



Ніна Клименко
доктор філол. наук,
член-кореспондент НАН України,
пров. наук. співр. Інституту
мовознавства ім.О.О.Потебні
НАН України



Євгенія Карпіловська
доктор філол. наук,
професор, в.о. завідувача
відділу Інституту мовознавства
ім.О.О.Потебні НАН України



Лариса Кислюк
канд. філол. наук,
наук. співр., докторантка
Інституту мовознавства
ім.О.О.Потебні НАН України

З КОМП'ЮТЕРОМ — ДО ГЛИБИН МОВИ

*з нагоди 80-річчя Інституту
мовознавства ім.О.О.Потебні НАН України*

*Ювілеї дають нагоду оглянути пройдений шлях,
оцінити зроблене й накреслити нові перспективи.
От і ми з нагоди ювілею маємо привід осмислити внесок Інституту
в становлення і розвиток в Україні такого новітнього напрямку
сучасного мовознавства, як комп'ютерна лінгвістика.
Історія науки, як і історія взагалі — річ не безіменна.
Її хід, перипетії розвитку певних напрямів і ідей, причини дискусій
довкола певних проблем стають зрозумілими на прикладі діяльності
конкретних наукових осередків і їхнього доробку.
Тому й історію застосування комп'ютерів у дослідженні
української мови, розв'язанні фундаментальних теоретичних проблем
українського мовознавства спробуємо подати через історію
того наукового колективу, з яким упродовж років нерозривно
пов'язана наша людська й професійна доля, — відділу структурно-
математичної лінгвістики Інституту мовознавства ім.О.О.Потебні.
Наш відділ стояв у витоків української комп'ютерної лінгвістики,
а тому в його вже майже півстолітній історії
неминуче відбилися всі злети і втрати,
пошуки і проблеми цього наукового напрямку.*

1. У далекі 1960-ті...

Початок 1960-х років для України, як і для всього Радянського Союзу, був означений сподіваннями на краще, які збурила післясталінська "відлига". Позитивні зрушення відбувалися й у мовознавчій науці. У цю галузь наукових досліджень, щоправда, дозовано, під пильним контролем компартійної верхівки, було допущено ті нові ідеї, якими жила тоді світова лінгвістика, і не помічати які було вже просто неможливо. З початком науково-технічної революції в СРСР гостро постає проблема оперативного опрацювання інформації, зокрема мовної. Так у теорію та практику радянського мовознавства стали проникати ідеї структуралізму, який активно розвива-

ли наукові осередки в Європі та США. На очах одного покоління в українському словнику утвердилося англійське запозичення *комп'ютер*, а для галузі мовознавства, яка його використовує для розв'язання теоретичних і практичних завдань, назва *комп'ютерна лінгвістика*. Використання комп'ютера для дослідження мови відкривало заманливі перспективи одержання нових відомостей про її будову та функціонування, про закономірності її розвитку.

1962 р. в складі Інституту мовознавства ім.О.О. Потебні НАН України (тоді — АН УРСР) був створений новий відділ теорії мови та структурно-математичної лінгвістики на чолі з відомим українським мовознавцем *М.А. Жовтбрюхом*. Очолити групу

фахівців зі структурно-математичної лінгвістики в складі цього відділу керівництво Інституту, його тодішній директор — академік *І.К. Білодід* запросили завідувачку кафедри іноземних мов Черкаського педагогічного інституту (нині — Черкаський національний університет імені Богдана Хмельницького) *Валентину Сидорівну Перебийніс*. Вона на той час уже здобула високий авторитет у колах структуралістів в Україні і поза нею своїми працями з лінгвостатистики та моделювання мови. Перед групою було поставлено масштабне завдання описати за допомогою новітніх формально-логічних, структурних та математичних методів систему української мови, а також створити статистичні моделі її функціонування в текстах різних стилів. Виконання таких досліджень заклало надійне теоретичне й фактичне підґрунтя для запровадження комп'ютерних технологій опрацювання не лише української, а й інших національних мов. Для кадрового забезпечення цього новітнього напрямку мовознавчих досліджень в Україні в Інституті мовознавства в 1964 р. було відкрито аспірантуру зі спеціальності 10.02.21 — структурна, прикладна та математична лінгвістика. Першими аспірантами з цієї спеціальності стали учениці Валентини Сидорівни — *Людмила Антонівна Алексієнко*, яка пізніше, перейшовши на роботу до Київського державного (нині — національного) університету імені Тараса Шевченка, створила там першу у вищих України лабораторію комп'ютерної лінгвістики при кафедрі сучасної української мови, та *Ніна Федорівна Клименко*, нині — член-кореспондент НАН України, яка в 1986 р. прийняла від Валентини Сидорівни керівництво відділом, очолювала його до 2004 р. і нині продовжує в ньому працювати.

У 1962-1968 рр. до складу групи входили крім *В.С. Перебийніс*, *М.П. Муравицька*, *М.М. Пешак*, *Т.О. Грязнухіна*, *І.Ф. Савченко*, *О.Ф. Савранчук*, *Л.А. Алексієнко*, *Н.Ф. Клименко*. Згодом до них приєдналися випускниці Київського державного університету ім. Т.Г. Шев-

ченка *Н.П. Дарчук*, *В.І. Критська*, *Л.В. Орлова*, які закінчили відділення структурної та математичної лінгвістики. У відділі з кафедрою математичної лінгвістики Київського університету, як і з іншими кафедрами й окремими фахівцями, що працювали в галузі структурної, математичної та прикладної лінгвістики, комп'ютеризації лінгвістичних досліджень в Україні та поза нею, встановилися тісні професійні й людські контакти. "Структуралісти" тодішнього Радянського Союзу всі були знайомі між собою, часто брали участь у міжрегіональних і міжвідомчих наукових проєктах, як, наприклад, група "Статистика мовлення", що об'єднала тоді фахівців з різних установ і регіонів Радянського Союзу й закордону. Для членів дружної структуралістської спілки стало доброю традицією, навідуючи Київ, обов'язково заглядати до Інституту для зустрічі з Валентиною Сидорівною та її колегами або й спеціально приїжджати до них у різних справах. Тоді у всіх була можливість поспілкуватися, незалежно від рангу, з москвичами *П.А. Соболевою*, *Г.П. Мель-*

никовим, *С.С. Хідекель*, *А.Я. Шайкевичем*, *М.В. Араповим*, *А.А. Полікарповим*, *Ю.А. Тулдавою* з естонського Тарту, лєнінградцями *Р.Г. Піотровським*, *Л.М. Бєляєвою*, *О.С. Гердом*, *Ю.К. Криловим*, *Ю.К. Орловим* з Тбілісі, *О.В. Зубовим* з Мінська, *Т.А. Якубайтис* з Риги та іншими провідними фахівцями у галузі структурної, математичної, інженерної лінгвістики та лінгвостатистики. Ці зустрічі й розмови важили більше, ніж стоси книжок і університетські курси. Це була наука в дії, практичний досвід, якого ще ніщо і нікому не замінювало. Попри тодішню ідеологічну боротьбу з ідеалізом "хомскіанства", що неминуче відображалася на ставленні до вітчизняних фахівців, які працювали в галузі структурної, математичної та прикладної лінгвістики, українські структуралісти жили насиченим життям, активно спілкуючись між собою, здійснюючи спільні наукові проєкти, а також перебуваючи в постійному контакті з філософами, кібернетиками, фізиками, математиками. Щороку у Будинку науково-технічної пропаганди по бульвару Тараса Шевченка, 16 (нині тут



Група структурно-математичної лінгвістики. Зліва направо стоять: *Л.М.Гриднева*, *В.С.Перебийніс* (керівник), *Н.Ф.Клименко*, *Л.А.Алексієнко*, *М.П.Муравицька*; сидять: *О.Ф.Савранчук*, *І.Ф.Савченко*, *М.М.Пешак*. Київ, 1965 р.



міститься Міністерство освіти та науки України) відбувалися семінари з проблем штучного інтелекту, які проводив молодий академічний Інститут кібернетики на чолі з *Віктором Михайловичем Глушковым*. Там лунали незвичні для філологів "алгоритми", "ряди Колмогорова", "ланцюги Маркова" й "інтелектуальні роботи", але ці високоерудовані, неординарні, розкуті й дотепні, захоплені своєю справою люди робили те, що сприяло виходу нашої країни й науки на рівень світових потреб і вимог, руйнувало стереотип гуманітарія як суто кабінетного вченого-дивака, що порпається у вицвілих, малощікавих загалу папірцях або такої собі училки-зубрилки — "синьої панчохи".

Комп'ютер висував мовознавцям свої "правила гри", вимагав принципово іншого опрацювання й подання мовного матеріалу. Для сприйняття мови іншим адресатом — технічним засобом, що не має людського інтелекту, інтуїції, мовної компетенції, — її слід було структурувати й формалізувати відповідно до можливостей математичного й програмного забезпечення тогочасного комп'ютера. Необхідно було створити штучний аналог природної мови, або мовний складник штучного інтелекту комп'ютера. Показовою для кола проблем, над якими в ці роки інтенсивно працювала група,

є назва виданої 1972 р. науково-популярної книжки *В.С. Перебийніс* — "Коли робот заговорить (Мовне спілкування людини з електронною машиною)".

Ужинок цього періоду дослідницької праці "структурних дівчат", як їх жартома почали іменувати в Інституті, дає уявлення про те наукове поле, на якому дедалі зосереджувала свої дослідження світова комп'ютерна лінгвістика — створення формальних моделей системи мови та продуктів її реалізації у мовленні: письмових і усних текстах. Ці дослідження мали ґрунтуватися на розв'язанні фундаментальних теоретичних проблем структурування й формалізації мовних явищ, конструювання нових лінгвістичних об'єктів як складників штучного інтелекту комп'ютера. Саме ця проблематика, опрацюванням якої й мала зайнятися фундаментальна академічна наука, визначила місце відділу структурно-математичної лінгвістики Інституту в історії української, як і світової, комп'ютерної лінгвістики.

Зусилля членів групи в 1960-і роки спрямувалися на підготовку структурної граматики української мови та розроблення статистичного апарату для аналізу українських текстів різних функціональних стилів ("Структурно-математична лінгвістика" (К., 1965), "Статистичні та структурні лінгвістичні моделі" (К., 1966), "Методи

структурного дослідження мови" (К.: 1968)). Найвагомим здобутком цього етапу стала колективна монографія за редакцією *В.С. Перебийніс* "Статистичні параметри стилів" (К.: 1967). У ній обґрунтовано поняття "нульового" або нейтрального стилю мови, певної точки відліку, щодо якої виразнішими ставали формальні показники лексичних і граматичних особливостей інших стилів української мови: художнього, наукового, публіцистичного, офіційно-ділового, розмовно- побутового. Ця монографія сьогодні привертає дедалі більшу увагу дослідників у галузі функціональної стилістики, корпусної лінгвістики та статистичної лексикографії.

2. Нове обличчя лінгвістики

1 листопада 1968 р. на карті Інституту з'явився окремий відділ структурно-математичної лінгвістики. Його завідувачкою цілком закономірно призначили *В.С. Перебийніс*. Маленька на зріст, завжди гранично зібрана, акуратна, стримана, іронічна, вона ніколи не підвищувала голосу. Проте, якщо хтось проштрафлювався, особливо, в роботі, вона цим тихим голо- сом могла вичитати так, що пропасниця була винуватця ще довго після такої нагінки. У відділі не було поділу на науковців і їхню обслу-гу,



Співробітники і вихованці відділу структурно-математичної лінгвістики різних років на ювілеї Н.Ф.Клименко. Стоять зліва направо: Т.І.Недозим, В.І.Критська, Л.А.Алексієнко, Н.П.Дарчук, Н.Ф.Клименко, В.С.Перебийніс, Л.М.Гриднева, Т.О.Грязнухіна, М.Ю.Федурко, Л.П.Кислюк, Є.А.Карпіловська. Київ, 2009 р.

старших лаборантів, або, як на інститутському ідіолекті їх жартома називали, "старших рабів". (У цьому каламбурі обіграно скорочення, прийняте в робочих планах Інституту, "ст. лаб."). Зайвим доказом цього було те, що у лаборантів були свої власні робочі столи, як і у наукових співробітників, а отже, була можливість нормально працювати, чого не можна сказати сьогодні про умови роботи не лише лаборантів, а й науковців Інституту.

В.С. Перебийніс була і є рідкісним науковим редактором. Просвічувала текст наскрізь, як рентгенівськими променями, моментально вловлюючи логічні несумісності, поверховість в аналізі і безпорадність в осмисленні його результатів. Терпіти не могла наукоподібності, безпредметного розумування й халтури. Відчувала від цього раз і назавжди всіх, хто з нею працював. Якщо у тексті вона натрапляла на якийсь закручений пасаж, на берегах сторінки з'являвся знак оклику або питання. Знайшовши щось таке, Валентина Сидорівна питала: "Що Ви цим хотіли сказати?" (вона незмінно до всіх підлеглих і молодших за віком зверталася на "ви"). Вислухавши пояснення, радила: "Добре. А тепер сядьте і напишіть все це простими людськими словами". Так само відпочувала від сплесків емоцій у науковому викладі. У вступному рефераті до аспірантури нашої колеги кожен учений, на праці якого вона покликала, здобув свій "ярлик": Сосюр як батько структуралізму, звичайно, був великим, *Потебня* і *Бодуен де Куртене* — видатними, інші — хто відомим, хто талановитим, а хто й просто — молодим. На берегах негайно ж з'явився відгук Валентини Сидорівни "Не створи собі кумира!". А в розмові згодом вона жартома порадила виробити класифікацію, щоб нікого не образити й розставити все по місцях: хто видатний, хто наш — хто не наш, а хто ще просто молодий. Натрапляючи в деяких текстах на таку шкалу оцінок, згадуємо цей урок.

У мудрій іронії, неметушливості, принциповості, мужності мати власну думку і обстоювати її, немінуче ускладнюючи тим своє життя, теж була і є незвичність "структуралістів" старшого покоління, цих жінок з гнучким розумом, глибокими знаннями і негнучким хребтом.

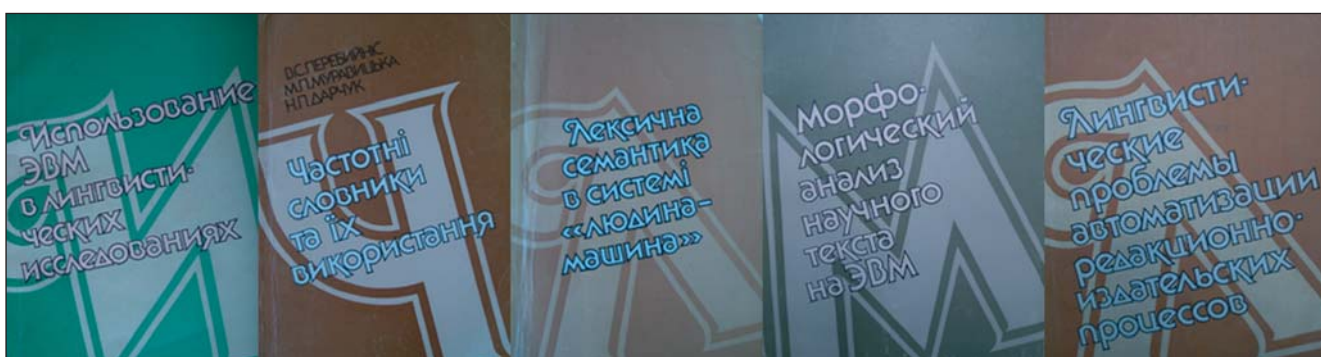
Відділ, який займався справами, що вимагали, здавалось би, чоловічого розуму, склався суто жіночий. Таким він залишається й донині. Поодинокі чоловічі постаті, які з'являлися зрідка на його небосхилі, надовго у відділі не затримувалися. Мабуть, важкувато було чоловікам-філологам виживати серед цього жіноцтва. Незвичного, нетипового, таких різних і зовні, і за характерами "філологинь", які аж ніяк не відповідали стереотипу жінки в науці. А можливо, це було спричинено й тим, що для ретельного аналізу великих масивів мовних фактів, які вимагали уваги до найменших дрібниць, і потрібні були саме жіноча терплячість, всидливість і хазяйновитість. У будь-якому разі ці люди зі своїми непростими біографіями, вдачами, слабкостями й особливостями незмінно були зайняті справою. Вона їх єднала, перед необхідністю її зробити завмирили всі житейські хвилі, немінучі у будь-якому, тим паче, творчому жіночому колективі. Не було тут "болотяності", "надування шік", бо для всього цього просто не лишалось ні часу, ні бажання.

Надихнений підтримкою з боку керівництва Інституту й Академії наук, колектив відділу взявся до праці, без перебільшення, новаторської як для української структурно-математичної лінгвістики, так і для української лексикографії. Досвід, набутий за попередні роки в лінгвостатистичному дослідженні українських текстів, вирішили узагальнити в новому для української мови типі словника — "Частотному словнику сучасної української художньої прози". Реалізувати такий масштабний задум на належному професійному рівні міг лише відданий справі, високопрофесійний колектив,

адже в Інституті тоді ще не було комп'ютерів, тож співробітники відділу уклали словник, опрацювавши (!) вручну вибірку текстів у 500 тис. слів.

Для підготовки словника відібрали твори 25 письменників, майстрів української прози 1945-1970 рр., різноманітні за будовою та тематикою, мовною палітрою, які б були показовими для стану української літературної мови цього періоду. 1981 р. "Частотний словник сучасної української мови" у 2-х томах вийшов друком у видавництві "Наукова думка". Він подав "портрет" української мови в художній прозі на базі близько 33 тис. різних слів — вірогідний активний лексикон українського мовця. 1998 р. на базі цього словника був укладений уже за допомогою комп'ютера та вийшов у паперовому вигляді однотомний "Обернений частотний словник сучасної української художньої прози" — різновид граматичного словника сучасної української мови з розгорнутою інформацією про особливості вживання у художніх прозових творах тих чи інших словоформ. Публікації цієї праці у київському видавництві "Спалах" посприяв фонд "Відродження" у рамках програми "Трансформація гуманітарної освіти в Україні".

Наступні десятиліття (1968-1988) позначені поєднанням фундаментальних теоретичних досліджень з експериментальною роботою: розроблення лінгвістичного та алгоритмічного забезпечення для систем автоматичного опрацювання текстової інформації, укладання українських словників нових типів. Ілюстрацією переходу від формалізованого моделювання мовних явищ, створення баз мовних даних до створення нових технологій опрацювання мови стала колективна монографія "Морфемна структура слова" (К., 1979), підготовлена під керівництвом *М.М. Пешак*. Ця оригінальна праця подала формально-логічні моделі морфемної будови слів різних частин мови, які уможливили встановлення їхніх словозмінних класів. Цю формальну модель української



словозміни згодом було покладено в основу алгоритмів і програм автоматичного редактора українського тексту РУТА та двомовного (українсько-російського й російсько-українського) перекладача ПЛАЙ. Їх у середині 1990-х років створили співробітники відділу *Т.О. Грязнухіна, Л.В. Орлова, Н.П. Дарчук, В.І. Критська, Т.К. Пуздирева* разом з програмістами *Т.І. Недозим, Г.В. Коленивим* та *В.М. Сорокіним*. Ці програми добре знають сьогодні всі, хто працює з українським офісом операційної системи WINDOWS фірми Майкрософт. Цю ж класифікацію змінюваних слів, доповнену розглядом їх унікальних, нерегулярних форм, використано й у "Граматичному словнику української мови: Словозміна", роботу над укладанням якого під керівництвом *Н.Ф. Клименко* щойно завершила група співробітників відділу *В.І. Критська* (керівник авторського колективу), *Л.В. Орлова, Т.К. Пуздирева, Ю.В. Романюк* і *Т.І. Недозим*. Відмітну рису цього словника з реєстром понад 145 тис. слів становить його операційність, тобто можливість завдяки формалізованим моделям словозміни автоматично генерувати всі властиві певному слову форми або, навпаки, зводити певну форму слова до його вихідної, словникової, форми. Словник укладено за допомогою комп'ютера, тож він становить не лише продукт багаторічного вивчення української словозміни, а й потужну базу знань про неї, комп'ютерну систему, придатну виконувати інформаційно-довідкові, навчальні, редакційно-видавничі, дослідницькі функції. Такий словник уперше з'являється в українській лексикографії.

3. 1980-ті: створюємо комп'ютерний фонд української мови!

Якісно новому витку комп'ютеризації мовознавчих досліджень у відділі в 1980-і роки чимало посприяла ідея створення комп'ютерного, або машинного фонду української мови. Її активно підтримав тодішній директор Інституту мовознавства — академік

В.М. Русанівський. Такі фонди мислилися як, з одного боку, скарбниці відомостей про будову та функціонування мови, а з другого, — як якісно нове фактичне підґрунтя для мовознавчих досліджень. Одним із перших ідей комп'ютерного фонду національної мови як фундаментальної автоматизованої системи, що містила б повний словник та генератор словоформ, а також формалізований тлумачний словник (тезаурус), висловив академік *А.П. Єршов*, фахівець у галузі створення діалогових систем спілкування з комп'ютером.

Відділ із 15 співробітників активно включився в роботу зі створення машинного фонду української мови. Група на чолі з *В.С. Перебийніс* почала активне формування баз російських та українських текстів і налагодження на їхній основі систем автоматичного опрацювання текстової інформації, передусім систем автоматичного редагування та перекладу. У цій групі впродовж кількох десятиліть працювали *М.П. Муравицька, Т.О. Грязнухіна, Н.П. Дарчук, Л.В. Орлова, В.І. Критська, Л.І. Комарова, Т.К. Пуздирева, Т.І. Недозим*, згодом до неї приєдналася *Ю.В. Романюк*. Ідеологію створення такої комплексної, багатоаспектної системи аналізу тексту, осмислення проблем, пов'язаних з підвищенням її ефективності, викладено в низці монографій 1980-1990-х років: "Закономірності структурної організації науково-реферативного тексту" (1982), "Автоматизація аналізу наукового тексту" (1984), "Лінгвістичні проблеми автоматизації редакційно-видавничих процесів" (1986), "Морфологічний аналіз наукового тексту на ЕОМ" (1990), "Синтаксичний аналіз наукового тексту на ЕОМ" (1999). З огляду на співпрацю з колегами за межами України, виконання спільних створення автоматизованих систем опрацювання текстової інформації ці праці видано російською мовою.

Інша група працівників відділу на чолі з *М.М. Пецак*, до складу її входили також *Н.Ф. Клименко, І.Ф. Савченко,*

Г.М. Ярун, Є.А. Карпіловська, Н.В. Сніжко, В.А. Шкурюв, І.В. Цимбалюк, від кінця 1970 рр. активно розробляла проблему формалізації аналізу семантики слів. Такі дослідження спиралися на ретельний аналіз структури метамов словників різного типу, передусім тлумачних. Їх результати узагальнено в колективних монографіях "Формалізовані основи семантичної класифікації лексики" (1982) та "Лексична семантика в системі "людина-машина" (1986).

За цими сухими скупими фактами — наукові будні й свята. Так, 1982-й — рік виходу зразу двох монографій, коли всі у відділі, від завідувачки до лаборантів, дружно різали, клеїли, нумерували, вичитували, креслили графіки й блок-схеми алгоритмів. Лаборанти поперемінно друкували обидві праці з нових напрямів, на яких тоді були зосереджені зусилля структуралістів і прикладників не лише України, а й тодішнього Радянського Союзу: створення формальних моделей семантики мовних одиниць та автоматизованих систем опрацювання текстової інформації.

Друкуючи монографії лаборанти ходили до спеціально обладнаного машбюро. Тут збиралася "могутня купка" інститутських лаборантів з усіх відділів. Колектив склався строкатий, язикатий, не відриваючись від перодруку наукових праць, встигали обговорити все на світі. Якщо в Інсти-туті і був "глас народу", то це був він, та, як казав тодішній директор Інституту академік *Іван Костевич Білодід*, "коридорна думка", яка часом істотно різнилася від "думки трибунної". У цьому середовищі теж були свої писані й неписані правила. Отож, неофітів вихували як згори, керівництво і старші колеги-науковці, так і знизу, колеги-рівні і ровесники. Цій школі життя, з роками розумієш це дедалі краще, вдячний кожний, хто її пройшов. Багато хто з колег — тодішніх лаборантів — сьогодні піднялися до вершин мовознавчої науки, стали докторами наук, професорами, визнаними фахівцями у своїх галузях, завідують кафедрами і відділами.



Для професійного зростання багато давали й щомісячні методологічні семінари, засідання відділу. Не менш цікавими від самих доповідей, що їх виголошували не лише науковці, а й лаборанти, були питання до доповідачів і дискусії. У таких дискусіях початківці опановували знаменитий "академічний стиль", коли сказати можна було все, що вважаєш за потрібне, але дівравши відповідної форми і за відповідної ситуації. Прикметно, і це було стилем ведення будь-якого засідання у відділі, що Валентина Сидорівна завжди починала саме з молоді. Висловлюватися ж мали всі співробітники. Відсидітися в кутку не вдавалося нікому, та й не хотілося. Навпаки, всі готувалися якнайретельніше. У питаннях філософії авторитет мали *М. П. Муравицька* й *Н. Ф. Клименко*. Знання іноземних мов (*В. С. Перейбініс*, за базовою освітою германіст, знала англійську, *Н. Ф. Клименко* — німецьку й новогрецьку, *М. М. Пецак* — угорську) допомагало стежити за іншомовною фаховою літературою (наскільки це було можливим у тодішній ситуації пильного партійного нагляду за "чистотою лав"), передусім, за дослідженнями в галузі структурної, математичної та прикладної лінгвістики. Питання були часом неочікувані, особливо славилася такими уділивими зауваженнями, які змушували замислюватися над речами, на які доти не звертав уваги, шукати переконливіших аргументів на оборону своєї думки, *Т. О. Грязнухіна*. Проте, яка б гаряча не розгорталася дискусія, вона завжди відбувалася в академічному стилі, з повагою до думки опонента, відсутністю крику й зацитування незгідних. Тон задавала Валентина Сидорівна й інші старші колеги. Сьогодні бракує таких зібрань, бракує такого постійного середовища для дискусій, обміну думками, апробації нашої праці, адже без цього наука мертвіє і, головне, роз'єднуються ті, хто її робить. І варто багато чого корисного зі старих часів повернути сьогодні до інститутського наукового життя, передусім те, що стосується

підготовки наукової зміни. Скажімо, раз на два роки в Інституті проводили конкурси молодіжних наукових праць. Авторитетна комісія з провідних фахівців різних відділів їх оцінювала. Найкраща робота здобувала рекомендацію для друку у журналі "Мовознавство", ті, які займали друге і третє місце, одержували премії дирекції. Такі конкурси чергувалися з конференціями молодих учених, які проводив Інститут мовознавства, де активно працювала Рада молодих учених під опікою дирекції і вченої ради Інституту. Виголошені найкращі доповіді виходили в академічному видавництві "Наукова думка" спеціальним збірником під грифом Інституту.

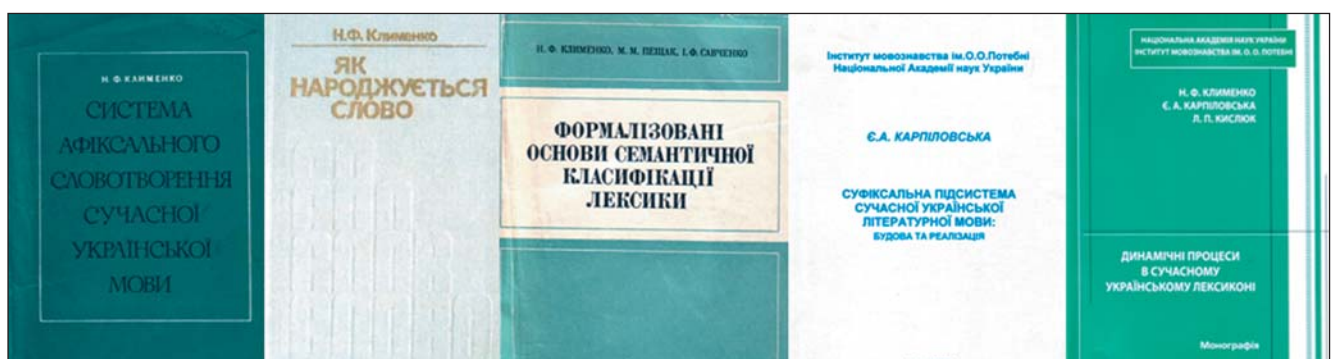
Накопичений досвід опрацювання проблем формалізованого опису лексики було втілено в проєкт комп'ютерного семантичного словника української мови ("Український семантичний словник: Проспект" (К., 1990, рос. мовою)). На жаль, цей масштабний задум залишився нереалізованим через відсутність стратегії формування основи семантичного словника — комп'ютерної версії 11-томного тлумачного "Словника української мови", придатної для його автоматичного укладання, а формувати його (як показав досвід наступних років) доцільно було починати з опрацювання тих рівнів системи мови, які найбільше надаються до формалізованого опису.

Саме таким шляхом з 1988 р. пішов новий колектив дослідників відділу структурно-математичної лінгвістики, очолений *Н. Ф. Клименко*. До його складу ввійшли *Є. А. Карпіловська*, *Л. І. Комарова*, *Н. В. Сніжко*. Пізніше до них приєдналася *Л. П. Кислюк*. У різний час з цим колективом працювали математики-програмісти *Т. І. Недозим*, *В. С. Карпіловський*, *Г. В. Коленов*, *С. Г. Бурігін*, *М. А. Перельмутер*. Дослідники поставили завдання: створити базу даних про будову сучасного українського слова та розробити засоби автоматизованого укладання на її основі нових типів українських словників. Створюваний комп'ютерний фонд дістав назву *морфемно-словотвірний*

фонд української мови. На початку 1990-х років завершено формування першої версії його стрижня — генерального реєстру українських слів, зведеного за матеріалами 5-ти авторитетних словників, різних за способом опису українського лексикону: 11-томного тлумачного "Словника української мови" (К., 1970–1980), 2-томного словника-довідника *І. Т. Яценка* "Морфемний аналіз" (К., 1980–1981), 2-томного "Частотного словника сучасної української художньої прози" (К., 1981), "Словника іншомовних слів" за редакцією *О. С. Мельничука* (К., 1974) та орфографічної частини "Словника-довідника з правопису та слововживання" *С. І. Головащюка* (К., 1989). На сьогодні у складі реєстру 171 304 слова з відомостями про їхню морфемну будову, частиномовну належність, кількість властивих їм значень, абсолютну частоту вживання у півмільйонній текстовій вибірці сучасної української художньої прози. На базі морфемно-словотвірного фонду було укладено за допомогою комп'ютера низку словників, нових для української мови: "Словник афіксальних морфем української мови" *Н. Ф. Клименко*, *Є. А. Карпіловської*, *В. С. Карпіловського*, *Т. І. Недозим* (К., 1998), "Кореневий гніздовий словник української мови" *Є. А. Карпіловської* (К., 2002), "Шкільний словотвірний словник української мови" *Н. Ф. Клименко*, *Є. А. Карпіловської*, *Л. П. Кислюк* (К., 2005).

4. Дорогу здолає той, хто йде

Упродовж майже півстоліття змінився і сам відділ, і загальна картина досліджень у галузі української комп'ютерної лінгвістики. Але в усі ці роки колектив, окрім значного доробку (7 словників, близько 50 колективних, індивідуальних монографій та збірників, понад 10 підручників та посібників для вищої школи, сотні статей в авторитетних українських і зарубіжних наукових виданнях) єднали і спільна справа, і спільні життєві інтереси. Усі знали, що тут знайдуть розуміння, почують мудру пораду: і як



дитину полікувати або наставити на путь істинну, і що зготувати на обід смачненького, і як статтю краще назвати. Відділ постійно жив якимось захопленням: то всі вивчали французьку мову в групі, яку організував для співробітників профком Інституту, то в обід усі оздоровлювалися, плаваючи у басейні стадіону "Динамо", знову ж таки під опікою профкому (тоді звичнішою була назва місцевком), то всі бралися до сиродіння і обідній стіл повнився вином з кульбаб, салатом з кореня коріандру, пророслим зерном і варенням із соснових голок, а бутерброди з ковбасою Валентина Сидорівна називала "провокацією". І ці спільні обіди, дні народження й свята, культпоходи до кіно, музеїв та театрів, просто задушевні розмови про все на світі теж були важливим компонентом життя відділу, тим, що його згуртовувало і виробляло той стрижень, який дав йому змогу вижити в глухії 1990-і роки, і не тільки вижити, а повноцінно працювати, готувати й видавати книжки, створити системи РУТА і ПЛАЙ, стрижень, що тримає його і нині.

1991 р. завдяки проекту, який за дорученням вченої ради Інституту підготували співробітники відділу, на його базі було створено науково-допоміжну інституцію в складі Відділення мови, літератури, мистецтвознавства НАН України — Мовно-інформаційний фонд (МІФ, нині — УМІФ) на чолі з *В.А. Широковим*. Основним завданням цієї нової академічної структури було створення надійної фактографічної основи для виконання фундаментальних мовознавчих досліджень, а саме: формування корпусів текстів, словникових баз, бібліотек і архівів, а також комп'ютерних технологій їх опрацювання. З цієї метою фонд передусім упорядкували кадрами інформатиків і математиків-програмістів, а з відділу працювати до нього перейшли науковці *М.М. Пещак* і *Г.М. Ярун*, програміст *І.В. Цимбалюк* та лаборант *В.А. Шкуров*. Відділ підтримав

створення цього фонду і завжди був відкритий для професійного спілкування з ним, як і з іншими колегами.

Не все в українській комп'ютерній лінгвістиці склалося за ці роки так, як бачилося в романтичну еру її становлення. Багато планів так і залишилися нереалізованими. Українське суспільство на сьогодні ще, на превеликий жаль, не одержало тих потужних українськомовних продуктів в комп'ютерному середовищі, на які воно з повним правом очікувало від академічної науки. Проте відділ попри всі перешкоди продовжує рухати вперед свою справу. Триває дослідження інноваційних процесів у сучасній українській мові, вивчення можливостей їх комп'ютерного моделювання й створення українських словників нового покоління: йде укладання нового для української лексикографії словника: "Словника активних ресурсів сучасної української номінації". Базу для його створення становить сформований у відділі комп'ютерний фонд інновацій в сучасній українській мові (параметризована база нової української лексики та корпус мікроконтекстів). На сьогодні фонд охоплює близько 18 тис. слів і словосполук, не засвідчених словниками радянської доби. Щоб з'ясувати стійкість й значущість кожного нового найменування для свідомості сучасного українця, розроблено оригінальну методику визначення активності інновацій у системі мови і в текстах. Практичній роботі над словником передувало теоретичне осмислення змін у лексиконі та граматиці української мови, з'ясування тенденцій її розвитку в нових суспільно-політичних обставинах у монографії *Н.Ф. Клименко*, *Є.А. Карпіловської* та *Л.П. Кислюк* "Динамічні процеси в сучасному українському лексиконі" (2008) з покажчиком у понад 4 тис. слів, що не ввійшли до реєстрів словників радянської доби.

Співробітники впроваджують ідеї комп'ютерної лінгвістики у систему підготовки філологів у вищій школі,

співпрацюють з університетськими кафедрами прикладної лінгвістики, яких нині в Україні вже майже два десятки: читають лекційні курси, видають підручники і посібники з комп'ютерної лінгвістики ("Використання ЕОМ у лінгвістичних дослідженнях" (рос. мовою, 1990), "Нариси з комп'ютерної лінгвістики" *М.М. Пещак* (1999), "Статистичні методи для лінгвістів" *В.С. Перебийніс* (2002), "Вступ до комп'ютерної лінгвістики" *Є.А. Карпіловської* (2003, 2006), "Комп'ютерна лінгвістика" *Н.П. Дарчук* (2008), "Традиційна та комп'ютерна лексикографія" *В.С. Перебийніс* і *В.М. Сорокіна* (2009), "Прикладна морфологія" *Ю.В. Романюк* (2010)). Де б не працювали нині наші колишні співробітники, вони усвідомлюють себе членами спілки "пташенят Перебийніс", зберігають вірність відділу, не поривають із ним зв'язків (фото).

На початку 2010 року ВАК України затвердила нову програму кандидатського іспиту зі спеціальності 10.02.21 — структурна, прикладна та математична лінгвістика, яку розробив колектив відділу. За нею вже почали складати іспити аспіранти кафедр прикладної лінгвістики університетів, отже, з'явилася реальна можливість відновити й захист дисертацій з цієї спеціальності. Відділ очолив робочу групу при Науково-методичній раді напряму "Філологія" Міністерства освіти та науки України у підготовці стандарту зі спеціальності "Прикладна лінгвістика" для кваліфікаційного рівня "Бакалавр". Таке відродження нашої спеціальності в університетах України після років її відвертого занеухання дає надію на кращі часи для української комп'ютерної лінгвістики.

Дорогу здолає той, хто йде.

Жодна справа не посувається вперед без перешкод, проблем, складнощів, а запорука успіху — злагоджена, дружня праця всіх зацікавлених у реальному її результаті.

