



## How to get started with EXMARaLDA tools on the basis of the Demo Corpus

---

This document explains how to install the EXMARaLDA tools and how to use the EXMARaLDA demo corpus to get a first impression of the EXMARaLDA system. More documents and literature can be found in the section “Help & Documentation” on the EXMARaLDA homepage ([http://www.exmaralda.org/en\\_hilfe.html](http://www.exmaralda.org/en_hilfe.html)).

### Contents

A. Installation of the tools .....	2
1. Download.....	2
2. Java Runtime installation.....	2
3. Installation .....	2
B. Demo corpus .....	3
C. Partitur Editor .....	3
D. Corpus Manager .....	5
E. EXAKT .....	6

## A. Installation of the tools

### 1. Download

You can download the EXMARaLDA tools from [www.exmaralda.org/en\\_downloads.html](http://www.exmaralda.org/en_downloads.html). The versions you find there are the most recent releases. Previews of upcoming versions can be found at [www.exmaralda.org/previews2.html](http://www.exmaralda.org/previews2.html).

Which and how many files you download depends on your operating system:

- Windows: one installation file which contains all EXMARaLDA tools
- Macintosh: three icons each of which contains one EXMARaLDA tool
- Linux: one gzipped tarball containing all EXMARaLDA tools

### 2. Java Runtime installation

In order for the tools to work, your computer needs a Java Runtime Environment (JRE). On Macintosh, a JRE is preinstalled. On Windows or Linux, you can find out whether a JRE is installed by typing "java -version" in a terminal window.



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\>java -version
java version "1.6.0_11"
Java(TM) SE Runtime Environment (build 1.6.0_11-b03)
Java HotSpot(TM) Client VM (build 11.0-b16, mixed mode, sharing)

C:\>_
```

If no JRE is installed, you can freely download the newest JRE from **Sun/Oracle** at <http://java.com/en/download/manual.jsp>. Please note that some Linux distributions come with a preinstalled JRE which is *not* **Sun/Oracle**'s JRE. EXMARaLDA does not work well with these preinstalled JREs. Replace them with the most recent **Sun/Oracle** JRE.



### 3. Installation

Once you've downloaded the files appropriate for your system, installation should be straightforward:

- Windows: double click "setup.exe" and follow the instructions
- Macintosh: mount the disk images and drag the included application file to some appropriate place on your system
- Linux: unzip the tarball to some appropriate place on your system

## B. Demo corpus

A good way to get a first impression of the EXMARaLDA system is to have a look at the EXMARaLDA demo corpus at [http://www.exmaralda.org/corpora/en\\_demokorpus.html](http://www.exmaralda.org/corpora/en_demokorpus.html).

**EXMARaLDA Demo Corpus**

[Browse demo corpus](#) [Download demo corpus \(ZIP-file, 39MB\)](#)

The EXMARaLDA demo corpus is a small corpus with which you can test the functionality of the EXMARaLDA system.

**1** **2**

**Transcription Rudi**

EXMARaLDA: [Transcription] [Se] [1.5s] then I just left school and I st

Visualisation: [Partiture] [RTF] [P] [1.5s] then I just left school and I st

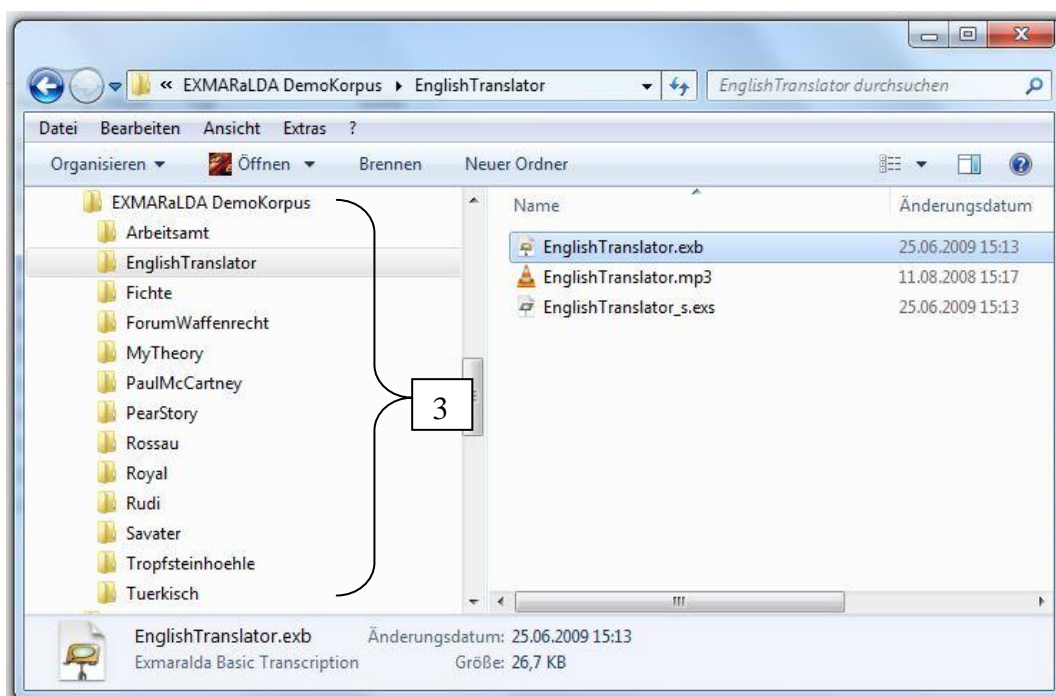
Export: [TEI] [AG] [EAF] [Praat] [C] [1.5s] then I just left school and I st

The demo corpus contains metadata for communications and speakers. All data are available in several transcription and presentation formats. Transcriptions are linked to the audio. The demo corpus is multilingual (English, German, Turkish, French, Spanish).

The demo corpus contains the following recordings and transcriptions:

- Beckhams - An English talk show excerpt with David Beckham and his wife
- Monty Python: My Theory - An English transcription of Miss Elk presenting her theory on the brontosaurus (Any resemblance with living linguists is purely coincidental)
- Paul McCartney: Interview - an English transcription of an interview with young, ambitionless Paul McCartney
- Pear Story - An English transcription of a retelling of Wallace Chafe's Pear Story
- Rudi Völler: Wutausbruch - a legendary fit of rage by the manager of the German national football team
- Helge Schneider: Tropfsteinhöhle - a short audio play by the German artist Helge Schneider which contains many speaker overlaps
- Hubert Fichte: Interview - an interview which was conducted by the German writer Hubert Fichte in Hamburg's StPauli district
- Helge Schneider: Arbeitsamt - another short audio play by Helge Schneider
- Studio Braun: English Translator - a telephone conversation between an unsuspecting seller of erasers and members of Studio Braun
- Deutsche Bahn: Rossau - a telephone conversation between a man from Saxony and the Deutsche Bahn railway timetable computer
- Studio Braun: Forum Waffenrecht - another telephone conversation from Studio Braun
- Hart aber fair - an excerpt from the TV talk show 'Hart aber fair'
- Anne Will: Halbes Wahlrecht - an excerpt from the TV talkshow 'Anne Will'
- Serif Issi: Expertenaukunft - a Turkish transcription of a telephone conversation between an expert on social matters and a somewhat confused caller
- Royal: Débat - a French transcription of an excerpt of the TV debate between Nicolas Sarkozy and Ségolène Royal
- Savater: Hermandad - a Spanish transcription of an excerpt of a radio interview with the writer Fernando Savater
- Gasparini: Calcio - an Italian transcription of an excerpt of a TV interview with footballer Christian Panucci

You can start by browsing the online version of the corpus (1) – there are transcription examples in several languages (see also "How to use an EXMARaLDA corpus"). You can then also download a ZIP archive of the demo corpus (2). In the following three sections, it will be assumed that you downloaded the demo corpus and unpacked it onto your system. The unpacked folder structure (3) should look something like this:

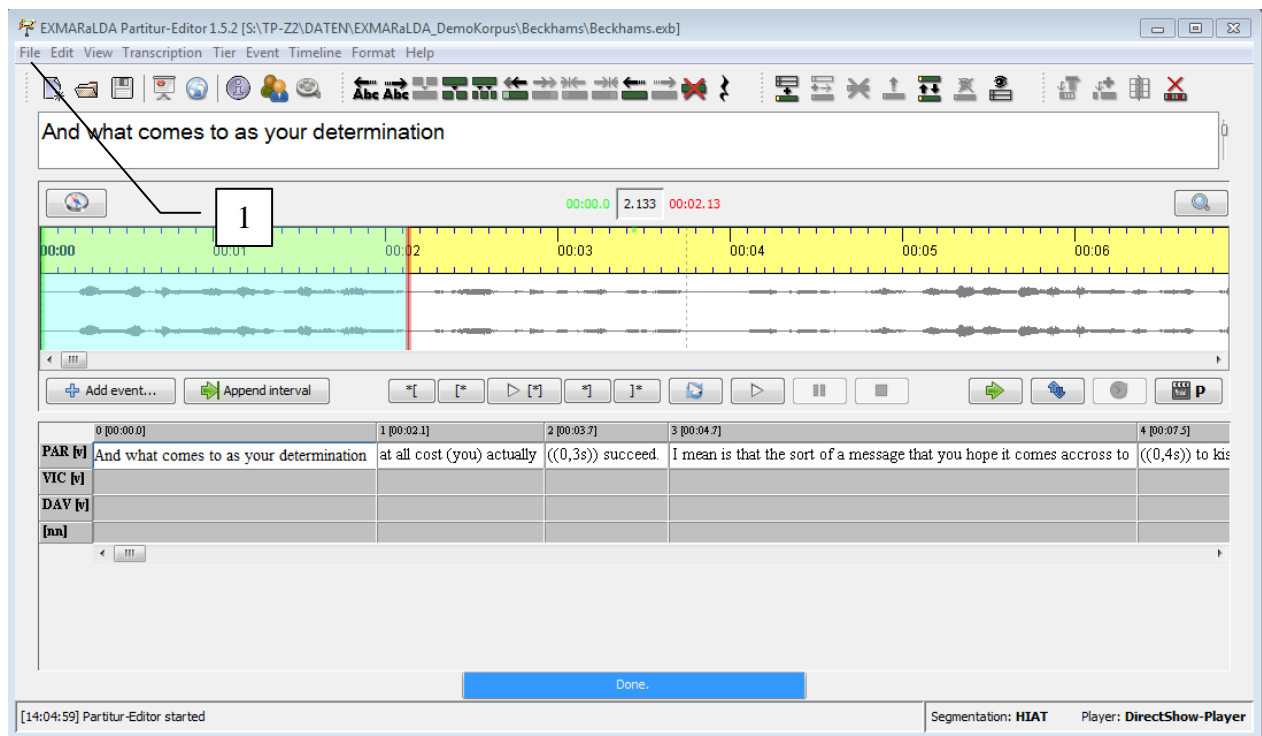


## C. Partitur Editor

To get a first impression of the Partitur-Editor, start it

- on Windows by selecting the appropriate entry in the start menu or by double clicking the symbol on your desktop.
- on Macintosh by double clicking the application symbol
- on Linux by double clicking "partitureditor.sh" or by calling this shell script from a terminal window (you may have to make the file executable before the first start)

and open a basic transcription file from the demo corpus via **File > Open** (1). Basic transcriptions can be found in all directories on the first level (i.e. 'Arbeitsamt', 'EnglishTranslator' etc.). They have the ending ".exb". For example, you could open "Beckhams/Beckhams.exb" in the Partitur-Editor:

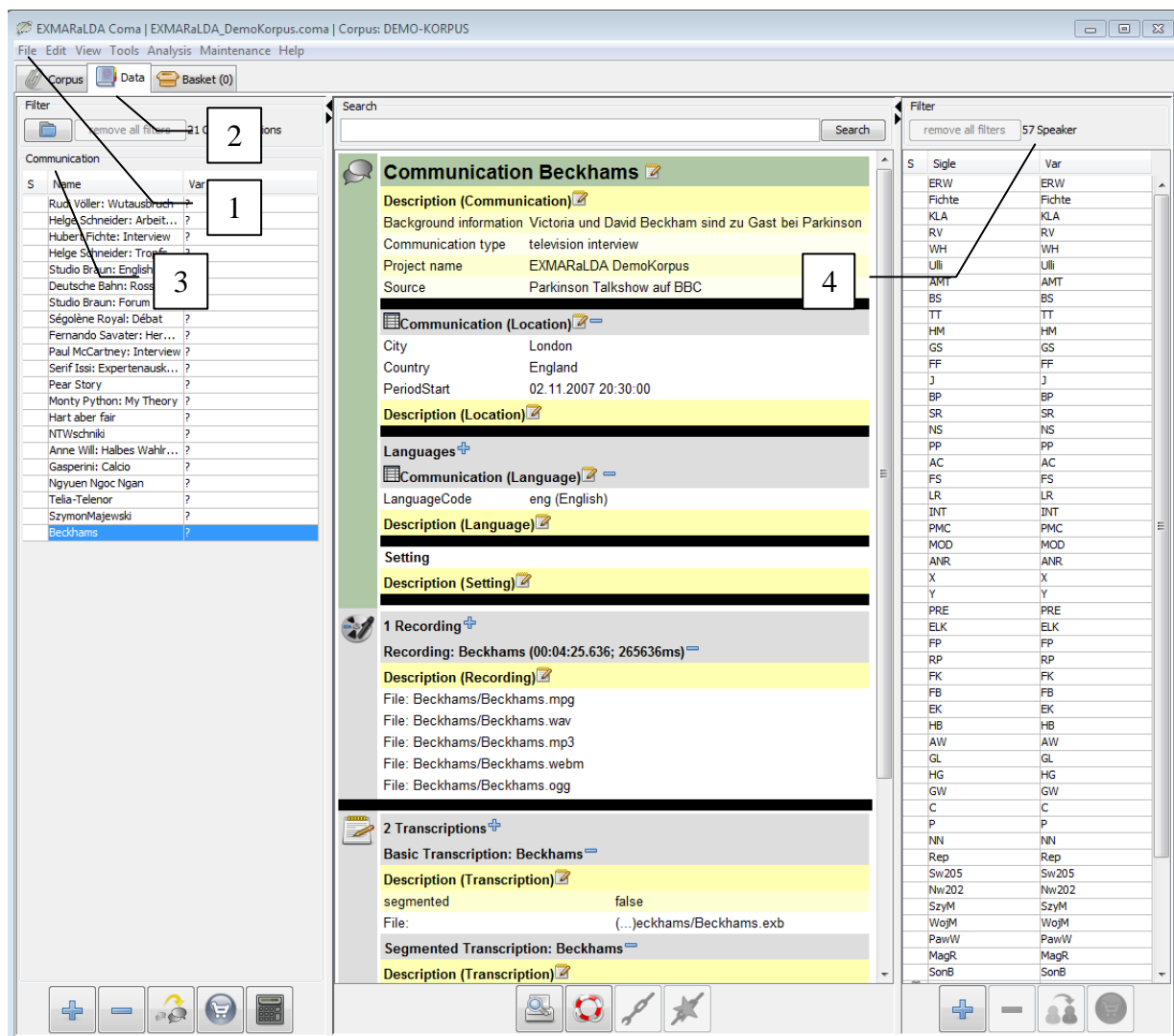


To continue from here, have a look at following documents and videos:

- Partitur-Editor video tutorial
- Quickstart documents for the Partitur-Editor
- Partitur-Editor Manual

## D. Corpus Manager

To get a first impression of the Corpus Manager, start it (see above) and open the corpus file of the demo corpus via **File > Open Document(1)**. The corpus file is in the top level directory and has the name "EXMARaLDA\_DemoKorpus.coma". After opening it, switch to the **Data** tab (2) and browse through the lists on the left and right to see meta-data about communications (3) and speakers (4).

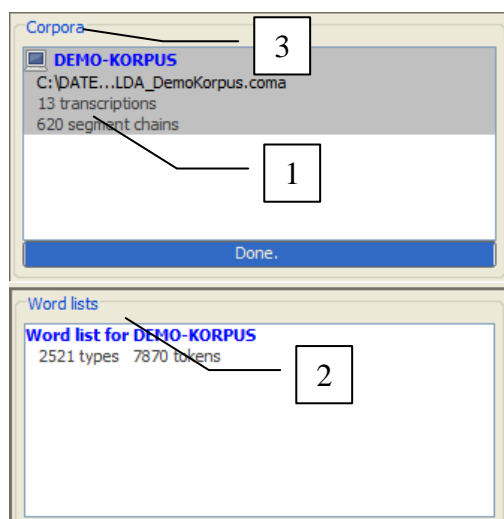


To continue from here, have a look at following documents:

- Corpus Manager Manual
- Quickstart documents for Corpus Manager

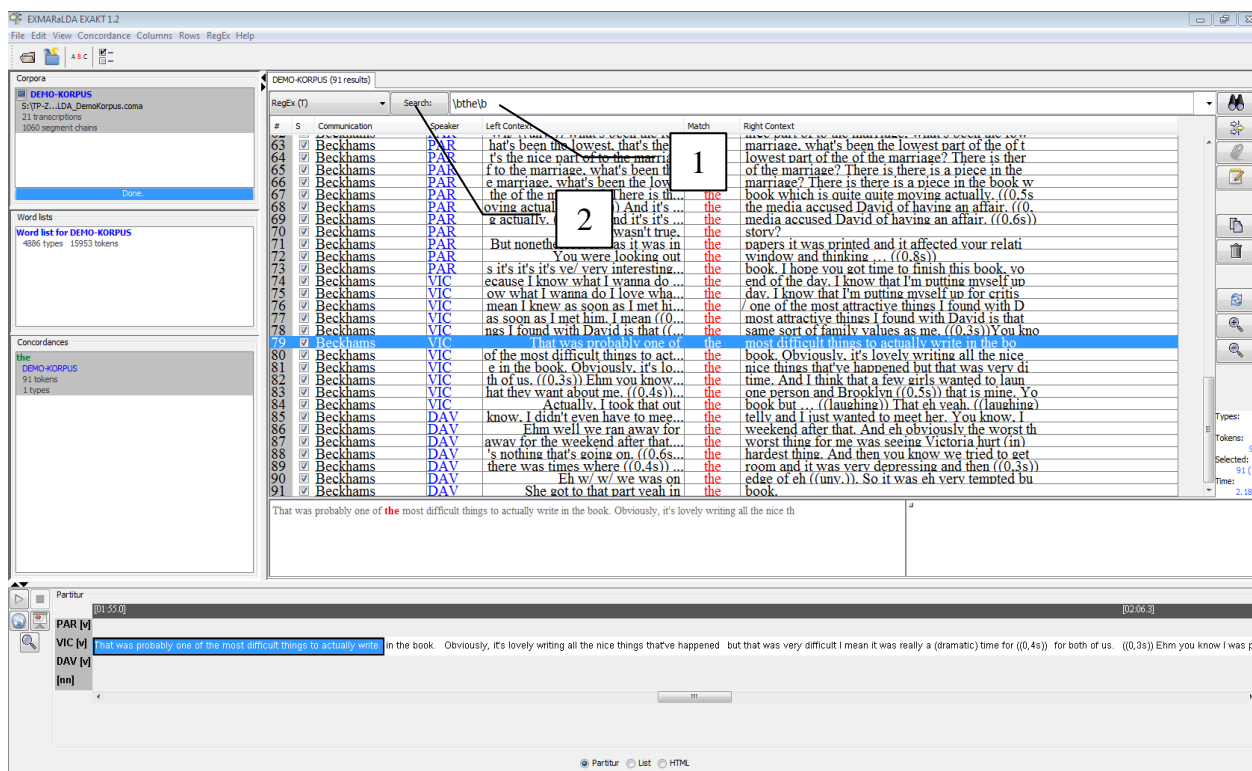
## E. EXAKT

To get a first impression of the query tool EXAKT, run it and open the corpus file of the demo corpus via **File > Open Corpus**:



The corpus file is in the top level directory and has the name "EXMARaLDA\_DemoKorpus.coma". Opening the corpus may take several seconds because EXAKT first creates an index, then a wordlist. The opened corpus (1) and the wordlist (2) will be displayed in the section titled "Corpora" (3).

A new concordance will be opened. Make a test query for the word 'the' by entering the regular expression <sup>1</sup>\bthe\b (1) in the text field beside the **Search** (2) button and hit enter to carry out the query:



To continue from here, have a look at following documents:

- EXAKT manual
- Document “Quickstart Regular Expressions”

<sup>1</sup> The characters “\b” indicate a word boundary. Searching using the word “the” would also yield results such as “them” or “bother”.