

Lær samisk med språkteknologi

Lene Antonsen, Saara Huhmarniemi, Trond Trosterud

Samisk har mye morfologi, og det kreves derfor trening for å kunne bøye ordene godt nok til en flytende samtale. Datamaskinen hjelper eleven å trene på ord og å bøye dem, og eleven kan også forme egne setninger i dialoger og få tilbakemelding på språklige feil. Datamaskinen lager selv varianter av noen av oppgavene. Dermed vil eleven kunne bruke programmene på nytt og på nytt.

Språklæringsprogrammet OAHPA! er laget av Senter for samisk språkteknologi ved Universitetet i Tromsø, og er et supplement til undervisninga på skolen og universitetet.



The screenshot shows the OAHPA! website interface. At the top, the title "OAHPA!" is written in large, stylized letters. Below it, a brief description states: "OAHPA! er program for ungdom og voksne som vil trene seg på nordsamisk. I høyre marg kan du velge hjelpespråk (leksikon er foreløpig bare oversatt til norsk), og om du vil ha østlig eller vestlig dialekt. Nedenfor velger du program." The main content area features six program icons: MORFA-S (Bøy enkeltord), VASTA (Svar på spørsmål), LEKSA (Øv på ord og stedsnavn), MORFA-C (Bøy ord i setninger), SAHKA (Dialoger), and NUMRA (Øv på tall). A sidebar on the right contains navigation links: "Hjelpespråk" (Norsk, Nordsamisk, Vestlig), "Veiledning for lærere og studenter", "VISL for nordsamisk", "Om OAHPA!", and "Tilbakemelding". The footer includes the copyright notice: "Copyright 2009 Universitetet i Tromsø Ta kontakt oahpa@hum.uit.no".

<http://oahpa.uit.no/>

```
<entry>
<lemma>monni</lemma>
<pos class="N"/>
<translations>
<tr xml:lang="nob">egg</tr>
<tr xml:lang="fin">muna</tr>
</translations>
<semantics>
<sem class="FOOD-GROCERY"/>
</semantics>
<stem class="bisyllabic" diphthong="no"
gradation="yes" soggi="i" rime="ø"/>
<dialect class="NOT-KJ"/>
<sources>
<book name="d1"/>
<book name="sara"/>
<book name="algu"/>
</sources>
</entry>
```

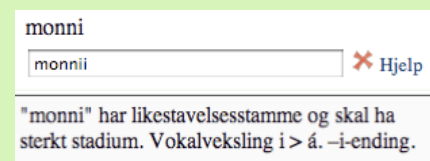
Leksikonet

Programmene har en del felles ressurser, f.eks. et leksikon og en automat som brukes til å lage bøyingsformene som kommer til syne i programmene. I leksikonet er det informasjon om dialektvariasjon, semantiske klasser og bøyning.

Feil bøyning?

Hvis eleven ikke har bøyd ordet riktig i øvingene, kan hun be om hjelp og prøve en gang til istedenfor for å få det riktige svaret med en gang. Meldinga formes på grunnlag av en kombinasjon av informasjon i leksikonet til venstre, og selve bøyingsoppgaven.

Bøy *monni* i illativ. (Riktig svar er *monnái*.)

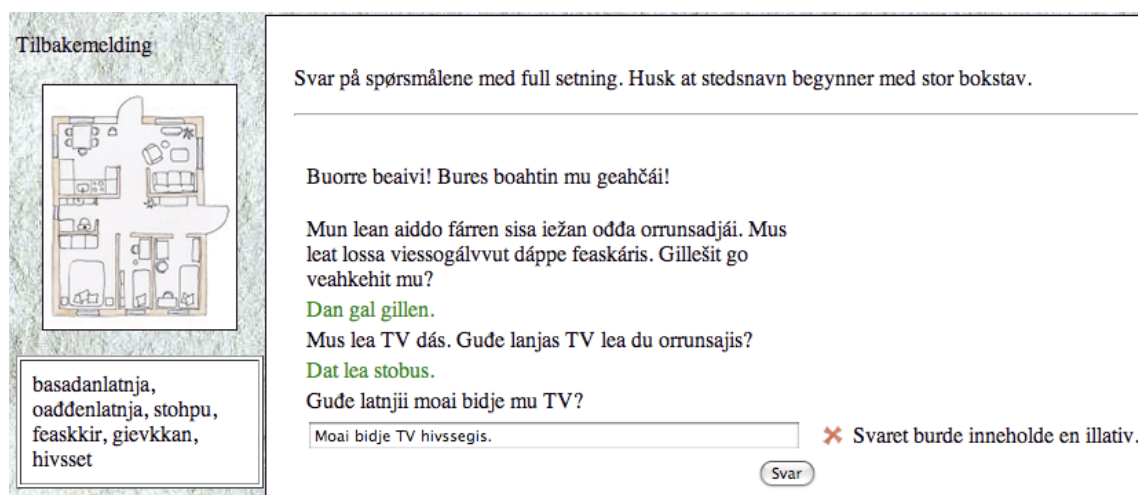


The screenshot shows the leksikon interface. It features a search box with the word "monni" entered. Below the search box, the word "monnii" is displayed. To the right of the search box is a red "X" icon and the word "Hjelp". Below the search box, a message reads: "monni har likestavellesstamme og skal ha sterkt stadium. Vokalveksling i > å. -i-ending."

Artikkel

Antonsen, Lene, Saara Huhmarniemi and Trond Trosterud 2009: Interactive pedagogical programs based on constraint grammar. Proceedings of the 17th Nordic Conference of Computational Linguistics. *Nealt Proceedings Series* 4. <http://dSPACE.utlib.ee/dspace/handle/10062/9206>

Interaktive dialoger med eleven

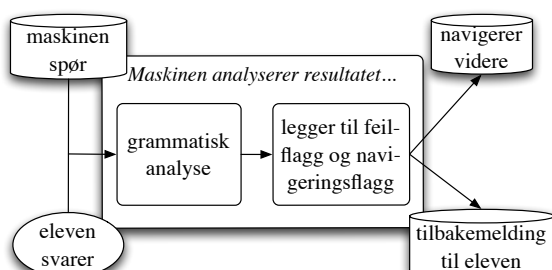


The screenshot shows an interactive dialog interface. On the left, there is a "Tilbakemelding" (Feedback) box containing a floor plan image and the text: "basadanlatnja, oaddenlatnja, stohpu, feaskkir, gievkkan, hivsset". The main dialog area contains the following text: "Svar på spørsmålene med full setning. Husk at stedsnavn begynner med stor bokstav." "Buorre beaivi! Bures boahhtin mu geahčái!" "Mun lean aiddo fáren sisa iežan odda orrunsjái. Mus leat lossa viessogálvvut dáppe feaskáris. Gillešit go veahkehit mu?" "Dan gal gillen." "Mus lea TV dás. Guđe lanjas TV lea du orrunsjáis?" "Dat lea stobus." "Guđe latnjii moai bidje mu TV?" Below this text is a text input field containing "Moai bidje TV hivssegis." and a "Svar" button. To the right of the input field is a red "X" icon and the text "Svaret burde inneholde en illativ."

Vasta og *Sahka* lar eleven formulere sine egne svar. Analysatoren inneholder 100.000 ord/navn og kan dessuten bøye dem. Maskinen analyser setningen og legger til feilflagg og navigeringsflagg.

Ovenfor er den første delen av en dialog i *Sahka* om å møblere en leilighet. Nedenfor er analysen av det tredje spørsmål/svar i dialogen.

```
<Gude>
"guhte" Pron Interr Sg Gen &grm-missing-Ill
<latnjii>
"latnja" N Sg Ill
<moai>
"mun" Pron Pers Du1 Nom
<bidje>
"bidjat" V TV Ind Prs Du1
<mu>
"mun" Pron Pers Sg1 Gen
<TV>
"TV" N ACR Sg Acc
<Agst>
"asahka" QDL gosa_bidjat_TV &dia-hivsset
<Moai>
"mun" Pron Pers Du1 Nom
<bidje>
"bidjat" V TV Ind Prs Du1
<TV>
"TV" N ACR Sg Gen
<hivssegis>
"hivsset" N Sg Loc
<.>
"." CLB
```



Grammatiske feil



Maskinen gir tilbakemelding om grammatiske feil. I det tredje spørsmålet i dialogen til venstre, spør maskinen: "I hvilket rom skal vi sette TVen?" Eleven svarer *Moai bidje TV hivssegis* ("Vi setter TVen på do."), med lokativformen *hivssegis* istedenfor den korrekte illativformen *hivssegii*.

Maskinen analyserer spørsmålet og svaret og setter på et feilflagg (&grm-missing-Ill) fordi spørsmålet krever at svaret skal være i illativ.

Feilflagget aktiverer en tilbakemelding på det språket som er valgt (i dette eksemplet er det norsk), og eleven får meldinga *Svaret burde inneholde en illativ*.

Navigering



Navigeringsflagget gir semantisk informasjon, f.eks. om svaret er positivt eller negativt, eller om det inneholder informasjon som skal lagres, slik at maskinen kan bruke elevens navn og informasjon om bosted eller bilmerke i et spørsmål eller en kommentarer.

I eksemplet til venstre er spørsmålet: "I hvilket rom setter vi TVen?" Svaret vil være bestemmende for hva som skjer videre. Forslaget om å plassere TVen på do, gir flagget &dia-hivsset, som gjør at svaret i dialogen blir "Det er ingen god ide. Prøv igjen!"

Når eleven svarer på spørsmål alder, vil maskinen navigere til en passende grein av dialogen. F.eks. vil man ikke spørre en person under 18 år om hun har bil.