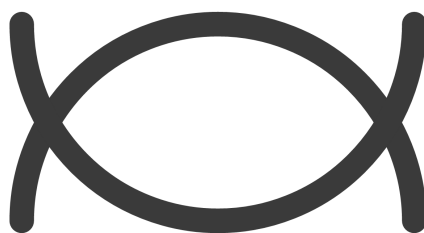

Karp

Övningar
Språkbankens höstworkshop 2017

<https://spraakbanken.gu.se/karp>
sb-karp@svenska.gu.se

17 oktober 2017



ÖVERSIKT

I nuläget finns det över 25 olika lexikon i Karp. När du går in på <https://spraakbanken.gu.se/karp> får du välja mellan olika lexikonlägen. Vissa lägen innehåller flera lexikon. Du kan välja vilka du vill söka i med hjälp av lexikonväljaren högst upp på sidan. För att byta lexikonläge klickar du på lägesväljaren bredvid Karp-ikonen.



I de här övningarna kommer vi använda flera lägen och lexikon.

ATT SÖKA I KARP

Enkel sökning är en fritextssökning som låter en söka efter ett eller flera ord. Sökningen görs i all text i de valda lexikonerna och resultaten presenteras i relevansordning (TF/IDF).

Vill man göra en snävare sökning kan man använda Utökad sökning. Exakt vad som är sökbart här avgörs av lexikonets struktur och uppmärkning. Att söka efter "*vad som helst*" motsvarar en fritextssökning, men resultatet för utökade sökningar presenteras alltid i bokstavsordning.

Knappen "Sammanställ på..." öppnar upp en meny där man kan välja olika attribut som man vill se en mer kompakt statistik för. Tryck på Sök när du gjort dina val. Resultatet kommer att visas under fliken "Statistik", bredvid "Träffar".

Om du håller musen över ett attribut visas en liten ikon till höger om kryssrutan. Med hjälp av denna kan du flytta attributet och därmed ändra ordning på kolumnerna i sammanställningen.

ÖVNING 1

Saldo beskriver modern svenska och är ett av våra största lexikon med ca 130 000 ingångar. Här har alla ingångar grundform och lemgram (en slags identifierare), så för att hitta alla ingångar kan du t.ex. göra en sökning efter ”där grundform finns”. Vi ska nu göra ett antal övningar i *Saldos morfologi* med hjälp av statistiken från ”Sammanställ på...”. Öppna lexikonväljaren och ställ in så att enbart morfologin är vald.

- 1.1 Vilka ordklasser förekommer i lexikonet?
(Se Appendix A för en översikt av förkortningarna.)
- 1.2 Klicka på den lilla miniräknaren över kolumnen med ordklasser, längst till vänster, för att se en mer kompakt sammanställning av de förekommande värdena. Är fördelningen mellan ordklasserna som du förväntat dig?
- 1.3 Gör en ny sökning och hitta statistik över vilka inherent drag som finns för substantiv i lexikonet. (Se Appendix B för en översikt av förkortningarna.)
- 1.4 Kan du hitta en ordform som tillhör fler än fyra ordklasser?
(Tips: sammanställ på ordform och därefter ordklass.)

ÖVNING 2

Resursen *Konstruktikon* samlar uttryckssätt och språkliga strukturer. Det lättaste sättet att navigera i lexikonet är att använda listningen till höger i fönstret. Under listningen kan du även välja att slå av Förenklad visning. Gör det!

- 2.1 Slå upp konstruktionen stå_där_med_X. Kolla på fältet Struktur och på Exempel för att få en bättre bild av vilket språkligt fenomen som beskrivs.

Den här konstruktionen länkar till det svenska frasnätet, SweFN (FrameNet). Det är en resurs som beskriver betydelse i förhållande till semantiska ramar, d.v.s. händelser, relationer etc. En länk innebär i det här fallet att konstruktionens betydelse beskrivs av FrameNet-ramen.

- 2.2 Vilken ram i frasnätet länkar stå_där_med_X till?
Genom att klicka på länken kommer du till SweFN. Gör det!

Nu befinner vi oss i Frasnätets visningsläge. Under Exempel hittar du nu en samling meningar som belyser den aktuella semantiska ramen. Under Lexikala enheter finns ord och flerordsuttryck (vilka är länkade till Saldo) som är kopplade till ramen.

- 2.3 Gruppera de lexikala enheterna per ordklass och välj ut din favorit!
- 2.4 SweFN länkar i sin tur tillbaka till Konstruktikon (SweCxn). Ta dig tillbaka till konstruktikonet genom att följa en av länkarna och passa på att se om du kan lära dig några nya bra uttryck från den valda konstruktionens exempel.

ÖVNING 3

Hellqvists *Svensk etymologisk ordbok* är från 1930-talet och är en av de resurser vars digitala version är under utveckling. Vi ska titta på hur man kan använda den information som finns redan nu. En del av informationen – såsom ordklass och grundform – är uppmärkt och därmed sökbar. Större delen av texten i ingångarna har däremot ingen uppmärkning och kan därför bara sökas i genom fritextsökning.

- 3.1 Sök efter "karp", "korp" eller "strix" och läs om ordets ursprung! Ordet "karp" har ordklass "e". Det betyder att digitaliseringsprocessen inte automatiskt kunde bestämma vilken ordklass ordet hör till, och att den manuella bearbetningen av resursen ännu inte hunnit hit.
- 3.2 Sök efter "höst". Här hittar du även en länk till Runebergs faksimil, där du kan läsa och bläddra i Hellqvist som pdf.

Hellqvist ger anvisningar om ords ursprung och använder förkortningarna *grek.* för (forn-) grekiska, *fsv.* för fornsvenska och *lat.* för latin. Förutom ursprung anger Hellqvist även besläktade ord och deras ursprung, vilket gör att sökningar efter ord med ett visst ursprung resulterar i ett visst brus. En manuell uppmärkning av etymologiska länkar finns påbörjad, om än bara i sin linda.

- 3.3 Gör en sökning efter substantiv som härstammar från latin. Använd sökfältet etymon-språk för detta.

Kanske hittar du något fel i lexikonet under sökningarna. Om du gör det kan du välja att lämna ett förbättringsförslag genom att klicka på det lilla kugghjulet bredvid den aktuella ingången.

ÖVNING 4

Dalins ordbok utkom på 1850-talet. Den kan vara till god hjälp om man läser äldre litteratur. Om du är intresserad av ords böjning kan du söka i *Dalins morfologi*, men för den här övningen kan du välja bort Morfologin och Grundmaterialet, med hjälp av lexikonväljaren. Som ett experiment ska vi ta en titt på ord som beskriver olika slags människor.

- 4.1 Sök fram alla adjektiv som börjar med antingen "bond" eller "adel". Ögna igenom resultatet och bilda dig en uppfattning om hur dåtidens bönder porträtteras.
- 4.2 Sök efter adjektiv där texten (*vad som helst*) innehåller "flicka" och sedan efter adjektiv som berör "pojke".
- 4.3 Om du vill fortsätta, så gör en fritextsökning efter "qvinna" eller "karl".

APPENDIX A

ORDKLASSER I SALDO

nn	substantiv
av	adjektiv
vb	verb
pm	egennamn
ab	adverb
in	interjektion
pp	preposition
nl	numeral
pn	pronomen
sn	subjunktion
kn	konjunktion
al	artikel
ie	infinitivmärke
mx	flerordsprefix
sx	prefix
abh	adverbsuffix
avh	adjektivsuffix
nnh	substantivsuffix
nnm	substantiv, flerording
nna	substantiv, förkortning
avm	adjektiv, flerording
ava	adjektiv, förkortning
vbm	verb, flerording
vba	verb, förkortning
pmm	egennamn, flerording
pma	egennamn, förkortning
abm	adverb, flerording
aba	adverb, förkortning
pnm	pronomen, flerording
inm	interjektion, flerording
ppm	preposition, flerording
ppa	preposition, förkortning
nlm	numeral, flerording
knm	konjunktion, flerording
snm	subjunktion, flerording
kna	konjunktion, förkortning
ssm	flerording, sats

APPENDIX B

INHERENTA DRAG I SALDO

Substantiv

n	neutrum
u	utrum
v	neutrum-utrum
p	plural

Pronomen

sg	singular
p1	första person
p2	andra person
p3	tredje person
pl	plural

För egennamn

aa	artifakt		
ac	dator		
ae	mat		
af	flygtransport		
ag	marktransport		
am	medicinsk artifakt	op	politisk organisation
ap	priser	os	sportorganisation
aw	vattentransport	pa	djur
ec	kulturhändelse	pc	kollektiv
eh	historisk händelse	ph	människa
en	naturlig händelse	pm	mytologisk person
er	religiös händelse	tb	botanik
es	sport	tm	medicinsk taxonomi
h	suffix	tz	zoologi
f	femininum	wa	konst
la	astronomisk plats	wb	böcker
lf	plats, facilitet	wc	föreställningar
lg	geografisk plats	wm	media
lp	politisk plats	w	neutrum-plural
ls	gator	wn	nyheter
m	maskulinum	wp	föreställningar
oa	flygindustri		
oc	kulturorganisation		
oe	utbildningshändelse		
og	regeringsorganisation		
om	mediaorganisation		