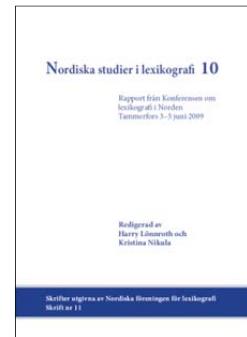


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Grammatisk informasjon i middels store tospråklege ordbøker - Prinsipp og utfordringar på bakgrunn av arbeidet med Norsk-litauisk ordbok
Forfatter:	Sturla Berg-Olsen & Eglė Išganaitytė
Kilde:	Nordiska Studier i Lexikografi 10, 2010, s. 103-112 Rapport från Konferens om lexikografi i Norden, Tammerfors 3.-5. juni 2009
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive



© Nordisk forening for leksikografi

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsrettsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for ’optical character recognition’ og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

STURLA BERG-OLSEN & EGLĖ IŠGANAITYTĖ

Grammatisk informasjon i middels store tospråklege ordbøker

Prinsipp og utfordringar på bakgrunn av arbeidet med Norsk-litauisk ordbok

This article deals with certain questions pertaining to the inclusion of grammatical information in middle-sized bilingual dictionaries. While lexicographers often would like to include quite a lot of information on phonology, morphology and syntax, at least for some users the amount of information can become so large as to hamper the practical use of the dictionary. This is especially the case in dictionaries that aim to be useful for both L1 and L2 speakers. We look at how the editors of the Norwegian-Lithuanian Dictionary tackle this balance in practice, especially focusing on some features that are particular for this language pair, such as variant forms in Norwegian and accentuation and aspect in Lithuanian.

Nøkkelord: tospråklig leksikografi, bidireksjonale ordbøker, grammatisk informasjon, baltiske språk

1. Innleiing

I denne artikkelen tek vi for oss ein balansegang som forfattarar av tospråklege ordbøker må ta stilling til: På den eine sida ønskjer leksikografen gjerne å gje ordboksbrukarane mest mogleg informasjon om materialet i ordboka – både om lemma og ekvivalentar, men på den andre sida er det ei grense for kor mykje informasjon det er plass til (i ei papirordbok) og også for kor mykje informasjon ein kan gje før det verkar meir forstyrrande enn opplysende for mange brukarar. Etter å ha drøfta denne balansegangen mellom prinsipp og praksis, går vi gjennom konkrete døme på val som redaksjonen i *Norsk-litauisk ordbok* (NLO) har gjort når det gjeld presentasjon av grammatisk informasjon.¹ Vi viser at norsk og litauisk byr på til dels nokså ulike utfordringar på dette området.

1 Vår definisjon av *grammatisk informasjon* i denne samanhengen er nokså vid, og omfattar ikkje berre morfolgiske og syntaktiske eigenskapar, men også informasjon om aksentuasjon. Som det går fram i del 6, må ein i litauisk ha kjennskap til aksentuasjonen av eit ord for å kunne bøye det rett.

2. Mellom prinsipp og praksis

Det er eit velkjent faktum i tospråksleksikografien at dei to brukargruppene – dei som har høvesvis kjeldespråket (L1) og målspråket (L2) som morsmål – har bruk for ulike typar informasjon i ei ordbok mellom dei to språka det er tale om (sjå t.d. Svensén 2004: 17 ff.). Samstundes veit vi også at ein berre unntaksvise har ressursar til å lage fleire ordbøker som er spesielt tilpassa dei ulike brukargruppene, særleg når det er snakk om høvesvis små språk som dei skandinaviske og baltiske. I praksis vil ein leksikograf som ønskjer å ta omsyn til begge gruppene difor ofte måtte prøve å gjere dette innanfor rammene av éi ordbok, som då blir bidireksjonell, det vil si av typen L1→L2/L2→L1. Dette er også tilfellet med NLO. I ei ordbok av denne typen vil ein gjerne ønskje å ha med grammatiske informasjon om ord i begge språk, sjølv om denne informasjonen om L1 er overflødig for dei som har L1 som morsmål, og omvendt. I vårt konkrete tilfelle ønskjer vi å gje litauiske brukarar opplysningar om ulike eigenskapar ved dei norske lemmaa, samstundes som vi vil gje norske brukarar tilsvarende informasjon om dei litauiske ekvivalentane. For begge brukargruppene er den grammatiske informasjonen truleg i første hand nytig ved produksjon av tekst og ytringar på framandspråket.

I ei lærebok om tospråksleksikografi set Valerij P. Berkov opp følgjande grunnprinsipp for kva grammatiske informasjon han meiner ein tospråkleg ordbok bør innehalde:

Som regel må ei tospråkleg ordbok gje både ein aktiv og ein passiv brukar [...] informasjon som tillèt han å danne kvart element i paradigmet (dvs. kvar form) til eit ord – lemma og ekvivalent – og bruke det grammatiske rett i ein kontekst. (Berkov 1996: 109, omsetjing S.B-O.)

Berkovs ideal er altså at ordboka skal gje begge brukargrupper *all* informasjon som trengst for å danne kvar form av lemmaet eller ekvivalenten, avhengig av kva morsmål dei har. Dette prinsippet er følgt til dømes i Berkovs store ordbøker mellom norsk og russisk (Berkov 1987, 1994, 2003). Her er alle lemma og ekvivalentar aksentuerte (det same gjeld bruksdøme og omsetjingar av dei), og både lemma og ekvivalentar er utstyrt med kodar som gjev informasjon om genus, bøyingsklasse, aksentuasjonsklasse (for russiske ord) og aspekt (for russiske verb).² Dette gjer ordbøkene til veritable gullgruver av informasjon som til dømes studentar kan ha stor nytte av, som illustrert i figur 1.

2 I *Stor norsk-russisk ordbok* er lemma med uregelrett bøyning, t.d. sterke verb, utstyrt med grunnformer, ikkje kodar.

vr̄ist m1/f1 подъём ^{la} (<i>nogu</i>); han har høy ~ у него высокий подъём
vr̄iste v1 ① vt вырывать ^{1/} вырвать ^{6aΔ} ; hun ~t kníven fra ham она вырвала у него нож ② vr ~ seg вырываться/ться; hun ~t seg løs fra ham она вырвалась от него
vr̄æl [æ:L, e:l] n1 рёв ^{1a} , (громкий) крик ^{3a} , вопль ^{2a} , вой ^{6a} ; han sätte i et sint ~ он издал яростный рёв
vr̄ælle [æ:L, è:l] v2 реветь ^Δ ; ùngene ~er дёти ревут; han fikk hèle forsám- lingen til å ~e av látter он застáвил все собráние кататься со смéхом {сó смеху}; han ~te av smèrte {sinne} он надрывался от боли {ярости}

Figur 1. Døme på strukturen
i Stor norsk-russisk ordbok
(Berkov 2003)

Ei av ulempene ved dette systemet, er at det stiller høge krav til brukarane: For å få full nytte av den grammatiske informasjonen som blir gjeven i ordboksartiklane, må dei stadig slå opp i tabellar andre stader i boka, eventuelt lære seg systemet utanboks. Brukarar som av ulike grunnar ikkje gjer dette, sit etter alt å dømme att med heller lite utbytte av den rike basen av informasjon som ord-boka inneheld. Ein kan heller ikkje sjå bort frå at den store informasjonsmengda lett kan verke meir forstyrrende enn opplysende på ein del brukarar. Bo Svensén skriv om ”den tvåspråkiga lexikografins problem”:

Artiklarna i en tvåspråkig ordbok är ofta överfulla av information som den enskilde användaren i en viss konsultationssituation inte behöver. Detta är ett störande moment som medför att många användare ger upp innan de funnit den sökta informationen. (Svensén 2004: 311.)

Med andre ord må ordboksforfattarens ønske om å gje brukarane så fullstendig informasjon som mogleg, vegast opp mot omsynet til brukarvennligheit. I tillegg til dette må ein tenke på kor mykje plass ein har til rådvelde (i ei papirordbok), og det kan også vere tids- og ressursmessige tilhøve som set grenser for kor mykje informasjon ein kan gje i ordboka. I det følgjande presenterer vi korleis desse spørsmåla blir takla i prosjektet *Norsk-lituansk ordbok* og ser på nokre døme på løysingar derfrå.

3. Prosjektet Norsk-litauisk ordbok

Norsk-litauisk ordbok er eit prosjekt ved Senter for skandinavistikk ved Vilnius universitet. Dei fleste prosjektmedarbeidarane held til der, men også leksikografar ved Universitetet i Oslo tek aktivt del på fleire stadium i arbeidet med ordboka. Arbeidet med boka tok til i 2005, og ein tek siktet på å ha boka ferdig i 2011. Med 45 000–50 000 lemma blir truleg NLO større enn noka anna ordbok mellom eit skandinavisk og eit baltisk språk, sjølv om ho i ein vidare samanheng likevel ikkje høyrer til blant dei største tospråklege ordbökene (til samanlikning inneheld *Stor norsk-russisk ordbok* cirka 90 000 lemma). Ordboka skal kunne brukast med utbytte både av litauarar og nordmenn, men ut ifrå dagens situasjon, vil ho truleg bli brukt av fleire litauiske enn norske brukarar.³ Om forfatarane ser seg nøydde til å velje å ta mest omsyn til éi av brukargruppene, er det altså mykje som talar for å sjå på dei litauiske brukarane som primærgruppa. Målet med prosjektet er å produsere ei papirordbok, men i og med at boka blir skriven i XML, ligg alt til rette også for elektronisk publisering.

4. Grammatisk informasjon om norsk

Ved norske lemma blir det gjeve informasjon om ordklasse, og ved substantiv også informasjon om genus, ved hjelp av forkortingar. For lemma med regelrett bøyning har ein valt å indikere bøyingsklasse med kodar, slik det er gjort til dømes i Berkovs ordbøker mellom norsk og russisk og i *Bokmålsordboka* (red. Wangensteen) og *Nynorskordboka* (red. Hovdenak et al.). I sjølve lemmaa er informasjon om trykk og tonem gjeven med aksentar plasserte over vokalen i den trykksterke stavinga, på same måte som i Berkovs ordbøker og i *Litauisk-norsk ordbok* (Jakaitienė & Berg-Olsen 2001). Dette representerer eit avvik frå den norske tradisjonen, der ein markerer trykk, tonem og vokallengd med teikn etter vokalen, eventuelt etter den påfølgjande konsonanten. Ved ein del lemma – primært framord – er det gjeve uttaleopplysningar i hakeparentesar.⁴

3 Interessa for å studere norsk i Litauen er høg og stabil, medan det i Noreg ikkje har vore tilbod om studium i litauisk sidan 1990-talet. Endå viktigare for samansettjinga av brukargruppa er det store talet på litauarar som dei siste åra har busett seg i Noreg for kortare eller lengre tid.

4 Her og vidare i artikkelen er ordboksmanuskriptet sitert slik det ligg føre i skrivende stund. Delar av artiklane er stundom utelatne om dei ikkje er viktige for det poenget som skal illustrerast.

- váffel** *m1a*
váffelhjerte *n2*
váffeljern *n1*
vag *a1*
vakánse *m1 [ŋ]*

Bruken av kodar for bøyingsklassar gjer det mogleg å gje mykje informasjon i komprimert form, men er truleg også med på å minske brukarvennlegheita til ordboka (jf. merknaden i samband med figur 1 over). For grunnord med uregelrett bøyning, til dømes sterke verb, blir grunnformene gjevne i fullskriven form, men for samansetjingar der den andre komponenten har uregelrett bøyning, blir ikkje desse formene gjentekne – brukaren må då slå opp på grunnordet (eller i en tabell om det gjeld eit sterkt verb, markert med koden v*) for å finne ut korleis ho skal bøye samansetjinga.

- bror** *brøren, brødre*
ta *v tok, tatt*
voldta *v**
v penbror *m*

Ein sparar plass i ordboka ved å ikkje fullskrive dei uregelrette formene ved lemma som *voldta* og *v penbror*, men samstundes føreset ein då at ein litauar som skal produsere norsk, tek seg bryet med å finne ut av kvar denne informasjonen er å finne.

5. Avspegling av variantrikdomen i norsk

Som kjent er norsk prega av ein h vesvis stor grad av variantrikdom når det gjeld bøyning og skrivem tar. Den norske spr kstoda, med to offisielle skriftnormer som begge har mykje valfridom og i tillegg fleire uoffisielle normer, som riks l  og h gnorsk, verkar lett uoversiktleg for folk som skal l re seg  forst  og bruke norsk. I ei tospr kleg ordbok som involverer norsk, blir ein difor stilt overfor valet mellom  avspegle ei eller fleire normer heilt ut, eller  avgrense mengda av variantar. Eit problem ved den f rstnemnde l ysinga er at dei offisielle normene (og kanskje s rleg bokm lsnorma) omfattar ei stor mengd former som i praksis er lite brukte og/eller er stilistisk markerte. D  kan ein lett risikere at framandspr klege brukarar i ein produksjonssituasjon vel norske former som, sj lv om dei ligg innanfor norma, sjeldan eller aldri ville blitt brukte av ein norsk morsm lstalar.

TILSLAGSORD	ARTIKKEL FRA BOKMÅLSORDBOKA (offisiell rettskriving)
bredbånd	~band el. ~bånd høyhastighetsnett for telekommunikasjon
	breiband el. ~bånd el. bred~ s d

Figur 2. Døme på variantrikdomen i bokmål: Både *bredband*, *bredbånd*, *breiband* og *breibånd* er tillatne former. Frå nettversjonen av Bokmålsordboka

I NLO har ein valt ein mellomveg. Berre former som fell innanfor den offisielle bokmålsnorma er med – altså finn ein ikkje riksmålsformer som *sne* eller *sprog* her. Samstundes er offisielle bokmålsformer som i røynda sjeldan blir brukte, ofte heller ikkje med. I ein del tilfelle blir slike lite brukte former oppgjevne ved grunnordet, men ikkje i samansetjingar. Når eitt ord har fleire moglege skrivemåtar, har ein valt ut éin hovudskrivemåte (ut frå bruksfrekvens, men òg i nokre tilfelle ut frå kva som kjem først i alfabetet), som det så er vist til frå dei andre oppgjevne variantane.

bred a1 (ir brei)

brèdbånd n1 (ir brèibånd)

brei a1 → bred

brèibånd n1 → brèdbånd

brott II n1 → brudd

brudd n1 (ir brott)

brùddflate f1/m1

Døma over viser at dei tillatne variantane *bredband* og *breiband* ikkje er tekne med (men *band* er med som variant til grunnordet *bånd*). Når det gjeld *brudd*, er varianten *brott* med ved grunnordet, men dukkar ikkje opp ved alle samansetjingar med *brudd*.

Når det gjeld variasjon mellom genus og/eller bøyingsklassar for eit og same lemma, er den redaksjonelle praksisen i NLO at ein speglar bokmålsnorma fullt ut.

6. Grammatisk informasjon om litauisk

6.1 Aksentuasjon

Aksentuasjonstilhøva i litauisk er etter måten kompliserte. Språket har fritt og mobilt trykk, det vil si at ein student for kvart leksem må lære ikkje berre kvar trykket ligg i utgangsforma, men også kva aksentuasjonsklasse lemmaet høyrer til, for å kunne aksentuere det rett i dei ulike kasus- og numerusformene. I tillegg finst det ein opposisjon mellom to tonem i trykksterke lange stavingar. Trykk og tonem blir ikkje markerte i vanleg litauisk tekst, men det er vanleg å markere det i lærebøker og ordbøker mynta på utlendingar. Utan nokon informasjon om aksentuering blir ei ordbok med litauisk som L2 klart mindre nyttig for dei med L1 som morsmål (i vårt tilfelle nordmenn) i produksjonssituasjonar.

Om ein skulle følgje Berkovs prinsipp om at ordboka skal gje brukarane all informasjon som er naudsynt for å danne kvar form av lemma og ekvivalentar, burde ein i NLO ikkje berre aksentuere ekvivalentane, men også oppgje kva aksentuasjonsklasse dei høyrer til. Dette kunne gjerast ved hjelp av det kode-systemet som er i bruk i litauiske eittspråklege ordbøker, med kombinasjonar av tal og bokstavar (t.d. 1, 2, 3^a, 3^b). Når ein har valt å ikkje gjere nett dette, er det i hovudsak for å unngå at informasjonsmengda i ordboksartiklane blir for stor og for å unngå forvirring blant brukarar utan særleg lingvistisk skolering. Hadde ein valt å oppgje aksentuasjonsklassar, ville aksentuasjonskodane måtte stå etter kvar ekvivalent, og i tilfelle der ekvivalenten er ein frase, ville hovedet i frasen ha kode, men ikkje komplement (kongruerande attributt, derimot, kunne i prinsippet få kode). Her ser vi nokre døme på korleis dette kunne sjå ut (jf. òg figur 1 frå *Stor norsk-russisk ordbok*, som oppgjev aksentuasjonsklassar for russiske ekvivalentar):

ildvann -et šnek. ugninis² vanduõ^{3a}
injúriesøksmål n1 ieškinÿs^{3a} dėl garbës īr̄ orùmo įžeidìmo
intím a1 1. (fortrolig) intymùs⁴, ařtimas^{3b}, asmeninis², nuoširdùs⁴ [...]

I staden har ein valt ein mellomveg, der ein aksentuerer ekvivalentane, men ikkje oppgjev aksentuasjonsklassar. Brukarane får altså vite korleis ordet blir aksentuert i den oppgjevne forma (til vanleg identisk med utgangsforma, dvs. nominativ eintal eller infinitiv), men vil måtte konsultere andre kjelder for å aksentuere andre former rett.

6.2 *Bøyning*

Om ein ser bort frå aksentuering, er böyinga av litauiske ord i stor grad føreseileg ut frå utgangsforma. I nokre få tilfelle treng ein likevel tilleggsinformasjon for å kunne bøye ord rett. Det gjeld til dømes substantiv på *-is*, som blir bøygde ulikt avhengig av om dei er feminina eller maskulina:

váltis ”(liten) båt”: femininum; Gen. sg. *válties*, Nom. pl. *válty*

brólis ”bror”: maskulinum; Gen. sg. *brólio*, Nom. pl. *brólai*

Alle ekvivalentar som er substantiv på *-is*, bør i lys av dette markerast. Éin måte vil vere å ta med genitiv eintal ved alle desse orda, men ein smidigare måte å få fram det same på, er truleg å markere genus. Det held å gjere dette ved feminina på *-is*, ettersom desse er i mindretal:

båt m1 1. (*lite fartøy*) *váltis*^F

bror m *brøren*, *brødre* *brólis*, dim. *broliukas* [...]

6.3 *Aspekt*

I verbsystemet har litauisk ein delvis grammatikalisert opposisjon mellom imperfektiv og perfektiv aspekt. I korte trekk kan nokre (imperfektive) verb berre uttrykkje tilstandar og handlingar som varer ved, medan nokre (perfektive) verb berre kan uttrykkje handlingar som er førde fram til eit resultat. Perfektive verb er avleidde av imperfektive ved hjelp av ymse prefiks.⁵ For ein nordmann vil det i mange tilfelle vere nyttig å vite både om eit gjeve verb har ein prefigert, perfektiv variant (eller fleire slike variantar, som i artikkelen *bore* under), og i så fall kva prefiks som er brukt. I NLO er perfektive variantar av litauiske ekvivalentar førde opp. Ofte blir det også gjeve bruksdøme med omsetjingar som illustrerer skilnaden mellom aspektane:

blø v4 1. (*miste blod*) (nu)*kraujúoti*, netékti *kraūjo* ► *såret blør žaizda krau-*

5 Situasjonen blir komplisert ved at prefigerte verb ofte har ei tilleggstyding i høve til grunnverbet, utan at dei er perfektive. Ein har heller ikkje noko gjennomført mønster med sekundære imperfektive verb, slik ein har i t.d. russisk (jf. russisk *pisat'* ”skrive [imperfektivt]” : *zapisat'* ”skrive ned [perfektivt]” : *zapisyvat'* ”skrive ned [imperfektivt]”).

juoja; [...] *blø seg i hjel* mirtinai nukraujuoti 2. [...]
bòre v1 (iš)gręžti, pragręžti ► *bòre hull (i noe)* (iš)gręžti (kame) skylę [...]

6.4 Valens

Når det gjeld valens, har det vore eit poeng for redaksjonen i NLO at ein bør markere det når tilhøva i dei to språka ikkje fell saman. Det vil konkret seie at ein ikkje markerer det særskilt når verb tek objekt utan preposisjon i norsk og akkusativobjekt i litauisk, ettersom dette er normaltilstanden (jf. første tyding ved *iaktta* i artikkelen nedanfor). Om eit litauisk verb blir brukt med andre objektskasus, blir dette markert ved ekvivalenten, og gjerne også illustrert med bruksdøme. Ved andre tyding i dømet under er det spesifisert at verba *laikytis* og *paisyti* blir brukte med genitiv, medan *kreipti dèmesj* ”rette merksemd” blir brukt med preposisjonen *j* og akkusativ. Valensen blir markert med pronomenet *kas* ”kva, kven” i rett kasus, slik praksis er i litauiske ordbøker som har med informasjon av denne typen.

iáktta v 1. (*observere*) stebéti, sèkti akimìs [...] 2. (*overholde*) laikytis (*ko*), páisyti (*ko*), kreipti dèmesj (*j kq*) ► *det er visse lòver og régler som må iákttas* esama tam tikrū įstatymū ir taisyklių, kurių reikia laikytis {paisyti}

For norsk blir det markert når verb blir brukte med visse preposisjonar, særleg om denne bruken skil seg frå litauisk. I artikkelen nedanfor er bruken av preposisjonen *med* i uttrykket *få bukt med* spesifisert, samstundes som det er markert at dei litauiske verba *jveikti* og *nugalėti* tek akkusativobjekt.

bukt f1/m1 1. geogr. [...] 3. (*krumning*) linkis, vǐngis, riñgis, kilpa [...] • *få bukt (med noe/noen)* jveikti, nugalėti (*kq*)

7. Konklusjon

Som vi har sett, stiller norsk og litauisk ordboksforfattarane overfor nokså ulike utfordringar når det gjeld å formidle grammatiske informasjon. For norsk må ein i tillegg til å gje informasjon om bøyning, ta stilling til i kva mon og korleis ein skal spegle av den store variantrikdomen i språket. Her er det gode argument for at ein ikkje blindt bør følgje norma, men primært ta med former som faktisk er i bruk. Når det gjeld litauisk, er det eit poeng å gje informasjon som

kan hjelpe norske brukarar til å bøye og bruke dei litauiske ekvivalentane rett. Samstundes må ein vere klar over at informasjonsmengda også kan bli for stor for ein del brukarar. Når redaksjonen i *Norsk-litauisk ordbok* har teke sine val på dette feltet, har det vore med det målet for auget å finne ein middelveg mellom informasjonsrikdom og brukarvennlegheit.

LITTERATUR

- Berkov, Valerij P., 1987: Russisk-norsk ordbok. Moskva.
- Berkov, Valerij P., 1994: Russisk-norsk ordbok. 2. utgave. Oslo.
- Berkov, Valerij P., 1996: Dvujazyčnaja leksikografija. St. Petersburg.
- Berkov, Valerij P., 2003: Stor norsk-russisk ordbok. Oslo.
- Hovdenak, Marit et al. (red.): Nynorskordboka. http://www.dokpro.uio.no/ordboksoek_nynorsk.html/
- Jakaitienė, Evalda & Berg-Olsen, Sturla (red.), 2001: Lietuvių-norvegų kalbų žodynas/ Litauisk-norsk ordbok. Vilnius.
- Svensén, Bo, 2004: Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik. 2 uppl. Stockholm.
- Wangensteen, Boye (red.): Bokmålsordboka. <http://www.dokpro.uio.no/ordboksoek.html/>