

Framdriftsrapport for Divvun

Nr 17, mai 2006

by Sjur N. Moshagen

Table of contents

1 Gjennomført arbeid.....	2
2 Framdrift.....	2

1 Gjennomført arbeid

I mai vart vi nesten ferdige med å velja underleverandør, og korpusinnsamlinga sig seint men fint framover. Det lingvistiske arbeidet går jamt vidare. Arbeidet har vore fordelt på desse hovudpunkta:

- tekstinnsamling: vi har no fått mykje tekst av Min Áigi, lovnad om ordbok frå Olavi Korhonen, Davvi Girji er positive, NSI har signert, og vi ser fruktene av det tidlegare arbeidet vårt - det er triveleg!
- korpusorganisering: framleis meir arbeid med forfining og feilretting av verktya våre. Dette arbeidet ligg no i hovudsak på Universitetsprosjektet, og vi nyt godt av det dei gjer (og dei nyt godt av innsamlingsarbeidet vårt).
- Lingvistisk arbeid:
 - lulesamisk: arbeidet har vore konsentrert om å leggja til resten av orda frå Kintel si ordbok, og forfina orddelings
 - nordsamisk: det vart lagt ned ein god del arbeid i å forfina den lingvistiske informasjonen i leksikonet, for å gjera orddelinga betre. Alle samansette ord er no merka med ordgrense, og alle unntak frå reglane er òg merka. I tillegg vart det halde ein lingvistisk "workshop" i samband med at Maaren har vore heile mai i Tromsø.
- namneleksikon: arbeidet med å føra namnebasen inn i ein fleirspråkleg XML-base i risten.no går sakte framover, og ikkje på langt nær så fort som planlagt. Sjå kommentar til sist.
- Orddeling: orddelinga blir stadig betre, og ved utgangen av mai var unntaksmerkinga i både nordsamisk og lulesamisk i all hovudsak ferdig. Dei regelbaserte orddelingspunkta er òg så godt som feilfrie no, og det einaste som manglar enno er leksikonstøtta orddeling av nokre problematiske ordgrenser.
- Utlysing av delar av prosjektet: i mai vart vi nesten ferdige med å velja underleverandør, men for mange uklåre punkt i tilboda gjorde at vi måtte utsetja avgjerda nokre veker for å be om meir informasjon. Endeleg val vil bli gjort i juni.

2 Framdrift

Prosjektet har for det meste gått fint framover i mai òg, men arbeidet med namnebasen går seint. Orsaken er i hovudsak konfidensiell, og vil bli diskutert direkte med prosjekteigar.