



EIN LEXIKALISCH-SEMANTISCHES WORTNETZ

GermaNet Wortnetz

Hyperonymie, Kohyponymie, Implikation, Lemmata, Literale, Nomina, Wörter, Literale, lexikalische Einheiten, Hyponymie, Holonymie, lexikalische Relationen, Frames, konzeptuelle Relationen, hierarchische Struktur, Meronymie, Verben, Kausativierung, künstliche Konzepte, Synonymie, Beispielsätze, Assoziation, Graph, Kreuzklassifikation, Paraphrasen, Lesarten, Antonymie, Eigennamen, Synset, Pertonymie

GERMANET

GermaNet ist ein lexikalisch-semantisches Wortnetz, in dem die bedeutungstragenden Kategorien Nomina, Verben und Adjektive modelliert sind. Jede Wortart ist dabei in mehrere semantische Felder subklassifiziert. Die zentrale Repräsentationseinheit ist das Synset, in dem lexikalische Einheiten, die das gleiche Konzept ausdrücken, zusammengefasst werden. Zwischen den Konzepten sind semantische Relationen definiert, die auf das gesamte Konzept oder auf einzelne lexikalische Einheiten Bezug nehmen können. GermaNet kann als Online-Thesaurus oder als „light-weight“ Ontologie betrachtet werden.

Seit 1997 wurde GermaNet im Rahmen verschiedener Projekte der Forschungsgruppe „Allgemeine Sprachwissenschaft und Computerlinguistik“ des Seminars für Sprachwissenschaft an der Universität Tübingen entwickelt. Das Basisvokabular von GermaNet ist in die multilinguale lexikalisch-semantische Datenbank EuroWordNet integriert worden.

DIE IDEE HINTER GERMANET

Mit GermaNet wurde ein lexikalisch-semantisches Wortnetz für das Deutsche entwickelt, das als Grundlagenressource für Anwendungen der maschinellen Sprachverarbeitung dient. Die Konzeption von GermaNet basierte auf einer weitgehenden Lesartendisambiguierung, die eine wichtige Voraussetzung für verschiedene Anwendungen wie maschinelle Übersetzung, Informationserschließung oder die Erstellung von Sprachtechnologietools bildet.

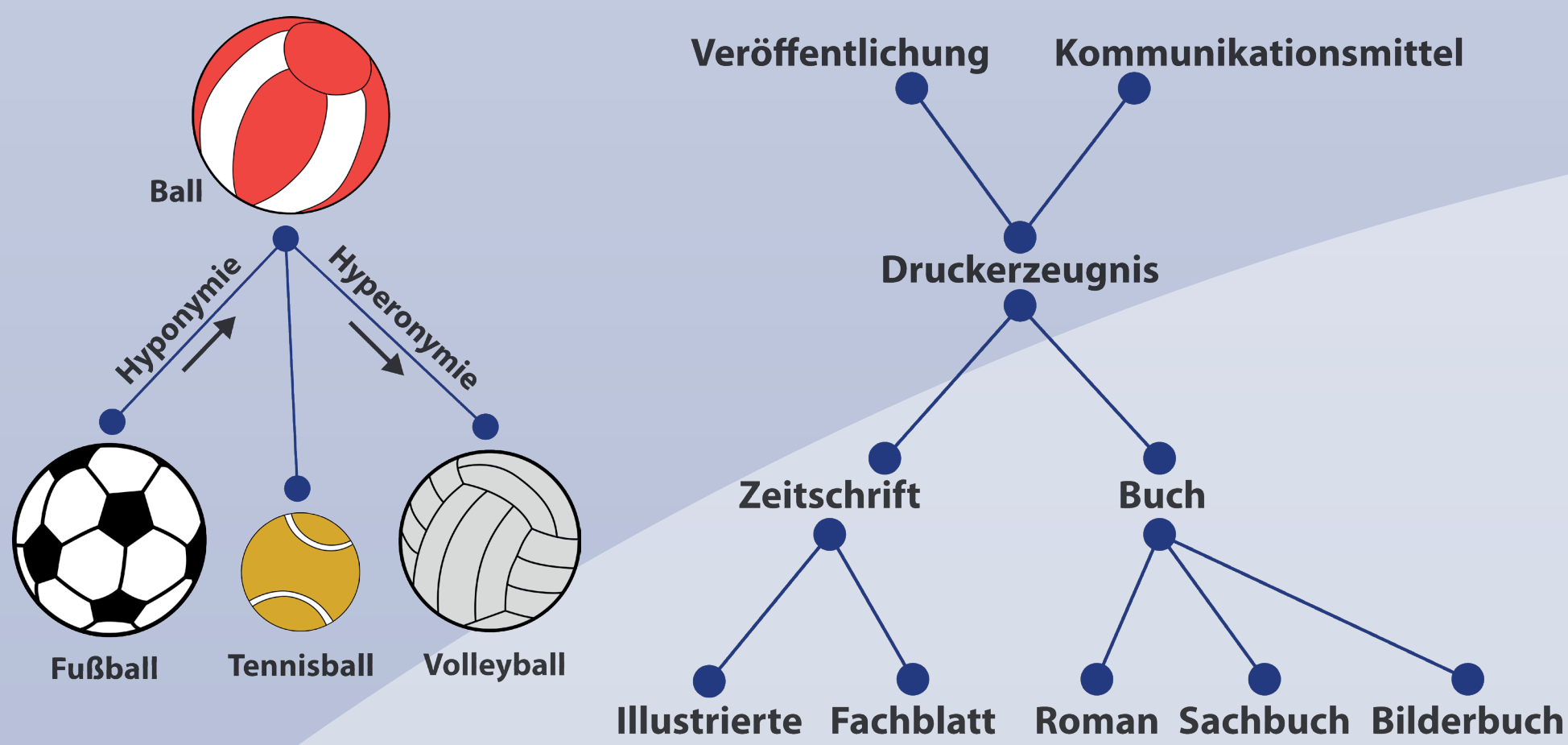
GERMANET & WORDNET®

Das deutschsprachige Wortnetz orientiert sich im Wesentlichen am Datenbankformat und den Strukturierungsprinzipien des englischen Princeton WordNet®. Dennoch unterscheidet sich GermaNet in einigen wesentlichen Aspekten von dem für das Englische entwickelte Wortnetz; so gelten folgende Punkte bspw. für GermaNet:

- Adjektive sind hierarchisch strukturiert
- GermaNet stellt einen komplett verbundenen Graphen dar
- die Relation der Kausativierung ist wortklassenübergreifend zulässig
- einheitliche Behandlung der Meronymie

DIE HIERARCHISCHE STRUKTUR

Durch die Hyperonymie-Hyponymie Relation ergibt sich die hierarchische Struktur von GermaNet, wie folgende Abbildungen verdeutlichen.



ABDECKUNG

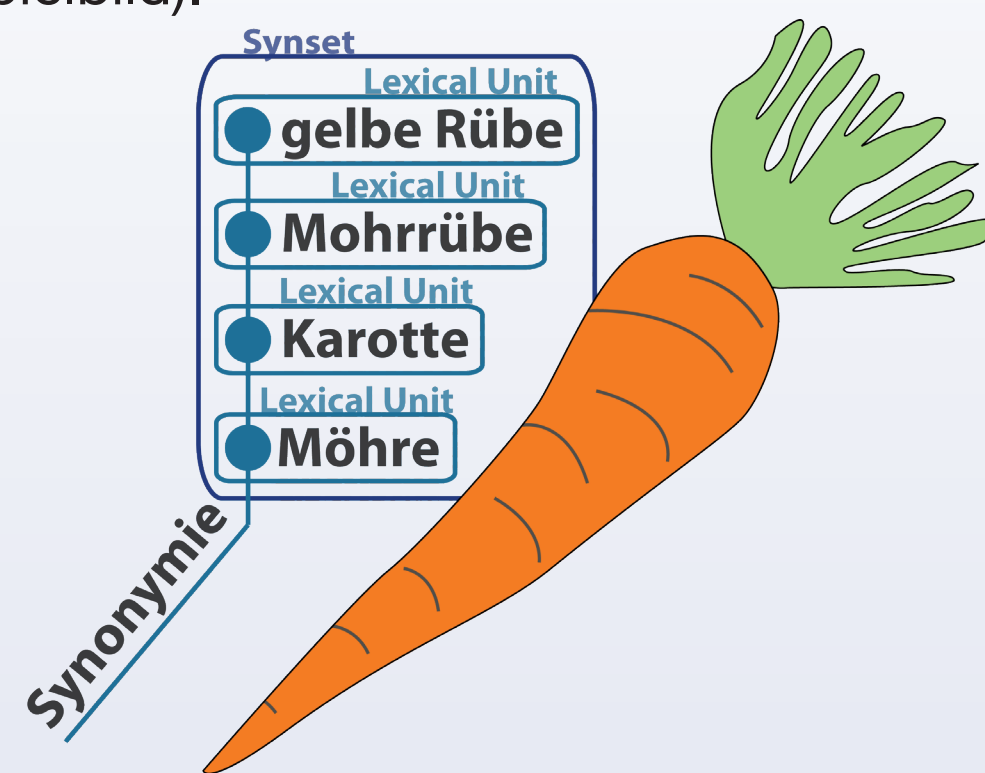
In seiner anfänglichen Version deckte GermaNet das deutsche Grundvokabular ab. Dieses Grundvokabular wurde durch statistische Analysen von Referenzkorpora sowie weiterer lexikografischer Quellen ermittelt. Da lexikalische Arbeiten unter Berücksichtigung der semantischen Felder stattfinden und nicht nur auf der Basis von Lemmatafrequenzen, ist es oft weder einfach noch gerechtfertigt, Lemmata unterhalb eines bestimmten Frequenzwertes nicht zu betrachten. Aus diesem Grund deckt GermaNet inzwischen deutlich mehr als nur das Grundvokabular ab.

Folgende Zahlen zeigen die aktuelle Statistik von GermaNet in der Version 5.2 (Release vom Dezember 2009):

- Anzahl der Literale: 76981
- Lesarten pro Literal: 1,10
- Anzahl der Synsets: 61575
- Anzahl der lexikalischen Einheiten: 84859
- Lexikalische Einheiten pro Synset: 1,38

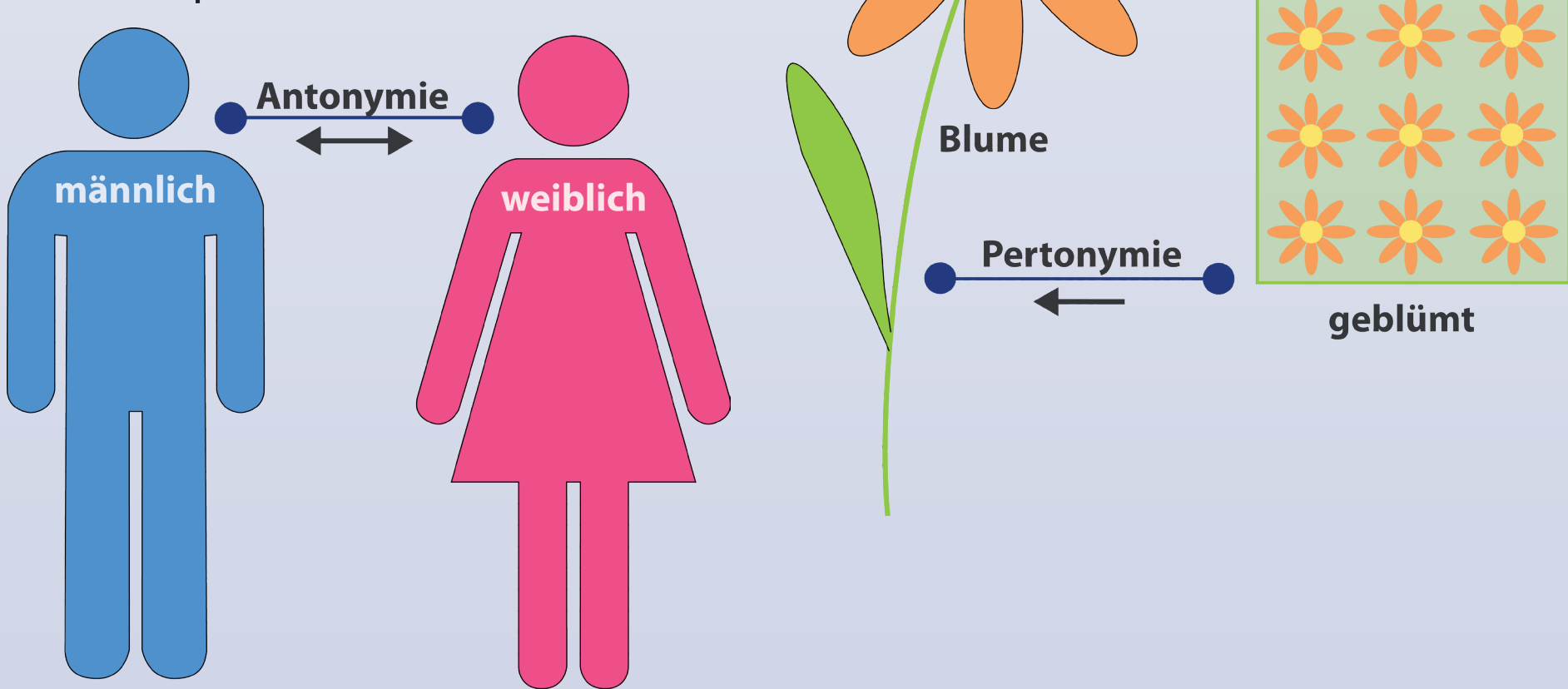
SYNSETS, LEXIKALISCHE EINHEITEN & RELATIONEN

Synsets bilden die grundlegenden Einheiten in GermaNet. Ein Synset stellt eine Menge von Synonymen dar, die das gleiche Konzept ausdrücken (siehe Beispielbild).



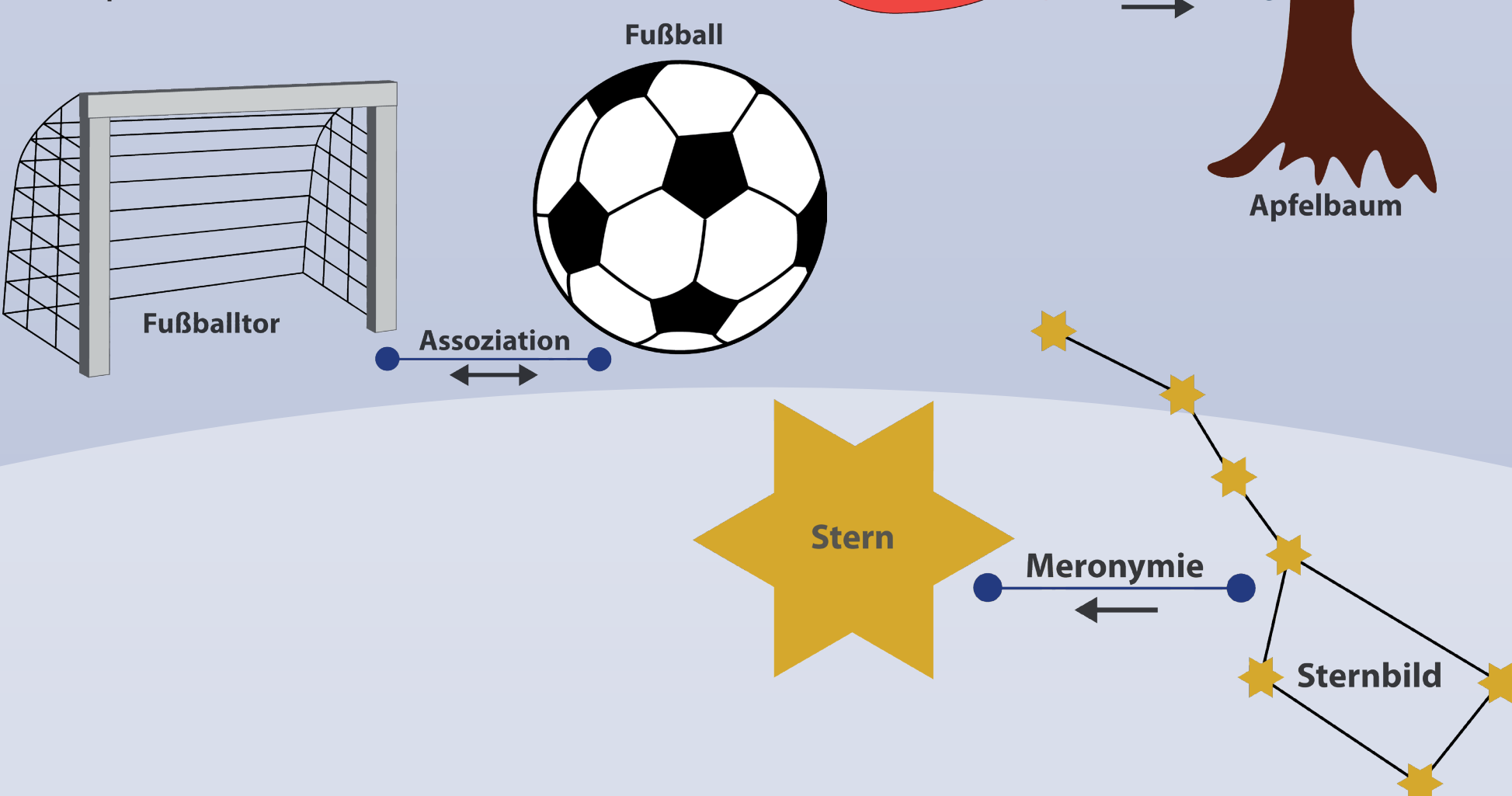
Es gibt zwei verschiedene Arten von Relationen: konzeptuelle und lexikalische. Lexikalische Relationen verbinden lexikalische Einheiten miteinander. Neben der Synset-internen Synonymie-Relation gibt es folgende lexikalische Relationen in GermaNet:

- Antonymie (siehe Beispielbild)
- Pertonymie (siehe Beispielbild)
- Partizipien

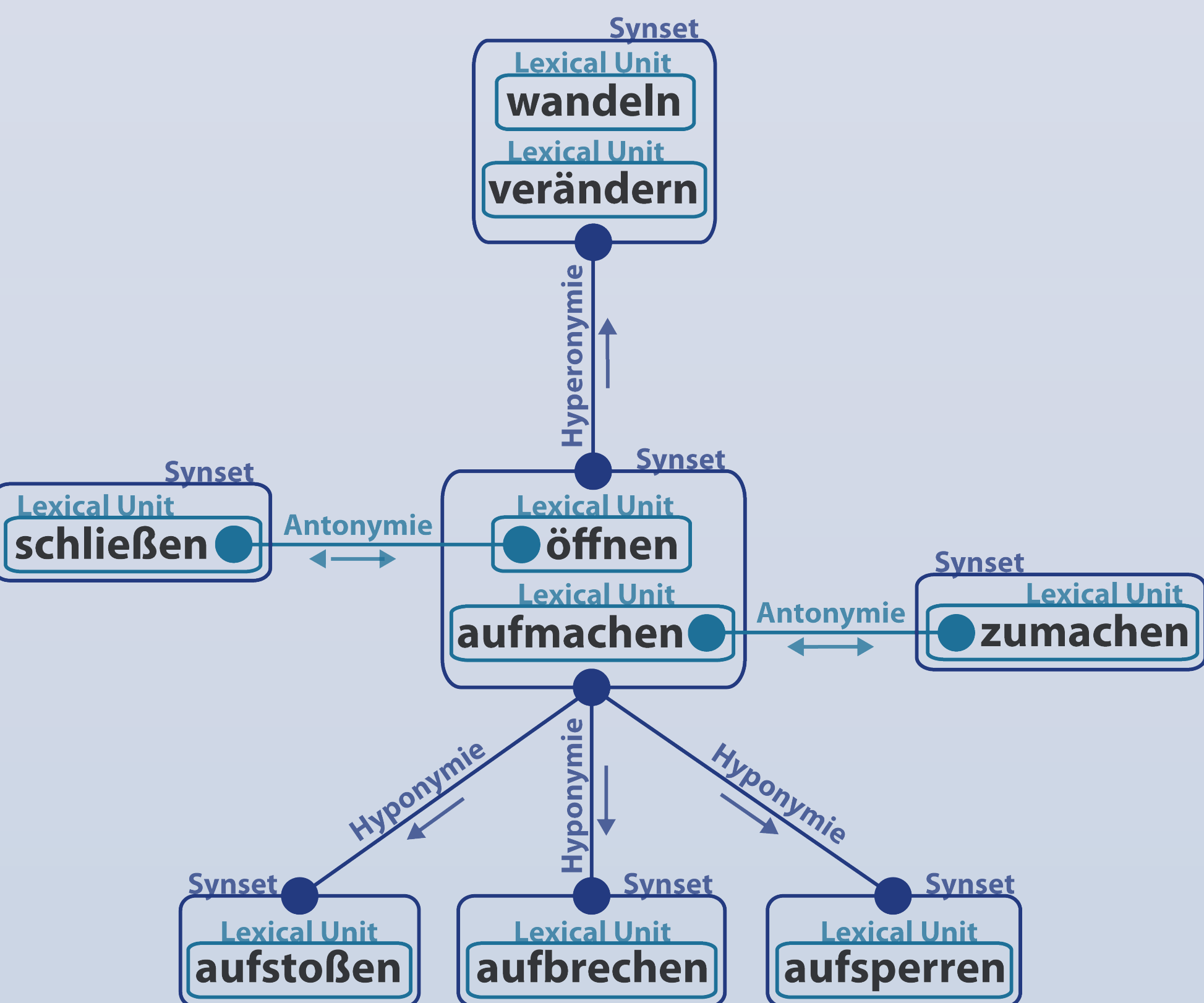


Konzeptuelle Relationen verbinden Synsets miteinander. Die folgenden konzeptuellen Relationen finden in GermaNet Verwendung:

- Hyperonymie/Hyponymie (siehe Beispielbilder links)
- Holonymie (siehe Beispielbild rechts)
- Meronymie (siehe Beispielbild)
- Assoziation (siehe Beispielbild)
- Kausativierung
- Implikation



Folgende Abbildung zeigt das Zusammenspiel lexikalischer und konzeptueller Relationen und verdeutlicht gleichzeitig den Unterschied zwischen einer lexikalischen Einheit und einem Synset.



LESARTEN

Lesarten für: fegen

- (v) [fegen] -- [fegen] -- NN,AN,BD -- Sie fegt die Zeitung vom Tisch. (NN,AN,BD)
 Hyperonyme: [Instrument zum Anstoß]
- (v) [rasen, düsen, preschen, sausen, brettern, heizen, fegen, flitzen]
 • [rasen]
 • [düsen]
 • [preschen]
 • [sause]
 • [brettern]
 • [heizen]
 • [fegen] -- NN,BD -- Der Wind fegt über das Feld. (NN,BD)
 • [flitzen]
 Hyperonyme: [eilen, spurten, beeilen, pressieren]
 Assoziationen: [Raser, Raserin]
- (v) [kehren, fegen, feudeln] -- mit einem Besen von Schmutz/Staub befreien
 • [kehren]
 • [fegen] -- NN,An -- Der Hausmeister hat den Hof gefegt. (NN,An)
 • [feudeln]
 Hyperonyme: [säubern, reinigen, putzen] [entfernen]

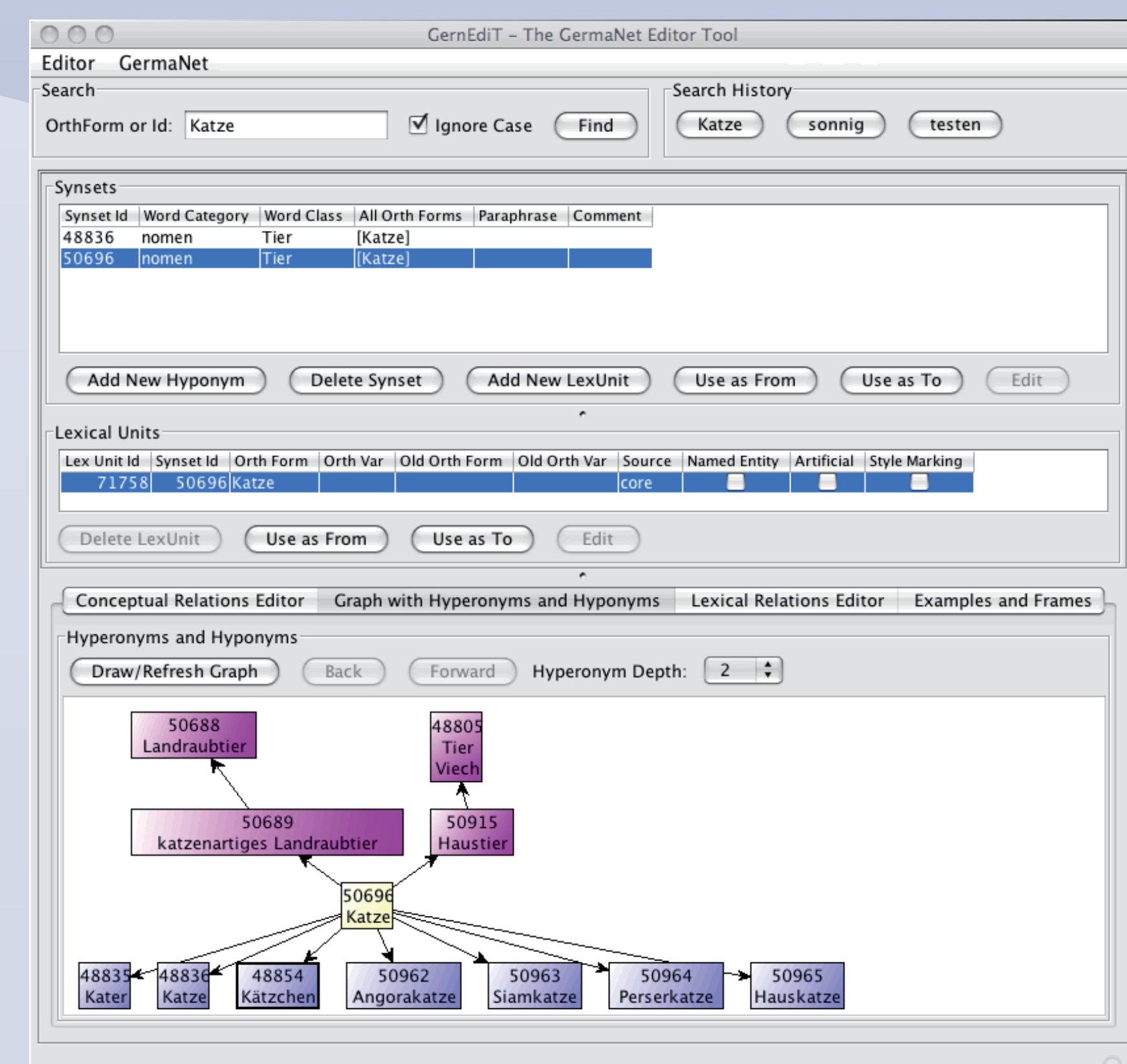
Lesarten für: unterhalten

- (v) [unterhalten, pflegen] -- über etwas verfügen
 • [unterhalten] -- NN,AN,Pp -- Sie unterhalten gute Beziehungen zu ihren Nachbarn.
 • [pflegen]
 Hyperonyme: [haben, besitzen]
- (v) [unterhalten] -- sich auf angenehme Weise die Zeit vertreiben
 • [unterhalten] -- NN,AR,BM -- Sie hat sich blendend unterhalten. (NN,AR,BM)
 Hyperonyme: [vergnügen]
- (v) [unterhalten] -- für Zerstreuung/Zeitvertreib sorgen
 • [unterhalten] -- NN,AN,Bs -- Er unterhielt seine Gäste mit Musik. (NN,AN,Bs)
 Hyperonyme: [vergnügen, amüsieren]
- (v) [unterhalten] -- etwas halten/einrichten/betreiben und dafür aufkommen
 • [unterhalten] -- NN,AN -- Er unterhält einen Reitstall. (NN,AN)
 Hyperonyme: [führen]
 Hyponyme: [instandhalten] [bewirtschaften]
- (v) [unterhalten] -- ein Gespräch führen
 • [unterhalten] -- NN,AR,Pp,Bo -- Er unterhielt sich den ganzen Abend über seine Prüfungen. (NN,AR,Pp)
 Er unterhielt sich nur mit mir. (NN,AR,Bo)
 Hyperonyme: [austauschen]
 Hyponyme: [klönen] [labern] [palavern] [philosophieren] [plauschen] [politisieren] [plaudern, schwatzen, schnattern]
- (v) [unterhalten, alimentieren] -- für jmds. Lebensunterhalt aufkommen
 • [unterhalten] -- NN,AN -- Er unterhält eine sieben-köpfige Familie. (NN,AN)
 • [alimentieren]
 Hyperonyme: [ernähren, nähren]

APPLIKATIONEN & TOOLS

Es gibt mehrere Applikationen für GermaNet:

- Application Programming Interfaces: stehen für Java und für Perl zur Verfügung.
- GermaNet-Explorer: Eine Software zur Visualisierung von GermaNet.
- GernEdiT (GermaNet Editing Tool): ist ein grafischer Editor, der ein benutzerfreundliches Bearbeiten der GermaNet-Inhalte ermöglicht. Dieses Programm wird von den Lexikographen verwendet, um Inhalte in die GermaNet-Datenbank einzupflegen oder diese zu bearbeiten (siehe Screenshot unten).
- Web Interface: um über das Internet auf GermaNet zuzugreifen.



LIZENZ

Wir bieten drei Arten von Lizenzabkommen für verschiedene Anwendungen und Benutzer an:

- Akademische Forschungslizenz
- Lizenz für Forschung & Entwicklung
- Kommerzielle Lizenz

Bitte schauen Sie auf unserer Homepage nach weiteren Informationen.



KONTAKT

Eberhard Karls Universität Tübingen
Seminar für Sprachwissenschaft
Wilhelmstr. 19
71072 Tübingen

Telefon: +49 7071 2974279
Fax: +49 7071 295214
E-Mail: germanetinfo@uni-tuebingen.de
Web: www.sfs.uni-tuebingen.de/GermaNet/