

**Transcription systems and UNICODE  
Special characters in a standardised world**

**CALL FOR PAPERS**

Venue: Vienna, Wohllebengasse 12-14, SR 1  
Date: 5 November 2008, expected to start around 10 am  
Organiser: Cultural Diversity Focus Group (CDFG)  
Institut für Österreichische Dialekt- und Namenlexika (I DINAMLEX)  
Contact: Eveline Wandl-Vogt ([eveline.wandl-vogt@oeaw.ac.at](mailto:eveline.wandl-vogt@oeaw.ac.at))

**Invitation**

We kindly invite you  
either to get actively involved in the programme of our workshop,  
or simply to join us during the workshop for debate and discussion.

Method of presentation: oral/poster presentation  
Conference language: English  
Submission date: 10 October 2008

Topics:

- transcription systems and UNICODE
- *Teuthonista* – a transcription system and its use, significance, history, development etc.
- problems caused by special characters in the digital domain
- presentation of problem-solving approaches

## Introduction

Special characters often cause difficulties when it comes to the digitisation or electronic edition of scholarly works, since those characters necessary for a scientifically correct rendering are not included in ordinary fonts.

For several years now I DINAMLEX has been working on a solution to overcome this problem in its own projects, especially in the dictionary project *Wörterbuch der bairischen Mundarten in Österreich (WBÖ)* and in the database project *Datenbank der bairischen Mundarten in Österreich (DBÖ)*. In 2006, a working meeting by the name of *5. Arbeitstreffen deutschsprachiger Akademiewörterbücher* also featured a workshop held on the topic of *Interfaces and Concordances*. This workshop (*Special Characters: Limits of Interconnectivity*) resulted in a web-based initiative appropriately named *WienInitiative* (after its place of origin); its aims are as follows:

1. Bringing together institutions willing to face the questions and problems of the digital rendering of special characters.
2. Assembling and providing a collection of characters neglected by UNICODE.
3. Assembling a collection of UNICODE characters which are relevant to numerous (linguistic) projects and hence may or should be of importance to font manufacturers.

In the past few years I DINAMLEX has completed these lists in respect of the Dictionary project, i.e. especially the list of characters that were still missing. Within the *dbo@ema* project this list was taken into account in the development of the *dinamlex* font, which resulted in 2 formats and 4 font styles.

With two years having passed, it is now time to take stock and, as a consequence, introduce further concrete steps. Above all, we would like to promote once more – on a broader basis and with reference to the results that have been, and can be, achieved – active participation in the initiative.

### Aims of the workshop:

- Bringing together different approaches in using the *Teuthonista* transcription system within the digital domain
- Setting up a working group focusing on the *Teuthonista* transcription system and on ways of including its characters in UNICODE
- Pooling European efforts with a view to transcribing linguistic texts using transcription systems other than *Teuthonista*

- Informing about, and raising awareness of, the problems of special characters in the digital domain.
- Informing about the compilation of a concrete, project-specific (wish) list of special characters to be handed over to the UNICODE consortium
- Developing problem-solving strategies across projects
- Establishing a working group for special characters: *WienInitiative*
- Realising a jointly defined project, e.g. the further development of a font with as many UNICODE-based characters as possible, such as *dinamlex*
- Discussing problem-solving approaches with regard to possible funding

### Workshop participants are requested:

- to draw up a (digital) list of special characters, i.e. characters which may cause problems when rendered digitally, or which are relevant to the publication of the work(s) to be edited.
- This must be preceded by checking whether or not the characters are defined in UNICODE.
- At the very best, the respective characters are entered in the lists (New Characters, Defined Characters for font manufacturers) within the framework of *WienInitiative* (which also provides the formats – name, description, comments, link to scanned page, contact person – necessary for the successful integration of characters into UNICODE).
- within the framework of one's project, to determine the resources (time, personnel, finances) required for this initiative
- to explore possibilities of financial support (third-party funds, special funding)

### Links:

- CDFG: <http://www.cdfg.org/index.html>
- I DINAMLEX: <http://www.oeaw.ac.at/dinamlex>
- workshop special characters: <http://www.oeaw.ac.at/dinamlex/Tagung2006.html>
- **WienInitiative:** <http://wiki.cdfg.org/WienInitiative>
- **Font *dinamlex*:** <http://www.wboe.at/de/font.aspx>
- **dbo@ema:** <http://www.wboe.at>