

**ABaC:us – Austrian Baroque Corpus**

**ABaCus** is a digital collection of printed texts in a variety of Upper German dating from the baroque era, in particular the years from 1650 until 1750.

Books of religious instruction concerning ‘last things’ had a central position in baroque culture and were a best-selling genre during the 17<sup>th</sup> century. The project group focuses on creating a thematic research collection of several texts specific to this genre (not samples, but full versions of every single book). The corpus, which currently yields 120.000 running words, contains texts associated with death and dying and includes sermons, illustrated devotional books, works related to the dance of death theme as well as compilations of prayers and songs for daily use.



**ABaC:us** is a rich source for the study of popular imagery of the Baroque era and allows for further insight into the way of life and the attitudes towards death during the period under investigation.

In the creation of this corpus, the working group has emphasized an interdisciplinary approach and strived to bring together traditional philological expertise with up-to-date text technology.

# Textual analysis

The scarcity of digital material available from the period makes reusability an important issue. For this reason all XML annotations were carried out according to the guidelines of the Text Encoding Initiative (version P5). The mark-up applied has been designed to capture several layers of information:

- Structural annotation:  
chapters, headings, paragraphs, etc.
  - Linguistic annotation:  
PoS and lemmas – a new tool is currently being tested, which allows the researchers to manually verify and correct wrongly assigned PoS tags and lemmas. We are planning to use the list of lemmas as the basis for generating a computational lexicon, which will be supplemented with automatically extracted corpus data.
  - Named entities: place names, personal names
  - Biblical quotations
  - Salient semantic features

One particular feature of the specific genre in this corpus is that ‘Death’ plays a major role in the countless variations of the topic. Authors were extremely inventive in modifying the image of the reaper figure, which was the most apparent personification of death, and has been preserved in art, literature and in people’s minds to the present day. Thus, one integral part of our annotation approach was to make such conceptual diversity more easily discernible and to support access to comparable digital resources. The goal was to apply appropriate domain specific information and to develop a shallow taxonomy. A few occurrences of the death-related vocabulary are listed on the right >>>

# Cooperation

CLARIN.AT / DARIAH.AT

Deutsches Forschungszentrum  
für Künstliche Intelligenz (DFKI)

ENBaCH – European Network  
for Baroque Cultural Heritage

Germanistisches Seminar Heidelberg,  
„Grimmelshausen digital“

Wolfenbütteler Digitale Bibliothek

*Aschen=Mann*  
*Auffsetzer*  
*Blumen=Feind*  
*Bock*  
*Bücher=Feind*  
*Capell=Meister*  
*aprioln=Schneider*  
*entre Admiral Tod*  
*Bereuter*  
*Dieb*  
*eb der Fröhlichkeit*  
*Herr Doctor Tod*  
*Donnerkeil*  
*Dürerer*  
*Fieber=Tinctur*  
*Fischer*  
*Gall=Aepffel=Brey*  
*Gärtner*  
*Gast*  
*Gefreitter*  
*ral Haut und Bein*  
*Gesell*  
*Grammaticus*  
*Haffner*  
*schel= und Maus=*  
*Fallen Jubilirer*  
*er von Beinhäusen*  
*und Sichelberg*  
*Holtzhacker*  
*ker=Meister aller*  
*efangenen Seelen*  
*Knoll*  
*Kräuter=Salat*  
*Lebens=Feind*  
*Mader*  
*auf langen Füßen*  
*Marode-Reuter*  
*Menschen=Fischer*  
*Jenschen=Mörder*  
*nschen=Schnitter*  
*Jenschen=Würger*  
*Mißgebirth*  
*Oppositum der*  
*Schnecken=Post*  
*Pedell*  
*Perforce-Jäger*  
*Postillion*  
*Reuter auf dem*  
*fahlen Pferd*  
*Rippen=Kramer*  
*Schnitter*  
*Schütz*  
*Schützen=Meister*  
*Sensentrager*  
*etc.*

# ABaC:us Activities and Progress

## Conferences 2013

### 26 January, Vienna

Workshop: Korpuslinguistik in Forschung und Lehre  
(in cooperation with Ludwig-Maximilians-Universität München, Universität Wien, Wr. Sprachgesellschaft)  
› Organized by ABaC:us / ICLTT

### 18-19 February, Berlin

DTA-/CLARIN-D-Conference und -Workshops: Historische Textkorpora für die Geistes- und Sozialwissenschaften. Fragestellungen und Nutzungs-perspektiven: „ABaC:us – Austrian Baroque Corpus: Aufbau, Annotationen und Anwendungen“

### 3-5 June, Saarbrücken

NooJ 2013 Conference: „Porting Historical Linguistic Resources to NooJ – The ABaC:us Corpus“

### 12 September, Hissar

Natural Language Processing and Linked open Data @ RANLP: „Porting Elements of the Austrian Baroque Corpus onto the Linguistic Linked Open Data Format“

### 22-24 November, Salzburg

Workshop „Korpusbasierte Linguistik“ @ 40. ÖLT:  
„ABaC:us für LinguistInnen: Morphosyntaktische Annotation im Austrian Baroque Corpus“

› Organized by ABaC:us / ICLTT

### 5-6 December, Luxembourg

Reading historical sources in the digital age: „Introducing the Austrian Baroque Corpus: Annotation and application of a thematic research collection“

### 9 December, Hildesheim

3. Arbeitstreffen zum Stuttgart-Tübingen-TagSet, STTS: Linguistische Annotation von älterem Neu-hochdeutsch - Zur Anwendung des STTS-Tagsets im Projekt ABaC:us

## Publications 2013

Czeitchner, Declerck, Resch (2013): Porting Elements of the Austrian Baroque Corpus onto the Linguistic Linked Open Data Format. In: Proceedings of the joint NLP&LOD and SWAIE Workshops, located at RANLP 2013, 12-16.

Czeitschner, Resch (2013): Repräsentationen von | in barocken Buchtentänzen im digitalen Medium. In: Kulturforschungen Nr. 2. Denkschriften der phil.-hist. Klasse. Verlag der ÖAW (forthc.).

Resch, Dressler (2013): Zur Pragmatik der Diminutive in ausgewählten Erbauungstexten der Barockzeit – Eine korpusbasierte Studie. In: Lingua Historica Germanica. Berlin (forthc.).

## ABaC:us Team

Ulrike Czeitschner

Thierry Declerck

Barbara Krautgartner

Charly Mört

Claudia Resch

Eva Wohlfarter

## For more information

Claudia Resch

Austrian Academy of Sciences

Institute for Corpus Linguistics and Text Technology (ICLTT)

1010 Vienna, Sonnenfelsgasse 19 / 8

e-mail: claudia.resch@oeaw.ac.at

[www.oeaw.acat/icltt/abacus](http://www.oeaw.acat/icltt/abacus)

[www.oeaw.acat/icltt/abacus-project](http://www.oeaw.acat/icltt/abacus-project)

## Austrian Baroque Corpus



Text-technological methods for the analysis of Austrian Baroque texts

A Research project hosted by the Austrian Academy of Sciences, Institute for Corpus Linguistics and Text Technology (ICLTT)



ICLTT

Institute for  
Corpus Linguistics and  
Text Technology

Supported by funds of the Österreichische Nationalbank – Anniversary Fund, project number: 14738, duration: 2012 – 2014