

Summary of the Tagsets used in the German SemEval data

TüBa-D/Z is originally a constituent treebank ([Telljohann et al., 2009](#)). For the SemEval dataset, the syntactic information has been converted to a dependency scheme (see section 3).

1 Parts of Speech

Part-of-speech tags and morphological information use the *Stuttgart-Tübingen Tagset* (STTS; [Schiller et al., 1999](#)).

English summary of the STTS POS tags

Nouns

NN	common noun (der <i>Baum</i> , die <i>Katze</i>)
NE	proper noun (<i>Hans</i>)

Adjectives, Adverbs, Prepositions

ADJA	attributive adjective (das <i>große</i> Haus)
ADJD	adverbial or predicative adjective (er fährt <i>schnell</i> , er ist <i>groß</i>)
ADV	Adverb (<i>schon</i> , <i>bald</i> \approx already, soon)
APPR	Preposition, left part of circumposition in [der Stadt], ohne [mich]
APPRART	Preposition with article folded in im [Haus], zur [Sache]
APPO	Postposition (ihm <i>zufolge</i> , der Sache <i>wegen</i>)
APZR	Right part of circumposition (von/APPR jetzt an/APZR)
PROAV/PROP	Pronominal adverb (<i>davon</i> \approx thereof) PROP is used in the TüBa-D/Z treebank while PROAV is used by the Negra/Tiger treebanks)

Pronouns, Articles, etc. The “attributive pronouns” usually take a similar role to articles, although *PIDAT* can also occur as a pre-determiner or as a prenominal modifier.

ART	definite or indefinite article <i>der, die, das, ein, eine</i>
PDAT	attributive demonstrative pronoun (<i>dieser</i> Mensch)
PDS	substituting demonstrative pronoun (<i>das</i> habe ich geahnt)
PIAT	attributive indefinite pronoun that cannot occur with a determiner (<i>etliche</i> Dinge, <i>zu viele</i> Fragen)
PIDAT	attributive indefinite pronoun that can occur with a determiner (<i>die beiden</i> Brüder, <i>solch</i> eine Frage \approx both books, such a question)
PIS	substituting demonstrative pronoun (<i>man, keiner, viele</i>)
PPER	non-reflexive personal pronoun (<i>ich, er</i>)
PRF	reflexive pronoun (<i>mich, sich</i> \approx myself, himself/herself)
PPOSS	substituting possessive (<i>meins</i> \approx mine)
PPOSAT	attributive possessive (<i>mein</i> Buch \approx my book)
PRELAT	attributive relative pronoun (<i>der Mann, dessen</i> Hund ... \approx the man whose dog ...)
PRELS	substituting relative pronoun (<i>der</i> Hund, <i>der</i> ... \approx the dog that ...)
PWS	substituting interrogative pronoun (<i>wer, was</i> \approx who, what)
PWAT	attributive interrogative pronoun (<i>welche</i> Farbe \approx which color)
PWAV	adverbial interrogative or relative pronoun (<i>warum</i> \approx why)

Verbs

VVFIN	full finite verb (<i>er geht</i> \approx he goes)
VVIMP	full imperative (<i>geh!</i> \approx go!)
VVINFIN	full infinitive (<i>gehen, ankommen</i> \approx [to] go, [to] arrive)
VVIZU	full infinitive with “zu” (<i>anzukommen, loszulassen</i>)
VVPP	full past participle (<i>gegangen, angekommen</i>)
VAFIN	auxiliary finite verb (<i>du bist, wir werden</i>)
VAIMP	auxiliary imperative (<i>sei ruhig!</i> \approx be quiet!)
VAINFIN	auxiliary infinitive (<i>haben, werden, sein</i>)
VAPP	auxiliary past participle (<i>gewesen</i>)
VMFIN	modal finite verb (<i>wir dürfen</i> \approx we may)
VMINFIN	modal infinitive (<i>wollen</i>)
VMPP	modal past participle (<i>er hat gewollt</i> \approx he has wanted)
PTKVZ	particle part of separable verb (<i>er kommt/VVFIN an/PTKVZ</i>)
PTKZU	“zu” before infinitive (<i>zu gehen</i> \approx to go)

Everything else

KOUI	subordinating conjunction with “zu” and infinitive (<i>um zu leben, anstatt zu fragen</i> ≈ in order to live, instead of asking)
KOUS	subordinating conjunction with sentence (<i>weil, damit, daß</i> ≈ because, so that, that)
KON	coordinating conjunction (<i>und, oder, aber</i> ≈ and, or, but)
KOKOM	comparative conjunction (<i>als, wie</i>)
CARD	cardinal number (<i>zwei Männer</i> ≈ two men)
ITJ	interjection (<i>mhm, ach, tja</i>)
PTKNEG	Negation particle (<i>nicht</i>)
PTKANT	answer particle (<i>ja, nein, danke, bitte</i>)
PTKA	particle associated with adverb or adjective <i>am</i> [schönsten], <i>zu</i> [schnell]
TRUNC	first member of compound noun (<i>An- und Abreise</i>)
FM	foreign word (<i>die Queen war not amused</i>)
XY	Non-word (<i>3:7, H2O, D2XW3</i>)
\$,	comma ,
\$.	sentence ending punctuation . ? ! ; :
\$(other sentence signs; sentence internal - ,

2 Morphology

Gender (gend)	masc	masculine
	fem	feminine
	neut	neuter
Number (num)	sg	singular
	pl	plural
Person (per)	1	first
	2	second
	3	third
Case (cas)	n	nominative
	a	accusative
	d	dative
	g	genitive
Tense (temp)	pres	present
	past	past
Mood (mood)	ind	indicative
	konj	conjunctive (non-factual)

3 Dependencies

The dependency representation is oriented towards the guidelines of [Foth \(2003\)](#). We deviate from Foth’s guidelines in two points, for technical reasons: (i) all punctuation gets a `-PUNCT-` label and is attached to the last non-punctuation token and (ii) all roots (e.g., main clauses) have a `ROOT` dependency label.

Verb arguments

SUBJ	subject
OBJA	direct (accusative) object
OBJD	indirect (dative) object
OBJG	indirect (genitive) object
OBJP/PN	prepositional object (in/OBJP Betracht/PN ziehen)
SUBJC	subject clause
OBJC	clausal object
PRED	predicate (e.g., in a copula clause)
S	subordinated full sentence

Modifiers

PP/PN	prepositional adjunct (die Katze auf/PP dem Sofa/PN)
ADV	adverb
ATTR	noun attribute (<i>zehn</i> Autos, das <i>rote</i> Auto)
GMOD	genitive (<i>Peters</i> Auto \approx Peter’s car)
KOM/CJ	comparative (<i>schöner als/KOM der Sommer/CJ</i>)
ZEIT	temporal adjunct (<i>1994</i> geboren \approx born [in] 1994)
DET	determiner
KON/CJ	coordination (Peter, Paul/KON und/KON Maria/CJ)

References

- Foth, K. (2003). Eine umfassende Dependenzgrammatik des Deutschen. Technical report, Fachbereich Informatik, Universität Hamburg.
- Schiller, A., Teufel, S., Stöckert, C., and Thielen, C. (1999). Guidelines für das Tagging deutscher Textcorpora mit STTS (kleines und großes Tagset). Technical report, Universität Stuttgart / Universität Tübingen.
- Telljohann, H., Hinrichs, E. W., Kübler, S., Zinsmeister, H., and Beck, K. (2009). Stylebook for the Tübingen Treebank of Written German (TüBa-D/Z). Technical report, Seminar für Sprachwissenschaft, Universität Tübingen.