

Г. ГЕНТШЕЛЬ

Університет ім. Карла фон Оссецького (Інститут славістики)
м. Ольденбург, Німеччина
Електронна пошта: gerd.hentschel@uol.de
<https://orcid.org/0000-0003-2034-1577>

О. О. ТАРАНЕНКО

Інститут мовознавства ім. О. О. Потебні НАН України
м. Київ, Україна
Електронна пошта: oataran99@gmail.com
<https://orcid.org/0000-0002-0269-8954>

ДВОМОВНІСТЬ ЧИ ТРИКОДОВІСТЬ: УКРАЇНСЬКА МОВА, РОСІЙСЬКА МОВА І «СУРЖИК» В УКРАЇНІ (аналіз і лінгвістично-географічне картографування)¹

Пропоноване дослідження має на меті показати, що різну частотність використання в повсякденному житті України української мови, російської мови і «суржику» можна описати й картографувати як континуум мовної «трикодовості». У цьому плані стаття відрізняється від попередніх досліджень на теми мовної ситуації в сучасній Україні, автори яких звичайно відштовхуються від факту історико-політичного розмежування субрегіонів країни і, як правило, говорять про її чіткий, однозначний поділ на переважно українськомовний (з охопленням Заходу і Центру) і переважно російськомовний (з охопленням Сходу і Півдня) регіони. Підхід, запропонований у статті, використовує емпіричні дані щодо використання трьох мовних кодів (за самооцінками респондентів) як основу для структурування мовної ситуації в Україні. За допомогою кластерного аналізу розглядаються найвищі частотності використання мовних кодів в окремих областях, які потім групуються за більшими регіонами, що супроводжується поясненнями їхнього внутрішнього варіювання. Результатом дослідження є не тільки більш континуальна (цілісна й послідовна), а й більш логічна й зрозуміла картина мовного ландшафту сучасної України.

¹ Пропонована стаття є дещо переробленою версією попередньої статті англійською мовою: [Hentschel, Taranenko 2021], яка також містила певну інформацію, відому українському читачеві. Вона ґрунтується на результатах двох проєктів: а) «Варіативність і стабільність у змішаному субстандарті в умовах широких та стабільних у часі мовних контактів: українсько-російський “суржик” в Україні» під керівництвом Г. Гентшеля (м. Ольденбург, Німеччина), профінансованого Фондом імені Фріца Тіссена (Кельн, Німеччина); б) «Гібридизація з двох сторін: українсько-російське і російсько-українське змішування кодів у контексті соціолінгвістичної ситуації на півдні України вздовж узбережжя Чорного моря» під керівництвом Т. Ройтера (м. Клагенфурт, Австрія) і Г. Гентшеля (Ольденбург), який спонсорується в рамках австрійсько-німецького партнерства FWF — Der Wissenschaftsfonds (№ I 4189-G30) і Deutsche Forschungsgemeinschaft (DFG, № 419468937), у співпраці з Т. Кузнецовою (м. Одеса, Україна). Автори висловлюють вдячність О. Палінській (Ольденбург) за різні коментарі до попередньої версії статті та К. Берд (Берлін) за переклад. Карти створені Т. Робберсом (Ольденбург).

Ключові слова: мовна ситуація в Україні, українсько-російська двомовність, українсько-російський «суржик», соціолінгвістика, лінгвістична картографія, мовні контакти.

1. Вступ: змішаний ідіом (українсько-російський «суржик» — УРС) як українська спадщина російського колоніалізму. Україну традиційно розглядають як країну з двома основними мовами спілкування — українською і російською. Крім того, як наслідок тривалих контактів між цими — історично спорідненими та структурно досить подібними східнослов'янськими мовами — сформувався особливий ідіом — змішане українсько-російське мовлення, що його звичайно іменують «суржи́ком» (це його розмовна назва) [див. докладніше про соціальні умови та історичні обставини формування та поширення УРС протягом кількох останніх століть, зокрема: Тараненко 2007; Масенко]. Ставлення різних верств українського суспільства до УРС і, відповідно, тією чи іншою мірою й до його носіїв варіюється в досить широкому діапазоні — від різкого засудження його як колоніальної спадщини російського панування, «кровозмісного дитяти двомовності» (за образною характеристикою Юрія Андруховича) з боку національно-культурних еліт [див., наприклад, огляд: Ставицька] до загалом спокійнішого його сприймання в широких колах, зокрема, «простого народу» [див., наприклад: Bilaniuk; див. за даними соціологічного опитування в центральних регіонах країни: Hentschel, Zeller 2016 : 652–657] і до легкої поблажливості щодо його використання в колах ліберальної інтелігенції (сама назва «суржик» виникла на основі позначення невисокої якості суміші зерна різних культур, а також борошна з такої суміші)².

Співіснування й, відповідно, та чи інша функціональна конкуренція між українською [національною / літературною] і російською [в основному літературною] мовами в межах України, а також формування на основі цих двох мов-донорів тих чи інших змішаних структурно-функціональних утворень має в країні, як відомо, тривалу й досить складну історію. Російська мова, поширюючись в Україні ще з другої половини XVII століття й витісняючи українську мову, стала домінувати в різних сферах соціально-політичного, культурного, релігійного життя, у сферах науки, виробництва, армії, спорту. Цей процес значно прискорився в період радянської індустріалізації, яка супроводжувалася зростанням темпів урбанізації, що призвело також до того, що в побуті стало набувати значно ширших обсягів функціонування українсько-російського змішаного мовлення (однак зі значно меншим ступенем проникнення елементів російської мови порівняно з «трасянкою») [див.: Menzel, Hentschel; Zeller; Hentschel 2018]. Вирішальну роль у цьому відігравали два чинники. По-перше, сільське населення, яке в основному спілкувалося місцевими українськими говорами, переїжджало на проживання до міст, з переважанням російської мови. Мігранти із сільської місцевості адаптувалися до мовних норм міського середовища, чому сприяла, з одного боку, генетична і структурна близькість обох мов, а з другого, — відсутність необхідності повного переходу на іншу мову спілкування, адже, навіть якщо перехід відбувався лише частково, люди могли розуміти одне одного [див. Taranenko 2014, 274–276]. По-друге,

² Подібна мовна ситуація спостерігається, як відомо, в сусідній Білорусі [див., наприклад: Hentschel 2017; Hentschel 2018], де змішане білорусько-російське мовлення стали називати «трасянкою» (далі будуть наводитися окремі дані в плані порівняння «суржику» і «трасянки»).

мігранти із сіл розмовляли таким змішаним мовленням також з дітьми та онуками, що сприяло його усталенню [див. Hentschel 2017 : 19–21]. Українська літературна мова, отримавши певну державну підтримку й, отже, право на офіційне використання лише з 1917/18 рр. (у різних формах української державності) і в 1920-х роках у період українізації, використовувалася в житті країни, звичайно, протягом усього радянського часу, але фактично в дуже обмежених обсягах. І тільки зі здобуттям 1991 року державної незалежності України українська мова стала розширювати діапазон свого соціального функціонування, що має своїм наслідком не тільки поступове витіснення російської мови з її попередніх соціальних позицій в Україні, а й помітне поширення вжитку вже російсько-українського «суржику»: мовці з базовою російською мовою користування тепер нерідко переходять на постійне або, частіше, ситуативне користування українською мовою, і це може призводити на індивідуальному рівні до утворення змішаного російсько-українського мовлення (деякі дослідники називають це «неосуржиком» на основі російської мови [див., наприклад: Flier]).

Однак, оскільки постійне й автоматичне користування українською літературною мовою поки що явно недостатньо поширилося на побутовий рівень спілкування широких верств населення, це продовжує сприяти неоднокодовості його мовного спілкування, із залученням у тому числі й УРС.

У пропонованому дослідженні буде показано з використанням статистичних даних, як три коди — українська мова, російська мова й УРС — розподіляються за різними регіонами країни як засоби звичайного спілкування в повсякденному житті різних груп населення. Увагу в цьому аналізі зосереджено на мовній неоднокодовості населення досліджуваних регіонів України, у тому числі з використанням «суржику», і обрано [за: Hentschel, Taranenko 2015; Гентшель, Тараненко 2015] кількісний соціолінгвістичний підхід як передумову для подальшого квалітативного аналізу.

Ми, звичайно, повністю усвідомлюємо, що УРС навіть у його більш або менш усталених на сьогодні формах існування жодним чином не можна прирівнювати за статусом до двох названих мов: він функціонує (переважно на рівні мовлення — усного й уже навіть писемного³) в периферійних зонах однієї з цих мов — української національної мови. Як один з кодів мовного спілкування українськомовного населення він залучається до зіставлення з певною умовністю (тим більше, що це робиться з урахуванням самооцінок мовної поведінки самих респондентів). По-перше, частотність використання в своєму мовленні трьох кодів респонденти визначають самостійно, а по-друге, межа між «суржиком» та його мовами-донорами ґрунтується також на суб'єктивних оцінках респондентів (тому надалі у протиставленні УРС і «української мови» під останньою слід розуміти «власне» українську мову).

Звичайно, «суржик» являє собою значною мірою варіабельний мовний код. Добре відомо, що в змішаному ідіомі або, у загальнішому сенсі, у змішаних мовних кодах стабілізаційне змішування (fusion) і спонтанне змішування накладаються в мовленні доти, доки в суспільстві використовуються донорські коди.

³ Неофіційну писемну форму використання «суржику» можна знайти насамперед в інтернеті [див., зокрема: Брага; Курохтіна]. У праці [Bilaniuk] також ідеться про використання суржику в інтернеті та про тенденції до його «дестигматизації» в цій сфері функціонування.

Для української і російської мов в Україні така ситуація лишається актуальною дотепер ⁴.

2. Попередні соціологічні дослідження щодо використання української та російської мов і «суржику». Однією з найбільш диференційованих картин мовної ситуації в Україні останніх років стало дослідження Київського міжнародного інституту соціології (далі скорочено КМІС) 2003 року [див.: Хмелько]. Це, здається, поки що єдине соціологічне дослідження в Україні, в якому, крім української та російської мов спілкування населення, брали до уваги також «суржик» (як третій код) ⁵. Однак у цьому соціологічному дослідженні і в нас застосовано різні (відмінні) підходи до констатації наявності УРС, як і двох інших мовних кодів, у респондентів і, відповідно, до визначення кількості (пропорції) його носіїв: у дослідженні [Хмелько] це робилося за визначенням інтерв'юєрів (до речі, це були соціологи, а не мовознавці), що проводили опитування в телефонному режимі, тоді як у нас — за самооцінкою самих респондентів у безпосередньому спілкуванні з ними. Це, звичайно, могло вплинути на розходження у визначенні такого співвідношення мовців, причому досить істотним чином. Вирішальним недоліком відзначеного соціологічного підходу є те, що багато жителів України є двокодовими, якщо не трикодовими: залежно від комунікативних потреб вони можуть переключатися між принаймні однією з двох названих мов і УРС [див.: Гентшель, Тараненко 2015; Hentschel, Zeller 2017]. Отже, дослідження [Хмелько], скоріше за все, реєструє лише тих носіїв «суржику», чий мовний репертуар обмежується цим одним кодом.



Карта 1. Розподіл трьох кодів відповідно до основного використання (згідно з дослідженням [Хмелько])

⁴ Щодо співвідношення у структурі УРС відносно стабільного ядра і варіабельних периферійних зон див.: [Тараненко 2013; Taranenko 2014; Menzel, Hentschel 2017; Hentschel 2018].

⁵ Акцентування на наявності в реальному мовному житті України такої трикодовості роблять останнім часом також у деяких мовознавчих дослідженнях. Наприклад: «Фактично є усі підстави визначити сучасну мовну ситуацію країни як тригліосію: з двома мовами (українська і російська) та мовним кодом (суржик), який має функціональне обмеження вжитку <...>» [Белей: 34–35, 136].

Як і багато інших «мовних карт» України, карта 1 ілюструє розподіл функціонування мов за регіонами: з одного боку, це Схід (у широкому розумінні), розділений на Східний Центр і Схід (у вузькому розумінні), де в повсякденному спілкуванні переважно використовується російська мова⁶, а з другого, — Захід (у широкому розумінні), розділений на Західний Центр⁷ і Захід (у вузькому розумінні). У цьому мовному розподілі неважко розпізнати історико-політичне розмежування країни: Схід (у широкому розумінні) значною мірою відповідає регіону України, який з другої половини XVII століття перейшов під владу Московської держави, а Південь — регіону, який з кінця XVIII століття ввійшов до складу Російської імперії, хоча раніше тут не було скільки-небудь помітних східнослов'янських поселень (ні українських, ні російських). Ще одним недоліком такої презентації є те, що ці два основні регіони з їхніми двома чи трьома субрегіонами були явно визначені як історико-політична схема з метою підрахунку середніх значень для цих субрегіонів.

Слід також згадати тут два інших нещодавніх опитування. Перше з них було проведене авторитетним Центром Разумкова в грудні 2015 року [див.: Ідентичність громадян...] ⁸, а друге — знову КМІС, цього разу в 2019 році [див.: Думки і погляди...]. Обидва вони охоплюють приблизно той самий комунікативний простір: ставили запитання «Якою мовою ви спілкуєтеся вдома?» або «Яку мову ви використовуєте для спілкування з близькими родичами?». Жодне з них не пропонує як варіант відповіді «суржик». Відповіді даються за шкалою «тільки українська — переважно українська — українська й російська (приблизно порівну) — переважно російська — тільки російська». Територіальна структура, представлена в цих дослідженнях, частково відрізняється від дослідження КМІС 2003 року. Дослідження Центру Разумкова передбачає «лише» один Центр, який охоплює Західний Центр (з дослідження КМІС 2003 року), а також Східний Центр, але без Дніпропетровської області, яка віднесена до Сходу. Крім того, Схід також включає Запорізьку область, яка в дослідженні КМІС 2003 року була частиною Півдня. В опитуванні КМІС 2019 року використовується така сама територіальна диференціація (принаймні номінально), хоча деякі деталі лишаються не зовсім зрозумілими: Дніпропетровська область вважається частиною Центру, як у дослідженні КМІС 2003 року (тобто Східним Центром) чи Сходу, як у дослідженні Центру Разумкова 2016 року? І куди віднесена Запорізька область — до Сходу або Півдня?⁹ Крім того, постає питання, чому в жодному з

⁶ Ми не приділяємо уваги так званім рідним мовам, як це роблять у багатьох інших оглядах. На запитання про рідну мову переважна більшість відповідей має символічне значення, яке досить слабо пов'язане з фактичним використанням мовцями того чи іншого коду. У таких країнах, як Україна і Білорусь, запитання про рідну мову постійно дає різні результати — на відміну від запитань про мову, яка в основному використовується [див. щодо України: Гентшель, Тараненко 2015; щодо Білорусі: Hentschel, Kittel : 110–117].

⁷ У [Хмелько] Західний Центральний регіон названо просто «Центром», і ми зберегли цей підхід на карті.

⁸ Інше дослідження Центру Разумкова 2016 року [див.: Консолідація...], проведене цією організацією в листопаді, ми тут не розглядаємо, бо воно менш диференційоване, ніж попередній огляд, а результати є подібними.

⁹ Це справді незвично для інтернет-публікацій КМІС. В іншому його опитуванні (2020 року), яке не стосується мови [див.: Кого українці хотіли б...],

досліджень Центральний регіон України не розділено на схід і захід від Дніпра, оскільки в дослідженні КМІС 2003 року результати для цих двох регіонів дуже сильно відрізнялися.

Останнє запитання стосується загальної проблеми всіх трьох згаданих тут аналізів: «великі регіони» представлені з різним ступенем прозорості щодо історичних і/або політичних мотивів їх розмежування. Відмінності в частотності використання в невеликих ареалах (наприклад, в окремих областях цих регіонів) належать до середніх значень. У нашому дослідженні ми використовуємо інший підхід. Звичайно, ми також застосовуємо політично зумовлене структурування. Ми узагальнюємо результати різних опитувань щодо адміністративної одиниці — області. Але їхні кордони мають тільки статус системи координат. Потім області групуються на основі частотності використання там відзначених мовних кодів. Це означає, що в дослідженні немає заздалегідь визначених великих регіонів — натомість кластери вказують на регіональне структурування на основі частотності використання мовного коду. Для того щоб прояснити відмінності й подібності між нашим підходом і підходом з традиційно визначеними великими регіонами, коротко представляємо результати досліджень Центру Разумкова 2016 року і КМІС 2019 року (див. рис. 1а і 1б)¹⁰.

Немає потреби і навіть було б неправильно обговорювати кількісні відмінності між цими двома опитуваннями. Можна припустити, що в тому, як проводилися опитування та як представлені дані, існують певні відмінності (частково не вказані). В обох підходах простежуються, однак, три тенденції. По-перше, Захід з його значенням у понад 90% для «тільки українська / переважно українська» є найбільш стабільним регіоном, який також чітко відрізняється від інших. На Півдні і Сході (а також на території Донбасу, яка не окупована Росією і яку ми не будемо розглядати в цьому дослідженні) частотність використання української мови як єдиного або основного засобу спілкування вдома або з родиною різко знижується: значення опустилися до близько 20% в опитуванні 2015 року і навіть нижче — в опитуванні 2019 року (на неокупованій частині Донбасу взагалі ніхто не заявляє, що розмовляє виключно або переважно українською. Центр країни посідає середнє становище між трьома останніми згаданими регіонами і Заходом. Його значення є більш наближеними до значень Заходу за опитуванням Центру Разумкова 2016 року, ніж за даними КМІС 2019 року.

розподіл областей за регіонами дуже схожий на той, що був у Центрі Разумкова 2016 року, за винятком Хмельницької області: в останньому вона віднесена до Центрального регіону, в першому — до Західного. Це означає, що КМІС 2003 року і КМІС 2019 року застосовують дуже різні структурування.

¹⁰ У дослідженні КМІС 2019 року [див.: Думки і погляди...] цифри для категорій «тільки українська» і «переважно українська» (аналогічно для російської мови) варіюються, у дослідженні Центру Разумкова 2016 року [див.: Ідентичність громадян...] — ні. Для того щоб результати у схемі були більш розбірливими, ми об'єднуємо категорії «тільки ...» і «переважно ...» в одну: т/п У (тільки або переважно українська) або т/п Р (тільки або переважно російська) vs. УіР (українська і російська), тобто приблизно збалансовано між цими двома категоріями. Категорія «н.і. — немає інформації» означає інші мови або відсутність відповіді.

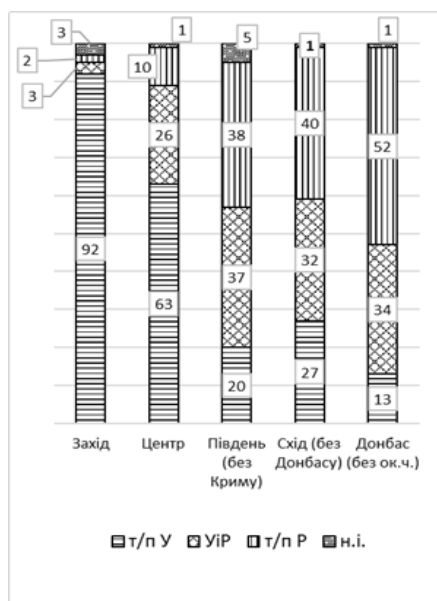


Рис. 1а. Розподіл української і російської мов (у %) відповідно до дослідження Центру Разумкова 2016 року

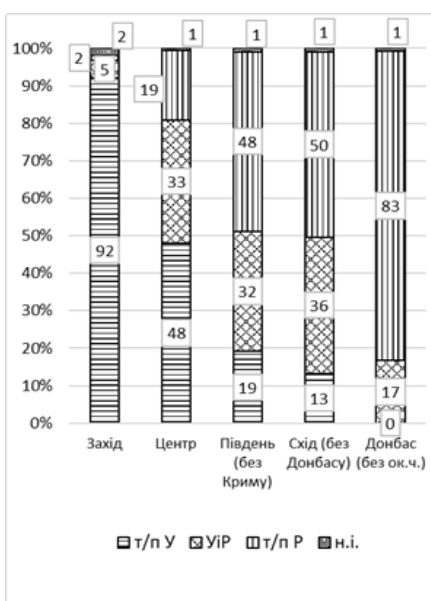


Рис. 1б. Розподіл української і російської мов (у %) відповідно до дослідження КМІС 2019 року

Що ж до російської мови як єдиної або основної мови повсякденного спілкування, то порівняння цих двох досліджень відображає дзеркальне зростання її використання із Заходу на Південь і Схід, що особливо помітно на неокупованій частині Донбасу, де понад вісім із десяти респондентів заявили, що розмовляють виключно або переважно російською мовою, а менше двох із десяти спілкуються російською та українською приблизно в однакових пропорціях.

Однак лишається головне запитання: який з варіантів (трьох або п'яти), представлених у двох опитуваннях, обрали респонденти, що розмовляють удома «суржиком», але з огляду на відповідну структуру опитування не змогли дати таку відповідь? Це стосується не лише людей, які звичайно спілкуються тільки «суржиком» і адекватно не володіють жодною з літературних (стандартних) мов¹¹, а й тих, хто використовує «суржик» на додаток до хоча б однієї з цих двох мов. Кількісний аналіз, як-от у дослідженні [Гентшель, Тараненко 2015 : 9], показав, що таких людей є досить багато і що значна частина їх має вищу освіту [пор. Hentschel, Zeller : 45]. Тому ще більше дивує той факт, що українські соціологічні опитування й надалі не включають «суржик».

Вражає те, що в обох відзначених опитуваннях за останні п'ять років (і особливо в останньому з них) дуже велика частина респондентів стверджує, що використовує українську й російську мови приблизно порівну вдома та в колі сім'ї. Добре відомо, що в незалежній Україні українці, які раніше спілкувалися російською, тепер хоча б частково звертаються до української мови в громадських місцях. Безумовно, є також сім'ї, які переходять у

¹¹ Деякі українські й білоруські мовознавці навіть використовували характеристику «недостатній» рівень володіння українською або білоруською мовами у своїх визначеннях «суржик» і «трасянки» [див. критику такого підходу: Hentschel 2017 : 22–23].

спілкуванні з української на російську і навпаки. З іншого боку, приватний і неформальний характер таких комбінацій приводить до того, що зміна та змішування коду трапляються досить часто. Як би там не було, респонденти, які в основному використовують суржик удома та в колі сім'ї, змушені обирати інші варіанти, доступні в двох останніх опитуваннях. Залишається неясним, як саме вони зробили свій вибір: вони могли керуватися мотивами політкоректності, вони могли також оцінювати свій «суржик» як ближчий до української або російської чи як збалансований між цими двома мовами [див.: Hentschel, Zeller : 51–53]; вони також могли зробити вибір навмання.

3. Аналіз.

3.1. Дані. Дані для аналізу було зібрано в межах двох проєктів, зазначених у примітці 1. Обидва дослідження були зосереджені на українському «суржику», а так звані закриті інтерв'ю щодо соціолінгвістичного «тла» респондентів проводилися Центром соціальних та маркетингових досліджень «SOCIS» (м. Київ)¹². Після цього інтерв'юєри заповнювали анкети. Перший проєкт — збір даних на початку 2014 року — був зосереджений на одинадцяти областях Центральної України, включаючи одну, що межує із Заходом (Хмельницька), і одну, що межує зі Сходом (Харківська). Як уже відзначалося вище, ці дві області зараховували в різних територіальних структуруваннях відповідно до Заходу і до Сходу, а не до Центру (див. карти нижче). За первісним планом, дві східні області мали ввійти до регіону Донбасу (це Луганська й Донецька області), але вже 2014 року це стало неможливим через розгортання збройної агресії з боку Російської Федерації. Однак у Центрі, за даними соціологічних опитувань, респонденти частіше, ніж в інших регіонах, характеризують своє мовлення як «суржик» [Гентшель, Тараненко 2015] (див. нижче).

У найбільших містах центральних регіонів російська мова поки що переважає в повсякденному житті, хоча в офіційних сферах дедалі більше поширюється українська мова. У сільській місцевості Центрального регіону більш поширена діалектна форма української мови (з деякими домішками російської). Наскільки сильним тут є вплив російської мови, не зовсім ясно, оскільки традиційна діалектологія звичайно ігнорує впливи «ззовні». Через це, а також через обмежений бюджет проєкту було вирішено не збирати дані в найбільших містах і населених пунктах сільського типу Центрального регіону. Таким чином, дослідження проводилося в містах будь-якого розміру (крім вищезазначених найбільших міст, які вважаються переважно російськомовними), а також у населених пунктах зі статусом «селище міського типу», оскільки вважалося, що там найімовірніше зустріти суржик. У кожній з одинадцяти областей було опитано від 100 до 150 респондентів, усього 1400 осіб.

У межах другого проєкту в Чорноморському регіоні в березні 2020 року було проведено дуже подібне опитування з частково ідентичною анкетною. Однак цей проєкт мав дещо іншу мету. Якщо проєкт у Центральній Україні досліджував «класичний суржик» на основі української мови, чорноморський проєкт на Півдні, що вважається переважно російськомовним, розглядав обидва типи змішаного мовлення — як українсько-російський, так і російсько-український «суржики». Як і в Центрі, до аналізу не було залучено переважно російськомовні міста-мільйонники: тут, на Півдні, це стосувалося

¹² На основі цих закритих інтерв'ю було відібрано респондентів з відповідним профілем використання трьох мовних кодів, яких було запрошено на відкрите інтерв'ю, що було або буде проаналізовано в якісному аспекті. Матеріал, зібраний у відкритих інтерв'ю, також становить основу для корпусно-лінгвістичного аналізу.

лише Одеси. Оскільки області Чорноморського регіону не вважаються ареалами, де розмовляли старими автохтонними східнослов'янськими діалектами (за винятком деяких північних і північно-західних регіонів Одеської і Миколаївської областей), до опитування були включені також села. Дослідження проводилося в трьох областях Чорноморського регіону (див. нижче)¹³. У кожній області було опитано 400 випадково обраних респондентів¹⁴ віком від 18 до 80 років з рівномірним розподілом за віком. Співвідношення чоловіків і жінок є приблизно рівним. Присутні всі рівні освіти, характерні для населення в цілому.

Дослідження використання мови в неодномовних суспільствах (а отже, і це опитування) не «вимірюють» її конкретного, або ж фактичного, використання. Така «об'єктивна» кількісна оцінка надзвичайно варіабельних змішаних форм мовлення на зразок «суржику», що походять від двох структурно подібних і тісно пов'язаних мов-донорів, була б теоретично складною й методологічно трудомісткою. Опитування ґрунтуються на «суб'єктивних» оцінках і твердженнях респондентів. У праці [Гентшель, Тараненко 2015 : 18–23] обговорюється ця «суб'єктивність» мовної ситуації в Україні. За підсумками цього обговорення можна припустити, що в цілому оцінки респондентів загалом досить добре відповідають реальному стану використання мови і що неправильні оцінки в будь-який бік взаємно компенсуються. Звичайно, може виникнути додаткова проблема, наприклад, у зв'язку з емоційно вмотивованими заявами щодо використання мовних кодів, які можуть навіть суперечити самооцінці респондента. В умовах напруження, яке нині спостерігається в значній частині населення України у ставленні до Росії, іноді рекомендується виявляти особливу обережність при інтерпретації кількісних спостережень (у подальшому аналізі звертається увага на такі окремі проблеми інтерпретації). Однак, як правило, ми просто говоримо про «використання», усвідомлюючи, що маються на увазі твердження про використання трьох кодів.

3.2. Поширеність окремих кодів та їх комбінацій.

3.2.1. Національність респондентів і вибір мови інтерв'ю. Національний склад респондентів у Центральному і Чорноморському регіонах є дуже подібним. У трьох областях Чорноморського регіону 90% респондентів заявили, що мають українську національність, у Центрі — 96%; відповідно 10% і 4% опитаних указали свою національність як російську. Респондентів інших національностей до опитування не було включено. Аналогічна картина спостерігається в батьків і членів подружжя респондентів. У Чорноморському регіоні батьки 80% респондентів були українцями, а 13% — росіянами; в Центральному регіоні значення розподілилися так: 88% — українці і трохи менше 10% — росіяни. Решта батьків були іншої національності (трохи частіше на Півдні, ніж у Центрі) або відповіді були відсутні. Українцями були до 88% членів подружжя респондентів на Півдні і до 95% — у Центрі, а росіянами — до 8% і 4% відповідно; у дуже невеликого залишку члени подружжя були вказані як представники інших національностей.

¹³ Запорізьку область, яка розташована далі на схід України, не було включено до опитування, оскільки під час планування проєкту існувала загроза того, що туди можуть перекинутися воєнні дії. Було визнано недоцільним збирати мовні дані (сімейні бесіди, інтерв'ю) як доповнення до початкового опитування на території, близькій до зони бойових дій.

¹⁴ Обидві анкети (українсько- та російськомовну) можна знайти на вебсайті Ольденбурзького університету: <https://uol.de/slavistik/forschung/sprachwissenschaft/hybridisierung-von-zwei-seiten>.

Незважаючи на те що склад обох груп респондентів дуже подібний у плані національного співвідношення українців і росіян, спостерігалися значні відмінності у виборі мови інтерв'ю¹⁵. У Центральному регіоні 83% респондентів обрали українську мову як мову інтерв'ю і, відповідно, 17% обрали російську мову; у чорноморських областях лише 46% обрали українську мову, а 54% віддали перевагу російській. Однак у двох областях Центрального регіону респонденти також обрали російську як мову інтерв'ю: 65% у Дніпропетровській і ще чіткіше, 83%, — у Харківській¹⁶. У Чорноморському регіоні спостерігається яскраво виражена відповідність між національністю і вибором мови інтерв'ю: якщо респонденти української національності обирали для інтерв'ю українську та російську мови в рівних пропорціях, 92% зі 122 респондентів російської національності вирішили провести інтерв'ю російською. У Центральному регіоні «відношення маркованості» були протилежними: 85 респондентів української національності обрали українську як мову інтерв'ю (три чверті тих, хто обрав російську мову, були з Дніпропетровської та Харківської областей). І навпаки, невелика кількість респондентів російської національності (570) обирали український і російський питальники в рівних пропорціях. Вибір мови для інтерв'ю дає вже зрозуміти, що російська мова має більше значення для респондентів на Півдні і (частково) на Сході¹⁷, навіть якщо майже половина їх обрала питальник українською мовою, тоді як у Харківській області в 2014 році цей показник становив усього 20%.

3.2.2. Первинні коди. Під первинним кодом тут мається на увазі код, який респонденти обрали у відповідь на запитання: «Який із цих кодів ви найчастіше використовуєте в повсякденному житті?». Відповідь можна було дати за п'ятибальною шкалою, наведеною на рис. 2.

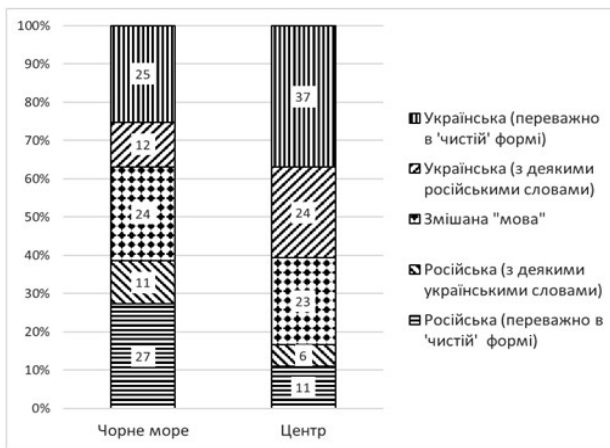


Рис. 2. Первинні коди за п'ятибальною шкалою

¹⁵ Окремі запитання ставив інтерв'юєр, який уводив відповіді до української або російської анкети. При цьому, крім використання респондентами мовних кодів (наприклад, щодо мовних біографій, ставлень до мови / мов тощо), в анкеті було багато інших запитань, які будуть проаналізовані в наступних статтях.

¹⁶ У всіх інших центральних областях, включених до опитування 2014 року, як мінімум у 80% випадків інтерв'ю проводилися українською мовою.

¹⁷ Звичайно, у дослідженні [Гентшель, Тараненко 2015] також підтверджується поступове зменшення значення української мови із заходу на схід для Центральної України, але також указується на подібне явище в напрямі від центру до периферії (див. нижче).

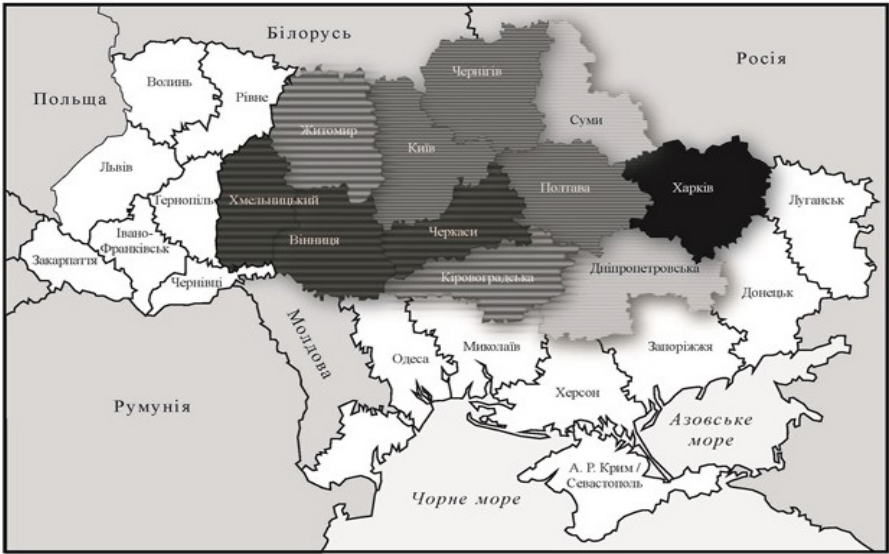
Перш ніж представити результати, потрібно сказати декілька слів про шкалу. Щоб уникнути появи «інфляційного» розширення «суржику», було обрано п'ятибальну шкалу, яка розрізняє два випадки у змішуванні елементів української і російської мов, а саме — «з деякими словами» іншої мови або без них. Наявність поняття «суржику» в нашому огляді вже істотно ширша, ніж в інших подібних опитуваннях, що, на нашу думку, пов'язане з тим, що ми уникаємо стигматизованої назви «суржик» у самій анкеті, натомість називаючи його нейтральніше як «змішана мова»¹⁸. Цілком імовірно, що багато сучасників не хочуть визнавати, що використовують саме «суржик», з огляду на поширене ставлення до цього мовного явища як до особливостей мовлення недостатньо освічених людей. Використання в опитуваннях термінів з негативною конотацією є більш ніж сумнівною практикою. П'ятибальну шкалу було обрано тут для того, щоб уникнути критичної оцінки українського мовлення з різними спорадичними (*ad hoc*) запозиченнями з російської мови або російського мовлення з такими ж запозиченнями з української мови і, таким чином, підвищити значення для «суржику». Якщо порівняти «переконаних» носіїв нібито «чистих» національних мов з мовцями, «невпевненими» щодо чистоти їхньої мови (української і російської мов в обох регіонах), то ми виявимо, що майже чверть респондентів у Чорноморському регіоні і трохи більше однієї чверті в Центральному регіоні заявили, що, говорячи однією з цих визнаних мов, вони використовують деякі слова з іншої. Ми не спинятимемося тут докладно на цьому питанні, оскільки буде проведено додаткове дослідження реальних мовленнєвих даних для подальшого диференціювання (це планується на пізнішому етапі проєкту). Надалі ми обмежимося трирівневою диференціацією, похідною від вищевказаної п'ятибальної шкали: українська мова, російська мова і «суржик». Останній охоплює лише відносно збалансоване або сильно змішане мовлення. Частка респондентів в обох регіонах, які вважають «суржик» своїм основним засобом спілкування, становить від однієї п'ятої до однієї чверті.

На рис. 2 вже показано (що не дивно), що російська мова має більше значення в Чорноморському регіоні і що українська мова є більш поширеною в Центральному регіоні¹⁹, тоді як частка використання «суржику» є приблизно однаковою в обох регіонах. Однак, як уже показано у праці [Гентшель, Тараненко 2015], спостерігається значне варіювання в п'яти субрегіонах Центрального регіону (це проілюстровано на карті 2)²⁰.

¹⁸ Звичайно, використання терміна «змішана мова» в анкеті не означає, що ми розглядаємо «суржик» як якусь особливу мову (див. про це у Вступі). Ледве чи хто-небудь із респондентів має які-небудь лінгвістичні знання, а назва мова, як відомо, охоплює в повсякденному вжитку, крім власне поняття «мова» в розумінні мовознавців, також «мовлення», «територіальний діалект», «соціальний діалект» тощо (не мало б сенсу називати нефахівцям такі терміни). Так само використання назви «слово» в анкеті є не власне термінологічним, оскільки в більшості випадків тут доречнішим був би термін «словоформа».

¹⁹ У цій статті «Центром» завжди позначається регіон, що був об'єктом дослідження в першому проєкті (див. примітку 1) та описаний вище.

²⁰ Розподіл на п'ять регіонів базується на частотності використання цих трьох кодів. Причини такої диференціації докладніше описано нижче.



Карта 2. Визначення поширеності трьох мовних кодів
[за: Гентшель, Тараненко 2015]

- А – Хмельницька, Вінницька, Черкаська області
В – Київська, Чернігівська, Полтавська області
С – Житомирська, Кіровоградська області
D – Сумська, Дніпропетровська області
Е – Харківська область

Ці п'ять субрегіонів Центрального регіону можна порівняти з трьома областями, що виходять до Чорного моря, за трирівневою шкалою. Ліва частина рисунка 3а не включає села на Півдні (як згадувалося раніше, в селах Центрального регіону дослідження не проводилося), а рисунок 3б, на якому представлено лише Південь, села включає.

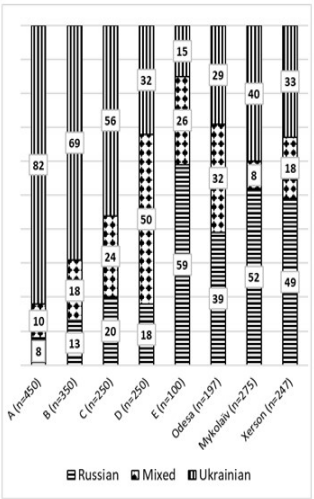


Рис. 3а. Первинні коди в трьох чорноморських областях і п'яти центральних регіонах у %

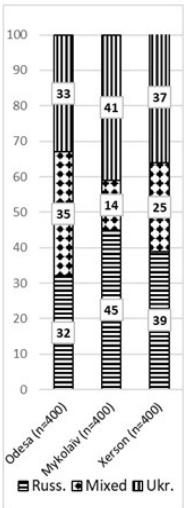


Рис. 3б. Первинні коди в трьох чорноморських областях (включно із селами) у %

Між трьома областями Чорноморського регіону немає значної розбіжності: приблизно одна третина респондентів назвали своєю основною мовою українську, від 40% до 50% – російську. Російська мова переважає, але не такою мірою, як це було показано у дослідженні КМІС 2003 року [див.: Хмелько]²¹. Пізніші дослідження Центру Разумкова 2016 року [див.: Ідентичність громадян...] і КМІС 2019 року [див.: Думки і погляди...] також указують на розширення наявності української мови на Півдні²². Що ж до «суржику», то між цими трьома областями існують значні відмінності: своїм основним кодом його називає лише кожен дванадцятий респондент у Миколаївській області, кожен третій — в Одеській і майже кожен п'ятий — у Херсонській. У нашому опитуванні «суржик» згадується частіше, що пояснюється описаною вище схемою опитування.

Загалом три обстежені області Чорноморського регіону є найбільш порівнянними з Центральним регіоном Е, який включає лише Харківську область. Однак в останньому регіоні українська мова відіграє набагато меншу роль як первинний код і є ще менш поширеною, ніж «суржик». Одеська область подібна до Центрального регіону D (Сумська, Дніпропетровська області), принаймні приблизно з погляду присутності «суржику», який, однак, є поширенішим у регіоні D. Значення для української мови є майже тими самими. Російська мова в Дніпропетровській області є набагато менш поширеною не лише порівняно з Одеською областю, а й з Харківською. Це свідчить про сумнівність віднесення Дніпропетровської області, як і Харківської, до Сходу. За даними [Гентшель, Тараненко 2015], регіони А, В і С Західного Центру відрізняються від усіх інших, оскільки в них українську мову назвала своїм основним кодом переважна більшість респондентів. Відмінності тут мають градуальний характер і виражаються в зниженні важливості української мови та відповідному підвищенні значення російської мови і «суржику». Хоча ці результати можна порівняти з результатами КМІС 2003 року [Хмелько], представлення в останньому дослідженні «Західного Центру» як території зі специфічною комбінацією використання кодів видається вельми сумнівним. Це є ще більш актуальним для ширшої одиниці «Центр» у двох вищезгаданих дослідженнях — Центру Разумкова 2016 року [див.: Ідентичність громадян...] і КМІС 2019 року [див.: Думки і погляди...], які також включають регіон D (див. рис. 3а).

Зрештою, слід зазначити, що для Півдня частка «суржику» збільшується в разі включення даних опитування по селах, як показано на рис. 3б. Частка ж російської мови знижується, що підкреслює її важливість у містах.

3.3. Трикодовість — співіснування і поєднання трьох кодів, типи мовців і їх розподіл за регіонами. Фокусування уваги на первинному мовному коді як на основному в повсякденному спілкуванні, як у розділі 3 та в наведених вище українських опитуваннях (що обмежуються домашньо-приватною сферою), може дати лише обмежене уявлення про мову, яку респонденти використовують у побуті в місцях свого проживання. Це стосується також нещодавніх досліджень, проведених українськими соціологами, у яких

²¹ Особливо низька частка російської мови в Одеській області зумовлена тим, що наш проєкт не враховував найбільше місто регіону — Одесу (як переважно російськомовне).

²² Імовірно розширення функціонування української мови в південних і східних регіонах за останні два-три десятиліття не є предметом цього дослідження, але буде розглянуто в наступному, також у рамках цього проєкту.

два коди, визнані стандартними (літературними) мовами, вважаються еквівалентними. У праці [Гентшель, Тараненко 2015 : 7–9] запропоновано інший підхід щодо центральних регіонів України, який зосереджений більше на співіснуванні мовних кодів у різних ареалах і менше — на сконструйованій бінарній комплементарності. Тут показано, що більшість мовців у Центральній Україні використовує щонайменше два коди в повсякденному житті, а менша кількість — усі три. Крім того, тут визначено ступінь частотності їх використання, зокрема для «суржику». Причому в цій праці, як і в пропонованій тут, огляд не обмежується домашньо-приватною сферою, а охоплює повсякденне життя в цілому, зокрема й робочу сферу.

Карта 2 (див. вище) ілюструє регіональні подібності і відмінності, виявлені при аналізі, описаному в [Гентшель, Тараненко 2015]. Розподіл на п'ять змодельованих ареалів, таким чином, не ґрунтується на зовнішньому (політично та історично вмотивованому) поділі України, а, як згадувалося раніше, використовує області як «систему координат»²³. Розподіл на п'ять регіонів є скоріше результатом кластерного аналізу, базованого на частотності визначення кодів як таких, що використовуються постійно або часто. На відміну від традиційних тверджень, у статті [Гентшель, Тараненко 2015] розглянуто всі три коди при моделюванні карти використання мови в повсякденному житті в Центрі України, яка демонструє однокодовість і трикодівість з приблизно таким самим значенням, які по-різному виражаються в областях і регіонах.

При розгляді спочатку «співіснування» кодів (лишаючи осторонь комбінації, що їх використовують окремі мовці) картина співіснування трьох мовних кодів на узбережжі Чорного моря дуже схожа на описану в [Гентшель, Тараненко 2015 : 8] для Центрального регіону²⁴. Єдина відмінність полягає в тому, що на Півдні значення для української і російської мов у цілому подібні, тоді як для російської мови в Центральному регіоні вони були нижчими (див. рис. 4).

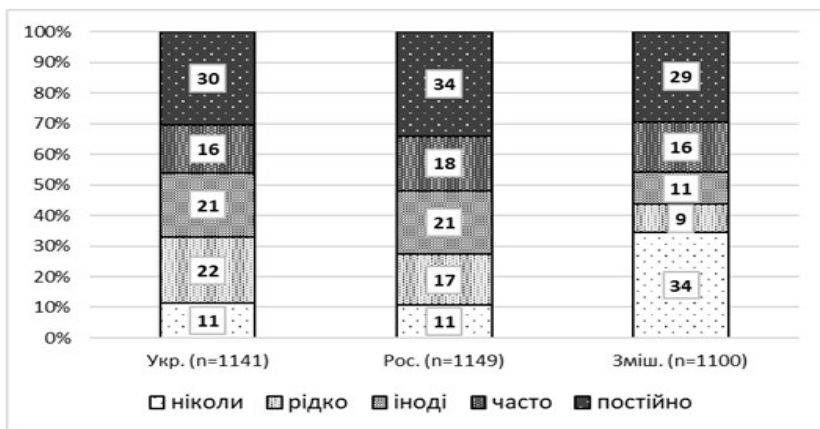


Рис. 4. Співіснування трьох кодів у Чорноморському регіоні та інтенсивність їх використання

²³ Звичайно, в межах областей існують певні відмінності, але для того, щоб зафіксувати їх усі, знадобилися б значні емпіричні зусилля, незалежно від мети їх виявлення.

²⁴ У цьому початковому огляді відношень немає необхідності повторювати ці результати. Подібності і відмінності стануть очевидні нижче.

Приблизно половина респондентів використовує всі три коди часто, а приблизно третина — завжди. За їхніми словами, російська мова використовується частіше, ніж українська або «суржик». Однак найбільше вражає те, що лише кожен десятий респондент говорить, що ніколи не використовує кожну з цих двох мов, а приблизно третина твердить, що ніколи не використовує «суржик». Різну кількість випадків (n) у трьох стовпцях можна пояснити можливістю відповіді «важко сказати». Ці «негативні» відповіді було виключено з розрахунків частотності, показаних на рис. 4 (докладнішу інформацію див. нижче). Респонденти обирали цей варіант відповіді удвічі частіше для «суржику» (рівно 100 разів), ніж для двох визнаних мов. Таку поведінку можна пояснити рефлекторністю, спричиненою свосередним соромом, пов'язаним з використанням часто висміюваного змішаного мовлення.

Із цифр на рис. 4 чітко видно, що мовці часто комбінують коди, тобто обирають їх залежно від ситуації, співрозмовника тощо ²⁵. Лишається ще простежити, які комбінації і в яких пропорціях використовуються, тобто як мовці «комбінують» коди, або, інакше кажучи, наскільки їхнє спілкування багатомовне чи, точніше, багатокодове.

Далі представлено більш диференційовану модель мовної дво- і трикодовості в Україні, ніж модель, представлена в дослідженні [Гентшель, Тараненко 2015]. Вона базується на використанні трьох кодів кожним окремим мовцем, а не на загальній поширеності мовних кодів у регіонах. Крім того, така модель належним чином ураховує частотність використання кодів, рідшу за «часто». Респонденти можуть не лише «переоцінювати» частотність використання коду, але й «недооцінювати» її, наприклад з міркувань «політкоректності». Таким чином, модель має враховувати якомога більше даних. Тому увага приділяється більше різним формам індивідуальних комбінацій, або типам мовців, і менше — співіснуванню (частому або постійному використанню) кодів в областях і регіонах. На другому етапі частотність цих типів мовців в окремих областях буде використано для того, щоб зробити начерк карти трикодовості, або, радше, карти трикодовості мовців у Центральній Україні та в Чорноморському регіоні.

За допомогою кластерного аналізу (з використанням SPSS у версії 26: англ. Statistical Package for the Social Sciences — Статистичний пакет для соціальних наук) спочатку досліджується, які комбінації багатокодовості можна використовувати як основу для групування мовців (кластерів), коли розглядаються всі п'ять рівнів частотності використання коду: від «завжди» до «ніколи». Кластери — це підкласи респондентів, які з погляду частотності використання трьох кодів є максимально однорідними, але максимально відрізняються від усіх інших підкласів. З огляду на велику кількість респондентів (усього понад 2 тис.) доречним є використання аналізу кластерних центрів із застосуванням методу k -середніх (Quickcluster). У цьому методі кількість досліджуваних кластерів (але, звичайно, не тип) визначається апріорі [пор. Bühl : 635–696]. Кількість кластерів можна визначити з, наприклад, квалітативних міркувань. Ми вибрали сім, ґрунтуючись на логічних можливостях використання трьох кодів: три однокодові, три двокодові та один трикодовий кластери. Тест виявить, які комбінації є найпомітнішими, на основі кожної з п'ятирівневих частотностей для кожного

²⁵ Див. [Hentschel, Zeller 2017] щодо Центрального регіону. У майбутньому дослідженні будуть представлені результати для Чорноморського регіону.

з трьох кодів. Кластери, сконструйовані таким чином, репрезентують певні типи мовців, тобто типи мовців з особливим профілем щодо ступеня використання ними трьох кодів. Вони не обов’язково мають відповідати семи логічним можливостям; вони також можуть бути градаціями однієї частини із семи комбінацій: наприклад, якщо одна або більше логічно можливих комбінацій не підтверджені емпірично.

Кластерний аналіз типів мовців спочатку було проведено окремо для Чорноморського регіону і для Центральної України, тому що, як уже згадувалося раніше, вибір респондентів для кожного з проєктів трохи відрізнявся: у Центральній Україні було виключено села.

Постає така картина (див. табл. 1) — спочатку окремо для Чорноморського регіону (ліворуч) і Центральної України (праворуч):

Таблиця 1

Кластери мовців на основі частотності використання трьох кодів: порівняння двох великих регіонів									
Як часто укр.?	Як часто рос.?	Як часто змішане?	Тип	%	Як часто укр.?	Як часто рос.?	Як часто змішане?	Тип	%
4.92	2.01	1.60	U	23	4.90	1.50	1.65	U	19
1.99	4.94	1.45	R	26	2.17	4.82	1.80	R	12
4.61	4.16	1.31	UR	5	4.66	3.20	2.41	UR	13
1.57	1.54	4.96	S	4	2.43	2.36	4.75	S	11
3.79	1.74	4.72	SU	7	4.51	2.22	4.09	SU	30
1.79	4.06	4.40	SR	10	2.49	4.24	4.43	SR	6
3.67	3.69	4.44	SUR	24	4.32	4.15	4.13	SUR	10
				N=					N=
				1,030	Центр				1,400

У цих двох окремих аналізах залежними змінними були п’ятирівневі частотності використання трьох кодів, надані респондентами (які становлять незалежну змінну). Якщо код ніколи не використовується, він отримує значення 1, якщо завжди, — 5. Стовпець із числами під запитанням «Як часто...?» містить середні значення частотності для цього коду. Комбінація всіх трьох середніх частотностей у кожному рядку таблиці становить так званий центр кластеру. У першій (ліворуч) частині таблиці для Чорноморського регіону значення в лівій верхній клітинці кластеру становить 4,92. Це означає, що респонденти в цьому кластері майже весь час використовують українську мову. У клітинці праворуч указано значення середньої частотності використання російської мови: 2,01, тобто «рідко». У наступній клітинці праворуч — значення для «суржик», яке становить 1,6. Воно навіть нижче, ніж для російської мови, і близьке до невикористання (інакше кажучи, дуже багато респондентів у цьому кластері справді не використовують його). Ми класифікуємо цей кластер, або, скоріше, тип мовця, як практично «однокодових» українських мовців, хоча інколи, але рідко вони звертаються до російської мови і ще рідше — до «суржик».

У цілому кластери відповідають семи можливим логічним умовам використання кодів часто, якщо не постійно. Кластер U-R у Центральній Україні — це більш очевидний виняток. Тут середнє значення для російської

мови майже на 1,5 бала менше, ніж для української, і є наближеним до «суржику». Взаємовідношення між українською та російською мовами більш збалансоване в типі U-R у Чорноморському регіоні. Це означає, що кластер U-R у Центральній частині України не так чітко відрізняється від типу S-R, як у Чорноморському регіоні.

Якщо інтерпретувати три середні значення в кожному рядку таким чином, то сім згаданих вище комбінацій можна підтвердити як типи мовців: (I) три в основному однокодові, (II) три в основному двокодові і (III) один в основному трикодовий типи. Тип визначається за найвищими середніми значеннями; в кожному із семи кластерів хоча б одне середнє значення більше від 4 (ці клітинки виділено темно-сірим). Якщо кластер є скоріше багатокодовим, два або три середні значення є більшими від 4 або принаймні близькими до цього значення (такі клітинки виділено світло-сірим).

Слід зазначити, що були розглянуті не всі респонденти з Чорноморського регіону: з 1200 респондентів було включено ті 1030, які прямо відповіли на всі три запитання про частотність використання ними трьох кодів, навіть якщо вони заявили, що ніколи не використовують один або більше кодів. Якщо респондент не відповідав на одне з трьох запитань, його було виключено з цього аналізу. Неможливо автоматично прирівняти відсутність відповіді до невикористання коду. Відсутність відповіді може бути мотивована тим, що респондент не хотів казати, що використовує певний код з будь-якою частотністю ²⁶.

Для того, щоб краще порівняти дані щодо Чорноморського регіону з даними щодо Центральної України в подальшому аналізі, необхідно виключити певну підгрупу респондентів з областей Чорноморського регіону. Як уже згадувалося раніше, дані в селах Центрального регіону не збиралися. Тому в подальшому аналізі враховувалися лише респонденти з населених пунктів, у яких проживає не менше людей, ніж у найменшому поселенні, включеному до Центрального регіону. Порогове значення — 4 тис. жителів. Це дозволило скоротити кількість респондентів до 2029 (Чорноморський регіон: 629, Центральний регіон: 1400) і наблизити їхню кількість у кожній області Чорноморського регіону до кількості респондентів в областях Центральної України.

Було проведено новий кластерний аналіз, але цього разу без відмінностей між Чорноморським і Центральним регіонами (див. табл. 2). Результати дуже подібні до результатів двох незалежних моделей для Чорноморського регіону з трохи відмінними середніми значеннями для центрів кластерів. Що ще важливіше, аналіз виявив ті самі типи мовців.

²⁶ Ми можемо тільки припускати, чому відповіді на ці три запитання відсутні тільки в Чорноморському регіоні, але не в Центральному. Дані щодо Центрального регіону було зібрано на початку 2014 року, коли відносини між Україною і Росією були досить напруженими, але ні трагічні події на Майдані незалежності в Києві, ні падіння уряду В. Януковича, ні окупація Криму Росією, ні воєнні дії на сході України ще не стали нагальними проблемами. Як уже згадувалося, дані щодо Чорноморського регіону було зібрано на початку 2020 року.

Таблиця 2

Кластери мовців на основі частотності використання для трьох кодів: об'єднані дані з двох регіонів					
Як часто українська?	Як часто російська?	Як часто змішане мовлення?	Тип	N	%
4.86	1.62	1.89	U	437	21.5
2.00	4.88	1.26	R	298	14.7
4.70	3.42	1.49	UR	154	7.6
2.41	2.33	4.80	S	213	10.5
4.43	1.82	4.44	SU	225	11.1
2.29	4.46	3.96	SR	241	11.9
4.43	3.52	3.93	SUR	461	22.7
загалом:				2,029	100,0

У цілому результати можна схарактеризувати таким чином: однокодові мовці, орієнтовані на українську або російську мови (U і R), становлять «тільки» цілу третину респондентів, причому тип мовців U в півтора рази більше поширений, ніж тип R. У середньому носії цих двох типів рідко, якщо взагалі коли-небудь, використовують інший відповідний «визнаний» мовний код або ж «суржик». Мовці, орієнтовані на обидві мови (тип U-R), які використовують «суржик» у найкращому разі периферійно, трапляються відносно рідко й становлять лише 8% від усіх респондентів. Більшість респондентів типу U-R віддають дещо більшу перевагу українській мові. Ми можемо інтерпретувати ці три типи як один блок з пуристичною орієнтацією, що характеризується вираженням «утриманням» від «суржику». В цілому такі мовці-«пуристи», орієнтовані на національні мови, становлять приблизно 42%.

Це означає, що, навпаки, мовці, для яких «суржик» відіграє відносно важливу роль, становлять більшість. Типи S, S-U, S-R і S-U-R разом становлять 58%, навіть незважаючи на те, що, як зазначено вище, ставлення до «суржику» з боку значної частини еліти країни є, м'яко кажучи, «нешанобливим». Ось чому є підстави вважати, що деякі респонденти применшували його використання або вважали за краще не згадувати про це. Тому можна цілком припустити, що деякі респонденти мінімізували або опустили своє використання «суржику», а це означає, що він навіть більш поширений, ніж представлено тут. З цих чотирьох типів мовців S-U-R є найпоширенішим, становлячи (подібно до однокодового типу U) більше п'ятої частини всіх респондентів. Інші три типи з літерою «S» є приблизно однаковими — на кожний припадає ціла десята частина респондентів, включно з однокодовим типом S. В опитуваннях, які не включають «суржик» до категорії відповідей, респонденти однокодового типу S, а також мовці типу S-U-R (усього одна третина респондентів) мали б давати неточні відповіді на запитання про мову, якою вони спілкуються вдома або в сімейному колі. У цій ситуації мовці типу S-U або S-R могли обрати українську чи російську мови, але ці мови, як правило, використовуються в публічному або офіційному середовищах і менше — вдома.

Якщо в такий спосіб додати мовців, що подавали себе як носіїв двох названих національних мов (незалежно від того, однокодові вони чи

багатокодові), вимальовується аналогічна картина: приблизно 63% з них розмовляють українською і 59% — російською мовами. Отже, всі три мовні коди присутні в українському суспільстві рівною мірою, хоча «суржик» є, звичайно, ідіомом з явно нижчим соціальним статусом, що, однак, не заважає йому використовуватися з такою ж частотою, як українська й російська мови. І останнє: в кластерах, де українська і російська мови об'єднані як широко використовувані коди (середнє значення > 3), саме українська мова має вищі значення.

Ситуація, яку ми щойно описали, звичайно, є узагальненням для всіх охоплених областей. У праці [Гентшель, Тараненко 2015] уже відзначено істотні відмінності між областями Центральної України в раніше згаданій простішій моделі, але при цьому можна було б також кластеризувати області. Наступним кроком, таким чином, є отримання результатів кластеризації окремих мовців і їх диференціація за областями. По-перше, ми описуємо ситуацію в 14 розглянутих областях. На рис. 5 представлено частку семи типів мовців у кожній області. Області представлені в приблизному порядку, починаючи з високої частки української мови ліворуч. Ця частка зменшується зліва направо на користь збільшення частки «суржику» в чотирьох згаданих комбінаціях (S, S-U, S-R, S-U-R), представленої на сірому фоні в стовпцях. У свою чергу, ця частка зменшується зліва направо на користь збільшення частки російської мови. Лише українська і російська мови або їхня комбінація без «суржику» представлені на білому тлі у стовпцях. З погляду географії можна спостерігати приблизне збільшення присутності «суржику» із заходу на схід, де пік припадає на Сумську та Дніпропетровську області. Присутність «суржику» знову зменшується на схід і південь (див. рис. 5).

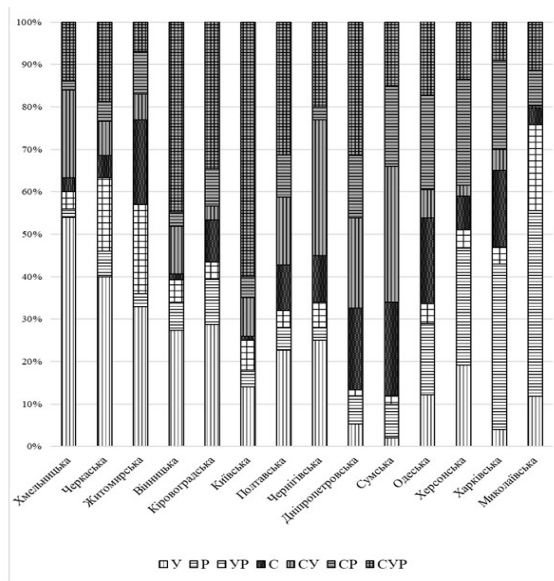


Рис. 5. Одно- і трикодовість в областях
(поля стовпчиків на білому тлі — у нижній частині — означають поєднання кодів без «суржику», на сірому тлі — з участю «суржику»)

При цьому слід навести лише кілька вибірових спостережень. Примітно, що в десяти областях «суржик» відіграє важливу роль, оскільки ним досить часто спілкується половина респондентів або навіть більше

(звичайно в поєднанні принаймні з одним із двох інших кодів). Особливо це помітно в Дніпропетровській і Сумській областях, де лише кожен десятий респондент не часто розмовляє «суржилом» у повсякденному житті. Лише в Хмельницькій, Черкаській і Житомирській областях (це більш центральні області на заході досліджуваної території), а також у Миколаївській (середня область у Чорноморському регіоні) «суржик» відходить на задній план, оскільки ним спілкується менше половини респондентів. Миколаївська область вирізняється несподівано високою часткою респондентів, які, за їхніми словами, не використовують «суржик». Комбінація S-U практично відсутня (частка становить менше одного відсотка), що є найнижчим показником «суржику» на всій досліджуваній території. У двох сусідніх областях на узбережжі Чорного моря частка «суржику» є значно вищою. Харківська область вирізняється відносно високою часткою однокодових російськомовних респондентів. Разом із Харківською, три області Чорноморського регіону традиційно вважаються переважно російськомовними регіонами України (див. карту 1 у: [Хмелько]). Це твердження не безпідставне, але його не можна назвати цілком правильним. Два інші коди в різних комбінаціях також поширені значною мірою. Частка однокодових російськомовних респондентів у Харківській і Миколаївській областях досягає максимуму «всього» на рівні 40%. Ці цифри, безумовно, можна порівняти з показниками Центру Разумкова 2016 року або також КМІС 2019 року. Як ці дослідження, так і два наші демонструють подібні пропорції однокодових українськомовних респондентів у цих чотирьох областях — як правило, менше 20%. Тоді як опитування двох українських соціологічних інститутів за останні п'ять років показують, що частка носіїв з індивідуально збалансованим співвідношенням використання української і російської мов становить приблизно від 40 до 50%, наші дані дають підстави вважати, що в цих цифрах «приховано» велику кількість носіїв «суржику». Наскільки їхній «суржик» більш наближений до української чи російської мов, можна з'ясувати лише тому, що частка респондентів, які сильніше орієнтовані на «суржик» і російську мову і, отже, «суржик» яких має бути «більш російськоподібним», досить висока порівняно із західнішими та північнішими областями.

Решта десять областей, крім Харківської та трьох областей Чорноморського регіону, має лише одну загальну характеристику: відсутність великої кількості однокодових російськомовних респондентів. В іншому вони дуже різні. Останнє вибіркове спостереження стосується однієї з них — Київської: тут більшість респондентів (близько 60 %) свідчить, що вони використовують усі три коди більш-менш рівномірно з високою частотністю. До цих показників близька лише сусідня Вінницька область (трохи більше 40%).

Однак опис ситуації в 14 областях, представлених на рис. 5, слід уточнити як загалом, так і в деталях. Одна з центральних цілей нашого дослідження — описати карту України з погляду трьох основних мовних кодів (української мови, російської мови та «суржику»), інакше кажучи, з погляду значення трьох кодів на даній території. Рисунок 5 не відображає наступного аспекту, який буде проілюстровано простим прикладом: для мовців, які заявляють, що дуже часто використовують один код (скажімо, українську мову), а інші коди — у найкращому випадку рідко (тип мовця U), українська мова відіграє набагато важливішу роль, ніж для мовців, які дуже часто використовують не лише українську, але й два інші коди — тут: російську мову й «суржик» (тип

мовця S-U-R). У цих двох випадках конкретне значення, наведене в оцінці частотності (тут: «дуже часто»), різниться залежно від комбінації, в якій воно виступає. Цей аспект буде розглянуто в подальшому аналізі.

На першому етапі кодам у кожній області було нараховано рейтингові бали. За кожного однокодового мовця типів U, R або S область отримала три бали за відповідний код, за кожного двокодового — по півтора бала за кожний код, за кожного трикодового — по одному балу за кожний код. Інакше кажучи, за кожного мовця було нараховано щонайбільше три бали, які мали бути надані одному, двом або трьом кодам, які використовуються регулярно й часто. Через різну кількість респондентів в областях на другому етапі бали було нормалізовано, щоб їхня сума становила 100. Таким чином, їх можна було розглядати аналогічно до відсоткових значень. Ці показники є індексом значення трьох кодів з урахуванням одно- і багатокодовості мовців.

Розраховані таким чином значення для кожного коду в певній області потім стали предметом кластерного аналізу. Для невеликої кількості прикладів (тут це 14 областей) найкраще підходить ієрархічний кластерний аналіз [див.: Bühl : 635-696]. Кількість розрахованих кластерів варіювалася від 2 до 7 (результати див. у табл. 3).

Таблиця 3

Комбінації масштабованих кластерів відстані (scaled distance clusters) (від 2 до 7 кластерів)						
Кластери: Область:	7	6	5	4	3	2
1: Хмельницька	1	1	1	1	1	1
2: Черкаська	1	1	1	1	1	1
3: Житомирська	2	2	2	1	1	1
4: Кіровоградська	2	2	2	1	1	1
5: Київська	2	2	2	1	1	1
6: Полтавська	2	2	2	1	1	1
7: Вінницька	2	2	2	1	1	1
8: Чернігівська	2	2	2	1	1	1
9: Дніпропетровська	3	3	3	2	2	1
10: Сумська	3	3	3	2	2	1
11: Одеська	4	3	3	2	2	1
12: Херсонська	5	4	4	3	3	2
13: Харківська	6	5	4	3	3	2
14: Миколаївська	7	6	5	4	3	2

Якщо розрізняти лише два кластери, то Херсонська, Харківська й Миколаївська області були б згруповані разом і, таким чином, відокремлені від усіх інших 11 областей (останній стовпець праворуч). На наступному етапі для трьох кластерів із цих 11 областей було б відокремлено ще три: Сумська, Дніпропетровська й Одеська, після чого залишилося б вісім областей. Якщо розрізняти чотири кластери, то Миколаївська область стає відокремленою

від Харківської й навіть сусідньої Херсонської областей — більш відмінною від будь-якої з восьми областей (Хмельницької, Черкаської, Житомирської, Кіровоградської, Київської, Полтавської, Вінницької, Чернігівської) та найбільш відмінною від решти семи областей. Останні утворюють досить компактний блок, який буде розділено лише на наступному етапі методу з п'ятьма кластерами. У цьому разі Хмельницька та Черкаська області будуть відокремлені від решти шести.

Цей аналіз можна розширити до 14 «кластерів», кожний з яких міститиме лише одну область. Таким чином, після застосування бритви Оккама кластери втрачають сенс. Як показано в таблиці 3, із семи кластерів кожна з трьох південних областей і Харківська область уже формують власний кластер. Отже, ми вирішили застосувати метод із п'ятьма кластерами у вигляді карт або рисунків, щоб проілюструвати безперервність мовної ситуації в Україні. З п'ятьма кластерами лише одна область формує власний кластер. Для ясності п'ять кластерів позначатимуться тут і далі римськими цифрами від I до V.

Відношення близькості між областями і кластерами областей найкраще проілюструвати за допомогою графічних засобів. У результаті кластерного аналізу створено дендрограму (тип деревоподібної діаграми), що показує відстань і подібність між аналізованими одиницями (областями) (див. рис. 6).

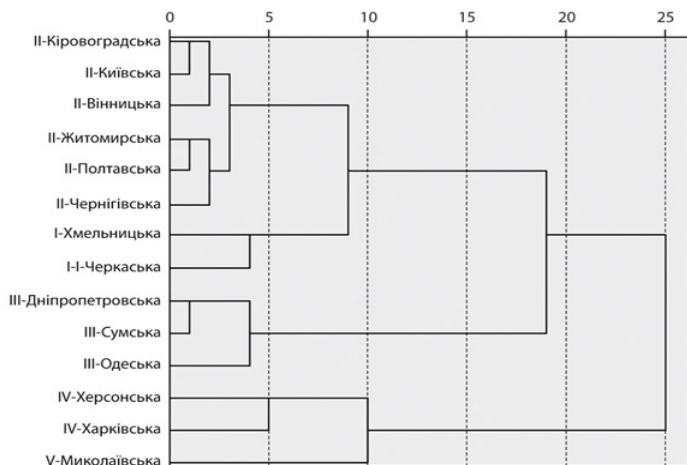


Рис. 6. Дендрограма ієрархічного кластерного аналізу на основі частотності типів мовців

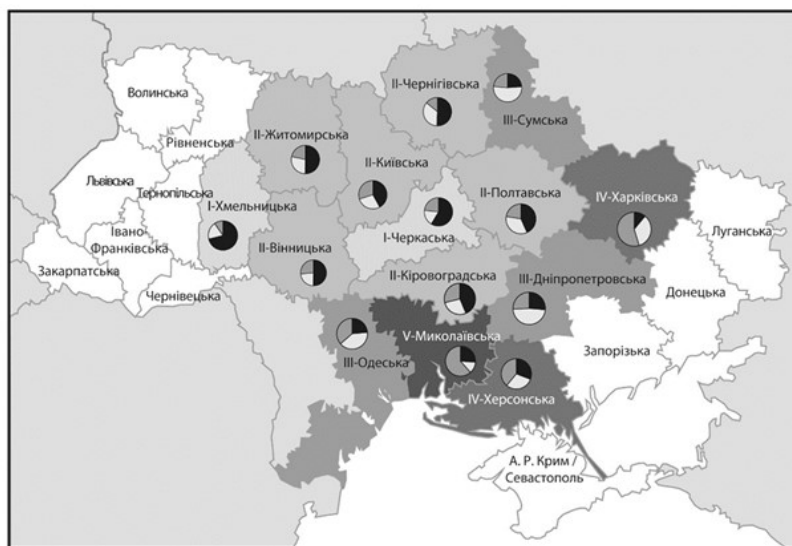
На рис. 6 місце, де з'являється розгалуження, відображається на горизонтальній лінії вгорі за шкалою від 0 до 25, що вказує на рівень відстані. Таким чином, перше розгалуження на відстані 25 виділяє блок Херсонської, Харківської та Миколаївської областей серед інших, відображаючи двокластерне розв'язання з правого стовпця таблиці 3 (див. вище). На відстані 19 (трикластерне розв'язання) Одеська, Сумська та Дніпропетровська області відгалужуються від інших восьми (перерахованих першими в вертикальному списку областей), і так далі. Розгалуження триває до кінцевої точки 14 областей. Таким чином, легко побачити, де ситуація є найбільш подібною. Однак слід урахувати не тільки об'єднання в одну лінію, а й відстань, на якій відбувається це розгалуження: наприклад, три пари областей розділені лише на мінімальній відстані 1: Кіровоградська і Київська, Житомирська і Полтавська, Дніпропетровська і Сумська. Повсякденне використання трьох кодів респондентами, які проживають у цих трьох парах областей, максимально подібне. За даними кластерного аналізу, в межах останнього

розгалуження є ще дві інші пари областей: Хмельницька і Черкаська, а також Херсонська і Харківська. Однак тут розгалуження відбувається на відстані 4 і 5 відповідно. Отже, області в останніх двох парах відрізняються одна від одної сильніше, ніж у трьох парах, названих першими. Більше того, ситуація у Хмельницькій і Черкаській або в Херсонській і Харківській областях більш неоднорідна, ніж у перших шести областях у верхній частині списку. Різниця між останніми починається з відстані 3, а не з 5, як у Херсонській і Харківській областях.

На жаль, неможливо безпосередньо зробити висновки про подібність між кластерами, що містять кілька елементів (областей) лише на основі їхнього вертикального порядку в дендрограмі, створеній у статистичній програмі SPSS. Так, наприклад, кластер I у п'ятикластерному розв'язанні (тобто Хмельницька і Черкаська області) у жодному разі не схожий більше на кластер III, ніж на кластер II.

Ці відношення між кластерами й навіть тонші подібності та відмінності в межах кластерів можна візуалізувати на карті (див. карту 3). На ній показано п'ять ареальних кластерів, позначених зліва від назв областей римськими цифрами.

Карта 3 подає ілюстративну й точнішу картину ситуації. Кластери з I до V показано різними відтінками сірого. Крім того, зв'язки, особливо всередині кластерів, диференціюються за допомогою так званих секторних діаграм²⁷. Чорна частина в них сигналізує про вагомість української мови у відповідній області, сіра — російської, а біла — «суржику».



Карта 3. Внутрішня диференціація чотирьох ареальних кластерів

Карта 3 демонструє кількісні результати, що відповідають таблиці 3 і дендрограмі на рис. 6, але тут вони представлені лише схематично. Звичайно,

²⁷ В англomовній статті, згаданій вище [Hentschel, Taranenko 2021: 294], як подібності, так і відмінності між кластерами та областями всередині кластерів були відображені на кольоровій карті шляхом програмно-керованого змішування трьох основних кольорних тонів, розрахованих безпосередньо на основі значень для трьох кодів. На жаль, це неможливо реалізувати в необхідній диференціації в чорно-білій версії.

на цій карті контрасти не такі чіткі, як, наприклад, на картах у дослідженнях Центру Разумкова 2016 року або КМІС 2019 року. Представлена тут карта більш диференційована, менш схематична. Кольорова карта в [Hentschel, Taranenko 2021: 294] показує це ще виразніше. Тут кожна область отримує забарвлення відповідно як до свого індивідуального соціолінгвістичного профілю, так і до належності до кластеру. Це в основному пов'язано з тим, що в тих дослідженнях до уваги брався історико-політичний територіальний розподіл, а потім обчислювалися середні значення для трьох із п'яти великих регіонів. Отже, територіально значні відмінності, особливо в надзвичайно неоднорідному Центральному регіоні, замовчувалися, а натомість відображалася більш очевидна відмінність між великими регіонами.

Ми обрали інший підхід. Як згадувалося вище, для кожної області ми обрахували міру використання кожного коду в повсякденному житті (за словами самих респондентів). Ці три значення для кожної області стали вихідними даними для кластерного аналізу. Таким чином, соціолінгвістичні просторові кластери областей і, відповідно, карта 3 безпосередньо ґрунтуються на аналізі, а не на «апріорі» сформованих субрегіонах України. Нижче показано 42 значення, які стали основою для кластерного аналізу й карти 3 (див. рис. 7). Ці результати представлено тут лише тому, що тепер ми можемо впорядкувати області відповідно до їх подібності і відмінності на основі кластерного аналізу.

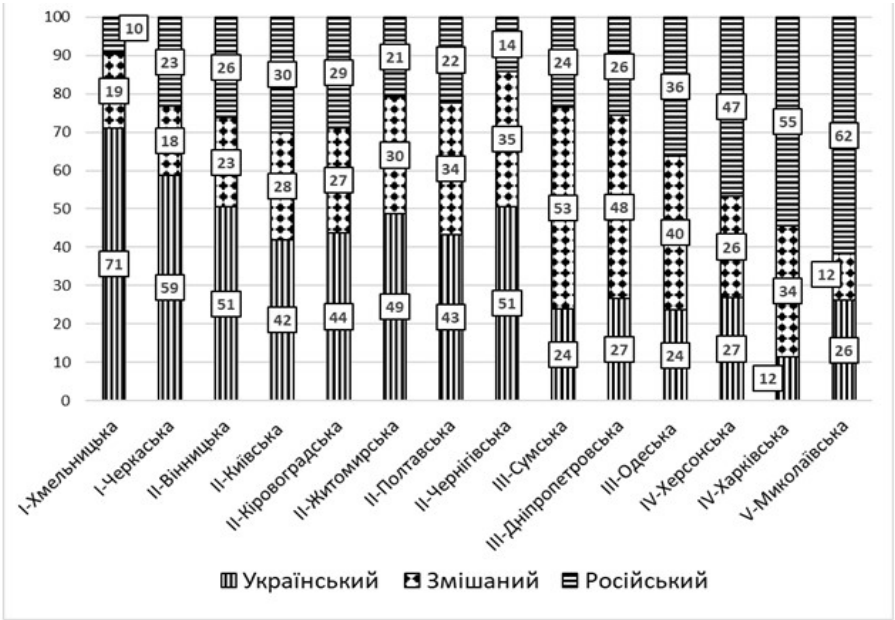


Рис. 7. Значення трьох кодів у кластерах і областях за шкалою від 0 до 100 балів

І карта 3, і рис. 7 по-різному ілюструють складне та неоднакове співвідношення в Центральній Україні і в трьох чорноморських областях. І і II кластери поєднують високий рівень використання української мови у повсякденному житті. У кластері I з Хмельницькою та Черкаською областями українська

мова має явну абсолютну перевагу (понад 50 пунктів, які можна розглядати як проценти). У найбільшому кластері II з Вінницькою областю, що розташована на «кордоні» з кластером I, у Київській, Кіровоградській, Житомирській, Полтавській та Чернігівській областях українська мова зберігає відносну перевагу: вона представлена сильніше, ніж будь-який з двох інших кодів, але не сильніше від двох інших разом узятих. Поступово значення «суржику» зростає із заходу на схід. Російська мова демонструє підвищені значення лише в Київській області, що можна пояснити як наслідок впливу мегаполіса Києва, та в Кіровоградській — через близькість до причорноморських областей, де російська мова має сильні позиції (див. нижче).

Чіткий розрив відбувається між кластерами II і III. В останньому, із Сумською та Дніпропетровською областями, а також з Одесою, яка наближається до кластера IV, «суржик» має абсолютну чи принаймні відносну перевагу в повсякденному житті людей. Тут значно зменшується значення української мови як коду повсякденного спілкування. Російська мова демонструє підвищену вагомість у порівнянні між кластерами II і III лише в Одеській області.

Між кластерами III і IV знову спостерігається чіткий розрив. Об'єднує їх лише відносно нижче значення української мови в повсякденному спілкуванні. Розрив в основному ґрунтується на тому, що в кластері IV російська мова виявляє відносну (Херсонська область) або навіть абсолютну (Харківська область) перевагу. Частка «суржику» є помітною в тому розумінні, що вона цілком порівнювана з ситуацією в кластері II, але різко зменшується порівняно з кластером III. Нарешті, кластер V, який складається лише з Миколаївської області, посилює тенденцію кластерів IV і III. Це означає, що частка російської мови знову зростає — до найвищого значення з усіх 14 зафіксованих областей, а частка «суржику» падає далі, до найнижчого значення з усіх областей. Таку динаміку, тобто той факт, що респонденти оцінювали російську як свою повсякденну мову ще вище, а «суржик» нижче, ніж у сусідній Херсонській області, розташованій далі на схід, чи в Харківській області на північному сході, ми не можемо пояснити на основі даних цього опитування²⁸. Однак подібність мовної ситуації в цих трьох областях загалом не викликає подиву.

4. Висновки. Скрупульозні соціолінгвістичні дослідження із застосуванням статистичного аналізу, кількісних методів у висвітленні функціонування різних ідіомів у межах однієї країни та ставлення до них з боку різних верств населення, безперечно, допомагають скласти якомога об'єктивнішу картину реального мовного життя у відповідних регіонах (незважаючи на скептицизм у цьому плані окремих дослідників)²⁹. Кількісна оцінка польових спостережень, звичайно, багато в чому залежить від передумов і методів збору даних і конкретного статистичного аналізу, які, якщо вони чітко вказані, слід брати до уваги при інтерпретації результатів.

Соціолінгвістична модель повсякденної мовної трикодовості в сучасній Україні, яка пропонується в цій статті, не дає (принаймні в її картографії)

²⁸ Це закриті кількісне опитування доповнюється відкритими глибинними інтерв'ю, які пізніше будуть проаналізовані квалітативно. Таким чином ми сподіваємося отримати додаткову інформацію.

²⁹ Пор., наприклад: «Оскільки “суржикові” різновиди не є окремим мовним кодом, усілякі мовні статистики, які пропонують окрему категорію “суржик”, викликають великий скепсис» [Мозер : 27].

чіткої й повної картини більших регіонів України. Інші аналогічні в тому чи іншому плані описи мовної ситуації в Україні відштовхувалися від історико-політичного структурування країни, щоб потім обраховувати середню частотність мовних кодів (звичайно первинних кодів, якщо не рідних мов) без урахування істотної диференціації в межах цих ареалів. Історико-політичне підґрунтя сучасної мовної ситуації в країні відбито, звичайно, і в нашому аналізі (див. карту 3). Просторові кластери IV і V є частиною Півдня і Сходу (в широкому сенсі), які відрізняються від решти України в усіх дослідженнях (див., зокрема, карту 1 у: [Хмелько]). У будь-якому разі ареал, охоплений кластерами IV і V, значною мірою відповідає ареалу України, де слов'янські поселення як з української, так і російської території було засновано відносно пізно. Ми також виявили, що в цьому ареалі російська мова має особливо велике значення, але значно менше порівняно з даними зі згаданих вище досліджень, оскільки ми брали до уваги мовну трикодовість у реальному побутовому спілкуванні місцевого населення. Так, Одеська область з Півдня групується разом із Дніпропетровською та Сумською областями в кластер III з урахуванням кількісного обсягу трикодовості. Кластери III, IV і V покривають переважну частину території, згрупованої КМІС (2003) за межами «Центру», тобто на Сході (включно зі «Східним Центром») або Півдні України, разом з Полтавською та Чернігівською областями. Останні дві області, однак, групуються разом з центральними регіонами далі на захід, якщо, як у нашому аналізі, до уваги береться не лише первинна мова й не ігнорується «суржик».

Ця територія (наші кластери I і II) є неоднорідною³⁰. Це, зрештою, й не дивно з огляду на різний час заселення різних регіонів країни, їхню залежність від різних сусідніх держав і як наслідок — ті чи інші відмінності їх не тільки в суто географічному плані, а й у політичних, конфесійних, культурних і, звичайно, мовних уподобаннях населення. З іншого боку, причини різної одно, дво і трикодовості серед мовців, звичайно, не лише історико-політичні. Соціобіографічні характеристики респондентів і відповідний соціально-демографічний склад областей також відіграють при цьому значну роль (так, у [Hentschel, Zeller 2017] описано взаємозв'язок між мовною орієнтацією респондентів і такими соціобіографічними характеристиками для Центральної України). На основі даних з Центральної України у праці [Hentschel, Zeller 2016 : 652–657] виділено чотири типи ставлення до «суржику», які впливають на те, чи респондент вживає «суржик» і чи він це визнає. Політичні погляди загальнішого плану, як-от щодо політики Росії стосовно Криму й Донбасу, можуть вплинути на використання російської мови російськомовними громадянами України або принаймні на твердження респондентів про частотність використання ними російської.

³⁰ Вона навіть більш неоднорідна, ніж це було показано в праці [Гентшель, Тараненко 2015]. Її внутрішня структура трохи відрізняється від структури, поданої в дослідженні шестирічної давності, хоча в обох цих дослідженнях в основному використовується та сама база даних щодо Центрального регіону. Обраний тоді метод був менш чутливим у тому плані, що мовна трикодовість спостерігалася, але лише на основі рівнів використання «завжди» і «часто». Зокрема, недавнє дослідження підтверджує взаємозв'язок між «старими» регіонами від А до С (тепер це кластери I і II), з одного боку, і областями в D (тут це кластер III, без Одеської області, яка не розглядалася 2015 року), з другого боку (і в цілому), — неоднорідність Центрального регіону.

Такі залежності і пов'язані з ними відмінності в межах кластерів і між ними буде розглянуто в подальшому дослідженні.

Карта українсько-російської двомовності в Україні є значно складнішою, якщо брати до уваги реальну мовну трикодовість різних груп населення — з урахуванням досить поширеного в країні українсько-російського змішаного мовлення. При такому соціолінгвістичному картографуванні контраст між великими регіонами країни, такими як Захід, Центр (можливо, розділений на західну і східну частини) і Південь, звичайно, не зникає, але значно послаблюється. На відміну від простіших протиставлень між великими регіонами, які в кінцевому підсумку припускають дихотомічний поділ України за переважанням чи української, чи російської мов, наведена тут модель мовної ситуації на основі враховування мовної трикодовості різних груп населення демонструє не різку, а поступову, градуальну диференціацію реального мовного простору країни.

ЛІТЕРАТУРА

- Белей Л. Українська мова на початку ХХІ ст.: параметризація соціально-лінгвістичних змін. Київ : Люта справа, 2019. 144 с.
- Брага І. «Клуб бойового суржика», або Українсько-російський суржик у соціальній мережі Facebook. *Дивослово*. 2015. № 5. С. 32–34.
- Думки і погляди населення щодо викладання російської мови в українськомовних школах і надання непідконтрольним територіям Донбасу автономії в складі України. *Київський міжнародний інститут соціології : Прес-релізи та звіти: березень 2019 року*. URL: <https://www.kiis.com.ua> (дата звернення: 12.08.2020)
- Ідентичність громадян України в нових умовах: стан, тенденції, регіональні особливості. *Національна безпека і оборона*. 2016. № 3–4. С. 3–117.
- Кого українці хотіли б бачити предстоятелем об'єднаної церкви. *Київський міжнародний інститут соціології : Прес-релізи та звіти: липень 2020 року*. URL: <https://www.kiis.com.ua> (дата звернення: 12.08.2020)
- Консолідація українського суспільства: виклики, можливості, шляхи. *Національна безпека і оборона*. 2016. № 7–8. С. 2–106.
- Курохтіна Т. Суржик в інтернеті. *Вестн. Моск. ун-та : Сер.9. Филология*. 2009. № 2. С. 101–112.
- Масенко Л. Суржик: між мовою і язиком. Київ : Видавн. дім «Києво-Могилянська академія», 2011. 135 с.
- Мозер М. «Суржик» чи «суржики»? *Українська мова*. 2016. № 1. С. 27–54.
- Ставицька Л., Труб В. Суржик: суміш, мова, комунікація. *Українсько-російська двомовність : Лінгвосоціокультур. аспекти* / за ред. Л. О. Ставицької. Київ : ПУЛЬСАРИ, 2007. С. 31–120.
- Тараненко О. О. Суржик. *Українська мова : Енциклопедія* / редкол.: В. М. Русанівський, О. О. Тараненко (співголови) та ін. — 3-тє вид., зі змін. і допов. Київ : +Укр. енциклопедія ім. М. П. Бажана, 2007. С. 689–692.
- Тараненко О. Варіантність vs. стабільність у структурі українсько-російського «суржика» (УРС): сукупність ідіолектів vs. соціолект. *Variation und Stabilität in Kontaktvarietäten : Beobachtungen zu gemischten Formen der Rede in Weißrussland, der Ukraine und Schlesien* (Studia Slavica Oldenburgensia, 21) / G. Hentschel (ed.). Oldenburg, 2013. С. 27–62.
- Хмелько В. Є. Лінгво-етнічна структура України: регіональні особливості та тенденції змін за роки незалежності. *Наук. зап. Нац. ун-ту «Києво-Могилянська академія»: Соціол. науки*. 2004. Т. 32. С. 3–15.

- Bilaniuk L. Purism and Pluralism : Language use trends in popular culture in Ukraine since independence. *Harvard Ukrainian Studies*. 2018. Vol. 35. N 1–4. P. 293–309.
- Bühl A. SPSS — Einführung in die moderne Datenanalyse ab SPSS 25. Hallbergmoos : Pearson, 2019 (16. Aufl.). 1056 S.
- Flier M. The rules of engagement. *Harvard Ukrainian Studies*. 2000. Vol. 22. P. 113–136.
- Hentschel G. Eleven Questions and Answers about Belarusian-Russian Mixed Speech («Trasyanka»). *Russian Linguistics*. 2017. Vol. 41(1). P. 17–42.
- Hentschel G. Die weißrussische «Trasjanka» und der ukrainische «Suržyk»: grundlegende quantitativ-qualitative soziolinguistische Ähnlichkeiten und Unterschiede. *Deutsche Beiträge der deutschen Slavistik zum XVI. Internationalen Slavistenkongress (Belgrad, August 2018)* / S. Kempgen, M. Wingender, L. Udolph (Eds.). München : Harrassowitz, 2018. S.127–138.
- Hentschel G., Kittel B. Weißrussische Dreisprachigkeit? Zur sprachlichen Situation in Weißrussland auf der Basis von Urteilen von Weißrussen über die Verbreitung «ihrer Sprachen» im Lande. *Wiener Slavistischer Almanach*. 2011. Bd. 67. S. 107–135.
- Hentschel G., Taranenko O. Die Sprachenlandschaft der zentralen Ukraine: Ukrainisch, Russisch, «Suržyk». Verwendung — Kompetenz — nationale Positionierung. *Die Welt der Slaven*. 2015. Jg. 60. Heft 2. S. 248–275.
- Hentschel G., Taranenko O. Bilingualism or tricolalism: Ukrainian, Russian and «Suržyk» in Ukraine: analysis and linguistic-geographical mapping. *Die Welt der Slaven*. 2021. Jg. 66. Heft 2. P. 268–299.
- Hentschel G., Zeller J. P. Meinungen und Einstellungen zu Sprachen und Kodes in zentralen Regionen der Ukraine. *Ztschr. für Slavistik*. 2016. Bd. 61(4). S. 636–661.
- Hentschel G., Zeller J. P. Aspekte der Sprachverwendung in zentralen Regionen der Ukraine. *Wiener Slavistischer Almanach*. 2017. Bd. 79 (1). S. 37–60.
- Menzel Th., Hentschel G. Flexionsmorphologische Irregularität im innerslavischen Sprachkontakt. Sprachinhärente Präferenzen oder politisch-soziale Dominanz: Russisch vs. Weißrussisch / Ukrainisch — Polnisch vs. Kaschubisch / Lemkisch. Frankfurt a. M. : Peter Lang, 2017. 371 S.
- Taranenko O. Ukrainisch-russischer Suržyk: Status, Bewertungen, Tendenzen, Prognosen. *Trasjanka und Suržyk — gemischte weißrussisch-russische und ukrainisch-russische Rede. Sprachlicher Inzest in Weißrussland und der Ukraine?* / G. Hentschel, O.Taranenko, S. Zaprudski (Eds.). Frankfurt a. M. : Peter Lang, 2014. S. 265–288.
- Trasjanka und Suržyk — gemischte weißrussisch-russische und ukrainisch-russische Rede. Sprachlicher Inzest in Weißrussland und der Ukraine? / G. Hentschel, O.Taranenko, S. Zaprudski (Eds.). Frankfurt a. M. : Peter Lang, 2014. 394 S.
- Zeller J. P. Zum ukrainisch-russischen Sprachkontakt: Phonische Variation im ukrainischen «Suržyk» im Vergleich mit der weißrussischen «Trasjanka». *Deutsche Beiträge der deutschen Slavistik zum XVI. Internationalen Slavistenkongress (Belgrad, August 2018)* / S. Kempgen, M. Wingender, L. Udolph (Eds.). München : Harrassowitz, 2018. S. 365–374.

REFERENCES

- Bilaniuk L. (2017–2018). Purism and Pluralism. Language use trends in popular culture in Ukraine since independence. *Harvard Ukrainian Studies* 35, (1–4), 293–309.
- Bühl A. SPSS — Einführung in die moderne Datenanalyse ab SPSS 25. Hallbergmoos. Aufl. 16.
- Dumky i pohliady naselennia shchodo rosiiskoi movy v ukrainskomovnykh shkolakh i nadannia nepidkontrolnym terytoriiam Donbasu avtonomii u skladi Ukrainy. URL: <https://www.kiis.com.ua/?lang=ukr&cat=reports&id=832&page=2> (last accessed: 12.08.2020) [In Ukrainian].

- Flier M. (2000). The rules of engagement. *Harvard Ukrainian Studies* 22. 113–136.
- Hentschel G. (2017). Eleven Questions and Answers about Belarusian-Russian Mixed Speech («Trasyanka»). *Russian Linguistics*, 41(1), 17–42.
- Hentschel G. (2018). Die weißrussische «Trasjanka» und der ukrainische «Suržyk»: grundlegende quantitativ-qualitative soziolinguistische Ähnlichkeiten und Unterschiede. *Deutsche Beiträge der deutschen Slavistik zum XVI. Internationalen Slavistenkongress, Belgrad, August 2018* (Reihe Die Welt der Slaven. Sammelbände — Sborniki). S. Kempgen, M. Wingender, L. Udolph (Eds.). München, 127–138.
- Hentschel G., Kittel B. (2011). Weißrussische Dreisprachigkeit? Zur sprachlichen Situation in Weißrussland auf der Basis von Urteilen von Weißrussen über die Verbreitung «ihrer Sprachen» im Lande. *Wiener Slawistischer Almanach*. Band 67, 107–135.
- Hentschel G., Taranenko O. (2015). Die Sprachenlandschaft der zentralen Ukraine: Ukrainisch, Russisch, 'Suržyk'. Verwendung — Kompetenz — nationale Positionierung. *Die Welt der Slaven* LX. 248–275.
- Hentschel G., Zeller J. P. (2016). Meinungen und Einstellungen zu Sprachen und Kodes in zentralen Regionen der Ukraine. *Zeitschrift für Slawistik*, 61(4), 636–661.
- Hentschel G., Zeller J. P. (2017). Aspekte der Sprachverwendung in zentralen Regionen der Ukraine. *Wiener Slawistischer Almanach* 79, (1), 37–60.
- Hentschel G., Taranenko O. (2021). Bilingualism or tricolism: Ukrainian, Russian and «Suržyk» in Ukraine: analysis and linguistic-geographical mapping. *Die Welt der Slaven*. Jg. 66. Heft 1. 268–299.
- Identychnist hromadih Ukrainy v novykh umovakh: stan, tendencii, rehionalni osoblyvosti (2016). *Nacionalna bezbeka i oborona*, (3–4), 161–162. URL: http://razumkov.org.ua/uploads/journal/ukr/NSD161-162_2016_ukr.pdf, pp. 7–8, 70–72, 86 (last accessed: 12.08.2020). [In Ukrainian].
- Konsolidatsiia ukrainskoho suspilstva: Vyklyky, mozhlyvosti, shliakhy (2016). = *Nacional'na bezbeka i oborona*, (7–8), 165–166. URL: http://razumkov.org.ua/uploads/journal/ukr/NSD165-166_2016_ukr.pdf (last accessed: 12.08.2020). [In Ukrainian].
- KIIS. (2020). Koho ukraintsi khotily b bachyty predstoiatlem ob'yednanoi tserkvy. URL: <http://kiis.com.ua/?lang=ukr&cat=reports&id=958&page=1> (last accessed: 12.08.2020). [In Ukrainian].
- Kurohtina T. (2009). Surzhyk v internete. *Vestnik Moskovskogo Universiteta*, seriya 9, Filologiya, (2), 101–112. [In Russian].
- Menzel Th., Hentschel G. (2017). *Flexionsmorphologische Irregularität im innerslavischen Sprachkontakt. Sprachinhärente Präferenzen oder politisch-soziale Dominanz: Russisch vs. Weißrussisch / Ukrainisch — Polnisch vs. Kaschubisch / Lemkisch*. Frankfurt a.M.
- Moser M. (2016). «Surzhyk» chy «surzhyky»? *Ukrainska mova*, (1), 27–54.
- Taranenko O. (2013). Variantnist vs. stabilnist u strukturi ukrainsko-rosiiskoho 'surzhyku' (URS): sukupnist idiolektiv vs. sotsiolekt. *Variation und Stabilität in Kontaktvarietäten: Beobachtungen zu gemischten Formen der Rede in Weißrussland, der Ukraine und Schlesien* (Studia Slavica Oldenburgensia, 21). Gerd Hentschel (Ed.). Oldenburg, 27–62.
- Taranenko O. (2014). Ukrainisch-russischer Suržyk: Status, Bewertungen, Tendenzen, Prognosen. *Trasjanka und Suržyk — gemischte weißrussisch-russische und ukrainisch-russische Rede: Sprachlicher Inzest in Weißrussland und der Ukraine?* Hentschel G., O. Taranenko, S. Zaprudski (Eds.). Frankfurt a.M., 265–288.
- Trasjanka und Suržyk — gemischte weißrussisch-russische und ukrainisch-russische Rede. Sprachlicher Inzest in Weißrussland und der Ukraine? (2014). Hentschel G., Taranenko O., Zaprudski S. (Eds.). Frankfurt a.M.
- Zeller J. P. (2018). Zum ukrainisch-russischen Sprachkontakt: Phonische Variation im ukrainischen «Suržyk» im Vergleich mit der weißrussischen «Trasjanka». *Deutsche Bei-*

träge der deutschen Slavistik zum XVI. Internationalen Slavistenkongress (Belgrad, August 2018). S. Kempgen, M. Wingender, L. Udolph (Eds.). München, 365–374.

G. HENTSCHEL

University of Oldenburg (Slavic department)
Oldenburg, Germany
E-mail: gerd.hentschel@uol.de
<https://orcid.org/0000-0003-2034-1577>

O. O. TARANENKO

O. O. Potebnia Institute of Linguistics of the National Academy of Sciences
of Ukraine
Kyiv, Ukraine
E-mail: oataran99@gmail.com
<https://orcid.org/0000-0002-0269-8954>

**BILINGUALISM OR TRICODALISM: UKRAINIAN, RUSSIAN AND
«SURŽYK» IN UKRAINE (analysis and linguistic-geographical mapping)**

The study aims to show that the different usage frequencies of Ukrainian, Russian and «Suržyk» in everyday life in Ukraine can be described and mapped as a continuum of «tricodalism». This sets it apart from previous representations that usually assume a historical-political demarcation of sub-areas and consequently tend to indicate a linguistic bisection of the country into a Ukrainian-speaking area encompassing the west and the central region, and a Russian-speaking area in the east and south. The approach proposed here employs empirical data on usage of the three codes as the basis for its structuration. Starting from the mean frequencies of code usage in individual oblasts, a cluster analysis groups these into larger areas and clarifies their internal variation. This result is not just a more continuous picture of the linguistic landscape in Ukraine, but also a more coherent one.

Keywords: language situation in Ukraine, Ukrainian-Russian bilingualism, Ukrainian-Russian «suržyk», sociolinguistics, linguistic cartography, language contact.

Дата надходження до редакції — 29.11.2021

Дата затвердження редакцією — 12.12.2021