

# KORPUSLINGUISTIK UND LEXIKOGRAPHIE

# Digitale Korpora in Lexikographie und Phraseologie

- Das Arbeiten mit Korpora hat gerade in der Lexikographie eine lange Tradition. Selbst bei gegenwartssprachlichen Wörterbüchern würde sich kein Lexikographenteam anmaßen, eine vollständige lexikographische Beschreibung allein auf der Basis der eigenen Sprachkompetenz auszuarbeiten. Vielmehr exzerpieren und analysieren seriöse Wörterbuchprojekte Belege aus Quellentexten und konsultieren andere Wörterbücher als Sekundärquellen. Unumgänglich sind Korpora für die Beschreibung älterer Sprachstufen, für die zeitgenössische Lexikographen ja keine muttersprachliche Kompetenz mitbringen. Die Wörterbuchforschung hat die Prozesse der Erarbeitung von gedruckten Wörterbüchern sehr detailliert erfasst und beschrieben (vgl. Wiegand 1998).

# Digitale Korpora in Lexikographie und Phraseologie

- Die folgende stark vereinfachte Skizze der „prädigitalen“ Korpusnutzung soll dazu dienen, den qualitativen Sprung deutlich zu machen, der durch die Verfügbarkeit digitaler Korpora entsteht. Beim prädigitalen Vorgehen werden aus Quellenkorpora, die eine möglichst vielfältige und ausgewogene Auswahl von Texten zum jeweils relevanten Sprachausschnitt enthalten, Belegstellen exzerpiert und in einem Belegarchiv alphabetisch nach Stichwörtern geordnet. Diese Belegarchive sind in verschiedenen Phasen des lexikographischen Prozesses wichtig: bei der Entscheidung, welche Stichwörter ins Wörterbuch aufgenommen werden, bei der Bestimmung, wie viele semantische Lesarten für ein Stichwort angesetzt werden und bei der Formulierung der lexikographischen Angaben zu Form und Bedeutung. Manche Wörterbücher integrieren auch ausgewählte Belege in die Wörterbuchartikel.

# Digitale Korpora in Lexikographie und Phraseologie

- Die Vorteile digitaler Korpustechnologie für lexikographische Arbeitsprozesse liegen auf der Hand: (1) Aus digitalen Korpora kann man flexibel Trefferlisten generieren; die zeit- und kostenaufwändige Exzeption von Belegen und die „Verzettelung“ in prädigitale Belegzettellarchive entfällt. (2) Digital verwaltete Belege können quantitativ ausgewertet werden; insbesondere lassen sich Daten zur Frequenz und zum gemeinsamen Auftreten von Wortvorkommen (Kollokationen/Kookkurrenz) berechnen (vgl. Geyken 2004). Natürlich unterscheiden sich die Trefferlisten, die von einem Korpusrecherchesystem erzeugt werden, vom prädigitalen Zettelarchiv:

# Digitale Korpora in Lexikographie und Phraseologie

Wie im zweiten Abschnitt erläutert, operiert die automatische Suche in digitalen Korpora vornehmlich über Wortformen und formbasierten Suchmustern und nicht über Lexemen in einer bestimmten Bedeutung. Linguistische Annotationen können zwar die Präzision der Suchanfragen deutlich verbessern, dennoch enthalten die automatisch erzeugten Trefferlisten oft auch Pseudotreffer, die manuell aussortiert werden müssen. Diesen Beschränkungen zum Trotz bietet bereits die aktuelle Korpustechnologie einem methodisch und technisch kompetenten Lexikographen Optionen zur Recherche und Analyse, die in einem prädigitalen Zettelarchiv nicht oder nur mit sehr hohem Zeitaufwand möglich wären.

# Digitale Korpora in Lexikographie und Phraseologie

- Die Nutzung digitaler Korpora in der Lexikographie ist in der korpuslinguistischen Literatur bereits gut beschrieben. Ein englischer „Klassiker“ ist Sinclair (1991), einen aktuellen Überblick geben u. a. Lemnitzer/Zinsmeister (2006:143ff.), McEnery/Xiao/Tono (2006:80ff.) und Heid (2008). Digitale Korpusrecherchesysteme, in denen man sehr flexibel nach Wortkombinationen suchen kann, sind insbesondere für die Phraseologieforschung attraktiv. Interessante Ergebnisse aus korpusgestützten Projekten zu Idiomen und Kollokationen sind u. a. dokumentiert in Moon (1998) (für das Englische) und Fellbaum (2007) (für das Deutsche).

# Digitale Wörterbücher und Korpora

- Digitale Medien und das Internet verändern nicht nur die Prozesse der Wörterbuchherstellung, sondern auch die dabei entstehenden lexikographischen Produkte, die als Wörterbuchportale bzw. lexikalische Informationssysteme direkt im Internet angeboten werden (vgl. Engelberg/Lemnitzer 2009; Storrer 2010). Beim Aufbau digitaler Wörterbücher müssen sich die Wörterbuchmacher nicht mehr darum bemühen, möglichst viele Informationen auf einer Druckseite unterzubringen; die lexikographischen Angaben können deshalb übersichtlicher präsentiert und durch mehr Belegbeispielangaben angereichert werden (vgl. Storrer 2001).

# Digitale Wörterbücher und Korpora

- Von dieser Option, Wörterbuchartikel um Korpusbelege anzureichern, machen zwei digitale Wörterbuchportale zur deutschen Gegenwartssprache Gebrauch: Das eLexiko-Wörterbuch, das im Wörterbuchportal „OWID“ des Instituts für deutsche Sprache abrufbar ist [eLexiko-OWID], und das „Projekt deutscher Wortschatz“ der Universität Leipzig, das Wörterbuchartikel semi-automatisch aus digitalen Korpora und Wörterbüchern generiert [Deutscher-Wortschatz].

# Digitale Wörterbücher und Korpora

Beide Wörterbücher bieten außerdem Angaben zur Frequenz der Stichwörter und zu typischen Wortverbindungen (Kollokationen, Kookkurrenzen), die automatisch aus den zugrunde liegenden Korpusdaten generiert werden: Das eLexiko-Wörterbuch ordnet alle Stichwörter einer Frequenzschicht zu und verlinkt diese mit automatisch erzeugten Kookkurrenzprofilen. Das Projekt deutscher Wortschatz gibt zu jedem Stichwort an, wie häufig dieses im zugrunde liegenden Korpus belegt ist, und ordnet es einer Häufigkeitsklasse zu, die relativ zur Häufigkeit der hochfrequenten Wortform „der“ berechnet wird. Außerdem werden typische Kollokationspartner aufgelistet und als Netzgraph dargestellt.

# Digitale Wörterbücher und Korpora

Einen Schritt weiter gehen digitale lexikalische Informationssysteme: Sie integrieren Wörterbuch- und Korpusressourcen durch eine Nutzeroberfläche, mit der man sowohl in Wörterbüchern als auch in Korpora recherchieren kann. Für das Deutsche wird ein solches System für den DWDS entwickelt (Klein 2004; Geyken 2005); auf die Funktionalität dieses Systems beziehen sich auch die folgenden Fallbeispiele.

# Digitale Wörterbücher und Korpora

Zentral für den Umgang mit der DWDS-Nutzeroberfläche sind das Konzept der Sichten und das Konzept der Panels: Als 'Sicht' bezeichnet man eine Kombination von Ressourcen (Wörterbücher, Korpora, Statistikwerkzeuge), mit der ein Nutzer arbeiten kann. Jede Ressource wird in einem als 'Panel' bezeichneten Arbeitsfenster angezeigt, das bei Bedarf vergrößert werden kann. Wer auf der Startseite ein Suchwort, z. B. das Wort Ampel, eingibt, erhält die in Abb. 4 gezeigte Standardsicht mit der folgenden Panelkombination:

# Digitale Wörterbücher und Korpora

- (1) Das DWDS-Wörterbuch (Panel oben links) basiert inhaltlich auf dem „Wörterbuch der deutschen Gegenwartssprache“ [WDG], einem 6-bändigen Printwörterbuch, das von 1952 bis 1977 auf der Basis eines umfangreichen Quellkorpora erarbeitet wurde (vgl. Malige-Klappenbach 1986). Im Rahmen des DWDS-Projekts wurde dieses Wörterbuch digitalisiert, strukturell aufbereitet, durch vertonte Ausspracheangaben ergänzt und an die neue Rechtschreibung angepasst.

# Digitale Wörterbücher und Korpora

- (2) Das Etymologische Wörterbuch (Panel oben rechts) ist auf Informationen zur Wortgeschichte spezialisiert. Es basiert auf der zweiten Auflage des „Etymologischen Wörterbuchs des Deutschen“ [Etym-WB], das in den 80er Jahren von einer Lexikographengruppe unter der Leitung von Wolfgang Pfeifer erstellt und im Rahmen des DWDS-Projekts digital aufbereitet wurde.

Ampel Suche (in meinem Korpus)

Mein Korpus Suchhistorie Hilfe anzeigen

Aktuelle Sicht: bald-standard

DWDS-Wörterbuch

### Ampel

Grammatik: die; -, -n  
Herkunft/Gebrauch: *Latein*

- kleine Hängelampe  
/bildlich/

**Belege** | gesamte (Unter)lesart | Anzeige schließen

**Belege:**  
G. Hauptm. Quint [1,45] Die strahlende Sonne war seine Ampel

- über einer Straßenkreuzung aufgehängtes Verkehrssignal
- hängendes Gefäß für Blumen

Eintrag: Ampel Zusammensetzungen | Eintrag | Belege | Beispiele

Etymologisches Wörterbuch des Deutschen (nach Pfeifer)

### Ampel

Grammatik: f.  
Definition: (schalenförmige) Hängelampe.

Ahd. *amp(ul)ā* (10. Jh.), mhd. *ampulle*, *ampel*, mnd. *ap(p)ulle*, *appulle*, mnl. *ampulle*, *appulle*, nl. *ampul* bezeichnen ein 'kleines Gefäß für Öl oder andere Flüssigkeiten', meist zu geweihtem Gebrauch bestimmt, im Mittelalter vor allem die 'Ewige Lampe' in der Kirche (ein von der Decke herabhängendes, fortwährend brennendes Öllämpchen), vom 14. Jh. an auch die Lichtquelle im Hause, seit dem 16. Jh. mit *Lampe* (s. d.) konkurrierend. Eine Neubelebung erfährt *Ampel* als Kurzwort für *Verkehrsampele* 'Lichtanlage (anfangs hängend über einer Kreuzung angebracht) zur Regelung des Straßenverkehrs' (20er Jahre 20. Jh.). Es handelt sich um eine Entlehnung aus lat. *ampulla* (eigentl. *\*amp(ul)hulla* > *dupl. \*ampulla* > *\*ampulla*) 'kelchförmiges Gefäß mit zwei EtmWB - V21-01

DWDS-Kernkorpus

Treffer: 176

1 2 3 4 ...

- ...en, die absurdenweise vor **Ampeln** auf Rot anhalten, Komment...
- ...umwinkender Polizist; die **Ampel** ist ausgefallen. Raine...
- ...interherfährt, immer eine **Ampel** nach ihnen, tunlichst dar...
- ... und sauste über die rote **Ampel**, ich mußte halten und ve...
- ...erten eine Tasse und eine **Ampel**. Nicht besonders attr...
- ...tete die Agentur Grey die **Ampel** dann auf Grün. Eine ju...
- ...ruppe hineinfuhr, als die **Ampel** für ihn grünes Licht zeig...
- ...f der Gegenseite, weil die **Ampel** Rot zeigte. Vielleicht...
- ...tz, überfahren dabei rote **Ampeln** und blockierten den Verke...
- ...Gepäck. An jeder roten **Ampel**, wenn das Taxi hielt, st...

OpenThesaurus

synonyme Wortgruppen für: **Ampel**

**Ampel**, Lichtsignalanlage, Lichtzeichenanlage, Verkehrsampele [ Technik ]

Sicht laden

Neue Sicht

Panel

# Digitale Wörterbücher und Korpora

- Der OpenThesaurus (Panel unten rechts) listet Synonyme und sinnverwandte Wörter. Die Einträge des von Daniel Naber initiierten kollaborativen Wörterbuchprojekts sind im DWDS-System als externe Ressource eingebunden.

# Digitale Wörterbücher und Korpora

- Im DWDS-Kernkorpus (Panel unten links) kann man gezielt nach Suchwörtern und Suchmustern recherchieren. Wegen seiner ausgewogenen Streuung über die Dekaden des 20. Jahrhunderts und über Textsortenbereiche hinweg eignet sich dieses Korpus sehr gut dazu, Entwicklungen und Veränderungen im Wortschatz des 20. Jahrhunderts zu untersuchen.

# Digitale Wörterbücher und Korpora

- Neben dieser Standardsicht bietet das DWDS-System weitere vordefinierte Sichten mit Kombinationen von Korpora, Korpusstatistiken und Wörterbüchern an. In den folgenden Beispielen verwenden wir zusätzlich zur Standardsicht das Zeitungskorpus der ZEIT, die Wortverlaufstatistik zum DWDS-Kernkorpus (vgl. Abb. 5) und das Statistikwerkzeug 'Wortprofil' (vgl. Abb. 6); diese und weitere Ressourcen kann man im DWDS-System in Panels dazuschalten.

# Digitale Wörterbücher und Korpora

- Registrierte Nutzer können Panelkombinationen dauerhaft als nutzerspezifische Sichten speichern. Der Aufwand für die kostenlose Registrierung lohnt sich nicht nur, weil die Definition eigener Sichten für linguistische Untersuchungsfragen oft die effizienteste Option ist, sondern weil registrierte Nutzer zudem die Möglichkeit haben, eigene Belegsammlungen anzulegen, nach Kategorien zu klassifizieren und in einer späteren Sitzung unter dem Menüpunkt 'Mein Korpus' wieder abzurufen. Für den Einstieg in die korpusgestützte Sprachanalyse stehen damit rudimentäre Funktionen eines lexikographischen Arbeitsplatzes direkt online zur Verfügung. Die folgenden einfachen Fallbeispiele sollen illustrieren, wie die Ressourcenkombination für korpusgestützte Untersuchungen zum deutschen Wortschatz genutzt werden kann.

# Frequenzinformationen und Frequenzverläufe: Analysebeispiel Streß/Stress

- Früher hatte man weniger Stress! Ob diese oft gehörte Behauptung stimmt, kann man sicher nicht durch eine Korpusanalyse klären. Allerdings zeigt die Recherche im DWDS-Kernkorpus, dass das Suchwort Streß erst seit den 70er Jahren belegt ist. Am automatisch generierten Frequenzverlaufdiagramm, das zu den 86 Treffern im DWDS-Korpus auf der Basis der Metadaten erstellt wird (vgl. Abb. 5), lässt sich weiterhin ablesen, dass das Wort zunächst überwiegend in Gebrauchstexten und wissenschaftlicher Fachliteratur vorkommt, ab den 90er Jahren aber zunehmend auch in der Belletristik und in Zeitungstexten verwendet wird. Bei der relativ geringen Treffermenge muss man diese Verteilung über die Textsortenbereiche hinweg sehr vorsichtig bewerten.

# Frequenzinformationen und Frequenzverläufe:

## Analysebeispiel Streß/Stress

Wenn man die überschaubare Trefferliste intellektuell analysiert, kann man jedoch sehr gut erkennen, wie sich das aus der Fachsprache der Psychologie stammende englische Lehnwort auch in nicht-fachsprachlichen Kontexten etabliert hat und wie sich dabei neue alltagssprachliche Formulierungsmuster und Kollokationen ausgebildet haben (z. B. Streß haben/machen, in Streß kommen/geraten, voll/total im Streß sein, etwas artetin Streß aus etc.).

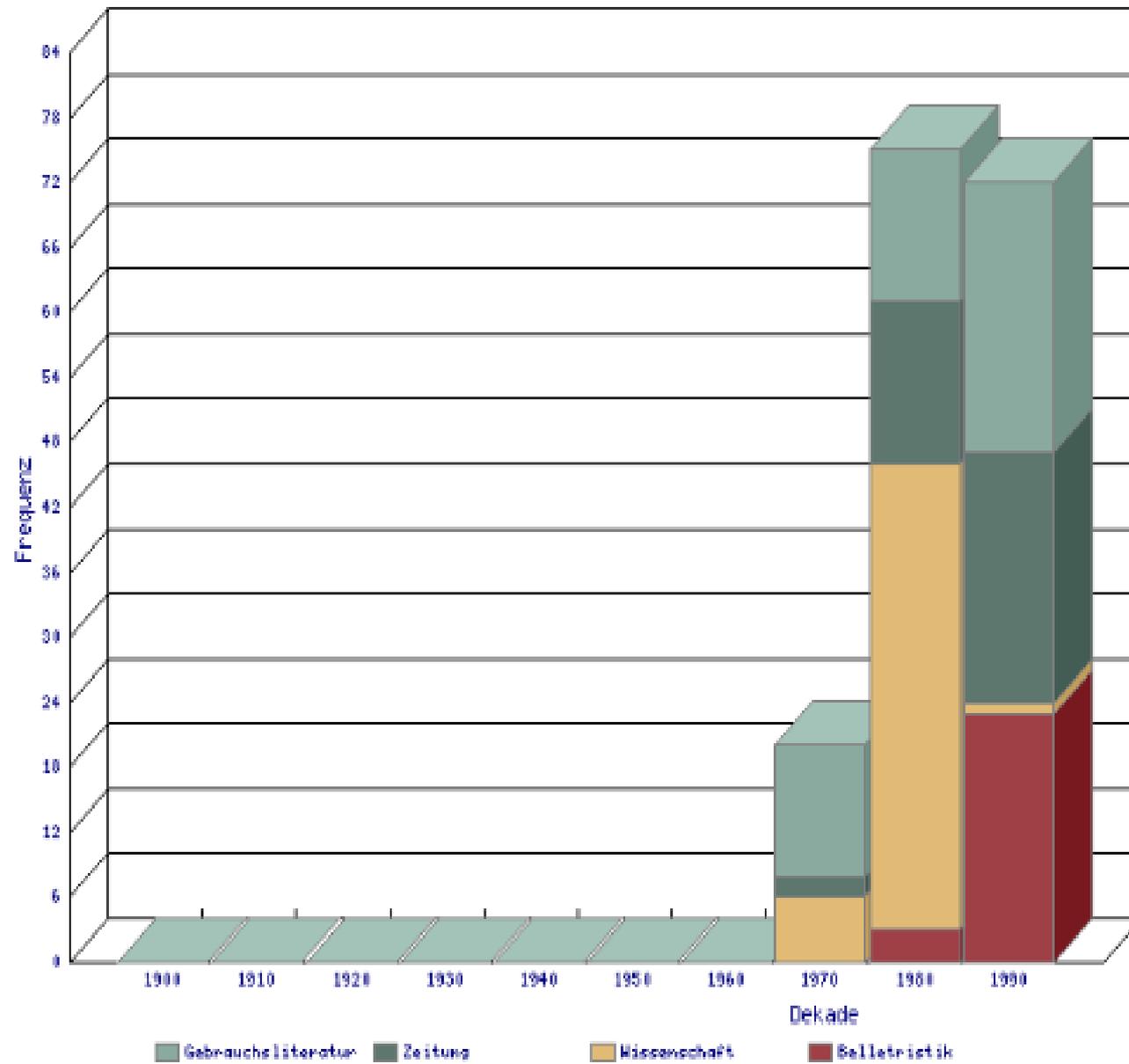
# Frequenzinformationen und Frequenzverläufe: Analysebeispiel Streß/Stress

Die Trefferliste zur Anfrage Streß enthält keinen Pseudotreffer; listet aber nicht alle relevanten Belege für das Lexem, denn dieses kommt auch in der Schreibvariante Stress vor, also in der regelkonformen Schreibvariante nach der Orthographiereform. Mit der Abfrage „Streß || Stress“ (vgl. Abschnitt 2.2) kann man nach beiden Varianten suchen, die Trefferzahl auf 106 erhöhen und eine interessante Beobachtung zur Verteilung der beiden Schreibvarianten machen:

# Frequenzinformationen und Frequenzverläufe: Analysebeispiel Streß/Stress

- Die meisten Belege zur Schreibvariante Stress stammen aus den Jahren 1971–1976; zunächst wurde also die englische Schreibform auch im Deutschen verwendet. Danach überwiegt die Schreibvariante Streß, die bis zur Orthographiereform 1998 regelkonform war; diese Schreibung ist auch noch in Texten belegt, die in den 90er Jahren, also nach der Orthographiereform, erschienen sind. Das Beispiel zeigt generell, dass es für eine vollständige Trefferausbeute im Kernkorpus oft erforderlich ist, alle im 20. Jahrhundert zulässigen Schreibvarianten zu kombinieren; im DWDS-Wörterbuch sind die vor und nach der Reform zulässigen Varianten zu allen Stichwörtern verzeichnet.
- Abb. 5: Frequenzverlaufdiagramm zum Suchwort Streß im DWDS-System

Verlaufstatistik für 'Streß' im DWDS-Kerncorpus



# Frequenzinformationen und Frequenzverläufe: Analysebeispiel Streß/Stress

- Am Beispielwort Streß/Stress kann man auch sehen, dass sich das DWDS-Korpus wegen seiner ausgewogenen Textauswahl zwar sehr dazu eignet, Sprachwandelprozesse im 20. Jahrhundert zu verfolgen, dass man die Datierung der Erstbelege aber vorsichtig interpretieren muss. Die beiden ersten Belege für die Varianten Streß und Stress im Kernkorpus stammen aus dem Jahr 1971. Das Etymologische Wörterbuch datiert die Übernahme des Lehnworts Stress aber bereits auf die 50er Jahre. Tatsächlich finden sich im Zeitungskorpus der ZEIT, das im DWDS-System als Panel hinzugefügt werden kann, sieben Belege aus dem Jahre 1958 und über dreißig weitere Belege aus Jahrgängen zwischen 1960 und 1970, die meisten davon in Artikeln zu medizinischen, biologischen oder psychologischen Themen. Es empfiehlt sich also, im Zweifelsfall den Datierungen im Etymologischen Wörterbuch zu vertrauen, zumindest solange man keine früheren Belege in den Korpora findet.

# Bedeutungsentwicklung und Bedeutungsveränderung: Analysebeispiel Ampel

- In Abschnitt 2.3 wurde erläutert, dass aktuell verfügbare große Korpora nicht semantisch annotiert sind, weshalb man nicht direkt nach speziellen semantischen Lesarten für ein Lexem suchen kann. Aus diesem Grund lassen sich auch Frequenzen in Korpora nur „semantisch blind“ berechnen, was dazu führt, dass auch die automatisch generierten Frequenzverlaufsdigramme nicht zwischen verschiedenen semantischen Lesarten einer Wortform differenzieren. Wenn man zum Suchwort Ampel ein Frequenzverlaufsdigramm erzeugt, kann man zwar erkennen, dass die Wortform über das gesamte 20. Jahrhundert vor allem in der Belletristik und in der Gebrauchsliteratur belegt ist. Die formorientierte Frequenzzählung gibt aber keine Hinweise auf Verschiebungen in der Gebräuchlichkeit der drei semantischen Lesarten, die im DWDS-Wörterbuch zu diesem Stichwort verzeichnet sind: (1) 'Hängelampe', (2) 'Verkehrssignal', (3) 'Blumengefäß' (vgl. Abb. 4, Panel links oben).

# Bedeutungsentwicklung und Bedeutungsveränderung: Analysebeispiel Ampel

- Erst die intellektuelle Durchsicht der 176 Treffer zum Suchwort Ampel im DWDS-Kernkorpus bringt zum Vorschein, dass in der ersten Hälfte des 20. Jahrhunderts die Lesart (1) ‚Lampe‘ dominiert, während sich in der zweiten Hälfte fast nur noch Belege für die Lesart (2) ‚Verkehrssignal‘ finden. Die frühesten Belege im Kernkorpus für die Lesart ‚Verkehrssignal‘ stammen aus einem Text von Kurt Tucholsky aus dem Jahre 1933. Dieser Text enthält drei von vier Belegen für diese Lesart, die in den insgesamt hundert Treffern des Zeitabschnitts 1900–1956 zu finden sind; der vierte stammt aus dem Jahre 1951.

# Bedeutungsentwicklung und Bedeutungsveränderung: Analysebeispiel Ampel

- Die restlichen 96 Treffer vor 1956 belegen ganz überwiegend die Lesart ‚Lampe‘, einige die Lesart ‚Blumengefäß‘. Nach 1956 verändert sich die Verteilung sehr rasch: In den insgesamt 76 Treffern aus dem Zeitraum 1956–1999 ist die Lesart ‚Lampe‘ nur noch zweimal belegt (1964 und 1977), die restlichen 74 Treffer belegen nur noch die Lesart ‚Verkehrssignal‘. Ein aktuelles Wörterbuch würde auf dieser Grundlage vermutlich die Lesart ‚Lampe‘ als ungebräuchlich markieren, um Missverständnisse bei der Textproduktion von Nicht-Muttersprachlern zu vermeiden.

# Bedeutungsentwicklung und Bedeutungsveränderung: Analysebeispiel Ampel

- Schwieriger ist es, Aussagen über die Bedeutung (3) (= Ampel als 'Blumengefäß') zu treffen, die sich vermutlich als Kurzform aus dem Kompositum Blumenampel entwickelt hat. Zwar kann man nachweisen, dass diese Lesart schon in den ersten Dekaden des 20. Jahrhunderts bekannt war; allerdings ist die Beleglage auch hier sehr dünn: Man findet insgesamt nur 16 Belege, acht davon stammen aus demselben Text (Paul Scheerbarts „Lesabéndio“). Nach 1956 ist die Bedeutung ‚Blumengefäß‘ im Kernkorpus nicht mehr belegt; wegen der niedrigen Frequenz in der ersten Jahrhunderthälfte sollte man daraus aber keinesfalls ableiten, dass diese Lesart nicht mehr gebräuchlich ist.

# Bedeutungsentwicklung und Bedeutungsveränderung: Analysebeispiel Ampel

- Generell sollte der Befund, dass eine Wortform oder eine Lesart im Korpus nicht belegt ist, nicht als Nachweis dafür interpretiert werden, dass die betreffende Lesart oder Wortform in der untersuchten Zeit noch nicht existiert hat. Die Neubedeutung von Ampel als Bezeichnung einer Koalition bundesdeutscher Parteien war schon in den 90er Jahren bekannt, auch wenn sie im DWDS-Kernkorpus nicht belegt ist. Auch in diesem Fall lohnt es sich, ergänzend im laufend aktualisierten Zeitungskorpus der ZEIT zu recherchieren: Dort findet man mehrere Belege aus den 90er Jahren, der früheste stammt aus dem Jahr 1991.

# Bedeutungsentwicklung und Bedeutungsveränderung: Analysebeispiel Ampel

Das Kompositum Ampelkoalition, aus dem die Neubedeutung vermutlich durch Kurzwortbildung entstanden ist, wird in diesem Korpus erstmals 1988 verwendet. Durch die Analyse der umfangreichen Trefferliste im ZEIT-Korpus kann man verfolgen, wie sich die Neubedeutung semantisch ausdifferenziert (schwarze Ampel, Schwampel) und auch immer häufiger gebraucht wird: Im Jahrgang 2009 des ZEIT-Korpus aktualisieren bereits 79 der 143 Treffer die Neubedeutung ‚Ampelkoalition‘; die übrigen Treffer 64 belegen die Lesart ‚Verkehrssignal‘; kein einziger Beleg findet sich für die Lesarten ‚Lampe‘ oder ‚Blumengefäß‘.

# Typische Umgebungen/Kollokationen: Analysebeispiel zeitigen

- Deutsche Muttersprachler haben meist ein gutes Gefühl dafür, welche Lexeme sich miteinander kombinieren lassen. Beispielweise wissen sie, dass man einen Brand legen und etwas in Brand setzen kann, dass aber die Verbindungen einen Brand setzen oder in Brand legen aber ungebräuchlich sind. Deutschlerner müssen solche kombinatorischen Präferenzen, man spricht auch von 'Kollokationen' oder 'Kookkurrenzen', die einzelsprachspezifisch sind und sich auch nicht aus der Bedeutung der kombinierten Lexeme ableiten lassen, oft im Wörterbuch nachschlagen.

# Typische Umgebungen/Kollokationen: Analysebeispiel zeitigen

- Korpuslinguistik und Lexikographie experimentieren seit längerem mit statistischen Verfahren, um Kollokationen bzw. Kookkurrenzen aus Korpusdaten zu gewinnen und für die lexikographische Sprachbeschreibung nutzbar zu machen (vgl. Lemnitzer/Zinsmeister 2006:145ff.; McEnery/Xiao/Tono 2006:208ff.). Auf solchen statistischen Verfahren basiert auch das Wortprofil im DWDS-System, das Kollokationen aus dem DWDS-Kernkorpus und dem ZEIT Korpus ermittelt nach

relevanteste syntaktische Relationen für **zeitigen** als **VVF**IN Frequenz: 518

Erfolg Ergebnis **Ergebnis** Folge Folge **Frucht** Wirkung

Belege für **zeitigen** als **VVF**IN in **v\_obj\_complement** mit **Frucht** als NN

Treffer: 11 [Anzeige schließen](#)

1	Wissenschaft   1955-12-31	Das Resultat dieser Nachahmung war, daß die nationalen Befreiungsbewegungen merkwürdigerweise meist mit einer philologischen Renaissance begannen, die oft wunderbare <b>Früchte zeitigte</b> , da sie ja im Dienste politischer Interessen stand, aber im ganzen doch sehr fruchtbar war.
2	Belletristik   1910-12-31	Mag sich die allzu große Willfährigkeit und Leichtgläubigkeit der edlen Dame in Sachen der Religion einmal auf diese Weise ein wenig rächen, und mag sie zur Erkenntnis gelangen, daß das von ihr geförderte Laienwesen in Sachen der Religion manchmal auch solche <b>Früchte zeitigt</b> .
3	Belletristik : Roman   1933-12-31	Nach einem halben Jahrhundert noch <b>zeitigte</b> die Energie Awetis Bagradiaus des Alten hier volle <b>Frucht</b> , die Liebe eines einzigen unternehmenden Mannes, die sich stürmisch auf diesen Heimattleck Erde konzentriert hatte, aller Weltlockung zum Trotz.
4	Wissenschaft   1909-12-31	Diese psychische Vergewaltigung <b>zeitigte</b> ihre <b>Früchte</b> ; in schandbarer Verhehlung ist die mißhandelte Erotik wiedergekehrt: als Zote.
5	Zeitung : POLEN.   1956-10-27	Die harte Anstrengung der Arbeiterklasse und der gesamten Nation <b>zeitigte</b> nicht die erwarteten <b>Früchte</b> .
6	Wissenschaft : Geschichte   1936-12-31	Die Zeitschriften sind voll von Fundchroniken und Fundberichten; aber es fehlt hier an einem höheren Plan, welcher die Übersicht erleichtert, und so <b>zeitigt</b> dieser Aufwand nur geringe <b>Früchte</b> .
7	Wissenschaft   1971-12-31	Diese vielgestaltigen Evangelisationsbemühungen <b>zeitigten</b> nur langsam <b>Früchte</b> .
8	Zeitung : Morgen-Ausgabe   1922-03-11	Ich weiß aus Erfahrung, daß gerade eine derartige persönliche Einwirkung von Mensch zu Mensch oft überraschend gute <b>Früchte zeitigt</b> .
9	Zeitung : Morgen-Ausgabe   1922-03-08	Auf der Grundlage seiner medizinischen Bildung und seiner philosophischen Veranlagung <b>zeitigte</b> seine dichterische Begabung mancherlei literarische <b>Früchte</b> , insbesondere über die Mechanik seelischer Vorgänge: so die Abhandlung "Von der Seele" und das vielgelesene Werk.
10	Gebrauchsliteratur : Landwirtschaft   1918-12-31	Ein Baum, der eine übergroße Menge <b>Früchte</b> auf einmal <b>zeitigt</b> , erschöpft sich auch auf dem besten Boden für einige Jahre und ist danach, bis er sich wieder gekräftigt hat, sehr empfindlich.
11	Gebrauchsliteratur   1925-12-31	Als daher später der Adel nicht in der Lage war, diese Umformung des Staatsorganismus zu erkennen, sondern unvernünftige Herrscher nur desto schlimmer hausten, ging diese Drachensaat blutig auf und <b>zeitigte</b> fürchterliche <b>Früchte</b> .

# Typische Umgebungen/Kollokationen: Analysebeispiel zeitigen

- Was man aus einem solchen Profil entnehmen kann, möchte ich am Beispiel des Wortprofils für das Verb zeitigen illustrieren (vgl. Abb. 6). Im DWDS-Wörterbuch wird die Hauptbedeutung von zeitigen mit dem Synonym hervorbringen beschrieben. Gerade weil diese Bedeutung der gehobenen Stilschicht zugeordnet ist, kann man sich vorstellen, dass auch muttersprachliche Schreiber unsicher sind, welche Nomina bei diesem Verb als Subjekt bzw. Akkusativkomplement in

# Typische Umgebungen/Kollokationen: Analysebeispiel zeitigen

- Im Gegensatz zu ähnlichen Funktionen in anderen digitalen Wörterbüchern, z. B. den sehr ausführlichen Kollokationsinformationen in den Artikeln des Projekts deutscher Wortschatz [Deutscher-Wortschatz] der Universität Leipzig, sind die Wortprofile des DWDS-Systems mit den zugrunde liegenden Korpus Treffern und ihren Metadaten verknüpft. Abb. 6 unten zeigt beispielsweise die Trefferliste zum Kollokationspartner Frucht als Akkusativkomplement. Die

# Fazit und Ausblick

- Linguistisch aufbereitete digitale Korpora bieten vielfältige Möglichkeiten, authentische Sprachdaten quantitativ und qualitativ zu analysieren. Die einfachen Analysebeispiele in Abschnitt 4 geben hier nur einen ersten Einblick, was man bereits ohne computerlexikographische Ausbildung in Online-Korpora entdecken kann. Die einfachen Beispiele dürften aber auch bereits deutlich gemacht haben, dass die Korpusdaten umsichtig interpretiert werden müssen, dass also digitale

# Fazit und Ausblick

- – Welche Lexeme werden als Stichwörter aufgenommen?
- – Wie viele Lesarten setzt man für ein Stichwort an?
- – Was sind typische und was sind ungewöhnliche Verwendungskontexte?  
müssen auf der Basis der sachkundigen Auswertung von Korpusdaten beantwortet werden. Sachkundig bedeutet einerseits, dass man den Quellenwert der Korpusbelege richtig einzuschätzen weiß:

# Fazit und Ausblick

- Die linguistische Aufbereitung (Lemmatisierung, Wortartenannotation etc.) in großen digitalen Korpora erfolgt mit automatischen Verfahren und ist deshalb nicht fehlerfrei. Wer häufiger mit Korpora arbeitet, wird allerdings schnell Strategien entwickeln, mit denen sich die Menge der Pseudotreffer reduzieren lässt. Die Korpuslinguistik arbeitet an der Verbesserung der Verfahren und an Werkzeugen, mit denen man sehr große Treffermengen für hochfrequente Wörter

# Fazit und Ausblick

- Neue computerlexikographische Funktionen – z. B. die Suche nach interessanten und ungewöhnlichen Belegen, die Entdeckung und Entwicklung von Neubedeutungen, das Aufspüren von Metaphern – sollten im Dialog zwischen Lexikographie und Korpuslinguistik entwickelt und in konkreten Wörterbuchprojekten evaluiert werden. Die korpusgestützte Lexikographie wird damit sicherlich in nächster Zeit ein sehr spannendes interdisziplinäres