

License Agreement for the “Tübingen Partially Parsed Corpus of Written German“ (TüPP-D/Z)

This agreement is made and entered into as of _____ by and between the parties:

a) Eberhard Karls Universität Tübingen
Seminar für Sprachwissenschaft, Abt. Computerlinguistik
Wilhelmstr. 19
D-72074 Tübingen, Germany
“LICENSER“

b) _____

email: _____
“LICENSEE“

1. Product

TüPP-D/Z is a collection of articles from the taz newspaper which have been automatically annotated with clause structure, topological fields, and chunks, in addition to more low level annotation including parts of speech and morphological ambiguity classes. All texts are processed automatically, starting from paragraph, sentence and token segmentation. Tokens include information about some regular types of named entities, including dates, telephone numbers, and number/unit combinations.

The TüPP-D/Z data of the current release are taken from the 1999 HTML distribution (scientific edition) of the taz, which includes newspaper articles from September 2, 1986 up to May 7, 1999 and which amounts to more than 200 million word tokens of text.

Copyright of the linguistic annotation of TüPP-D/Z is to the Seminar für Sprachwissenschaft at the Eberhard-Karls Universität Tübingen.

2. Pre-Condition

The raw text of 'die tageszeitung' as used in the corpus is copyrighted by contrapress media gmbh, Berlin. LICENSEE certifies by signing this license agreement that LICENSEE has a valid license for the taz-CD edition or the taz-Wissenschafts CD-ROM (obtainable at <http://www.taz.de>).

3. License

LICENSER grants LICENSEE a non-exclusive license to use the TüPP-D/Z corpus.

LICENSEE agrees

- to use the corpus only for non-commercial, non-profit research purposes;
- to make no changes to the corpus;
- and to acknowledge the use of the corpus in all publications reporting on results produced with the help of the TüPP-D/Z corpus.

Use of the corpus or use of data derived from the corpus for any commercial purposes requires explicit written agreement of LICENSER.

4. Non-Disclosure

The TüPP-D/Z corpus will be held in confidence by LICENSEE and will not be disclosed by LICENSEE to third parties, except for those parts of the corpus that are explicitly marked for public use, and except for example sentences in scientific publications.

LICENSEE shall and will employ all necessary precautions to ensure that no persons or institutions other than persons as are in the employ of LICENSEE or in the same research project as LICENSEE will get access to the TüPP-D/Z corpus or parts thereof. Other persons or institutions desiring access to the corpus should be directed to LICENSER to obtain separate license agreements.

5. Commencement and Duration

The agreement will take effect on the date given above and will continue until terminated either by LICENSEE or LICENSER in accordance with section 6 below.

6. Termination

LICENSER may terminate the agreement forthwith if the LICENSEE commits any material breach of this agreement or if the copyright holder raises an objection against the terms of the license agreement.

LICENSEE may terminate this agreement within 30 days of written notice.

On termination howsoever caused, LICENSEE shall erase all copies of the TüPP-D/Z corpus.

7. Fee

As compensation for costs incurred by LICENSER in the preparation and shipment of the DVD data medium, this license is granted by LICENSER to LICENSEE for a fee of € 50

8. Disclaimer

The TüPP-D/Z corpus is provided “AS IS“, with no guarantee of its veracity or accuracy. No liability is accepted for any damage caused by its use.

LICENSER hereby informs LICENSEE that the distributed version of the TüPP-D/Z corpus is annotated automatically so that annotation errors are inevitable.

9. Acknowledgement

The raw text of 'die tageszeitung' (taz) as used in the corpus is copyrighted by contrapress media gmbh:

contrapress media gmbh
Kochstr. 18
10969 Berlin, Germany
Tel: +49 – 30 – 25902-212
Fax: +49 – 30 – 2515028
email: abu@taz.de
url: <http://www.taz.de>

Tübingen, _____
LICENSER
Signature _____
Name _____

_____, _____
LICENSEE
Signature _____
Name _____

Before sending the license, please make sure that you have included the copy of the taz CD invoice!