

Lennart Lönngren - Jan Rak - Jens Skov-Larsen

OM FRAMSTÄLLNING OCH ANVÄNDNING AV ORDLISTOR I TJECKISKA

1. Inledning

Vid Slaviska institutionen i Uppsala har inom ramen för ett år 1971 igångsatt projekt, benämnt MIR-projektet (= målanalys i ryska), utarbetats principer för systematisk uppbyggnad av det ryska ordförrådet vid läsning av ryska texter. När nämnda projekt avvecklades (1973-74) hade enligt dessa principer framställts dels en s k basordlista, omfattande ca 300 ord, dels ordlistor för den skönlitterära delen av A-nivåns i studieplanen fastställda textkurs. I basordlistan finns de ord som de flesta studerande kan förväntas ha som gemensamt förråd när de lämnar gymnasietadiet. Basordlistan bygger på ett par ryska frekvenslexika samt den i gymnasieskolan (tidigare) använda nybörjarboken, Slovo za slovom. Ordlistorna till de enskilda texterna upptar inte ord som förekommer i basordlistan.

MIR-projektet har också haft som målsättning att bedöma texters lexikaliska svårighetsgrad - varvid man särskilt tagit fasta på ordriken - och att ordna texterna inbördes med hänsyn härtill. Därför bygger alla textordlistorna utom den första inte bara på basordlistan utan också på föregående texter och deras ordinnehåll. Avgörande för de resultat som uppnåtts har varit tillgången till dator och ett system för automatisk morfologisk analys (på flexionsnivå), vilket utarbetats av FD Anna-Lena Sågvall, som också varit huvudansvarig för MIR-projektet.

Ovan beskrivna basordlista och textordlistor finns samlade i ett läromedelspaket, benämnt "Text och ord" (Slaviska institutionen, Uppsala). En beskrivning av de bakomliggande principerna återfinns i Anna-Lena Sågvall och Berith Brännström: "MIR-projektet, en sammanfattning av verksamheten 1.9.1971 - 31.1.1974" (stencil, PU-enheten, Uppsala universitet).

De mer generella av MIR-projektets principer kan utan vidare tillämpas på andra språk än ryskan. Vilket behov föreligger då av läromedel av detta slag? Vi skall här inte närmare diskutera argumenten för och emot textordlistor över huvud taget. Såsom nackdelar har anförts bl a att de studerande riskerar att få för liten vana att använda lexikon och att inlärningseffekten är större vid det mödosamma släendet i lexikon än vid användning av ordlistor. Såsom en fördel kan man peka på det förhållandet att undervisningen kan avlastas viss ren faktainlärning och koncentreras på grammatisk analys, speciella översättningssvårigheter m m. Vi förutsätter att åtminstone inom det här aktuella området - de lägsta stadierna av undervisningen - fördelarna överväger framför eventuella nackdelar. Behovet av goda självinstruerande läromedel av ifrågavarande slag torde förent av vara större i språkämnen med få studerande och följaktligen blygsamma undervisningsresurser, i synnerhet om det därtill saknas fylliga lexika till och från den studerandes modersmål.

I tjeckiska har vid vår institution sedan flera år förekommit sporadisk produktion av ordlistor till texter avsedda för A- och B-nivån. Sådana ordlistor har emellertid visat sig bli otillfredsställande - oavsett kvaliteten i övrigt - på grund av det subjektiva och godtyckliga urvalet av ord, vilket ofta uppfattas som inadekvat av den studerande. Sålunda händer det ofrånkomligen att ordlistan å ena sidan innehåller en del kända ord men å andra sidan inte upptar en del okända ord, varför den studerande ändå måste vända sig till lexikon. Det är emellertid orimligt att tänka sig att en ordlista skulle uppta alla, även de mest frekventa orden; sådana ordlistor skulle bli alldeles för långa och otympliga och därmed motverka sitt syfte. Det förefaller som om ett system med en basordlista åtminstone delvis skulle kunna lösa detta problem. En basordlista möjliggör nämligen ett automatiskt, objektigt urval av ord till de enskilda texternas ordlistor.

Mot denna bakgrund uppkom höstterminen 1973 tanken på att konstruera en basordlista på A-nivån i tjeckiska och att därefter igångsätta systematisk framställning av ordlistor till i första hand A-nivåns i studieplanen fastställda textkurs. I juni 1975 förelåg såväl en basordlista som ordlistor till 70 normalsidor text i färdigt skick.

Tillsammans med texterna har detta material samlats till ett läromedelspaket, benämnt "Tjeckisk textkurs A 1" (Slaviska institutionen, Uppsala). Basordlistan, som är försedd med ett förord och en kort instruktion, är utarbetad av Lennart Lönngren, Jan Rak och Jens Skov-Larsen (Jan Rak kunde dock inte delta i den slutliga utformningen) och textordlistorna är sammanställda av Olga Klauber och Lennart Lönngren.

Ehuru avsikten från början var att tillämpa MIR-projektets principer visade det sig snart att vi på flera punkter fann anledning att modifiera dessa. Detta berodde delvis på tjeckiskans särart och på att vi inte hade till vårt förfogande MIR-projektets tekniska hjälpmedel, men även vissa allmänna överväganden ledde oss till att i några avseenden ompröva MIR-principerna. Bl a av denna anledning kan kanske en presentation av vårt tillvägagångssätt ha ett visst intresse och möjligen stimulera till en diskussion inom detta område. I nedanstående beskrivning görs inte någon detaljerad jämförelse med MIR-projektets principer, men motiveringen för vårt tillvägagångssätt kommer att vara extra utförlig på de punkter där vi skiljer oss från MIR-projektet.

2. Basordlistan

I första hand är basordlistan, som redan antytts, avsedd att utgöra ett grundlexikon till texterna på A-nivån. Detta innebär också att den måste användas av eleven under textläsningen - vid sidan av de enskilda textordlistorna - i den mån som eleven inte har lärt in den. I andra hand kan basordlistan fylla funktionen att utgöra riktmärke för undervisning och examination även inom de övriga språkfärdighetskurserna. Det rör sig ju om garanterat centrala glosor, vilka det är motiverat att inlära aktivt på ett tidigt stadium. Basordlistans betydelse härvidlag ökar givetvis med graden av integration mellan språkfärdighetskurserna, speciellt vad gäller undervisningsmaterial.

Man kan dock sannolikt inte gå så långt att basordlistan uttömmande får definiera det ordförråd som skall inläras aktivt på A-nivån i dess helhet; de häri ingående språkfärdighetskurserna har ju var och en sin speciella karaktär med avseende på undervisning och läromedel.

Det är exempelvis rimligt att anta att det ordförråd som kommer till användning inom kursen Muntlig språkfärdighet delvis är ett annat än det som ryms i en basordlista, speciellt komponerad för kursen Tjeckisk litteratur. Å andra sidan: för att basordlistan på bästa sätt skall kunna fylla sin uppgift inom den sistnämnda kursen är det av vikt att den uppfattas som så central som möjligt, så att den studerande känner sig starkt motiverad att inlära den på ett tidigt stadium.

I basordlistan har en kolumn reserverats för vardera a) de tjeckiska orden, b) grammatisk information till dessa, samt c) svenska ekvivalenter. Vi skall nedan behandla var och en av dessa komponenter för sig.

a) De tjeckiska orden ges i alfabetisk ordning. Urvalet av ord har gjorts på följande sätt. Som en första komponent har medtagits de 290 mest frekventa orden i den tjeckiska frekvensordboken J Jelínek - J Bečka - M Tešitelová, Frekvence slov, slovníků druhů a tvarů v českém jazyce, Prag 1961. Emellertid visar det sig att ordlistan vid tillämpning av enbart denna urvalsprincip kommer att sakna många för vardagsspråket centrala ord, främst substantiv. För att kompensera denna brist har 151 sådana ord tillagts, vilka subjektivt valts ur en tjeckisk nybörjarkbok, nämligen M Šára - J Šárová - A Bytel, Čeština pro cizince (Czech for English-speaking students), Prag 1970. Dock har även dessa ord en betryggande frekvens enligt ovannämnda frekvensordbok. - Sammanlagt omfattar således basordlistan 441 ord.

Basordlistan innehåller endast de öppna, s k "autosemantiska" ordklasserna, dvs substantiv, adjektiv, verb och adverb. De slutna, s k "synsemantiska" ordklasserna, dvs främst pronomina, räkneord, prepositioner och konjunktioner, skulle enligt de ursprungliga planerna "dragas ut" till separata listor. Det har emellertid vad gäller dessa ordklasser visat sig mer ändamålsenligt att helt enkelt hänvisa till elementära nybörjarkböcker och grammatikor, i vilka dessa ordklasser i allmänhet redovisas i sin helhet. De slutna ordklasserna förutsätts till största delen inläras inom ramen för nybörjarkursen samt kurserna i muntlig och skriftlig språkfärdighet.

En fördel med att basordlistan består av enbart öppna ordklasser är att den studerande blir uppmärksam på den viktiga distinktionen mellan öppna och slutna ordklasser samt även starkt motiverad att lära in de som regel högfrekventa orden i de slutna ordklasserna. En annan motivering för ifrågavarande begränsning av ordlistan ligger däri, att denna av den studerande bör upplevas som så adekvat, överskådlig och homogen som möjligt.

Vissa svårigheter uppstår oundvikligen när det gäller avgränsningen av öppna resp slutna ordklasser. Sålunda har i basordlistan medtagits orden mnoho, málo, tolik, několik och dost, trots att dessa ord - med undantag av dost - av morfologiska skäl brukar klassificeras som obestämda räkneord. Även den relativt slutna gruppen av pronomiella adverb (nikdy, někam etc) har medtagits i basordlistan.

b) Den grammatiska informationen upptar ordklasstillhörighet endast då denna inte - såsom i de flesta fall - sammanfaller med den svenska ekvivalentens. Uppgift om genus, böjningsmönster och böjningsformer ges endast i den mån som ordet i fråga avviker från vissa "normalfall", vilka anges i instruktionsdelen till basordlistan. Till normalfallen har förts endast ett relativt litet antal av de mest generella reglerna och paradigmen. Avsikten med att presentera den grammatiska informationen på detta sätt är att göra den studerande motiverad att ständigt hålla huvuddragen av grammatiken aktuell, så att dessa inte går förlorade bland mängden av mindre generella regelmässigheter.

Uppgift om syntaktisk konstruktion, främst verbrektion, ges endast om denna inte lätt kan förutses utifrån den svenska ekvivalentens konstruktion. Vidare anges för verben eventuell partner i aspektpar. Härvid tillämpas samma princip som återfinns i Ožegovs ryska ordbok. Bland de tjeckiska orden upptas således den ordbildningsmässigt enklare aspektformen, medan den deriverade formen ges i informationskolumnen. Till t ex dělat ges informationen "ipf; pf: udělat; till dát ges informationen "pf; ipf: dávat". För en alfabetiskt ordnad ordlista torde detta vara den bästa modellen (beträffande aspektinformation i textordlistorna, se nedan). - Att helt avstå från att sammanföra verb i aspektpar, såsom görs i MIR-projektets basordlista,

måste anses mindre tillfredsställande.

c) Den svenska översättningen till varje uppslagsord har i stort sett utformats på vedertaget sätt. En strävan har varit att göra denna del så koncis som möjligt. Endast de viktigaste betydelseerna har medtagits. Vi har vidare undvikit att rada upp synonymer; sådana har endast medtagits då de fyllt en klart preciserande funktion.

3. Textordlistorna

I textordlistorna upptas orden i den ordning i vilken de förekommer i ifrågavarande text. I övrigt har textordlistorna i stort sett samma utseende som basordlistan. En skillnad består i informationen om aspektpartner: i textordlistorna upptas som utgångsord den aspektform som förekommer i den aktuella texten, även om detta verb råkar utgöra den i ordbildningshänseende mer komplicerade aspektformen. Varje textordlista bygger direkt på basordlistan, dvs den upptar alla ord som inte finns i denna. Dessutom kan ord som visserligen förekommer i basordlistan, men som i den givna texten har en mycket speciell betydelse, medtagas i textordlistorna.

Vid framställningen av ordlistorna har de tekniska förutsättningarna för en automatisk textbehandling saknats. Det har t ex inte gått att på grundval av en direkt objektiv analys ordna texterna inbördes med hänsyn till ordinnehållets svårighetsgrad. Den enda möjligheten har varit att låta varje textordlista bygga direkt på basordlistan (jfr MIR-projektets system). Detta medför ett par nackdelar: dels blir ordlistorna längre än om man låter varje textordlista bygga på den närmast föregående, dels kan ordlistorna (utom den första textens ordlista) uppfattas som mindre adekvata av den studerande. Dessa olägenheter torde minskas något om man undviker att välja alltför många och korta texter. Man kan emellertid också peka på den fördelen, att texter som ej är ordnade inbördes är lätt utbytbara och alltså inte bildar ett så slutet och definitivt paket som ordnade texter. Detta är naturligtvis särskilt väsentligt om man är osäker beträffande de valda texternas lämplighet för nivån i fråga.

I detta sammanhang kan också påpekas att det när ordlistorna väl är framställda givetvis finns ett kriterium för att fastställa i vilken ordning texterna bör läsas: texter med relativt korta ordlistor i förhållande till textmängden har ett mer centralt ordinnehåll (i relation till basordlistan) än texter med längre ordlistor i förhållande till textmängden. De förra bör därför läsas före de senare.

Ett sätt att reducera textordlistornas längd skulle kunna vara att utesluta sådana ord som står i mycket frekventa och regelbundna ordbildningsförhållanden till basordlistans ord - främst diminutiver, iterativer, verbalsubstantiv och desubstantiviska adjektiv. Det torde dock vara lämpligt att uppskjuta en sådan reduktion till närmast högre nivå, då elementa i tjeckisk ordbildning bör ha hunnit genomgås.

Ord i textordlistorna, vilka enligt den ovan nämnda frekvensordboken uppnår en viss frekvenströskel, motsvarande ungefär de 1000 vanligaste tjeckiska orden, markeras med plustecken. Avsikten är att den studerande skall ägna särskild uppmärksamhet åt dessa ord, vilka bör inläras passivt, dvs så att den studerande behärskar dem i riktningen tjeckiska - svenska, något som knappast är realistiskt att kräva för övriga ord.

En annan avsikt med "plusorden" är att de kan utgöra en av byggstenarna i en eventuell basordlista på närmast högre nivå, dvs B-nivån. Vid komponerandet av en sådan basordlista får önskvärdheten av att den studerande mött ordet i kontext (såsom fallet är med plusorden) vägas mot kravet att basordlistan skall vara central, alltså att orden skall ha en garanterat hög frekvens i språket och att alla sådana ord (upp till en viss gräns) finns med. B-nivåns basordlista kan således förutom A-nivåns plusord lämpligen innehålla alla ord i den tjeckiska frekvensordboken upp till en viss frekvenströskel; därjämte kan det visa sig lämpligt att även på denna nivå lägga till vissa "vardagsord". B-nivåns basordlista bör i likhet med A-nivåns tjäna som riktmärke för vad som på nivån i fråga skall krävas som aktivt ordförråd (= ord som behärskas i riktningen svenska - tjeckiska). Det torde vara praktiskt för den studerande att den i utskrivet

skick inkluderar A-nivåns basordlista. För övrigt får erfarenheterna av A-nivåns ordlistor avgöra storleken och utformningen av basordlistan på B-nivån.

I bilaga följer nedan såsom prov en sida från vardera basordlistan och en textordlista. Textkursen i sin helhet kan rekvireras från Slaviska institutionen i Uppsala.

TJECKISK BASORDLISTA

ano		ja
asi		ungefär; troligen, kanske
aspoň el alespoň		åtminstone
ať		må, låt
barva	gen pl: <u>barev</u>	färg
bát se	ipf; pres: <u>bojím se</u> ; pret: <u>bál se</u> ; kstr: <u>koho/čeho</u> (1:a bet); <u>o koho/co</u> (2:a bet)	vara rädd (för); vara rädd om
běžet	ipf; fut: <u>poběžím</u>	springa
bílý	komp: <u>bělejší</u>	vit
blízky	komp: <u>bližší</u> ; adv: <u>blízko</u> ; komp: <u>blíže</u> el <u>blíž</u>	nära; närbelägen; adv: nära
boj		kamp
bota		sko; känga
brambor	i pl oftast f-böjn	potatis
brát	ipf; pres: <u>beru</u> ; pf: <u>vzít</u> ; pres pf: <u>vezmu</u> ; pret: <u>vzal</u>	taga
bratr		broder
brzo el brzy		snart; tidigt
bůh (Bůh)	gen sg: <u>boha</u> ; dat o lok sg: end <u>bohu</u> ; nom pl: <u>bozi</u> el <u>bohové</u>	gud
bydlit el bydlet	ipf	bo
být	ipf; pres-böjn, se gram; fut: <u>budu</u>	vara
celý		hel
cena		pris, värde
cesta		väg
cítit	ipf; pf: <u>po-</u> el <u>u-</u>	känna, förnimma
cizí		främmande; utländsk
cukr		socker
čas	lok sg: end <u>čase</u>	tid; tempus

Karel Čapek: Svatá noc

+ svatý		helig
divit se	ipf; pf: <u>po-</u>	förvåna sig
+ křičet	ipf; pf: <u>křiknout</u>	ropa, skrika
pořádný		ordentlig
starosta		borgmästare
+ takhle		så här, på det här sättet
žebrota		tiggeri
Šimonovic		familjen Simon
ujímat se	ipf; pf: <u>ujmout se</u>	ta sig an
ausgerechnet	(tyska)	just
+ copak		verkligen, då
holota		patrisk
+ pustit	pf; ipf: <u>pouštět</u> ; pres 3 p pl: <u>pouštějí</u>	släppa in
zahazovat se	ipf; pf: <u>zahodit se</u>	nedlåta sig
bručet	ipf; pf: <u>za-</u>	muttra
+ pravit	ipf	såga, yttra
pozvedat	ipf; <u>pozvednout</u>	höja
nevidáno		man har inte sett maken
to by tak hrálo		det skulle se ut, det
pípnout	pf; ipf: <u>pípat</u>	pipa
+ vůle	f	vilja
kvůli	kstr: <u>komu/čemu</u>	på grund av; för ...
tulák		skull
+ povídat	ipf; pf: <u>povědět</u> ; pres pf: <u>povím</u> , 3 p pl: <u>povědí</u>	luffare
poběhlík		såga, berätta
+ chodit	ipf	luffare
		här: bruka vara