

Framdriftsrapport for Divvun

Nr 15, mars 2006

by Sjur N. Moshagen

Table of contents

1 Gjennomført arbeid.....	2
2 Framdrift.....	3

1 Gjennomført arbeid

I mars gjekk arbeidet med det lulesamiske leksikonet vidare, og mange opne normeringsspørsmål kring nordsamisk vart avklara. Vi har laga ein fyrste prototyp av ein samisk orddelar, og korpusinnsamlinga har gått vidare, med både framsteg og baksteg. Arbeidet har vore fordelt på desse hovudpunkta:

- tekstinnsamling: diskusjonen med I#ut+grafisk partner gjekk vidare, dårlegare og dårlegare. Som eg alt har rapportert direkte, vil dei ikkje gje oss dei einaste filene dei har, og vil ha - etter vår meining - grovt overbetalt for å konvertera filene dei har til noko som vi kan bruka. Argumenta dei har for ikkje å gje filene sine til oss, held ikkje, anna enn at det er eit prinsipp for dei. Derimot er andre forlag meir positive enn vi hadde venta oss, og både Áššu og Min Áigi er meir eller mindre klare til å gje oss tekstar i massevis.
- korpusorganisering: meir arbeid med forfining og feilretting av verktya våre - dei blir stadig betre!
- Lingvistisk arbeid:
 - lulesamisk: i mars vart det lulesamiske leksikonet bygt ut vidare, og er no på drygt 18 000 ord + namn.
 - nordsamisk: nye ord vart lagt til, og vi har oppdatert normeringsdokumentet vårt med alle vedtak frå det siste SGL-møtet, samt frå tidlegare møte så langt det er relevant for arbeidet vårt.
- namneleksikon: arbeidet med å utvikla risten.no for formåla til Divvun-prosjektet har gått vidare, og den grunnleggjande infrastrukturen for å støtta fleire parallelle termsamlingar og ordbøker er no på plass, og fungerer. Det står framleis att ein god del koding før vi har noko som kan visast fram, men det går framover.
- Stavekontroll: ingen ting i mars heller, vi prioriterer arbeidet med den fleirspråklege namnebasen (inkl. risten.no) før vi går vidare med stavekontrollar (det er Tomi som skal gjera meirparten av arbeidet med begge).
- Orddeling: på tampen av mars dukka det opp fleire spørsmål kring orddeling i samband med utlysinga av delar av prosjektet. Desse spørsmåla spora oss, og ikkje minst Trond frå Universitetsprosjektet, til å laga ein avstavingsmodul som gjev alle potensielle orddelingspunkt i samiske ord. Det står enno att litt arbeid før orddelinga er heilt korrekt, men alt no er det mogleg å dela ord på [ei av sidene til Universitetsprosjektet](#). Med dette arbeidet kan vi no anten pakka inn vår eigen språkteknologi til ein ferdig orddelar, eller vi kan generera data som andre kan bruka til å laga ein eigen orddelar ut i frå.
- Utlysing av delar av prosjektet: i mars har arbeidet med utlysinga konsentrert seg på å svara på spørsmål frå interesserte firma. Til no er det berre eitt firma som har kome med spørsmål. Det er heilt i det uvisse om det er fordi dei andre firmaa har all informasjon dei treng, eller fordi dei ikkje er interesserte. 18. april veit vi svaret.

2 Framdrift

Prosjektet har i all hovudsak gått fint framover i mars.

Last modified: \$Date: 2006-11-23 12:13:18 +0100 (tor, 23 nov 2006) \$, by \$Author: boerre \$