
УДК 811.161.2'373.4

Василь Старко

м. Луцьк

КОРПУСНІ ДАНІ В ДОСЛІДЖЕННІ УКРАЇНСЬКИХ КОЛОРАТИВІВ

У статті розглянуто електронний корпус як джерело мовних даних у реконструкції референційних моделей українських колоративів блакитний, голубий і синій. Показано, що корпуси дають змогу встановити недоступні для інших методів аспекти функціонування кольороназв. Зроблено висновки щодо моделей досліджуваних колоративів. Проаналізовано переваги й окреслено можливі шляхи вдосконалення використаного корпусу української мови.

Ключові слова: корпус, корпус української мови, категоризація, колоратив, кольороназва, колірна категорія, референтний взірць, референційна модель.

Дослідження кольороназв має багату історію. У розрізі вивчення природномовної категоризації виховою стала праця Б. Берліна й П. Кея [6], яка дала поштовх інтенсивному вивченню колірних категорій у різних мовах. У сучасній когнітивній лінгвістиці один із перспективних напрямків дослідження колоративів передбачає виявлення референтних взірців (еталонів, прототипів), на які спираються мовці, імпліцитно чи явно, коли вживають назви кольорів. Заведено, що такими еталонами найчастіше є культурно важливі об'єкти (природні чи артефактні). Опертя на референтні взірці дає змогу пояснити, чому менш чи більш подібні колірні категорії в різних мовах мають неоднакові межі. На категоризацію кольору можуть впливати одночасно кілька чинників. Зокрема, у різних мовах фокальна спектральна зона для певного набору кольороназв-відповідників може великою мірою збігатися завдяки нейрофізіологічним властивостям

© В.Ф. СТАРКО, 2014

зорового апарату людини, однак межі категорії залишаються мовноспецифічними, і їх найкраще можна пояснити зв'язком із еталонними референтами.

Колірна палітра української мови особлива тим, що для «синьої» частини спектру вона має аж три базові кольороназви: *блакитний*, *голубий* і *синій*. На відміну від таких номінацій, як *волошковий* чи *небесний*, ці колоративи не мають явних, виражених у самій лексичній формі еталонів — взірці, що стоять за ними, імпліцитні, й встановити їх важче. В україністиці вже зроблено спроби розв'язати це завдання. С.В. Мартінек на основі результатів вільного асоціативного експерименту дійшла висновку, що прототиповим референтом для всіх трьох колоративів є небо, однак *блакитний* і *голубий* частіше викликають асоціацію з небом, ніж *синій*. Крім того, для категорії *синій* більшу роль відіграють водні об'єкти як вторинний референтний взірець; цей колоратив частотніший й охоплює ширшу частину спектра [4]. На основі аналізу контекстів, взятих із Лексичної картотеки Інституту української мови НАНУ, Г.М. Яворська дійшла висновку, що кольороназва *синій* має два референти-еталони: основний — небо й додатковий — море (переважно з огляду на усталену сполуку *синє море*). Натомість вода є «базовою», а небо — другорядною референцією для *голубий*; взірцем *блакитного* є «небо в ясний, сонячний день» [7: 30].

Послугуючись методом тренованої інтроспекції та результатами «неформального» опитування носіїв мови, А. Вежбицька [2] проаналізувала низку колоративів із різних мов, які позначають відтінки синього кольору. Усі вони по-різному пов'язані з уявленням про небо. Дослідниця постулює, що зв'язок колоратива зі взірцем може мати різні ступені, наприклад, пол. *blekitny* тісніше, ніж пол. *niebieski*, асоціюється з кольором неба, причому зв'язок цей непрямий. Важать також умови спостереження, наприклад, рос. *голубой* та *синий* розрізняють за рівнем освітленості спостережуваного референтного взірця (неба).

Отже, щодо об'єкту таких досліджень можна зауважити: реконструкція має бути зорієнтована на встановлення не окремих взірців, а цілих референційних моделей, які охоплюють умови спостереження, можуть складатися з кількох взірців різного ступеня важливості та відбивають вплив низки чинників категоризації. Під час дослідження залучено дані мовного корпусу, щоб здійснити докладнішу реконструкцію референційної моделі.

Поява мовних корпусів відкриває набагато ширші можливості текстових досліджень кольороназв. Окрім значно більшого масиву контекстів дослідник дістає змогу проводити кількісні й статистичні оцінювання. Україністи мають у своєму розпорядженні поки що один такий повноцінний інструмент дослідження — онлайнний «Корпус української мови» (КУМ) [3], що має статус дослідницького. Його розробили працівники лабораторії комп'ютерної лінгвістики Київського національного університету імені Тараса Шевченка під керівництвом

Н.П. Дарчук. На сьогодні у вільному онлайн-доступі викладено частину корпусу обсягом 13 млн. слововживань [1]. Він складається з чотирьох підкорпусів: художня проза, нехудожні тексти, поетичні тексти, фольклорні тексти. Розміри виставлених онлайн-підкорпусів суттєво різняться. Найбільша частка припадає на художню прозу, яка матиме для нашого дослідження визначальну вагу. Другий за обсягом підкорпус поетичних текстів відіграватиме допоміжну роль. Розмір третього і четвертого надто малий для того, щоб робити певні висновки про вживання колоративів. Всі наведені нижче ілюстрації, а також вихідні дані до них можна знайти в корпусі через наданий інтерфейс.

Методика опрацювання корпусних даних мала кілька етапів. Спершу ми зробили повну вибірку контекстів з досліджуваними колоративами. Інтерфейс корпусу дає змогу шукати лексему одразу в усіх словозмінних варіантах. Далі в межах одного жанру контексти проаналізовано й розподілено на групи за потенційним референтним взірцем. Групування спиралося на поняттєвий критерій, завдяки чому такі номінації, як *в блакитному зеніті*, *блакитні бані небесні*, *блакитне небесне поле*, потрапляли в одну групу, що відповідає взірцю «небо». Далі сформовано перелік потенційних взірців та встановлено частку кожного з них щодо загальної кількості контекстів у певному жанрі. На нашу думку, цей відсоток тісно корелює з відносною вагою відповідного взірця в референційній моделі. Дальший текстовий аналіз зосереджено на виявленні умов спостереження, переносних вживань, порівнянь тощо.

У підкорпусі художньої літератури *синій* трапляється приблизно вдвічі частіше, ніж *блакитний*, і вчетверо, ніж *голубий*. Домінантним взірцем усіх трьох колоративів є небо. Очевидно, що відмінності між ними приховані в інших аспектах референційних моделей.

Колоратив *блакитний* зафіксований у корпусі 770 разів: 530 у художній прозі, 100 в нехудожніх текстах, 140 у поезії, 0 разів у фольклорних текстах. Нехудожнім текстам, відібраним до корпусу, — це переважно публіцистика та певна кількість наукових текстів — не властиве вживання колоративів в описовій функції: найчастотнішою сполукою (41%) є *блакитне паливо* та інші сполуки-перифрази на позначення газу. Поширені також сполуки на кшталт *блакитні екрани* (*блакитний ящик*, *блакитний портал*) — 9%. У поетичних текстах виразно окреслено колір води (9%) та неба (7%). Повторюються сполуки *блакитна кров* (4%), *блакитні очі* (3%), *блакитний світ* (3%), позначення холодного світла, блиску (*блакитні блискавиці*, *блакитне полум'я*, *блакитні ятагани ковзанів* тощо — 3%) та далечіні (*блакитні гори*, *блакитне безконеччя* тощо — 3%). Зафіксовано також низькочастотні сполуки *блакитне море*, *блакитний день* (*вечір*, *ранок*), *блакитна весна* (*зима*). Підкорпус художньої прози містить найбагатший матеріал й уможливорює докладнішу реконструкцію. Тут домінують дві сполуки: *блакитні очі* (12%) та *блакитне небо* (11%). Колір очей не може слугувати еталоном кольору, однак прикметно, що очі уподібнено за кольором до потенційних взірців, наприклад:

Кругленьке та рум'яне обличчя, з блакитними, як ранкове небо, очима, чорними, вузькими бровами та червоними, як коралі, губоньками... (А. Кашенко, Зруйноване гніздо);

Мала вона світле довге волосся, що вкривало її наче плащ, блакитні очі сяяли, наче джерела, освітлені сонцем (Ю. Винничук, Красуня зі студні);

Блакитні очі можуть бути такими сонячними! (Л. Денисенко, Сарабанда банди Сари).

Відзначмо виразний зв'язок із сяйвом сонця у двох останніх реченнях, що його спостерігаємо й у таких ілюстраціях до вірця «небо»:

Він оббігав поглядом своїх дивно сивих очей тісний форум, різонуло його по самому серцю буйністю барв на святкових шатах, м'якою осінньою позолотою покривало довколишні будівлі сонце з дивно блакитного неба... (П. Загребельний, Диво);

Біленькі хмарки пливли по блакитному морю, купаючися в теплому світлі (Б. Грінченко, Під тихими вербами);

В тих степах, де небо чисте, аж сяюче, вражає вам душу своєю незбагненною блакитною неосяжністю... (О. Гончар, Твоя зоря).

Ще з візантійських часів блакитну барву використовували як колір небес у розписах бань церков, іконах тощо. Сильну позитивну асоціацію блакитного з небесами як оселею Бога й людських душ відбито в корпусних даних, зокрема зафіксовано уподібнення блакитного неба до бані храму й навпаки:

Чого він так страшно дивиться, коли довкрузь — такий чистий польовий простір і такі ясні блакитні бані вгорі небесні? (Є. Гуцало, Голодомор);

Навіть Свято-Троїцька церква горіла і падала на його очах не раз, ворон її пам'ятає ще дерев'яною із блакитними банями у срібних зорях (як поглянеш було на неї згори вночі, то таке, наче небо, лежить під тобою) (В. Шкляр, Залишенець/Чорний ворон).

Тісна асоціація з яскравим сонячним світлом, а також релігійні уявлення про небеса стали, вочевидь, джерелом майже без винятку позитивних конотацій *блакитної* барви, наприклад:

Утікали хвилі од вітру. ... А воно [небо] сміялось до них широким, безмежним блакитним усміхом, і той усміх спадав на їх золотим дощем сонячним, несучи їм життя й силу (Б. Грінченко, Під тихими вербами).

Контекстів, де блакитними виступають водойми (канали, озера, річки, хвилі тощо) й просто вода, у підкорпусі зафіксовано 4%:

А мокра після дощу земля? А веселі калюжі, які щойно аж кипіли од рясного краповиння, були каламутні, аж сиві, а тепер ясні та блакитні, і в кожній полощеться сонце? (А. Дімаров, На коні й під конем);

Скоро Гнідко вже лисніє від сонячного молока, вже його грива пливе в чистій блакитній воді, зграйка дрібних рибок спурхує попереду над водою (Є. Гуцало, Княжа гора);

Ген унизу видніє Дніпро, прослав блакитні, повиті тоненьким маревом полтна блакитної води, яким кінця й краю не видно (Є. Гуцало, Княжа гора).

Із підвищої вибірки цитат видно, що характерною ознакою візуальної ситуації, в якій фігурує *блакитна* вода, є наявність яскравого сонячного світла. Пункт спостереження може бути близький або віддалений (в передостанньому й останньому прикладі, відповідно). Вислів *блакитне море* трапляється в підкорпусі рідко (1%), однак навіть ці нечисленні контексти засвідчують ситуацію яскравого природного освітлення:

Він не міг годинами стояти перед пюпітром і розучувати гами в той час, коли за вікнами світило сонце, сяяло блакитне море (П. Загребельний, Учитель);

Ви мене розбудили. А мені снилося море — блакитне, тихе-тихе. І коттедж на березі. Пальми, пляж (О. Бердник, Чаша Амріти).

Аналіз корпусних даних дав змогу виокремити ще одне коло вживань колоративу *блакитний*. Йдеться про низку відмінних на лексичному рівні контекстів, які передають колірний ефект маси повітря, освітленої сонцем. На цей референтний взірць припадає 3,5% вживань:

Задкуючи, він вийшов з церкви, заплющився від яскравості блакитного кийського дня... (П. Загребельний, Диво);

Ранок сліпучо сонячний, здивовано блакитний... (В. Винниченко, Сонячна машина);

Йому [Андрію] хочеться бодай спиною відчувати волю, безмежний простір, кілометри блакитного повітря (П. Загребельний, Дума про невмирущого).

Той самий природний ефект лежить в основі асоціацій із далечіню, й у цьому ж ряду стоїть словосполучка *блакитні гори*, наприклад,

Я вже підходив до вікна, тримаючись за стіну. Видно було ліс, блакитні зубці гір на обрії, білопінний потік недалеко від хати (О. Бердник, Чаша Амріти).

Ситуація вдивляння в далечінь, очевидно, є джерелом уживань, де *блакит* асоціюється із ностальгією переважно за ідилічним минулим:

А зараз викликав своєю обдертістю щось схоже на співчуття, був чимось ближчий до її дитячих мрій, до отих блакитних планет її дитинства (О. Гончар, Собор).

Блакитний уживається також із цілим рядом лексем на позначення холодного світла, блиску (3%): *блакитна блискавиця, блакитні сніги, блакитна ніч, блакитне місячне сяйво* тощо. Холодний відтінок *блакитного* властивий і диму, імлі, туману, серпанку (2%), наприклад, *блакитний димок смеркання, блакитна світанкова мла* тощо.

Отже, домінантним взірцем *блакитного* кольору є небо (11%), а найвиразнішими допоміжними — повітряний простір (3,5%), водойми й море (4% і 1%). Прикметною особливістю ситуацій спостереження цих взірців є яскрава природна освітленість. Кореляція з теплим світлом пояснює домінування позитивних конотацій, властивих *блакитній* барві. Тісними є також асоціації *блакитного* з холодним світлом чи блиском (3%) та імлою (2%). Джерелом конотації ностальгії є ситуація вдивляння в далечінь (взірець «повітряний простір»).

Як бачимо, матеріал корпусу дає змогу не лише встановити референтні взірці кольороназви, а й визначити їхню відносну вагу в референтній моделі. Розгорнуті контексти вживання уможливають реконструкцію характеристик еталонних візуальних ситуацій і пояснення переносних уживань.

Аналіз контекстів із кольороназвами *голубий* і *синій* показав, що вони мають такі самі еталонні референтні ситуації. Відмінності ж полягають в інших аспектах моделей. Колоратив *голубий* зафіксований у корпусі 478 разів: 281 у художній прозі, 175 у поезії, 22 в нехудожніх текстах, 0 разів у фольклорних текстах. У ключовому для нас підкорпусі художньої прози хотілося б мати більше ілюстративного матеріалу, адже йдеться про з'ясування тонких нюансів вживання й аспектів референтної моделі. Для порівняння можна навести такий факт: укладаючи статтю до тлумачного словника англійської мови, західний лексикограф має опрацювати, у випадку досить часто вживаного слова, до 1000 контекстів. Щоб дійти висновків стосовно кольороназви *голубий*, нам потрібно мати в розпорядженні корпус щонайменше в 3–3,5 рази більший. Щоправда, важить не лише обсяг, а й наповнення, наприклад, менший корпус із текстів, насичених колоративною лексикою, може виявитися достатнім для дослідження.

Аналіз наявних у корпусі контекстів показав, що найвиразнішим еталоном голубого кольору виступає небо (10%), однак на відміну від *блакитного* зв'язок із сонячним світлом значно слабший і проступає в меншій кількості контекстів. Натомість виразні інші характеристики, як-от чистота, свіжість, просторість тощо:

...він уперше помічав небо — ще не спечне, голубе, свіже (В. Дрозд, Катастрофа);

Глянув Звірко в голубу безодню неба, і йому видалося, що сонце пригасло, а небо враз потьмарилось (Р. Іванченко, Сіті життя і смерті);

Весна йде та йде. Ось вже і небо голубе і чисте, і вода голуба, про-світчаста. Сонечко блищить і горить; гаї розвиваються; садки зацвітають (Марко Вовчок, Три долі).

Значно менш частотний взірець «повітряний простір» (3%) теж асоціюється переважно із чистотою й прозорістю:

Туман осідав на дно гаварецького ізвору, немов каламуть у збовтаному плесі, голубі хвилі чистого повітря спадали із зелених схилів, наповнювали велетенську чашу... (Р. Іванчук, Манускрипт з вулиці Руської).

На інший колірний взірець, водойми, припадає 6% вживань лексеми *голубий* у підкорпусі художньої літератури. Голуба вода уявляється чистою, прозорою, однак може бути й темною:

Той Дунай, про який ми стільки передумали, — чистий, голубий від неба... (О. Гончар, Прапороносці).

Досвіток, золоті зорі вишиті на сірому полотні неба. ... Церква стоїть на високому горбі коло ставу, що дримає темною голубою водою в густих верболозах (Є. Гуцало, Голодомор).

Прикладів зі сполукою *голубе море* зафіксовано лише три (1%). На контексти, пов'язані з холодним світлом і блиском різних об'єктів (*голубі сніги, голуба ніч, голубе проміння зір* тощо), припадає 5%, а на ті, що позначають особливий голубий відтінок туману, імлі, диму, серпанку, — 4% усіх вживань, наприклад, *голубе марево, голубий туман, голубий серпанок* тощо. Пов'язані з небом конотації *голубого* (спокій, задумливість) надто нечисленні. Частіше проступає ностальгія, причому вона пов'язана, очевидно, з відчуттями мовця-спостерігача, який вдивляється в далечінь:

Очі Зіньчини близько до Давидових із сутіні... Як в оті давні голубі літні ночі десь на межі... (А. Головка, Бур'ян).

Голубому властиві й певні негативні конотації (біль, журба, жаль, тривога), які зрідка трапляються в поезії й, схоже, не мають прозорої мотивації. Корпус фіксує й вживання *голубий* на позначення гомосексуалістів, тоді як *блакитний* майже не вживано в цьому значенні.

Відомо, що українці, які пильнують своєї мови, в наш час віддають перевагу *блакитному* як «своєму» слову перед *голубим*, що його вважають росіянізмом. Тобто на вжиток впливає й соціолінгвістичний чинник. Було б цікаво простежити динаміку вживання прикметника *голубий*, зокрема порівняти роки незалежності з попередніми періодами. Це стане можливим, коли КУМ досягне збалансованості за цими часовими відрізками, а в інтерфейсі з'явиться змога задавати пошук за періодами.

Отже, корпусне дослідження вживання колоративу *голубий* показує, що домінує взірець «небо» (10%), а взірці «водойми» (6%) та повітряний простір (3%) перебувають на других ролях. В усіх цих випадках зв'язок із сонячним світлом набагато слабший, ніж у колоративу *блакитний*, а натомість переважають такі ознаки, як чистість і прозорість. Досить чіткий зв'язок зафіксовано з об'єктами, що відбивають чи випромінюють холодне світло (5%), та еталоном із умовною назвою «імла» (4%). Деякі конотації *голубого* пояснено через віднесення до ситуації вдивляння в далечінь.

Кількість фіксацій кольороназви *синій* розподілено так: художня література — 1090 разів, нехудожня проза — 73, поезія — 441, фольклорні тексти — 2. У підкорпусі поезії спостерігаємо цікаву ситуацію: найчастотнішою є сполука *синє море* (69 вживань, 16% від загальної кількості), однак її розподіл вкрай нерівномірний. У творах Тараса Шевченка зафіксовано 105 вживань *синій* і 43 рази *синє море*, тобто частка номінацій моря становить 41%. У поезіях Лесі Українки відповідні показники такі: 18 і 14 (78%). Решта підкорпусу поезії містить 318 вживань *синій* і лише 12 *синє море* (4%). Як бачимо, на цих двох авторів припадає 28%, і врахування їхніх творів збільшує частотність номінації *синє море* з 4% до 16%. Причина такого феномену лежить, очевидно, в близькості мови обох поетів до українського фольклору. Вказана незбалансованість фактично не дає змоги оцінити з певністю частотність сполуки *синє море* в українській поезії. Для цього потрібен набагато

більший і краще збалансований корпус поетичних творів. Без нього також не обійтися в дослідженні конотацій синьої барви: наявний підкорпус фіксує лише одиничні контексти, переважно з поезій В. Стуса, на позначення ностальгії, смутку, жалю, туги, тривоги.

У підкорпусі художньої літератури ми вивели уже знайомі нам референційні взірці. Першість незмінно тримає «небо» (12%), причому в контекстах чітко зафіксовано зв'язок між світлими й темними відтінками синього й відповідними умовами спостереження. З одного боку, світло-синє небо описують як яскраве, сліпуче, незахмарене, освітлене сонцем, осяйне, ясне тощо. Ситуації, де небо спостерігають в умовах максимального денного освітлення, домінують. З іншого боку, небо нерідко наділяють епітетом *синє* тоді, коли освітлення слабше або взагалі мінімальне — вдосвіта, надвечір або вночі. У таких випадках небо має виразно темніший відтінок, а вночі може наближатися до чорного, наприклад:

Виходимо на безлюдну вуличку, поволі вечоріє: синє небо все дужче густішає і темніє (А. Дністровий, Пацики).

Сполука *синє море* трапляється лише 25 разів (2,3%), причому деякі вживання мають виразно фольклорний характер. Позначення водойм та води фігурують у 4,7% контекстів і разом із морем становлять важливий компонент референційної моделі (7%). Цікаво, що в цій групі річки фактично дорівнялися за частотністю згадок до моря. Багатство наявних у підкорпусі контекстів дає змогу реконструювати пункт спостереження: на березі водойми, з висоти, у самій водоймі, при яскравому світлі чи в сутінках.

На взірці «повітряний простір» припадає 5% контекстів. Знову ж таки виразно окреслюються дві ситуації спостереження: 1) високий рівень освітленості, коли синя барва варіює від блідо-синього (*синій ранок*) до яскравого насиченого тону (*синій, чистий небокрай*) до темніших відтінків (*сині гори*); 2) понижений чи низький рівень освітленості у вечірній і нічний час (*сині смерки, синього зимового вечора, лагідна синя ніч* тощо). Взірець «імла» фігурує в 3% контекстів, а об'єкти з блиском чи світлом синюватого відтінку — у 2%. Синіми видаються також різні квіти: волошки, барвінок, петрові батоги, терен, льон, проліски тощо. Із конотацій у художній прозі найвиразніше окреслена ностальгія (корелює зі спогляданням далечіні) й трапляються поодинокі асоціації з мрійливістю, смутком та життєрадісністю, навіяні виглядом яскраво-синього неба.

Отже, проведений аналіз дав змогу встановити колірні взірці, що входять до референційних моделей трьох українських колоративів, та розставити їх у певній ієрархічній послідовності. Результати показують, що значну роль відіграє не лише сам об'єкт-референт, а й умови його спостереження. Зокрема для кольороназв *синій, блакитний* і *голубий* важить ступінь освітлення: *синьому* властивий широкий діапазон, а *блакитний* виявляється дуже тісно прив'язаним до наявності яскравого

сонячного світла. Крім того, тепер ми маємо підстави твердити, що колірний взірець не конче мусить мати сталий колір. Точнішею виявляється кореляція з **еталонними візуальними ситуаціями**, в яких взірець спостерігають за певних умов. Колірний взірець має бути цілісним. Саме з цієї причини мусимо відмовити у цьому статусі об'єктам із холодним блиском чи світлом: група залишається гетерогенною, без чіткого переважання одного чи кількох об'єктів. У цьому випадку йдеться, швидше, про типове застосування, прикладення кольороназв, ніж про джерело їхньої референційної здатності. Ці міркування стосуються й моря. Сполука *синє море* — це усталена дескрипція, особливо у фольклорі, але позаяк море не є компонентом щоденного довілля більшості українців, воно може відігравати хіба що допоміжну роль як референтний взірець. Основні висліди нашої розвідки подано в нижчій таблиці.

Насамкінець підкреслимо, що корпус дає досліднику унікальний інструмент вивчення мови, уможливаючи, зокрема, кількісні виміри й порівняння. Корисність і реалістичність корпусу прямо залежать від ступеня його збалансованості й репрезентативності. Шляхи підвищення цих характеристик корпусу, зокрема для української мови, ми докладно розглянули в іншій праці [5], а тут вкажемо на особливості КУМ, які потребують дальшого розвитку. Дослідникові, який працює з цим корпусом, значну частину роботи досі доводиться виконувати вручну, адже програмний комплекс поки що не має низки функцій: формування статистики за запитом, показ обсягу (під)корпусу, одноразовий пошук у всьому корпусі, задавання відстані між шуканими словами. Бракує також виведення результатів у поширеному й зручному форматі KWIC (Key Word In Context), коли посередині подано шукане слово, а по обидва боки — безпосередній контекст. Понад те, знайдені в КУМ уривки бувають закороткі, особливо в підкорпусі поезії, а змоги побачити ширший правий і лівий контексти немає. Хоча параметр «стать» прописаний у параметрах корпусу, тексти за ним не збалансовані, й твори авторів-жінок представлені вкрай скупо. Наприклад, у підкорпусі художньої літератури ми виявили (з необхідності — шляхом підрахунку вручну) лише 28 вживань слова *небо* в «жіночих» текстах і майже 4,5 тисячі в «чоловічих». Наскільки нам відомо, автори КУМ

Таблиця 1. Референтні взірці кольороназв блакитний, голубий і синій за даними підкорпусу художньої прози КУМ, у відсотках до загальної кількості вживань.

Взірець \ Колоратив	блакитний	голубий	синій
небо	11	10	12
повітряний простір	3,5	3	5
море	1	1	2,3
інші водойми	4	6	4,7
імла	2	4	3

планують вдосконалювати інтерфейс і наповнення корпусу, а тому є підстави сподіватися, що вже найближчим часом дослідники матимуть змогу скористатися оновленою, збалансованішою й функційно багатшою версією цього надзвичайно корисного інструмента дослідження української мови.

1. *Алексієнко Л.* До двадцятиріччя лабораторії комп'ютерної лінгвістики / Л. Алексієнко, Н. Дарчук // Українське мовознавство. — Вип. 43. — Київ, 2013. — С. 3–10; режим доступу: http://www.philolog.univ.kiev.ua/library/zagal/Ukr_movoznavstvo_2013_43/3-10.pdf
2. *Вежбицкая А.* Обозначения цвета и универсалии зрительного восприятия / А. Вежбицкая // А. Вежбицкая. Язык. Культура. Познание. — М.: Русские словари, 1997. — С. 231–290.
3. *Корпус української мови*, режим доступу: <http://www.mova.info/corpus.aspx>.
4. *Мартінек С.В.* Концептуалізація синього кольору (за результатами асоціативних експериментів) / С. В. Мартінек // Науковий щорічний журнал «Мова і культура». — К.: Вид. дім Дмитра Бураго, 2005. — Вип. 8. — Т. II. Психологія мови і культури. Мова і засоби масової комунікації. — С. 31–39.
5. *Старко В.* Параметризація корпусу як спосіб підвищити його репрезентативність та збалансованість / В. Старко, Н. Чейлітко // Українське мовознавство. — Вип. 43. — К., 2013. — С. 87–94; режим доступу: http://www.philolog.univ.kiev.ua/library/zagal/Ukr_movoznavstvo_2013_43/87-94.pdf
6. *Berlin B.* Basic Color Terms: Their Universality and Evolution / B. Berlin, P. Kay. — Berkeley: University of California Press, 1969. — 166 p.
7. *Javorska G.* O podstawowych nazwach barw w języku ukraińskim / G. Javorska // Studia z semantyki porównawczej: [Red. R. Grzegorzycykowa, K. Waszakowa]. — Warszawa: Wydawnictwo Uniwersytetu Warszawskiego, 2000. — S. 29–38.

Статтю отримано 17.10.2013

Vasyl Starko

Lutsk

CORPUS DATA IN THE STUDY OF UKRAINIAN COLOUR NAMES.

The article considers an electronic corpus of Ukrainian as a source of linguistic data in the reconstruction of referential models underlying the Ukrainian colour names *blakytnyi*, *holubyi* and *syniy*. It is shown that corpus data permits studying aspects of colour names that are inaccessible to our methods. Conclusions are reached about the models of the colour names in question that function as categorization principles for the respective colour categories. The advantages of the corpus used are analyzed, and possible ways of improvement are suggested.

Key words: corpus, corpus of Ukrainian, categorization, colour name, colour category, reference point, reference model.