

Framdriftsrapport for Divvun

Nr 8, august 2005

by Sjur N. Moshagen

Table of contents

1 Gjennomført arbeid.....	2
1.1 Samling og kurs.....	3
1.2 Demonstrasjon.....	3
2 Framdrift.....	3

1 Gjennomført arbeid

August hadde enno litt feriepreg i starten, men har deretter vore ein effektiv månad, der det viktigaste er at vi har fått fram den aller fyrste prototypen av ein nordsamisk stavekontroll. Det arbeidet som har vorte gjort har vore fordelt på desse hovudpunkta:

- Dokumentasjon: nettstaden til prosjektet vart flytta over på ein ny server i august, for å retta til nokre problem vi hadde med den opphavlege servaren. Det er framleis nokre småsaker som står att, men i all hovudsak fungerer nettstaden vår bra, og blir jamleg oppdatert. Arbeidet med dokumentasjonen handlar no om å skriva ny og oppdatera eksisterande dokumentasjon.
- tekstinnsamling:
 - i august har arbeidet med kontraktane for tekstlisensieringa gått framover, og vi har no fått tre av fire kontraktar som i hop utgjer grunnlaget for korpuslisensmodellen til Helsingfors universitet. Kontraktane vil bli omsett til norsk, og deretter sendt vidare til jurist for korrekt formulering og kvalitetssikring. Dei ferdige kontraktane vil bli lagt ut på nettet i lag med dei finske versjonane slik at andre kan bruka kontraktane som modell for eigne prosjekt. Alt arbeidet med kontraktane skjer i samarbeid med Helsingfors universitet, Lingvistiska institutionen.
 - Børre har vore i kontakt med Anders Kintel om den lulesamiske ordboka - han vil ha ein vanleg lisens (jf det førre punktet), men det hadde vore meir fruktbart for alle partar om vi kunne få til eit praktisk samarbeid om vidareutviklinga av det lulesamiske leksikonet. Dette må diskuteras nærmare med alle involverte partar, fyrst dei formelle sidene ved og interessa for eit slikt samarbeid, og seinare dei praktiske detaljane.
- korpusorganisering: vi har vore missnøgd med eit punkt i den modellen vi har brukt tidlegare, og vi la ned ein del tid på samlinga i Helsingfors med å koma fram til eit betre alternativ. I tillegg har verktya våre for å bruka korpusa i det lingvistiske arbeidet vorte utvikla vidare.
- Lingvistisk arbeid: Arbeidet med lulesamisk har gått kraftig framover i august, og heile bøyingsmorfologien er no sjekka og korrigerert. Arbeidet med nordsamisk har òg gått vidare, der har det mest handla om å leggja til nye ord og retta feil som har vorte oppdaga.
- Termbasen: arbeidet med å retta opp feil som har vorte oppdaga etter opninga har halde fram, men ganske lite har vorte gjort i august
- Stavekontroll: vi laga ein aller fyrste versjon av ein nordsamisk stavekontroll for intern testing og vidare utvikling. Stavekontrollen vart demonstrert på konferansen FSMNLP-05 i Helsingfors i månadskiftet august/september.

1.1 Samling og kurs

Vi hadde ei svært nyttig samling ved utgangen av august, i forkant av workshopen og konferansen [FSMNL](#)P. Fleire både praktiske og lingvistiske problem vart løyste eller tok viktige steg framover.

Sjølve konferansen var lærerik, men var innimellom for abstrakt og matematisk til at vi kunne få noko ut av han. Heldigvis fanst det trådløst nett der, slik at vi kunne arbeida i det stille når det vart for avansert for oss:-)

1.2 Demonstrasjon

Vi gjennomførde demonstrasjonen som planlagt, og med rimeleg godt resultat. Lingsoft var der og gjorde seg kjende med prosjektet vårt og resultatata så langt – andre samarbeidsaktuelle firma fanst ikkje på konferansen.

Det var fleire andre teknologidemonstrasjonar der, og fleire språkteknologiprojekt som blir gjennomført som open kjeldekode vart presentert. Vi hadde samtalar med fleire av dei, og særleg med hovudmannen bak Stuttgart Finite State Transducer-pakka ([SFST](#)). Desse prosjekta er viktige med tanke på framtida og dei banda vi i dag har til Xerox og teknologien deira, og SFST er pr i dag det mest lovande alternativet til Xerox-verktya.

2 Framdrift

Prosjektet har gått fint framover i august, og vi er i rute (gjeve at dei endringane eg føreslår for styringsgruppa blir akseptert), og til og med litt føre planen når det gjeld den fyrste stavekontrollen, som eigentleg var planlagt til seint på hausten eller føre jul.

Last modified: \$Date: 2005-09-13 17:39:28 +0200 (tir, 13 sep 2005) \$, by \$Author: sjur \$