність Есту-Зіпфа (3).  
Частотно-ранговий розподіл суфіксів іменників та прикметників за спадною валентністю (x - індекс валентності суфіксу, y - ранг суфіксу)

Виконане дослідження уможливило такі **висновки**:

1. Послідовно застосований у роботі формально-морфологічний підхід виявився ефективним для створення логічно несуперечливого якісного і кількісного опису словотвірної системи мови у синхронічному зрізі.
2. Опрацьований широкий спектр проблематики з морфології на основі методологічних засад формально-морфологічного підходу (повнота, простота, однорідність, формалізація) свідчить про перспективність цієї концепції для розв'язання лінгвістичних питань у суміжних зі словотвором ділянках.
3. Серед дослідних завдань було передбачено укласти інвентарій словотвірних формул сучасної перської мови. Цей інвентарій створено, подано його квантитативні характеристики.
4. Засобами формальної морфології стало можливим виявлення основних закономірностей структури питомих засобів перського словотвору в будові ініціалей та медіалей, а також кількісно-рангової, морфемно-валентної та кількісно-морфемної залежностей словотвірних формул у морфологічних класах.
5. Отримані результати відповідають загальнолінгвістичним проблемам структурної лінгвістики: проблемам кібернетики та лінгвістичного моделювання (яким чином можливості людської мови та мислення співвідносяться з можливостями машин, що переробляють інформацію), проблемам морфологічної типології мов, виявлених на матеріалі парадигматичного типу зв'язків мовних одиниць, проблемам потужності морфологічних груп та проблемам лінгвістичної ймовірності (пов'язаної з валентністю морфем).

#### **Основні положення дисертації висвітлено у таких публікаціях:**

##### **Статті у наукових фахових виданнях:**

1. Криконюк К.М. Семантика, словотвір і етимологія оронімів Ірану / К.М. Криконюк // Вісник Львівського університету. – Львів, 2013. – Випуск 59. – С. 131–143. – (Серія філологічна).
2. Криконюк К.М. Формально-морфологічний підхід до вивчення перського словотвору / К.М. Криконюк // Вісник Львівського університету. – Львів, 2014. – Випуск 61. – С.39–44. – (Серія філологічна).
3. Криконюк К.М. Число словотвірних моделей і лексична потужність частин мови (до типології перської та слов'янських мов) / К.М. Криконюк // Компаративні дослідження слов'янських мов і літератур: зб. наукових праць. Пам'яті академіка Леоніда Булаховського. – К., 2014. – Випуск 24. – С.94–98.
4. Криконюк К.М. Сполучуваність словотвірних морфем сучасної перської мови у світлі лінгвостатистики / К.М. Криконюк // Мова і культура: науковий журнал. – К.: Видавничий дім Дмитра Бураго, 2014. – Випуск 17. – Т. I (169). – С.298–307.
5. Криконюк К.М. Лингвостатистические и морфологические закономерности в персидском словообразовании / К. М. Криконюк // Филологические науки. Вопросы теории и практики. – Тамбов: Изд-во «Грамота», 2014. – № 11 (41). – Часть 1. – С. 119–125.

۶. کریکونیوک کاترین. بررسی نظام واژه سازی به روش تجزیه تکواژی صوری // نامه فرهنگستان. – شماره ۴/۱۳. – تهران: ۱۳۹۳. – ص. ۵۷-۶۷.

## АНОТАЦІЯ

**Крикониук К.М. Формальна морфологія питомих засобів словотвору сучасної перської мови. – На правах рукопису.**

Дисертація на здобуття наукового ступеня кандидата філологічних наук за спеціальністю 10.02.13 – мови народів Азії, Африки, аборигенних народів Америки та Австралії. Інститут сходознавства імені А. Ю. Кримського НАН України, Київ, 2014.

У дослідженні побудовано логічно несуперечливий якісний і кількісний опис питомих засобів словотвору сучасної перської мови у синхронічному зрізі засобами формальної морфології. Застосовано основні методологічні засади формально-морфологічного підходу, якими є: повнота, простота, однорідність, формалізація. На основі матеріалу тлумачного словника перської мови М. Каріба (29 162 слова) укладено словотвірний інвентарій та подано його квантитативні характеристики, виявлено кількісні співвідношення морфологічних класів та прямих і опосередкованих запозичень, а також обраховано процентні кореляції рекурентного та унікального корпусів словотвірного інвентарію. За допомогою макросів оптимізовано найбільш громіздкі ділянки морфологічного аналізу. Було помічено і сформульовано кілька морфологічних закономірностей, за якими відбувається організація морфологічних складових у словотвірних формулах та цілій системі. Встановлено продуктивні словотвірні моделі та види дистрибуцій, наявні у лексемах словотвірного реєстру, визначено ступені префіксації та суфіксації у морфологічних класах та описано потужність морфологічних груп коренів теперішнього та минулого часу дієслова перської мови.

**Ключові слова:** словотвір, формально-морфологічний підхід, морфологічна закономірність, словотвірний інвентарій, продуктивна словотвірна модель.

## АННОТАЦИЯ

**Крикониук К.М. Формальная морфология исконных средств словообразования современного персидского языка. – На правах рукописи.**

Диссертация на соискание ученой степени кандидата филологических наук по специальности 10.02.13. – языки народов Азии, Африки, аборигенных народов Америки и Австралии. Институт востоковедения им. А. Е. Крымского НАН Украины, Киев, 2014.

В исследовании построено логически непротиворечивое качественное и количественное описание исконных средств словообразования современного персидского языка в синхроническом срезе средствами формальной морфологии. Рассмотрены современные теоретические подходы к изучению словообразования, понимание механизмов которого помогает создать целостное представление о структуре языка. Применены основные методологические принципы формально-морфологического подхода, которыми являются: полнота, простота, однородность, формализация. На основании материала толкового словаря персидского словаря М. Карива (29162 слова) составлен словообразовательный инвентарий и представлены его квантитативные характеристики, выявлены количественные соотношения морфологических классов и прямых и опосредованных заимствований, а также рассчитаны процентные корреляции рекурентного и уникального корпусов инвентария. С помощью макросов оптимизировано описание наиболее массивных участков морфологического анализа. Сформулированы морфологические закономерности организации компонентов в словообразовательных моделях и целой системе. Выявлены продуктивные



словообразовательные модели и виды дистрибуции лексем словообразовательного реестра, определены степени суффиксации и префиксации в морфологических классах, описана мощность морфологических групп основ настоящего и прошедшего времени глагола.

**Ключевые слова:** словообразование, формально-морфологический подход, морфологическая закономерность, словообразовательный инвентарий, продуктивная словообразовательная модель.

## SUMMARY

**Krykonyuk K. M. Formalized Morphology of the Native Word Formation in Modern Persian. – In accordance with the right of Manuscript.**

A thesis submitted for a Candidate degree in Philology on scientific branch 10.02.13. – Languages of people of Asia, Africa, natives of America and Australia. A. Yu. Krymskyi Institute of Oriental Studies of the National Academy of Sciences of Ukraine, Kyiv, 2014

The purpose of this dissertation is to build a description model of Word Formation System (WFS) of Modern Persian by means of formalized morphology. A study of the word formation mechanisms is a necessary step in a holistic understanding of a language structure. The complexity of the word formation system is caused by its close connections to the other levels of a language, mobility and lack of a clear distinction between its potency and actual implementation, between derivation and inflection, as well as by the diversity and communicative significance of its functions. In modern linguistics the following main approaches to a study of word formation are distinguished: i) *onomasiological approach* proposed by M. Dokulil (1962) treats word formation within the framework of the Nomination Theory with its notions of onomasiological basis, attributes, relations, and category; the purpose of this concept is to find out how cognitively grounded mental categories are linguistically represented in each individual word formation process; ii) *generative approach* offered by N. Chomsky (1957) on a base of the mathematical automata theory interprets syntax and derivatives as generating consequence of the different syntax constructions and the Word Formation Rule as a chain of transitions that underlie transformation of the initial syntax construction into a corresponding derivative; iii) *formalized-morphological approach* allows creating logically consistent descriptive linguistic models of the inflectional and derivational system of a language in a synchronic dimension by applying the formalized procedures to the morphological analysis. Basis of this theory has been developed for the needs of machine translation by the Russian scientists I. Bratchikov (1958), G. Tseytin (1958) and S. Fitalov (1961). During 1963-2003 K. Tyschenko (Kyiv University), proposed a description of a few inflectional models in French, Italian, Portuguese, Spanish, Persian, Finnish and Ukrainian.

Formalized morphological approach includes a number of methodological guidelines: *completeness, simplicity, uniformity* and *formalization*. Completeness of formalized morphological description is achieved by the involvement in the initial stage of all linguistic phenomena of the studied area without exception that are recorded in an authoritative source or group of sources. Simplicity of the morphological description is provided by selection of the economical metalanguage, as well as commitment to transparency of the analysis procedure available for being consistently repeated, checked or verified. It is important to remember that the metalanguage of the morphological description should not be more complicated than the described linguistic object itself. Formalization of the study is assured by neat use of the laconic symbol system of the morphological categories (for example, N – root morpheme of a noun, Aj – root morpheme of an adjective, V– present tense root morpheme of a verb, Pron –

root morpheme of a pronoun; pl - plural, etc.). The principle of uniformity is aimed to avoid mechanical and eclectic mix of the mutually incommensurable tools and techniques of the research worked out on other methodological grounds.

The notions of *metalanguage*, *inventory of word-formative models*, *morphological classes*, *productivity* and *valence*, *initiale*, *mediale* and *finale*, *hapax*, *segments of unique and recurrent word-formative formulas* form the core of the theoretical structure of the formalized morphological approach. In this research metalanguage is regarded as universal formalized semiotic code (or “second-level” language) used in the morphological description of the modern Persian language (“first-level language”) which allows further classification of the obtained material, establishing various kinds of relations between its components and making calculations in order to reveal statistical regularity in WFS. A set of word-formative formulas acquired by performing the morphological analysis on all dictionary lexemes was coined by the term word-formative inventory a study of which make it possible to formulate a number of linguistic laws in the Persian WFS. Word formation inventory is considered to be a kind of a model of final states (morphological grammar of word formation) that enables presenting a morphological description of the first level. To mark the word-formative models which compose only one lexeme the notion hapax has been chosen (gr. ἄπαξ λεγόμενον – meaning “something said only once”). Hapaxes do not require the morphological analysis, they can be considered as vocabulary (and not grammar) units.

A group of hapaxes constitute a section of the unique word-formative formulas of inventory. A section of the recurrent models of inventory consists of the word-formative units which form more than two lexemes.

This research has been conducted on the Dictionary of the Persian Language and Literature compiled by Mohammad Qharib (29162 words). Inventory analysis has revealed quantitative correlations of morphological classes in the Persian language, as well as of direct and indirect loanwords from other languages. According to the performed research the morphological classes in the examined lexicon are distributed as follows: 55.44% of the total number of words account for nouns, 23.6% – for adjectives and 12.91%, 3.72%, 2.26% and 1.3% account respectively for verbs, adverbs, nouns-adjectives and adjectives-adverbs. The quantitative characteristics of the morphological classes’ correlations are estimated separately for simple and compound words. For better understanding of the obtained numbers we used similar correlation data from other languages comparison of which revealed a significant role of nouns in the Persian WFS.

The total number of loanwords (direct and indirect) in the examined lexicon is 52.4%. The largest share of lexical borrowings is from Arabic (46.4% of the examined lexicon); 2.8% of loanwords account for French, 1.04% – for Turkish, 0.2% – for English. The largest number of indirect loanwords belongs to the Greek words that were borrowed into Persian via Arabic (0.23%).

It was shown that word-formative inventory consists of seven main morphological classes (noun, adjective, verb, adverb, numeral, pronoun, verboid) and six auxiliary morphological classes (preposition, article, conjunction, interjection, interrogative words, and particle). Some words of the lexicon which under certain conditions change their grammatical role and convert from one part of speech to the other were ascribed to the conversive morphological classes (most of such words were borrowed from Arabic); overall we distinguished 17 conversive morphological classes in word-formative inventory. The morphological class of nouns is formed by 407 word-formative formulas, adjectives – by 241, adverbs – 71, verbs – 15. The total number of word-formative formulas in the conversive morphological classes is 80. Thus the Persian Native World Formation System as a whole includes 814 word-formative models. The composition of the word-formative units of the inventory in percentage is distributed as

follows: 50% of the inventory account for nouns, 29.61% – for adjectives, 8.73% – for adverbs, 1.84% – for verbs and 9.82% – for conversive morphological classes.

The WFS of the specific morphological tools in Modern Persian was comprehended as a dynamic analytical model which has a multi-faceted structure with distinct internal regularities and relations. The components of this model are organized on a basis of nature of initials and finales – whether they are formed by affixes or the morphological classes (division of the system into four morphological subsystems), productivity or non-productivity of the word-formative models (division of inventory into segments of the unique and recurrent word-formative formulas; differentiation of three Bradford zones in the recurrent segment of word-formative formulas), as well as on a basis of attributing a word-formative model to a certain part of speech (distribution of word-formative formulas in the morphological classes). Shown tendencies in the Persian WFS concern both its inner and outer structure (regularities of the initials' and finales' organization in word-formative formulas; Estoup-Zipf law 1, 2, 3).

All the regularities of the inner word-formative structure were formulated by means of matrixes which final version became possible due to the method of optimization. During matrix optimization of the WFS's noun six word-formative plurals were defined ( $\{a-\emptyset-C\}$ ,  $\{a-C-a\}$ ,  $\{C-\emptyset-a\}$ ,  $\{C-C-a\}$ ,  $\{C-a-a\}$  та  $\{C-aC/Ca/aaC-a\}$ , where a – affix, C – the morphological class,  $\emptyset$  – zero morpheme). Every word-formative plural consists of a certain number of structural types.

The major findings of this research include the following key results:

1. Consistently applied formalized morphological approach has proved to be effective in creating a quantitative and qualitative description model of WFS of Modern Persian in the synchronic dimension.
2. A wide range of morphological problems solved by applying the main methodological principles of the formalized morphological approach (such as completeness, simplicity, uniformity, formalization etc.) demonstrate the sufficiency of this concept to solve linguistic issues in the domains related to word formation.
3. By means of formalized morphological approach the main word formation rules have been formulated. They depict the inner and the outer structure of the Persian WFS.
4. The obtained results correspond with the general linguistic problems of the structural linguistics viz. problems of cybernetics and linguistics modeling (in which way the possibilities of the human language and intellection correlates with the possibilities of the machines that process information), problems of the morphological typology of languages revealed on a basis of the paradigmatic type of relations in linguistic units, problems of the morphological group capacity and the problems of the linguistic probabilities (associated with the valence of morphemes).

**Key words:** word formation, formalized morphological approach, morphological rule, word-formative inventory, productive word-formative model.