

Framdriftsrapport for Divvun

Nr 9, september 2005

by Sjur N. Moshagen

Table of contents

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

1 Gjennomført arbeid

September har gått i vanleg takt. Eit viktig framsteg har vore arbeidet med lulesamisk. Det arbeidet som har vorte gjort har vore fordelt på desse hovudpunkta:

- Dokumentasjon: nettstaden til prosjektet blir no automatisk oppdatert, ein gong i halvtimen, slik at ny og oppdatert dokumentasjon automatisk blir tilgjengeleg.
- tekstinnsamling:
 - Korpuskontraktane: i september har alle kontraktane vorte omsett til norsk, og omsetjingane har fått ein fyrste gjennomgang og sjekk. Vi har prøvd å få tilbakemeldingar frå Kimmo Koskenniemi til omsetjingane, og vi prøver no å få ein jurist til å sjå på dei.
 - Innsamling: vi har arbeidd med å få tak i bibeltekstar, og byrja å byggja opp infrastruktur som skal gjera det lettare å leggja til tekstar til korpussamlinga.
- korpusorganisering: arbeidet med å omorganisera korpussamlinga, som vi byrja på i Helsingfors, har gått vidare i september, særleg gjeld dette arbeidet med å konvertera frå innkomande format (Word, PDF, m.fl.) til ein felles XML-representasjon.
- Lingvistisk arbeid:
 - lulesamisk: arbeidet har gått vidare framover i september, og heile bøyingsmorfologien er no sjekka og korrigert.
 - nordsamisk: vi har fortsett med å leggja til nye ord og retta feil som har vorte oppdaga.
 - anna: vi har identifisert fleire opne saker som vi må arbeida med framover, og særleg gjeld det arbeidet med å byggja eit felles namneleksikon. Eit felles namneleksikon vil samla alle namn i ein felles base, slik at vi sikrar oss at alle språkverkty dekkjer dei same namna. Dei fleste namn er like på tvers av språk (det gjeld til dømes alle firmanamn, dei fleste utanlandske og nordiske stadnamn, organisasjonar m.m.). Derimot vil bøyinga av namna, og måten dei lagar samansetjingar m.m. på vera spesifikt for kvart språk. Vi har i september starta eit internt prosjekt for å utvikla eit felles namneformat som tek omsyn til dette (fleire detaljar i [møtereferata våre](#), særleg [dette](#)).
- Termbasen: det har ikkje vorte gjort vesentleg arbeid på termbasen i september.
- Stavekontroll: i etterkant av konferansen og demonstrasjonen i månadskiftet august/september har vi sett nærare på alternativ til den stavekontrollmotoren vi brukte til demonstrasjonen. Ein lovande kandidat er [Hunspell](#). Demoversjonen vi presenterte i Helsingfors har òg vorte betre.

2 Framdrift

Prosjektet har gått fint framover i september òg.

Last modified: \$Date: 2010-11-11 22:54:48 +0100 (tor, 11 nov 2010) \$, by \$Author: boerre \$