Words in Contexts. Digital Editions of Literary Journals in the "AAC - Austrian Academy Corpus"

Hanno Biber, Evelyn Breiteneder, Karlheinz Mörth

AAC - Austrian Academy Corpus, Austrian Academy of Sciences
Sonnenfelsgasse 19/8, Vienna, Austria
hanno.biber@oeaw.ac.at, evelyn.breiteneder oeaw.ac.at, karheinz.moerth oeaw.ac.at

Abstract

In this paper two highly innovative digital editions will be presented. For the creation and the implementation of these editions the latest developments within corpus research have been taken into account. The digital editions of the historical literary journals "Die Fackel" (published by Karl Kraus in Vienna from 1899 to 1936) and "Der Brenner" (published by Ludwig Ficker in Innsbruck from 1910 to 1954) have been developed within the corpus research framework of the "AAC - Austrian Academy Corpus" at the Austrian Academy of Sciences in collaboration with other researchers and programmers in the AAC from Vienna together with the graphic designer Anne Burdick from Los Angeles. For the creation of these scholarly digital editions the AAC edition philosophy and edition principles have been applied whereby new corpus research methods have been made use of for questions of computational philology and textual studies in a digital environment. The examples of the digital online editions of the literary journals "Die Fackel" and "Der Brenner" will give insights into the potentials and the benefits of making corpus research methods and techniques available for scholarly research into language and literature.

1   "AAC - Austrian Academy Corpus"

The "AAC - Austrian Academy Corpus" is a corpus research unit at the Austrian Academy of Sciences in Vienna concerned with establishing and exploring large electronic text corpora and with conducting scholarly research in the field of corpora and digital text collections and editions. The variety of texts integrated into the collections of the AAC are predominantly German language texts of historical and cultural significance from the last 150 years.

Figure 1: AAC logo

The AAC has collected thousands of texts from various authors, representing many different text types from all over the German speaking world. Among the sources, which systematically cover various domains, genres and types, are newspapers, literary journals, novels, dramas, poems, advertisements, essays on various subjects, travel literature, cookbooks, pamphlets, political speeches as well as a variety of scientific, legal, and religious texts, to name just a few forms.

The AAC provides resources for investigations into the linguistic and textual properties of these texts and into their historical and cultural qualities. More than 350 million running words of text have already been scanned, digitized, integrated and annotated. The selection of texts is made according to the AAC's principles of text selection that are determined by specific research interests as well as by systematic historical, empirical, technological and text typological parameters. The annotation and mark-up schemes of the AAC, based upon XML related standards have in the first phase of corpus build-up been concerned with the application of basic structural mark-up and some selective thematic annotations. In the next phase of application development more specific thematic annotations are being made exploring questions of linguistic and textual scholarly research as well as experimental and exploratory mark-up.

As journals are regarded as highly interesting sources for corpus research because they usually comprise a great variety of texts and text types over a long period of time, two model digital editions of literary journals have been developed within the framework so far: The digital edition of "Der Brenner", published in October 2007, and the digital edition of "Die Fackel", online since January 2007.

2   "Der Brenner"

The literary journal "Der Brenner" was published between 1910 and 1954 in Innsbruck by Ludwig Ficker. The satirical journal "Die Fackel" published by Karl Kraus in Vienna from 1899 until 1936 had set an example for Ludwig Ficker and his own publication. The text volume of "Der Brenner" is 18 volumes and 104 issues, which is just a small segment of the AAC's overall holdings, is about two million running words of carefully corrected text, annotated and provided with additional philological information. Contrary to the more widely read, very influential critical, polemical and satircal journal "Die Fackel" the more quiet journal "Der Brenner" deals primarily with themes of religion, nature, literature and culture.
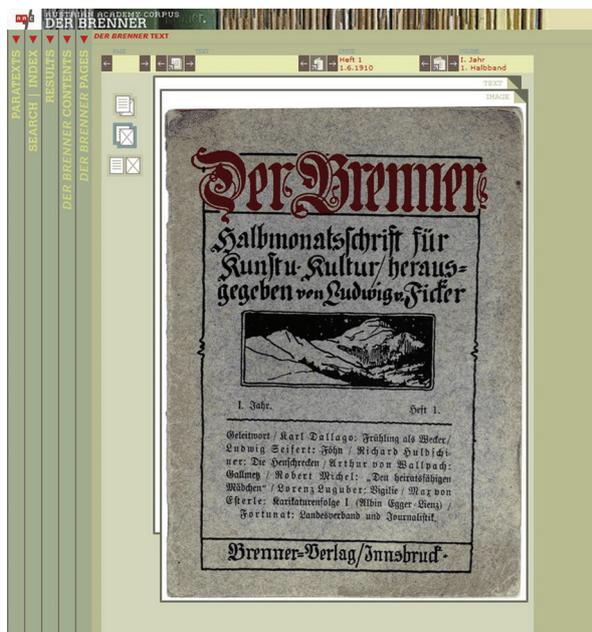

Figure 2: BRENNER ONLINE Interface

Among its famous contributors is the expressionist poet Georg Trakl, the writer Carl Dallago, the translator and cultural critic Theodor Haecker, translator of Søren Kierkegaard and Cardinal Newman into German, the moralist philosopher Ferdinand Ebner and many others. The philosopher Ludwig Wittgenstein was affiliated with the group around the journal's editor-in-chief Ludwig Ficker and participated in some of the debates launched by the journal. The journal covers a great variety of

subjects and is to be regarded as an important voice of Austrian culture in the pre and post second world war periods. It is not available in print anymore, another important fact which, apart from our initiative to create an experimental research environment in the field, has brought us to prepare an electronic version of it.

The digital edition of "Der Brenner" has been made by the AAC in collaboration with the Brenner-Archive of the University of Innsbruck. Both institutions, the Austrian Academy of Sciences and the University of Innsbruck have committed themselves to establish a valuable digital resource for the study of this literary journal.

3   "Die Fackel"

The AAC digital edition of the journal "Die Fackel" (AAC-FACKEL) offers free online access to the 37 volumes, 415 issues, 922 numbers, comprising more than 22.500 pages and six million word forms and also contains a fully searchable database of the entire journal with various indexes, search tools and navigation aids in an innovative and highly functional graphic design interface, in which all pages of the original are available as digital texts and as facsimile images.

"Die Fackel" ("The Torch") was published from 1899 until 1936 and almost entirely, from 1912 on entirely written by the famous language, culture and social critic Karl Kraus. The work of Karl Kraus in its many forms and variations, of which the journal is the core, can be regarded as one of the most important, although almost untranslatable, contributions to world literature. It is a source for the history of the time, for its language and its moral transgressions. In his journal Karl Kraus covers in a typical and idiosyncratic style in thousands of texts the themes of journalism and war, of politics and corruption, of literature and lying. The journal "Die Fackel" comprises in its enormous thematic variety a great variety of different forms, such as essays, notes, commentaries, aphorisms, poems, drama and many other forms.
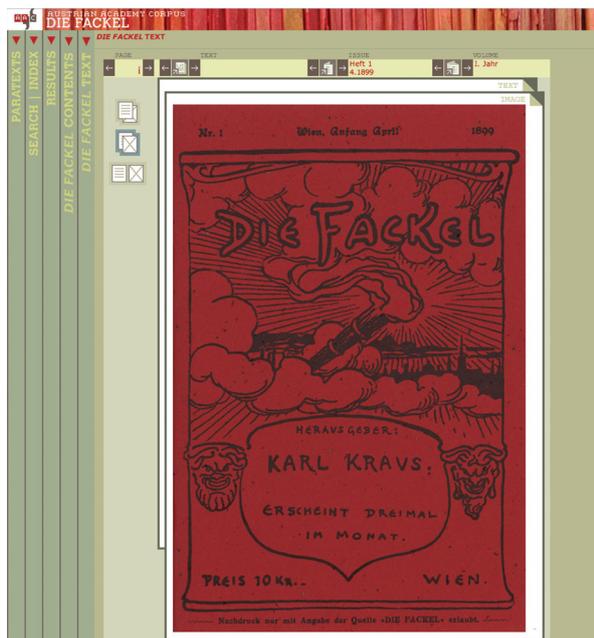
Figure 3: AAC-FACKEL Interface

"Die Fackel" has been made available in electronic form, which allows new ways of philological research and analysis and offers new perspectives for literary studies. The AAC is publishing this complex digital edition of "Die Fackel" in the World Wide Web. The basic elements and features of such an approach of corpus research in the field of textual studies are shaped by our edition principles.

4   AAC Digital Editions Principles

The philological and technological principles of scholarly digital editions within the AAC are determined by our conviction that the methods of corpus research will enable us to produce valuable resources and tools for scholars. The AAC has developed such model digital editions to meet these aims. We have constructed a well designed web interface for these purposes. The AAC digital editions of journals of crucial significance for literary scholars and researchers, provide well structured accesses to their contents. The main task of developing these applications was to devise a design and a web interface so that access to the texts and the related information could be fully utilized. All pages are accessible as electronic text and as facsimile images of the original pages. Various indices and search facilities are provided

so that word forms, personal names, foreign text passages, illustrations etc. can be searched for in various ways. The search mechanisms for personal names have already been implemented in the interface for "Der Brenner" and will be done in the coming months for "Die Fackel".

The interface is designed to be easily accessible also to less experienced users of new media. Multi-word queries are possible. The search engine supports left and right truncation. The interface provides also search mechanisms for linguistic searches and offers also some experimental features. It allows, for instance, to perform lemma queries. Instead of searching for particular word forms queries for all the word forms of a particular lemma are possible. The same goes for the POS annotation. The web-sites of both editions are entirely based on XML and cognate technologies. On the character level use of Unicode has been made throughout. All of the content displayed in the digital edition is dynamically created from XML data. Output is produced through means of XSLT style sheets. This holds for the text section as well as the contents overview and the result lists. We have adopted this approach in order to be able to cope with future challenges and to ensure the viability of our data for as long a period as possible. These digital editions of AAC-FACKEL and BRENNER ONLINE have been optimized for use with recent browser versions. One of the basic requirements is that the browser i able to handle XML-Dom and the local system should be furnished with a Unicode font capable of displaying the necessary characters.

A sophisticated contents display has been developed, in order to show the reader the whole range of the journals ready to be explored. The edition interface has synchronized five individual frames within one single window. The frames can be alternatively opened and closed as required. The "Paratext"-section situated within the first frame provides additional information about the background of the edition and scholarly essays about the journal, its editor and contributors. The "Index"-section gives access to a variety of indices, databases and full-text search mechanisms. The results of these queries and lists are displayed in the adjacent "Results"-section. The "Contents"-section provides access to the whole run of issues

and all of the contents of the journal in chronological order. The "Text"-section has a complex navigational bar at the top so that the reader can navigate and read within the collections either in text-mode or in image-mode from page to page, from text to text, from issue to issue and with the help of hyperlinks. These digital editions will function as models for similar applications. The AAC's scholarly editions of "Der Brenner" and "Die Fackel" will contribute to the development of digital resources for research into language and literature.

5   References

AAC - Austrian Academy Corpus:
   http://www.aac.ac.at/fackel

AAC - Austrian Academy Corpus: AAC-FACKEL,
   Online Version: Die Fackel. Herausgeber: Karl Kraus,
   Wien 1899-1936, AAC Digital Edition No 1, 2007,
   (http://www.aac.ac.at/fackel)

AAC - Austrian Academy Corpus and Brenner-Archiv:
   BRENNER ONLINE, Online Version: Der Brenner.
   Herausgeber: Ludwig Ficker, Innsbruck 1910-1954,
   AAC Digital Edition No 2, 2007,
   (http://www.aac.ac.at/brenner)