

Workshop CorpuScript

Ghent University, 19-12-06

In the context of the educational project "CorpuScript: Digitale corpora als tool voor het trainen van schriftelijke vaardigheden in het vreemde-talenonderwijs" (UGent) Ghent University organises a 1-day workshop about constructing and tagging corpora (learner as well as reference corpora) that can be used both in FLE and in SLA research.

The following topics will be dealt with in the workshop:

Methodology of constructing corpora

- the criteria of good corpora: defining specific goals, size of the corpus, suitable texts, selection of learner texts, ...
- how to create a database?; elaborating a research protocol (necessary if several collaborators work within the same project); which programmes to use?; which (research) questions to ask?; ...

Efficient tagging with a view to automatic treatment

- which annotation tools?; which software?; ...
- defining the grammatical categories to tag; defining the errors to be marked within learner corpora (+ learners profiles)
- the integration of a corpus parser

Application of data in teaching and research

- learners and corpora: pedagogical applications of learner and reference corpora
- creating exercises based on corpora: possibilities within writing skills
- results of present research in SLA and writing skills > transfer to teaching practice?
- problems in teaching practice > new research perspectives?

Atelier CorpuScript

Université de Gand, 19-12-06

Dans le cadre du projet « CorpuScript : corpus électroniques dans l'enseignement de la compétence écrite en langue étrangère » (UGent) se déroulera à l'Université de Gand un atelier consacré à la constitution et à l'exploitation de corpus de référence ainsi que de corpus d'apprenants.

L'objectif de l'atelier est de traiter les questions suivantes :

La constitution de corpus: problèmes méthodologiques

- quels critères retenir dans la définition d'un bon corpus : déterminer les besoins spécifiques, quels textes retenir ? (exemples et productions réelles d'apprenants), ...
- comment créer une base de données performante ? ; quel protocole utiliser ? (indispensable au fonctionnement parallèle des membres de l'équipe) ; quels programmes utiliser ? ; déterminer les questions de recherche ; ...

Le traitement de corpus : le « tagging »

- quels outils d'annotation ? ; quels logiciels ? ; ...
- quelles catégories grammaticales / quelles erreurs sélectionner pour l'encodage ? ; déterminer les profils d'apprenants
- intégrer un parseur

L'exploitation des données : aspects pédagogiques et scientifiques

- apprenants et corpus : applications pédagogiques de corpus de référence et d'apprenants
- création d'exercices à partir des corpus : comment stimuler la compétence écrite ?
- intégration des résultats de recherches actuelles en acquisition / apprentissage > quelles implications pédagogiques ?
- la réalité de classe > nouvelles perspectives / hypothèses de recherche ?

Workshop CorpuScript

Universiteit Gent, 19-12-06

Naar aanleiding van het onderwijsproject "CorpuScript: Digitale corpora als tool voor het trainen van schriftelijke vaardigheden in het vreemde-talenonderwijs" (UGent) wordt aan de Universiteit Gent een 1-daagse workshop georganiseerd rond het aanleggen en bewerken van tekstcorpora (leerders- en referentiecorpora) die zowel bij het vreemde-talenonderwijs als bij het onderzoek naar vreemde taalverwerving kunnen ingezet worden.

Tijdens de workshop zullen volgende aspecten aan bod komen:

Methodologie van het samenstellen van corpora

- criteria waaraan een goed corpus moet voldoen: bepalen van specifieke noden, omvang corpus, geschikte teksten, welke leerdersteksten weerhouden?, ...
- hoe een database creëren; een onderzoeksprotocol opstellen (noodzakelijk indien meerdere medewerkers op hetzelfde project werken); welke programma's gebruiken; welke (onderzoeks)vragen stellen, ...

Het efficiënt bewerken (taggen) met het oog op automatische verwerking

- welke annotatie-tools?; welke software?; ...
- welke (grammaticale) categorieën markeren; hoe fouten markeren binnen leerderscorpora (+ leerdersprofielen)
- de integratie van een corpusparser

Toepassing van gegevens in onderwijs en onderzoek

- leerders en corpora : pedagogische toepassingen van leerders- en referentiecorpora
- oefeningen genereren aan de hand van corpora : mogelijkheden binnen de schrijfvaardigheid
- resultaten van huidig onderzoek rond schrijfvaardigheid en vreemde taalverwerving > bijdrage tot de pedagogische praktijk ?
- problemen in de praktijk > nieuwe onderzoeksperspectieven?

Programme

- 8.45 Reception
9.00 Presentation of the workshop
9.30 Corpora for CorpusCALL
(H. Paulussen/P. Desmet - KULAK)
11 Break
Pedagogical applications of native and
11.15 learner corpora
(F. Meunier - CECL, Louvain-la-Neuve)
12.45 Lunch
Bilingual corpora as a research and
educational tool in contrastive language
14.30 studies. Focus on the Dutch-Swedish
parallel corpus
(G. Rawoens - UGent)
15.15 Break
Using CHILDES in Second Language
15.30 Acquisition research
(A. David - University of Newcastle)
17 Conclusion

Registration

The workshop is organised for language teachers who are interested in corpora for their teaching practice and for researchers focusing on CALL-applications.

The registration fee is 10 euro and can be paid at the beginning of the workshop. The fee includes a snack and coffee during lunch and coffee breaks.

You can register by sending an e-mail message to a.demeulenaere@UGent.be with 'Corpuscript' as subject in which you mention your name, address and academic affiliation before December 5.

We will confirm your registration.

Contact

Organisation

Prof. Pascale Hadermann
Prof. Godelieve Laureys
Dr. Torsten Leuschner
Prof. Ilse Logie
Prof. Annemieke Van Herreweghe

Freek Adriaens
Geert Bonamie
Alex Demeulenaere

Address

CorpuScript – Vakgroep Frans
Blandijnberg 2
9000 Gent

Internet

<http://www.schrijven.ugent.be/Corpuscripteng>

Telephone

+32/(0)92644045

Mail

a.demeulenaere@UGent.be