

Lexikalisches und enzyklopädisches Wissen

Claudia Kunze

kunze8@gmail.com, kunze@c1.uni-heidelberg.de

Computational Lexicography



Themen

- Begriffsklärung
- Lexikon-Enzyklopädie-Debatte im Kontext verschiedener Theorien
- Lexikalische und enzyklopädische Informationen in Lexika

Begriffsklärung

Die Unterscheidung *lexikalisches* vs. *enzyklopädisches Wissen* war und ist aktuell in der Sprachwissenschaft, Sprachphilosophie, Kognitionswissenschaft, Künstlichen Intelligenz etc. Beispiel:

Wiedehopf

- semantisches Wissen: ist ein Vogel, ist keine Amsel, Drossel etc.
- enzyklopädisches Wissen: rackenartiger Vogel, Angaben über Aussehen, Herkunft und Verhalten, episodisches und prozedurales Wissen sowie Faktenwissen...

keine klaren bzw. eindeutig akzeptierten Kriterien für Abgrenzung der Wissenstypen

Begriffsklärung

Weitere übliche Gegensatzpaare für die Lexikon-Enzyklopädie-Unterscheidung

- Semantik vs. Pragmatik
- Sprachwissen vs. Weltwissen
- definitorisches vs. enzyklopädisches Wissen über Wörter
- essenzielle vs. kontingente Eigenschaften
- analytische vs. synthetische Aussagen
- lexikalisches Wissen über Wörter vs. nicht-lexikalisches Wissen über Dinge
- Eigennamen vs. definite Beschreibungen

Abgrenzungsproblematik bleibt bestehen

Lexikon vs. Enzyklopädie im theoretischen

Strukturalismus

- Sprache als System: strikte Orientierung am beobachtbaren sprachlichen Zeichen
- Annahme einer semantischen Struktur des Lexikons
- autonomer Stellenwert der Semantik
- Wortbedeutungen werden über ein Merkmalsinventar erschlossen
- Kontextvariation und Weltwissen sind nicht Gegenstand der Analyse

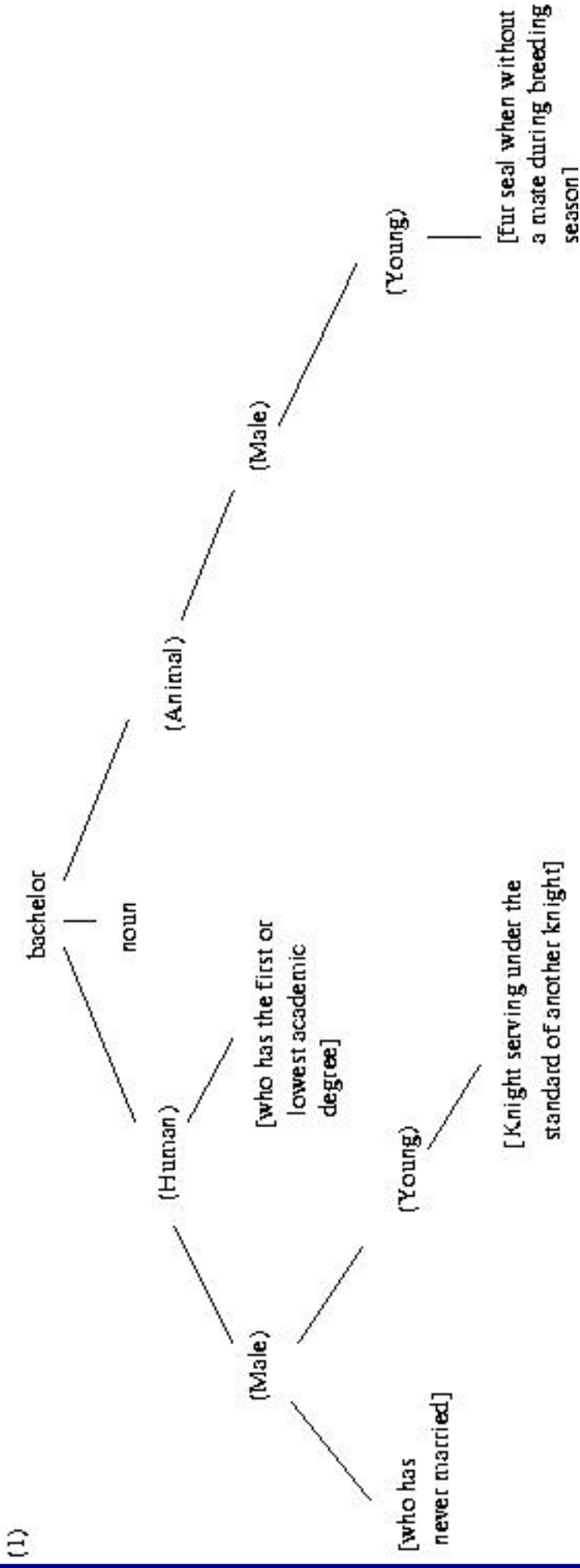
Lexikon vs. Enzyklopädie im theoretischen

Generative Grammatik

- Katz und Fodor (1966) führen Dichotomie semantisches vs. pragmatisches Wissen ein
- Semantikebene erhält Stellenwert innerhalb der syntax-dominierten Generativen Grammatik
- semantische Theorie zur Interpretation der syntaktisch generierten/ generierbaren Sätze
- Beschränkung auf satzsemantische Bedeutung, Ausklammerung von Kontextwissen
- Redeweise vom linguistisch relevanten Wissen, “Rest” ist enzyklopädisches Wissen
- Lexikon soll Verknüpfung zwischen Form und Bedeutung herstellen

Lexikon vs. Enzyklopädie im theoretischen

Semantische Marker bei Katz und Fodor



Zwei unterschiedliche Beschreibungsentitäten:

- Marker (runde Klammern): kennzeichnen sprachsystematisch relevante Merkmale
- Distinguisher (eckige Klammern): beliebige weitere semantische Eigenschaften

Lexikon vs. Enzyklopädie im theoretischen

- Bolinger zeigt, wie man über Distinguisher vermittelte semantische Information auf Marker reduziert werden können - Unterscheidung wird damit hinfällig
- ab den 70er Jahren: Katz & Fodor nutzen Marker und Distinguisher zur Repräsentation linguistischen Wissens
- einflußreicher Ansatz für die Entwicklung der generativen Semantik, kognitiven Semantik und Prototypentheorie

Lexikon vs. Enzyklopädie im theoretischen

Kognitive Semantik

Vorwiegend unikal enzyklopädische
Semantikauffassung: lexikalische Konzepte als
integraler Bestandteil menschlicher Kognition
(Geeraerts 1988, auch Lakoff, Fillmore)

- keine Annahme einer spezifisch linguistischen
Semantikebene: lexikalische Konzepte seien
aus interdisziplinärer kognitiver Sicht der KI,
(Psycho-)Linguistik, Anthropologie zu
beschreiben
- Frame als Repräsentationsformat für
lexikalische und enzyklopädische Information

Lexikon vs. Enzyklopädie im theoretischen

Aber: Unterscheidung zwischen lexikalischem und enzyklopädischem Wissen durch Wierzbicka (kognitive Semantikerin und Universalienforscherin) mit weit gefasstem Lexikonbegriff; man beachte den Beispielerintag *Maus*:

charakteristische Größe, Form und Farbe von Mäusen, ihre Gewohnheiten und Fortbewegungsart sowie ihre Geräusche. Mäuse seien ängstlich, ruhig und unauffällig, sie werden von Katzen gejagt; sie essen gern Käse und leben in der Nähe von Menschen; sie werden als Plage aufgefasst und die Menschen wollen sie loswerden. (Wierzbicka, zit. nach Peeters:2000, S. 16f.)

Lexikon vs. Enzyklopädie im theoretischen

Rechtfertigung durch linguistische Auftretenskontexte von

Maus

- eine Katze kann ein guter Mäusefänger sein,
- eine schüchterne, ängstliche oder unauffällige Person kann als *Maus* bezeichnet werden, z.B. als *graue Maus*,
- Idiome, z.B. *arm wie eine Kirchenmaus*.

Enzyklopädische Fakten über Mäuse, z.B. genaue geographische Verteilungen, Länge der Schwangerschaft, Körpergewicht etc. für Wierzbicka kein linguistisches Wissen. Schwierigkeit der Grenzziehung nach Wierzbickas Kriterien liegt auf der Hand.

Lexikon vs. Enzyklopädie im theoretischen

Fortentwicklungen der Generativen Grammatik

- Zwei-Ebenen-Semantik (Bierwisch, Lang, Wunderlich): mit der Ebene der Semantischen Form (SF), welche die grammatisch relevante Bedeutung repräsentiert und der Ebene der Konzeptuellen Struktur (CS), welche die außersprachliche Wirklichkeit repräsentiert
- Jackendoff (1983, 1990): autonome Syntaxkomponente, jedoch wird rein konzeptuelle Prägung der semantischen Struktur angenommen, direkte Abbildung auf das mentale Lexikon, keine unabhängige Lexikalische Semantik

Lexikon vs. Enzyklopädie im theoretischen

Generatives Lexikon Pustejovski (1991) unterscheidet zwischen linguistischem Wissen und Allgemeinwissen ('common sense knowledge')

- SELECTIVE BINDING und TYPE COERCION werden als lexikalische Prozesse aufgefasst, die über die Qualiastruktur von Nomen gesteuert werden
- SELECTIVE BINDING: Adjektive verbinden sich kontextsensitiv mit einer semantischen Rolle des modifizierten Nomens: guter Pilot/ gutes Essen/ guter Film
- TYPE COERCION: Typeshift vom Eigennamen zum geschriebenen Objekt: *Ich lese gern Walser.*

Klinkenberg: Prozesse auf enzyklopädischer Ebene

Lexikon vs. Enzyklopädie im theoretischen

Zusammenfassung

- wird Grenzziehung überhaupt postuliert: ja (generative Semantik) oder nein (kognitive Linguistik)
- wenn Grenzziehung, dann wird sie in verschiedenen Lagern, ggf. von Autor zu Autor, unterschiedlich getroffen
- meist Akzentuierung einer Ebene zu Ungunsten der jeweils anderen Komponente (z.B. Aufblähung des Lexikons bei Wierzbicka und Pustejovsky)
- ausführliche Debatte bei Peeters (2000)

Lexikalische und enzyklopädische Informa

Vagheit der Unterscheidung spielt für Wörterbuchklassifikation zwar ebenfalls eine Rolle, doch es gibt in der Praxis relativ gut erfassbare Unterschiede auf der Beschreibungsebene:

- Wörterbuch (Sprachwissen): mehr und ausführlichere sprachliche Information
- Enzyklopädie (Sachwissen): wenig sprachliche Information, mehr Hintergrundwissen, andere Stichwortauswahl (mehr Fachausdrücke und Eigennamen)
- intelligenter Nutzer entscheidet, welche Informationen er benötigt, ggf. Abbruch der Lektüre des Artikels oder Konsultation weiterer Nachschlagewerke

Definitonen aus dem DUW

1. **Wörterbuch**, **das** Nachschlagewerk, in dem die Wörter einer Sprache nach bestimmten Gesichtspunkten ausgewählt, angeordnet und erklärt sind; ein ein-, zweisprachiges, etymologisches W.; ein W. der deutschen Umgangssprache.
2. **Lexikon**, **das** 1. nach Stichwörtern geordnetes Nachschlagewerk für alle Wissensgebiete oder für ein bestimmtes Sachgebiet: ein L. in fünfzehn Bänden; ein L. der Kunstwissenschaft; L. für Wirtschaft und Handel; L. zur Sexualität; er ist ein wandelndes/lebendes L. (ugs. scherzh.: weiß auf allen Gebieten Bescheid; 2. (veraltet) Wörterbuch 3. (Sprachw.) a. Gesamtheit der selbstständigen bedeutungstragenden Einheiten einer Sprache; Wortschatz im Unterschied zur Grammatik einer Sprache; b. (in der generativen Grammatik) Sammlung der Lexikoneinträge einer Sprache.
3. **Enzyklopädie**, **die** Nachschlagewerk, in dem der gesamte Wissensstoff aller Disziplinen oder nur eines Fachgebiets in alphabetischer oder systematischer Anordnung dargestellt ist.

Praktische Unterscheidungen für die Lexik

- **Wörterbuch:** ein Nachschlagewerk, das sprachliches (= formbasiertes und inhaltsbezogenes) Wissen zu Wörtern oder anderen lexikalischen Einheiten enthält, z.B. das Duden Universalwörterbuch;
- **Enzyklopädie:** ein Nachschlagewerk, das Sach- und Fachwissen zu lexikalischen Einheiten und Begriffen enthält, z.B. die Enzyklopädie von Brockhaus oder auch die Online-Enzyklopädie *Wikipedia*;
- **das enzyklopädische Wörterbuch**, das sowohl Sprach- als auch Sachwissen zu lexikalischen Einheiten aufführt, z.B. Brockhaus in vier Bänden (1968);

Praktische Unterscheidungen für die Lexik

- das Übersetzungswörterbuch, das die Äquivalente zu lexikalischen Einheiten in der Zielsprache bereitstellt;
- das Fachwörterbuch, das überwiegend sachbezogene Informationen zu den Fachtermini einer Disziplin bereitstellt, z.B. das sprachwissenschaftliche Standardwerk von Bussmann;
- das Fremdwörterbuch, in dem nicht-native Wörter aufgelistet und beschrieben werden.

Zwei Definitionen zum Stichwort *Banane*

(1) Ba-na-ne *die*: -, -n; e-e längliche, gekrümmte, tropische Frucht mit gelber Schale u. e-m weißen, süß schmeckenden Fruchtfleisch || Abbildung unter *Obst* || K-: Bananen-, -schale, -staude

(2) *Banane* (Musa), tropische baumartige großblättrige Staude mit palmenartigem Aussehen und deren gelbschalige Frucht. Die vitaminhaltigen Früchte werden grün geerntet und erhalten ihre Reife während des Transports zum Verbraucher. In den Anbaugebieten (Tropen, zum Teil auch Subtropen) ist die Banane Grundnahrungsmittel. Bekannt sind auch die *Faser-* (Manilahanf) und die *Zierbanane*. – Bananenkulturen wurden bereits von den Geographen Alexanders des Großen beschrieben, das heutige Europa kennt sie seit der Entdeckung Amerikas.

Beobachtungen zu den Einträgen

Eintrag aus Langenscheidts Großwörterbuch Deutsch als Fremdsprache (1998):

- formbasierte Information: Silbentrennung, Genus, Pluralbildung
- kurze attributreiche Definition als *Frucht*
- Angabe prototypischer Eigenschaften: *länglich, gekrümmt, gelbe Schale, süß schmeckendes Fruchtfleisch*
- einzig *tropisch* ist enzyklopädische Angabe
- Verweis auf Abbildung im semantischen Feld *Obst*
- Beispiele für Komposita mit *Banane*

Beobachtungen zu den Einträgen

Eintrag aus der Enzyklopädie von Trautwein (2000)

- keine Angaben zur sprachlichen Form
- Angabe des botanischen Fachausdrucks
- Einordnung als *Stau*de und deren *gelbschalige Frucht*
- keine Angaben zur Form und zum Geschmack
- Angaben zu Anbau, Anbaugelieten, Transportweg und Reifungsprozess,
- zum Status als Grundnahrungsmittel in Anbaugelieten
- Nennung zweier (vermutlich nicht essbarer) Unterarten
- historischer Verweis auf Bekanntheitsgrad in Europa
- Abbildung (hier nicht sichtbar)

Ausblick

- Hybridformen der enzyklopädischen Wörterbücher sind zunehmend beliebt, vor allem bz. Lernerwörterbücher, wo zu einem Stichwort neben der Schreibung und Verwendung auch fachsprachliches Wissen und kulturelle Besonderheiten aufgeführt werden;
- elektronische Ressourcen unterliegen nicht den in Printlexika üblichen Beschränkungen und eignen sich besonders für die Adaption an Nutzerbedürfnisse;
- KI-Projekt Cyc stellt umfangreiches Weltwissen zur Verfügung;
- auch das OpenSource-Projekt Wikipedia lässt sich für lexikographische Anwendungen nutzen