

Iryda Grek-Pabisowa

BADANIE SŁOWNICTWA GWAROWEGO PRZY POMOCY METOD SYSTEMOWYCH
(NA PRZYKŁADZIE BADANIA SŁOWNICTWA
ROSYJSKIEJ GWARY STAROBRZĘDOWCÓW W POLSCE)

1. WSTĘP

Gwara enklawy starobrzędowców-Rosjan istniejącej w Polsce od ok. 200 lat budzi zainteresowanie nie tylko ze względu na zachowane archaizmy, lecz także ze względu na zmiany, jakie zaszły w niej pod wpływem obcego otoczenia językowego. Gwara nie straciła przy tym swych cech charakterystycznych pozwalających dokładnie określić jej typ i przynależność do odpowiedniej grupy narzeczy wielkoruskich¹.

Słownictwo gwary jest niezwykle bogate²; charakteryzuje się dużą wariacyjnością postaci zewnętrznej wyrazów, obocznym występowaniem łańcuchów równoznaczników oraz żywym procesem tworzenia się i eliminowania homonimów. Wszechstronne przedstawienie słownictwa gwary wymagało, prócz ustalenia inwentarza leksemów, ukazania stopnia nasycenia gwary zapożyczeniami dawnymi i nowymi, ustalenia przyczyn wahanía postaci zewnętrznej wyrazów oraz stopnia zbliżenia gwary do języka ogólnorosyjskiego. Należało także ustalić przyczyny i źródła wariacyjności wyrazów, rozwiniętej synonimii i homonimii i wskazać mechanizmy, które te zjawiska wywołały. Niewystarczające dla tego celu byłyby badania typu

¹ Opisowi gwary, określeniu jej typu i przynależności jest poświęcona książka autorki: I. G r e k - P a b i s o w a, *Rosyjska gwara starowierców w województwach olsztyńskim i białostockim*, Wrocław 1968.

² Słownik tej gwary wydany w 1980 r. liczy ponad 10 000 haseł - zob. I. G r e k - P a b i s o w a, I. M a r y n i a k o w a, *Słownik gwary starowierców mieszkających w Polsce*, Wrocław 1980.

słownikowego. Dlatego też w pracy poświęconej pokazaniu wyżej wymienionych cech słownictwa gwary starowierców³ posłużyłam się metodą, którą poniżej przedstawiam.

2. METODA

Rosjanie-starobrzędowcy mieszkają w Polsce w trzech skupiskach, w Suwalsko-sejneńskim, Augustowskim i na Mazurach⁴. We wszystkich skupiskach mówi się tą samą gwarą, ale różne otoczenie językowe i różne warunki historyczne i cywilizacyjne odbiły się w sposób zauważalny na ich mowie. Dlatego przy opracowywaniu materiału należało uwzględnić jego geografię. Do badań zostało wzięte słownictwo z zakresu tematycznego "Człowiek, jego życie fizyczne i duchowe". Materiał liczy 4280 leksemów, z uwzględnieniem powtarzających się w różnych skupiskach - 6500 (z S pochodzą 2032 wyrazy, z A - 2145, z M - 2290). We wszystkich skupiskach powtarza się 2189 wyrazów. Część wyrazów powtarza się tylko w dwu skupiskach. Materiał pochodzi głównie z tekstów ciągłych nagranych na taśmę magnetofonową, w małej części - z odpowiedzi kwestionariuszowych.

Operowanie tak znaczną liczbą fiszek i konieczność wielokrotnego klasyfikowania materiału dla zrealizowania założeń pracy wpłynęły na wybór metody. Należało bowiem całe zebrane słownictwo rozpatrzeć kompleksowo badając współzależności zbiorów słów. W takim ujęciu przedmiot badań można było traktować jako system. Odwołałam się więc do metod teorii systemów, której przedmiotem jest badanie systemów i ich wzajemnego oddziaływania. System jest to zbiór obiektów (w naszym wypadku zbiór słów), którego strukturę tworzą związki między tymi obiektami (tu - wzajemne związki między wyrazami danego zbioru⁵). Metody teorii systemów pozwalają na dokonywanie różnych ich podziałów w zależności od celów pracy. Badany system można rozpatrywać jako umiejscowiony wśród

³ I. G r e k - P a b i s o w a, *Słownictwo rosyjskiej wyspy gwarowej starobrzędowców w Polsce. Rozwój i stan dzisiejszy*, Wrocław 1983.

⁴ W dalszym ciągu skupiska te będą oznaczane skrótami: S = Suwalsko-sejneńskie, A = Augustowskie, M = Mazury.

⁵ Związki między wyrazami mogą być wielorakie. Tu wystarczającym kryterium wzajemnego powiązania wyrazów danego zbioru jest fakt, że opisują one ten sam temat.

innych systemów tworzących jego otoczenie. W danym wypadku otoczenie tworzą języki obcego terytorium etnicznego, na którym znalazła się gwarą. W skupisku S jest to język polski i częściowo litewski, w A - język polski, w M - język polski i język niemiecki. Jedną z cech otoczenia jest poziom cywilizacji. W wyniku oddziaływania otoczenia mamy taką sytuację, że możliwe jest w tej samej gwarze w różnych skupiskach oboczne lub wyłączne używanie różnych nazw tego samego desygnatu, np. o tej samej chorobie mówi się *t'if*, w A - *o'if*, a na M - *t'yfus*; 'bratanek' w S i A to *brat'en'ik*, w M *ad br'ate syn*; 'denaturat' w S to *s'i-n' ak||s'in''uxə*, w A - *denat'ura*, w M - *brand'uxə brand'ul'ə*.

Przeprowadzona analiza uzasadnia zastosowanie metod teorii systemów, a ponad sześciotysięczny zbiór fiszek - skorzystanie z techniki komputerowej. Technikę komputerową zastosowano do wyodrębniania różnego rodzaju podzbiorów (np. podzbiory według części mowy, podzbiory wyrazów zapożyczonych, wyrazów i ich wariantów formalnych itp.), sortowania oraz przeprowadzania na nich innych operacji. Pozwoliło to m. in. na:

- ograniczenie subiektywnej oceny badanego materiału;
- zachowanie jednolitości oceny;
- kilkusetkrotne skrócenie czasu sortowania fiszek;
- otrzymanie gotowych zestawów słownictwa żądanej kategorii, dzięki czemu można było uniknąć przepisywania tych zestawów.

3. SPOSÓB

Wstępną czynnością było ustalenie zestawu danych, które należało uzyskać z materiału. Dla tego zestawu został opracowany odpowiedni program. Tu konieczna była współpraca z fachowymi programistami. W programie ujęto konieczność uwzględniania wielowariantowych postaci poszczególnych leksemów, podawania szeregów synonimów i równoznaczników, podawania homonimów, sortowania według skupisk, pochodzenia, kategorii gramatycznych itp. Wprowadzono też zestaw znaków do drukowania wyników w transkrypcji.

Po opracowaniu zestawu cech niosących określone informacje potrzebne do udokumentowania założeń pracy ocechowano nimi każdy wyraz całego zbioru. Uwzględniłam następujące cechy: 1) przynależność do dziedziny mieszczącej się w zakresie tematycznym "Człó-

wiek, jego życie fizyczne i duchowe" - weszły tu takie dziedziny, jak: części ciała, zdrowie człowieka, czynności i stany fizyczne, stany duchowe, choroby i ich leczenie, żywienie, nazwy pokrewieństwa, przezwiska, cechy fizyczne człowieka, życie religijne itp., łącznie 17 dziedzin; 2) numer kolejny karty (fiszki) w dziedzinie - otrzymuje się w ten sposób liczbę wszystkich wyrazów z danej dziedziny; 3) czy wyraz ma synonimy i równoznaczniki; 4) numer kolejny synonimu lub równoznacznika w szeregu synonimicznym; 5) rodzaj synonimu, tzn. czy jest to synonim właściwy, czy równoznacznik; 6) czy wyraz ma warianty (fonetyczne lub morfologiczne); 7) numer kolejny wariantu danego lekseru; 8) jakiego pochodzenia jest wyraz - ogólnorosyjski, gwarowy, polonizm, germanizm, inny; 9) czy wyraz zapożyczony jest zasymilowany czy nie; 10) czy wyraz ma jedno czy więcej znaczeń; 11) w jakim skupisku wyraz występuje - uwzględniono następujące możliwości: wyraz występuje w M, A i S, występuje tylko w M, tylko w A, tylko w S, występuje w M i A, w M i S, w A i S; 12) czy wyraz niezależnie od pochodzenia: a) wykazuje zniekształcenia fonetyczne w stosunku do systemu fonetycznego gwary, np. *l'ig'istra-v'acse* (wobec *reg'istrav'acse*), *mask* (wobec *muak* 'mózg') itp.; b) czy miejsce akcentu wyrazu zapożyczonego odbiega od odpowiedniego wyrazu w języku dawcy, np. *v'id'ok* (wobec *wídok*), *l'istan'oš* (wobec *listónosz*), *ban'of* (wobec *Báhnhof*); c) czy wyraz zapożyczony zachowuje rodzaj gramatyczny odpowiedniego wyrazu w języku dawcy, np. *padn'eb''en'* m (pol. *podniebienie* n), *sanat'or'ije* ž (wobec pol. *sanatorium* n, ros. *sanatorij* m, niem. *das Sanatorium*), *r'sv'ije* ž 'rewir' (pol. *rewir* m, niem. *das Revier*) itp.; 13) czy wyraz występujący w gwarze wykazuje przypadkową zbieżność brzmienia przy różnym znaczeniu z wyrazem ogólnorosyjskim, np. *atk'i-nut'* 'podwieźć' (wobec ros. *otkinut'* 'odrzucić'), *d'afke* 'dawka, doza' (wobec ros. *davka* 'tłok, ścisk'), *trut* 'trucizna' (wobec ros. *trud* 'praca') itp.; pozwoliło to na ustalenie homonimów; 14) nacechowanie chronologiczne wyrazów: a) wyraz jest archaizmem; b) wyraz pojawił się w gwarze niedawno (po drugiej wojnie światowej); c) wyraz jest nienacechowany; 15) podział wyrazów na części mowy; 16) podział wewnątrz części mowy na poszczególne kategorie gramatyczne - rodzaj, liczba, aspekt itp.; 17) czy w formach zależnych wyraz wykazuje odchylenia od norm ogólnorosyjskich; 18) czy akcent w wyrazach ogólnorosyjskich i gwarowych

jest zgodny z akcentem odpowiednich wyrazów we współczesnym języku ogólnorosyjskim (por. też wyżej 12b); 19) czy wyraz jest zdrobnieniem; 20) czy w wyrazie znajdujemy afiks mało charakterystyczny lub nie występujący w języku ogólnorosyjskim w danej kategorii wyrazów, np. *pa'ub'ovni'ice* (ros. *ljubovnica*), *pa'av'a-r'išš* (ros. *tovarišč*), *s'iv'u'l'a*, *ikt'afka* (ogros. *ikota*), *pa'za-str'el'ivat'* (ros. *zastrelit'*), *pa'v'er'et'* (ros. *poneret'*), *pa-s-p'or'ice* (ros. *posporit'*), *ma'uš'pnyj* (ros. *malen'kij*, *maljusen'kij*) itp.

W wyniku wprowadzenia do komputera zbioru wyrazów z zakodowanymi informacjami otrzymałam listę generalną, czyli pełny wydruk badanego słownictwa. Po skorygowaniu listy na jej podstawie można było otrzymywać dalsze wydruki. Wydruki te stanowiły odpowiedzi na zadane komputerowi polecenia. Przy każdym poleceniu konieczne było podanie sposobu sortowania. Oto kilka przykładowych poleceń: np. wydruk wszystkich wyrazów z podziałem na dziedziny, wydruk wyrazów z podziałem na skupiska, pochodzenie wyrazów w poszczególnych skupiskach, wyrazy mające synonimy i równoznaczniki, wydruk szeregów synonimicznych, zapożyczenia dawne i nowe w poszczególnych skupiskach itp.

Analiza i porównanie otrzymanych wydruków umożliwiły sformułowanie szeregu stwierdzeń i wniosków dotyczących słownictwa badanej gwary.

4. REZULTATY

Na podstawie otrzymanych wydruków można było określić skład i pochodzenie słownictwa całej gwary i słownictwa w poszczególnych skupiskach.

Okazało się, że słownictwo rosyjskie (ogólnorosyjskie i gwarowe) wynosi w gwarze średnio 67%, a w poszczególnych skupiskach waha się od 73% w A poprzez 69% w S do 59% w M. Słownictwo ogólnorosyjskie wynosi w gwarze średnio 50%. Świadczy to o tym, że leksyka gwary jest w dużym stopniu zbliżona do języka ogólnonarodowego. Udział dialektyzmów w gwarach rosyjskich na terytorium Rosji szacuje się na ok. 32,5%⁶, a tzw. dialektyzmów

⁶ Zob. N. A. L u k j a n o v a, *Sistemye otnošenija v leksike govorov Syzynskogo i Ordynskogo rajonov Novosibirskoj oblasti*, Novosibirsk 1966.

lokalnych na 4-6%⁷. Dane uzyskane przeze mnie są dość zbliżone. W gwarze staroobrzędowców w Polsce mamy 22% dialektyzmów i 4,3% dialektyzmów lokalnych. Jak widzimy, w obcym otoczeniu językowym gwara zachowuje dialektyzmy lokalne, natomiast wyrazy o wspólnych z ogólnorosyjskimi rdzeniami tracą częściowo nacechowanie gwarowe.

Gwara w znacznym stopniu wzbogaciła swe zasoby leksykalne dzięki czerpaniu zapożyczeń z języka otoczenia, głównie z języka polskiego. Zapozyczenia polskie sięgają w M i S 30%, zaś w A przekraczają 25%. Składają się na to zarówno zapożyczenia nabyte po opuszczeniu terenów macierzystych, jak i te, które w gwarze istniały już dawniej. Jak widać, język polski był, przynajmniej od wieku XVII, źródłem wzbogacającym gwary rosyjskie zachodniej strefy obszaru narzeczy wielkoruskich, zwłaszcza gwary pskowskie. Wśród zapożyczeń dawnych pożyczki polskie znacznie przewyższają zapożyczenia innego pochodzenia. Wynoszą one 17% wszystkich zapożyczeń, podczas gdy dawne germanizmy i bałtyzmy - zaledwie ułamek procenta.

Z otrzymanych danych można było wysnuć wniosek, że zasób leksykalny gwary bardzo szybko się zmienia. W ciągu ostatnich 170 lat pojawiło się w gwarze bardzo dużo nowych wyrazów, zarówno zapożyczeń, jak i wyrazów rosyjskich, szacowanych na 33%. Najwięcej zmian mamy na Mazurach, gdzie do zapożyczeń polskich doszły zapożyczenia niemieckie. Wpływ języka niemieckiego na mowę staroobrzędowców na Mazurach nie był tak znaczny, jak można było tego się spodziewać. Germanizmy wynoszą w tym skupisku 11%, a polonizmy aż 29%. Bariera językowa była tu, jak widać, skuteczną obroną przed napływem wyrazów niemieckich.

W gwarze trwa wzbogacanie się zasobu leksykalnego. Wyrazów nowych, które mogły pojawić się w gwarze po drugiej wojnie światowej, mamy 8%.

Odpowiednie wydruki wykazały, że gwara jest szczególnie bogata w synonimy i równoznaczniki. W badanej próbce, zawierającej ok. 40% słownictwa gwary, znalazło się 578 szeregów, w tym 131

⁷ Zob. G. G. M e l n i č e n k o, *K voprosu o zadačach i putjach processov, nabljudaemych v leksike govorov russkogo jazyka na sovremennom etapie ich razvitiia v sovetskom obščestvie*, [w:] *Tezisy dokladov na X dialektologičeskom soveščanii 11-14 maja 1965 g.*, Moskva 1965, s. 21-23.

szeregów synonimów i 447 równoznaczników. Przy porównywaniu tych danych z danymi z rosyjskich gwar na terenie Rosji wynika, że synonimiczność i obecność równoznaczników jest cechą szczególnie charakterystyczną dla wysp gwarowych. Dla różnych gwar w Rosji podaje się⁸ bowiem od 80 do 215 szeregów synonimów, a opracowany słownik synonimów i równoznaczników gwar z okolic Nowosybirsk⁹ liczy 300 szeregów.

Wydruki pokazały także szczególną wariancyjność, tj. niestabilną postać zewnętrzną wyrazów gwarowych. Łańcuchy wariantów mogą liczyć po osiem ogniw. Są to obocznie występujące warianty fonetyczne, słowotwórcze, fleksyjne i ekspresywne.

Pod względem danych o wariancyjności, obecności synonimów, równoznaczników i homonimów w słownictwie poszczególnych skupisk wyniki komputerowe były podobne.

Otrzymałam dane, na podstawie których można było stwierdzić, że w gwarze nie rozwija się polisemia. Przeważają wyrazy o jednym znaczeniu - jest ich 92%. Wśród wyrazów polisemicznych większość ma tylko dwa znaczenia. Tę ograniczoną polisemię (uwarunkowaną zresztą takimi czynnikami, jak brak kontaktu z rodzimym językiem literackim, brak pisemnej postaci gwary itp.), rekompensuje gwara czerpiąc z języków otoczenia oraz rozwijając synonimię.

Porównanie odpowiednich wydruków umożliwiło wskazanie źródeł wariancyjności i synonimii w gwarze, opracowanie sposobów przyswajania przez gwara zapożyczeń i udokumentowanego przedstawienia szeregu zjawisk w gwarze.

Pokazana tu metoda może być stosowana nie tylko do badania słownictwa. Może być wykorzystana do badań różnych poziomów języka, jeśli tylko jest to uzasadnione obszernością materiału. Opracowany do moich badań program może służyć również innym pracom

⁸ Zob. S. R. Kačinskaja, *O nekotorych osobennostjach leksičeskich sinonimov v južnych govorach Gorkovskoj oblasti*, [w:] *Problemy russkogo jazykoznanija*, Moskva 1970; Ł. K. Żyżova, *Ob osobennostjach sinonimov raznych častej reči*, "Izvestija Voronežskogo pedagogičeskogo instituta" LXVIII, 1969; V. V. Palagina, *Sovremennij govor starožiličeskogo naselenija zapadnoj časti Tomskogo rajona Tomskoj oblasti*, Tomsk 1951; S. S. Sol'tan, *Pričiny vozniknovenija dialektnoj sinonimii*, [w:] *Sinonimija v jazyke i reči*, Novosibirsk 1970, s. 187-192.

⁹ Zob. przypis 6.

leksykalnym, a także, z małymi uzupełnieniami - słowotwórczym i fleksyjnym. Program ten może być udostępniony osobom zainteresowanym.

Przedstawiając zastosowaną przeze mnie metodę i technikę chciałam pokazać, iż poza dyscyplinami tradycyjnymi dialektologia może owocnie współpracować także z dyscyplinami nowszymi. Technika komputerowa jako instrument metod systemowych daje się w dialektologii stosować z powodzeniem. Pomaga ona i znacznie przyspiesza otrzymywanie rezultatów badań w postaci czytelnych dla lingwisty danych, gdyż prócz wskaźników statystycznych uzyskujemy żądane listy wyrazów.

Polska Akademia Nauk
Warszawa

Iryda Grek-Pabisowa

SYSTEM ANALYSIS METHOD OF DIALECTAL LEXICON RESEARCH
(EXAMPLE: A STUDY OF THE VOCABULARY
OF THE RUSSIAN ORTHODOX CHURCH DIALECT IN POLAND)

The paper presents a system analysis and computer aided research of the lexicon of a Russian dialectal packet (in Poland). The vocabulary referring to a chosen discipline (covering over 6000 words) was indexed with respect to various indispensable features like lexical geography, etymology, chronology, phonetics, word-formation, inflection, stability of the surface form etc. The participation of words in a given cluster of synonyms, their stylistic and other features were also indexed. In all, every word was assigned over 20 indices. Thus prepared material was then fed into a computer and sorted out according to a program worked out for this purpose.

The process of sorting out yielded a few dozen printouts (lists) of words. The lists contain sets of words grouped according to a preprogrammed distribution pattern. Thus, in addition to production of quantitative data, the whole material was adequately arranged. The lists made possible a relatively precise determination of lexical differences in particular dialectal centers, of the quantity and type of innovations, of the degree to which the native vocabulary was preserved and the degree of influence from neighbouring languages. Especially visible was the lack of stability of the external form of words (variability) as well as the presence of numerous synonyms and words designating the same objects. A comparative analysis of particular printouts has revealed the mechanisms of linguistic processes in dialects.