

Språkinläring: mänsklig och datorstödd

Ola Knutsson
knutsson@csc.kth.se
Språkteknologikursen
KTH

Ola Knutsson
knutsson@csc.kth.se
Språkteknologikursen
KTH

Innehåll

- Några olika typer av system för datorstödd språkinlärning
- Vad handlar språkinlärning om?
- Språkteknologins roll
- Undersökning av språkteknologins roll
- Exemplet Grim

2

- Några olika typer av system för datorstött språkinläring
- Vad handlar språkinläring om?
- Språkteknologins roll
- Undersökning av språkteknologins roll
- Exemplet Grim

Varför datorer i språkutbildning?

- Datorn som lärare? Assistent?
- Tillgänglig, uttröttlig, blir inte arg
- Multimodal återkoppling
- Individualisering
- Variation i material och stöd för olika inlärningsstilar/strategier
- Upptäckande i enorma mängder språkliga data.

- Datorn som lärare? Assistent?
- Tillgänglig, outtröttlig, blir inte arg
- Multimodal återkoppling
- Individualisering
- Variation i material och stöd för olika inlärningsstilar/strategier
- Upptäckande i enorma mängder språkliga data.

Olika typer av program för datorstödd språkinläring

- CALL
- ICALL
- Intelligent Tutoring Systems

4

- CALL
- ICALL
- Intelligent Tutoring Systems

[illegible]

Safari File Edit View History Bookmarks Window Help
Glossier-Web
omns 14.37.13
http://edut.let.rug.nl/abu/glossier/Glossier/ GURT 2005

Home	Help	Phylogenie	Dictionnaire	Exemples
Find	con	off	con	off

DE LA TERRE A LA LUNE

Trajet Direct en 97 Heures 20 Minutes

par Jules Verne

I
LE GUN-CLUB

Pendant la guerre fédérale des États-Unis, un nouveau club énergique s'établit dans la ville de Baltimore, en plein Maryland. On sait que toute énergie instinctive se développe chez ce peuple d'armateurs, de marchands et de mécaniciens. De simples négociants, leur honneur pour comptoir par improviser capitaines, cobolons, généraux, sans avoir passé par les écoles d'application de West-Point (École militaire des États-Unis); ils égaient bientôt dans «l'art de la guerre» leurs collègues du vieux continent, et comme eux ils remportent des victoires à force de prodigier les boulets, les millions et les hommes.

Mais en qu'il les Américains surpassent singulièrement les Européens, ce fut dans la science de la ballistique. Non que leurs armes atteignent un plus haut degré de perfection, mais elles offrent des résistances insolites, et eurent par conséquent des portées incommensurables. En fait de tir, les Anglais, les Français, les Prussiens, n'ont plus rien à apprendre; mais leurs canons et leurs mortiers ne sont que des pistolets de poche auprès des formidables engins de l'artillerie américaine.

Ceci ne doit étonner personne. Les Yankees, ces premiers mécaniciens du monde, sont ingénieurs, comme les Italiens sont musiciens et les Allemands métrologues, de naissance. Rien de plus naturel, dès lors, que de leur voir apporter dans la science de la ballistique leur ancienne ingéniosité. De là ces canons géants, beaucoup moins utiles que les machines à coudre, mais aussi étonnantes et encore plus admirées.

Analysis:
Federal=Fems+Adj

fédéral [fɛdɛral] 10.1 bonds- > bondgenootschappelijk, federal 0.2 verbond- > federale- 1.1 la République
Fédérale allemande des Bundesrepublik Dalsland 7.1 les Fédéraux de nondeelingen, federalen

... directement sous les feux des Confédérés. Rangons d'aussi près que possible la drôte de la rade, quitte à recevoir la bordée des batteries fédérales. Vous avez un homme sûr à la barre ?—Oui, capitaine.—Faites descendre vos canaux et les feux du bord. C'est déjà trop, beaucoup trop, des ...

Les forcenés de blues - Jules Verne Hetzel, Paris
http://web.cnam.fr/ABU/abu_server.html
pub/ABU/anteABU/bluics.txt

WERTi - Working with English Real Text

- University of Ohio
- <http://prospero.ling.ohio-state.edu/WERTi/index.py/main>

7

REAP - Reader-Specific Lexical Practice for Improved Reading Comprehension

- <http://reap.cs.cmu.edu/demo/>

8

Språkinläring: typer

- Förstaspråk
- Andraspråk
- Främmande språk
- Tillägnande
- Lärande

Teorier för andraspråksinläring

- Input-hypotesen (Krashen)
- Interaktionshypotesen (Long)
- Outputhypotesen (Swain)
- Focus on form (Long)
- Sociokulturell teorier (Lantolf mfl)

10

Språkteknologi i datorstödd språkinläring (CALL)

- Språkgranskning är ganska vanligt
- Vanlig text – gör om meningar till frågor och svar.
- Textsammanfattning
- Talteknologi – uttalsträning
- Automatisk betygsättning (med LSA)
- ”Känna av begrepp” som är svåra för studenterna
- Eliza-liknande program

Möjligheter och risker med språkteknologi

- Språkteknologiska program kan ge “feedback” på språkbrukarens obegränsade produktion.
- Falsk “feedback” kan lura användaren
- Brist på “feedback” kan leda användaren fel

Baskomponenter i andraspråksinlärning (Gass 1997; Chapelle, 1998)

- Input (of target language)
- Apperception (noticing aspects in TL)
- Comprehension (+semantic -syntax)
- Intake (+semantic +syntax)
- Integration (into learners' interlanguage)
- Output (L2 production)

Exempel: Grim – en lärmiljö med fokus på svenska

Språkliga datorstöd och

andraspråksinlärning: Tessy Cerratto,
Kerstin Severinson Eklundh, Stefan Westlund

CrossCheck: Viggo Kann, Jonas Sjöbergh,
Johnny Bigert, Lars Borin, Beb Lindberg, m.fl

Språkverktyg utvecklade på KTH Nada

- Stava (Kann & Hollman, 1992)
- Granska (1998-2001)
- CrossCheck (2002-2004) (inkl. ProbGranska)
- GTA – a shallow parser for Swedish (2003)
- Swesum + andra verktyg

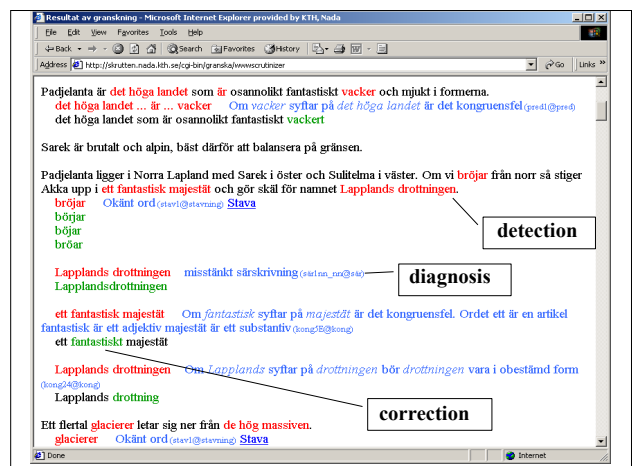
Varför fokusera på fel?

“The learners’ errors are a register of their current perspective of the target language” (James, 1998).

“Noticing a problem ‘pushes’ the learner to modify his/her output” (Swain & Lapkin, 1995): – syntactic processing mode – important for the development of IL

Datainsamling

- * Deltagande observation
- * Studenternas texter (orig.- och slutversion)
- * Enkät
- * Intervjuer
- * Användarnas bedömningar av Granska



5. **Utmärkt** – t.ex. jag förstår precis vad Granska menar
4. **Bra** – t.ex. Granska är till ganska stor hjälp
3. **Godkänt** – t.ex. jag har svårt att ta ställning, men jag chansar ...
2. **Dåligt** – t.ex. jag har svårt att ta ställning, måste slå i lärobok
1. **Obegripligt** – t.ex. jag förstår inte vad Granska menar

Antal uppgifter, vilken typ Ordförande inledde diskussionen och deltagarna hade olika uppfattning om vad en uppgift och **ett ämnet** är.
ett ämnet Om ett syftar på **ämnet** är det kongruensfel (kong22E@kong)
 (M5,D3,E2) Här borde det vara ett **ämne**
 det ämnet

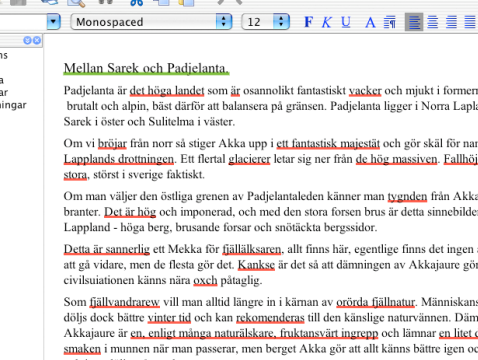
De föreslog att ett prov med flera delar / uppgifter betyder flera skriftliga uppsatser eller övningar (t. ex. grammatik eller ordkunskap) som måste göras under provet. Ett ämne beskrevs som en uppsatssfråga (t. ex. pappealighet). Efter en tolkning av vad en uppgift och ett ämne egentligen betyder samt en kort diskussion beslutade mötes **majoritetet** att en uppgift på provet är tillräcklig. (M5,D3,E5) **Här**
majoritetet Okänt ord (stav1+staving) Stava
borde det vara: beslutade majoriteten...
 majoriteten
 majoritetet

Bedömningar av falska alarm

Error type	detection mean	diagnosis mean	correction mean
Agr. NP	5	2	2
Spelling	4	3.3	2.6
Word order	5	5	5
Missing X	2.3	1.7	1.7

Resultat

- Brist på "feedback" och vilseledande "feedback"
- Olika källor med lingvistisk information – ett verktyg räcker inte.
- Fokus på form
- Att lita på programmet
- Metaspråk och grammatisk kunskap
- Genomskinlighet
- Interaktion and integration



Grim v.1.21 - Mellan Sarek och Padjelanta

default Monospaced 12 F K U A

Feltyper
 Kongruens
 Ordföljd
 Sannolika
 Stavningar
 Särskrivningar
 Verb
 Övriga

Mellan Sarek och Padjelanta.

Padjelanta är det höga landet som är osannolikt fantastiskt vacker och mjukt i formerna. Sarek är brutalt och alpin, både därför att balansera på gränsen. Padjelanta ligger i Norra Lapland med Sarek i öster och Sullelma i väster.

Om vi bröjar från norr så stiger Akka upp i ett fantastiskt majestät och gör skäl för namnet Lappland drotsningen. Ett flertal glacierer letar sig ner från de hög massiven. Fallhöjden är stora, störst i Sverige faktiskt.

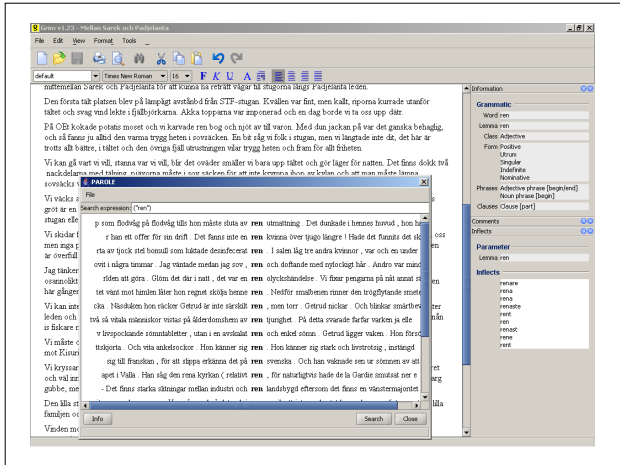
Om man väljer den östliga grenen av Padjelantaleden känner man tynden från Akka massivets branter. Det är hög och imponerad, och med den stora fören brus är detta sinnebild av Lappland - hög berg, brusande forsar och snötäcka bergssidor.

Detta är samerle ett Mekka för fjällnäsaren, all finsns här, egentligen finns det ingen anledning att gå vidare, men de flesta gör det. Kamkse är det så att dämningen av Akkajaurer gör att civilisationen känns nära oxch påtaglig.

Som fjällandrew vill man alltid längre in i kärnan av prörda fjällnatur. Människans ingrepp döljer både vinter tid och kan gekomederas till den känslige naturlivn. Dämningen av Akkajaurer är en enligt många naturläkare, fruktansvärt ingrepp och känner en litet djält smaken i munnen när man passerar, men berget Akka gör att det alltså blir en lugn och snön och isen döljer så mycket.

Med endast en SJ maskin i magen nådde jag blod socker fattig inlänzen området nära Akkastagan. Jag hade kommet med eftermiddagsbussen till Ritsem, där mötte O o g S upp, de

[illegible]



Sammanfattande slutsatser

- Det är viktigt att ge återkoppling på olika aspekter i inlärares språk och i målspråket.
- En ensemble av verktyg reducerar negativa effekter av det enskilda verktyget

Framtida arbete

- Användarstudier med Grim (en del avslutat, mer pågår)
- Statistisk återkoppling
- Innehålla en kommunikativ del? (kollaborativt lärande)

<http://skrutten.nada.kth.se/grim>

Över 6000 personer har testat
programmet

Fungerar under Windows, Linux,
Solaris, MacOSX och FreeBSD