

Het gebruik van taaltechnologie in het onderwijs

Citation for published version (APA):

Van Bruggen, J., & Van Rosmalen, P. (2009). *Het gebruik van taaltechnologie in het onderwijs*.

Document status and date:

Published: 06/01/2009

Document Version:

Peer reviewed version

Document license:

CC BY-SA

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 20 Jul. 2021

Open Universiteit
www.ou.nl





Language Technologies for Lifelong Learning

Het gebruik van taaltechnologie in het onderwijs

Jan van Bruggen
Peter van Rosmalen

Maastricht, 21 Oktober 2008

OpenUniversiteitNederland

Het gebruik van taaltechnologie in het onderwijs

- Achtergrond
- Toepassing: “het beantwoorden van vragen”
- ‘Taaltechnologie’: LSA
- Project: Language Technologies for Lifelong Learning
- Discussie

Het gebruik van taaltechnologie in het onderwijs

Begeleidingstaken (te) arbeidsintensief:

- Onderwijs met focus op actieve studentinbreng
- Elektronische leeromgevingen en tools

Voorbeelden:

- Werkstukken (plagiat)
- Studievoortgang (eportfolio feedback)
- Leerproces (forum)
- Inhoudelijke vragen



1. Een toepassing: het beantwoorden van vragen



Het beantwoorden van vragen

Kritische support activiteit:

- Arbeidsintensief: frequent, werkonderbrekend
- Belangrijk voor de student

Doelen

- “sustainable” support faciliteit
- Studenten activeren en in contact brengen

Het beantwoorden van vragen

Vraag:

Ik gebruik Norton Antivirus voor het schoonhouden van mijn PC
Daarbij zit een spamfilter, bovendien heb ik nog een spamfilter van mijn
provider OrangeCable. Toch komen er gemiddeld per dag nog zo'n 30
spammiltjes door. Heeft iemand een tip voor mij?

Vraag:

Erik de Jonghe schreef zijn visie op digitale televisie in 2005. Zou zijn visie
anno 2007 genuanceerd kunnen worden nu websites als uitzending gemist alle
kanalen.nl en oldtvseries.com zijn gekomen?

Vraag:

Laatst was ik aan het chatten en in de chatbox waar ik in zat hadden mijn
medechatters allemaal mooie avatars. Ik vroeg me af hoe je aan deze avatars
komt en kan je ze ook zelf maken? Is er misschien software waarmee je je
eigen avatar kunt ontwerpen?



Hoofdstappen

1. *A student* poses a question.
2. The *system* determines:
 - text fragments to help answering the question;
 - the topic(s) of the question;
 - the most suitable peer-learners.
3. The *system* sets up a wiki with the question, the text fragments and guidelines.
4. The selected *peer-students* receive an invitation to assist.
5. *The questions poser and his peers* discuss and phrase an answer in the wiki.
6. The *question poser* closes the discussion and rates the answer.

Hoofdstappen (details)

Text fragmenten als ondersteuning bij het beantwoorden:

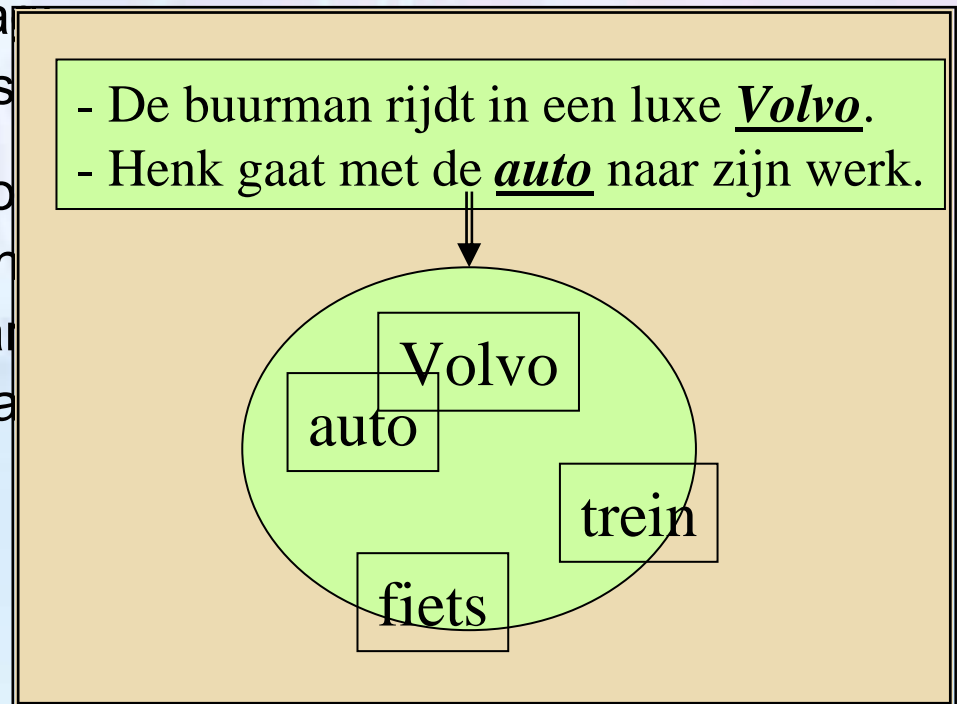
- Latent Semantic Analysis selecteert tekst in het studie materiaal

Het onderwerp van de vraag

- Latent Semantic Analysis

De 'juiste' peer-studenten op

- content competency (inhoud)
- availability (beschikbaar)
- Eligibility (gelijkheidsmaat)



Question Answer wiki

Algemene informatie

Je bent ingelogd als Peter van Rosmalen (Log uit)

[Minicursus](#) » [Algemene Info](#) »

Deze wiki is opgezet namens "Peters Adminaccount" om bijgaande vraag te beantwoorden:

- Lees de vraag en de tekstsuggesties aandachtig voordat u een antwoord formuleert.
- Stel zonodig in de wiki een vraag ter verduidelijking.
- Probeer in overleg zo snel mogelijk tot een antwoord te komen.
- Selecteer "Bewerk" om uw antwoord in te voeren.
- Tot slot, "Peters Adminaccount" rondt de vraag af door een beoordeling van het antwoord te geven. Als een vraag afgerond is, wordt dit aangegeven door een vinkje voor de vraag.

Vraag:

Laatste was ik aan het chatten en in de chatbox waar ik in zat hadden mijn mede-chatters allemaal mooie avatars. Ik vroeg me af hoe je aan deze avatars komt en kan je ze ook zelf maken? Is er misschien software waarmee je je eigen avatar kunt ontwerpen

De cursus bevat de volgende teksten die mogelijk relevant zijn bij het beantwoorden van de vraag. Klik op de bijbehorende links om de teksten te bekijken:

[Tekst 1](#)

[Tekst 2](#)

[Tekst 3](#)

Toon

Bewerk

Links

Geschiedenis

Experiment data

- Learning network met 11 topics; 8 weken, 22 studie-uren
- 110 studenten in 2 groepen: 78 actief (40 : 38)
- Data: *loggings; ratings*; questionnaire; 2 docenten
- Vragen: 101 → 59 vs 42
- Opgelost: 60% → 71% vs 45% beter
- Tijd: 7 dagen → 5.6 vs 9.6 sneller
- Eerste uitnodiging: 67% → 80% vs 50% actiever
- 47 vraagstellers; 65 beantwoorders → 68 betrokken



Experiment data (gebruikersfeedback)

Almost all respondents (n=57) agreed that answering a question is a good investment of time, motivations:

- “I am aware that other students also have questions” (n=24)
- “It improved my knowledge and understanding” (n=29)

Usefulness: 26 experimental; 17 control

Usability: 22 experimental; 16 control

Use it again: 25 experimental; 16 control

2. Latente Semantische Analyse

Jan van Bruggen

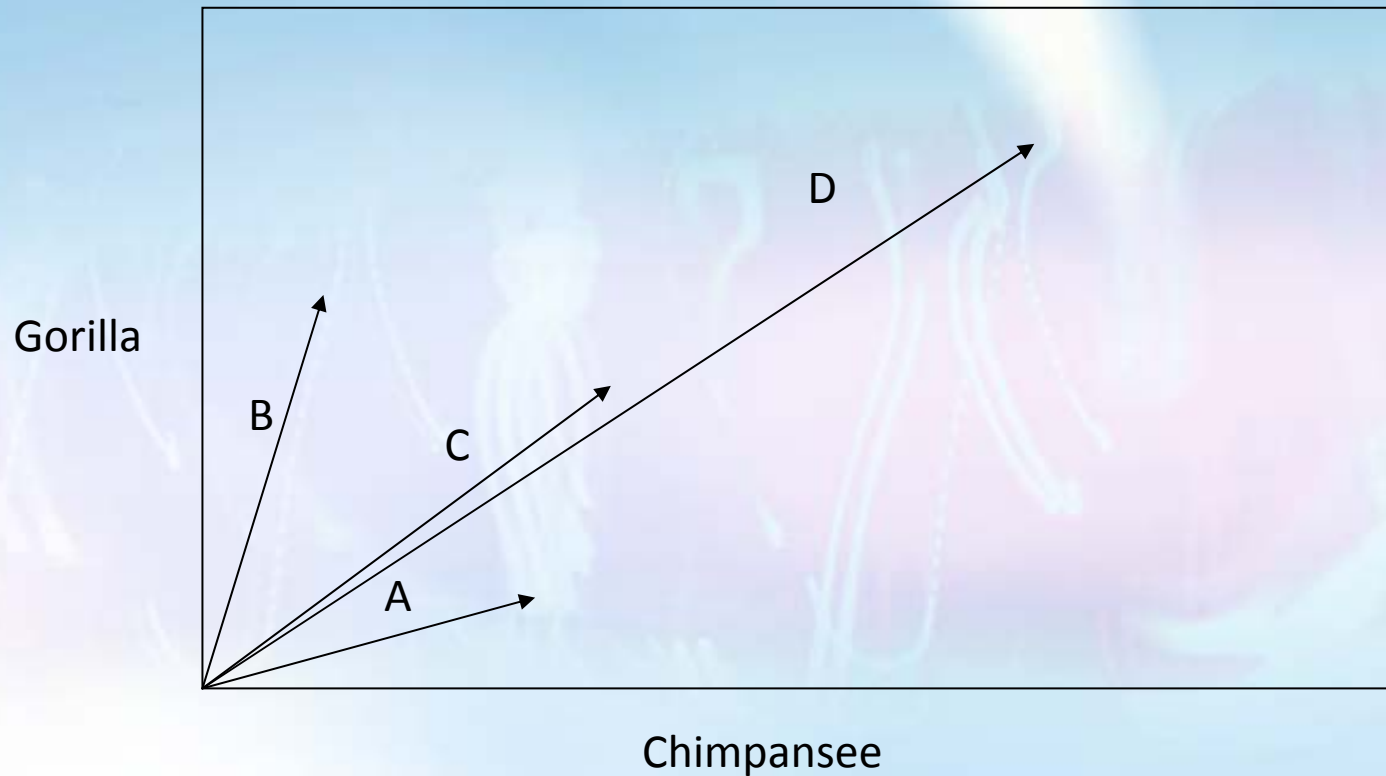
Achtergrond: Tekst-Vector Model

- Tekst representeren als vector
- Term-Document matrix met
 - Term-rijen
 - Document-kolommen
 - Termfrequenties in de cellen
- Toepassingen
 - Document retrieval
 - Documenten vergelijken

Voorbeeld

Documenten	A	B	C	D
Chimpansee	2	1	3	6
Gorilla	1	3	2	4

Documentvectoren



Compacte weergave: PCA en LSA

- Sterke gelijkenis met principale componenten analyse
- PCA:
 - Symmetrische matrix **M** (bijvoorbeeld correlaties)
 - Eigenwaarden en eigenvectoren
 - **$M = U \Lambda U'$**
 - **Λ** matrix eigenwaarden
 - Diagonaalmatrix
 - Afnemende waarden: $\Lambda_1 > \Lambda_2$
 - Orthogonale assen
 - Herschreven als stelsel lineaire vergelijkingen

LSA, SVD en dimensiereductie

- Singuliere Waarden Ontbinding (SVD)
- Asymmetrische matrix (TD-matrix)
- **$D = L S R'$**
- S zijn orthogonale vectoren (dimensies)
- Stelsel lineaire vergelijkingen met $S_1 > S_2$ enz
- *LSA reproduceert TD-matrix*
 - *Dimensie-reductie (cf factorenanalyse)*
 - *Latente factoren (contexten / concepten)*
 - *Theory of mind (Landauer & Dumais)*

Toepassingen

- Document retrieval
- Vergelijk documenten (Pearson KT)
 - Essays beoordelen
 - Plagiaat controle
 - Samenvattingen leren schrijven
 - CareerMap
- Conceptverwerving
 - Conceptuele groei
 - Feedback in ITS (Interactive Fishics)
- Semantiek van tekst(Kintsch)
- <http://www.pearsonkt.com>

3. Project: Language Technologies for Lifelong Learning (LTfLL)

Problem Statement & Approach

The availability of support is crucial for effective task performance. Current e-learning and personal development environments provide too little effective support to the users in their various tasks.

LTfLL fully concentrates on offering the learners a set of next-generation support services that enhance individual and collaborative building of competences and knowledge creation.

The services will run (semi) automatically and thus require no or only limited tutor-based support and will make extensive use of *language technologies*.





Research Lines

1. Modelling and relating the learner and the domain (*WP4*)
 - analysis of learner eportfolios
 - modelling conceptual development
2. Contemporary pedagogical models and support (*WP5*)
 - analysis of interactions (chat, forum)
 - analysis of textual output
3. Collaborative and social dimensions in learning and knowledge structuring (*WP6*)
 - knowledge extraction
 - ontologies + social tagging
 - support knowledge construction / negotiation

Onze focus

- Onderwijs georiënteerde scenario's
 - Conceptuele Ontwikkeling
- +
- Verdieping inzicht in mogelijke toepassingen

Diagnosis Conceptual Development

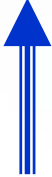
Theoretical Background

- **How to support medical students so they have the level of expertise required for a particular domain in a particular level**
 - Identify the current status of the learner's conceptual development (*person's knowledge of a domain by looking on how she organizes the concepts of such domain*)
 - → identify potentially weak knowledge areas and recommend remedial actions.
- Distinction between expert, novices and intermediate levels
 - Differ in their problem-solving skills, knowledge use, information processing, the time the use to provide diagnosis, and their well-organized knowledge structures

Diagnosis Conceptual Development

Theoretical Background

Arts et al. (2006), Boshuizen & Schmidt (1992; 2008), Jonassen et al. (1993), Nievelein et al., (in press)

Novices	Intermediates	Experts
<ul style="list-style-type: none"> • Use every-day conceptual knowledge, knowledge less hierarchically structured • Novices structure concepts less well than experts. • Present problems in terms of literal characteristics. • Limited, incomplete knowledge, short texts, expressed in everyday language, both in the verbs and nouns. 	<ul style="list-style-type: none"> • Much detail arguments, and long protocols • Have great difficulty with the activation of relevant knowledge when confronted with problems that are not in professional terminology (i.e. raw material) 	<ul style="list-style-type: none"> • Larger numbers of ready-made illness scripts when dealing with a case • Present problems in terms of abstract principles • Give extensive explanations in which the links between different disciplines are given and activate knowledge through many roads (observation data, specific tasks, other hypothesis, ..) • jargon, instead of detail arguments



Diagnosis Conceptual Development

Measuring Conceptual Development

- Structural approach (Goldsmith et al., 1991) to assess the person's knowledge of a domain by looking on how she organizes the concepts of such domain

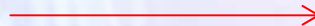
3 steps

1. **Knowledge elicitation:** Measures an individual's understanding of the relationships among a set of concepts.
 - Categorization methods: card sorting, word association
 - Graphical: **concept maps**, semantic networks
 - Verbal reporting: thinking-aloud, **essay**, essay questions
2. **Knowledge representation methods:** cluster analysis, three constructions, dimensional representation, pathfinder nets, verbal test.
3. **Evaluation** of an individual's knowledge representation relative to some standard (e.g., expert's organization of the concepts in the domain)
 - Qualitative assessment of derived representations
 - Quantifying the similarities between a student representation and a derived structure of the content of the domain
 - **Compare the cognitive structures of 'experts' and novices.**



LTfLL onderzoek

Knowledge
Evaluation



Conceptual Map
vs.
Text output (LSA)

Comparison



Keuze tekstmaterialen

Tekstmateriaal studenten:

- Specifiek onderwerpen
- Omvang (minimaal ½ A4)
- Verschillende 'niveaus' / tijdstippen
- Gebaseerd op opdrachtomschrijving
- Voldoende deelnemers (>20)

Overig:

- Referentiemateriaal per onderwerp (tutor)
- Cursustekst voor textcorpus
- Tutor voor validatie uitkomsten

Voorbeelden:

- werkstukken, casus beschrijvingen (symptomen, diagnose, behandeling)

Tot slot

- 3 onderzoeks & ontwikkelronden (halverwege rond 1): inventarisatie mogelijke tekstbronnen
- alles open source
- Talen: Engels, Duits, Frans, Nederland, Roemeens, Bulgaars
- Project site: www.ltfll-project.org

vragen

OpenUniversiteitNederland