

Hauke Bartels

## Ku koncepciji historisko-dokumentěrujucego informaciskego systema za dolnoserbški slowoskład\* Plany k wótpóranju nuznego slěžeńskego deziderata

### 1. Wuchadnišćo

Jo znate, až dolnoserbščina licy k nejžmócnjej wobgrozonym rěcam w Europje. Stakim njejo jeje wobšyrna dokumentacija jano za slawistiku wažna. Chósebuska wótnožka za dolnoserbške slěženja Serbskego instituta zaběra se togodla južo lěta dlujko z tym, aby tak wjele ako móžno informacijow wó toš tej rěcy nazběrała a za wšake zaměry wobžělała. Na pólu leksiki dej se pšez systematiske wugódnósenje wšykných k dispoziciji stojecých žrědłow dojspíš pó móžnosći poľne leksikografiske wopisanje, kótarež dokradnje dokumentěrujo hyšći znate abo něži zapisane dolnoserbške slownistwo. Koncept togo pšedewzeša, kótaremuž se nimski groni *SORBLEX NIEDERSORBISCH*<sup>1</sup>, pšedstajijo se how w głowných cerach.

We zachadnem lětzasetku smy se krađu procowali wó natwarjenje datoweje bazy, ako jo notna (mjazy drugim) za nadrobnu leksikografisku dokumentaciju dolnoserbškeje rěcy. Wšake projekty slědnych lět su južo direktnje slůžyli k pšigótowanju tudy wopisanego pšedewzeša (pšeglěd dajo BARTELS 2012). Ako wuslědk móžomy se za *SORBLEX* mjaztym zepěraš na pisnorěcny elektroniski tekstowy korpus<sup>2</sup> a digitalne wobžělanja wažnych dolnoserbško-nimskich slownikow (ZWAHR 1847; MUKA 1911–15/1926, 1928; ŠWJELA 1961; STAROSTA 1999). K tomu pšišo wjeliki datowy wobstatk Nimsko-dolnoserbškego slownika w interneše (DNW).<sup>3</sup> Dalej žělamy na zukowem a tekstowem archiwje dialektalneje dolnoserbščiny<sup>4</sup>, ako jo žinsa hyšći słyšaš na jsach na pódpóľnoc wót Chósebuza (wěcej k tomu w BARTELS/THORQUINDT-STUMPF 2013).

\* Toš ten tekst jo pšeložk nimski napisanego artikla, kótaryž se wózwajijo we zběrniku nimskich pšinoskow za 15. mjazynarodny kongres slawistow: KEMPGEN, S.; FRANZ, N.; JAKIŠA, M.; WINGENDER, M. (Hrsg.): *Deutsche Beiträge zum 15. Internationalen Slavistenkongress, Minsk 2013*. München, Berlin: Verlag Otto Sagner. – Žěkujom se Gunteroju Spießoju za plodne diskusije pši wužělanju koncepcije a Fabianoju Kaulfürstoju za porážowanje pši pšestajenju teksta.

<sup>1</sup> Aby jadnorrej byľo, pišom w slědujućem jano hyšći *SORBLEX* (mimo dodanka *NIEDERSORBISCH*) a „slownik“ (město wuraza „informaciski system za slowoskład“), lěcrownož se zapšimješe „slownik“ skerzej njezwěžujo z dodatnymi kakosćami, kótarež pomjenjony informaciski system móžo póbitowaš.

<sup>2</sup> Nětejšny korpus (pšir. BARTELS 2010) južo wopšimjejo wětšy žěl zachowaných tekstow a móžo se w pšiducnych lětach nejskerrej kšac pó kšacu wutwarjowaš na poľny korpus, ako cele znate pismojstwo reprezentěrujo. Dalšne techniske a rěcywědne wobžělanje korpusa jo notne za jago efektiwne wužywanje a jo togodla kšuty wobstatk wopisanego projekta.

<sup>3</sup> Wše naspomnjete resurse su pšisťupne pód internetoweju adresu <http://www.dolnoserbški.de>.

<sup>4</sup> Pši tom se jedna wó slěžeński projekt, kótaryž financěrujo se wót kónca 2011 pšez zaľožbu VolkswagenStiftung we ramiku naglědneho spěchowańskego programa „Dokumentation bedrohter Sprachen (DoBeS)“.

Z tymi a dalšnymi<sup>5</sup> resursami mamy abo skóro změjomy písamem optimalny datowy zakład za planowany projekt. Stakim *SORBLEX* by był přědny słownik dolnoserbšćiny, kótaryž móžo póbitowaš písamem<sup>6</sup> dopolne wopisanje leksiki dolnoserbškeje pisneje rěcy a kótaryž wušej togo tak daloko ako móžno teke zapšimjejo słowoskład dolnoserbških dialektow.

## 2. Ku koncepciji *SORBLEXa*

Projekt *SORBLEX* pšedstajijo słownik nowego typu, kótaryž drje by se přědny raz wuželał za „mału“ europejsku rěc ako dolnoserbšćinu. Wón jo koncipěrowany ako wědomnostne leksikografisko-leksikologiske wopisanje dolnoserbškeje rěcy a by był přědny dolnoserbški słownik, kenž móžo se zepěraš na wugódnosnje písamem celego pismojstwa we formje elektroniskego tekstowego korpusa. Teke wše donětejšne inventarizacije a wopisanja dolnoserbškego słowoskłada (słowniki, słowne lišćiny atd.; pšir. 2.3.1.) se zapšimjeju. Dodatnje se žiwa na zaklaže wšykných k dispoziciji stojecých žrědłow na dolnoserbške dialekty. Pšedmjat projekta jo pótakem dolnoserbška rěc wót casa nejstaršých rěcnych pomnikow z přědneje polojcy 16. stolěša až do pšibytnosći.

*SORBLEX* zwězujšo pšiznamjenja historiskego słownika z tymi synchroniskego wopisanja dolnoserbšćiny, ako znajomy ju někak wót kónca 19. stolěša. Z tym móžo słownik přědny raz registrěrowaš słownistwo we swójej diachroniskej dimensiji, rownocasnje pak teke dowólijo pšez pšezěžlanje a rozšyrjenje na pš. Starostowego słownika (1999), taki pšistup na žinsa wužywany słowoskład, ako jo znaty z powšykneho dolnoserbško-nimskego słownika.

Internet bužo primarny medium za prezentaciju słownikowych informacijow. Cely datowy wobstatk se wopisujšo we pšez kompjuter cytajobnem formaše XML ([www.w3.org/XML](http://www.w3.org/XML)). Z togo slěduju wšakorake móžnosći pšedstajenja datow, kótarež comy do dalokeje měry wužywaš, jolic su leksikografiski zmysłapólne.

### 2.1. Makrostruktura

Zapšimješe makrostruktura implicěrujo mimo wobstatka lematow teke wěsty leksikografiski řěd, na pš. alfabetiski. Dokulaž bazěrujo how wopisany słownik na wšakich, mjazy sobu kompatibelnych řědowańskich strukturach, tematizěrujo se toš ten aspekt makrostruktury w kapitlu 2.2.

Status ako lema we *SORBLEXu* změju (cysto pokazkowe lemata sobu licone):

1. Wše dolnoserbške wuraze, kótarež se ako lema wšakorakego typu (to groni głowny, pódlański, gnězdowy, pokazkowy) namakaju w digitalizěrowanych słownikach a słowných lišćinach, a to w originalnem pisanju žrědła (na pš. *hopazny / hopacny / wopacny*). Gaž jo notne, dodajomy normalizěrowane pisanje pó žinsajšnej ortografiji.

<sup>5</sup> To nastupa na pš. Serbski rěcny atlas (SSA 1965–96), ako we swójom leksikaliskem žěle wopšimjejo górnjo- a dolnoserbške (dialektalne) wótpowědniki jadnab 1300 nimskich słow. Pšir. teke nožku 21.

<sup>6</sup> Njezglědajucy na celo zgubjone pismojstwo dajo na pš. někotare doněta wědomnostnje nje-edeřowane rukopise, ako pó něntejšnem stawje njamóžomy zapšěgnuš.

2. Wše dolnosěrbske wuraze, kótarež w digitalizěrowanych słownikach njejsu lemata, kótarež pak dejali taki status dostaš (tak pomjenjone „schowane lemata“).<sup>7</sup>
3. Wše dolnosěrbske wuraze, kótarež we słowniku DNW služe ako ekwiwalenty za nimske lemata. (K słownym wězbam w toš tej funkciji gl. dołojce.)
4. Wše dolnosěrbske wuraze z elektroniskich korpusow dolnosěrbsčiny<sup>8</sup>, kótarež dejali dostaš status ako lema (pó za *SORBLEX* płašecych směrnicach, ako muse se w přédnej fazy projekta hyšći precizěrowaš).
5. Žěl we słownikach zidentificěrowanych a we korpusach wustupujucych kšutych słownych wězbow<sup>9</sup>, kótarež su (pó žělach) idiomatiske.

Kšute słowne wězby, kótarež pomjeniju zapšimješe a graju we dolnosěrbsčinje wažnu rolu ako wótpowědniki nimskich nominalnych kompozitumow (na pš. *pisańske blido* za *Schreibtisch*) a teke (druge) kolokacije se tuchylu zapšimjeju jano tak daloko, ako su we wugódnosonych dolnosěrbsko-nimskich słownikach zapisane (nejcesćej we formje „pšikłada“, rědšej ako lemata) abo wustupuju we słowniku DNW ako ekwiwalenty za nimske lemata (pšedewšym za kompozitumy).<sup>10</sup> Toš te wuraze pak zwětšego njedostanu status ako lema. Njejo lažko, dokradnje wótgranicowaš idiomatiske wurazy ze statusom ako lema (jasne pady su celo njetransparentne słowne wězby) wót takich, kótarež wopišomy w mikrostrukturje artikla wó jadnom z elementow takeje wězby. To dej se w ramiku celkownego pšedewzeša realizěrowaš.

## 2.2. Wó ředowanju slowoskłada

We šišćanych słownikach jo doněta notne było se rozsuziš mjazy alfabetiskim ředom a drugimi móžnosćami ředowanja, kótarež glědaju na imanentne kakosći celkownego słownistwa (słowne familije, słowne póla, rědšej: słowne družyny). Tradicionelnje jo se wubrał nejcesćej alfabetiski princip, w jadnotliwych padach teke j a d n a z drugih móžnosćow. Toš ta trěbnosć se rozsuziš we swěše digitalneje (internetoweje) leksiko-

<sup>7</sup> Tak stoj na pš. w artiklu „brumak“ (nimski ‘der Brummer’) w Mukowem słowniku (MUKA 1911–15/1926, 1928) pšikładowa sada, w kótarež se namakajo słowo *brumar* (z wóznamom ‘Drohne’). To słowo howac niži we słowniku njestoj, jo we pšikłaze schowane a jano pšipadnje na njo starejjoš. W toš tom paže by słowo *brumar* teke namakali pšez pšikładowy korpus, ako jo naspomnjety we pšiducnej nožce. Take „schowane“ lemata pak se teke namakaju na drugih městnach struktury digitalizěrowanych słownikow, na pš. w komentarach.

<sup>8</sup> K tomu ako pódkorpus teke lice pšikłady, wustupujuce w digitalizěrowanych słownikach. Tam se namakaju wuraze, kótarež njejsu niži ako samostatne lemata zapisane, lěcrownož „se za to góže“. Dokulaž maju take wuraze ako dokłady ze sekundarnych žrědłow leksikografiski hynakšy status ako wuraze z korpusow, budu wóni wósebnje markěrowane.

<sup>9</sup> Kšuta słowna wězba (feste Wortverbindung) se how wužywa ako pógromadne zapšimješe za wšykne (pó žělach) idiomatiske wuraze a rowno tak za wuraze, kótarež drje su z perspektiwy slucharja (do dalokeje měry) transparentne, z perspektiwy powědarja pak dejali se naglědaš ako we mentalnem leksikonje skłazone. Powšyknje slědujo how wužywana terminologija tuchylu BURGER 2010 a BURGER ET AL. 2007, 1–10. Zapšimješe „kolokacija“ wužywa BURGER (2010: 52) za „den ganzen Bereich der festen Wortverbindungen, die nicht oder nur schwach idiomatisch sind“. Teke konstrukcije z funkcionalnym werbom (Funktionsverbgefüge; na pš. *eine Entscheidung treffen*) su pó Burgeru kolokacije. – Konsistentne wobchadanje z kšutymi słownymi wězby jano jedno z nejwětšych wupominanjow projekta.

<sup>10</sup> Kompletne wugódnosenje korpusow jo glědajucy na to jano móžne ako dlujkodobne pšedewzeše.

grafije wěcej njeeksistěrujo. Pomjenjone zasady ředowanja se wěcej njewuzamkuju, mógu se skerzej samo mjazy sobu wudopólniš. Wużywarje słownika mógu wuběraś mjazy wšakimi móžnosćami ředowanja a z tym prezentacijami, jolic se pśi koncepciji a wužěłanju słownika w pśiměrjonej formje žiwa na wótpowědne wopśimješowe a techniske opcije. Take póstupowanje njepomina sebje jano rozšyrjenje, ale teke nowe wobžěłanje donětejšnego formata mikrostruktury a wótpowědne wobšyrne wopisanje kuždego słowa. Dodatne žěło pak se wuplašijo, dokulaž wużytnosć słownika se z tym radnje pówušyjo. Pódlu rozšyrjonych móžnosćow pśistupa, ako se w slědujucem pśedstajaju, jo se wě teke alfabetiski pśistup zarucony.

### 2.2.1. Słowna familija

Pśi nowem wobžěłanju celkownego słowoskłada se póstupuju primarnje pó principje słownych familijow. Z tym žiwa se južo pśi leksikografiskem wopisanju jednotliwych leksemow na toś ten genuiny princip rozředowanja słownistwa (HUNDSNURSCHER 2002: 679). Dokulaž jo za polne wopisanje słowneje familije<sup>11</sup> semantiska analiza wšyknych (potencielných) čłonkow notna, zachopijo se wobžěłanje nejppjerwej z kupku leksemow, zestajonej pó cysto formalnych kriterijach (zgromadny słowny kórjeń), kótaraž se pótom dalej analizěrujo.<sup>12</sup>

Z definicije pśedmjata *SORBLEXA* a z rozsuda, až njocomy rownocasnje teke wužěłaś etymologiski słownik dolnoserbšćiny (pśir. 2.3.2.), slědujo, až se „diachroniska słowna familija“ n j e rozmějo ako rěcywědnje rekonstruěrowana etymologiska słowna familija, kótaraž se póšěgujo na (zrekonstruěrowane) słowa z pśedpismowego casa až do indoeuropejskich abo nanejmnjej prasłowjańskich kórjenjow. Skerzej rezultěrujo maksimalne přédne zestajenje słowneje familije (w zmysle kšaca analize, kenž zepěra se hyšći celo na formalny bok) ze zwiska ze zgromadnym (indigenym abo póžyconym) dolnoserbškim kórjenjom. Na toś tej bazy pón slědujo semantiska analiza čłonkow tuchylnje wobčerjoneje słowneje familije a pśeslěžowanje a koděrowanje semantiskich póšěgow<sup>13</sup> mjazy nimi. Z tym dožyomy ku skónčnej formje diachroniskeje słowneje familije resp. rezultěruju z togo kšaca někotare žěłne słowne familije, kótarež póšěgujo se na wšake wóznamy jědrowego słowa.

<sup>11</sup> Zapśimješe „słowna familija“ (Wortfamilie) se njewużywa w literaturje jednotnje. Tak se na pś. rozeznowatej (a) słowna familija ako diachroniski definěrowana jadnota słowotwórbnych produktow, kótaraž zepěra se na etymologiske zwiski (na pś. FLEISCHER/BARZ <sup>2</sup>1995: 71; SCHIPPAN 1984: 40) a (b) słow(otwórb)ne gnězdo ako synchroniski definěrowana struktura z formalnje a semantiski identiskim zakładnym morfemom (jědrowym słowom gnězda) (t.). Toś to rozeznowanje pak se wšuži njenamaka. AUGST na pś. we swójom słowniku „Wortfamilienwörterbuch der deutschen Gegenwartssprache“ (1998) definěrujo słowne familije na zaklaže relatiwneje motiwěrowanosći, ako wiži ju „normalny“ wyžywař rěcy (VII), powědajacy w toś tom nastupanju wó synchroniskej etymologiskej kompetency (IX). Wón wuraznje wuzwigujo, až njejsu wót stawiznarjow rěcy wopisowane historiske zwiski pśedmjat jogo słownika (VIII).

<sup>12</sup> Wótpowědna procedura njamóžo se how wobgranicowanego městna dla dokradnej wopisaś; pśir. pokazki we pśiducnem wótrězku.

<sup>13</sup> How zo wó słowotwórbne kategorije ako nomen agentis, a teke wó pśirědowanje słowotwórbnych produktow abo jich wóznamow, jolic až su polysemne, k wótpowědnej bazy resp. jadnomu wóznamuju.

Diachroniska słowna familija reprezentěrujo cełe rěcnohistoriske wuwise danego wurězka dolnosěrbskego słownictwa a wopšimjejo wšykne jej formalnje a semantiski pširědowane słowa ze wšych znatych žrědłow. Pši tom se pódawaju informacije wó zastupjenju słowa do familije (prědny dokład abo zapisk; nimski: Erstbeleg resp. Erstbuchung) a ewentualnje teke wó casowym wobgranicowanju pšislušności (gaž dajo slědny dokład; nimski: Letztbeleg) a dalšne za diachroniske pšedstajenje relewantne informacije.

Synchroniska słowna familija wopisujo dolnosěrbsčinu wót něži kónca 19. stolěša. Wóna jo wósebnje pšepytowana a pšedstajona žěl diachroniskeje słowneje familije. Wobej warianse familije móžotej byś identiskej, gaž w běgu rěcnych stawiznow ze synchroniskeje perspektiwy žeden leksem abo žeden wóznam danego cłonka njejo se zgubił. Za wětšy part pak jo synchroniska słowna familija glědajuce na licbu leksemow a jich wóznamow žěl diachroniskeje. Wóna se how rozmějo ako słowotwórnbne gnězdo,

- (a) kótarež se zestajijo jano z leksemow, ako se we žinsajšnej dolnosěrbsčinje naložuju abo nanejmnjej hyšći pasiwnje derje rozměju a
- (b) kótarež se konstituěrujo na zaklaže „rěcneho wědobnja lajkow” („Laienbewusstsein“, pšir. AUGST 2002: 681) pšez spóznaše motiwěrowanych a z tym formalnje a semantiski transparentnych słowotwórbnych zwiskow.<sup>14</sup>

Generelnje póbituju kuždy pšedstajenje (synchroniskeje abo diachroniskeje) słowneje familije primarnje pšegłěd wó celkowej familiji. To groni, až informacije wó jadnotliwych cłonkach (leksemach) su mócnje skrotcone. Zasadnje jo how móžno, wótwisnje wót zajma wužywarja wšake stopnje dokradności pokazaš. W takem pšegłěže słowneje familije nalicone leksemy wjedu ako hyperlinki k wobšyrnym wopisanjam jadnotliwych leksemow. Wót nich jo móžno se wrošís k pšegłědoju daneje familije a wóttam zasej dojs k wšyknym drugim cłonkam familije. Na taki part jo pšedstajenje słowneje familije systematiski zwězane z nadrobnym semaziologiskim wopisanim wšykných jeje cłonkow. Wětšy part słownikow słowných familijow, na pš. styrizwězkowy pólski SŁOWNIK GNIAZD SŁOWOTWÓRCZYCH (2001, 2004), póbituju jano relatiwnje abstraktne pšedstajenje słowotwórbnych strukturow. Lěcrownož su toś te informacije mimo cwiblowanja krađu wužytne, wóstanjo take želne pšedstajenje za wužywarja weto fragmentariske. Pólski słownik drje bazěrujo na monoseměrowanju jědrowego słowa a na wótpowědnem semantiskim pširědowanju deriwatow, wóni sami pak se we słowniku wěcej njewopišu. Město tego se pokazuju na žrědłowe słowniki, kótarež by musał wužywar k ruce mės, aby połne wopisanje dostał. Take wobgranicowanje na žěl pšistupnych informacijow pši prezentaciji, ako jo we šiścanych słownikach cesto trěbne, w projekše *SORBLEX* njejo notne. Togodla jo móžno mjaz sobu zwězane informacije (na pš. wěsty deriwat<sup>15</sup>, jogo relaciju z jědrowym zdonkom a jogo wóznam(y)<sup>16</sup>) teke

<sup>14</sup> Njebužo pšecej lažko how se rozsuziš a musymy togodla licyš z njewěstymi padami. Weto se póstupowanje na zaklaže rěcneje wědy wó motiwěrowanosći cłonkow jadneje słowneje familije zasadnje gózi. K wótpowědnej metodice pšir. AUGST 2002: 681.

<sup>15</sup> Zawěšće se žiwa teke na kompoziciju (na pš. *glukužycenje* ‘Glückwunsch’) a druge wašnje słowotwórbny.

<sup>16</sup> SPLETT (2009) to na pš. zarucyjo. – Celkowna struktura jadnotliweje słowneje familije se we pšedlažecyach słownikach wušej tego rozdźělne pšedstajijo. AUGST (1998) pódajo jědrowe słowo familije a pšegłěd wó jogo wóznamach. K tomu prezentěrujo (w rědowanju, ako se za formu orientěrujo) deriwaty a dodatnje jich pšez jědrowe słowo motiwěrowane wóznamy we

nagromadu pokazaš. Móžności modularnego pšedstajenja we wótwisności wót zajmow wużywarjow we internetowem słowniku wušej tego zmóžniju za- abo wublendowanje wěstych informacijow. – How gropnje pšedstajona koncepcija słownych familijow mu- sy se dalej wužęłaš.<sup>17</sup>

### 2.2.2. Semantiske relacije a słowne póla

Pši leksikografiskem wopisanju se žiwa na semantiske relacije ako (quasi)synonymiju, hyperonymiju, (ko)hyponymiju, polarnosc (antonimiju, komplementarnosc, konwersitu, holonymiju a meronymiju. Dokulaž su (a) toš te relacije (pódlá drugih) konstitutiwne za etablěrowanje słownych pólow a dokulaž (b) dej se zmóžniš pši prezentaciji datow we online-słowniku teke onomaziologiski pšistup pšez wótpowědne pšawidłowne pširědowanje jednotliwych wóznamow k słownym pólam, njejo hyšći dowujasnjone, žo dokradnje w mikrostrukturje se trěbne informacije zapišu a na kaki part mógu se zwęzaš k zgromadnemu, pó móžności pšewidnemu celkownemu wobrazuju. (pšir. 2.3.3.).<sup>18</sup> Njejo pšedwizone wužęłaš swójsku kategorizaciju zapšimješow. Skerjej comy pšespytowaš klasifikaciske systemy, kótarež se nałožuju w drugih słownikach ako na pš. *Historical Thesaurus of the Oxford English Dictionary* 2009, *Roget's International Thesaurus* <sup>7</sup>2010, *Dornseiff: Der deutsche Wortschatz nach Sachgruppen* <sup>8</sup>2004). „Semantic hierarchy“ słownika *Historical Thesaurus* wopšimjejo 236400 kategorijow. Toš ta hierarchija jo nejwobšyrnjejša klasifikacija z górnjeje naspomnjetych słownikow a pokazuju z nagromadu wósym rowninami wjeliku hierarchisku dłyminu. Taka diferencěrowanosć pak jo za rěc ako dolnoserbšćinu, kótaraž ma relatiwnje mały slowosklad, nanejmnjej pó žělach pšeliš nadrobna. Napšęsiwo tomu wujžo słownik Dornseiffa z jano 970 kategorijami a z tym póbituju system, ako dajo se bžez problemow za *SORBLEX* wużywaš. Na drugem boce pak ma toš ten system jano dvě rowninje, což casy wježo do bejnje wobšyrnych słownych pólow, kótarež by se mimo šěžkosćow dalej diferencěrowaš dali.

### 2.3. Mikrostruktura

Charakteristiska za *SORBLEX* jo kombinacija artiklow, ako se pósęguju na jednotliwe lemata, z nadřędowanymi pšedstajenjami, ako zepěra se na jednotliwe artikle a na do-

---

formje tabele. Na taki part jo cely wobstatk formow jano raz pokazany (synchronisko-polysemne pšedstajenje); kuždu pšez wěsty wóznam jědrowego słowa motiwěrowanu žělnu słownu familiju móžoš pak indirektnje z pomocu tabele teke wopóznaš. We pólskem *Słowniku gniazd slowotwórczych* 2001, 2004 pšedstajuju napšęsiwo tomu žělne słowne familije hynac, to groni kuždu za se (synchronisko-monosemne pšedstajenje). Pši takem póstupowanju pak musyš se wótnamakaš z tym, až wóspjetuju se te formy, kótarež su motiwěrowane pšez wšake wóznamy jědrowego słowa, to groni kótarež su teke polysemne a z tym slušaju k wšakim słownym familijam.

<sup>17</sup> K tomu se zapšęguju dalšne relewantne słowniki ako SPLETT 2009, typologiskeje bliskosci dła pak wósebnje take słowjańskich rěcow (mimo naspomnjetejo *Słownika gniazd slowotwórczych* na pš. ТИХОHOB 1985).

<sup>18</sup> Teke móžnu orientaciju za systemami semantiskeje kategorizacije ako WordNet (<http://wordnet.princeton.edu>) abo FrameNet (<http://framenet.icsi.berkeley.edu>) bužomy pšępytowaš. Toš ten žěl projekta drje jo nejmócnjej „eksperimentalny“ aspekt a njejo hyšći jasne, do kakeje měry dajo se wón realizěrowaš.

datne, w XML-owych datajach koděrowane informacije. Wósebná gódnota *SORBLEXA* napěšejiwo staroznatym słownikam lažy teke we móžnej „změnje perspektiwy“ mjazy diachroniskim a synchroniskim předstajenim a mjazy jednotliwym artiklom (lema) a předstajenim słownje familije resp. słownego póla. Pši tom wótpowědne systematiske zwězanja zwidobnjiju kompleksne formalne a semantiske pósęgi we słowoskłaže. Wušej tego zmóžnjijo wašnja koděrowanja datow wjeleserake pšistupy k informacijam, ako su we słowniku zapisane.

Pódlá wěcej abo mjenjej zwuconych pozicijow – lemaforma, lemowe warianty, gramatiske informacije; (k jednotliwym wóznamam:) walenca (žož relewantna), wóznamowe pódaše<sup>19</sup>, dokłady –, kótarychž dokradne wugótowanje how wopisaš njamóžomy, wopšimjejo mikrostruktura teke w slědujucych wótrězkach wopisane pozicije (3.3.1–3.3.3).<sup>20</sup>

### 2.3.1. Leksikografiske zapiski

Wušej naspomnjete styri digitalizěrowane słowniki fungěruju ako standardowe žrědła<sup>21</sup> za informacije wó leksikografiskich zapiskach danego leksema z danym wóznamom. Jolic lema resp. wěsty jogo wóznam jo zapisany w toš tych žrědłach, se wóni nalicuju ako skrotconki (sigle), a to w chronologiskem rěže, což dokumentěruju stawizny leksikografiskego registrěrowanja. Skrotconka jo rownocasnje hyperlink a twóri direktny zwisk z internetoweju wersiju žrědła (słownika). Na taki part jo pšecej pšistupne połne doněntejšne leksikografiske wopisanje, a to w nowo wugótowanej formje na wobrazowce a wušej tego teke ako faksimile. Take póstupowanje pówušyjo informacisku gódnotu teje pozicije mikrostruktury a dodatnje garantěruju transparencu glědajucy na tradiciju leksikografiskego wopisanja.

### 2.3.2. Etymologija

Ako jo se južo groniło, dej *SORBLEX* předstajiš leksikaliske wuwise dolnosěrbsčiny wót zachopjenka pismojstwa. Etymologiske informacije su togodla předwizone jano w dwěma padoma:

- (a) Za indigene jědrowe słowa słownych familijow, mjazy kótarymiž rěcneho wuwisa dla južo na zachopjenku pismojstwa – na zaklaže lajkowego wědobnja (pšir. nožku 11) – formalny a(bo) wopšimješowy zwisk njedajo se južo spóznaš, kótarež pak daju se pšez nutšikownu rěcnohistorisku rekonstrukciju wótwóziš wót zgromadnego dolnosěrbskego žrědła, pódaajo se toš ta informacija w předstajenju wótpowědnych separatnych słownych familijow. Pši tom zepěramy se předewšym na SCHUSTER-ŠEWCA (1978–89), a jolic jo notne, teke na dodatne slěženja. Za toš tu a za dalšne informacije jo předwizona fakultatiwna pozicija mikrostruktury „diskurs“.

<sup>19</sup> Za pódaše wóznama se wužywa nimski ekwiwalent, gaž wón jo (pšisamem) synonymny z wóznamom dolnosěrbskego słowa. Za precizěrowanje (pši semantiskich wóthylenjach wót ekwiwalenta) abo jolic direktnego ekwiwalenta scelego njejo, parafrazěruju se dolnosěrbske wóznamy (teke w nimskej rěcy). Pši tom comy se zepěraš, tak daloko ako móžno, na wobgranicowany wopisowański wokabular.

<sup>20</sup> K semantiskim relacijam pšir. kap. 2.2.2.

<sup>21</sup> Manuelnje se zapšěguju SCHUSTER-ŠEWCA 1978–89 a, jolic jo trjeba, dalšne słowniki, kótarež njepředlaže w digitalnej formje.

- (b) Pši cuzych słowach resp. póżyconkach se pó móžnosći pódawaju informacije wó direktnej žrědłowej rěcy a pśedłozje. Za toś te pódáša trěbne pozicije mikrostruktury su: žrědłowa rěc, direktna pśedłoga, wóznam pśedłogi (jolic njejo identiski), žrědło informacije (jolic jo trjeba) a wujasnjenje (gaž jo za rozměše notne), za dalšne komentary stoj k dispoziciji wušej naspomnjeta pozicija „diskurs“.

### 2.3.3. Informacije za generěrowanje nadržědowanych pśedstajenjow

Koncepcija *SORBLEXA* pomina sebje teke koděrowanje informacijow, kótarež su trěbne za generěrowanje nadržědowanych pśedstajenjow (pśedewšym słownych familijow a pólow) a zmóžnju wótpowědne pśistupy k datam. Kake informacije su za to notne, skicěruju se we slědujucem na pśiklaže słowneje familije:<sup>22</sup>

- (a) synchroniska morfematiska segmentacija (na zaklaže synchroniskeje morfematiskeje transparence);
- (b) jolic jo trjeba: diachroniska morfematiska segmentacija;
- (c) njepósrědna baza deriwata;
- (d) jědrowy zdonk wótpowědneje synchroniskeje słowneje familije,
- (e) jolic jo trjeba: jědrowy zdonk wótpowědneje diachroniskeje słowneje familije,
- (f) pširědowanje wóznama deriwata k wóznamoju deriwaciskeje bazy;
- (g) stopjeń słowotwórbneho wótwóženja glědajucy na jědrowe słowo familije;
- (h) morfologiski typ deriwacije, na pś. prefigacija, sufigacija atd.;
- (i) semantiski typ deriwacije ako wuslědk słowotwórbneho procesa, na pś. nomen agentis, finitiwna wašnja cyna atd.

Wšake (formalizěrowane) pódáša, kótarež se namakaju we słownikach słownych familijow, na pś. stopjeń wótwóženja, ewentualnje w zwisku ze změnami słowneje družyny, mógu se z wušej naspomnjetych a dalšnych we pšislušnych artiklach koděrowanych informacijow (na pś. słowna družyna) awtomatiski generěrowaš. Ako pśikład za j a d n u móžnu konwenciju: V,V,S = substantiw (*wótgráše*) na zaklaže werba (*wótgráš*) na zaklaže werba (*gráš*).

## 3. Wuznam projekta

Pšez wjelerakosć za *SORBLEX* nazběranych a w formaše XML koděrowanych informacijow nastanjo wobšyrny informaciski system wó dolnoserbškem słowoskłaže, w kótaremž se zjadnošuju wšake, pó zwuconej wašni rozdžělnje wózwajwone słowniki (alfabetiski řědowane, take wó słownych familijach abo pólach, wó póżyconkach atd.). Wašnja informacisko-techniskego koděrowanja datow zmóžnju jich wjelebočne wužywanje a zarucyjo móžnosći řědowanja a pśistupow w bogatej libcje (polyakcesiwny słownik). Na pś. dowólju za pširědowanje k słownym familijam trěbne registrěrowanje segmentacije leksemow pši wobžělanju słowoskłada dodatny morfematiski pśistup k datam, kótaryž póbituju wó wjele wěcej móžnosćow ako retrogradny słownik. Lažko bužo togodla zestajiš wše leksemy, kótarež wopšimuju wěšte słowotwórbne morfemy, na pś.

<sup>22</sup> Nalicone informaciske pozicije njesu dapołne juž togodla, až pósěguju se wósebne na deriwaciju. Weto su někotare pozicije, na pś. (a) a (b) teke w dalšnych padach, na pś. pla kompozitumow, dosega juce. – Wótpowědne pozicije su teke notne glědajucy na konstituciju słownych pólow (pšir. 2.2.2.).

dwójnje prefigěrowane werby na *do-na-* (na ps. *donatwariš*) abo substantiwu na *-ar'* (na ps. *twarjař*).

Pšez konsekwente koděrowanje za SORBLEX nazběranych informacijow we formaše XML su toš te daty, rowno tak ako z nimi zwězane daty (z DNWa, z digitalizěrowanych słownikow atd.) polyfunkcionalne a neutralne glědajuce na móžnu medialnu prezentaciju. Polyfunkcionalność se zložuju na młogosc a wjeleserakosc za projekt zezběranych informacijow. Medialna neutralnosť datow zmóžnijo mimo glowneje prezentacije informacijow na wobrazowce dalšne formaty wudaša, na ps. šišćarske wudaše (k individualnemu wótwołanju) mjenjej abo wěcej wobšyrnych žělow słownika (na ps. jadnotliwych artiklow abo pšedstajenjow słowneje familije). Wušej togo póbituju reprezentacija datow w formaše XML dosegajuce zakład za zwězanje datowego wobstatka z drugimi, pó pódobnych metodach póstupujucymi projektami (druge internetowe słowniki, słownikowe portale abo seši).

Pó móžnosći wupócerajuce wobžěłanje dolnosěrbskego słowoskłada, kakež we projekće SORBLEX na zakłaže maksimalneje datoweje bazy přědny raz do statka stajis comy, by było wjeliki póstup, wósebnje glědajuce na to, až móžo se z tym wažny žěl rěčno-kulturnego derbstwa toš teje wěcej ako 1000 lět w pódzajtšnej Nimskej domorodnej rěcy dokumentěrowaš a reprezentěrowaš. Wušej togo mógu wuslědki pšedewzeša wótpóraš dlujkólětne dezideraty sorabistiskego slěženja a twóriš nowy zakład za wšakorake pšepytowanja. Aby jano pór pšikładow naspomnjeł:

- (a) SORBLEX stwórijo bazu za diachronisku a synchronisku wucbu dolnosěrbskeje słowotwórby a z tym zmóžnijo dalšne pšepytowanja na toš tom pólu.
- (b) Móžnej bužotej dopołna inventarizacija a dokradne wopisanje póžyconkow w dolnosěrbsčinje.
- (c) „Nutsikowne“ stawizny dolnosěrbskeje rěcy njejsu hyšći napisane. Nanejmenjej za słowoskład by mógal planowany słownik tam wjele pšinjasć. Teke wuwijanje dolnosěrbsčiny pó 1945, ako jo wobwliwowane pšez wšakorake faktory, móžo se pšawje wopisowaš jano z dobreju znajobnosću pšnorěčneje tradicije.

Gaž comy rozměš how wopisane pšedewzeše ako moderny „tezawrus“, dejali pokazaš na přědny wopyt, taku twórbu wužělaš. Južo wót lěta 1959 jo se w Institúše za slawistiku tencasneje Nimskeje akademije wědomnosći zachopiło žěło na „Thesaurus des Sorbischen“, pó tegdejšej koncepciji za wobej serbskej rěcy: „Wir planen ein sorbisches Gesamtwörterbuch, in dem der Wortschatz aller Zeiten von den ältesten Spuren bis zur jüngsten Gegenwart [...] enthalten sein wird“ (BIELFELDT 1961: 125, pšir. teke MÜLLER 1967). Cil, wužělaš „wjeliki pomnik“ („ein großartiges Denkmal“ (t.)) serbsčiny, njejo pak se mógal zrealizěrowaš. Projekt su musali zazdašim južo pšed dokóńcenim fazy zběranja materiala kónc 1960ych lět wótlamaš, nejskerrej z wědomnostno-politiskich pšicyn. Lěc how wopisane pšedewzeše, kótarež by mógało slůžys ako model teke za górnosěrbsčinu a jo charakterizěrowane pšez celo hynakše techniske móžnosći ako něgajšne akademiske pšedewzeše, bužo se w swójej celosći realizěrowaš daš, njejo hyšći jasne.

## Literatura

- AUGST, G. (2002): Typen von Wortfamilien, w: CRUSE, D. A. ET AL. (Hrsg.): *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*. 1. Halbband. Berlin-New York, 681–688.
- AUGST, G. (1998): *Wortfamilienbuch der deutschen Gegenwartssprache*. Tübingen.
- BARTELS, H.; THORQUINDT-STUMPF, K. (2013): Ein neues Ton- und Textarchiv des muttersprachlich-dialektalen Niedersorbischen, w: *Lětopis* 60 (2013) 1, 38–59.
- BARTELS, H. (2012) [w šišću]: Maßnahmen zur Dokumentation des Niedersorbischen, w: *Slavia Occidentalis* 69 (2012).
- BARTELS, H. (2010): Das (diachrone) Textkorpus der niedersorbischen Schriftsprache als Grundlage für Sprachdokumentation und Sprachwandelforschung, w: HANSEN, B.; GRKOVIĆ-MAJOR, J. (eds.): *Diachronic Slavonic Syntax. Gradual Changes in Focus*. München-Berlin-Wien (Wiener Slawistischer Almanach – Sonderbände; 74), 7–18.
- BIELFELDT, H. H. (1961): Die sorabistischen Arbeiten des Instituts für Slawistik der Deutschen Akademie der Wissenschaften zu Berlin, w: *Lětopis A* 8 (1961), 124–126.
- BURGER, H. (2010): *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin.
- BURGER, H. ET AL. (2007): Phraseologie: Objektbereich, Terminologie und Forschungsschwerpunkte, w: BURGER, H. ET AL. (Hrsg.): *Phraseologie. Ein internationales Handbuch zeitgenössischer Forschung*. 1. Halbband. Berlin-New York, 1–10.
- FLEISCHER, W.; BARZ, I. (1995): *Wortbildung der deutschen Gegenwartssprache*. Tübingen.
- HUNDSNURSCHER, F. (2002): Das Wortfamilienproblem in der Forschungsdiskussion, w: CRUSE, D. A. ET AL. (Hrsg.): *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*. 1. Halbband. Berlin-New York, 675–680.
- MUKA, A. (1911–15/1926, 1928): *Wörterbuch der nieder-wendischen Sprache und ihrer Dialekte*. Zw. 1: St. Petersburg 1911–15, Prag 1926: Verlag der russischen und tschechischen Akademie der Wissenschaften; zw. 2/3: Prag 1928: Verlag der böhmischen Akademie für Wissenschaften und Kunst. (Fotomech. Neudruck. Bautzen 2008.)
- MÜLLER, K. (1967): Die Bedeutung des Sorbischen Thesaurus, w: *Forschungen und Fortschritte* 9 (1967). Berlin, 283–285.
- SCHIPPAN, TH. (1984): *Lexikologie der deutschen Gegenwartssprache*. Leipzig.
- SCHUSTER-ŠEWIC, H. (1978–1989): *Historisch-etymologisches Wörterbuch der ober- und niedersorbischen Sprache*. Bautzen.
- SPLETT, J. (2009): *Deutsches Wortfamilienwörterbuch. Analyse der Wortfamilienstrukturen der deutschen Gegenwartssprache, zugleich Grundlegung einer zukünftigen Strukturgeschichte des deutschen Wortschatzes*. Berlin-New York.
- STAROSTA, M. (1999): *Niedersorbisch-deutsches Wörterbuch*. Bautzen.
- SŁOWNIK GNIAZD SŁOWOTWÓRCZYCH (2001, 2004) – *Słownik gniazd słowotwórczych współczesnego języka ogólnopolskiego*. Pod redakcją naukową H. JADACKIEJ. Kraków, zw. 1: T. VOGELANG: *Gniazda odprzymiotnikowe*, 2001; zw. 2: H. JADACKA ET AL.: *Gniazda odrzeczownikowe*, 2001; zw. 3: M. SKARŻYŃSKI ET AL.: Cz. 1+2: *Gniazda odczasownikowe*, 2004; zw. 4: M. SKARŻYŃSKI ET AL.: *Gniazda motywowane przez liczebniki, przysłówki, zaimki, przyimki, modulanty, onomatopeje, wykrzykniki*, 2004.

ŠWJELA, B. (1961): *Dolnoserbsko-němski słownik*. Budyšyn.

ТИХОНОВ, А. Н. (1985): Словообразовательный словарь русского языка, т. 1–2, Москва.

ZWAHR, J. G. (1847): *Niederlausitz-wendisch-deutsches Handwörterbuch*. (Fotomechanischer Neudruck. Bautzen 1989.)