

IDS INSTITUT FÜR DEUTSCHE SPRACHE



Rainer Perkuhn, Programmbereich Korpuslinguistik (IDS Mannheim)

Valenz und Kollokation – Zwischen empirischen Befunden und abstrakter Beschreibung

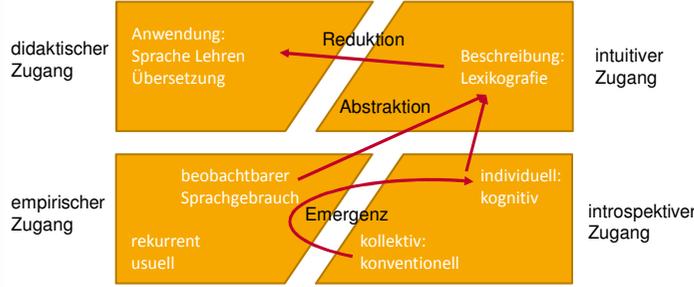
GAL-Sektionentagung, Valenz und Kollokation im (digitalen) Wörterbuch, Frankfurt/Oder, 24.09.2015

Wolfgang Iser
Leibniz
Leibniz-Gesellschaft

IDS INSTITUT FÜR DEUTSCHE SPRACHE

DER BEGRIFF „KOLLOKATION“ IN VERSCHIEDENEN DOMÄNEN

- Kollokationen als überraschende / beschreibenswerte / lernenswerte Wortkombinationen

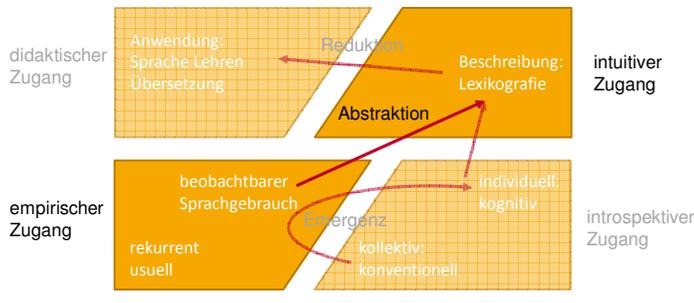


2

IDS INSTITUT FÜR DEUTSCHE SPRACHE

DER BEGRIFF „KOLLOKATION“ IN VERSCHIEDENEN DOMÄNEN

- Kollokationen als überraschende / beschreibenswerte / lernenswerte Wortkombinationen



3

IDS INSTITUT FÜR DEUTSCHE SPRACHE

WUNSCH NACH EINFACHER BESCHREIBUNG VS. VIELFALT IN DER AUSPRÄGUNG

<p>Lexikographie</p> <p>Mehr oder weniger feste typische Wortverbindungen</p> <p>plus ggf. weitere Einschränkungen:</p> <ul style="list-style-type: none"> wortklassenbezogene Asymmetrie nicht-freie Kombination keine Übersetzungsäquivalenz syntaktische Einheit 	<p>Empirische Ansätze</p> <p>präferentielle statistisch berechenbare Kombinationen von Wörtern und/oder Wortformen? auf festen Positionen? mit festem Abstand? auf variablen Positionen innerhalb eines Textfensters?</p> <p>→ was ist berechenbar ← wie viel kann zusammengefasst werden zu welcher Nennform</p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

4

WANN SPRECHEN WIR VON
BLONDEN HAAREN, BLONDEM HAAR ODER EINEM BLONDEN HAAR?



The slide features two images. On the left, a woman with long, wavy blonde hair is shown from the chest up, holding a hairbrush to her hair. On the right, a single, long, thin blonde hair strand is shown against a dark background.

WANN SPRECHEN WIR VON
GRAUEN HAAREN, GRAUEM HAAR ODER EINEM GRAUEN HAAR?



The slide features three images. On the left, a man with short grey hair is smiling. On the top right, a man with dark hair is shown. On the bottom right, a woman with dark hair is shown.

WANN SPRECHEN WIR VON
BLAUEN AUGEN ODER EINEM BLAUEN AUGE?



The slide features two images. On the left, a close-up of a baby's face with bright blue eyes. On the right, a man with a purple bruise on his eye.

WIR PUTZEN NUR AUSNAHMSWEISE EINZELNE ZÄHNE
ABER PUTZEN, PUTZE, PUTZT, GEPUTZT ... IST EGAL



The slide features two images. On the left, a woman brushing her teeth. On the right, a baby's face with a red arrow pointing to a tooth.


 INSTITUT FÜR DEUTSCHE SPRACHE

WAS SOLLEN WIR EMPIRISCH ZUGRUNDE LEGEN? BIGRAMME?

ein Eigentor. Sie wollen ihre **Zähne putzen** und schützen, erreichen aber genau das Viele tun wir täglich. **Zähne putzen**, über das Wetter reden, die Tagesschau in dem dereinst Kinder die **Zähne putzen** werden. Beim Trakt B (für die Technik Chips oder den Süßigkeiten die **Zähne putzen** müssen. Und zum Schluss verabschieden ist.» Also zum Beispiel **Zähne putzen** statt Schokolade schlecken. Oder die Nämlich essen, waschen, **Zähne putzen** und schlafen – jawohl, auch miteinander die Kartoffeln schälen, die **Zähne putzen**, die Telefonnummern notieren oder den von ihrer besten Seite zeigen. **Zähne putzen**, Haare kämmen, Sonntagsgewand anziehen. Mass an Pflege wie duschen, **Zähne putzen**, ein frisch rasiertes Gesicht, saubere tonisieren und eincremen, **Zähne putzen**, mehrmals wöchentlich die Haare waschen an die Morgentoilette, kämmen, **Zähne putzen**, rasieren. «Das häsch du no nie gmacht richtig ernähren und sorgfältig **Zähne putzen** Zahnärzte stellen seit einiger Zeit der Nacht seien zu vermeiden. **Zähne putzen** mit fluoridhaltiger Kinderzahnpaste, Spielend **Zähne putzen** Richtige Zahnpflege beginnt bereits im Kinder dann schon selbst ihre **Zähne putzen**. Dazu eignet sich am besten die regelmäßig und gründlich ihre **Zähne putzen**, wird auch für die Kinder die tägliche und am besten auch damit **Zähne putzen**. Offene Getränke und Eiswürfel Ritual im Bad Zahnseideln, **Zähne putzen**, Nase spülen, Gesicht waschen,

→ Kollokation „Zähne putzen“?

9


 INSTITUT FÜR DEUTSCHE SPRACHE

WAS SOLLEN WIR EMPIRISCH ZUGRUNDE LEGEN? TEXTFENSTER (HIER: +/-5 WÖRTER)?

Zahnpasta ein Eigentor. Sie wollen ihre **Zähne putzen** und schützen, erreichen aber werden alle zu Multi-Taskern, **putzen Zähne** und Schuhe. Die Denkanfrage kann auch Badezimmer: "Vergesst nicht, euch die **Zähne zu putzen** und die Hände vor dem Essen Meinung müssen nicht nur Menschen ihre **Zähne** regelmäßig **putzen**, um sie frei von Wer eine Spange trägt, sollte seine **Zähne** übrigens besonders gründlich **putzen**. Jahrhundert wurden Männer, die sich die **Zähne putzten**, sogar als dekadent bezeichnet in Deutschland **putzt** sich falsch die **Zähne**. Das geht aus einer Studie des Kölner Zahnbürste besitzt, **putzt** seine **Zähne** lieber, länger und mit mehr Inbrunst Hoffmann. **Putzt** ein Patient optimal die **Zähne** und hat keine Beschwerden, reiche

- Wie **putze** ich denn nun richtig **Zähne**? Generell gilt: eine Zahnbürste mit Honig und Marmelade **putze** ich mir die **Zähne**. Der sogenannte Bass-Stil soll die mehr, auch keinen Fernet. **Putze** deine **Zähne** länger als eine Minute. Miß deinen jetzt deinen Pyjama an und **putz** dir die **Zähne!**" an den Neunjährigen um einen Grad gern, telefoniere oder **putz** mir die **Zähne**. Texte schreiben gehört nicht dann folgenden neun Monaten sollten die **Zähne** besonders sorgfältig **geputzt** und süsse Alkohol und Nikotin. **Geputzt** werden die **Zähne** wie immer, Wundreinigung erfolgt Mit rüttelnden Bewegungen versuche ich, **Zahn für Zahn zu putzen**. Ich komme mit der werden wir uns noch eifriger **putzen**, **Zahn um Zahn**. Frischer Atem macht

→ Kollokation „Zähne & putzen“?

10


 INSTITUT FÜR DEUTSCHE SPRACHE

ÜBERSICHT DER TREFFERZAHLEN

Suchanfrage über Cosmas II in DeReKo	Trefferzahl
Zähne /+w1,s0 putzen	1291
Zähne /+w1,s0 &putzen	2178
Zähne /w5,s0 &putzen	3763
&Zahn /w5,s0 &putzen	3809

11


 INSTITUT FÜR DEUTSCHE SPRACHE

LANGENSCHIEDT DAF (1993 – 2008) BENUTZERHINWEISE

typische Kombinationen

7.4. Kollokationen: < >

Kollokationen sind typische Verbindungen von Wörtern, die eine syntaktische Einheit bilden. Sie sind für den Lernenden von besonderer Bedeutung, denn sie zeigen ihm „Partner“, mit denen das Stichwort häufig zu finden ist. Das ist wichtig für die Sprachproduktion, bei der es ja besonders darauf ankommt, die unterschiedlicher Wortarten + Substantiv, Verb + Objekt usw.) so zu verknüpfen, daß sie zusammenpassen (kompatibel sind).

Der Begriff Kollokation wird in diesem Wörterbuch relativ weit gefaßt, so daß dazu auch durchaus lose Kombinationen zählen, die aufgrund ihrer semantischen Verträglichkeit eine Einheit bilden und somit für den Benutzer relevant sind. Die Kollokationen erscheinen hier in spitzen Klammern < >.

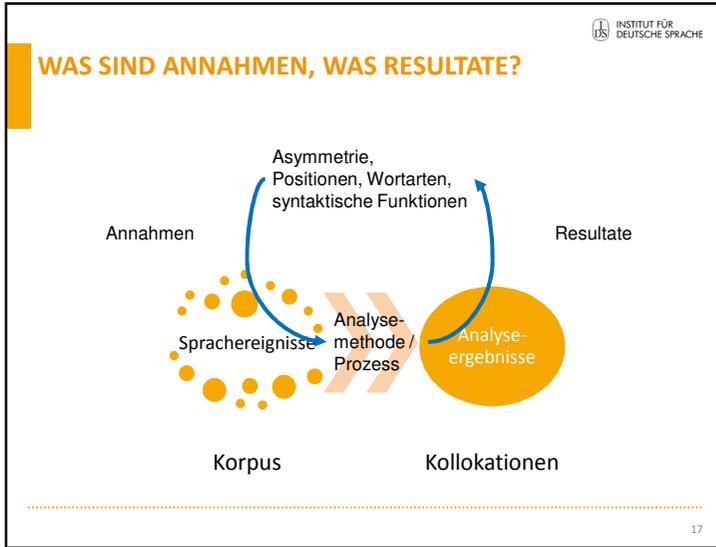
gekennzeichnet mit spitzen Klammern

weitere typische Kombinationen werden als Beispiele angegeben

Wenn das Stichwort in der Regel einem bestimmten Wort oder wenigen anderen Wörtern verwendet werden kann, wird dies durch den Hinweis *mit* (= meist, meistens) innerhalb der spitzen Klammern verzeichnet:

aufzuleihen ... [1] 1 etw. *zieht* auf ... (mit Neben-, zu Gewitter)

12



- IDS INSTITUT FÜR DEUTSCHE SPRACHE
- ### ZU: AUF WELCHER POSITION UND IN WELCHER FORM?
- Kookkurrenzanalyse des IDS bietet
- Autofokus: liefert Hinweis auf präferierte Position des primären Partnerwortes
 - syntagmatische Muster
 - die dominierenden Formen des Suchausdrucks
 - die hervorstechende Reihenfolge des Suchausdrucks und der kollokierenden Wortformen
 - sowie die häufigsten Füllungen dazwischen
- 20

IDS INSTITUT FÜR DEUTSCHE SPRACHE

KOOKKURRENZEN ZU &HAAR FORMEN *BLOND* UND POSITIONEN

Kookkurrenzprofil zu Lemma <i>Haar</i>		Autofokus	LLR	Partnerwörter primär/weitere	Häufigkeit
		-1	-1	19942	blonden schulterlangen blauen
-1	-1	19942	blonden schulterlangen	45	
-1	-1	19942	blonden blauen	81	
-1	-1	19942	blonden	1568	
...	
-3	-1	14575	blonde schulterlange	64	
-3	-1	14575	blonde blaue	91	
-3	-1	14575	blonde	1331	
...	
-2	-1	11271	blondes kurzes schulterlanges	1	
-2	-1	11271	blondes kurzes	115	
-2	-1	11271	blondes schulterlanges	62	
-2	-1	11271	blondes langes	92	
-2	-1	11271	blondes	780	
...	
-4	5	4477	blond gefärbte	43	
-4	5	4477	blond	507	

19

IDS INSTITUT FÜR DEUTSCHE SPRACHE

KOOKKURRENZEN ZU &HAAR SYNTAGMATISCHE MUSTER

AF	LLR	Partnerw.	Ant.	syntagmatisches Muster
-1	-1	19942	blonden	51% den/mit langen blonden [...] Haaren und ...
-3	-1	14575	blonde	83% hat blonde [...] Haare und
-2	-1	11271	blondes	99% hat blondes [...] Haar und
-4	5	4477	blond	41% die Haare [...] blond gefärbt/färben

20

INSTITUT FÜR DEUTSCHE SPRACHE

BUHOFER FESTE WORTVERBINDUNGEN (2014) (PRINT)

Zahn m

ANWENDUNGSBEISPIELE

- **kräftige** | **saubere** | **gesunde** Zähne Ihre zersenden Zähne sind das Ergebnis täglicher Pflege. | **schöne** | **spitze** | **stumpfe** | **scharfe** (messerscharfe Zähne Der Hai hatte messerscharfe Zähne, die sich bei dem Angriff in das Schlagen) | **vorstehende** | **schiefe** Zähne Schiefe Zähne sind nicht immer nur unglücklich, sondern werden häufig durch Karies verursacht. ■ **weiße** | **gelbe** Zähne Der Kesseln von Nikotin verursacht gelbe Zähne. | **schwarze** Zähne Die Folgen von Karies sind schwarze Zähne. | **schlechte** Zähne | **lockerer** | **kranker** | **schmerzender** | **kaputt** | **hohler** | **tauler** Z. Wenn sie taubst, sieht man die faulen Zähne. ■ **die ersten** Zähne Mit sechs Jahren hat man nach die ersten Zähne. | **die zweiten** Zähne | **die eigenen** Zähne | **natürliche** | **echte** | **künstliche** | **falsche** Zähne Falsche Zähne haben den Vorteil, dass man sie zum Reinigen herausnehmen kann. | **die dritten** Zähne künstliche Zähne

VERBUNDEN

 - **gebisslos**, die ersten/zweiten Zähne kommen (ein Z. bricht ab) Mit etwa 6 Jahren brechen bei den Kindern die ersten bleibenden Zähne ab. | Zähne **lokalisieren** ■ **entfernen**

die Zähne. | die Zähne **entstehen** | mit den Zähnen **hinschauen** Wer im Schlaf mit den Zähnen knirscht, schlägt auf Dauer das Gebiss. | durch die Zähne **pfleifen** ■ **kaufen** mit den Zähnen nach etw. **schnappen** | etw. mit den Zähnen **zerbeissen** | mit den Zähnen an etw. **wagen** Aus Nervosität nagte er mit seinen Zähnen an der Unterlippe. | die Zähne **zusammenbeißen** 1. Vor Anstrengung bis der Gesichtshalter die Zähne zusammen. 2. Wenn man von dem haben | die Zähne in eine **graben** | **schlagen** Gierige Menschen schlagen mit die Zähne in das Fleisch Brod. **schlecken** Ihre **klitzten** Aus dem strahligen Gesicht blitzten weiße Zähne. | **schlecken** die Zähne **reinigen** Einmal jährlich werden die Zähne von der Zahnärztin/Ärztin intensiv gereinigt. In den Zähnen **stechern** Ich stecherte mit einer Nähnadel in den Zähnen, um die Heilkeime **herauszubekommen** | sich die Zähne **machan** lassen In Mundspiegel ist es besser, sich in Zahnspange **helfen** die Zähne **machen** lassen. ■ **ausziehen**, einen Z. **ziehen** | **ausziehen** | **reißen** Bei meinem Zahnarzt wurde der faule Zahn **gerissen**. | einen Z. **ausreißen** | **entfernen** ■ **entschlagen** Ja, ein Z. **auschlagen** | **entschlagen** Im Streit schlug der eine Mann dem anderen die Zähne ein.

25

INSTITUT FÜR DEUTSCHE SPRACHE

BUHOFER FESTE WORTVERBINDUNGEN (2014) (ONLINE)

KOLLOKATIONENWOERTERBUCH

Feste Wortverbindungen des Deutschen. Kollokationswörterbuch für den Alltag

Zahn

Bsp./Infos ganze Form

die **Zähne reinigen** Wegen der Conference Aktivitäten Projekt

Karies Schäden konnte sie ihre Zähne nicht einfach mit der Zahnbürste putzen, sondern musste sie speziell reinigen. tv

Zahn m (Stichwort nicht)

Adjektiv/Adverbien Verben

gelb - **schief** **graben** - **reinigen** - **schlagen** **Goldzahn** - **Löwenzahn**

Zahnarzt/Zahnärztin - Zahnradbahn - Zahnschmerzen

Zahn m

Adjektiv/Adverbien Verben

gelb - **schief** **graben** - **reinigen** - **schlagen** **Goldzahn** - **Löwenzahn**

Zahnarzt/Zahnärztin - Zahnradbahn - Zahnschmerzen

26

INSTITUT FÜR DEUTSCHE SPRACHE

VALENZ ALS ABSTRAKTE ZUSAMMENFASSUNG KOLLOKATIVEN VERHALTENS?

- Valenz beschreibt das systematische Zusammenspiel von Klassen von Wörtern
- semantisch motiviert, aber auch eingeschränkt
- syntaktisch realisiert
- ausgehend von dem, was wir haben, haben wir ein empirisches Kriterium entworfen, wie die Klassen von Wörtern als Gruppierungen von Kollokationen entstehen können
- unter allen Vorbehalten der Nicht-Berechenbarkeit und der Abhängigkeit von der Datengrundlage
- diese empirisch hervorgetretenen, lexikalisch gefüllten Klassen könnten als zusätzliche Angabe zu den Rollenbeschreibungen angeboten werden (evtl. mit Verweis auf die zugrundeliegenden Texte)

27

INSTITUT FÜR DEUTSCHE SPRACHE

GRUPPIERUNG ANHAND KOOKKURRENZ, POSITIONALER PRÄFERENZ UND PROFILÄHNLICHKEIT

Kookkurrenzanalyse

positionale Präferenz

-5 ← ... → -1

Flugzeug [...] abheben

Ähnlichkeit im Kookkurrenzprofil (im Gesamtkollokationsverhalten)

↑ ↓

Hubschrauber [...] abheben

-5 ← ... → -1

DP14 1s erstes reguläres **Flugzeug** von der Startbahn **abheb**, atmete die Luftfahrtr DP13 Das **Flugzeug** sollte eigentlich bereits 2012 **abheben**, doch die Genehmigung DP13 sagter: «Ich bin dagegen, dass dieses **Flugzeug** **abhat**, bevor nicht alle Frau MNR9 einen explosionsartigen Knall, als das **Flugzeug** **abgehoben** hatte, dann sahen s MNR4 agen. Sie lassen das kleinste **Flugzeug** der Welt **abheben**. "Caroli" heißt das 3

BAZ07 vralien. Bevor eine Crew mit einem **Hubschrauber** **abhat**, müssen auch sie mehr 898 d nicht. Wenn der **Hubschrauber** zu einem Einsatz **abhat**, haben wir uns entlang 800 nd die Bilder, wenn die übervollen **Hubschrauber** **abheben** und dutzende Menschen M01 hin soll, und dann der **Hubschrauber** noch einmal **abheben** muss. Hinzu kommt die M02 ter Zeit hatte der **Hubschrauber** vor ihren Augen **abgehoben** und war verschwunde

28

VALENZ DATENGETRIEBEN

INSTITUT FÜR DEUTSCHE SPRACHE

The screenshot shows a software window titled 'TreeView Classic' with a menu bar (File, Settings, Analysis, Export, Window, Help). The main area displays a complex hierarchical tree structure. A red highlight is visible on a branch in the lower right portion of the tree. On the right side of the window, a list of nodes is visible, including 'Konkurrenzprodukten [-1,-1]', 'Massenprodukten [-1,-1]', 'Konkurrenzen [-1,-1]', 'Konkurrenzen [-1,-1]', 'Wettbewerbern [-1,-1]', 'Anbietern [-1,-1]', and 'Billiganbietern [-1,-1]'. The number '33' is located in the bottom right corner.

VALENZ DATENGETRIEBEN

INSTITUT FÜR DEUTSCHE SPRACHE

The screenshot shows a software window titled 'TreeView Classic' with a menu bar (File, Settings, Analysis, Export, Window, Help). The main area displays a complex hierarchical tree structure. A red highlight is visible on a branch in the lower right portion of the tree. On the right side of the window, a list of nodes is visible, including 'Hubschrauber [-5,-1]', 'Paläste [-5,-1]', 'Flugzeug [-5,-1]', 'Flugzeuge [-5,-1]', 'Jakt [-5,-1]', 'Maschine [-5,-1]', and 'Maschinen [-5,-1]'. A second window is overlaid on top, showing a zoomed-in view of a branch with nodes 'Flughafen [-5,-1]' and 'Startbahn [-5,-1]'. The number '33' is located in the bottom right corner.

VALENZ DATENGETRIEBEN

INSTITUT FÜR DEUTSCHE SPRACHE

The screenshot shows a software window titled 'TreeView Classic' with a menu bar (File, Settings, Analysis, Export, Window, Help). The main area displays a complex hierarchical tree structure. A red highlight is visible on a branch in the lower right portion of the tree. On the right side of the window, a list of nodes is visible, including 'Daoh [-5,-1]' and 'Daohs [-5,-1]'. A smaller window titled 'leben.cdf' is open on the right, showing a list of nodes including 'Hörax [-1,-1]' and 'Teilekatheter [-2,-1]'. The number '36' is located in the bottom right corner.

VALENZ DATENGETRIEBEN

INSTITUT FÜR DEUTSCHE SPRACHE

The screenshot shows a software window titled 'TreeView Classic' with a menu bar (File, Settings, Analysis, Export, Window, Help). The main area displays a complex hierarchical tree structure. A red highlight is visible on a branch in the lower right portion of the tree. On the right side of the window, a list of nodes is visible, including 'markant [-5,-1]', 'deutlich [-5,-1]', 'deutlicher [-5,-1]', 'wertigkeit [-5,-1]', 'angenehm [-5,-1]', 'wohlwend [-5,-1]', 'schwarzen [-2,-2]', 'dunkeln [-5,-2]', 'grünen [-5,-2]', 'dunkeln [-5,-1]', 'optisch [-5,-1]', 'plastisch [-5,-1]', 'kontrastreich [-3,-1]', and 'leuchtend [-5,-1]'. The number '36' is located in the bottom right corner.