

# Learning cuneiform the modern way

Timo Homburg<sup>1</sup>, Christian Chiarcos<sup>1</sup>, Thomas Richter<sup>2</sup>, Dirk Wicke<sup>2</sup>

<sup>1</sup> Institute for Computer Science      <sup>2</sup> Institute for Archaeology  
Goethe University, Frankfurt, Germany  
`timo.homburg@stud.uni-frankfurt.de`,  
`{chiarcos|thomas.richter|wicke}@em.uni-frankfurt.de`

**Keywords:** Assyriology, cuneiform, input method engines (IME), flash card learning

With our poster and the accompanying demo, we present current progress on the information-technological support for scholars and students of cuneiform. For a period of about 3000 years, cuneiform was the dominant writing system of the Ancient Near East, with a rich literary tradition in several languages, and an extensive amount of texts preserved in tens of thousands of clay tablets.

Despite this wealth of data and a strong academic tradition in their analysis, the numerous specific challenges of cuneiform have only partially been addressed so far. Here, we propose adapting input method engines (IMEs) and learning strategies commonly used for Asian languages according to the needs of Assyriology.

*Typing cuneiform* Cuneiform writing for Akkadian, Sumerian and Hittite is ideosyllabic, i.e., combining syllabic and ideographic elements, often for the same sign. Up until this date there is no free and convenient way of typing Unicode cuneiform characters other than utilizing the Unicode code tables directly, i.e., to copy and paste from online dictionaries. Not all online dictionaries, however, use Unicode symbols, some use legacy fonts, some represent signs by images, and some do not provide a cuneiform representation at all.

To accommodate this deficit, we developed an input method which is based on the transliteration concept of Chinese Pinyin, the most common way of typing non-alphabetical languages on a computer. To achieve an equivalent input for the aforementioned languages we utilized a given char transliteration to cuneiform table<sup>1</sup> to create transliteration to cuneiform mappings of Akkadian, Sumerian and Hittite CDLI<sup>2</sup> corpora respectively. Organized as a tree, thus minimizing latency, word and char-based input method engines were created for Java (JIMF,

---

<sup>1</sup> <http://www.acoli.informatik.uni-frankfurt.de/resources/cuneiform/signs-final.xml>

<sup>2</sup> <http://cdli.ucla.edu/>

Fig. 1.a)<sup>3</sup>, JQuery(Fig. 2)<sup>4</sup>, SCIM<sup>5</sup> and Ibus (Fig. 3)<sup>6</sup>, thereby covering the most important input method engines on Linux, Web and Java environments.

*Learning cuneiform* Because of limited technological support, digital resources in assyriology often focus on transliteration or transcription as means of representation whereas students are required to acquire the necessary knowledge on cuneiform characters on their own. Clearly, none of those practices are satisfying or easily adaptable for text processing and therefore not useful for computer-aided teaching methods. For conveniently typing and learning cuneiform characters, words and phrases for the Akkadian, Sumerian and Hittite language, we present an adaptation of Anki, a common tool for flash card learning (Fig. 1.b). We utilized the existing character/word table to create flash card sets consisting of more than 50000 words for the Anki and AnkiDroid<sup>7</sup> flash card learning program. Subsequent extensions may exploit existing corpora, e.g., the Open Richly Annotated Cuneiform Corpus (ORACC),<sup>8</sup> to create flash cards for words and phrases. Anki schedules learning content according to a spaced repetition learning method having proven its positive learning effect over a longer period of time to maximize learning success. Because of its usability in both mobile and desktop environments and its ability to share flashcards online Anki suits not only the students but also simplifies sharing lecture specified flash card sets for the lecturers.

*In our presentation*, we demonstrate how an input method engine can act as a suitable tool for solving the mentioned input and compatibility problems while at the same time being useful for education and language learning purposes. With the tools described above, teachers can now easily create their own flash cards according to the pace and content of their lectures. Students may enjoy a convenient and scientifically proven way of learning cuneiform vocabulary, as well as a way to prove their learning by utilizing the input method engine to create their own cuneiform texts. In conclusion, a notable improvement in writing and in learning the concerned languages has been realized and is in general perceived well.

Both tools and the accompanying sign/word table have been created in the context of on-going experiments on word segmentation and transliteration in cuneiform languages. In this regard, we are thus primarily working on *processing cuneiform*. In addition to demonstrating input methods and learning tools, we will include early results with respect to these aspects in demo and presentation, as well.

<sup>3</sup> <http://docs.oracle.com/javase/7/docs/technotes/guides/imf/overview.html>

<sup>4</sup> Sourcecode: <https://github.com/situx/webime>

Livedemo: <http://www.web-ime.de.vu>

<sup>5</sup> <http://sourceforge.net/projects/scim/>

<sup>6</sup> <https://code.google.com/p/ibus/>

<sup>7</sup> <http://ankisrs.net>

<sup>8</sup> <http://http://oracc.museum.upenn.edu>

