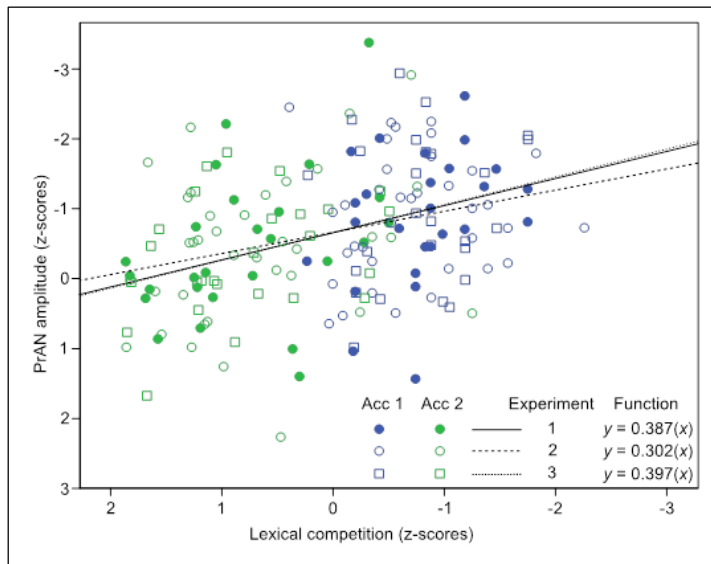


Humanistlaboratoriet vid Lunds universitet är involverat i flera Swe-Clarín-relaterade forskningssamarbeten. I detta nyhetsbrev fokuserar vi på två av dem.

I det första projektet hjälper vi en grupp neurolinguister (Pelle Söderström, Mikael Roll och Merle Horne) genom att utveckla en metod för beräkning av en lexikal konkurrens. Metoden består av att konstruera en databas med ordinitiala fragment (WIFs) och stavelser som extraherats från det svenska NST lexikonet och sedan frekvensviktats med ordfrekvenser från den svenska PAROLE-korpusen. Detta användes sedan i en studie av ERP-effekter. Studien presenterar resultat från tre experiment där vi undersökte en föreslagen ERP-effekt, pre-aktiveringsnegativitet (PrAN). PrAN har tidigare visat sig som respons på ordinitiala fragment och hypotesen var att det är ett index på prediktiv visshet om hur ett ord kommer att sluta. Det konstaterades att PrAN är känslig för lexikal konkurrens, särskilt när det gäller antalet ordfortsättningar som kan följa ett visst ords inledande fragment. Detta stödjer tanken att det speglar en pre-aktiveringsmekanismen i ord.



Linjära regressionslinjer för de tre experimenten. Den heldragna linjen och punkter representerar data från Experiment 1, den streckade linjen och runda ofyllda punkter representerar data från experiment 2, och den prickade linjen och kvadraterna representerar data från Experiment 3. Kurvan visar att PrAN-amplituder blir mer negativa med minskande lexikal konkurrens för alla experiment, vilket bekräftar hypotesen i föreliggande studie.

```
b_ao+ NN SINGLE ACC1 8 boys båg bål bår bård bås båt båts
b_ao+ NN INITIAL ACC1 9 båken bålen bälens bålet båren båsen båset båte
båtens
b_ao+ NN INITIAL ACC2 83 bågar bågarna bäge bågskytta bågskyttarna
bågskytte bågsåg bälur bälverk bälverket bårar bårarna bårbararna bårde
bårhus bårhusen bårhuset bårtäcke bårtäcket batar båtarna båtarnas
båtbenet båtbyggen båtbyggen båtbyggare båtbyggaren båtbyggarna
båtbyggeri båtbyggeriet båtflyktingar båtflyktingarna båtfolk båtfolket
båtfärd båtfärden båtförare båtförbindelse båthamn båthamnarna båthamne
båthus båthuset båtklubb båtklubbar båtklubbar båtklubben båtlast
båtlaster båtliv båtmotorer båtmassa båtmässan båtmässor
båtmossa båtmössan båtmössor båtnad båtolucky båtplats båtplatser
båtplatserna båtregetret båtresan båtresor båtskatt båtsåsong
båtsåsongen båttjuvar båtttrafik båtttrafiken båttur båtturen båtturer
båttyp båttypen båtttyper båtvarv båtägare båtägaren båtägarna
```

Exempel på lexikala konkurrenter för WIF bå [bo:] (kodat som b_ao + i lexikonet). På varje rad, från vänster till höger, finns fälten (separerade med mellanslag) WIF, ordklass (NN = substantiv), stavelseposition, ordaccent och antalet konkurrenter. Detta följs av en lista av konkurrenterna.

En demoversion av databasen finns på

goo.gl/kDi04J och ett manuskript som beskriver forskningen har lämnats in till tidskriften Brain Research.

I det andra projektet samarbetar vi med projektet MEPAC som drivs av Anna W Gustafsson och Charlotte Hommerberg. Detta projekt undersöker användningen av metaforer i upplevelsen av vård i livets slutskede. Det handlar om en systematisk analys av de metaforer som används av olika intressentgrupper (patienter, obetalda anhörigvårdare och vårdpersonal) i en korpus bestående av intervjuer och bidrag till online-forum. Hittills har vi främst varit inblandade i datainsamlingen, men vi har genomfört en explorativ ämnesmodelleringsanalys med hjälp av LDA. Ett sätt att visualisera detta är genom LDAvis, där ord och ämnen kan utforskas interaktivt (demo på goo.gl/gIBRK3) Det är ett något trubbigt verktyg, men man kan identifiera vissa ämnen som t ex 'behandling', 'familj' och 'kost'. En ansökan om extern finansiering av denna forskning har lämnats in till Crafoordska stiftelsen.

Samlingarna och forskningen

I slutet av februari 2016 arrangerade Digisam tillsammans med Riksbankens Jubileumsfond <http://www.rj.se> Nordiska museet <http://www.nordiskamuseet.se> och SWE-CLARIN <https://sweclarin.se> en seminariedag för de projekt som beviljades stöd i RJ/KVHAA-satsningen på Samlingarna och forskningen 2015 <http://www.vitterhetsakad.se/kalendarium/aktuellt>

Bakgrunden till dagen är att vi på Digisam ser stora möjligheter i en fördjupad dialog mellan minnesinstitutioner och universitet. Vi har med intresse följt RJ:s och KVHAA:s initiativ för att utveckla forskningsmöjligheterna på samlingar inom de svenska minnesinstitutionerna, och ser också en växande entusiasm för digitala humaniora i universitetsvärlden. Därför kändes det motiverat att ordna ett seminarium där de olika projekten skulle kunna lära känna varandra och kanske hitta nya samarbetsmöjligheter. Dessutom ville vi ta tillfället i akt att visa hur språkteknologin genom SWE-CLARIN kan utgöra ett stöd för humanistiska forskare. Lars Borin presenterade satsningen som sådan och han och hans kollegor fick under dagen många frågor om sådant som *topic modelling* och *name entity recognition*.

Dagen inleddes av Pelle Snickars, som presenterade ”Digitala modeller”, ett samarbete mellan HumLab Umeå och Tekniska museet. I tre delprojekt kommer man att arbeta med olika typer av material: Polhems mekaniska alfabet, arkivet efter Carl Sahlin och museets årsbok Daedalus. Material kommer att digitaliseras och bearbetas med olika tekniker och i planerna finns både 3D-scanning, Minecraft och kvalificerad språkteknologi. Se vidare: <http://pellesnickars.se/2016/02/forsta-projektpresentationen-av-digitala-modeller/>

Patrik Granholm och Roger Andersson presenterade ”TTT: Text till tiden!”, där KB samarbetar med UB Uppsala och Centrum för medeltidsstudier vid Stockholms universitet. Projektet syftar till att skapa en fullständig katalog över de 234 fornsvenska handskrifter som förvaras på biblioteken. Det inkluderar även TEI-kodning <http://www.tei-c.org/index.xml> av innehållet och inblandade forskare fokuserar på sådant som språkväxling i texterna, form, kontext och influens. Se vidare: <http://1drv.ms/1RzhT2X> & <http://www.kb.se/aktuellt/nyheter/2015/Sveriges-medeltida-handskrifter-lyfts-fram-i-ljus>

Claes Gejrot, Christian Tomner och Carl-Niclas Odenbring presenterade ”TORA (*TOPografiskt register på RA*)”, där Riksarkivet samarbetar med Institutet för språk och folkminnen, SLU, Stockholms och Uppsala universitet. Målet är att skapa en nationell standard för att sammankoppla

information om geografiska enheter (gårdstomter, ortsnamn) i olika historiska källor och databaser. Se vidare: <https://riksarkivet.se/Nyhetsarkiv?item=108498>

Jonas Engman och Jonas Hedberg presenterade ”Vardagens bilder”, ett samarbete mellan Nordiska museet och Stockholms universitet. Projektet utgår ifrån ett par större samlingar i museets arkiv – Gunnar Lundhs fotografier (300 000) och transkriberade frågelistsvar (500 000 sidor) – och undersöker bl a hur de i olika former tecknar historien om folkhemmets framväxt och hur de olika källorna förhåller sig till varandra.

Tyvärr saknade vi representation från Arosenius-projektet som drivs av Centrum för digital humaniora vid Göteborgs universitet i samverkan med bl a UB Göteborg, Göteborgs konstmuseum och Nationalmuseum. Projektet syftar till att sammanställa alla slags material om Arosenius – målningar, skisser, fotografi, brev, föremål etc – för att undersöka hur digitaliseringen kan göra det möjligt att ”frammana konstnären ur arkiven”. Mer info finns här: <https://www.facebook.com/arsenius/?fref=ts> & <http://www.gu.se/omuniversitetet/aktuellt/nyheter/detalj//digitalt-museum-over-ivar-arsenius-skapas-vid-goteborgs-universitet.cid1314928>

Under dagen fick vi också veta att Riksbankens Jubileumsfond fått in 40 fullständiga ansökningar på årets utlysning för Samlingarna och forskningen. Bedömningen är att ansökningarna generellt håller hög kvalitet, och man hoppas denna gång kunna stödja kanske åtta eller tio projekt. Beslut kommer till midsommar, och vi är många som ser fram emot nya forskningsfrågor, nya samarbeten och nya givande seminariedagar för gemensam utveckling av den digitala humanioran.

Johanna Berg, Digisam

Kalender

9 maj: Swe-Clarin på turné-stopp i Stockholm. Temat är Nordiska museets frågelistor.

<https://sweclarin.se/swe/evenemang/swe-clarin-pa-turne-I>

September: Lansering av Swe-Clarins verktygslåda

November: Ett Swe-Clarin på turné-stopp planeras i anslutning till SLTC, den svenska språkteknologi konferensen, i Umeå.

Framtida nyheter

Meddela gärna samordningsgruppen (info@sweclarin.se) om ni har information till nyhetsbrevet. Om man inte får nyhetsbrevet skickat kan man anmäla sig till e-postlistan här:

http://lists.sweclarin.se/mailman/listinfo/news_lists_sweclarin.se

