

# License Agreement for the “Tübingen Treebank of Written German” (TüBa-D/Z)

This agreement is made effective as of \_\_\_\_\_ [date] by and between the parties:

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

## “LICENSER“

b) Name: \_\_\_\_\_

Institution: \_\_\_\_\_

Address: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Email: \_\_\_\_\_

## “LICENSEE“ (institution, if applicable)

### 1. Product

The TüBa-D/Z treebank is a syntactically annotated, German newspaper corpus based on data taken from the daily issues of the 'die tageszeitung' (taz). The treebank currently comprises approximately 65,500 sentences (ca. 1,200,000 tokens). It is annotated on four levels: the lexical level, the phrasal level, the level of topological fields, and the clausal level. In addition to constituent structure, annotated trees contain edge labels between nodes. These edge labels encode grammatical functions (as relations between phrases) and the distinction between heads and non-heads (as phrase-internal relations).

The 65,500 sentences of the treebank (about 3,000 articles) are enriched with anaphoric and coreference relations referring to nominal and pronominal antecedents.

For selected connectives, the instances occurring in the treebank have been annotated with the discourse relation(s) conveyed by the connective instance.

Named Entities are classified as organisation, person, location, geo-political entity, and other.

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

### 2. License

LICENSER grants LICENSEE a non-exclusive license to use the TüBa-D/Z treebank. LICENSEE agrees

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

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

### 3. Non-Disclosure

The TüBa-D/Z treebank will be held in confidence by LICENSEE and will not be disclosed by LICENSEE to third parties, except for those parts of the treebank 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 TüBa-D/Z treebank or parts thereof. Other persons or institutions desiring access to the treebank should be directed to LICENSER to obtain separate license agreements.

### 4. 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 5 below.

### 5. 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 TüBa-D/Z treebank.

### 6. Fee

This license is granted by LICENSER to LICENSEE free of charge.

### 7. Disclaimer

The TüBa-D/Z treebank 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 TüBa-D/Z treebank is work in progress, and although large efforts are spent on removing annotation errors, some of them still remain in the treebank.

### 8. Acknowledgement

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

contrapress media GmbH  
Rudi-Dutschke-Straße 23  
10969 Berlin, Germany  
Tel: +49 – 30 – 25902-300  
Fax: +49 – 30 – 2515028  
Email: [abu@taz.de](mailto:abu@taz.de)  
URL: <http://www.taz.de>

Tübingen, \_\_\_\_\_

LICENSER

Signature \_\_\_\_\_

Name \_\_\_\_\_

\_\_\_\_\_, \_\_\_\_\_

LICENSEE

Signature \_\_\_\_\_

Name \_\_\_\_\_