

Konstituierung des Textes und Ausdruckssyntax

- Harald Schweizer -

Dienstags, 17-19 Uhr - Hörsaal 2 (Sand) (Vorverlegung nach Semesterbeginn auf 15 Uhr wird angestrebt)

Unter: <http://www-ct.informatik.uni-tuebingen.de/index.htm> → LEHRE → LEHRANGEBOT ist dieses file (Gliederung/Literaturliste/Materialien) zugänglich

Fakultät für Informations- und Kognitionswissenschaften
Arbeitsbereich Textwissenschaft

Sand 13
72076 Tübingen
harald.schweizer@uni-tuebingen.de
Fon: 29-75248
Fax: 29-5060

Sprechstunden: Mi 11-13 (bzw. nach Vereinbarung)

1. Hinführung / Rekapitulierung

- 1.1 Kurzerläuterung der einzelnen Aspekte
- 1.2 Ort der Fragestellungen im gesamten Interpretationskonzept
- 1.3 Dichotomie in der Zeichendefinition: signifiant vs signifié
- 1.4 Hermeneutische Spirale
- 1.5 Das SLANG-Konzept - Lernende, nicht-normative Grammatik = Search for a Learning Non-normative Grammar

A. KONSTITUIERUNG DES TEXTES

2. Texterfassung / Maschinenlesbarkeit
 - 2.1 Verschiedene Modalitäten der Eingabe
 - 2.2 Speicherung des Textes
 - 2.21 Seitenbeschreibung/-format - Text Encoding Initiative
 - 2.22 Zeichensätze/Sonderzeichen/SGML
 - 2.3 Korrekturprozeduren am eigenen Text
3. Beurteilung des fremden Textes
 - 3.1 Ausgangspunkt: Kritische Aneignung des Endtextes (≠Textkritik, Handschriften-/Druckausgabenkunde, Editionsphilologie), Wahl der weiter interessierenden Textfassung
 - 3.11 Einzelnes Manuskript / einzelner Druck
 - 3.12 Kritische Edition (vs. Faksimile)
 - 3.13 Vergleich verschiedener Fassungen → begründete Wahl
 - 3.14 Textstammbaum, Textfamilien
 - 3.15 Echtheitsüberprüfung
 - 3.2 Aufbereitung des Textes
 - 3.21 Übersetzungsbegründung
 - 3.22 Textsegmentierung
 - 3.23 Transkription / Präsentation
 - 3.3 Textbildungsprozeß: Welche synchrone Schicht - jede inzwischen textkritisch überprüft - soll der späteren Beschreibung / Interpretation zugrundegelegt werden?
 - 3.31 Probleme auch noch in textkritisch geklärten Texten?
 - 3.32 Soziolinguistische Informationen
 - 3.4 Theorie und Praxis der Methode zur Identifizierung von Zusätzen (Literarkritik)
 - 3.5 Übersicht über gängige Motive und Techniken von Interpolatoren/Redaktoren
 - 3.6 Präsentation des Ergebnisses für die folgenden Interpretationsschritte

B. INTERPRETATION

4. (Ausdrucks-) SYNTAX
 - 4.1 Zeichentheorie - Distributionsanalyse - Computer
 - 4.11 Zeichentheorie
 - 4.12 Stichwort: "Syntax"
 - 4.13 Stichwort: "Distribution"
 - 4.2 Was Grammatiken unter "Morphologie" verstehen
 - 4.21 Inkonsistenzen
 - 4.22 Suche nach Grundlagen
 - 4.3 Neuansatz: Inhaltsfreie Morphologie -
 - 4.31 Kursorische Beobachtungen
 - 4.32 Ausdrucksbasierte Formenlehre (Morphologie)
 - 4.33 Automatische Analyse
 - 4.331 Überlegungen zum Algorithmus
 - 4.332 Eigener Ansatz (Rekursion)
 - 4.333 Häufigkeit und Wortlänge
 - 4.34 Neologismen
 - 4.35 Dialekte
 - 4.4 Einzeltextbezogen: Algorithmen, Darstellung und Auswertung einer *textinternen* Befunderhebung
 - 4.41 Beispiele, Motivation
 - 4.42 Worthäufigkeit / Distribution
 - 4.43 Wortketten / *distant-reading*

- 4.44 Autoridentifikation
- 4.45 Sprachbarrieren
- 4.46 Bewertung der Daten / Ausblick / Ähnlichkeit
- 4.47 Übersetzung ohne Bedeutungswissen?
- 4.5 Einzeltextbezogen: Automatische Erhebung *textexterner* Befunde
- 4.51 Phrasendreschmaschine
- 4.52 Intertextuelle Befunde
- 4.53 Algorithmus
- 4.54 Relative Chronologie
- 4.55 Interpretation
- 4.6 Bewußtheitsgrade bei der Textrezeption und ausdrucks syntaktische Realisierungen (Zitate, Anspielungen, Gruppensprache, Formeln, vergleichbare Konstruktionen, relative Chronologie usw.)
- 4.61 Begriffsbildung
- 4.62 Feste Wortketten
- 4.63 Wortnachbarschaften und ihr Ertrag
- 4.7 Phonetik/Prosodie (ausgeklammert)
- 4.8 Sprachfamilien - ausdrucksformal erhoben
- 5.0 Literatur zum Gesamtbereich

zu Ziff. 0.0: vgt0.0 vsem0.0 vprag0.001 vsyst0.0

Einschlägigkeit der Vorlesung

Die Vorlesung kann - alternativ - in drei Formen im Rahmen des Studiums angerechnet werden.

1. Für "Textwissenschaft" im Rahmen der "Praktischen Informatik" (Diplom oder BA).
2. Für "Medieninformatik".
3. Für "Informatik und Gesellschaft" bei Diplom; seit WS 09/10 **nicht mehr** als "Schlüsselqualifikation" bei Bachelor / Master
Vgl. im Internet:

<http://www-ct.informatik.uni-tuebingen.de/sd/index.htm>

Im letzteren Fall gilt: Am Ende des Semesters wird in einer mündlichen Prüfung (15 min.) der Vorlesungsstoff, der an einer Stelle vertieft wurde, geprüft.

Bitte generell folgendes Verfahren beachten:

- Pro 2 SWS wird *ein* Unterthema von Ihnen ausgewählt und anhand von Literatur vertieft.
- Sie wählen aus den Literaturangaben (in die Materialien integriert) einen oder mehrere Titel, die ein Thema Ihres Interesses behandeln. Umfang im Fall "I&G": nicht unter 40 Seiten, im Fall "Praktische Informatik" laut Absprache. Es dürfen auch andere, zum Thema passende Literaturen genannt werden - sofern sie wissenschaftlichen Charakter haben.
- Unter "Spezialthema" kann z.B. ein Unterpunkt eines der Kapitel verstanden werden, auch eine Fragestellung, die von einem Zitat behandelt wird. Für die Prüfung **genügt es nicht**, als "Literatur" nur auf Vorlesungsmaterialien zu verweisen. Es **genügt auch nicht**, als Thema nur das zu wählen, was man als Vortrag im Rahmen der "Übung" schon einmal geboten hatte.
- Bevor Sie sich festlegen: vergewissern Sie sich in punkto Beschaffbarkeit und Lesbarkeit/Verstehbarkeit der anvisierten Literatur.
- Standard ist: Sie kommen mit einer Vorauswahl in die Sprechstunde und lassen sich vor der definitiven Festlegung beraten.
- Geben Sie spätestens 1 Woche vor der vereinbarten Prüfung ein Papier ab (bzw. email), das Ihren Namen, Geburtstag, Typ der Prüfung und die ausgesuchte Schwerpunktliteratur (in exakter Zitierung, einschließlich Bibliothekssignatur) enthält.

Sie können sich darauf verlassen, daß der Schwerpunktstoff auf jeden Fall Gegenstand der Prüfung sein wird - neben dem sonstigen Stoff (in dem - für sich genommen - wenigstens ein "ausreichend" erzielt werden soll. **Nur mit Spezialthemen kann man die Prüfung nicht bestehen**).

Hinweis zum Literaturverzeichnis: Es ist nach der Vorlesungsgliederung geordnet. Steht am Schluss einer Literaturangabe "[+Be-zifferung]", so verweist dies auf den *Standort in der Informatik-Bibliothek*. Im Übrigen verarbeiten die Materialien oft Literatur, die über das Lit-Verzeichnis hinausgeht (auch sie kann herangezogen werden - sofern sie wissenschaftlichen Charakter hat).

Hinweis zur Zählung der Materialien: Sie folgt der Bezifferung der Vorlesungskapitel und -abschnitte. Es kann beim Hochzählen zu Sprüngen kommen. Solche sind beabsichtigt, stellen also keinen Fehler dar. Auch gibt es nicht zwangsläufig zu jedem Unterpunkt ein Material - die mündliche Erläuterung braucht ja auch noch Freiraum...

Die Materialien stellen **kein Vorlesungsskript** dar. Aber anhand der Stoffe wird in Zustimmung oder Kritik bzw. flankierender Erläuterung die Linie der Vorlesung entwickelt. **Markieren, unterstreichen, ergänzen Sie nach Herzenslust! Nur wer so mit den Materialien arbeitet, ist auch innerlich bei der Sache. Dafür wurde auch immer wieder Leerraum gelassen. Bloßes Zuhören ist die ineffektivste Form des Wissenserwerbs.**

In ungefilterter Vollversion sind die Materialien unter der homepage zugänglich:

<http://www-ct.informatik.uni-tuebingen.de/ct/lehrangebot.htm>

In der Vorlesung können nur Teile daraus ausgeteilt und behandelt werden.

Eine **ausführliche Lektüre** der Materialien ist im Vorlesungsrahmen nicht möglich. Sie wird aber **nachdrücklich dem Selbststudium empfohlen.**

Tipps zur Prüfungsvorbereitung:

Der Stoff ist umfangreich und differenziert. Man möchte folglich hinter den vielen Bäumen den Wald wieder erkennen. Auf unserer homepage - oben schon erwähnt -, Stichwort "FORSCHUNG", findet sich zu den methodischen Ebenen eine kleine Auswahl von Materialien: an ihnen kann positiv und gerafft abgelesen werden, wie wir die jeweilige Methode verstehen.

Zweite Möglichkeit um komprimiert nachzulesen, was wir zu den einzelnen Punkten für richtig halten:

<http://www.alternativ-grammatik.de>

Viele Vorlesungsstoffe sind dort im Blick auf "Schule" vereinfacht zusammengefasst [die theoretischen Hintergrunddebatten fehlen].

Bald nach Semesteranfang wird die Vorlesung eher im Stil eines **Kolloquiums** durchgeführt: Diskussionen und gemeinsames Beitragen von Beispielen aktivieren die Aufmerksamkeit.

Zwischen den Doppelstunden: nur kurze Pause. Dafür ist früher Schluss.

zu Ziff. 0.01:

Zu unserer Aversion gegen Powerpoint-Präsentationen

zu: vprag0.01 vsem0.01 vsyst0.01

aus: SPIEGEL-ONLINE (17.6.2009): Frank Patalong, "Beamer an, Hirn aus"...

... Kaum wird das Licht gedimmt, kaum wirft der Beamer das Bild der ersten Folie an die Wand, kaum hat der Referent angehoben, uns die Mühe abzunehmen, die Folie selbst lesen zu müssen, da legt sich eine Art wuscheliger, muffeliger aber warmer Flokatiteppich um unsere grauen Zellen. Er isoliert das Hirn vom Rest des Körpers, der Aufmerksamkeit simulierend, in Wahrheit in eine Art Winterschlaf-Stasis, ein temporäres Wachkoma verfällt. ...

An der Wand blättern dynamisch die Powerpoint-Folien, zoomen und waschen, gehen fließend ineinander über. Ein hübscher blauer Pfeil rauscht von rechts ins Bild, von oben fällt eine Beschriftung herab, bremst neben der Pfeilspitze, bevor beide von den von unten ins Bild drängenden Balken eines Graphen verdrängt werden. Das Licht geht an, irgendetwas zieht den wuscheligen Flokatiteppich aus unserem Kopf, wir drücken den Rücken durch und klopfen anerkennend auf die Tischplatte: Gut gemacht, schicke Präse, sehr erkenntnisreich. Anerkennendes Nicken reihum.

Willkommen in der Zwischenwelt des Powerpoint, seit Mitte der Achtziger das weltweit am weitesten verbreitete Werkzeug, um komplexe Sachverhalte zu bebildern, ohne sie damit klarer zu machen. ...

Die der Präsentations-Software zugesprochenen Wirkungen scheint es nicht zu geben ...

Schon 2004 kam eine Studiengruppe an der University of New South Wales (Australien) zu dem Ergebnis, dass Powerpoint uns schlicht und einfach überfordert: Informationen in einem Mix aus akustischer und visueller, im schlimmsten Fall sogar noch bewegter Form zu vermitteln, sei völlig kontraproduktiv. Studienleiter John Sweller brachte es damals in die Weltpresse, weil er die Ergebnisse seiner Forschungsgruppe in eine griffige Formulierung goss: "Die Nutzung von Powerpoint-Präsentationen war ein Desaster. Es sollte verboten werden."

Wir sind alle Videoten

Für Wahrnehmungspsychologen sind das alles Binsenweisheiten, die Mechanismen sind seit Jahrzehnten bekannt. ... denn wenn Bild und Text parallel laufen, dominiert das Bild und der Text verhallt und wird vergessen. Deshalb sind Zeitungen, Magazine, Online-Seiten und Bücher gute Informationsmedien und das Fernsehen ein gutes Unterhaltungsmedium: Was etwa die Abendnachrichten an Informationen übermitteln, wird zum größten Teil bereits vergessen, bevor der Wetterbericht beginnt.

Das ist nichts Schlimmes, zeigt aber, dass auch der ach so multimediale Powerpoint-Vortrag vielleicht nicht das beste aller Mittel ist, seine Infos an den Mann zu bringen. Im Klartext: Wenn das Ding gut aussieht, kann man jeden Bockmist dazu erzählen, und niemand bemerkt es. Zyniker glauben, genau dafür sei Powerpoint ja auch da. Präsentationssoftware sei ein Chef-Beeindruckungswerkzeug.

Eine aktuelle, in der nächsten Ausgabe des Fachmagazins "International Journal of Innovation and Learning" veröffentlichte Studie behauptet nun, dass man nicht nur nichts lerne, wenn man derart bepowerpointet wird. Die Zuhörer würden vor allem dann, wenn dynamische, bewegliche Elemente ins Spiel kämen, sogar zusätzlich verwirrt. So ließe sich der Kenntnisstand eines uninformierten Publikums besser erhöhen, wenn man auf Animationen verzichte.

Das setzt allerdings voraus, dass man auch etwas mitzuteilen hat. Denn natürlich liegt der Verdacht nahe, dass Powerpoint nicht zuletzt dafür eingesetzt wird, simple Sachverhalte chic aufzublasen. Zur multimedialen Präsentation kommt es im akademischen wie im Arbeitsleben allzu oft, weil das halt erwartet wird: Beamer und Präse wirken besser als kompetenter Redner mit Tafel und Kreide. Die "gekonnte" Präsentation wird so zum beeindruckenden Feder schmuck des in den Krieg ziehenden Corporate-Indianers - sie sorgt dafür, dass er besser aussieht, als er ist.

zu Ziff. 0.1: vgt0.1 vsem0.1 vprag0.101 vsyst0.1

Bücher

Im Arbeitsbereich können zum Selbstkostenpreis u.a. folgende Bücher von H. Schweizer bezogen werden:

1. **"JOSEPH"** von 1993 (nicht mehr im Handel): Es enthält im Wesentlichen den ursprünglichen Text (d.h. von vielen nachträglichen Entstellungen befreit) der alttestamentlichen Josefsgeschichte ("Josef und seine Brüder" bzw. "Josef in Ägypten"; Gen 37-50). Es ist in der langen Forschungsgeschichte der erste gelungene Versuch, alle Überarbeitungen herauszunehmen und die spannende, freche ursprüngliche Erzählung zu bieten. Mit Photocollagen, Einführung und Kurzinterpretation. € 5.-. [Da in den Vorlesungsmaterialien immer wieder darauf angespielt wird, möchte manche/r vielleicht mal den Text im Ganzen nachlesen]. Ein auf diese Textgestalt bezogener längerer und jüngerer Essay kann aus dem Netz heruntergeladen werden:
<http://www-ct.informatik.uni-tuebingen.de/daten/pattl.pdf> dazu die passende Übersetzung:
<http://www-ct.informatik.uni-tuebingen.de/ct/joskm.html>
 Neu: Gut lesbare Zusammenfassung früherer Arbeiten, incl. Übersetzungstext:
<http://www-ct.informatik.uni-tuebingen.de/daten/jguebers.pdf>
2. **"...deine Sprache verrät dich!" Grundkurs Religiosität. Essays zur Sprachkritik** Münster 2002. Am Beispiel kirchlicher oder theologischer Sprache wird eine Einführung in sprachkritisches Denken gegeben. Viele Bezüge zu benachbarten Wissenschaften (Psychologie, Philosophie, Kunst[theorie], Soziologie). Kapitel daraus können auch für "I&G" verwendet werden. € 20.-. Informationen/Hintergründe/Rezensionen des Buches sind zu finden unter:
<http://www-ct.informatik.uni-tuebingen.de/ct/krs1.htm>
3. **Fantastische "Opferung Isaaks". Textanalyse in Theorie und Praxis (Beispiel Genesis 22)** Lengerich 2006. - Teil II des Buches enthält in kompakter Form das Analysekonzept, das in den verschiedenen Vorlesungen des Arbeitsbereichs vorgestellt wird (Hermeneutik, Konstituierung, Ausdrucks-Syntax, Semantik, Pragmatik). Garniert ist die Theorie durch verschiedene, oft amüsante Textbeispiele. - Teil I ("Isaaks Opferung") ist eine praktische Anwendung und Illustration, wie das textbeschreibende Vorgehen an einem Einzeltext aussieht und welche Ergebnisse erwartet werden können. Bei diesem oft Ratlosigkeit oder Proteste auslösenden Text sind sie jedenfalls faszinierend. - Beide Teile: Kompakter Einblick in unsere textwissenschaftliche Arbeitsweise. - Für Studierende: 12.- €.
4. **Krach oder Grammatik ? Streitschrift für einen revidierten Sprachunterricht - Kritik und Vorschläge** Frankfurt/M 2008 (erschienen Nov. 2007). - Es werden Begriffe, Definitionen und die ganze Grammatikkonzeption untersucht, die SchülerInnen heutzutage im Unterricht gelehrt werden. Vielfach erweisen sich die Inhalte des Grammatikunterrichts als Katastrophe. Kein Wunder, dass dieser Unterricht - bei Schülern wie Lehrern (!) - unbeliebt ist. Das ist besonders verhängnisvoll, da Sprache

unser entscheidendes Mittel ist, uns in der Welt zu orientieren. - Für Studierende: 12.- €.

Bei Interesse melden Sie sich nach Vorlesungen oder in Sprechstunden. Bücher auch im Sekretariat B108 beziehbar.

Viele Aufsätze von H.S. stehen im Rahmen der homepage zur Verfügung: Über den Button "FORSCHUNG", entweder über die einzelne methodische Ebene, oder bei "MANUSKRIPTE"/"SCHRIFTEN".

zu Ziff. 0.2:

Erstreckungsbereich der VORLESUNG

zu:

I. Traditionelle Grammatikgliederung	Arbeitsfelder der Sprachanalyse	II. Semiotisch basierter Dreischritt
Phonetik	Akustik/Optik (Spracherkennung)] (Ausdrucks-) SYNTAX
Morphologie /] SEMANTIK
Syntax	Satzanalyse (Parsing)] SEMANTIK
[? ? ? ? ? ? ? ? ? ? ? ? ? ?	Textanalyse (z.B. Dialogsysteme)] PRAGMATIK
	↓	

=====
Vorgesaltet: Bereitstellung des Textes
=====

Philologie Textkritik/Editions- philologie Literar-/Redaktions- kritik	Prüfung der überlieferten Textgestalt (Les- und Versteh- barkeit, digitale Codierung)	KONSTITUIERUNG DES TEXTES
--	--	--

=====
Flankierend betrieben:
=====

? ? ? ? ?	Methodenlehre zeichentheoretische Grundlagen Philosophie/Logik Kommunikations- theorie	SEMIOTIK SYSTEMTHEORIE
-----------------------	---	---

1. Hinführung / Rekapitulierung**Literatur**

- DÜRR, M; SCHLOBINSKI, P: Einführung in die deskriptive Linguistik. WV studium 163 1990 // I.2.4
- FELIX, S W; KANNGIEßER, S; RICKHEIT, G (ed./eds.): Sprache und Wissen. Studien zur Kognitiven Linguistik. Psycholinguistische Studien 1990

zu Ziff. 1.01:

Konstituierung: Höhlenmalerei / Archäologie

zu:

aus: M. LORBLANCHET, Höhlenmalerei. Ein Handbuch. ²2000 Stuttgart. (126) Grabung und graphische Dokumentation sind Vorgänge einer Entdeckung. In beiden Fällen dringt der Archäologe durch sorgfältige Arbeit tiefer in ein Dokument ein und erreicht so eine Ebene der Realität, die ohne diese Arbeit nicht direkt zugänglich wäre. Die Grabung erkundet (127) die Schichtenfolgen, fördert die Funde zutage, identifiziert sie und fertigt Verbreitungspläne an, um die Strukturen sichtbar zu machen. In gleicher Weise ermittelt die graphische Dokumentation die Motive und grenzt sie gegeneinander ab, sie arbeitet deren Verbindung untereinander und zur Umgebung heraus. Beide Verfahrensweisen sind einander so ähnlich, daß man sagen kann: Eine nicht aufgezeichnete Darstellung ist wie ein nicht ausgegrabener Fund.

Die Grabung erkundet die Aufeinanderfolge der Schichten, die im Laufe der Zeit erfolgte Anhäufung von Zeugnissen in ein und demselben Bereich, so wie die graphische Dokumentation die Überlagerungen der Linien und Figuren auf dem Felsen analysiert. Die Grabung besteht, allgemein gesprochen, in einer Lesung des Bodens, die graphische Dokumentation in einer Lesung der Wand. Grabung und graphische Dokumentation führen beide zu einer graphischen Wiedergabe der Dokumente. Die Grabung erschöpft sich dabei nicht in der Ausarbeitung eines Plans, der alle Spuren enthält, sondern sie umfaßt auch alle Vorarbeiten. Ebenso beschränkt sich die graphische Dokumentation nicht auf eine einfache Wiedergabe, sondern beinhaltet alle Forschungsarbeiten und Analyseschritte, die ihr vorausgingen.

Die Grabung zerstört ihren Untersuchungsgegenstand, während die graphische Dokumentation ihn bewahrt. Dieser grundlegende Unterschied gewährt der graphischen Dokumentation ein größeres Maße an Freiheit. Sie kann es sich im Prinzip erlauben, unvollständig, subjektiv und vorläufig zu sein, da sie wiederholt werden kann. Die Grabung hingegen kann kein zweites Mal vorgenommen werden, kann sich nicht auf die einfache Bestätigung einer Hypothese beschränken und alle anderen, nicht direkt verwertbaren Informationen außer acht lassen; sie muß vielmehr das Maximum aller Daten erfassen, die sich für eine weitere Verwendung eignen.

1.1 Kurzerläuterung der einzelnen Aspekte**Literatur**

ECO, U: Der Name der Rose. München ¹⁵1983

1.2 Ort der Fragestellungen im gesamten Interpretationskonzept**Literatur**

- JECHLE, T: Kommunikatives Schreiben. Prozeß und Entwicklung aus der Sicht kognitiver Schreibforschung. ScriptOralia 41. Tübingen 1992 // I.2.0
- RICKHEIT, G; STROHNER, H: Grundlagen der kognitiven Sprachverarbeitung. UTB 1735. Tübingen 1993: A. Francke // I.2.0
- SCHWARZ, M: Einführung in die Kognitive Linguistik. UTB 1636. Tübingen 1992 // I.2.4
- SCHWEISTHAL, K G (ed./eds.): Grammatik-Kybernetik-Kommunikation (FS A. HOPPE). Bonn 1971 // I.2.4
- SCHWEIZER, Harro: Sprache und Systemtheorie. Zur modelltheoretischen Anwendung der kybernetischen Systemtheorie in der Linguistik. TBL 121. Tübingen 1979: Narr // I.2.0
- WATZLAWICK, P (ed./eds.): Die erfundene Wirklichkeit. Wie wir wissen, was wir zu wissen glauben. Beiträge zum Konstruktivismus. München-Zürich 1985

zu Ziff. 1.21:

Kognition - systemtheoretisch eingeordnet

RICKHEIT, G; STROHNER, H: Grundlagen der kognitiven Sprachverarbeitung. UTB 1735. Tübingen 1993: A.FRANCKE
zu: vsyst5.001

(11f) "Die Komponenten auf der jeweiligen Ebene kooperieren so eng miteinander, daß durch diese Interaktion neue Verhaltensweisen des gesamten Systems ermöglicht werden, die ohne diese Interaktion nicht vorstellbar sind. Eigenschaften eines Systems, die durch die Interaktion der Komponenten neu entstehen, heißen *emergente Eigenschaften*.

In wissenschaftstheoretischer Hinsicht sind die emergenten Eigenschaften die Grundlage dafür, daß die einzelnen Teildisziplinen der Kognitiven Wissenschaft, die für unterschiedliche Ebenen des Gesamtsystems zuständig sind, nicht aufeinander reduziert werden dürfen, und zwar weder in die eine noch in die andere Richtung. Die drei wichtigsten Ebenen des menschlichen kognitiven Systems, die wissenschaftlichen Disziplinen entsprechen, sind die *biologische*, die *psychologische* und die *soziologische* Ebene:

- Der Gegenstand der *Biologie* ist der Organismus von Lebewesen mit seinen verschiedenen Organen, deren Zusammenwirken bei geeigneter Unterstützung durch die Umwelt des Lebewesens die emergente Eigenschaft des *Lebens* entstehen läßt. Aus dem großen Gegenstandsbereich der Biologie sind für die Kognitive Wissenschaft besonders diejenigen Disziplinen relevant, die sich mit dem zentralen Nervensystem beschäftigen, vor allem die Neurologie und die Neurophysiologie, die heute oft unter dem Begriff der *Neurowissenschaft* zusammengefaßt werden. Die Neurowissenschaft gibt Auskunft darüber, wie das zentrale Nervensystem den Beitrag des Organismus zur Kognition zustandebringt.
- Die *Psychologie* analysiert auf der Ebene darüber die Interaktion zwischen dem menschlichen Organismus und seiner Umwelt, deren emergentes Produkt der menschliche Geist mit all seinen sich positiv, aber auch negativ sich auswirkenden Fähigkeiten ist.
- Die *Soziologie* schließlich befaßt sich mit den durch die Interaktion der Menschen entstehenden emergenten Eigenschaften der menschlichen *Gesellschaft* und ihren verschiedenen *kulturellen* Hervorbringungen.

Leben, Geist und Gesellschaft sind drei für die Beschreibung und Erklärung von Kognition relevante Kategorien, die nicht untereinander austauschbar oder aufeinander reduzierbar sind."

1.3 Dichotomie in der Zeichendefinition: signifiant vs signifié

Literatur

- LUHMANN, N: Die Realität der Massenmedien. Opladen 23 2004: Westdeutscher Verlag // I.2.0
- POSNER, R; ROBERING, K; SEBOEK, Th A (ed./eds.): Semiotik / Semiotics. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur. HSK 13 / 1 1997: de Gruyter // I.2.7
- RADTKE, P: Die Kategorien des deutschen Verbs: zur Semantik grammatischer Kategorien. TBL 438. Tübingen 1998: Narr // I.2.4
- SCHWITALLA, J: Gesprochenes Deutsch. Eine Einführung. Grundlagen der Germanistik 33. Berlin 1997: Erich Schmidt // EB
- SOTTONG, H; MÜLLER, M: Zwischen Sender und Empfänger. Eine Einführung in die Semiotik der Kommunikationsanalyse. Berlin 1998: E. Schmidt // I.2.0
- VERSTEEGH, K; et.al. (ed./eds.): The emergence of semantics in four linguistic traditions: Hebrew, Sanskrit, Greek, Arabic. Studies in the history of the language sciences 82. Amsterdam 1996: Benjamins // EB

zu Ziff. 1.31:

Sprechen und Schreiben

zu: vgtl.311

aus: SCHWITALLA, J, Gesprochenes Deutsch. Eine Einführung. Berlin 1997: E. Schmidt

(14-16)

2.1. Gesprochene Sprache als Gegenstand der Linguistik

Am 1. Oktober 1899 hielt Otto Behaghel vor der Hauptversammlung des Deutschen Sprachvereins in Zittau einen Vortrag über das Thema "Geschriebenes Deutsch und gesprochenes Deutsch". Behaghel war der erste Sprachwissenschaftler, der die Unterschiede des gesprochenen zum geschriebenen Deutsch festgehalten hat: Betonung und Satzmelodie, das Lauter- und Leiserwerden der Stimme, die Art, wie man redet, ob "mit leiser gleichgültiger Stimme oder mit starkem Nachdruck [...] das anzudeuten ist die Schrift völlig unvernünftig". Die Verschränkung des Gesprochenen mit dem, wie wir heute sagen, Nonverbalen, die Implizitheit und Knappheit der Äußerungen im Vergleich mit dem vom Sprecher Gemeinten und auch vom Hörer Verstandenen, das notwendig vorausgesetzte Wissen von den Redeumständen, das syntaktisch Reihende und wörtlich Wiederholende - all das hatte Behaghel gut beobachtet und dazu auch die Übergänge zwischen Sprechen zum Schreiben bemerkt.

Warum hat es mehr als 60 Jahre gedauert, bis diese Beobachtungen zu einem richtigen Forschungszweig der Sprachwissenschaft wurden? Ein Grund unter mehreren war das Fehlen von geeigneten Tonaufzeichnungsmaschinen, mit denen man schnell beliebig lange Strecken gesprochener Äußerungen mehrmals nacheinander wiedergeben konnte. Nur beim wiederholten Hören kann man sich nämlich auf unterschiedliche Eigenschaften des Gesprochenen konzentrieren. Dies gelang erst mit der Rückschalttaste des Tonbandgeräts. Das Tonbandgerät wirkt ähnlich wie ein Mikroskop, mit dem man an altbekannten Gegenständen neue Eigenschaften entdecken kann.¹ Aber auch nach der Erfindung des Tonbands Mitte der 30er Jahre dauerte es (abgesehen von frühen Versuchen von E. Zwirner) noch ziemlich lange, bis man in den 60er Jahren begann, gesprochene Alltagsrede aufzuzeichnen und ihre sprachlichen Merkmale zu untersuchen...

Sprachverwendung in der Form des Produzierens von hörbaren Lauten ist eindeutig vom Produzieren sichtbarer, schriftlicher Zeichen zu unterscheiden. Die medialen Realisierungen nennt man "phonisch" und "graphisch"...

Dennoch können wir gesprochene Sprache als einen Forschungsgegenstand nicht anders als durch seinen einzigen Gegenpart, das Ge-

¹ Die Analogien zur Erfindung des Mikroskops sind erstaunlich: die Herkunft des Geräts aus der nicht-wissenschaftlichen Sphäre, die lange Latenzzeit bis zur wissenschaftlichen Anwendung: die Eröffnung eines bislang unbekanntes Mikroskosmos und die damit einhergehende Umorganisation eines wissenschaftlichen Weltbilds; häßliche und formlose Gegenstände erscheinen plötzlich als schön und geordnet etc. Daß man Laute festhalten könne, war vor Erfindung des Grammophons eine Wunschvorstellung wie der Traum vom Fliegen. Seit Plutarch wird die Metapher von den "gefrorenen Wörtern" tradiert, die, wenn man sie auftaut, zu Tönen anfangen. Rabelais hat daraus in "Gargantua und Pantagruel" (4. Buch, 56. Kap.) eine lustige Szene gemacht, die einen Gutteil seines eigenen literarischen Schaffens versinnbildlicht.

schriebene (und dessen Varianten: das Getippte, das auf dem Bildschirm Erscheinende), in den Blick bekommen...

Es gibt noch eine dritte Weise, in der Sprache materiell erscheinen kann, nämlich die Gebärdensprachen der Gehörlosen, welche Handbewegungen im dreidimensionalen Raum für die Vergegenständlichung sprachlicher Strukturen ausnützen.²

² Erst in den 60er Jahren hat man entdeckt, daß Gebärdensprachen vollgültige Realisierungsweisen von "Sprache" sind, mit einem offenen Repertoire an arbiträren Zeichen, mit Analogien zu Phonemen, Flexionssystemen und grammatischen Regeln der Lautsprachen.

zu Ziff. 1.32:

Was sind Medien?

zu: vprag7.5132 vsyst1.361

aus: Niklas Luhmann, Die Realität der Massenmedien. ²1996 Opladen: Westdeutscher Verlag.

A) Medien = Technische Kanäle

(10) Mit dem Begriff der Massenmedien sollen im folgenden alle Einrichtungen der Gesellschaft erfaßt werden, die sich zur Verbreitung von Kommunikation technischer Mittel der Vervielfältigung bedienen. Vor allem ist an Bücher, Zeitschriften, Zeitungen zu denken, die durch die Druckpresse hergestellt werden; aber auch an photographische oder elektronische Kopierverfahren jeder Art, sofern sie Produkte in großer Zahl mit noch unbestimmten Adressaten erzeugen. Auch die Verbreitung der Kommunikation über Funk fällt unter den Begriff, sofern sie allgemein zugänglich ist und nicht nur der telephonischen Verbindung einzelner Teilnehmer dient.

(13) Dennoch wollen wir die Arbeit dieser Maschinen und erst recht ihr mechanisches oder elektronisches Innenleben nicht als Operation im System der Massenmedien ansehen ... Es macht daher guten Sinn, die reale Realität der Massenmedien als die in ihnen ablaufenden, sie durchlaufenden Kommunikationen anzusehen.

B) Medien = Einbahnkommunikationen

(11f) Entscheidend ist auf alle Fälle: *daß keine Interaktion unter Anwesenden zwischen Sender und Empfänger stattfinden kann.* ... Durch die Unterbrechung des unmittelbaren Kontaktes sind einerseits hohe Freiheitsgrade der Kommunikation gesichert. Dadurch entsteht ein Überschuß an Kommunikationsmöglichkeiten, der nur noch systemintern durch Selbstorganisation und durch eigene Realitätskonstruktionen konfrontiert werden kann. Andererseits sind zwei Selektoren an Werk: die Sendebereitschaft und das Einschaltinteresse, die zentral nicht koordiniert werden können. Die Organisationen, die die Kommunikation der Massenmedien produzieren, sind auf Vermutungen über Zumutbarkeit und Akzeptanz angewiesen.

(13f) Während wir die technischen Apparaturen, die "Materialitäten der Kommunikation", ihre Wichtigkeit unbenommen, aus der Operation des Kommunizierens ausschließen, weil sie nicht mitgeteilt werden, schließen wir den (verstehenden bzw. mißverstehenden) Empfang ein. Eine Kommunikation kommt nur zustande, wenn jemand sieht, hört, liest - und so weit versteht, daß eine weitere Kommunikation anschließen könnte.

(33f) Für die Ausdifferenzierung eines Systems der Massenmedien dürfte die ausschlaggebende Errungenschaft in der Erfindung von Verbreitungstechnologien gelegen haben, die eine Interaktion unter Anwesenden nicht nur einsparen, sondern für die eigenen Kommunikationen der Massenmedien wirksam ausschließen. Schrift allein hatte diesen Effekt noch nicht, denn sie war zunächst nur als Gedächtnisstütze für primär orale Kommunikation konzipiert worden. Erst der Buchdruck multipliziert das Schriftgut so stark, daß eine mündliche Interaktion *aller* an Kommunikation Beteiligten wirksam *und sichtbar* ausgeschlossen wird. Die Abnehmer machen sich allenfalls quantitativ bemerkbar: durch Absatzzahlen, durch Einschaltquoten, aber nicht entgegenwirkend. Das Quantum ihrer Präsenz kann bezeichnet und interpretiert werden, wird aber nicht über Kommunikation rückvermittelt. Selbstverständlich bleibt mündliche Kommunikation als Reaktion auf Gedrucktes oder Gefunktes möglich. Aber

das Gelingen von planmäßiger Kommunikation hängt davon nicht mehr ab. So kann im Bereich der Massenmedien ein autopoietisches, sich selbst reproduzierendes System entstehen, das auf Vermittlung durch Interaktionen unter Anwesenden nicht mehr angewiesen ist. Erst damit kommt es zu einer operativen Schließung mit der Folge, daß das System die eigenen Operationen aus sich heraus reproduziert, sie nicht mehr zur Herstellung von interaktionellen Kontakten mit der gesellschaftsinternen Umwelt verwendet.

(164) Dies alles gilt auch für die Realität der Massenmedien. Auch hier ist es operativ nicht möglich, und dies kann man wissen, die Selektivität der publizierten Informationen in die Rekursivität der gesellschaftlichen Kommunikationen einzubeziehen. Man reagiert wie der bereits zitierte Horatio: "So I have heard, and do in part believe it." Es mag zwar manches Detail bezweifelt werden und jeder mag Gelegenheiten finden, sich selbst in die Kommunikation mit besonderen Meinungen einzubringen. Aber den Rahmen der Konsistenzprüfungen, der Rekursivität, kann die Kommunikation im Gesellschaftssystem nicht ausschalten. Sie verlore sonst fast allen täglich benötigten Sinn... Das vielleicht wichtigste Ergebnis dieser Überlegungen ist, daß die Massenmedien zwar die Realität, aber eine nicht konsenspflichtige Realität erzeugen.

(167) Man muß nur die eigene Art der Einstellung auf Realität akzeptieren - und unterscheiden können. Man muß sich nur davor bewahren, sie für allgemeingültig, für die Realität schlechthin zu halten.

(204) Wie das Theater versetzen auch die Massenmedien das Individuum in eine Szene außerhalb der Inszenierung. Wir hatten das als technische Bedingung der Ausdifferenzierung eines Mediensystems beschrieben. Auf die Individuen muß diese Distanz ambivalent wirken. Denn einerseits sind sie nicht selbst der Text, der ihnen vorgeführt wird; und wenn sie ihn geschrieben und publiziert haben wie Rousseau, sind sie es schon nicht mehr. Sie sehen sich auch nicht selber im Fernsehen, und wenn ausnahmsweise, dann mit einem Spezialvergnügen des Sich-selbst-Wiedererkennens, das man nur bei Ausnahmen findet. Andererseits produzieren die Massenmedien die Welt, in der die Individuen sich selber vorfinden... Wenn Individuen Medien als Text oder als Bild betrachten, sind sie draußen; wenn sie in sich deren Resultate erleben, sind sie drinnen. Sie müssen zwischen draußen und drinnen oszillieren, und dies so wie in einer paradoxen Situation: schnell, fast ohne Zeitverlust und unentscheidbar. Denn die eine Position ist nur dank der anderen möglich - und umgekehrt.

zu Ziff. 1.33:

Zeichen

zu: vsyst1.335 vgt2.135

aus: SOTTONG, H; MÜLLER, M, Zwischen Sender und Empfänger. Eine Einführung in die Semiotik der Kommunikationsanalyse. Berlin 1998.

(52) Unsere Beispiele und Ausführungen haben gezeigt, daß Zeichen erstens immer Objektivationen menschlicher Tätigkeiten sind, die zweitens intersubjektiv und sozial funktionieren müssen, also den Charakter kultureller Institutionen haben. Zeichen sind drittens - im Gegensatz zu Anzeichen - ablösbar und unabhängig von dem, was sie bezeichnen. Die Bezeichnungsfunktion ist Bestandteil eines jeden Zeichens, was aber nicht heißt, daß das Bezeichnete selbst Bestandteil des Zeichens ist. Demgegenüber ist die Bedeutung des Zeichens ein notwendiger Bestandteil: Sie ist in aller Regel nicht etwas, das wir aufgrund des reinen Vorhandenseins oder des Erscheinungsbildes des Zeichens folgern können, sondern etwas, das wir kennen müssen - jeder, der jemals eine Fremdsprache gelernt hat, weiß, daß man der Gestalt eines Wortes nicht ansehen kann, was es bedeutet; er mußte die Bedeutung mehr oder weniger mühevoll erlernen.

Alle Zeichen bestehen somit aus zwei Komponenten: einem Erscheinungsbild und einem "Inhalt", oder anders ausgedrückt, aus dem materiellen Zeichenträger und der Bedeutung...

Diesen materiellen Aspekt des Zeichens nennt man "Signifikant". In der Theorie der Zeichen kann der Signifikant als ein ideales Modell des Zeichenkörpers aufgefaßt werden, auf das sich jede Realisation dieses Zeichens in der Kommunikation beziehen läßt. Nehmen wir als Beispiel das geschriebene Wort "Semiotik": es ist im vorliegenden Text in einer bestimmten Type, einer bestimmten Größe geschrieben. Wir können seine Gestalt selbstverständlich variieren und fett **Semiotik** schreiben oder kursiv *Semiotik* usw. Jedesmal haben wir es mit einem speziellen Vorkommen des Signifikanten (der Buchstabenfolge S-e-m-i-o-t-i-k) zu tun, die sich untereinander zwar leicht unterscheiden, aber als Repräsentationen ein und desselben Modells, des Signifikanten "Semiotik" erkennbar sind. Der Signifikant ermöglicht und sichert also die Reproduzierbarkeit und (Wieder-)erkennbarkeit des Zeichens. Nun ist selbstverständlich nicht alles, was als Realisation eines Modells aufgefaßt und wiedererkannt werden kann, auch schon ein Zeichen. Mit dem Zeichenkörper, dem Signifikanten, muß eine bestimmte Bedeutung verknüpft sein, eine bestimmte Wissensmenge, die durch das Auftreten des entsprechenden Signifikanten regelmäßig abgerufen wird. Diesen Aspekt des Zeichens, seine Bedeutung, nennen wir mit dem semiotischen Fachbegriff das "Signifikat" des Zeichens. Erst die feste Relationierung einer bestimmten, lautlich, gestisch oder bildlich realisierbaren Gestalt mit einer definierten, intersubjektiv nachvollziehbaren Bedeutung macht ein Phänomen zum Zeichen...

(54) Zeichen bilden die Grundlage der Kommunikation. Anders als Anzeichen gehören sie ganz und gar der kulturellen Sphäre an und stehen in keiner kausalen, materiellen, temporalen Relation zu dem, was sie bezeichnen: in dieser Hinsicht sind sie arbiträr und autonom. Sie sind das Ergebnis einer ausdrücklichen oder unausgesprochenen, aktuellen oder tradierten sozialen Übereinkunft zwischen mindestens zwei Zeichenbenutzern, zwischen denen ein - bewußter oder bereits zur Selbstverständlichkeit gewordener - Grundkonsens über ihre Anwendung und ihre Bedeutung herrscht. Wenn wir also Zeichen betrachten, betrachten wir sie als Objektivationen sozialer Tätigkeit, wir betrachten und beschreiben etwas, was zwi-

schen Menschen ausgetauscht wird, nicht etwa das, was in den Köpfen von Individuen ausgelöst wird oder intendiert war. Zeichen bestehen immer aus zwei Grundelementen: dem Signifikanten - also dem materiellen Zeichenträger - und dem Signifikat, der Bedeutung des Zeichens. Mengentheoretisch ausgedrückt: Ein Zeichen ist eine Zweiermenge aus Signifikant und Signifikat (Signifikant, Signifikat). Ein Signifikat ist seinerseits eine Zweiermenge aus Denotat und Konnotat ($\{D,K\}$); die beiden Elemente dieser Zweiermenge sind Mengen von semantischen Merkmalen. Das Konnotat kann die leere Menge sein, wenn alle Benutzer des Zeichens diesem genau die gleichen semantischen Merkmale zuschreiben. Ein Zeichen, bei dem man von evtl. existierenden Konnotaten absieht und ausschließlich das Denotat im Signifikat betrachtet, nennen wir Basiszeichen. Verschiedene nicht leere Konnotate $K_1...K_n$ zu einem Basiszeichen ergeben die Signifikatsvarianten $S_1...S_n$.

zu Ziff. 1.34:

Zeichen / Bedeutung / Gebrauchsregeln

zu: vsyst1.334 vgt2.134 vprag1.961

aus: RADTKE, P, Die Kategorien des deutschen Verbs. Zur Semantik grammatischer Kategorien. TBL 438. Tübingen 1998.

(47ff) Erstens: Die repräsentationistische Zeichenauffassung geht - wie bereits gesehen - davon aus, daß Zeichen durch die "Stellvertreter-Relation" symbolisieren. Sie stehen für etwas in der Welt (ontologische Variante) oder für ein Konzept (epistemologische Variante). Derartige Zeichenauffassungen orientieren sich vornehmlich an der Bedeutung von Autosemantika; Synsemantika sind in diesem Rahmen weitaus komplizierter zu beschreiben. Die Bedeutung von *Tisch* beispielsweise wäre je nach Theorie ein tatsächlicher Tisch, die Menge aller existierenden oder möglichen Tische oder die Menge der Tisch-Konzepte der Sprecher einer Sprachgemeinschaft (beschreibbar über Merkmalsmengen oder Prototypenkonzepte). Eine Konjunktion wie *weil* ließe sich nicht über eine extensionale Bedeutung (48) beschreiben, sondern allenfalls als ein Konzept (das Konzept der Kausalität?).

Zweitens: Instrumentalistische Zeichenauffassungen gehen davon aus, daß Zeichen dadurch symbolisieren, daß ihre Verwendung konventionell geregelt ist. "Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache," schreibt Wittenstein. Dieser Auffassung gemäß ist die Bedeutung eines Zeichens genau das, was das Zeichen interpretierbar macht. Die Interpretierbarkeit eines sprachlichen Zeichens ergibt sich aus der Regelmäßigkeit des Gebrauchs in der Sprachgemeinschaft. Zeichen symbolisieren also nicht dadurch, daß sie für etwas stehen, sondern dadurch, daß sie Gebrauchsregeln folgen. Formuliert man die Gebrauchsregel aus, so erhält man eine Formulierung der Bedeutung des Zeichens. Entgegen weit verbreiteter Interpretation ist damit nicht gemeint, daß sich die Bedeutung eines Ausdrucks in Abhängigkeit vom jeweiligen Kontext verändert und mithin ein Parole-Ereignis sei. Die *Bedeutung* eines Ausdrucks ist - synchron betrachtet - konstant, der *Sinn* kann von Verwendung zu Verwendung variieren. Die Bedeutung (d.h. die Regel des Gebrauchs) ist rekonstruierbar aus den konkreten Verwendungsweisen; sie ist jedoch keinesfalls identisch mit den verschiedenen Verwendungsweisen, und sie ist auch nicht die Summe aller Verwendungsweisen. Innerhalb dieses Modells weisen sowohl Autosemantika wie auch Synsemantika Bedeutung auf. Die Bedeutung der Ausdrücke *Tisch* und *weil* erhält man, wenn man die Konvention ihres Gebrauchs beschreibt, d.h. die Kriterien der Verwendung von *Tisch* und *weil* ausformuliert. Entsprechendes gilt auch für grammatische Kategorien wie etwa die des Verbs. ...

(49) Die Bedeutung eines komplexen Ausdrucks (eines Satzes, eines finiten Verbs etc.) ergibt sich kompositionell aus der der einfachen Ausdrücke (des Verbs, der Verbalkategorie etc.). Der entscheidende Unterschied besteht in dem Bedeutungsbegriff: Unter der Bedeutung eines Ausdrucks werden weder Konzepte verstanden noch Wahrheitsbedingungen, sondern Gebrauchsregeln. Die Kenntnis von Wahrheitsbedingungen ließe sich allenfalls als ein Spezialfall der Kenntnis von Gebrauchsbedingungen ansehen. In die Regel können nämlich ganz unterschiedliche Parameter eingehen - neben (i) wahrheitsfunktionalen auch (ii) epistemische, (iii) soziale, (iv) diskursbezogene und (v) innersprachliche; Kombinationen sind ebenfalls möglich.

(50) Zu (i): Der wohl bekannteste Parameter, der, wie bereits ausgeführt, oftmals verabsolutiert und als Bedeutung schlechthin an-

gesehen wird, ist der wahrheitsfunktionale. Die Gebrauchsregel gibt an, wie die Welt beschaffen sein sollte, damit der Ausdruck sinnvoll verwendet werden kann. Paradebeispiel ist *Junggeselle*; der Ausdruck dient üblicherweise dazu, auf Personen zu referieren, die männlich, erwachsen und und unverheiratet sind.

Zu (ii): Die Gebrauchsregel kann jedoch auch epistemische Bedingungen zum Gegenstand haben. Für die Bedeutung von evaluativen Adjektiven wie *gut* und *schön* sind Objekteigenschaften weitgehend irrelevant; vielmehr müssen die jeweiligen Bewertungen und Einstellungen des Sprechers in die Gebrauchsregel eingehen.

Zu (iii): Eine ganze Reihe von Ausdrücken folgt sozialen Parametern. So kann beispielsweise der Gebrauch der Anredepronomen *du* und *Sie* nur mit Bezug auf das bestehende Sozialsystem expliziert werden. Der Hinweis, das Pronomen diene dazu, auf den Gesprächsteilnehmer zu referieren, allein reicht nicht aus.

Zu (iv): Des weiteren können diskursbezogene Parameter bei der Formulierung der Gebrauchsregel eine Rolle spielen. Modalpartikeln wie *mal* und *wohl* haben diskurssteuernde Funktion, was z.B. daran deutlich wird, daß sie bestimmte Sprechakttypen indizieren können. Eine wahrheitsfunktionale Semantik muß zwangsläufig an der Beschreibung derartiger Ausdrücke scheitern, da Modalpartikeln sich gerade dadurch auszeichnen, daß sie den Wahrheitswert einer Aussage unberührt lassen.

Zu (v): Nur mittels innersprachlicher Parameter kann beispielsweise die Bedeutung von Relativpronomen expliziert werden. Für die Verwendung des Relativpronomens *der* ist die sprachliche Umgebung, d.h. der Kontext, relevant - es muß ein Nominal auftreten, auf das es sich beziehen kann...

(51) Versteht man unter der Bedeutung eines Ausdrucks in Spät-Wittgensteinianischer Tradition eine Gebrauchsregel im dargelegten Sinne, so hat dies meines Erachtens die folgenden vier Vorzüge:

(i) Lexeme und grammatische Kategorien können einheitlich behandelt werden; diese Unterscheidung wird (unter bedeutungstheoretischen Überlegungen) hinfällig.

(ii) Historische Prozesse wie der der Grammatikalisierung, d.h. beispielsweise der Übergang eines freien Elements (z.B. eines Lexems) zu einem gebundenen Element (z.B. einer grammatischen Kategorie), können problemlos dargestellt werden.

(iii) Umständliche Mechanismen und Prinzipien, die zwischen Semantik (bzw. "semantischem Basiskonzept") und Grammatik (bzw. "grammatischem Basiskonzept") vermitteln, werden überflüssig.

(iv) Unterschiedliche Bezugssysteme erfordern keine eigenen Ontologien; das heißt, es bedarf nicht der Annahme einer Welt der Dinge für semantische Konzepte, einer Welt der Sozialbeziehungen für epistemische Konzepte, einer Welt der Sozialbeziehungen für soziale Konzepte, einer Diskurswelt für kommunikativ-pragmatische Konzepte und einer Welt der linguistischen Kategorien für grammatische Konzepte.

zu Ziff. 1.341:

Grammatik & Sprachgebrauch - Regel

zu: vsyst2.004

aus: M. MEYER, Grammatische Praxis. Probleme der grammatischen Theoriebildung und der Grammatikschreibung. Linguistik 43. Tübingen 2006.

(16) "Grammatik" war ... eng verbunden mit dem Studium von schriftlich überlieferten Texten und insofern befasst mit der Frage danach, welche "Sätze" in der jeweiligen "Sprache" vorkommen und welche nicht. Die Frage nach der "Wohlgeformtheit" war also im Wesentlichen eine Frage nach dem Vorkommen oder Nichtvorkommen von "Sätzen" in bestimmten Texten. "Grammatik" zeigte sich damit als ein Rekonstruktionsunternehmen (Bartsch: 1985, 80). Dies deutet auf eine entscheidende (17) Voraussetzung des Studiums von "Sprache" hin, die im Sinne einer philosophischen Begründung unhintergebar erscheint: "... die syntaktischen und semantischen Formen [ergeben sich] in reflektierender Analyse eines schon durch implizite Normen des Richtigen etablierten Gebrauchs." (Stekeler-Weithofer: 2002a, 208) Das Studium von "Sprache" ist aus dieser Perspektive also eine Beschreibung von etwas, das uns schon immer vertraut ist. Wir sprechen eben eine oder mehrere "Sprachen", bevor wir uns dem Studium der eigenen oder fremder "Sprachen" widmen. "Grammatik" nimmt vor allem Bezug auf beobachtbare Sätze - zunächst in Form überlieferter Schriften - und damit auf einen Bereich, der in Bezug auf seine Verwendung außerhalb der "Grammatik" als Wissenschaft liegt und z.B. als "sprachliche Norm" rekonstruiert wird (Stetter: 1999b, 93). Damit verbunden ist einer der wesentlichen Zwecke traditioneller grammatischer Forschungen: Es galt in belehrender Absicht darzustellen, was als normgerechte Schriftsprache seine Gültigkeit immer schon erlangt hatte (Bartsch: 1987, 187, Jellinek: 1968). Diese Orientierung an einer "Norm" ermöglichte zugleich eine Entscheidung darüber, was als "wohlgeformt" und was als "abweichend" zu bezeichnen war. ...

(80) Chomsky legt das Bild eines rechnenden Automaten seiner Erforschung syntaktischer Strukturen zugrunde. Alle wesentlichen Begriffe und Annahmen werden nicht nur gemäß eines solchen Bildes, sondern auch aus einem Bild heraus formuliert - die möglichst genaue Beschreibung "realer Verhältnisse" ist nicht Ziel seiner Forschungen. Der Begriff "grammatisch" wird als "Berechenbarkeit einer Strukturbeschreibung" definiert, ungeachtet der Frage, ob diese Berechnung durch "Regeln" und/oder durch so genannte "Prinzipien" gesteuert wird. "Grammatische Sequenzen" sind also solche, die gemäß den Berechnungsvorschriften eines Automaten erzeugt werden. Damit erweist sich der Begriff der Regel als fundamental für Chomsky. Mit dem Begriff der "Grammatikalität" wird kein Bezug zu einer (einzel-)sprachlichen Norm hergestellt, er wird also inhaltlich nicht gefüllt und bleibt rein formal-schematisch. Dennoch macht ein solcher Begriff nur Sinn vor dem Hintergrund der Annahme, dass die Menge "grammatischer Sequenzen der formalen Sprache "L" genau denjenigen natürlichsprachlichen Sätzen einer beliebigen (Einzel-)Sprache entsprechen, die von "kompetenten Sprechern" jeweils als "akzeptabel" bezeichnet werden, zumindest in Bezug auf den Kern "klarer Fälle". Sinnvoll kann daher ein solcher Begriff von "Grammatikalität" nur dann sein, wenn man immer schon weiß, was es heißt, dass ein bestimmter Ausdruck in einer bestimmten Sprache "wohlgeformt" ist. Damit wird eine Orientierung am vertrauten Umgang mit der jeweiligen Sprache deutlich, den ein formal-schematischer Begriff ausschließt:

Zumindest prima facie handelt es sich also bei diesen Regeln [der Generativen Grammatik, M.M.] nicht um "reale" Regeln der Spracherzeugung und Spracherkennung der Sprecher einer Sprache, sondern nur um ein auf bestimmte Weise geordnetes System der Darstellung konfigurativer Regelmäßigkeiten von sprachlichen Ausdrucksformen, die wir schon in Gebrauch haben. (Kam-bartel/Stekeler-Weithofer: 2988, 215)

(265) In Bezug auf diese Hypothese konnte am Beispiel der Grammatik Chomskys nachgewiesen werden, dass eine formale Grammatik die "Grammatikalität" von "Sequenzen" einer "formalen Sprache" nicht formal bestimmen kann, da eine Formalisierung als Rekonstruktion dessen aufzufassen ist, was in Bezug auf eine "natürliche Sprache" bereits als "grammatisch" immer schon vorausgesetzt ist. Formale Grammatik kann demnach als ein bestimmtes System von Aussagen aufgefasst werden, in dem bereits grammatisch relevante Unterschiede auf spezifische Weise notiert und in ein systematisches Verhältnis zueinander gesetzt werden.

(267) Es gibt sowohl theoretische als auch empirische Argumente, die es plausibel erscheinen lassen, die Grenze des Grammatischen als "unscharf" anzunehmen, d.h. "Grammatikalität" als relative oder gar graduelle Eigenschaft sprachlicher Ausdrücke anzunehmen.

1.4 Hermeneutische Spirale**Literatur**

STEIN, D (ed./eds.): Cooperating with written Texts. The pragmatics and comprehension of written texts. Studies in Anthropological Linguistics 5. Berlin 1992: de Gruyter // I.2.0

**1.5 Das SLANG-Konzept - Lernende, nicht-normative Grammatik = eS
arch for a Learning Non-normative Grammar****A. KONSTITUIERUNG DES TEXTES****Literatur**

BRADLEY, J: Thinking about interpretation. Pliny and scholarship in the humanities. Literary and Linguistic Computing 3 (2008) 263-279.

2. Texterfassung / Maschinenlesbarkeit**Literatur**

- ALBERT, R; KOSTER, C J: Empirie in Linguistik und Sprachlehrforschung. Ein methodologisches Arbeitsbuch. Tübingen 2002 ISBN 3-8233-4985-6: Narr // I.2.0
- ALTMANN, G; FENGXIANG, F (ed./eds.): Analyses of Script. Properties of Characters and Writing Systems. Quantitative Linguistics 63. Berlin 2008 978-3-11-019641-2: Mouton de Gruyter // I.2.7
- BÁTORI, I S u.a (ed./eds.): Computational Linguistics / Computerlinguistik. Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendungen. HSK 4. Berlin 1989 // I.2.8
- BURCH, T; FOURNIER, J; GÄRTNER, K; RAPP, A (ed./eds.): Standards und Methoden der Volltextdigitalisierung an der Universität Trier, 8-9 Oktober 2001. Stuttgart 2003: Franz Steiner Verlag // I.2.1
- BURNARD, L: The Text Encoding Initiative: A progress report in: LEITNER, G (ed./eds.): New Directions in English Language Corpora. Methodology, Results, Software Developments. Topics in English Linguistics 9. Berlin, New York 1992 97-107. // I.2.1
- FRAISTAT, N; JONES, S E: Editing Environments. The Architecture of Electronic Texts. Literary and Linguistic Computing 24 (2009) 9-18.
- GORMAN, G E (ed./eds.): Metadata applications and management. International Yearbook of Library and Information Management (IYLIM) 2003-2004. London 2004
- GRÜTER, M: Optimalitätstheoretische Modellierung von Groß- und Kleinschreibung. Eine Beispielanalyse englischer Texte um 1730. Zeitschrift für Sprachwissenschaft 28,2 (2009) 203-230.
- HANDLER, P (ed./eds.): E-Text: Strategien und Kompetenzen - Elektronische Kommunikation in Wissenschaft, Bildung und Beruf. Frankfurt / M 2001 ISBN 3-631-37206 X br: Peter Lang // I.2.8
- JUNKER, M-O; LUCHIAND, R: Developing Web Databases for Aboriginal Language Preservation. Literary and Linguistic Computing 22 / 2 (2007) 187-206.
- KAMZELAK, R (ed./eds.): Computergestützte Text-Edition. Beihefte zu editio, Band 12. Tübingen 1999: Niemeyer // I.2.1
- KAMZELAK, R (ed./eds.): Computergestützte Text-Edition. Beihefte zu editio, Band 12. Tübingen 1999: Niemeyer // I.2.1
- KÜSTER, M W: Geordnetes Weltbild. Die Tradition des alphabetischen Sortierens von der Keilschrift bis zur EDV. Eine Kulturgeschichte. Tübingen 2006 978-3-484-10899-8: Niemeyer // I.2.1
- KÜSTER, M W: Europäische Standardisierung oder der ganz normale Wahnsinn. Benutzer-Informationen (98 / 7-8) 6-10.: <http://www.uni-tuebingen.de/zdv/bi/bi98/bi98711-euro.html>
- McENERY, T; WILSON, A u.a.: Multilingual Resources for European Languages: Contributions of the CRATER Project. Literary and Linguistic Computing 12 (1997) 219-226.
- NELSON, G: Standardizing Wordforms in a Spoken Corpus. Literary and Linguistic Computing 12 (1997) 79-85.
- O'DONNELL, D P: Back to the future. what digital editors can learn from print editorial practice. Literary and Linguistic Computing 24 (2009) 113-125.
- PEKAR, V; EVANS, R: Discovery of Language Resources on the Web: Information Extraction from Heterogeneous Documents. Literary and Linguistic Computing 22 / 3 (2007) 329-343.
- ROBINSON, P: What text really is not, and why editors have to learn to swim. Literary and Linguistic Computing 24 (2009) 41-52.
- SPERBERG-McQUEEN, C M: How to teach your edition how to swim. Literary and Linguistic Computing 24 (2009) 27-39.
- STEHNO, B; EGGER, A et. al.: METAe - Automated Encoding of Digitized Texts. Literary and Linguistic Computing 18 / 1 (2003) 77-88.
- YONG-ho, Choi: Das koreanische Hangul und das lateinische Alphabet: Eine semiotische Untersuchung des Ursprungs zweier Buchstabenschriften. Zeitschrift für Semiotik (2007) 29 / 1 109-121.

zu Ziff. 2.01:

etexts / Korpusanalyse / tools / Übersetzungsmaschinen

Von unserer homepage aus:

<http://www-ct.informatik.uni-tuebingen.de/ct/infoquellen.htm>

Textarchiv:

<http://www.gutenberg2000.de>

links zu Textarchiven:

<http://www.ub.uni-siegen.de/epub/archive.htm>

<http://www.uni-duesseldorf.de/ulb/extexte.html>

<http://www.grass-gis.de/bibliotheken/buecher.html>

Zugang zu Texten und zusätzlichen Beschreibungen:

http://www.ksbjg.ch/ksbgalt/FachgruppeDeutsch/KSBG_FgDt_Adres-sen_AutorInnen.htm

Suchmaschine:

http://www.fh-hannover.de/bibl/ub/all_epub.htm

Korpusanalyse: Institut für deutsche Sprache - COSMAS II

<http://www.ids-mannheim.de/cosmas2>

Korpusanalyse: Romanische Sprachen

<http://www.uni-duisburg.de/FB3/ROMANISTIK/PERSONAL/Burr/hum-comp/tacthome.htm>

Vernetzung verschiedener Archive:

<http://www.linguistlist.org/olac>

zu Ziff. 2.011:

Korpustechnologie - COSMAS / IDS - Textbeschaffung

zu:

aus: PERKKUHN, R; BELICA, C; al-WADI, D; LAUER, M; STEYER, K; WEISS, C, Korpustechnologie am Institut für Deutsche Sprache in: Schwitalla, J; Wegstein, W, Korpuslinguistik deutsch: synchron - diachron - kontrastiv. Würzburger Kolloquium 2003. Tübingen 2005. S. 57-70

(58) Die Korpora geschriebener Gegenwartssprache des IDS

bilden mit knapp zwei Milliarden Wörtern die weltweit größte Sammlung elektronischer Korpora mit geschriebenen deutschsprachigen Texten aus der Gegenwart und der neueren Vergangenheit.

sind gebührenfrei zugänglich über COSMAS, das IDS-eigene Recherche- und Analyse-System, das speziell auf linguistische Bedürfnisse abgestimmt ist.

enthalten belletristische, wissenschaftliche und populärwissenschaftliche Texte, eine große Zahl von Zeitungstexten sowie eine breite Palette weiterer Textarten und werden kontinuierlich weiterentwickelt.

werden im Hinblick auf große Variabilität und Menge akquiriert und erlauben in der Nutzungsphase über COSMAS die Komposition virtueller Korpora, die repräsentativ oder auf spezielle Aufgabenstellungen zugeschnitten sein können. ...

Insofern sollte es eigentlich selbstverständlich, ja sogar eine Ehre sein, dass Autoren ihre Vorlagen zur Verfügung stellen. Die Rechteinhaber werden aber häufig von der Angst abgeschreckt, mit der Freigabe ihrer Texte für die Korpora lasse sich deren widerrechtliche Vervielfältigung nicht mehr kontrollieren. Darüber hinaus sind die meisten Vorlagen aus technischen Gründen nicht besonders gut geeignet. Die Textproduzenten sind selten willens oder in der Lage, ihre Quellen in einem Format aufzubereiten, das eine leichte Überführung in das Korpusformat ermöglicht. Andererseits verfügt das IDS nicht über die Kapazitäten, die ansonsten erforderliche aufwändige Aufbereitung der Quellen selber zu übernehmen. Aufgrund dieser Rahmenbedingungen gestalten sich die rechtlichen und finanziellen Verhandlungen meistens sehr schwierig.

Urheberrechte

Durch juristische Vereinbarungen mit Verlagen, Zeitungsredaktionen und Autoren war und ist das IDS in der Lage, urheberrechtlich abgesichertes Textmaterial derart zu beschaffen, dass alle Korpora IDS-intern und Teile dieser Korpora weltweit öffentlich genutzt werden können, und zwar ausschließlich zu wissenschaftlichen, nichtkommerziellen Zwecken. Die Textkorpora des IDS sind zudem nur über das COSMAS-System recherchierbar; kein Nutzer hat Zugriff auf vollständige Korpustexte, sondern nur auf begrenzte Kontexte zu Suchanfragen. ...

(59) Um Teil der IDS-Korpora zu werden, müssen sie in ein einheitliches, durch das IDS-Textmodell beschriebenes Format überführt werden. Das bedeutet, dass große Mengen sehr heterogener Daten in mehreren Arbeitsschritten analysiert und aufwärts konvertiert werden müssen. Zur maschinellen Unterstützung werden dazu verschiedene Parser, Konvertierer und Filter z.T. selbst entwickelt und eingesetzt. Probleme bereiten z.B. Inkonsistenzen bei Zeichensatzkodierung und Worttrennung sowie "Datenmüll" wie unmotivierte For-

matangaben, Tabellen oder Textdopplungen. Aber auch sinnvolle Formatangaben variieren je nach Quelle und müssen für eine optimale Datenüberführung eingehend analysiert und vereinheitlicht werden. Die Aufwärtskonvertierung der Quellen in das IDS-Format hat einen stark iterativen Charakter und ist wegen des damit verbundenen hohen Korrektur- bzw. Wartungsbedarfs sehr kosten- und zeitaufwändig.

zu Ziff. 2.012:

IDS - Textmodell

zu:

aus: PERKKUHN, R; BELICA, C; al-WADI, D; LAUER, M; STEYER, K; WEISS, C, Korpustechnologie am Institut für Deutsche Sprache in: Schwitalla, J; Wegstein, W, Korpuslinguistik deutsch: synchron - diachron - kontrastiv. Würzburger Kolloquium 2003. Tübingen 2005. S. 57-70

(61) Um virtuelle Korpuskompositionen, sinnvolle Quellenlisten bei der Ergebnispräsentation u.a.m. zu ermöglichen, werden die Quelltexte nach festgelegten Kriterien gegliedert und in eine hierarchische Struktur eingebunden, die folgende drei Ebenen umfasst:

Korpusebene (Korpusidentifikator, z.B. LES)

|→ Dokumentebene (Dokumentidentifikator, z.B.LES/ESS)

|→ Textebene (Textidentifikator, z.B.LES/ESS.20022)

Das IDS-Textmodell definiert Text als eine relativ selbstständige, inhaltlich kohärente Folge natürlichsprachlicher Äußerungen, die natürlichen Kommunikationssituationen entstammen. Sie bildet den Korpustext, die "kleinste" Einheit eines Korpus.

Jedes Korpus besteht aus einem oder mehreren Dokumenten; jedes Dokument setzt sich wiederum aus einem oder mehreren Korpustexten zusammen. In einem Dokument können mehrere Texte nach bestimmten Gesichtspunkten zusammengefasst sein, z.B. nach Quellen, chronologischer Abfolge, Themenbereichen und/oder Textarten. Ein Text beinhaltet je nach Korpusstruktur z.B. einen oder mehrere Zeitungsartikel oder eine als Ganzes aufgenommene Zeitung/Zeitschrift, einen Auszug aus einem selbstständigen Werk oder ein selbstständiges Werk als Ganzes.

Beispiel: Das Korpus Siegfried Lenz: Werkausgabe in Einzelbänden (20 Bde.) - Hamburg: Hoffman und Campe Verlag, 1996-1999

Texte	Dokument	Beschreibung	Bd
1	LES/HIL.00000	Es waren Habichte in der Luft. Roman	1
...
1	LES/ALE.00000	Die Auflehnung. Roman	12
77	LES/ERZ.13001[-16022]	[Erzählung]	13-16
3	LES/ECH.17001[-]	[Schauspiele]	17
4	LES/HOR.18001[-18004]	[Hörspiele]	18
98	LES/ESS.19001[-20032]	[Essays]	19+20

Korpustext-Bibliographie

Die IDS-Korpustexte sind von jeher mit Quellennachweisen versehen, die bei der Anzeige gefundener Belege mit angezeigt werden. Allerdings waren sie in den früheren Korpora unstrukturiert. So wurde in den Neunzigerjahren ein Korpustext-Bibliografiemodell als eine zentrale Komponente des IDS-Textmodells entwickelt, das korpus-

übergreifende automatische Zugriffe auf die nunmehr einheitlich strukturierten umfangreichen Quelldaten mit folgenden Zielen erlaubt:

automatische virtuelle Korpuskomposition nach Autoren, Textarten, Entstehungszeiten, Sachgebieten usw.; vorkommende Textarten sind z.B.:

- | | | | |
|-----------------------|---------------------|------------|-------------------|
| - Abhandlung | - Aphorismus | - Aufsatz | - Autobiographie |
| - Bericht | - Biografie | - Brief | - Denkschrift |
| - Erlass | - Erzählung | - Essay | - Flugblatt |
| - Fußnote | - Forschungsbericht | - Gebet | - Gebrauchsanw. |
| - Gedicht | - Handzettel | - Hörspiel | - Interview |
| - Klappentext | - Leitartikel | - Märchen | - Nachruf |
| - Nachwort | - Parteiprogramm | - Petition | - Presseerklärung |
| - Produktbeschreibung | - Protokoll | - Rede | - Rezension |
| - Roman | - Schauspiel | - Tagebuch | - Vorspann |
| - Werbung | | | |

zu Ziff. 2.013:

IDS - Methoden der Korpusanalyse

zu:

aus: PERKKUHN, R; BELICA, C; al-WADI, D; LAUER, M; STEYER, K; WEISS, C, Korpustechnologie am Institut für Deutsche Sprache in: Schwitalla, J; Wegstein, W, Korpuslinguistik deutsch: synchron - diachron - kontrastiv. Würzburger Kolloquium 2003. Tübingen 2005. S. 57-70

(66) Beide COSMAS-Systeme bauen auf einer von der Arbeitsgruppe konzipierten, reichhaltigen, auf linguistische Bedürfnisse abgestimmten Suchanfragesprache auf und ermöglichen

Recherche in vordefinierten oder frei zusammengestellten virtuellen Korpora,

Suche nach Wortformen und -segmenten, nach Satz- bzw. Sonderzeichen und Zahlen,

Suche mit Hilfe von Suchmustern (z.B. Grundformen) und Anzeige aller im ausgewählten Korpus belegten Wortformen eines Suchmusters, so dass relevante Formen selektiert werden können,

Einbeziehung morphosyntaktischer Merkmale und früherer Rechercheergebnisse in eine Suchanfrage,

Verknüpfung der verschiedenen Bestandteile über logische Operatoren sowie Abstandsoperatoren, verschiedene Sortier- und Präsentationsmöglichkeiten, z.B. grafische Darstellung der chronologischen Sortierung,

Anzeige der Quellenangaben, der KWICs und der Volltexte, verschiedene Exportmöglichkeiten. ...

(67) 3.3 Lemmatisierung

Die Lemmatisierung ermöglicht, dass Flexionsformen, Zusammensetzungen und/oder sonstigen Wortbildungsformen ihre Grundformen zugeordnet werden können. In diesem Zusammenhang sind Grundformen

unflektierte Simplizia verschiedener Wortarten,

unflektierte Ableitungen und Komposita,

Wortbildungsmorpheme. ...

3.4 Kookkurrenzanalyse

Die Kookkurrenzanalyse ist eine korpusanalytische Methode zur Strukturierung von Belegmengen. Sie

ermöglicht das Aufdecken von signifikanten Regelmäßigkeiten bei der Verwendung von Wortkombinationen in den Korpora,

wertet mit Hilfe mathematisch-statistischer Analyse- und Clusteringverfahren den definierbaren Kontext eines vorgegebenen Suchobjekts in beliebigen virtuellen Korpora aus,

liefert Hinweise auf systematisches gemeinsames Auftreten von Wörtern (Partnerwörter, Kollokatoren) und ein Maß für deren Affinität (Kohäsion),

fasst Belege, die ähnliches Kookkurrenzverhalten des Wortes dokumentieren, zu Gruppen/Clustern zusammen,

strukturiert diese Belegmengen ggf. hierarchisch,

bietet eine entsprechende synoptische Präsentation der Belege,

erfasst neben binären Wortrelationen auch usuelle phrasale Muster bis hin zu (idiomatischen) Mehrworteinheiten.

Die Arbeitsgruppe für Korpustechnologie stellt der Öffentlichkeit die Kookkurrenzanalyse seit 1995 integriert in COSMAS zur Verfügung. Die Kookkurrenzanalyse ist auf beliebige COSMAS-Suchobjekte anwendbar mit

optionaler Lemmatisierung,

variabler Kontextgröße,

ggf. automatischer Fokussierung auf den Kontext mit dem stärksten Kohäsionswert,

variabler Zuverlässigkeit (d.h. Signifikanz des ersten Kollokators),

variabler Granularität (d.h. Signifikanz der Kollokatoren, die für die Ermittlung von Mehrworteinheiten berücksichtigt werden),

variabler Zuordnung von Belegen bei Mehrworteinheiten.

zu Ziff. 2.02:

Korpuslinguistik

zu: vprag7.5111

aus: R. ALBERT; C. J. KOSTER, Empirie in Linguistik und Sprachlehrforschung. Ein methodologisches Arbeitsbuch. Tübingen 2002.

(68) Ein Korpus ist eine Sammlung von Wörtern, Sätzen oder Texten, die üblicherweise als elektronische Datenbank vorliegt; es kann ein paar Hundert oder viele Millionen von Wörtern umfassen (gemeint hier als sogenannte *tokens*). Aber auch eine Sammlung von Sätzen oder Texten, die die Forscherin selbst für ihre Untersuchung (aus tatsächlich beobachtetem Sprachmaterial, nicht aus selbst erfundenen Beispielen) erstellt hat, ist ein Korpus. Korpora, die in elektronischer Form vorliegen und die man für eigene Untersuchungen benutzen kann, sind meist sehr viel umfangreicher als eine selbst zusammengestellte Sammlung; das NEGRA-Korpus der Universität des Saarlandes (Saarbrücken) z.B. umfasst 176.000 tokens (in 10.000 Sätzen), die aus deutschen Zeitungstexten (*Frankfurter Rundschau*) stammen. Andere Korpora sind noch erheblich größer, und natürlich gibt es Korpora für sehr viele verschiedene Sprachen der Welt.

(69) Manche Korpora sind von den Forscherinnen für ein bestimmtes Projekt zusammengestellt worden; z.B. zum Zweck der Erstellung oder Überarbeitung eines Wörterbuchs können aktuelle Belege für die Verwendung der Wörter gesammelt werden. Jemand, der sich für Arzt-Patient-Interaktion interessiert, kann sich ein Korpus von Transkripten selbst aufgenommener Arzt-Patient-Gespräche zusammenstellen. Wer sich für stilistische Unterschiede zwischen seriösen Tageszeitungen und Boulevardzeitungen interessiert, wird eine Sammlung von Artikeln aus beiden Arten von Zeitungen anlegen; wer sich für die Unterschiede zwischen Reklame für Autos im Internet und in Zeitungen interessiert, wird Anzeigen aus dem Internet und aus Zeitungen sammeln und sie zum Beispiel daraufhin vergleichen, wie viele technische Details sie enthalten.

(71) Die Adresse

<http://ftp.ids-mannheim.de/dsav/korpora/korpusliste.html> bietet eine Übersicht über Korpora des gesprochenen Deutschen und erläutert auch für die einzelnen Korpora, wie man Zugang zu ihnen bekommt. Etwas Ähnliches zum geschriebenen Deutschen gibt es natürlich auch: <http://www.ids-mannheim.de/kt/corpora.html>.

Viele der dort aufgeführten Korpora sind für eine begrenzte Zeit ohne Nutzungsgebühr zugänglich über die Adresse <http://www.ids-mannheim.de/kt/cosmas.html>. Die Zugriffszeit ist auf 60 Minuten beschränkt. Das kann ausreichen, wenn man bestimmte Wörter in ihren Satzkontexten sucht, es reicht aber sicher nicht, wenn die Suche nicht ausschließlich vom Computer erledigt werden kann, wenn man z.B. ein Wort nur in einer bestimmten Bedeutung sucht, die man selbst aus dem Kontext erschließt, wenn man die Belegstellen im Korpus findet.

Zeitlich unbeschränkt kann man das bereits erwähnte Korpus der Computerlinguisten der Universität Saarbrücken kostenlos nutzen; es enthält allerdings ausschließlich Zeitungstexte: <http://www.coli.uni-sb.de/sfb378/negra-corpus>.

Eine sehr nützliche Adresse für Transkripte gesprochener Sprache ist übrigens die Mailliste zur Gesprächsforschung, auf die sehr viele (72) Sprachwissenschaftlerinnen und Sprachwissenschaftler abonniert sind, die sich mit gesprochener Sprache beschäftigen, und über die man u.a. Transkriptkorpora zum Tausch anbieten und

nach bestehenden - auch unveröffentlichten - Transkriptkorpora fragen kann (mailliste@gesprächsforschung.de).

zu Ziff. 2.03:

Buchwelt + Internet

zu:

aus: DER SPIEGEL 35/2005

- Amazon will Hunderttausende Bücher digitalisieren; Google 15 Millionen (bis 2015; Kosten ca. 200 Mio. \$)
- Digitalisierungszentrum der Uni-Bibliothek Göttingen hat 4 Mio. Seiten digitalisiert und auf Server überspielt
- In 2000 elektronische Form einer der Gutenbergbibeln
- <http://www.gutenbergdigital.de>
- erlaubt Volltextsuche: Fundstellen in Papierbüchern. = eigentliche Neuerung
- "gemeinfreie" Bücher, deren Urheberrechte abgelaufen sind
- Kosten sollen durch Werbung aufgefangen werden: Fundstücke können - via *links* - in Bibliotheken ausgeliehen oder bei Verlagen/Buchläden gekauft werden
- Vorstoß von Google wirkt rüde: "systematischer Verstoß gegen das Urheberrecht". Derzeit Klärung.
- Probleme: Dominanz der englischsprachigen Literatur. Buchsuche könnte kostenpflichtig werden. Suchergebnisse können zugunsten zahlungswilliger Verlage manipuliert werden. Ranking-Methode von Google befördert Monokultur.
- Deutschland: technischer Rückstand (obwohl mal Vorreiter). Gemeinsames Portal fehlt, ebenso der Blick auf Benutzerbedürfnisse. Im Vordergrund stand "konservatorisches" Scannen.
- moderne Forschungsbibliothek = Standortfaktor und Aushängeschild
- Bereits heute wertet Google die Bestände von fünf Bibliotheken in den USA und Großbritannien aus und überlässt diesen im Gegenzug Digitalkopien. = Beispiel für Kooperation.
- <http://www.amazon.de> mit Suchfunktion "Search Inside"
- <http://www.print.google.com> mit Suchfunktionen "Google Print" und "Google Library" - derzeitiger Umfang nicht genau bekannt
- <http://gallica.bnf.fr> 70.000 Werke als Volltext aus der Bibliothèque nationale de France
- <http://www.gutenberg.org> und <http://gutenberg.spiegel.de> = "Projekt Gutenberg", nichtkommerzielle Internet-Datenbanken, ca. 16.000 Bücher, davon 2.000 auf Deutsch
- <http://www.hki.uni-koeln.de/retrodig/index.html> - Übersicht über 90 kleinere Digitalisierungsprojekte; ca. 17 Mio Seiten und Bilder
- <http://www.ulib.org> - "Million Book Projekt" der Carnegie Mellon University; 50.000 Bücher bislang; 10.000.000 geplant

zu Ziff. 2.04:

Analyse alphabetischer Schriften

zu:

aus: ALTMANN, G, Towards a theory of script: G. ALTMANN; F. FENG-XIANG (eds.), Analyses of Script. Properties of Characters and Writing Systems. Quantitative Linguistics. 63. Berlin 2008. S.149-164.

(160)

1. **Set up the inventory of phonemes** Do not adhere to historically or morphologically based orthographies; realize that an allophone exists only if there is no independent phoneme, e.g. in German there is a phoneme /ŋ/ ("Wange") written as <ng>; here <n> in "Bank" does not represent an allophone of /n/ but the phoneme /ŋ/. In the German Words "Grad" and "Grat" the last phonemes are identical. The number of phonemes in the inventory is |P|. One should rely on the work of phonemicians if their work is not morphological or historical. One should decide in advance whether phonemes occurring only in foreign words are parts of the inventory, or set up two different inventories. Write phonemes in slant lines //.
2. For each phoneme, write all **letters and graphemes** which can represent it. This is the most difficult task. As shown in step 1, assimilations usually "change" one "phoneme" in another. ... Write letters or graphemes in brackets <>.
3. For each phoneme, write the **number of representing graphemes**. For phoneme x it is G_x . Set up the inventory of graphemes G (each grapheme occurs in it only once, even if it represents several phonemes).
4. Count all G_x to get the **number of graphemes** in a language, $|G|$, i.e. ascertain the cardinal number of G .
5. Take several text samples and count the **frequency** of phonemes or letters or graphemes or all. To this purpose one can also use a corpus.
6. Compute the relative frequencies of phonemes/letters/graphemes and use them to compute the **entropies** of the frequency distributions. Use the Shannon entropy but try to compute also others ones if you like (there are about 40 different definitions of entropy).
7. Compute the **phonemic load** of letters, PL_x .
8. Compute the **graphemic load** of letters, GL_x .
9. Compute the **orthographic uncertainty** of each phoneme, U_x .
10. Compute the **mean grapheme size** from a list of graphemes set up in step 2 (each grapheme counted only once).
11. Compute **letter utility** (useful only in some languages).
12. Compute the **complexity** of all letters. Use the connection method and the fractal dimension or develop new methods.
13. Compute the **L-distinctivity** of letters and the **mean distinctivity** of the script. Propose new methods if necessary.
14. If you have students ready to fill in questionnaires, compute the (162) **ornamentality** of different letter types (a first step has been made for Latin script types in MS WORD, cf. Best & Altmann in this volume).

15. Fit the negative hypergeometric distribution to the (absolute) frequency counts. Then try to fit simpler **distributions** called Zipfian.
16. Take the known **entropy curves** for phoneme frequencies and localize your language/script on this curve.
17. Prepare a table containing the P -load and the frequency of a letter. Test whether there is a correlation between them and try to find the shape of the dependence (P -load is the independent variable). Display the points graphically < **P -load, Frequency**>.
18. Prepare a table containing the G -load and the frequency of a letter and proceed as in step 17, < **G -load, Frequency**>.
19. Prepare a table containing the graphemic utility of a letter and its frequency. Proceed as in step 17, < **G -utility, frequency**>.
20. Prepare a table containing the complexity and the distinctivity of a letter. Proceed as in step 17, <**Complexity, distinctivity**>.
21. If you have computed ornamentality, prepare a table containing ornamentality and complexity of letters. Proceed as in step 17, <**Ornamentality, Complexity**>.
22. If you have computed ornamentality, prepare a table containing ornamentality and distinctivity of letters. Proceed as in step 17, <**Ornamentality, Distinctivity**>.
23. Present **distinctivity** as a function of **ornamentality and complexity**.

2.1 Verschiedene Modalitäten der Eingabe

Literatur

- FLEISSNER, P; HOFKIRCHNER, W; MÜLLER, H; POHL, M; STARY, C
(ed./eds.): Der Mensch lebt nicht vom Bit allein..
Informationen in Technik und Gesellschaft. Frank-
furt / M ²1997: Peter Lang // I.2.0
- HAMMÖHNER, R: Offene Hypertextsysteme: das Konstanzer Hyper-
textsystem (KHS) im wissenschaftlichen und techni-
schen Kontext. Schriften zur Informationswissen-
schaft 32. Konstanz 1997: UVK, Univ.-Verlag // I.2.5
- HEYN, M: Zur Wiederverwendung maschinenlesbarer Wörterbücher.
Eine computergestützte metalexikographische Studie
am Beispiel der elektronischen Edition des "Oxford
Advanced Learner's Dictionary of Current English".
Lexicographica. Series Maior 45. Tübingen 1992 //
I.2.7
- JARAUSCH, K H; ARMINGER, G; THALLER, M: Quantitative Methoden in
der Geschichtswissenschaft. Eine Einführung in die
Forschung, Datenverarbeitung und Statistik. Darm-
stadt 1985 // I.2.5
- KICHUK, D: Metamorphosis: Remediation in *Early English Books
Online (EEBO)*. *Literary and Linguistic Computing*
22 / 3 (2007) 291-303.
- KLEIN, H: Computer gestützte Inhaltsanalyse mit INTEXT. Aktu-
elle Medien- und Kommunikationsforschung 8. Münster
1996 // I.2.7
- LUDWIG, H-W: EDV für Literaturwissenschaftler. Arbeits- und Pro-
grammiertechniken für den PC. Literaturwissenschaft
im Grundstudium 16. Tübingen 1991: Narr // I.2.5
- ZWEIG, R W: Lessons from the 'Palestine Post' Project. *Literary
and Linguistic Computing* 13 (1998) 89-95.

zu Ziff. 2.11:

Datenaquisition

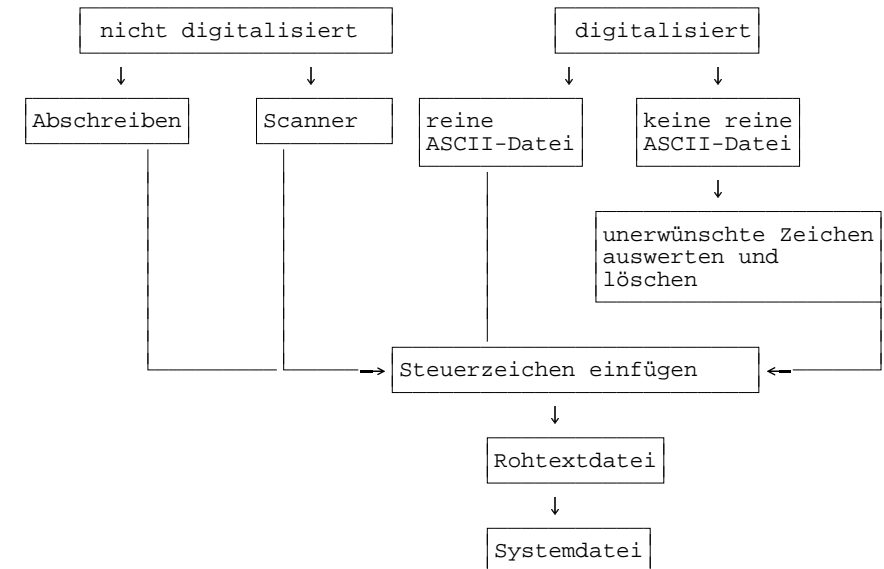
zu:

aus: H. KLEIN, Computerunterstützte Inhaltsanalyse mit INTEXT.
Münster 1996.

(38f) Unter Datenaquisition soll im folgenden die Beschaffung von
Texten bezeichnet werden, die bereits digitalisiert vorliegen oder
digitalisiert werden können. Vor einer Digitalisierung muß aller-
dings geklärt werden, ob dadurch Urheberrechte verletzt werden.

Die Texte können prinzipiell in mehreren Formen vorliegen:

1. nicht digitalisiert (z.B. gedruckt)
2. digitalisiert als reine Textdatei (ASCII-Datei)
3. digitalisiert als Datei mit unerwünschten Zeichen (Druck-
steuerzeichen) ...



aus: H. WIEGMANN, Und wieder lächelt die Thrakerin. Zur Geschichte
des literarischen Humors. Frankfurt/M 2006. S. 53:

Zur altdeutschen Dichtung: "Die frühesten Schriftzeichen sind in
runen (Ursprungsbedeutung Rune = Geheimnis) überliefert, wie man
die germanischen Schriftzeichen nennt.

Unser deutsches Wort "lesen" kommt ja vom Wort "auflesen". Der
runenkundige Drude des germanischen Stammes befragte vor einer
wichtigen Entscheidung die Götter, warf die teils mit Runen verse-
henen Buchenholzstäbchen (daher das Wort Buchstabe!) rückwärts
über die Schulter und las sie dann auf im Sinne von deutendem Auf-
lesen (daher das Wort lesen!), las damit auch aus der Zufallslage
der Buchenstäbchen bzw. ihrer gottgewollten Anordnung die sinnge-
bende Deutung ab, ob nun etwa Krieg oder Frieden gelten sollte.

zu Ziff. 2.12:

Optical Character Reading (OCR) bei arabischer Schrift?

zu:

- Arabische Schrift verläuft von rechts nach links
 - arabische Schrift ist kursiv, Buchstaben können miteinander verbunden sein und unterscheiden sich deshalb im Aussehen, abhängig von der Stellung im Wort (Anfang, Mitte, Ende oder isoliert):
 - nicht alle Buchstaben eines Wortes werden miteinander verbunden, deshalb kann ein Wort aus mehreren verbundenen Teilen bestehen
 - einige arabische Buchstaben haben Diakritiken (Punkte oberhalb oder unterhalb des Buchstabens)
 - arabische Buchstaben können besondere Zeichen haben, die zu einer besonderen Aussprache führen.
- ا ب ر ا ه ي م
ا ب ر ا ه ي م
- hinzu kommen alle weiteren, üblichen Merkmale von Druckschriften, wie unterstrichen, **dick**, *italic*, usw.

Handschrifterkennung (nach H. Amiri)

Zwei Schreibstile lateinischer Schrift: Block- oder Fließschrift

Aspekte der off-line-Erkennung (Erkennung bereits geschriebener Schrift [im Gegensatz zur on-line-Erkennung, die während des Schreibens erkennt]):

1. Erkennung isolierter, positionsgebundener Einzelzeichen
2. Erkennung von Blockschrift (z.T. verklebte Einzelzeichen, variable Abstände)
3. Fließschrifterkennung
4. Erkennung gemischter Fließ- und Blockschrift

Aufsteigender Schwierigkeitsgrad. Segmentierungsproblem: Zerlegung eines Wortes in Einzelzeichen entlang der Horizontalen (bei lateinischer Schrift). Ziff.1 $\hat{=}$ starke Einschränkungen bei der Anwendung (Benutzung von Kästchen in Blindfarben). Aber: Bei Ziffernbeiträgen (z.B. Eurocheck) ist mit verklebten Ziffern zu rechnen. Die relative Position der Ziffern zueinander ist dann nicht bekannt. Vorverarbeitung durch Segmentierungsalgorithmen notwendig. Fließschrift: Segmentierung gelingt dem Menschen oft nur mit dem Wissen um das geschriebene Wort. *Ohne Segmentierung keine Erkennung, ohne Erkennung keine Segmentierung.* Verschiedene Strategien.

Verwendung expliziter Segmentierung:

Zunächst Erarbeitung von u.U. mehreren Segmentierungsalternativen (in gerichtetem Graphen erfaßt); dann Abgleich z.B. mit einem Lexikon.

Verwendung impliziter Segmentierung:

Jede mögliche Aufteilung des Eingangsbildes = potentielle Segmentierung in Einzelzeichen. Überführung des Schriftzuges in eine Folge von Merkmalen, Merkmalsvektoren oder Symbolen. Aus der Sicht der Erkennung besteht ein Hidden Markov Modell aus mehreren Zuständen, von denen jeder für bestimmte Ereignisse entlang dieser Folge steht. Jeder Zustand kann dabei für jedes Merkmal oder Symbol einer Folge eine Bewertung liefern, die anzeigt, wie gut dieses zum jeweiligen Zustand "paßt". Zwischen den Zuständen existieren Übergänge, die für jeden Zustand mögliche Folgezustände festlegen. Gängige Suchalgorithmen bearbeiten die Merkmal-/Symbolfolge dann von links nach rechts und bestimmen dabei für ein HMM diejenige Folge von Zuständen, die am besten zu der Symbolfolge paßt.

Verwendung neuronaler Netze: Diese liefern an jeder horizontalen Position des Schriftzuges eine Bewertung für alle Buchstaben des Alphabets. Weiterverarbeitung mit HMM.

Verfahren ohne Segmentierung (Ganzworterkennung):

Statt verketteter Einzelzeichen werden ausschließlich Ganzwortmodelle verwendet. Lokale Merkmale + komplexe Suchstrategie (wie oben).

Lexikalisches Wissen:

Bei allen Verfahren zur Handschrifterkennung ist unbestritten, daß eine zuverlässige Erkennung nur unter Einbeziehung von Kontextwissen möglich ist, mindestens also des Wissens um gültige Wörter.

zu Ziff. 2.121:

SHARK - Shorthand Aided Rapid Keyboarding

zu:

aus: DIE ZEIT Nr. 36/2005

SHARK hat drei Vorteile:

die Schreibbewegung geht flüssiger von der Hand als das Antippen von Buchstabenfeldern

ist schneller und effizienter als Handschrift

und leichter zu lernen als traditionelle Stenoschriften

Tatsächlich kombiniert SHARK das Beste aus allen drei Schreibtechniken zu einem neuen System. Herzstück ist ein Buchstabenfeld, das an eine gequetschte Computertastatur erinnert. Doch anstatt die Zeichen nacheinander anzutippen, zieht der Nutzer auf einem berührungsempfindlichen Display einen Plastikstift über die Felder, ohne abzusetzen.

Die Software im Hintergrund errechnet aus der Kombination der überstrichenen Buchstaben das wahrscheinlichste Wort.

Ein Beispiel: Berührt man in einem Rutsch die Buchstabenfolge "wurtsieteing", bleibt das englische *writing* (schreiben) übrig.

Die durchgehende Handbewegung sei wichtig, sagt Z., weil sie dem Schreiben mit der Hand entspreche.

Wird das Wort nicht eindeutig erkannt, blendet die Software die anderen Möglichkeiten ein, der Nutzer kann wählen.

Die Anordnung der Buchstaben wurde zunächst für die englische Sprache so optimiert, dass vor allem häufig vorkommende kurze Wörter oder Buchstabengruppen - wie *the* oder die Endung *-ing* - mit einer knappen Handbewegung geschrieben werden können. Darin ähnelt das Verfahren der Stenografie.

Im Englischen machen die 100 häufigsten Wörter über 40 Prozent aller Texte aus. Sie bestehen in der Regel aus nicht mehr als vier, fünf Buchstaben, sodass ein Anfänger schon nach kurzer Zeit einen guten Teil intuitiv schreiben kann. Tests hätten gezeigt, dass geübte Nutzer auf bis zu 70 Wörter in der Minute kommen. Zum Vergleich: die durchschnittliche Geschwindigkeit für Handschrift liegt bei 15 Wörtern pro Minute.

Ohne Lernphase lässt sich SHARK allerdings ebenso wenig meistern wie andere Computerschreibsysteme, etwa die kryptischen Kürzel der Buchstabenerkennung Graffiti von Palm.

Tastatur und Software an die Besonderheiten anderer Sprachen anzupassen sei relativ einfach.

zu Ziff. 2.13:

Elektronischer Text

zu:

aus: HAMMWÖHNER, R, Offene Hypertextsysteme: das Konstanzer Hypertextsystem (KHS) im wissenschaftlichen und technischen Kontext. Schriften zur Informationswissenschaft 32. Konstanz 1997.

(30) Elektronische Speichermedien, aber auch Mikrofiche oder -film bedingen, daß ein Zugriff auf die Speicherinhalte und damit auch die Texte nur unter Vermittlung von Geräten stattfinden kann. Im Fall von mikroverfilmten Texten ist eine optische Vergrößerung erforderlich. Der Leser muß die Optik evtl. fokussieren und dann einen Bildausschnitt wählen. Der Abbildprozeß elektronischer Medien ist komplexer, da hier ein mehrfacher Recodierungsprozeß erforderlich ist. Physische Gravuren oder magnetische Ladungen des eigentlichen Speichermediums werden als binäre Zahlenfolgen interpretiert. Teilsequenzen dieses Bitstroms werden als Zeichen gedeutet, die dann nach einer weiteren Codierung und Formatierung auf einem externen Medium, etwa einem Bildschirm oder einem Ausdruck, zur Darstellung kommen. Es ist beim Lesen elektronischer Texte also immer zwischen einer primären und einer technisch vermittelten Repräsentation eines Textes zu unterscheiden.

Diese Distanzierung des Lesers, aber auch des Autors von der physischen Repräsentation des Texts, sowie die Existenz gleichzeitiger Abbilder für mehrere Leser nimmt (Landow 92, S. 18) zum Anlaß, von virtuellen Texten zu sprechen. Der Unterschied zwischen virtuellen und kopierten Texten scheint hier noch nicht vollständig geklärt zu sein. Was unterscheidet einen Text, der als Computerdruck vor dem Leser liegt, von demjenigen, der ihm als Fernkopie oder Fax zugeschickt wurde? Im ersten Fall ist das Original nicht direkt zugreifbar, weil es elektronisch ist, im zweiten deshalb, weil es nicht räumlich anwesend ist. Der signifikante Unterschied zwischen virtuellem Text und Kopie ist der, daß die Kopie durch einen Abbildprozeß entsteht, der nur optische Kriterien kennt. Der virtuelle Text kann aus seinem Urbild aber anhand komplexer Berechnungsvorgänge entstehen, die sowohl semantische als auch pragmatische Kriterien berücksichtigen. Der Text präsentiert sich dann anders, je nachdem, ob er zum Nachschlagen oder Argumentieren verwendet wird.

zu Ziff. 2.14:

Elektronische Texte / Probleme

zu: vprag7.5136

aus: FLEISSNER, P HOFKIRCHNER, W, u.a., Der Mensch lebt nicht vom Bit allein... Information in Technik und Gesellschaft. Frankfurt/M 21997. //I.2.0

(244)

- Standards für einen vernünftigen Gebrauch des Netzes sind noch nicht voll entwickelt. Zum Unterschied von Printmedien und bei wissenschaftlicher Publikationstätigkeit (Lektorate, Reviewing) bilden sie sich bei elektronisch vermittelter Kommunikation erst heraus. Die Verwendung von wissenschaftlichen Netzen in der Bundesrepublik Deutschland zum Transfer von pornographischen Farbbildern und Mailboxsysteme mit ähnlichen Inhalten sind öffentlich bekannt gewordene Beispiele dafür.
- Die Feststellung der Urheberschaft bei einem komplexen, womöglich unter Einbeziehung unterschiedlicher Medien oder mit Hilfe von online-Dokumenten erzeugten Produkt und ihr sachlicher und rechtlicher Schutz stecken in den Kinderschuhen. Damit ist die Kontrolle und daher die korrekte Remuneration einer Leistung der AutorInnen erschwert. In den USA stehen bisher zur Wahrung der Eigentumsrechte an gesetzlichen Titeln nur das Copyright (für Printmedien und Software), Patente (auf der Bundesebene), "Trademark" (geschützte Markenbezeichnung) auf Länderebene, und "unfaire Konkurrenz" (bei den Gerichten) zur Verfügung. Weitere, der Netzsituation besser angepasste Formen werden notwendig.
- Das Authentifikationsproblem, das im Bereich der Datensicherheit derzeit auf der Ebene der Netz- und Systembenutzer mittels Zugriffskontrollen und Paßwörtern einer Lösung zugeführt wird, stellt sich genauso für die LeserInnen von Netzprodukten. Wie können sie sicher sein, von wem das Produkt tatsächlich stammt?
- Die Entscheidung, wer vorhandene Produkte verändern, teilweise kopieren, zu neuen Produkten zusammenfassen, neu aussenden und - vor allem - verkaufen darf, erfordert eine spezielle rechtliche Fixierung. Dies umso mehr, je wichtiger netzvermittelte Information in der Praxis wird.
- Eine besondere Problematik bietet der Datenschutz, also der Schutz der Individuen vor dem Mißbrauch von personenbezogenen Informationen, immer dann, wenn - wie in den Netzen - private Kommunikation einfach und schnell in den öffentlichen Bereich (etwa in die NetNews) transportiert und vervielfältigt werden kann. Dazu kommt die Aufzeichnung von Benutzerdaten (wer hat welches Dokument wann gelesen), die ein Persönlichkeitsprofil des Lesers zu erstellen erlauben.
- Bei weiterem Wachstum der Zahl der Netzteilnehmer und der Netzdienste wächst die Gefahr der Informationsüberlastung, der Fehlinformation und der Zusendung unqualifizierter Äußerungen.
- Bei einem Wildwuchs von Informationspaketen, die ins Netz eingespielt werden, wird das Suchen einer bestimmten Art von Information immer schwieriger. Die Katalogisierung und Archivierung der vorhandenen Produkte verdient besondere Aufmerksamkeit, ebenso die Zugänglichkeit dieser Kataloge und Archive.
- Wie groß ist die Lebensdauer der heute in Archiven verwendeten Datenträger (CDs, Platten, Bänder etc.) und der entsprechenden Hardware?

- Wer hat die Verantwortung für die Informationen, die im Netz zirkulieren? Ist es der Netzbetreiber, der Anbieter von Netzdiensten, der Urheber oder der Absender der Nachricht? Wer trägt mögliche Verluste, die durch einen Ausfall des Netzes entstehen? (245)
- Wie wird die Zugänglichkeit generell geregelt? Hat jeder Bürger das Recht auf einen Zugang zur Netzinformation? Muß er dafür bezahlen? Trägt (wie bisher für die wissenschaftlichen Netze) der Staat die Kosten? Wird es Bürger erster und zweiter Klasse geben? Wird die informative (und daraus wahrscheinlich resultierende soziale) Ungleichheit innerhalb der WissenschaftlerInnen-gemeinde zu- oder abnehmen?

2.2 Speicherung des Textes**Literatur**

- BADER, W: SGML-ähnliche Textauszeichnung in der TUSTEP-Umgebung. Historical and Social Research 18 (1993) 92-112.
- BADER, W: Was ist die Text Encoding Initiative (TEI)? in: KAMZELAK, R (ed./eds.): Computergestützte Text-Edition. Beihefte zu editio, Band 12. Tübingen 1999 9-20.: Niemeyer // I.2.1
- BRYAN, M: SGML, An Author's Guide to the Standard Generalized Markup Language. Wokingham, England 1988 // I.7.2
- HARDMEIER, C; TALSTRA, E; GROVES, A (ed./eds.): Stuttgarter elektronische Studienbibel (SESB) CDROM (Texte, Datenbanken, Suchprogramm). Stuttgart 2004 // I.2.1
- HERWIJNEN van, E: Practical SGML. Dordrecht 1990 // I.7.2
- HICKEY, R: Corpus Presenter. Software for language analysis with a manual and A Corpus of Irish English as sample data. Amsterdam 2003 ISBN 90-272-2607-5: John Benjamins // I.2.1
- LI, N: Computational Research of Chinese Writing System-Han4-Z4. Literary & Linguistic Computing 9,3 (1994) 225-234.
- LUDWIG, H W; ROMMEL, T: Studium Literaturwissenschaft. Arbeitstechniken und Neue Medien. UTB 2332. Tübingen 2002 ISBN 3 8252 2332 9: A. Francke // I.2.5
- ORE, C-E; EIDE, O: TEI and cultural heritage ontologies: Exchange of information?. Literary & Linguistic Computing 24/2 (2009) 161-172.
- REHM, g; SCHONEFELD, O u.a.: Sustainability of annotated resources in linguistics: A web-platform for exploring, querying, and distributing linguistic corpora and other resources. Literary & Linguistic Computing 24/2 (2009) 193-210.
- SGARBAS, K N; FAKOTAKIS, N D u.a.S: A PC-KIMMO-based Bi-directional Graphemic / Phonetic Converter for Modern Greek. Literary and Linguistic Computing 13 (1998) 65-75.
- ZILLIG, B L P: TEI Analytics: converting documents into a TEI format for cross-collection text analysis. Literary & Linguistic Computing 24/2 (2009) 187-192.

2.21 Seitenbeschreibung/-format - Text Encoding Initiative**Literatur**

- RAHTZ, S; SCHREIBMAN, S . Literary and Linguistic Computing 24/3 (ed./eds.): [Special Issue 'Selected papers from Text Encoding Initiative'] (2009) 249-372.

zu Ziff. 2.211:

TEI-DTD/Interpretierende Elemente

zu:

aus: W. BADER, Was ist die Text Encoding Initiative (TEI)? in: KAMZELAK, R (ed.), Computergestützte Text-Editon. Tübingen 1999: Niemeyer. S.9-20.

(16) Es gibt prinzipiell zwei Arten von Elementen, mit denen ein Text beschrieben wird. Zum einen die konstruktiven Elemente, die die Struktur des Textes beschreiben, also 'front - body - back', oder innerhalb von 'body' die Einteilung in 'div' und 'seg', und das kann natürlich weiter unterteilt und feiner gegliedert werden. Diese Elemente konstituieren in ihren Zusammenhang den vollständigen Textbestand. Hängt man den Inhalt all dieser Elemente einer bestimmten Ebene zusammen, so ergibt sich jeweils der vollständige Text (die TEI-Richtlinien sprechen von 'chunk level elements').

Zum anderen sind es die sogenannten 'phrase level elements'. Das sind Elemente, die verschiedene kleine Teile des Textes interpretieren. Sie können beliebig innerhalb des Textes vorkommen, beschreiben aber niemals den kompletten Text - auch nicht wenn man alle aneinanderreicht -, sondern immer nur einzelne Teile. Z.B. ist die Auszeichnung eines bestimmten Textteils als Name ein solches interpretierendes Element. Ein oder zwei Wörter werden durch dieses Element als Name interpretiert. Namen können fast an beliebiger Stelle im Text vorkommen, sie werden aber nur in extremen Ausnahmefällen aneinandergereiht den ganzen Text bilden.

Solche interpretierenden Auszeichn(ung,H.S.)en sind allen Texten gemeinsam. Alle Texte sind dadurch ähnlich, daß sie 'Kleinteile' enthalten, die interpretierend ausgezeichnet werden können. Daher gehören auch diese Elemente wie Tomatensauce und Käse zum Kernbestand der DTD. Es sind dies im einzelnen u.a. die Elemente für Hervorhebungen (betont='emph', hervorgehoben='hi', Fremdwort='foreign', terminus technicus='term', Erklärung='gloss'), für Anführungen (Rede='q', Zitat='cit', Übertragene Bedeutung='socalled'), für Eigennamen (Name='name', Hinweis auf ein Individuum='rs' [reference string]), für Datums- und Maßangaben (Datum='date', Zeit='time', Anzahl='num', Maßeinheit='measure'), für editorische Eingriffe (korrigiert='corr', ursprünglich='orig', Lücke='gap'), für Abkürzungen (Abkürzung='abbr', Expandierte Abkürzung='expan'), für Querverweise (Referenz='ref') und Registereinträge ('index'). Solche Auszeichnungen sind prinzipiell auf alle Texte anwendbar.

An dieser Art von Auszeichnungen läßt sich die Eingangsthese, daß jede Art von Textauszeichnungen immer schon Interpretation ist, nochmals verdeutlichen. Ein in der gedruckten Vorlage kursives Wort kann z.B. einfach als 'hi'=hervorgehoben bezeichnet werden, oder noch neutraler als Segment abgeteilt werden, dem man das Attribut 'Schriftart=kursiv' zuspricht. Auf dieser Ebene liegt die Interpretation darin, daß ein Schrifttyp erkannt wurde (was ggf. bei einem einzelnen Buchstaben sich nicht ganz sicher erkennen läßt). Bei der Arbeit am Text kann aber erkannt werden, und das ist wirklich Interpretation, daß der Schrifttyp kursiv z.B. für Fremdwörter, für Namen, für Termini technici oder für Zitate verwendet wird, und entsprechend werden solche Textstellen mit 'foreign', 'name', 'term' oder 'cit' ausgezeichnet. Der Vorteil liegt auf der Hand: für die reine Weiterverarbeitung auf Papier lassen sich diese Diversifizierungen leicht auf die eine Typographie 'kursiv' reduzieren ('Downkonvertierung'). Für ein Abfragesystem steht aber die differenziertere Information zur Verfügung. In der Balance zwischen späterem Nutzen und jetzigem Aufwand bei der Er-

fassung muß die Dokumentenanalyse im Vorfeld des Projekts festlegen, welche dieser Unterscheidungen durchgeführt werden soll.

zu Ziff. 2.212:

TEI/SGML

zu:

aus: BADER, W, Was ist die Text Encoding Initiative (TEI)? in: KAMZELAK, R (ed.), Computergestützte Text-Editon. Tübingen 1999: Niemeyer. S.9-20.

core tag set: Textkorpus ('body'), einleitende Materialien (Titel-seite, Inhaltsverzeichnis, Widmung, Abstract, Vorwort - Element 'front'), Anhänge (Register, Glossar, Bibliographie, Index, Endnoten - Element 'back'). Metainformationen (Element 'teiHeader')³

Textkörper: unterschiedliche Gattungen haben unterschiedliche Grobstrukturen.⁴ Gemeinsamkeit: Tatsache der Unterteilung. TEI stellt Elemente 'div' (division, Unterteilung) und 'seg' (segmentation, Teilung) bereit. Der Individualität wird man durch Setzen eines Attributs gerecht (z.B. <div type='Szene'>).

Globale Attribute: Jedes Element hat grundsätzlich vier Attribute: 'ID' (identity) zur eindeutigen Identifizierung, 'NAME' zur Identifizierung innerhalb einer Hierarchie, 'TYPE' = Art des Elements, 'LANG' (language) Sprache des Elementinhalts.⁵

Konstitutive Elemente: Alles, was durch 'front - body - back' eingeschlossen ist, ergibt den vollständigen Textbestand ('chunk level elements').

Elemente der Detailinterpretation: Hervorhebungen,⁶ Anführungen,⁷ Eigennamen,⁸ Datums- und Maßangaben,⁹ editorische Eingriffe,¹⁰ Abkürzungen,¹¹ Querverweise (Refe-

³ Entspricht im Buch: Titelseite und Impressum, bei elektronischen Texten näherhin: Autor und Beschreibung des Dokuments, Encodierer, Titel der Vorlage (falls es sie gab), genaue bibliographische Daten (Element 'fileDesc'); Methode der Textauszeichnung mit der Zahl aller Elemente - hilft bei Integritätsprüfung (Element 'encodingDesc'); in welcher/n Sprache/n ist das Dokument erfasst (Element 'profileDesc'); wann, wer, was, warum korrigiert, hinzugefügt und verändert hat (Element 'revisionDesc').

⁴ Prosa ist in Abschnitte und Kapitel aufgeteilt. Bei älteren Autoren Aufteilung in Bücher, die oft modernen Kapiteln entsprechen. Drama: Akt und Szene. Briefedition: Briefe und Regeste. Lexikon: Stichwörter etc.

⁵ Durch 'ID' können Zeiger ('IDREF', also Referenz = Verweis auf 'ID') auf das Element verweisen und so Analysen und Zusammenhänge herstellen. 'TYPE' ist wichtig zur inhaltlichen Charakterisierung. 'LANG' löst sowohl Zeichensatzprobleme wie auch Verwechslungen bei der Indexerstellung.

⁶ betont='emph', hervorgehoben='hi', Fremdwort='foreign', terminus technicus='term', Erklärung='gloss'

⁷ Rede='q', Zitat='cit', übertragene Bedeutung='socalled'

⁸ Name='name', Hinweis auf ein Individuum='rs' [reference string]

⁹ Datum='date', Zeit='time', Anzahl='num', Maßeinheit='measure'

¹⁰ korrigiert='corr', ursprünglich='orig', Lücke='gap'

¹¹ Abkürzung='abbr', expandierte Abkürzung='expan'

renz='ref'), Registereinträge ('index'). Ein so aus-gezeichneter Text wird bei der Papierwiedergabe z.B. die Elemente 'hi', 'term' und 'foreign' alle mit "kursiv" deuten - die Downkonvertierung reduziert die differenzierte Information. Für ein Abfragesystem, die Registererstellung usw. stehen aber die benötigten Differenzierungen zur Verfügung.

Basisgruppe (zur Kerngruppe hinzugewählt): Abhängig von der Textgattung. Die TEI-DTD stellt bereit:
'TEI.prose' für Prosatexte
'TEI.verse' für poetische Texte
'TEI.drama' für Theaterstücke¹²
'TEI.spoken' für gesprochene Sprache¹³
'TEI.dictionaries' für Lexika
'TEI.terminology' für terminologische Datenbanken
'TEI.general' und 'TEI.mixed' für gemischte Texte

additional tag set: Zu den bisherigen Gruppen für die spezifische Interessenlage, mit der man Texte beschreibt.
'TEI.linking'=Querverweise, Unterteilung, Zuordnung
'TEI.analysis'=einfache Analysen
'TEI.fs'=strukturierte Analysen
'TEI.certainty'=Informationen über Sicherheitsgrad der Auszeichnungen
'TEI.transcr'=Manuskripte, analytische Bibliographien, Transkription von [Primär]Quellen
'TEI.textcrit'=kritische Editionen
'TEI.names.dates'=Namen und Kalenderdaten
'TEI.nets'=Graphen, Baumstrukturen, Netzstrukturen
'TEI.figures'=Graphiken, Figuren, Illustrationen, Formulare
'TEI.corpus'=Sprachkorpora

Gewährleistet: Austauschbarkeit und Kommunikabilität, Erweiterbarkeit, Anregung und Reflexion der eigenen Arbeit, Ergänzierbarkeit, Zwang zur Metadokumentation sichert das Überleben der Dokumente, macht sie zitierfähig. TEI = SGML-Anwendung.

¹² Man muß Sprecher oder Regieanweisungen auszeichnen können.

¹³ Wiedergabe des zeitlichen Ablaufs (auch Parallelität), Hinweise auf Verständlichkeit/Unverständlichkeit, Gestik.

zu Ziff. 2.213:

TEI-DTD/Basisgruppen/Verschiedenheit der Texte

zu:

aus: W. BADER, Was ist die Text Encoding Initiative (TEI)? in: KAMZELAK, R (ed.), Computergestützte Text-Edition. Tübingen 1999: Niemeyer. S.9-20.

(17) Verwendet man die TEI-DTD, so hat man auf jeden Fall die Elemente der Kerngruppe zur Verfügung, und wird sie auch brauchen, da sie auf alle Texte zutreffen. Zur Kerngruppe hinzu wählt man sich eine Basisgruppe. So wie die Füllung einer Schwäbischen Maultasche vegetarisch aus Gemüse sein kann, aus Fleischbrät besteht oder aus Fisch, was frei zu wählen ist, es aber keine Maultasche ohne Füllung gibt, und auch keine mit mehreren Füllungen, so ist es auch mit den Basisgruppen: man braucht genau eine.

Die Basisgruppe der TEI-Elemente ('base tag set') wählt man passend zur Textart oder Textgattung, die beschrieben werden soll. Jede Textgattung fordert unterschiedliche Elemente: bei einem Drama muß man z.B. Sprecher oder Regieanweisung auszeichnen können, bei gesprochener Sprache, die in einer Transkription erfaßt wird, sind Mechanismen zur Wiedergabe des zeitlichen Ablaufs, Hinweise auf Verständlichkeit oder Unverständlichkeit, Hinweise auf Gestik notwendig. Jede Textgattung, braucht so ihre spezifischen Elemente, die man bei anderen Texten nicht braucht. Die vorbereitende Dokumentenanalyse stellt fest, welche Art von Text vorliegt, entsprechend wird dann die Basisgruppe ausgewählt.

Die TEI-DTD stellt die folgenden Basisgruppen bereit: 'TEI.prose' für Prosatexte, 'TEI.verse' für poetische Texte und sonstige Texte mit Verseinteilung oder Sinnzeilen, 'TEI.drama' für Theaterstücke, 'TEI.spoken' zur Erfassung gesprochener Sprache, 'TEI.dictionaries' für die Wiedergabe gedruckter Lexika, 'TEI.terminology' zum Aufbau terminologischer Datenbanken, 'TEI.general' und 'TEI.mixed' sind zusätzliche Spezialfälle, wenn es sich um gemischte Texte handelt.

Die Zusatzgruppen: Gleiche Texte können verschieden beschrieben werden

So wie jede Pizza, die konstitutiv aus einem Boden in verschiedener Backart (Basisgruppen) und Tomatensauce mit Käse (Kerngruppen) besteht, mit beliebig vielen weiteren Zutaten in nahezu beliebiger Kombination belegt werden kann, so kann man sich aus der TEI-DTD neben der Kerngruppe, die man immer hat, und der einen Basisgruppe, die man für die zu beschreibende Textgattung auswählt, beliebig viele Zusatzgruppen ('additional tag set') auswählen für die Hinsicht oder Hinsichten, mit der man den Text beschreibt. Zusatzgruppen gibt es für 'TEI.linking' (Auszeichnungen für Querverweise, Unterteilung und Zuordnung), 'TEI.analysis' (Auszeichnungen für einfache Analysen), 'TEI.fs' (=feature structures', Auszeichnungen für strukturierte Analysen), 'TEI.certainty' (Auszeichnungen für Informationen über die Sicherheit und Wahrscheinlichkeit der Auszeichnungen), 'TEI.transcr' (Auszeichnungen für Manuskripte, analytische Bibliographien und die Transkription von [Primär]Quellen), 'TEI.textcrit' (Auszeichnungen für kritische Editionen), 'TEI.names.dates' (Auszeichnungen für Namen und Kalenderdaten), 'TEI.nets' (Auszeichnungen für Graphen, Baumstrukturen, Netzstrukturen), 'TEI.figures' (Auszeichnungen für Graphiken, Figuren, Illustrationen, Formulare), 'TEI.corpus' (Auszeichnungen für Sprachkorpora).

zu Ziff. 2.214:

SGML/TEI - Praktisches Projektbeispiel

CHRISTMANN, R, Books into Bytes: Jacob and Wilhelm Grimm's *Deutsches Wörterbuch* on CD-ROM and on the Internet. *Literary and Linguistic Computing*, Vol.16, No.2, 2001, S.121-133.

XML/TEI - Praktisches Projektbeispiel

RENNIE, S, *The Electronic Scottish National Dictionary (eSND)* : Work in Progress. *Literary and Linguistic Computing*, Vol.16, No.2, 2001, S.153-160.

Handbücher, DTDs und Beschreibung von TEI

<http://www.uic.edu/orgs/tei>

2.22 Zeichensätze/Sonderzeichen/SGML

Literatur

PANAS, E: Indexes of lexical richness can be estimated consistently with knowledge of elasticities: some theoretical and empirical results in: Grzybek, P; Köhler, R (ed./eds.): Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 521-532.: Mouton de Gruyter // I.2.5

zu Ziff. 2.2201:

Codierung von Texten

zu:

aus: Dossier von <kpa>Martin Schindele<kpe>

Grundproblem jedes Codierungsverfahrens ist aber, dass die Formenvielfalt graphischer Zeichen in der Regel grösser ist als die Bandbreite, die ein Code von der Bauart des ASCII-Codes für die direkte 1:1-Codierung bietet.

2 DIE "DRUCKBAREN" ZEICHEN DES 7-BIT-ASCII-CODES

2.1 Definition.- Die Menge der druckbaren Zeichen des 7-Bit-ASCII-Codes ist die folgende Liste von graphischen Zeichen, zu denen hier jeweils der zugehörige Zahlenwert in verschiedenen Zahlssystemen angegeben wird:

graph. Zeichen	Dezimalw.	Binärw.	Hexadezimalw.
BLANK	32	00100000	20
!	33	00100001	21
"	34	00100010	22
#	35	00100011	23
\$	36	00100100	24
%	37	00100101	25
&	38	00100110	26
'	39	00100111	27
(40	00101000	28
)	41	00101001	29
*	42	00101010	2A
+	43	00101011	2B
,	44	00101100	2C
-	45	00101101	2D
.	46	00101110	2E
/	47	00101111	2F
0	48	00110000	30
1	49	00110001	31
2	50	00110010	32
3	51	00110011	33
4	52	00110100	34
5	53	00110101	35
6	54	00110110	36
7	55	00110111	37
8	56	00111000	38
9	57	00111001	39
:	58	00111010	3A
;	59	00111011	3B
<	60	00111100	3C
=	61	00111101	3D
>	62	00111110	3E
?	63	00111111	3F
@	64	01000000	40
A	65	01000001	41
B	66	01000010	42
C	67	01000011	43
D	68	01000100	44
E	69	01000101	45
F	70	01000110	46
G	71	01000111	47

H	72	01001000	48
I	73	01001001	49
J	74	01001010	4A
K	75	01001011	4B
L	76	01001100	4C
M	77	01001101	4D
N	78	01001110	4E
O	79	01001111	4F
P	80	01010000	50
Q	81	01010001	51
R	82	01010010	52
S	83	01010011	53
T	84	01010100	54
U	85	01010101	55
V	86	01010110	56
W	87	01010111	57
X	88	01011000	58
Y	89	01011001	59
Z	90	01011010	5A
[91	01011011	5B
\	92	01011100	5C
]	93	01011101	5D
^	94	01011110	5E
~	95	01011111	5F
a	96	01100000	60
b	97	01100001	61
c	98	01100010	62
d	99	01100011	63
e	100	01100100	64
f	101	01100101	65
g	102	01100110	66
h	103	01100111	67
i	104	01101000	68
j	105	01101001	69
k	106	01101010	6A
l	107	01101011	6B
m	108	01101100	6C
n	109	01101101	6D
o	110	01101110	6E
p	111	01101111	6F
q	112	01110000	70
r	113	01110001	71
s	114	01110010	72
t	115	01110011	73
u	116	01110100	74
v	117	01110101	75
w	118	01110110	76
x	119	01110111	77
y	120	01111000	78
z	121	01111001	79
{	122	01111010	7A
	123	01111011	7B
}	124	01111100	7C
~	125	01111101	7D
	126	01111110	7E

Vgl. ergänzend das Thema "Codierung" unter verschiedenen Betriebssystemen:

<http://www.schoenitzer.de/encoding.html>

zu Ziff. 2.221:

Sonderzeichen

zu:

aus: TUSTEP-Handbuch (kleine Auswahl)

7-bit-TUSTEP-Zeichensatz

! ! Ausrufezeichen
/ / Schrägstrich

8-bit-TUSTEP-Zeichensatz

^! ; spanisches Ausrufezeichen
^a ä kleiner Umlaut ä




Mit Steuerzeichen "#" codierte Sonderzeichen

#! § Paragraph-Zeichen
#.% } Doppel-Alef

Mit Steuerzeichen "#" codierte Sonderbuchstaben

#.d đ d mit Querstrich (serbokroatisch)
#.^A Æ Ligatur AE

Mit Steuerzeichen "#(name)" codierte Sonderzeichen

#(AEH) ~ ähnlich
 #(DEAG)  doppelte eckige Klammer auf (groß)
 #(DMG)  Durchschnittsmenge (groß)
 #(INT)  Integral
 #(PFLD) <=< Pfeil links doppelt
 #(pf) £ Pfund-Zeichen

Griechische Schrift: #g+...#g-

a α Alpha
c ψ Psi

Akzente in der griechischen Schrift

%(ó Spiritus asper
%)? õ Spiritus lenis + Zirkumflex

Hebräische Schrift: #h+...#h-

a ׀ Alef
q ׀ Kof

Punktierung in der hebräischen Schrift

"#3 ׀ Segol
"#0#"5 ׀ Chatef Patach

Russische (cyrillische) Schrift: #r+...#r-

d ђ D

f ϕ F

Abweichende Zeichen im Bulgarischen

^w ш

Abweichende Zeichen im Makedonischen

%/g^w ђ

Abweichende Zeichen im Serbischen

^d đ

Abweichende Zeichen im Ukrainischen

^g ґ

Abweichende Zeichen im Weißrussischen

%)u ў

Cyrillische (alt-kirchenslavische) Schrift: #C+...#C-

a а Az

m м Myslite

Syrische Schrift: #Y+...#Y-

^b b B ^B ܒ ܝ ܚ ܝ Beth

^j j J ^J ܝ ܝ ܝ ܝ Teth

Arabische Schrift: #A+...#A-

^b b B ^B ٲ ٲ ٲ ٲ Bāʾ

^k k K ^K ڪ ڪ ڪ ڪ Kāf

Zusätzliche Zeichen im Persischen

#.^R #.R ړ ړ ړ Že

Phonetische Zeichen: #P+...#P-

^i ĭ i, barred: Vokal, high, central, unrounded

^? ʁ glottal stop reversed: Konsonant, fricative, pharyngeal, voiceless

Auszeichnungen, DruckeffekteAnfang bzw. Ende **Fettdruck** #f+ bzw. #f-Anfang bzw. Ende **S p e r r u n g** #s+ bzw. #s-Anfang bzw. Ende **Mittellinie** #0+ bzw. #0-Anfang bzw. Ende **Unterpunktierung** #4+ bzw. #4-

zu Ziff. 2.222:

TUSTEP und Unicode

zu:

aus: BI 99/7 3f - von M W KÜSTER.

Anders als Menschen können Computer nicht wirklich Zeichen verstehen und verwalten. Für sie ist jedes Zeichen nur eine arbiträre Bitfolge, der willkürlich eine bestimmte Semantik hinterlegt wird. Als Speicherpreise noch von zentralem Interesse waren, hat man sich bemüht, die pro Zeichen benötigte Bitfolge zu minimieren. Es gab Versuche mit 5-Bit-Code (Telex), dann auch mit 6- und 7-Bit-Codes (so z.B. in der ISO/IEC 646, deren bekannteste Anwendung ASCII ist). Die meisten der heute üblichen Verfahren gehen von einer 8-Bitfolge (= Oktet oder Byte) aus, was für $2^8 = 256$ Zeichen ausreicht. Traditionell wird die Zuordnung Bitfolge \leftrightarrow Zeichen in einer 16x16-Matrix wiedergegeben.

256 Positionen, von denen aus technischen Gründen obendrein nicht alle nutzbar sind, reichen selbst in dem Teil Europas, der die lateinische Schrift benutzt, nicht für alle Buchstaben und Sonderzeichen aus, von anderen Schriften ganz zu schweigen. Als Antwort darauf haben sich viele verschiedene, einander widersprechende Konventionen für die Zuordnung von Buchstaben zu Bitfolgen herausgebildet. So wird das ä z.B. allein in Westeuropa in sehr verbreiteten Konventionen durch drei verschiedene Werte repräsentiert: 0xE4 (»Latin-1«), 0x84 (DOS-Codepages) und 0x8A (Apple Macintosh). Sehr viel mehr Kodierungen sind zu berücksichtigen, wenn man auch seltenere Konventionen einbezieht.

Solange man Daten primär nur auf einem Rechner oder in einem eng umgrenzten Kulturkreis einsetzt, wirkt sich das nicht aus. Datenaustausch und langfristige Datenhaltung sind nach diesem Modell allerdings praktisch unmöglich, falls man nicht die auf dem Ausgangsgerät verwendeten Konventionen kennt und über geeignete Konvertierungswerkzeuge wie z.B. das #umwandle-Kommando in TUSTEP verfügt. In einer Zeit, in der Daten intensiv über das Internet ausgetauscht und verstärkt als wiederverwertbare Ressourcen eingesetzt werden, wird das Problem täglich drängender.

Das Universal Character Set (UCS)

Wenn eine 8-Bit-Folge nicht ausreicht, alle benötigten Zeichen zu kodieren, muss man entsprechend längere Sequenzen verwenden. Es wurden zwei konkurrierende Entwürfe erarbeitet, einer von der internationalen Standardisierungsorganisation ISO und ein weiterer vom Unicode Consortium, das im wesentlichen große amerikanische Softwarefirmen repräsentiert. 1993 wurden die Zeichenbelegungen der beiden Standards ISO/IEC 10646-1 und Unicode™ harmonisiert; sie werden seither in enger Kooperation erweitert. Verbleibende konzeptionelle Differenzen haben für den Endanwender praktisch keine Konsequenzen.

Letztendlich hat man sich auf eine zwei-Oktet-Folge geeinigt, mit der man $2^{16} = 65536$ Zeichen repräsentieren kann, was für die üblichen Schriftzeichen der momentan verwendeten Staatssprachen reicht. Über sogenannte *surrogate pairs* kann man weitere $1024^2 = 1.408.576$ Zeichen ansprechen, mit deren Hilfe man hofft, alle Schriftzeichen, die es je gegeben hat, erfassen zu können. In einer weiteren Ausbaustufe (»level-3-Implementierung«) erlauben es *combining diacritics* (fliegende Akzente), Diakritika mit (fast) beliebigen Grundbuchstaben zu kombinieren. Diese für viele wissen-

schaftliche Disziplinen essentielle Eigenschaft wird allerdings noch von fast keiner Software außer TUSTEP unterstützt.

Unicode läuft bereits jetzt in vielen Systemen, so z.B. in Windows NT, Aix und Java. Es setzt sich verstärkt auch als Transfermedium für Daten im Internet durch. Sowohl HTML 4 als auch XML rekurren auf Unicode als Grundzeichensatz, oft allerdings in einem speziellen normierten Transferformat UTF-8 (UCS Transformation Format, 8 bit form), das eindeutig in 16 Bit abbildbar ist. Alle üblichen Browser unterstützen in ihren aktuellen Versionen UTF-8. Die korrekte Darstellung der Resultate hängt dann natürlich davon ab, ob der Endnutzer geeignete Unicode-Fonts installiert hat. Insbesondere bei nichtlateinischen Schriften muss man obendrein oft deutliche Abstriche bei der typographischen Qualität hinnehmen.

Die Kodierung der modernen Schriften im UCS nähert sich inzwischen der Vollendung; die Bearbeitung der historischen Schriften beginnt momentan hingegen erst richtig. Leider mangelt es den zuständigen Bearbeitern in diesem Feld oftmals sehr an Expertise, so daß die so entstehenden Vorschläge, die z.B. auch die ägyptischen Hieroglyphen betreffen, mangelhaft sind. Auch im Eigeninteresse ist hier die intensive Mitarbeit der entsprechenden Fachwissenschaften sehr anstrengenswert. Bereits jetzt fungiert die Abteilung Literarische und Dokumentarische Datenverarbeitung hier als Schnittstelle zwischen Wissenschaftlern und Normungsorganisationen. Nachfragen aller Art sind herzlich willkommen.

Unicode-Unterstützung in TUSTEP

Für eine führende Anwendung im Bereich multilingualer Textkorpora wie TUSTEP ist Unicodeunterstützung hochrelevant. In Anbetracht des in TUSTEP konsequent angewandten Prinzips des *Script tagging* und der Kodierung von Akzentbuchstaben mit fliegenden Akzenten ist der Schritt hin zum Datenaustausch mit Unicode kleiner, als es vielleicht erscheint. Intern erlaubt Unicode beliebige Speicherformate, solange sie nur eindeutig in das UCS abbildbar sind, wobei *script tagging* ein für ernsthaftes Arbeiten mit nicht-lateinischer Schrift auf einer deutschen oder amerikanischen Tastatur bequemes und effizientes Vorgehen ist.

Bei der Arbeit mit TUSTEP ergeben sich keine Veränderungen. Man verwendet weiterhin die übliche TUSTEP-Zeichenkodierung.

Zum Datenaustausch im Unicode-Format gibt es ab der TUSTEP-Version 2000, die im Oktober 1999 freigegeben wurde, zwei neue Spezifikationswerte in *#umwandle*:
code=unicode liest und schreibt direkte 16-Bit-Zeichen,
code=utf8 wandelt die Daten aus dem und in das *UCS Transformation Format* um, das im WWW üblich ist.

So wird z.B. aus *#r+Materialy#r-* automatisch **Материалы** und aus *#g+Politeia#g-* automatisch **Πολιτεία** - und das sowohl im Druck als auch im Netz.

Nehmen wir an, wir wollten eine kleine UTF-8-kodierte Datei *sem.txt* folgenden Inhalts nach TUSTEP importieren:

```
Museum Graeco Latinum
Материалы семинара по Платона
Πολιτεία
```

Nach dem üblichen Anmelden müssen wir nur das Kommando *#umw, sem.txt,tu sem,co=utf8,lo=+* absetzen und erzeugen so die folgende Datei:

```
Museum Graeco Latinum
#r+Materialy seminara po Platona#r-
```

#g+Politeia#g-

Diese Datei kann man nun wie gewohnt bearbeiten. Mit *#umw, tu sem, sem.txt, co=utf8, lo=+* kann man das Ergebnis dann wieder in das Austauschformat bringen.

Zur Präsentation im Web sollte man dann auch die entsprechende »Meta-Information« im Header der HTML-Datei mitliefern:
`<meta http-equiv="Content-Type" content="text/html; charset=utf-8">`

Ganz analog arbeitet man mit dem Code *unicode*, dessen Ergebnis man in verschiedene Windowsprogramme übertragen kann. Hierbei hat man allerdings das Problem der Byteorder zu beachten, die sich zwischen sog. *little-endian* und *big-endian* Architekturen unterscheidet. Insgesamt dürfte daher zumindest noch UTF-8 die sichere Variante sein.

TUSTEP fungiert darüber hinaus als vollwertige *rendering engine* für die unterstützten Schriften (neben Lateinisch auch Griechisch, Kyrillisch, Hebräisch, Arabisch, Syrisch, Koptisch und Devanagari). In vielen Schriften wie z.B. im Arabischen ist diese Aufgabe komplex, da das Aussehen der Buchstaben von der Umgebung abhängt, in der sie stehen.

Auch der TUSTEP-Viewer, der beim Arbeiten mit dem Editor mit der Anweisung *mw+* gestartet werden kann, arbeitet intern mit Unicode und Unicode-Fonts. Er erlaubt dem Benutzer, gleich bei der Eingabe zu überprüfen, ob die Transliteration und die sonstige Kodierung korrekt ist. Auf diese Weise kombiniert TUSTEP die Bequemlichkeit der Dateneingabe mit einer Tastatur, die mindestens den ASCII-Zeichenvorrat unterstützt, mit den Vorteilen des *Universal Character Set*.

Vgl. auch: Themenheft "Literary & Linguistic Computing. Volume 15 Number 4 December 2000

zu Ziff. 2.223:

HTML - SGML

zu:

aus: W. BADER, Was ist die Text Encoding Initiative (TEI)? in: KAMZELAK, R (ed.), Computergestützte Text-Edition. Tübingen 1999: Niemeyer. S.9-20.

(10f) Plattform- und Softwareunabhängigkeit ist auch nicht banales 'Cross Media Publishing', das als Schlagwort in aller Munde ist, worin in der Tat eine große Zukunft und Herausforderung liegt. Oft heißt das leider, man wird statt nur durch eines gleich durch zwei Ausgabemedien diktiert.

Plattform- und Softwareunabhängigkeit läßt sich nicht durch die Verwendung von HTML erreichen. Zwar sind HTML-Dokumente plattformunabhängig auf der ganzen Welt anzuschauen, doch ist HTML - oder wird zumindest so verwendet - entgegen der Theorie nicht softwareunabhängig. Die zwei oder drei gängigen Browser bestimmen das Erscheinungsbild. HTML wird meist nicht zur Beschreibung einer Textstruktur verwendet, sondern zum Erzeugen typographischer Effekte, mit all den Einschränkungen und Zugeständnissen, wie man sie auch bei der Arbeit für die Papierausgabe machen muß. HTML ist eine in ihrer Struktur kleine Anwendung von SGML, die Texte hinsichtlich der Erscheinung am Bildschirm beschreibt, aber weit davon entfernt ist, ein Instrumentarium für die wissenschaftliche Beschreibung bereitzustellen.

Plattform- und Softwareunabhängigkeit bedeutet in einem tieferen Sinne - den ich hier favorisieren und als den richtigen betrachten möchte - eine neue Denkweise losgelöst von Medium und Software der Präsentation. Das heißt aber genaugenommen: Plattform- und Softwareunabhängigkeit ist eine alte Denkweise. So wie es die Wissenschaft früher nicht gekümmert hat, ob konkrete Setzer das handwerkliche Geschick hatten, eine bestimmte Type zu gießen und zu feilen, so sollte auch in der computergestützten Textedition diese strikte Teilung im Blick sein: inhaltliche Erarbeitung der Daten und ihre Interpretation - Übertragung der Daten - Präsentation der Ergebnisse...

(11) Die Übertragung der Daten ist dann nicht nur ein technischer Vorgang, sondern eine inhaltliche Auswahl und Umformung. Niemals kann hier noch Intelligenz und Interpretation hinzugefügt werden, sondern es wird weggelassen, ausgewählt, ungeordnet. Es handelt sich immer um eine 'Downkonvertierung', vom reichen Material der erarbeiteten Version, hin zur Auswahl für die spreizifische Präsentation. Solche 'Downkonvertierungen' können zu jedem späterem Zeitpunkt neu erarbeitet, entwickelt und vorgenommen werden und sind vom jeweiligen Bedürfnis der Anwendung abhängig.

Die Erarbeitung der Daten ist der eigentliche wissenschaftliche Schritt. Es gilt ein Modell zu finden, das diesen inhaltlichen Freiraum gewährt, der Wissenschaft alle notwendigen Möglichkeiten bietet und gleichzeitig offen für die weiteren Schritte ist.

2. Textauszeichnung

These: *Textauszeichnung ist eine Methode, eine oder mehrere Interpretationen eines Textes explizit zu machen.*

Nach der bisher Gesagten ist klar: Interpretation eines Textes ist hier im umfassenden Sinn gemeint, von der gesicherten Erarbeitung des Textbestandes, über Textvarianten, Anreicherungen durch Erklärungen, Kommentare, Analysen, bis hin zu Auslegungen und Ansichten über diesen Text. Dies ist Aufgabe der Wissenschaft. Solche Ergeb-

nisse können mit Hilfe von Auszeichnungen in den Text eingetragen werden.

zu Ziff. 2.224:

Metadaten

zu:

aus: V. HENZE, Metadata / Dublin Core. in: KAMZELAK, R (ed.), Computergestützte Text-Edition. Tübingen 1999: Niemeyer. S.29-40.

(29) Grundsätzlich besagt der Begriff *Metadaten* erst einmal nur ganz generell, daß es sich um Daten über Daten handelt. Metadaten enthalten Angaben über Form und Inhalt von Dokumenten oder Objekten. Im großen und ganzen sind Metadaten also gleichbedeutend mit einer bibliographischen Beschreibung. Sie finden nicht allein im Bibliotheksbereich Anwendung, sondern z.B. auch in Archiven, Museen oder den Dokumentverwaltungssystemen anderer öffentlicher Einrichtungen und privater Firmen.

Metadaten können Teil des Dokuments oder Objekts sein, auf das sie sich beziehen, oder aber auch getrennt davon vorliegen. Vertrautestes Beispiel aus dem bibliothekarischen Bereich sind die traditionellen Katalogaufnahmen gedruckter Publikationen, sei es nun in Form eines CIP-Eindrucks (Cataloguing in Publication) im Dokument selbst oder eines Datensatzes in der Katalogdatenbank eines Bibliothekssystems.

(30) Das Dublin Core Set entstand als Ergebnis eines von OCLC (Online Computer Library Center) im März 1995 in Dublin, Ohio veranstalteten Metadaten-Workshops. Die Absicht war dabei, Kernelemente für die Beschreibung von dokumentartigen Objekten im Fernzugriff festzulegen, um auf diese Weise ihre Identifizierung in einer Netzumgebung (Internet) und den Zugriff darauf zu erleichtern.

(31) 4.1. Inhaltsbezogene Elemente des Dublin Core Sets

- | | |
|----------------|--|
| 1. Title | Sachtitel |
| 2. Subject | Schlag- oder Stichwort, Schlagwortkette |
| 3. Description | Inhaltliche Beschreibung (Abstracts) |
| 4. Source | Quelle |
| 5. Language | Sprache |
| 6. Relation | Beziehung zu anderen Dokumenten / Objekten |
| 7. Coverage | Räumliche oder zeitliche Maßangaben |

4.2. Quellen- und urheberbezogene Elemente

- | | |
|------------------|---|
| 8. Creator | Verfasser, Urheber, Produzent |
| 9. Publisher | Verleger, Herausgeber |
| 10. Contributors | Sonstige beteiligte Personen |
| 11. Rights | Rechtliche Bedingungen, Copyrightangaben, Benutzungsbedingungen |

(32) 4.3. Elemente zur Spezifizierung der Quelle

- | | |
|----------------|---------------------------------------|
| 12. Date | Datum |
| 13. Type | Objekt-/Dokumenttyp, Ressourcenart |
| 14. Format | Format |
| 15. Identifier | Identifikation des Objekts, Dokuments |

Beispiel eines DC-Metadateneintrags:

```
Subject (scheme=LCSH) = Cataloguing of computer files
Subject (scheme=LCSH) = Internet (Computer network)
Subject (scheme=LCSH) = Information storage and retrieval systems
Subject = Network Information Discovery and Retrieval
Subject = Metadata
Subject = Dublin core set
```

```
Subject = Warwick framework
Title = Warwick framework and Dublin core set provide a comprehensive infrastructure for network resource description.
Title (Subtitle) = Report from the Metadata Workshop II, Warwick, UK, April 1-3, 1996
Author = Hakala, Juha
Author = Husby, Ole
Author = Koch, Traugott
Publisher = NORDINFO
Date = April 7, 1996
ObjectType (scheme=ACR2) = computer file
Form (scheme=IMT) = text/html
Identifier (scheme=URL) = http://www.ub2.lu.se/tk/dcwsrept.html
Relation (type=sibling) (identifier=URL) = http://www.oclc.org:5046/oclc/research/conferences/metadata2/
Language = Englisch
```

zu Ziff. 2.225:

SGML \supset XML

zu:

aus: SPERBERG-McQUEEN, C M, Perspectives on XML in: Schwitalla, J; Wegstein, W, Korpuslinguistik deutsch: synchron - diachron - kontrastiv. Würzburger Kolloquium 2003. Tübingen 2005. S. 287-310.

(288) But what is this SGML which these projects apply to corpus linguistics? And how does it relate to the XML mentioned in my title? SGML, or the Standard Generalized Markup Language, is an international standard defined by ISO 8879:1986. It offers a non-proprietary means of providing explicit markup of any collection of textual features which may be of interest. It does not, as one might at first expect, achieve this by defining a vocabulary or set of tags for marking up textual features; it defines nothing of the sort. Instead, it defines a meta-language by means of which the user of SGML may define an arbitrary set of tags (or, more precisely, of "element types") for making up documents of a given sort. The user decides what element types to define, and the meaning ascribed to them (what textual features they should describe); the selection of textual features to be marked up is limited not by the capabilities of a particular piece of software (as in word processors of formatting languages), nor by the decisions made in a committee defining a markup language (as in HTML and most other specific applications of SGML), but only by the user's interests and the user's ability to mark up the features of interest (both economic and intellectual limitations play a role). The meta-language provided by SGML takes the form of a /document type definition/, which contains what might be called a "document grammar", roughly in the style of a BNF (Backus/Naur Form) grammar.

SGML also allows the user to declare the coded character set they are using; this allows SGML to be used on virtually any kind of computer system. XML (the Extensible Markup Language) is by contrast not an International Standard in the narrow sense, although it is fast becoming a de facto standard. It is defined by a Recommendation issued by the World Wide Web Consortium (W3C) first issued in 1998 and re-issued with corrections from time to time since then. XML is a subset of SGML, both in the sense that the rules of XML are a subset of the rules of SGML and in the sense that every well-formed XML document is a legal SGML document. The subset was designed to be easier to handle on the World Wide Web (the responsible Working Group was originally called the *SGML on the Web* working Group), and in particular to be easier to parse than SGML. There are fewer optional features (the most prominent exception being that processing of the document grammar or DTD is not required of conforming XML processors), and the character set is constrained to be Unicode (although non-Unicode character encodings may be used, the repertoire of characters is defined exclusively with reference to Unicode / ISO 10646). By making XML easier to parse than SGML, the designers hoped to make it easier to write software for (289) XML than for SGML and thus to encourage software development. Since 1998, this hope has been amply fulfilled.

I should pause for a moment to describe the World Wide Web Consortium, as the body responsible for the XML specification. W3C is a member-supported organization which creates Web standards (in the form of Recommendations, which describe recommended practice). Its mission is to *lead the Web to its full potential*, and W3C is accordingly engaged in a wide variety of endeavors to ensure that

the World Wide Web is accessible to all, regardless of language, geography, script, visual impairment, or other disability; to improve the utility of the Web as a medium not just for human to human but also for program to program communication; to address social concerns confronting the Web, in particular the development of a Web of Trust which makes the Web a more successful collaborative environment; to improve the interoperability of Web-based software and the evolvability of the technical design of the Web; to encourage decentralization in the Web; and to provide better standards-based multimedia formats. ...

(292) A simple example may make this more concrete. Following the tradition of many computer languages, we use a trivial "hello, world" example which illustrates some of the most important features of SGML:

```
<!DOCTYPE greetings[
<!ELEMENT greetings (hello+)>
<!ELEMENT hello (#PCDATA)>
<!ATTLIST hello
    lang CDATA #IMPLIED>
<!ENTITY szlig "&#223;">
<!ENTITY uuml "&#252,">
]>
<greetings>
<hello lang="en">Hello, world!</hello>
<hello lang="fr">Bon jour, tout le monde!</hello>
<hello lang="no">Goddag!</hello>
<hello lang="de">Guten Tag!</hello>
<hello lang="de-franken">Gr&uuml;&szlig; Gott!</hello>
</greetings>
```

The document is divided into two parts: first the document type declaration, which runs from the first line (<!DOCTYPE greetings[) through the eighth (]>), and then the document instance, which runs from the ninth line (<greetings>) through the fifteenth (</greetings>). Let us consider the document instance first. It is encoded using a fairly simple form of labeled bracketing. The document is divided into structural units known as "elements"; (293) each element is delimited by a "start-tag" at its beginning and an "end-tag" at the ending, each giving the "element type" of the element. Here, the top-level element (there is always exactly one outermost element in any well-formed XML document) is of type "greetings", and it has five child elements, each of type "hello", each containing a greeting. The final "hello" element uses two "entity references" (the strings ü and ß) to represent characters not conveniently accessible from my US-oriented keyboard. The start-tag can also contain a list of attribute-value pairs showing attributes of the element type; in this example, each "hello" element is labeled with a code for its language. The document type declaration contains declarations for the two element types; the element declarations may be thought of as productions in a regular right-part grammar defining the set of valid documents. The declaration

```
<!ELEMENT greetings (hello+)>
```

indicates that valid elements of type "greetings" contain one or more elements of type "hello". The declaration

```
<!ELEMENT hello (#PCDATA)>
```

indicates that "hello" elements contain character data.

Attributes may also be declared: the declaration

```
<!ATTLIST hello
```

```
  lang CDATA #IMPLIED>
```

indicates that elements of type "hello" may but need not carry a "lang" attribute, which in turn contains a string of character data. The entities referred to in the document are also declared. Since entities are simply named strings of characters, the declaration for an entity consists, in the simple case, just of the entity name and its replacement value:

```
<!ENTITY szlig "&#223;">
```

```
<!ENTITY uuml "&#252;">
```

Here, the name /szlig/ is defined for the string consisting of the simple character whose decimal number in Unicode is 223 (U+00DF, in the usual Unicode notation), and the name /uuml/ is assigned a replacement string consisting of the character U+00FC. ...

(303) For purposes of such search, retrieval, and manipulation, the specifications of XPath (1.0 and 2.0), XSLT (1.0 and 2.0) and XQuery (1.0) are all of particular interest. In theory, at least, the core functions of these specifications are quite distinct. XPath provides a language for identifying (or, as some say, /addressing/) particular elements and attributes in an XML document, while XSLT is designed for use in document formatting or rendering systems, and XQuery is designed to provide, for XML-encoded data, data manipulation functionality analogous to that provided for relational data by SQL, the Structured Query Language. In practice, however, the three specifications are intimately related: XPath 2.0 is used as an expression language by both XQuery 1.0 and XSLT 2.0, and all three specifications use a common data model and are based on the same formal semantics. All three languages can be used for querying XML data (although XSLT and XQuery both allow for manipulation and elaboration of the data, whereas systems which use XPath as their query language generally confine themselves to presenting the results, without modification), and both XSLT and XQuery can be used for manipulation of XML tree structures.

zu Ziff. 2.2251:

SGML - Praxis

zu:

aus: M. KOLTES, Die Regestaussgabe der Briefe an Goethe. in: KAM-ZELÄK, R (ed.), Computergestützte Text-Edition. Tübingen 1999: Niemeyer. S.101-115.

(108) In der Reihenfolge von oben nach unten stehen die Kennungen für die Regestnummer (natürlich bis zur Fertigstellung des letzten Regests im Band noch unbesetzt), die Nummer der Mappe (zur Identifizierung), den Briefschreibernamen, das Datum des Briefs, den Absendeort, die Signatur, die Angabe zum Druck, den Bezugs- bzw. Antwortbrief Goethes, Tagebuchvermerke Goethes sowie Bemerkungen zur Vorlage für das Regest. Schließlich werden der Regesttext selbst sowie Anlagen und/oder Beilagen codiert. An den eigentlichen Text schließen sich die Registereinträge an, die je nach Registerart codiert sind. Die Codierungen bestehen jeweils paarig aus einer eindeutigen Anfangs- und einer Endekennung, in unserem Fall einer mnemotechnischen Abkürzung des Inhalts, in spitze Klammern eingeschlossen...

(112) Beispiel für ein durchcodiertes Regest (noch in Bearbeitung)

```
<numa><nume>
<mapa>0072<mape>
<korna>15<korne>
<auta>Boisserée, Johann Sulpiz Melchior Dominikus<aute>
<data>1814 Dezember 3<date>
<orta>Heidelberg<orte>
<sigas>28/206 St. 14<siges>
<drua>GB 2, 42-49<drues>
<beza>1814 November 19 (25, Nr. 6934)<bezes>
<anta>1815 Januar 2 (25, Nr. 6972)<antes>
<umfa>1 St. Bl. Kop.:<umfes>
<inzipa>{-} Wir sehen mit Freuden, daß Sie wohl und heiter
sind <inzipes>
<dabeia><dabeies>
<sonsta>auf Brief: rechte obere Ecke: "Augusts Nachl." <sonstes>
<regas>Freude über G.s Wohlfinden und Erinnerungen an den
gemeinsamen Aufenthalt in Heidelberg (24. September - 11.
Oktober), an den B.s Brief anknüpfen möchte. - Erwartungen an
G.s <zitas>durch die Betrachtung unserer Sammlung angeregten
Rückkehr zu Kunstgeschichte<zites>... (113) Grüße an G. beson-
ders von A.F.J. Thibaut, dessen Forderung nach einem deut-
schen Gesetzbuch ("Über die Notwendigkeit eines allgemeinen
bürgerlichen Rechts für Deutschland") von K.F. von Savigny
("Vom Beruf unserer Zeit") zurückgewiesen worden sei. Thibaut
habe darauf in den <zitas>hiesigen Jahrbüchern brav geantwor-
tet<zites> (Heidelbergische Jahrbücher der Literatur 1814 Heft
1, 526-528). <reges>
*erp Moller, Georg
*erp Boisserée, Melchior Hermann Joseph Georg
*erp Schiller, Johann Christoph Friedrich von
*erp Gretserus, Jacobus ...
```

(114)*gwr Wilhelm Meisters Lehrjahre. (2.Buch)

*awr+ Schiller, Friedrich von, Der Taucher. - In: Musenalma-
nach für das Jahr 1798.

*awr Gretserus, Jacobus. De imagibus non manufactis opus
(=Corpus Byzantinae historiae,8). Paris 1648.

*awr Freitag, Johann Gotthilf. De sudario Veronicae in templo Portensi depicto progr. - In: J. Bertuch. Chronicon Portense. Duobus libris distinctum. Leipzig 1739, Bd. 1, S. 207-212.

zu Ziff. 2.226:

SGML → Software

nu:

```

aus: TUSTEP-Handbuch (kleine Auswahl)
#ko,?1,?2,+,+,*
>lz      0123456789
        Kapitälchen
xx      \<kpa>\++\<kpe>\++{\
        Kursiv
xx      \<ka>\@\<ke>\@\{\
        Titelzeile: kleinste
xx      \<la>\ & \<le>\ & \
        Titelzeile: Zweitgrößte
xx      \<2a>\ && #f+\<2e>\#f- &&{ \
        Titelzeile: Größte
xx      \<3a>\ &&& #f+\<3e>\#f- &&&{ \
        Schrift größer Ende
xx      \<sgr>\ &!K14\</sgr>\&!K00\
        Schrift kleiner Ende
xx      \<skl>\ &!K12\</skl>\&!K00\
        Fett
xx      \<fa>\#f+\<fe>\#f-\
        $ = Absatz      $$ = Einschaltung      $$$ = Ende der E.
xx      \<ab>\ $ \<ein>\ $$ \</ein>\$$\ \
        $$$ = Neue Zeile stumpf
xx      \<z>\ $$$ \@I+<>%i-\ \
xx      \<z1>\ $$$=$$$ \
xx      \<z2>\ $$$=$$$=$$$ \
xx      \<z3>\ $$$=$$$=$$$=$$$ \
xx      \<z4>\ $$$=$$$=$$$=$$$=$$$ \
        negativer Durchschuss (einstellig)
xx      \<nd>/>\&!D-<=02 \
        positiver Durchschuss (zweistellig)
xx      \<pd>/>/>\&!D<=02<=03 \
        zentriert
xx      \<zen>\ @-z \
        Kopftext rechts / links / Ende
xx      \<kra>\ @/ \<kla>\ @= \<kte>\ @{\ \
        Fußnoten
xx      \<sa>\@f+\<se>\@f-\
        Griechisch
xx      \<gra>\#G+\<gre>\#G-\
        Hebräisch
xx      \<hea>\#H+\<hee>\#H-\
        Sperrung
xx      \<spa>\#s+\<spe>\#s-\
        Zeilenwechsel + VORSCHUB
xx      \<zv>\ $ \
        Zeilenwechsel + LEERZEILE
xx      \<lz>\ $$$=$$$ \
        Merkstelle n      Positionieren auf Merkstelle n
xx      \<M>/>\&!M<=02\<P>/>\&!P<=02\
        Einrücken der Folgezeilen auf erreichte Position
xx      \<&i>\ &=999\
        Einzug von links in Punkt
xx      \<e>/>/>/>\@-(>=03>=04>=05)\
        Einrücken ab erreichter Position um nnn Punkt
xx      \<er>/>/>/>\&=>=04>=05>=06\
        Unterstreichen

```

```

xx  \|#1+\|#1-\
xx  Unterstreichungsstrich von a nach b
xx  \<<>/U<>/>/>\&!-(>=02,<=02)\
xx  nach zuvor gemerkten Stellen
xx  \<p>/U<p>/>/>\&!-(p>=03,p<=02)\
xx  nach Bruchteilen der Satzbreite
xx  \<>/>//U//>/>\&!-(>=02/<=04,>=04/<=02)\
xx  Positionierung linker Rand
xx  \<l>/>\&=-00>=03 \
xx  \<l>/>/>\&=-0>=03>=04 \
xx  \<l>/>/>/>\&=->=03>=04>=05 \
xx  Hochgestellter Ber. o. Größenänderung: Punkt - Ende
xx  \<h>/>\#H(<=02)\</h>\#G:\
xx  Tiefgestellter Ber. o. Größenänderung: Punkt - Ende
xx  \<t>/>\#T(<=02)\</t>\#G:\
xx  AUTOR- oder REGISTEREINTRAG: Markierung weg
xx  \<aa>\<ae>\<ra>\<re>\
xx  Zeilenteilung: Rest wird nach rechts geschoben
xx  \<r>\ @-0 \
xx  NEUE SEITE
xx  \<ns>\ &&N&&{\ \
xx  rechtsbündig
xx  \<rba>\ @-0 \
xx  Auffüllen mit einer Linie
xx  \<lin>\@--0\
xx  Silbentrennung: aus - ein
xx  \&s-\&!s-\&s+\&!s+\
xx  Datum für Makro vorbereiten
xx  \<datum>\!datum!\
*eof

```

zu Ziff. 2.2261:

DTD + SGML

zu:

aus: H-W LUDWIG; T. ROMMEL, Studium Literaturwissenschaft. Arbeitstechniken und Neue Medien. Tübingen 2003.

(55) Denn solchen aufbereiteten Dokumenten sind Listen mit Definitionen, welche Codierung was bedeutet, in einem so genannten DTD (*Document type definition*) vorgeschaltet. Hierbei handelt es sich um rein technische Informationen.

Beispiel:

DTD enthält die Beschreibung/Definition der Codierungen, die im nachfolgenden Text verwendet werden

TEXT Primärtext, der mit Codierungen ausgezeichnet ist

Der elektronische "Text", der auf dem Bildschirm erscheint, ist lediglich der zweite, bei TEXT enthaltene Textteil. Das DTD stellt, wie ein Regelwerk beim Spiel, den formalen Rahmen für das Verständnis und die Bewertung des Nachfolgenden zur Verfügung. Das eine funktioniert nicht ohne das andere, aber nur Spezialisten interessieren sich für den konkreten Inhalt des DTD:

Beispiel:

```

<!ELEMENT anthology -- (poem+)>
<!ELEMENT poem -- (title?, stanza+)>
<!ELEMENT title - O (#PCDATA)>
<!ELEMENT stanza - O (line+)>
<!ELEMENT line O O (#PCDATA)>

```

Ein professionelles Leseprogramm wertet dieses (hier vereinfacht wiedergegebene) DTD als "Vorspann" aus und kann dann die formatierte Textdatei aufbereitet wiedergeben; ebenso erlaubt es ausgeklügelte Such- und Indexfunktionen.

Im Nachfolgenden das Gedicht von William Blake, "The Sick Rose" mit Auszeichnungen. Deutlich sind die individuellen Markierungen zu erkennen, die hier am jeweiligen Zeilenanfang eingefügt sind. Alle metatextuellen Informationen sind, wie bei SGML üblich, in Spitzklammern eingeschlossen. Deutlich wird auch, dass sich alle Codierungen einer strikten Logik unterwerfen; beispielsweise sind Anfang und Ende durch identische Codierungen markiert, bei denen beim Ende ein zusätzliches "/" eingefügt ist. (56)

Beispiel:

```

<anthology>
<poem><title>THE SICK ROSE</title>
<stanza>
<line> O Rose thou art sick.</line>
<line>The invisible worm,</line>
<line>That flies in the night</line>
<line>In the howling storm:</line>
</stanza>
<stanza>

```

```

<line>Has found out thy bed</line>
<line>Of crimson joy:</line>
<line>And his dark secret love</line>
<line>Does thy life destroy.</line>
</stanza>
</poem>
</anthology>

```

zu Ziff. 2.227:

Schriftsysteme

zu: vsyst6.11

aus: P. SCHLOBINSKI, Zum Prinzip des Relativismus von Schriftsystemen - die chinesische Schrift und ihre Mythen: Zeitschrift für Sprachwissenschaft 20.1 (2001) 117-146.

(141f) So zeigt sich, daß zum einen alle Schriften zu Beginn Semo-gramme, speziell Piktogramme, aufweisen, im historischen Prozess jedoch zunehmend und mehr minder stark phonetisiert wurden: Die altägyptische Konsonantenschrift ist in hohem Maße und frühzeitig phonetisiert, die morphosyllabische chinesische Schrift demgegenüber weniger stark. Eine rein pikto- oder logographische Schrift ohne Phonetika hat sich als Kodierungssystem natürlicher Sprachen nicht entwickelt (vgl. auch Abb. 3), wobei hier von Erfindungen wie der Bliss-Schrift etc. abgesehen wird.

Während eine morphosyllabische Schrift wie die chinesische am geringsten phonetisiert ist, steht die lateinische *scriptio continua* am anderen Ende der Phonetisierungsskala (s. hierzu Saenger 1997). Neben den Alphabetschriften, die als Konsonantenschrift und Konsonant-Vokal-Schrift vorkommen, gibt es reine Silbenschriften wie die Cherokee-Schrift oder morisch-syllabische Schriftsysteme wie die japanischen Syllabare Hiragana und Katakana. Das Koreanische ist eine Alphabetschrift, die in Syllabogrammen angeordnet ist (alphabetosyllabische Schrift). Die drei Grundtypen von Schriftsystemen, die morphosyllabische, die syllabische und die Alphabetschrift treten allein und in Hybridsystemen auf...

(143f) Ein Schriftsystem beginnt sich dann zu entwickeln oder wird von einer Geberkultur übernommen, wenn ein gesellschaftliches Erfordernis besteht, das Gedächtnis einer Kommunikation unabhängig von den Interaktionsteilnehmern zu fixieren: "Durch Schrift wird Kommunikation aufbewahrbar, unabhängig von dem lebenden Gedächtnis der Interaktionsteilnehmer, ja sogar unabhängig von Interaktion überhaupt" (Luhmann 1984: 127). Es findet eine Entsituierung vom Hier und Jetzt der Interaktionsteilnehmer statt, die Kommunikation wird (im Bühlerschen Sinne) von der primären Origo abgelöst, indem Produktions- und Rezeptionssituation entkoppelt sind. "Kommunikation wird, obwohl sie nach wie vor Handeln erfordert, in ihren sozialen Effekten vom Zeitpunkt ihres Erstauftretens, ihrer Formulierung abgelöst. [...] Man formuliert für unabsehbare soziale Situationen, in denen man nicht anwesend zu sein braucht" (ibid., S.128). Diese Funktionalität von Schrift mit all ihren gesellschaftlichen Differenzierungen - Fixierung von rituellem, kulturellem Wissen, kommunikativem Handeln, ökonomischen Sachverhalten etc. - ist unabhängig vom jeweiligen Schriftsystem: Eine Alphabetschrift leistet prinzipiell das Gleiche wie eine syllabische oder morphosyllabische Schrift.

zu Ziff. 2.2271:

Ausdruck + Bedeutung → Bedeutung → Bedeutung →→ .

zu:

aus: Italo Calvino, Herr Palomar. München 1988.

(111) In der altmexikanischen Archäologie stellt jede Figur, jeder Gegenstand, jedes Detail eines Flachreliefs etwas dar, alles bedeutet etwas, das etwas bedeutet, das seinerseits etwas bedeutet. Ein Tier bedeutet einen Gott, der einen Stern bedeutet, der ein (112) Element bedeutet oder eine menschliche Eigenschaft, und so weiter. Wir befinden uns in der Welt der Bilderschrift. Wenn die Tolteken schreiben wollten, zeichneten sie Figuren, aber auch wenn sie einfach nur zeichneten, war es, als ob sie schrieben: Jede Figur erscheint wie ein Bilderrätsel, ein zu entziffernder Rebus. Selbst noch die abstraktesten, rein geometrischen Friese auf einer Tempelwand können als Sonnenstrahlen gedeutet werden, wenn man darin ein Motiv mit unterbrochenen Linien sieht, oder man kann eine Zahlenabfolge in ihnen lesen, je nachdem, wie sich die Mäander verschlingen. Hier in Tula wiederholen die Flachreliefs stilisierte Tiere: Jaguare, Coyoten. Der mexikanische Freund erklärt Herrn Palomar jeden Stein, übersetzt ihn in kosmische Mythen Erzählungen, Allegorien, moralische Reflexionen. ...

(113) Ein Stein, eine Figur, ein Zeichen, ein Wort, die uns (114) isoliert von ihrem Kontext erreichen, sind nichts als eben nur dieser Stein, diese Figur, dieses Zeichen oder Wort; wir können nur versuchen, sie als solche zu definieren und zu beschreiben, aber mehr nicht; wenn sie hinter dem Antlitz, das sie uns zeigen, noch ein verborgenes Antlitz haben, muß es uns verborgen bleiben. Die Weigerung mehr zu begreifen als das, was diese Steine uns zeigen, ist vielleicht die einzig mögliche Art und Weise, ihr Geheimnis zu achten. Es erraten zu wollen, ist die Anmaßung, Verrat an ihrer verlorengegangenen wahren Bedeutung.

zu Ziff. 2.228:

European Ordering Rules / Marc Wilhelm Küster

Vgl. im Internet:

<http://www.uni-tuebingen.de/static/skripte/tustep/prot/prot782-eor.html>

Es werden anschaulich Sortierprobleme bei unterschiedlichen Alphabeten vorgestellt - lateinischen und nicht-lateinischen, sowie Strategien, hier zu internationalen Standards zu kommen. Es werden auch Web-Adressen für weitere Vertiefung genannt.

zu Ziff. 2.229:

(Deutsche) Gebärdensprache (DGS)

zu:

aus: S. KUTSCHER, Non-Verbale Prädikation in der Deutschen Gebärdensprache (DGS): Probleme der Abgrenzung von Attribution und Prädikation in einer kopulalen Sprache in: L. GEIST; B. ROTHSTEIN (eds.), Kopulaverben und Kopulasätze. Intersprachliche und intrasprachliche Aspekte. LA 512. Tübingen 2007. S. 91-113.

(92) Anders als in der Lautsprache werden in gebärdensprachlichen Äußerungen grammatische und inhaltliche Informationen über mehr als einen Übertragungskanal gleichzeitig vermittelt. Während in der Lautsprache sowohl segmentale als auch suprasegmentale Komponenten nur über die Stimme, den vokalen Kanal, übermittelt werden, werden sprachliche Informationen in der Gebärdensprache gleichzeitig über verschiedene Kommunikationskanäle übertragen. Daher werden in der DGS drei verschiedene Arten von Zeichen mit bedeutungstragenden Funktionen unterschieden: a) mit den Händen ausgeführte Zeichen, die als Gebärden bezeichnet werden, b) mit Mund, Gesicht, Kopf und Oberkörper ausgeführte Bewegungen, sogenannte non-manuelle Komponenten, und c) mit dem Mund geformte "Wörter" der deutschen Lautsprache, die in der Regel tonlos artikuliert werden und deren Artikulation sich häufig auf den sichtbaren Teil des lautsprachlichen Wortes beschränkt. (93) Diese werden als Mundbilder oder Ablesewörter bezeichnet. Gebärden im engeren Sinne, also mit den Händen ausgeführte, konventionalisierte Form-Bedeutungs-Relationen müssen von spontanen Gebärden, Fingeralphabetzeichen und Gesten unterschieden werden. Die Ausführung einer Gebärde kann als eine einhändige konventionalisiert sein oder als eine koordinierte Bewegung beider Hände. Die beiden Hände des Gebärdenden werden als dominante und nicht-dominante Hand spezifiziert. Die dominante Hand (für Rechtshänder in der Regel die rechte Hand) führt konventionalisierte einhändige Gebärden aus und ist diejenige Hand, die sich in asymmetrischen zweihändigen Bewegungsgebärden bewegt, während die nicht-dominante Hand still steht.

Neben diesem visuell-gestischen Aspekt ist vor allem der räumliche Aspekt ein spezifisches Charakteristikum von Gebärdensprachen. Syntagmatische Beziehungen zwischen Gebärden und die zeitliche Verankerung von Sachverhalten werden im wesentlichen über die Ausführung von Gebärden in bestimmten Teilen des Raums vor dem Oberkörper und dem Kopf des Gebärdenden (Gebärdenraum) ausgedrückt. DGS-Äußerungen sind daher stark an der Visualität von Situationen orientiert. Bevor die eigentliche Ereignisbeschreibung gebärdet wird, wird zunächst der Raumkontext etabliert. Nachdem der Ground (vgl. Talmy 1983) etabliert ist, folgen Figure und Prädikat. So beginnt in (1) die Schilderung einer anstrengenden Fußreise über einen Alpen-Pass mit der Setzung der topologischen Gestalten TAL und BERG als Ground.¹⁴

¹⁴ Notationskonvention: Großbuchstaben: Gebärde, Unterstrich: Dauer der Fragment-Boje, D: dominante Hand, ND: nicht-dominante Hand, B: Gebärdennamen ist beidhändig, fa: Fingeralphabetzeichen, gestrichelte Linie (----): non-manuellen (!) Komponente wird über mehr als eine Gebärde ausgeführt, Führungsstriche (z.B. "trüb"): Gebärde wird üblicherweise durch spezifische Mimik begleitet, intensiv: spezifische Mimik (zusammengesetzte Augen und gespannter Mund), +: reduplizierte Gebärde, IDX: Gebärde INDEX, eine mit den Zeigefinger ausgeführte auf virtuelle oder sichtbare Entitäten zeigende Bewegung mit grammatischen Funktionen (näheres s. z.B. Keller 1998). Gebärden-

[94] (1) Kotext: Früher musste man zu Fuß über die Berge gehen.

Brauen	neutral	-----	-----	-----	-----
Mimik	neutral	-----	-----	anstrengend	-----
Gebärde	TAL	BERG	LAUFFBERGHOCH	ANSTRENGEND	
Dominanz	D	ND:Boje	D	B	
Oberkörper	neutral	-----	-----	-----	-----
Kopf	neutral	-----	-----	-----	-----

Brauen	-----	hoch	-----	neutral	
Mimik	-----	neutral	-----	Negation	
Gebärde	SCHNELL	MÜDE	FLACH	LAUFFLACH	MÜDE
Dominanz	D	ND	D	ND	B
Oberkörper	-----	-----	-----	-----	-----
Kopf	-----	leicht	neutral	-----	geneigt
		links			schütteln

Brauen	-----	hoch	-----	-----	-----
Mimik	-----	-----	-----	anstrengend	-----
Gebärde	NICHT	FLACH	LAUFFLACH	LAUFFBERG	Geste:
			(verkürzt)	HOCH	puuh
Dominanz	D	D	ND	ND	D
Oberkörper	-----	-----	-----	leicht	rechts
Kopf	-----	neutral	-----	-----	-----

'Tal und Berg: Den Berg hinauflaufen ist sehr anstrengend. Man wird schnell müde. Im Tal laufen macht nicht müde. In der Ebene laufen nicht, aber den Berg hoch, puuh!" (NAR)

Typische verbale Kategorien wie Tempus, Modus und Aspekt werden, Lautsprachen des isolierenden Typs (z.B. Chinesisch) ähnlich, durch lexikalische Elemente und nicht durch Flexion kodiert. Die zeitliche Verankerung erfolgt über freie lexikalische Formen (GESTERN, FRÜHER, MORGEN, SPÄTER). Satzarten, Irrealis und Konditional werden über Mimik und Mundgestik ausgedrückt und evtl. durch lexikalische Mittel begleitet. Die DGS hat keinen grammatischen Aspekt (im Sinne einer systematischen perfektiv-imperfektiv Unterscheidung). Aktionsarten-Aspekt wird über Reduplikation von Gebärden, über die Mimik und über lexikalische Mittel (z.B. durch die Gebärden SCHON, FERTIG) ausgedrückt. Gemäß der von Dik (1989) aufgestellten "copula support theory", die die Kopula als einen semantisch leeren Träger für Tempus/Aspekt/Modus-Kategorien auffasst, ist für (95) die DGS daher zu erwarten, dass sie keine lexikalische Kopula aufweist, was, wie Abschnitt 3 zeigt, für präzisierende bzw. askriptive non-verbale Prädikationen tatsächlich der Fall ist.

In der DGS gibt es keine der deutschen Lautsprache entsprechenden Kategorien wie Kasus und Genus, es gibt keine Artikel und die Kennzeichnung des Numerus einer Nominalphrase ist nicht obligatorisch. Das Fehlen spezifischer nominaler Kategorien und weitestgehend auch verbaler Kategorien, wie sie aus Lautsprachen bekannt sind, führt dazu, dass die Wortartenklassifikation für die Forschung zur Zeit noch strittig ist.

namen sind zur Identifizierung der Gebärden in Transkriptionen gedacht und orientieren sich in der Regel an Übersetzungen in die deutsche Lautsprache und dem eventuell zur Gebärde artikulierten Mundbild. Die Semantik bzw. Funktion der Gebärden stimmen jedoch in vielen Fällen nicht mit den aus der Lautsprache herangezogenen Wörtern überein. Wie die Grammatik ist auch das Lexikon der DGS anders strukturiert als die deutsche Lautsprache.

2.3 Korrekturprozeduren am eigenen Text

zu Ziff. 2.31:

Korrektur des Textes mit Hilfe einer zweiten EingabeErster Eingabe-Text = a

- 1.1 GRASS, G, Danziger Trilogie. Die Blechtrommel. Katz und Maus.
 1.2 Hundejahre. Darmstadt 1980. [Die Blechtrommel]
 1.3 Stichwort: Tkr
 1.4 (249) "Matzerath entschloß sich auf Gretchen Schefflers Zu-
 reden
 1.5 hin, meine Geliebte zu heiraten. Wenn ich also ihn, meinen
 mut-
 1.6 maßlichen Vater, als Vater bezeichne, muß ich feststellen:
 mein
 1.7 Vater heiratete meine zukünftige Frau, nannte später meinen
 Sohn
 1.8 Kurt seinen Sohn Kurt, verlangte also von mir, daß ich in
 seinem
 1.9 Enkelkind meinen Halbbruder anerkannte und meine geliebte,
 nach
 1.10 Vanille duftende Maria als Stiefmutter in seinem nach Fisch-
 laich
 1.11 stinkenden Bett duldete."
 1.12 (315) "Da trat Oskar näher. Da zog es ihn heran. Da wollte
 er
 1.13 auf den Teppich, wollt nicht mehr auf den Fliesen stehen.
 Eine
 1.14 Altarstufe gab ihn an die nächste weiter. So stieg ich hin-
 auf
 1.15 und hätte ihn lieber hinabsteigen sehen. »Jesus«, kratzte
 ich
 1.16 einen Rest Stimme zusammen, »so haben wir nicht gewettet.
 1.17 Sofort gibst du mir meine Trommel wieder. Du hast dein
 Kreuz, das sollte
 1.18 dir reichen!« Ohne abrupt abzubrechen, beendete er die Trom-
 melei,
 1.19 kreuzte die Stöcke übertrieben sorgfältig auf dem Blech
 1.20 und reichte mir ohne Widerrede, was Oskar im leichtsinnig
 1.21 gepumpt hatte. Schon wollte ich ohne Dank und hastig wie
 zehn Teufel
 1.22 die Stufen runter und raus aus dem Katholizismus, da berühr-
 te
 1.23 eine angenehme, wenn auch befehlerische Stimme meine Schul-
 ter:
 1.24 »Liebst du mich, Oskar?« Ohne mich zu drehen, antwortete
 1.25 ich: »Nicht daß ich wüßte.«"

Zweiter Eingabe-Text = b

[Gleicher Text von anderer Person getippt. Hier nicht abgedruckt. Annahme, dass diese Person die Fehler der ersten Person nicht macht (dafür vielleicht andere). Hier angenommen: zweiter Text hat die beiden Problemstellen richtig geschrieben.

Durch Vergleich werden die Textunterschiede aufgelistet.

Korrekturleser muss nur diese lesen und entscheiden, welche Varianten er auswählt.

Die Auswahl wird vom Programm als *Korrekturanweisungen* verstanden, so dass aus Text a und Text b mit Hilfe der korrigierten Liste

3. Beurteilung des fremden Textes**Literatur**

- GRÉSILLON, A: Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editionswissenschaft 4. Bern, Berlin, Frankfurt 1999: Peter Lang // I.2.1
- HEIDELBERGER-LEONARD, I (ed./eds.): "Text-Tollhaus für Bachmann-Süchtige?": Lesarten zur kritischen Ausgabe von Ingeborg Bachmanns Todesarten-Projekt. Mit einer Dokumentation zur Rezeption in Zeitschriften und Zeitungen. Opladen 1998: Westdeutscher Verlag // I.2.1
- MOTSCH, W (ed./eds.): Ebenen der Textstruktur. Sprachliche und kommunikative Prinzipien. RGL 164. Tübingen 1996: Niemeyer // I.2.4
- RABE, N: Vom Gerücht zum Gericht. Revidierte Text- und Literaturkritik der Kundschaftererzählung Numeri 13.14 als Neuansatz in der Pentateuchforschung. THLI 8. Tübingen 1993: Francke // I.2.1
- RÖHRING, H-H: Wie ein Buch entsteht. Einführung in den modernen Buchverlag. Darmstadt 1992: Wiss. Buchgesellschaft // I.2.1
- SCHIFFER, E: Zwischen den Zeilen. Manuskriptänderungen bei Thomas Mann. Transkription und Deutungsversuche. Berlin 1982 // I.2.1
- SCHWEIZER, H: Die Josefsgeschichte. Konstituierung des Textes. Teil I: Argumentation. THLI 4 / 1. Teil II: Textband. THLI 4 / 2. Tübingen 1991: Francke // I.2.1

zu Ziff. 3.01:**Vorstufen eines Textes**

zu:

aus: HEIDELBERGER-LEONARD, I (ed.), "Text-Tollhaus für Bachmann-Süchtige?" Lesarten zur kritischen Ausgabe von Ingeborg Bachmanns Todesarten-Projekt. Opladen 1998: Westdeutscher Verlag

(161f) [Rezension von J. Kaiser]
Text-Tollhaus für Bachmann-Süchtige

Das "Todesarten"-Projekt in fünfbändiger philologischer Aufbereitung

Was also ist hier neu? Jetzt werden wir mit *Vorstufen, Zwischenstufen* und detektivischen *Datierungs-Versuchen*, Zuordnungs-Hypothesen, überhäuft. Die Herausgeber kommentieren philologisch entstehungsgeschichtlich. Prüfen Papier-Sorten und Selbst-Kommentare. Geben also überhaupt keine wertende *Interpretation*. Man betritt ein Text-Tollhaus, das viel Herrliches, bislang Unbekanntes, aber keineswegs mehr enthält als eine verrückterweise immer noch unvollständige Material-Darbietung. Saftiges Weideland für künftige Bachmann-Dissertanten, für wertende Biographen. Reinhard Baumgart hat in seiner klugen Rezension (Zeit vom 24.11.95) daraufhin gewiesen, wie "sträflich" es sei, daß hier infolge akademischen Reinheits- und Neutralitäts-Wahnes entscheidende (allen Interessierten halbwegs bekannte) Dinge unausgesprochen bleiben, als hätte Max Frisch nie gelebt. Dafür werden im ersten Band allen wissenschaftlichen Ernstes vier mehr oder weniger gleichlautende Anfänge der "Geschichte einer Liebe" mitgeteilt (S. 47, 51, 52, 54) und im zweiten Band mehr als zehn von einander gewiß nicht überaus heftig abweichende Einleitungs-Entwürfe, die Ingeborg Bachmann sich für alle möglichen öffentlichen Vorlesungen aus dem *Buch Franza* notierte...

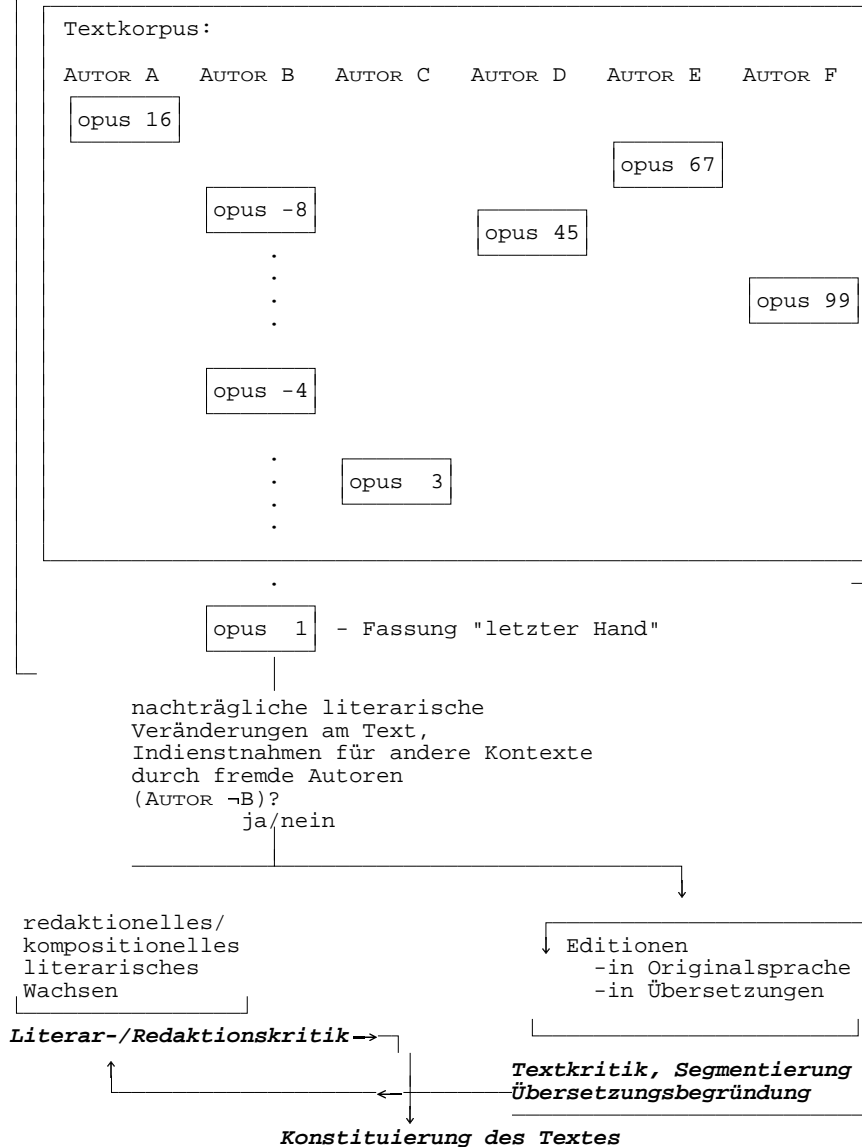
Und doch: Man gerät beim Lesen in den Sog der immer neuen Bachmann-Ansätze. Man verstrickt sich, wie die Dichterin, in einen Kampf, wo es keinen Sieg geben kann, sondern wirklich nur ein Überstehen. Es ist ja weithin unkritisierbar, was hier vorliegt - will man sich nicht in kleinliche Zuordnungs-Rechthaberei verwickeln, wo die Herausgeber aller Welt überlegen sein dürften, wenn auch vieles unlösbar bleiben muß, weil der Dichterin selber noch keine Lösung vorschwebte.

zu Ziff. 3.02:

Entstehung - Überarbeitung - Überlieferung von Texten

Zu: vkonst04.521 vsyst7.51 vgt3.431

→ **externe Syntax** (Verhältnis opus 1 : Textkorpus)



redaktionelles/
kompositionelles
literarisches
Wachsen

Editionen
-in Originalsprache
-in Übersetzungen

Literar-/Redaktionskritik →

Textkritik, Segmentierung
Übersetzungsbegründung

Konstituierung des Textes

Hier angenommen: *Opus 1* ist der Text eines Autors, auf den man sich bezieht, für den der Autor bekannt ist, z. B. GOETHES *Faust*, oder das Gedicht von BRECHT *An die Nachgeborenen* usw.

Wichtig: *Opus 1* meint die Textfassung, die ein Autor selbst geschaffen hat (geschrieben bzw. autorisiert). "Opus 1" ist nicht identisch mit einem frühen Druck, denn auf dem Weg von der Abfassung zur Drucklegung können bereits Veränderungen am Text stattfinden.

Bei einem gegebenen Werk - *Opus 1* als existierende Handschrift oder nur noch als erschließbare Größe (weil das Original verloren ist) - einsetzend, gilt zugleich: dieser Text hat auf jeden Fall eine **Vorgeschichte** (denn er fiel nicht vom Himmel); und er hat eine **Nachgeschichte**.

Alle drei Faktoren - Vorgeschichte/Verfassen des Textes/Nachgeschichte - werden als *Textbildungsprozess* bezeichnet und sind Gegenstand verschiedener wissenschaftlicher Disziplinen (mit je spezifischen Methoden).

Alle Stadien eines solchen *Textbildungsprozesses* können durch Zetfelsammlungen, Briefe, Handschriften, Typoskripte, Drucke bezeugt sein, oder aber nur durch einige späte Editionen - was die Rekonstruktion von *opus 1* entsprechend erschwert.

Ist durch die Schritte der **Konstituierung** ein Text ausreichend kritisch gesichtet, bereitgestellt, aufgearbeitet, steht also für weitere Deskription und Interpretation zur Verfügung, kann die (**Ausdrucks-**) **Syntax** anschließen - und zwar unter extensivem Rechereinsatz für die Datenerhebung :

- die **interne Syntax** analysiert die *string*-Kette des gegebenen Textes; die zuvor eingetragenen Segmentierungen des Textes, werden dabei übergangen;
- die **externe Syntax** macht sichtbar, welche *substrings* (z.B. Wortfolgen > 3) exakt/ähnlich in anderen Texten des (zuvor zu definierenden) Textkorpus vorkommen.

Die oben eingetragene **externe Syntax** kennzeichnet in der Grafik den Zuständigkeitsbereich, nicht eine Position im Ablauf der Methodenschritte. Zunächst muss die **Konstituierung** abgearbeitet werden, bevor es zur **Ausdrucks-Syntax** kommt.

zu Ziff. 3.021:

Aufbauend auf Text-Traditionsstrom

zu: vsyst7.23 vprag7.51103

aus: SCHÖNE, A (ed.) Johann Wolfgang Goethe. Faust. Texte. Frankfurt/M 1994. (= Bd.I) SCHÖNE, A Johann Wolfgang Goethe. Faust. Kommentare. Frankfurt/M 1994. (= Bd.II)

(II,11f) Alle Kunst aber beruht auf Vorgaben in der Kunst, keines ihrer Werke verdankt sich allein dem Ingenium eines einzelnen. Als sich unser Zeitgenosse Bertold Brecht dem Vorwurf ausgesetzt sah, er habe in seiner *Dreigroschenoper* Lieder von François Villon verwendet, ohne deren Herkunft kenntlich zu machen, erklärte er seine grundsätzliche "Laxheit in Fragen geistigen Eigentums". Der Weimaraner hatte es ihm vorgesagt: "alles, was die Vor- und Mitwelt geleistet" gehöre dem Dichter von Rechts wegen an, erklärte er am 17.12.1824 dem Kanzler v. Müller. "Nur durch Aneignung fremder Schätze entsteht ein Großes. Hab' ich nicht auch im <Prolog im Himmel für das Gespräch des Herrn mit> Mephistopheles den Hiob und <für die Nacht-Szene auf der Straße vor Gretchens Türe> ein Shakespearisches Lied mir angeeignet?" (Biedermann/Herwig 3.1, Nr 5564)

Der Verfasser des jüdischen Buches *Hiob* und der Dichter des englischen *Hamlet* stehen in einer kaum übersehbaren, vielsprachigen, weit in die Antike zurückreichenden und bis in Goethes Gegenwart führenden Reihe von Erzählern, Lyrikern und Dramatikern, deren Stimmen im *Faust* vernehmbar werden. Auch bestehen die fremden Schätze, die der Autor sich hier angeeignet hat, keineswegs nur aus identifizierbaren literarischen Vorgaben, wie er sie durch Zitate, Paraphrasen, Anspielungen und Verweise ins Spiel bringt. Daß man die Poesie aus diesem Werk "wiederherstellen" könne, will auch besagen, daß es seinen Bedeutungsreichtum durch eine Vielzahl von Aneignungen aus dem Formenarsenal gewinnt, welches sich im kollektiven Besitz der Dichter befindet. Im Hinblick etwa auf lexikalische Vorgaben, Satzbaupläne, rhetorische Figuren oder an bestimmten Motiven, Bildformeln, Symbolen und Allegorien, auch für typisierte Gestalten, Schauplatzmodelle oder Verhaltensmuster, ebenso beim vorgegebenen Instrumentarium der Bühnenkunst wird das in den nachfolgenden Kommentaren an vielen Einzelfällen sichtbar werden.

zu Ziff. 3.022:

Schreibprozess

zu:

aus: B. GLOOR, *staat - sex - amen*. 81 Sprachbeobachtungen. Zürich 2004.

Nr. 49

Vier Faktoren sind mit Sicherheit am Schreibprozess beteiligt: die Wahl des Schreibgeräts, sprachliche Konventionen, die Planung und die Überarbeitung.

1. Wer schreibt, muss sich für ein "Werkzeug" entscheiden: Bleistift oder Schreibmaschine? Diktaphon oder Computer? Jedes Medium erfordert andere Fähigkeiten, bei deren Fehlen dem Schreibprozess ein Teil der gedanklichen Energie entzogen wird. Goethe berichtet in "Dichtung und Wahrheit", er sei oft nachts erwacht, ans Pult gerannt und habe ein Gedicht heruntergeschrieben, mit Bleistift, damit nicht "das Schnarren und Spritzen der Feder mich aus meinem nachtwandlerischen Dichten aufweckte". Hätten diese Gedichte nach der Erfindung der Schreibmaschine überhaupt entstehen können? "Das Dichten auf der Schreibmaschine hat vermutlich mehr als alles andere zum Verfall der englischen Prosa beigetragen", sagt Edmund Wilson, ein pessimistischer Literaturkritiker der sechziger Jahre. Und was ist mit dem Computer? Entsprechen sein immateriell-anonymes Wesen und die immer weiter verfeinerte Tastatur dem Schreiben nicht am besten?

2. Grundlegende sprachliche Konventionen müssen beachtet werden: leserliche Handschrift, Beibehaltung eines Schriftsystems und einer Sprache, Beachtung einer Rechtschreibung etc.

3. In der Planungsphase entscheidet sich, was die Leser erfahren sollen, welche Wirkung der Text haben und wie er aufgebaut sein soll. In diesem Stadium wird auf der Stilebene auch klar, in welchem Ton der Text zu seinen Lesern sprechen will.

4. Der Löwenanteil des Schreibens ist das Überarbeiten. Bis ein Gedanke über den Entwurf, die Skizze, die Roh- und weitere Fassungen bis zum Gut-zum-Druck gediehen ist, verstreicht oft viel Zeit. Aussenstehende bringen Anregungen und eigene Gedanken ein, vielleicht wird das Ganze redigiert und hat in der Endfassung nur noch wenig mit der ursprünglichen Idee gemein.

Das sind vier wichtige Faktoren des Schreibens. Wenn wir den Text entstehen lassen, reihen wir Sätze aneinander. Wir setzen Platte an Platte, verfugen sie mit den Gesetzen der Logik und machen so das Nichts allmählich begehbar. Dabei bestimmt das, was wir sagen wollen, nicht immer das Geschriebene - manchmal ist es auch umgekehrt. Schreiben ist keine mechanische Wiedergabe von Gedanken, sondern ein Eintauchen in die Grauzone zwischen dem Gedanken und seinem grafischen Ausdruck, zwischen einem noch nicht ganz fassbaren Bild und einer Kette von Buchstaben mit einer nachvollziehbaren Bedeutung, zwischen Halbtraum und papierner Halbwirklichkeit.

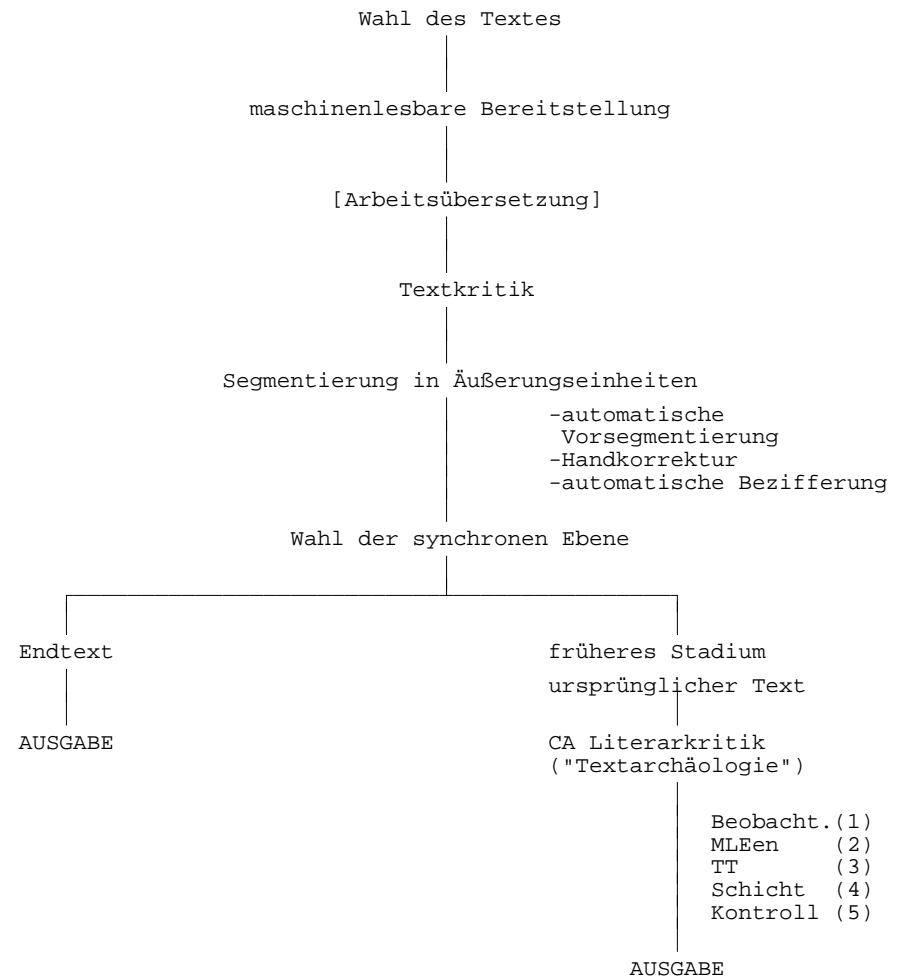
Wenn wir etwas zu Papier bringen wollen, können wir vom Denken oder vom Schreiben fortgerissen werden. Im ersten Fall sitzen wir am Ende mit vollem Kopf vor einem leeren Blatt, im zweiten betrachten wir das volle Blatt und denken vielleicht: "Das wollte ich eigentlich gar nicht sagen." Beim Schreiben versuchen wir ein Gedankengebäude voller Kreuz- und Querverbindungen, in dem wir uns sprunghaft und assoziativ bewegen, in eine lineare Bewegung zu bringen, indem wir einen Buchstaben hinter den anderen setzen. Am Ende steht ein sprachliches Gebäude, wiederum mit zahlreichen

Querbezügen, weniger kunstvoll und raffiniert zwar als das Original, aber auch hier können wir hin und her springen und unsere inneren Gegenden wieder wachrufen bzw. in unseren Lesern aufbauen, wo sie sich mit deren eigenen Erfahrungen verbinden.

Zuvor jedoch fließt alles säuberlich hintereinander durch den dünnen Kiel der Feder. Dabei kann viel schief gehen - oder alles. Dass uns vor der Niederschrift angst wird, ist logisch. Der Schreibstau ist die Befürchtung, das Ergebnis reiche nicht ans innere Original heran - und deshalb immer berechtigt.

zu Ziff. 3.03:

A. Konstituierung des Textes



AUSGABE = Präsentation des Ergebnisses für die folgenden Schritte der Deskription und der Interpretation:

(Ausdrucks-) SYNTAX

SEMANTIK

PRAGMATIK

zu Ziff. 3.04:

Rechte an Texten

zu:

aus: HEIDELBERGER-LEONARD, I (ed.), "Text-Tollhaus für Bachmann-Süchtige?" Lesarten zur kritischen Ausgabe von Ingeborg Bachmanns Todesarten-Projekt. Opladen 1998: Westdeutscher Verlag

(14) Die vierbändige Werkausgabe von Ingeborg Bachmann fällt unter eine Kategorie, die der Osnabrücker Germanist Winfried Woesler mit drei Kriterien gegen die "wissenschaftliche" Edition abgrenzt. Nach Woesler haben die Herausgeber

1. die Vorstellungen der Erben zu berücksichtigen.
2. Sie bekommen nicht jede Publikationserlaubnis.
3. Sie müssen vermeintlich Kompromittierendes weglassen.

"Vermeintlich Kompromittierendes" ist eine Variable, je nach Lebenswelt, Welterfahrung, Literaturerfahrung, Tabuisierung und Selbstzensur...

(16f) Zur Veranschaulichung des eingangs erwähnten Konflikts zwischen dem Anliegen der Diskretion und der "Preisgabe" eines Nachlaßtextes, soll jedoch anhand eines Einzelbeispiels, der "Briefe an Felician", gezeigt werden, welche Folgen sich aus der Zuordnung in die Kategorie entweder des freigegebenen oder des gesperrten Nachlasses herleiten.

Die 36 handgeschriebenen Blätter, die nach Datierung (Mai 1945 bis April 1946), Anrede und Inhalt als Liebesbriefe der jungen Ingeborg Bachmann zu erkennen waren, sollten auf Wunsch der Erben als *Privatbriefe* klassifiziert und demgemäß dem großen Korpus der Bachmannschen Privatkorrespondenz zugeordnet werden, die mit einer Sperrfrist von 50 Jahren belegt worden war. Gegen diese private Lesart gelang es, auf der Freigabe der Briefe zu bestehen, mit dem Argument, daß es sich um Briefe an einen *imaginären Geliebten* handelt, die der Autorin Bachmann, nicht der Privatperson zugeschrieben werden müßten. So konnten die Briefe als *Briefprosa* vor der Sperrung bewahrt werden.

Mitte der achtziger Jahre ergab sich aus der eingehenden Beschäftigung mit den Briefen der Vorschlag an den Piper Verlag zu einer kommentierten Ausgabe der Briefe, unter dem Titel "Briefe an Felician", der die Erben zustimmten. Es hatte sich gezeigt, daß wir es nicht mit einer beziehungslos für sich stehenden "Prosaarbeit" zu tun hatten, sondern daß den Briefen ein philosophisch-literarisches Fundament zugrunde liegt, aus dem die junge Autorin Inspiration und den sprachlichen Gestus gewonnen hat. Es sind die moraltheoretischen und ästhetischen Schriften Friedrich Schillers und der "Tugendkatalog" der Frau des 18. Jahrhunderts, dem die ekstatischen Liebesformeln der Briefe entlehnt sind: Die Tugend der *Unterwerfung*: "Du sollst kommen und einen Willen über mich werfen"; die Tugend der *Dienstbarkeit*: "Nie wäre ich besser bereit zu dienen als jetzt"; die Tugend des *Opfers*: "Was käme inniger aus mir als Dir zu opfern".

In der Folgezeit erhellt sich jedoch auch der bis dahin unerkannte biographische Ursprung der Briefe. Hinter dem Bild des "komponierten" Geliebten Felician trat der reale Geliebte hervor: Der Schriftsteller Josef Friedrich Perkonig (1890-1959), damals die zentrale Gestalt Kärntner Dichtung, Ingeborg Bachmanns Lehrer und berühmt als "der Dichter der Grenze".

zu Ziff. 3.05:

Komposition gegen den Autorwillen?

zu:

aus: HEIDELBERGER-LEONARD, I (ed.), "Text-Tollhaus für Bachmann-Süchtige?" Lesarten zur kritischen Ausgabe von Ingeborg Bachmanns Todesarten-Projekt. Opladen 1998: Westdeutscher Verlag

(26f) Warum wurde die Vielfalt der Entwürfe, der Fragmente, der Abbrüche, die aus radikal verschiedenen lebensgeschichtlichen Erfahrungen Ingeborg Bachmanns entstanden waren, eingebettet unter der Fiktion eines Projekts?

Daß die Entwürfe und Fragmente, die auf das sogenannte "Todesarten"-Vorfeld gehäuft wurden, inhaltlich weit entfernt sind von dem, was für Ingeborg Bachmann die "Todesarten" waren, liegt offen zutage. Das gilt für die Entwürfe "Ein Fenster zum Ätna" und die beiden Fragmente aus dem ersten Roman "Stadt ohne Namen" von 1950, den Ingeborg Bachmann überdies damals nicht "zurückgestellt", wie es in dem "Todesarten"-Projekt heißt, sondern vernichtet hat. Es gilt auch für die Entwürfe "Geschichte einer Liebe" und die Entwürfe zu dem sogenannten "Eugen-Roman I", entgegen der Ingeborg Bachmann unterstellten "übergreifenden", "vernetzten", "genetisch gut dokumentierten" Werkintention.

Ingeborg Bachmanns "Todesarten" beginnen erst mit der gescheiterten Beziehung zu Max Frisch und mit dem Malheur, sich gleich darauf mit jemand eingelassen zu haben, der in die Figur des Anton Marek eingegangen ist. Nichts, was sie bis dahin geschrieben hatte, dürfte rechtmäßig den Titel "Todesarten" führen oder unter dem Titel vereinnahmt werden.

Wie in Franz Kafkas Leben das Jahr 1912 die entscheidende, zum Tode führende Zäsur bezeichnet, die sein Werk *zerteilt*, so zerteilt das Jahr 1962 Leben und Werk von Ingeborg Bachmann. Sie war um ihre "Lebensselbstsicherheit" gebracht worden und seitdem, wie ihre Figuren, von pathologischer Angst und auch von wehrlosem Haß gepeinigt - was die neu eingefügten Textpassagen in der Edition von 1995 offenbaren.

zu Ziff. 3.061:

Disziplin: Textgenetik

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editionswissenschaft 4. Bern, Frankfurt 1999.

(13) Folgeschwerer für die Geschichte der Naturwissenschaften war das Schicksal der Kiste mit den Handschriften von Newton. Der Begründer der modernen Physik hatte 1796 (? , 1696 - HS), vor seiner Übersiedlung von Cambridge nach London, die Handschriften seiner nicht veröffentlichten Werke darin untergebracht. Seine Nachkommen überließen sie 140 (? 240 - HS) Jahre lang einem Dornröschenschlaf, bis sie sich 1936 zum Verkauf entschlossen. Der bekannte Wirtschaftswissenschaftler J.M. Keynes interessierte sich dafür, erwarb sie und vermachte sie schließlich dem Trinity-College in Cambridge. Zu seiner Verwunderung enthielt die Kiste jedoch weder physikalische noch mathematische Schriften, sondern der Theologie und Alchimie zugehörige spekulative Abhandlungen. Newton hatte sie wohlwissend seiner Mitwelt vorenthalten. Sein naturwissenschaftliches Werk, und nur dieses, wurde wegweisend für das Zeitalter der Moderne. Der Riegel, mit dem die Kiste verschlossen war, wurde zur Metapher eines in sich abgeschlossenen Weltbildes.

(22) Gegenstand der "critique génétique" sind geschriebene, im allgemeinen handschriftliche Dokumente, die, in bestimmte Zusammenhänge eingeordnet, die "Urgeschichte" eines Textes und die sichtbare Spur eines schöpferischen Prozesses darstellen. Diese Handschriften mit ihrer janusköpfigen Materialität haben seit langem eine seltsame Anziehungskraft ausgeübt: auf die Schriftsteller, die sie aufbewahren (als wollten sie sich so ihres eigenen Lebens versichern), auf die Archive, die sie erwerben (als Zeugen nationalen Erbes), auf die Liebhaber, die sie sammeln (als Kunstobjekt oder als Geldanlage), auf die Philologen, die sie erschließen (für den kritischen Apparat wissenschaftlicher Ausgaben)...

Es bedarf sicherlich einer recht außergewöhnlichen Neugier, bei der "passion" und "patience", Leidenschaft und Geduld, gleichermaßen erforderlich (23) sind. Die Leidenschaft, einem Text so nah wie möglich zu sein, da man ja gewissermaßen seiner Wieder-Geburt beiwohnt; die Leidenschaft, die Authentizität des Autographs sozusagen mit den Händen zu fassen und dem Akt der Einschreibung bei-zuwohnen; flüchtige und uneingestandene Leidenschaft, sich während des Hinabsteigens zu den Ursprüngen des Textes mit dessen Schöpfer zu identifizieren; Leidenschaft, in den verbotenen Raum der Kulissen einzudringen, und kriminalistisch angehauchte Lust, das Geheimnis der Herstellung zu enthüllen: Der Forscher wird auf der Hut sein müssen, um der Versuchung durch Psychologismus, Voyeurismus und Fetischismus zu widerstehen...

Neben der Leidenschaft ("passion") bedarf es jedoch auch einer grenzenlosen Geduld ("patience"). Die Geduld, sich tatsächlich auf die Suche nach einer bestimmten Handschrift zu machen, deren Standort nicht selten mit Wechselfällen der Geschichte zu tun hat oder aber in kleinere Geschichten von Verkauf, Erbschaft und Erbfolgerecht verstrickt ist; Geduld, die Manuskripte zu entziffern, zu ordnen und zu transkribieren. Demut vor der Fülle der Materialien samt ihren oft kaum entwirrbaren Problemen; Geduld des Gelehrten im Umgang mit einem Dokument, zu dem er die rechte Distanz finden muß, damit aus dem Objekt der Leidenschaft ein Objekt der Erkenntnis wird; Geduld schließlich des Herausgebers beim Versuch, die Genese des Textes zu rekonstruieren und sichtbar zu machen.

zu Ziff. 3.062:

Druck vs. Arbeitshandschrift

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editionswissenschaft 4. Bern, Frankfurt 1999.

(45) Literarische Handschriften weisen in der Tat eine materielle und semiotische Vielfalt auf, die nur wenig mit der einer Druckseite gemein hat. Einer regelmäßigen, rechts- und linksbündig angeordneten Druckseite, mit ihrer über das ganze Buch gleichbleibenden Zeilenzahl, ihren stets gleich gesetzten Druckbuchstaben und dem sich deutlich vom weißen Rand abhebenden Satzspiegel, stellt das Manuskript eine Schrift "im Rohzustand" gegenüber, wo jede Seite ihre eigene Form und Semiotik besitzt: freie Aufteilung des Schreibraums, freies Spiel mit der Schreibrichtung sowie mit Länge und Anzahl der Zeilen, Streichungen und Erweiterungen, deren Schreibspur nicht selten eine besonders impulsive und affektgeladene Verfassung des Schreibers verrät... Gegenstand der textgenetischen Forschung ist vielmehr die Arbeitshandschrift, welche die Spuren der "allmählichen Verfertigung" trägt, einer Handlung, eines Produktionsablaufs, eines Schöpfungsaktes mit seinem Vorankommen und seinen Blockaden, seinen Erweiterungen und Tilgungen, seinen ungezügeln Impulsen und Zurücknahmen, seinen Aufschwüngen und Zögerlichkeiten, seinen Maßlosigkeiten und Mängeln, seinem Gewinn- und Verlustgeschäft. Gegenstand ist das Manuskript mit all dem "Schlamm" (boue) und der "Aufwallung" (ébullition), worauf die doppelte (47) Etymologie des Wortes "Brouillon" hinweist. Gegenstand sind Aufzeichnungen aller Art, auch die "Kritzeleien" (griffouillis), wie Louis Aragon seine ersten, heimlich aufgezeichneten Schreibspuren nannte, Notate eines Kindes, das "niemals schreiben gelernt" hat.

zu Ziff. 3.0621:

Vorgeschichte eines Manuskripts / ambivalente Wertungen

zu: vprag2.1474

SCHIFFER, Eva: Zwischen den Zeilen. Manuskriptänderungen bei Thomas Mann. Transkription und Deutungsversuche. Berlin 1992.

(11) Man erhofft sich von der kritischen und einigermaßen systematischen Betrachtung eines mit Korrekturen versehenen Manuskripts einen Einblick in die Arbeitsweise, aber - was wichtiger ist - auch in stilistische und thematische Motivierungen des Autors, die einem die bereinigte (oder aber mit Druckfehlern behaftete) Druckseite verschweigt.

(12f) So nimmt z.B. wie jedes erzählerische Werk von Thomas Mann der Anfang des *Doktor Faustus* in nuce eine Fülle von Motiven des Romans vorweg. Sieht man zunächst von diesen ab und betrachtet statt dessen die ersten Seiten, in denen Zeitblom sich langatmig, vorstellt, rein vom Sprachlichen her, so fällt, die ausgesprochene Häufung von Formulierungen wie den im folgenden gesperrt wiedergegebenen auf:

" M i t a l l e r B e s t i m m t h e i t w i l l
i c h v e r s i c h e r n , d a ß e s k e i n e s w e g s
a u s d e m W u n s c h e g e s c h i e h t , m e i n e P e r s o n i n d e n V o r d e r -
g r u n d z u s c h i e b e n , w e n n i c h d i e s e n M i t t e i l u n g e n ü b e r d a s
L e b e n . . . A d r i a n L e v e r k ü h n s , d i e s e r e r s t e n u n d g e -
w i ß s e h r v o r l ä u f i g e n B i o g r a p h i e . . .
e i n i g e W o r t e ü b e r m i c h s e l b s t u n d m e i n e B e w a n d n i s s e
v o r a u s s c h i c k t e . E i n z i g d i e A n n a h m e b e s t i m m t m i c h
d a z u , d a ß d e r L e s e r - i c h s a g e b e s s e r :
d e r z u k ü n f t i g e L e s e r ; d e n n f ü r d e n A u g e n b l i c k b e s t e h t
j a n o c h n i c h t d i e g e r i n g s t e A u s -
s i c h t , d a ß m e i n e S c h r i f t d a s L i c h t d e r Ö f f e n t l i c h k e i t
e r b l i c k e n k ö n n t e , . . . e s s e i d e n n , d a ß
s i e d u r c h e i n W u n d e r . . . ; i c h b i t t e w i e d e r a n s e t z e n z u
d ü r f e n : n u r w e i l i c h d a m i t r e c h n e , d a ß m a n w ü n s c h e n
w i r d , ü b e r d a s W i e u n d W a s d e s S c h r e i b e n d e n b e i -
l ä u f i g u n t e r r i c h t e t z u s e i n , s c h i c k e i c h d i e s e n
E r ö f f n u n g e n e i n i g e w e n i g e N o t i z e n ü b e r
m e i n e i g e n e s I n d i v i d u u m v o r a u s , n i c h t o h n e
d i e G e w ä r t i g u n g f r e i l i c h , g e r a d e d a d u r c h d e m L e s e r Z w e i -
f e l z u e r w e c k e n , o b e r s i c h a u c h i n d e n r i c h t i g e n
H ä n d e n b e f i n d e t . . . "

Hier wird jede positive Wendung durch eine negative, jede bestimmte durch eine verunsichernde neutralisiert: "mit aller Bestimmtheit" durch "keineswegs", "sehr vorläufig" durch das alles andere als gewisse "gewiß", "einzig" durch das zurücknehmende "ich sage besser...", "nicht die geringste Aussicht" durch "es sei denn...". Die negative Formulierung, "daß es keineswegs aus dem Wunsche geschieht..." wird mit neuem, stark positivem Ansatz umformuliert: "nur weil ich..." - jedoch wird dieser positive Ansatz wieder ins Negative gezogen durch die Worte "nicht ohne die Gewärtigung...", "einige" wird diminuiert durch den Zusatz "wenige", "unterrichtet" durch "beiläufig", "in den richtigen Händen" verunsichert durch den Vorsatz "auch".

Daß diese ambivalente Ausdrucksweise charakterisierend für Zeitblom ist, liegt auf der Hand; daß gerade von den acht negierenden oder verunsichernden Wendungen drei im Manuskript als nachträglich vorgenommene - wenn in zwei Fällen auch nur umformulierte - Ein-

schübe erscheinen, mag bloß beweisen, wie bewußt Thomas Mann diese Ambivalenz Zeitbloms stilistisch betonte.

Es mag aber auch zur Nachdenklichkeit stimmen, daß diese Ambivalenz eine ins Negative tendierende ist, nicht etwa ein Nein - oder vielleicht doch Ja, sondern durchweg ein Unbedingt Ja - oder vielleicht doch eher Nein. Zu untersuchen, inwiefern sich eine solche - oder eine andere etwa für Thematik oder Charakterisierung der Personen aufschlußreiche - Tendenz konsequent nachweisen ließe, war das Ziel der vorliegenden Arbeit.

zu Ziff. 3.0622:

Ilias - Textgenese

zu:

aus: HOMER, Ilias. Übertragen von Raoul Schrott. Kommentiert von Peter Mauritsch. München 2008.

(XXIII) Daß wir es bei der *Ilias* mit einem Text zu tun haben, ver-raten auch die Quellen des Epos. Es arbeitet neben mündlichen Stoffen wie den *Kypria*, der *Aithiopsis* und anderen griechischen Sagen (die bei Bellerophon oder der Geschichte der Aitoler, Pylener und Epeier bereits kilikische Bezüge zeigen) Erzählgut auf, das wir auch im Alten Testament wiederfinden. Dazu fließen - als kilikisches Volksgut - hethitische und hurritische Mythentexte ein sowie die schriftlich fixierten Epen des Zweistromlandes: die Geschichte des weisen Mannes *Atrahasis*, der Weltschöpfungsmythos des *Enuma Elish*, die Erzählung vom Kriegsgott *Erra* und vor allem der *Gilgamesh* als Nationalepos des Ostens dienen Homer, seine eigene Erzählung auszugestalten. Darüber hinaus benützt er die in ihrem literarischen Ton bereits an die *Ilias* gemahnenden assyrischen Annalen als Materialbruch; sie verbreiteten damals die Ereignisse der jährlichen Militärkampagnen der Herrscher von Sargon bis Assurbanipal in alle Landesteile.

Das zeitliche Umfeld des *Gilgamesh*- dem nach Sargons Tod im Jahre 705 eine zwölfte Tafel ihrer Unterweltschilderung angehängt wurde, welche die Vorlage für das Erscheinen von Patroklos' Geist abgibt -, des *Erra* sowie die Übernahme von Motiven aus den assyrischen Annalen der ersten Hälfte des 7. Jahrhunderts, legen einen direkten Zugang Homers zu ihnen nahe. Eine Vermittlung all dieser Stoffe durch das bloße Hörensagen oder phönizische Übersetzungen nach Griechenland ist auszuschließen. Von den Phöniziern ist keine einzige literarische Zeile erhalten, und selbst wenn sie dies alles übertragen hätten, dann nicht in textgetreuer Form. Wie uns die hethitische Fassung des *Gilgamesh* oder die etwa zeitgleich zur *Ilias* entstandene Einarbeitung der Sintflutgeschichte durch die Judäer - als weiteres phönizisches Nachbarvolk - in die biblische Genesis demonstrieren, wurden Texte von einer anderen Kultur stets adaptierend umgeformt.

zu Ziff. 3.063:

Textkritik vs. Textgenetik

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editions-wissenschaft 4. Bern, Frankfurt 1999.

(125) Moderne literarische Handschriften fanden, historisch gesehen, in zwei Schüben Eingang in die Literaturwissenschaft: über die Philologie und über die "critique génétique".

Zunächst nur kulturgeschichtliches Objekt, wurden sie zum Gegenstand der Textkritik, und dies vor allem Siegeszuge der Philologie: in der italienischen Renaissance und dann insbesondere in der deutschen Philologie des 19. Jahrhunderts und den damit verbundenen ersten großen historisch-kritischen Ausgaben.

Neben dieser Tradition und relativ unabhängig von ihr bildete sich, wie erwähnt, in Frankreich in den siebziger Jahren des 20. Jahrhunderts eine weitere auf Handschriften bezogene Fachrichtung heraus: die hier zur Diskussion stehende "critique génétique". Ihr Hauptziel ist nicht die Herstellung des Textes, sondern das Freilegen der Schreibmechanismen, die dem Schaffensprozeß zugrunde liegen.

(126) Der Einfachheit halber nehmen wir also eine bestimmte Zahl von Fällen an, von denen behauptet werden darf, daß die textgenetische Dokumentation verhältnismäßig vollständig ist. Die gängigen Entstehungsmaterialien sind verschiedenster Art: Quellenexzerpte, Wortlisten, Notizbücher, Pläne, Szenarios, Entwürfe, Zusammenfassungen, mehr oder weniger fortgeschrittene Versuche der Textniederschrift, aufeinanderfolgende Textfassungen, letzte Revisionen, eigenhändige Abschriften, von einem Kopisten angefertigte Abschriften, Druckvorlagen, Korrekturfahnen und schließlich vom Autor revidierte Ausgaben mit handschriftlichen Korrekturen.

(131) Alle diese textgenetischen Schriftstücke gehören zu dem, was hier als *interne Werkgenese* bezeichnet werden soll. Sie stellt das Hauptforschungsfeld der "critique génétique" dar. Mehr oder minder systematisch ergänzt der Textgenetiker diese Schriftstücke durch andere Dokumente, die schriftlich oder mündlich, autograph oder allograph sein können. Gemeint sind Informationen, welche neben den eigentlichen Werkhandschriften dazu beitragen, die Genese aufzuhellen: Zeugnisse von Freunden, Erwähnungen im Briefwechsel, Interviews, Tagebücher, Autobiographien, kurz Dokumente, die zur *externen Genese* des Werkes gehören und die in Anlehnung an die Terminologie Gérard Genettes als "péri-avant-texte" bezeichnet werden könnten.

(133) Trotz dieser Vielfalt an genetischen Materialien kann man davon ausgehen, daß es, grob gesehen, zwei Arten gibt, einen Text zu produzieren: *das produktorientierte Schreiben* und *das prozeßorientierte Schreiben*. Ersteres ist bei Autoren belegt, die den Text nach einem vorher festgelegten Programm abfassen; ein typisches Beispiel dafür ist Zola. Letzteres findet sich bei Autoren, die sozusagen nichts wissen, bevor sie sich in das Abenteuer des Schreibens stürzen; die "inventio" fließt sofort in die Hand, die über das Papier eilt; ein typisches Beispiel hierfür ist Proust.

zu Ziff. 3.064:

Disziplin: Textgenese vs. Edition

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editions-wissenschaft 4. Bern, Frankfurt 1999.

(217) Wie die "critique génétique" den "avant-texte zum Gegenstand hat, so hat die Editions-wissenschaft den Text zum Gegenstand. Und nur wenn man diesen Unterschied ernst nimmt, werden an die Stelle der gegenseitigen Mißverständnisse und Fehleinschätzungen fruchtbare Dialoge treten können. Heißt das nun, daß der Textgenetiker nicht Herausgeber sein kann und der Herausgeber nichts von textgenetischen Verfahrensweisen versteht? Eine solche Rollenverteilung wäre absurd, und manches spricht geradezu für das Gegenteil...

So könnte man die Textgenetiker fragen, warum sie einerseits *den* einen Text "entweihen", indem sie nicht nur dessen wirkliche, sondern auch dessen mögliche oder unvollendete Formen und genetische Stadien aufwerten, und *im gleichen Atemzug* Editionsprojekte - zum Beispiel "Pléiade"-Ausgaben - mit der damit verbundenen Forderung nach einer allein gültigen Textgrundlage unterstützen. Und die Editoren könnte man fragen, warum sie darauf beharren, den Variantenapparat immer komplizierter zu gestalten, ihn mit textgenetischen Informationen zu überfrachten, obwohl eine solche Überfrachtung häufig an die Grenze der Lesbarkeit stößt. Der Herausgeber stellt dem Textgenetiker immer wieder dieselbe Frage: Wozu (218) befaßt man sich mit den Handschriften, wenn nicht zum Zwecke einer Edition? Und der Textgenetiker fragt den Herausgeber einer kritischen Ausgabe, warum und für wen er Variantenapparate veröffentlicht, die trotz aller Vollständigkeit den Verlauf der Textgenese nicht ausreichend wiedergeben.

zu Ziff. 3.065:

"editio princeps" vs. Ausgabe "letzter Hand"

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editions-wissenschaft 4. Bern, Frankfurt 1999.

(223) Doch nicht alle Vorhaben werden zum Abschluß gebracht, zumal das Prinzip "letzter Hand", bisher als testamentarischer Ausdruck, als letzte und vom Autor autorisierte Schreibspur hochgehalten, mehr und mehr umstritten ist. Stattdessen kommt den Frühfassungen immer größere Bedeutung zu. Zum einen durch die Wiedergabe der editio princeps, die als originaler Ausdruck der ersten vollendeten Textform angesehen wurde, und zum anderen infolge des wachsenden Interesses für handschriftliche Entwürfe und Skizzen, die - besser als der erstarrte Text - die Kreativität sozusagen im Urzustand zu zeigen vermögen. Der Kult der Ausgabe letzter Hand weicht dem Mythos der ersten handschriftlichen Spur. Zur Zeit Freuds ahnen die Editoren dunkel die verborgene Wahrheit der frühen Stufen der Handschriften, die Enthüllungen der ersten Skizze, die oft genug ausgestrichen und verworfen wurden. Vom Durchstreichen bis zum Freudschen Konzept der "Verneinung" war es nur ein Schritt...

zu Ziff. 3.066:

Textgenetische Transkription

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editions-wissenschaft 4. Bern, Frankfurt 1999.

(157) Ihr Zweck: ein leichtes Lesen dessen zu ermöglichen, was oft sehr mühevoll entziffert werden muß. Der Idealfall ist ein Paralleldruck, in dem Faksimile und Transkription quasi gleichzeitig lesbar sind. Das hat nichts mit Positivismus, weder mit altem noch mit neuem, zu tun, es handelt sich vielmehr um eine "Erleichterungsstrategie", wie es in der Terminologie der Psychologen heißt.

Wir plädieren hier für die diplomatische Transkription, nur sie bewahrt etwas von der Topographie des Originals. Hier einige Hinweise zur Erstellung diplomatischer Transkriptionen:

- Jede Transkription ist, wie es Jean-Louis Lebrave formuliert (*Proust à la lettre*, 1990, S.143), "zugleich reicher und ärmer als die zugrundeliegende Handschrift". Reicher, weil sie schon einen gewissen Grad von Analyse enthält. Denn um etwas wiederzugeben, genügt es nicht, einfach zu kopieren, (...162) man muß die graphischen Zeichen zunächst begreifen und dann in Schreibprozesse zurückübersetzen. Und ärmer, weil jede maschinenschriftliche Transkription unwiederbringlich verliert, was das handschriftliche Schreiben an affektiver Aufladung (Hast, Stocken, Angst, Jubel) und an Hinweisen auf die Schriftzüge (wechselnde Stärke der Linienführung, Wechsel zwischen Tinte und Bleistift, zwischen schwarz und farbig, Veränderung des Duktus usw.) enthält. Nichts kann also die Hinzuziehung des Originals ersetzen, und sei es nur zur Überprüfung.
- Jede Transkription muß die Gesamtheit des Originals buchstabengetreu wiedergeben, einschließlich seiner bisweilen seltsamen oder fehlerhaften Orthographie und Interpunktion sowie seiner metaschriftlichen Zeichen (Einfügungs- und Verweiszeichen, Sternchen, Unterstreichungen usw.). Eventuell vorhandene Skizzen oder Zeichnungen müssen mitgeteilt werden (entweder durch Reproduktionen oder entsprechende Anmerkungen).
- Das Ziel einer Transkription ist nicht Perfektion, sondern Perfektibilität. Mit anderen Worten, eine Transkription ist brauchbar, auch wenn sie ein paar Fehler enthält. Zwei Prinzipien sollten jedoch beachtet werden: 1 / die Dossiers haben Faksimile und Transkription im Paralleldruck wiederzugeben; 2 / die Transkription muß als Diskette verfügbar sein, damit jeder neue Benutzer Korrekturen einfügen kann.
- Der Nutzen einer Transkription beruht auf ihrer Transparenz; sie ist um so sicherer garantiert, als sie mit einem Minimum an Zeichen und Siglen auskommt. Dabei erlaubt es die Computertechnik inzwischen, die Transkription dem Original graphisch möglichst treu nachzubilden.
- Die Transkription ist in gewöhnlicher Antiqua vorzunehmen, Kursivschrift wird lediglich bei eventuellen Eingriffen des Transkriptors verwendet (s.u.: "nicht entziffert").
- Eine Streichung in der Handschrift wird durch eine Streichung wiedergegeben, z.B.:

"er hat sein gesamtes Erbe verkauft".

Gänzlich gestrichene Abschnitte sind genauso darzustellen: durch einen großen, mit der Hand ausgeführten Durchstreichungs-

- strich... Bei jedem anderen wichtigen metaschriftlichen Zeichen ist auf dieselbe Weise zu verfahren (z.B. Einfügungszeichen).
- Die Erweiterung ist im allgemeinen an der Position zu erkennen: und kleinen
"Er hat seine großen Fehler korrigiert".
- Die Varianten im zwischenzeiligen Raum und am Rand können in kleinerer Schrift wiedergegeben werden als der Text auf der Hauptzeile. Diese graphische Unterscheidung bedeutet nicht, daß das in einer kleineren Schrift Transkribierte "weniger wichtig" sei.
- Der Wechsel der Schreibwerkzeuge (Tinte / Bleistift; Farbstifte) sowie jede andere materielle Besonderheit (gefaltetes Blatt, verkehrt herum benutztes Heft) sind in Fußnoten zu vermerken.
- Nicht entzifferte Textstellen sind mit "nicht entziffertes Buchstabe", "nicht entzifferte Silbe", "nicht entziffertes Wort", "drei nicht entzifferte Wörter" usw. zu kennzeichnen.
- Überschreibungen von A durch B sind durch Schrägstrich zwischen A und B als solche auszuweisen:
"Er hat sein/ihr Vermögen verkauft".
- Unsichere Lesungen sind mit Sternchen zu versehen:
"Er hat seine *zahlreichen* Besitztümer verkauft".

zu Ziff. 3.067:

Textgenetischer Apparat

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editionswissenschaft 4. Bern, Frankfurt 1999.

(227) 1958 stellt Hans Zeller ein neues Apparatmodell vor, das in der Ausgabe von C.F. Meyer zur Anwendung kommt. Die wesentlichen Prinzipien Beißners werden darin beibehalten. Eine möglichst vollständige Variantendarstellung sowie das Prinzip der syntagmatisch und paradigmatisch fortschreitenden Textentwicklung. Gleichzeitig wird jedoch das bestehende Modell erweitert und modifiziert. Für Zeller geht es nicht mehr um das "ideale Wachstum", sondern um eine reale, materiell abgesicherte Textgenese. Folglich enthält die Edition alle Angaben, mit deren Hilfe nicht nur die Zeugen selbst, jedes einzelne Blatt sowie jede einzelne Zeile auf dem Blatt, sondern auch die genaue Stelle aller Umarbeitungen (darüber, darunter, am Rand usw.) identifiziert werden können. Damit fügt dieser Editionstyp - zur Zeit zweifelsohne der vollkommenste - der chronologischen die topographische Dimension hinzu. Außerdem wird mit besonderen Siglen gekennzeichnet, wenn eine zweite Lesart oder eine andere chronologische Folge der Varianten vertretbar ist oder wenn der Autor in einem bestimmten Stadium der Textgenese eine Alternative unentschieden beibehält. So liegt die ganze Textgenese einschließlich ihrer typographischen Wiedergabe offen vor Augen. Einen derartigen kritischen Apparat herzustellen heißt, alles, was man beim Lesen und Analysieren der Handschriften gesehen, begriffen und gefolgert hat, exakt und erschöpfend zu erfassen und im Druck wiederzugeben. Von daher auch die Feststellung, Zellers Apparat sei einem mit der Präzision eines Computers versehenen Algorithmus vergleichbar...

(228) Im Bereich der kritischen Edition kam die Textgenese in dem Augenblick wissenschaftlich zum Tragen, als eine technische Erfindung - das Faksimile - die Handschriften materiell reproduzierbar gemacht hat und so immer mehr Forscher sie wissenschaftlich untersuchen konnten. Dagegen scheint es heute fast, als versetze eine in der Apparatdarstellung auf die Spitze getriebene Technik der kritischen Ausgabe in Buchform den Todesstoß. Vielleicht ist es das Gebot der Stunde, wiederum auf eine neue Technik überzugehen und die kritischen Apparate ausschließlich als Computerprodukt herzustellen, wobei selbstverständlich der Text nach wie vor auch in Buchform verfügbar sein muß. Damit könnten jedenfalls die Kosten gesenkt und der Kommentar immer auf den neuesten Stand gebracht werden. Gleichzeitig käme dies einer ganzen Reihe von Lesern entgegen, die zwar Ansprüche an die Qualität der Textaufbereitung stellen, aber offen zugeben, daß sie mit den Variantenapparaten (229) nichts anzufangen wissen. Eine andere Möglichkeit bestünde darin, zwar schon alle Texte eines Autors zu edieren, jedoch für die vollkommene Darstellung der Genese einen bestimmten Text zu wählen, an dem sich die Arbeitsweise des Schriftstellers besonders anschaulich darstellen läßt.

zu Ziff. 3.068:

Genetische Computer-Edition

zu:

aus: GRÉSILLON, A, Literarische Handschriften. Einführung in die "critique génétique". Arbeiten zur Editionswissenschaft 4. Bern, Frankfurt 1999.

(244) Mit der Entdeckung des Hypertextkonzeptes schafft die Informatik in der Tat Möglichkeiten für die genetische Edition, von denen man vor zehn Jahren nicht einmal zu träumen gewagt hätte. Diese Neuerung überschneidet sich überraschenderweise mit den Ergebnissen theoretischer Überlegungen, die zur gleichen Zeit und völlig unabhängig davon in der Literaturwissenschaft angestellt wurden. Der von Gérard Genette geprägte Begriff des "Hypertextes" bezeichnet eine Reihe von literarischen Werken, die miteinander verbunden sind oder voneinander abstammen und durch ein gemeinsames, Hypotext genanntes Modell, von dem sich alle herleiten, zusammengehalten werden. Das Prinzip der elektronischen Hypercard-Software unterscheidet sich kaum davon. Es beruht seinerseits auf dem Begriff des elektronischen Hypertextes und (245) stellt nichts anderes dar als die vielfältige, flexible und stets veränderbare Vernetzung einer beliebigen Anzahl von Texten. Dadurch können im Speicher eine Reihe von Dokumenten durch ein Netz von Beziehungen und von Merkmalen miteinander verbunden werden.

Der Nutzen dieses Systems für die textgenetische Forschung liegt auf der Hand. Man braucht nur alle Teile eines "dossier génétique" als eine Ansammlung von Hyperdokumenten anzusehen, die im Speicher des Computers verfügbar und auf dem Bildschirm abrufbar sind. Damit ist die Dynamik des Schreibprozesses, die keine Buch-Edition zu vermitteln vermag, reproduzierbar geworden...

Darüber hinaus bietet die Informatik weitere, erst in jüngster Zeit erschlossene Möglichkeiten. Die Datenerfassung ist nicht auf den Text beschränkt, sie läßt sich ebenso auf Bild und Ton anwenden. Mit einem guten Scanner können Handschriften digitalisiert werden; bei Theaterstücken kann man Ton und Bild der Proben und der Premiere sowie die vom Autor durchgesehene und während der Proben revidierte Textfassung einspielen. Damit nicht genug. Man stelle sich zum Beispiel auf dem Gebiet der Oper die Genese des *Rosenkavaliers* vor, ein Werk, das in enger Zusammenarbeit zwischen Hugo von Hofmannsthal und Richard Strauss entstanden ist. Auf dem Bildschirm (246) läßt sich ein bestimmtes, dem Komponisten vorgelegtes Textstadium abrufen, dem der Komponist einen ersten Entwurf der Partitur beigibt. Von dieser wiederum kann man sowohl die Notenhandschrift sehen als auch die Vertonung hören. Alles Utopie? Keineswegs. Das Projekt ist auf dem Wege der Realisierung. Einer der beiden Herausgeber von Hofmannsthals *Rosenkavalier* entschloß sich, im Anschluß an die Buch-Edition eine textgenetische Multimedia-Edition in Angriff zu nehmen...

(247) Eine letzte technische Innovation, scheinbar reine "Spielerei", und doch von entscheidender Bedeutung, ist die Unterteilung des Bildschirms in mehrere Fenster, in denen man nunmehr *simultan* verschiedene Dokumente zeigen kann. Die Möglichkeit, gleichzeitig sechs Fenster zu aktivieren, um z.B. den ersten Wurf zu lesen - in einem Fenster als Faksimile, im nächsten in der Transkription und in einem dritten mit dem Kommentar des Herausgebers - und mit zwei entsprechenden Brouillons und dem endgültigen Manuskript (jeweils in einem eigenen Fenster) zu vergleichen, erweitert die Darstellung der Textgenese in der Tat um eine wesentliche Dimension: die der Zeit.

zu Ziff. 3.07:

Synopse

zu:

aus: WONNEBERGER, R, Normaltext und Normalsynopse. Neue Wege bei der Darstellung alltestamentlicher Texte: Zeitschrift für Sprachwissenschaft 3 (2/1984) 203-233.

(230) Da die Unterschiede zwischen ähnlichen Zeilen einer Synopse von verschiedener Art sind, stellt sich die Aufgabe, eine Skala der Ähnlichkeitsgrade zu entwickeln. ...

- i) *oberflächengleich*: die Texte haben ein identisches Erscheinungsbild. Wenn sie in unterschiedlichen Sprachen erfaßt sind, können sie dann als oberflächengleich angesehen werden, wenn sich sinnvollerweise keine Übersetzungsalternativen vertreten lassen (philologisch exakte Übersetzung).
- ii) *basisgleich*: die Aussagen unterscheiden sich nur in Kleinigkeiten, die am Grundbestand der Aussage nichts ändern. Dafür kommen in Frage kleine Einfügungen ... oder kleinere Änderungen in der Wortstellung.
- iii) *sinngleich*: die Aussagen sind zwar sprachlich unterschiedlich, formulieren aber offensichtlich denselben Sachverhalt.
- iv) *funktionsgleich*: die Aussagen unterscheiden sich sowohl in der Formulierung als auch im Sachverhalt, der Sachverhalt hat aber im Textgefüge dieselbe Funktion.

ungleich: alle anderen Fälle werden als ungleich angesehen
Bsp. Ri 6,8bβ-9a

"So hat gesprochen Jahwe, der Gott Israels: Ich selbst habe her-
ausgeführt *euch* aus Ägypten und habe gebracht euch aus dem Hause
der Sklaven und habe euch errettet aus der Hand Ägyptens und aus
der Hand aller *euch Bedrängenden*. "

1 Sam 10, 18aβγb

"So hat gesprochen Jahwe, der Gott Israels: Ich selbst habe her-
ausgeführt *Israel* aus Ägypten und habe euch errettet aus der Hand
Ägyptens und aus der Hand aller *Königreiche der bedrängenden
euch*."

Segmentierung und Programm zur Parallelstellung

(1) So hat gesprochen
Jahwe, der Gott Israels:
(2) Ich selbst habe
herausgeführt
Israel aus Ägypten
(3) und habe gebracht euch
aus dem Hause der Sklaven
(4) und habe euch errettet
aus der Hand Ägyptens
und aus der Hand aller
*Königreiche der bedrängenden
euch*.

(1) So hat gesprochen
Jahwe, der Gott Israels:
(2) Ich selbst habe
herausgeführt
euch aus Ägypten
(3) und habe gebracht euch
aus dem Hause der Sklaven
(4) und habe euch errettet
aus der Hand Ägyptens
und aus der Hand aller
euch Bedrängenden

3.1 Ausgangspunkt: Kritische Aneignung des Endtextes (Textkritik, Handschriften-/Druckausgabenkunde, Editionsphilologie), Wahl der weiter interessierenden Textfassung

Literatur

- ALAND, K; ALAND, B: Der Text des Neuen Testaments. Einführung in die wissenschaftlichen Ausgaben sowie in Theorie und Praxis der modernen Textkritik. Stuttgart 1982: Deutsche Bibelgesellschaft
- AUDENART, N; FURUTA, R: Annotated Facsimile Editions: Defining macro-level structure for image-based electronic editions. *Literary & Linguistic Computing* 24/2 (2009) 143-151.
- BEIN, T; NUTT-KOFOTH, R; PLACHTA, B (ed./eds.): Autor-Autorisation-Authentizität. Beihefte zu editio 21, Germanistik, Philosophie, Musikwissenschaft. Tübingen 2004 ISBN 3-484-29521-x: Niemeyer // I.2.1
- BOWMAN, A K; DEEGAN, M . *Literary and Linguistic Computing* 12 / 3 (ed./eds.): Special Issue on the Use of Computers in the Study of Ancient Documents (1997)
- BRAUN-RAU, A: William Shakespeares "King Lear" in seinen Fassungen. Ein elektronisch-dialogisches Editionsmodell. Beihefte zu editio 20. Tübingen 2004 ISBN 3-484-29520-1: Niemeyer // I.2.1
- CANFIELD, K: The Navajo Language Literature Project. A Case Study in Client-side Design Patterns Using Asynchronous Requests. *Literary and Linguistic Computing* 22/4 (2007) 395-403.
- CASSUTO, P: Qerê-Ketib et listes massorétiques dans le manuscrit B19a. *Judentum und Umwelt* 26. Frankfurt / M 1989: Peter Lang // I.2.1
- CIULA, A; LOPEZ, T: Reflecting on a dual publication: Henry III Fine Rolls print and web. *Literary & Linguistic Computing* 24/2 (2009) 129-141.
- COARTNEY, J S; WIESNER, S L: Performance as digital text: Capturing signals and secret messages in a media-rich experience. *Literary & Linguistic Computing* 24/2 (2009) 153-160.
- CRAIG, H . *Literary and Linguistic Computing* 20 / 4 (ed./eds.): [Themenheft zu Textkritik] (2005)
- DAVIES, M: The Corpus of Contemporary American English as the first reliable monito corpus of English. *Literary and Linguistic Computing* 25 (2010) 447-464.
- DOUBLES, M C: Toward the Publication of the Extant Text of the Palestinian Targum(s). *VT* 15 (1965) 16-26.
- FREEDMAN, D.N; SANDERS, J.A; BECK, A (ed./eds.): The Leningrad Codex 1995: Wm.B.Eerdmans Publishing Co.
- GALEY, A; RUECKER, S: How a prototype argues. *Literary and Linguistic Computing* 25 (2010) 405-424.
- GALEY, A; RUECKER, S: How a prototype argues. *Literary and Linguistic Computing* 25 (2010) 405-424.
- GIRON BLANC, L F: Pentateuco Hebraeo - Samaritano. Genesis. Edición crítica sobre la base de Manuscritos inéditos. Madrid 1975

- GOETHE, J W: Faust in: SCHÖNE, A (ed./eds.): J.W.Goethe, Sämtliche Werke. Briefe, Tagebücher und Gespräche. I. Abt: Sämtliche Werke Band 7 / 1 (Texte); Band 7 / 2 (Kommentare). Frankfurt / M 1994: Deutscher Klassiker Verlag
- GRÄTZ, K: Der Weg zum Lesetext. Editions kritik und Neuedition von Friedrich Hölderlin "Der Tod des Empedokles". Untersuchungen zur deutschen Literaturgeschichte 79. Tübingen 1995: Niemeyer // I.2.1
- HENKES, C; HETTICHE, W; RADECKE, G; SENNE, E (ed./eds.): Schrift - Text - Edition (FS H. W. Gabler). Beiheft zu editio 19. Tübingen 2003: Niemeyer
- HILTSCHER, M: Shakespeares Text in Deutschland. Textkritik und Kanonfrage von den Anfängen bis zur Mitte des neunzehnten Jahrhunderts. Münsteraner Monographien zur englischen Literatur 12. Frankfurt 1993: Peter Lang // I.2.1
- ISENBERG, S R: On the Jewish-Palestinian Origins of the Peshitta to the Pentateuch. JBL 90 (1971) 69-81.
- KRAFT, R: Computer assisted identification and reconstruction of fragmentary manuscripts (Papyrus, Leather, Paper): Chester Beatty Greek Papyrus 5 (Genesis) = Rahlfs 962. Actes du Second Colloque International Bible et Informatique: Méthodes, Outils, Résultats. Genève 1989 319-321. // I.2.0
- KRAFT, H: Editionsphilologie. Darmstadt 1990 // I.2.1
- KRETZSCHMAR, W A jr.; POTTER, W G: Library collaboration with large digital humanities projects. Literary and Linguistic Computing 25 (2010) 439-445.
- MARSDEN, A; MACKENZIE, A u.a.: Tools for Searching, Annotation and Analysis of Speech, Music, Film and Video - A Survey. Literary and Linguistic Computing 22/4 (2007) 469-488.
- MARTENS, G: Methoden der Textkritik und Textedition in: BRINKER, K; ANTOS, G; HEINEMANN, W; SAGER, S F (ed./eds.): Text- und Gesprächslinguistik, Linguistics of Text and Conversation. Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research. HSK 16 / 1. Berlin 2000 201-208. // I.2.4
- MCCARTHY, C: The Targum Sopherim and Other Theological Corrections in the Masoretic Text of the Old Testament. Göttingen 1981 // I.2.1
- MCCARTER, P K: Textual criticism. Recovering the text of the Hebrew Bible. Philadelphia 1984
- OTT, W: Edition und Datenverarbeitung in: KRAFT, H (ed./eds.): Editionsphilosophie. Darmstadt 1990 59-70. // I.2.1
- PLETT, H: Textwissenschaft und Textanalyse. Semiotik, Linguistik, Rhetorik. UTB 328. Heidelberg 1975
- PRICE, J D: The syntax of masoretic accents in the Hebrew Bible. Studies in the Bible and early Christianity 27. Melbourn 1990 // I.2.1
- RABE, N: Zur synchron definierten alttestamentlichen Textkritik. BN 52 (1990) 64-97.

- RABE, N: Questions of Textual Criticism: Diachronic vs. Synchronic Approach. Actes du Troisième Colloque International "Bible et Informatique: Interprétation, Herméneutique, Compétence informatique", Tübingen, 26-30 August 1991. Paris-Genève 1992 283-306. // I.2.0
- RABE, N: The semantic adequacy of the LXX translation of Hebrew units in Exodus 19. Textus (1998)
- RAHLFS, A: Septuaginta. I Leges et historiae. II Libri poetici et prophetici. Stuttgart 1935 // I.2.1
- ROUED-CUNLIFFE, H: Towards a decision support system for reading ancient documents. Literary and Linguistic Computing 25 (2010) 365-379.
- SCHWEIZER, H: Wahlen zur Josefsgeschichte. bibel heute 133 (1998) 132-133.
- SPENCER, M; HOWE, C J: Estimating Distances between Manuscripts Based on Copying Errors. Literary and Linguistic Computing 16 / 4 (2001) 467-484.
- STIPP, H-J: Das Verhältnis von Textkritik und Literarkritik in neueren alttestamentlichen Veröffentlichungen. Biblische Zeitschrift 34 (1990) 16-37.
- SUTHERLAND, K: Material text, immaterial text, and the electronic environment. Literary and Linguistic Computing 24 (2009) 99-112.
- TEEHAN, A; KEATING, J G: Appropriate Use Case modeling for humanities documents. Literary and Linguistic Computing 25 (2010) 381-391.
- TEEHAN, A; KEATING, J G: Appropriate Use Case modeling for humanities documents. Literary and Linguistic Computing 25 (2010) 381-391.
- TERRAS, M, P; ROBERTSON: Downs and Acrosses. Markup on a Stroke Level. Literary and Linguistic Computing 3 (2004) 397-414.
- TERRAS, M: Digital curiosities: resource creation via amateur digitization. Literary and Linguistic Computing 25 (2010) 425-438.
- TOV, E: Criteria for evaluating textual Readings. The limitations of textual rules. HTR 75 (1982) 429-448.
- TOV, E: Textual Criticism of the Old Testament. Assen / Maastricht 1992: van Gorcum // I.2.1
- ULMSCHNEIDER, H: Götz von Berlichingen. Mein Fehd und Handlungen. Forschungen aus Württembergisch Franken 17. Sigmaringen 1981: Jan Thorbecke
- WHEELLES, D: Testing NINES. Literary and Linguistic Computing 25 (2010) 393-403.
- WONNEBERGER, R: Leitfaden zur Biblia Hebraica Stuttgartensia. Göttingen 1984 // I.2.1

3.11 Einzelnes Manuskript / einzelner Druck

Literatur

PILZ, T; ERNST-GERLACH, A u.a.: The Identification of Spelling Variants in English and German Historical Texts: Manual or Automatic?. Literary and Linguistic Computing 23/1 (2008) 65-72.

zu Ziff. 3.111:

Textkritik: Autorisation↔Authentizität

zu: vgt4.02

aus: Siegfried Scheibe, Zur Abgrenzung der Begriffe Autorisation und Authentizität in: BEIN, T; NUTT-KOFOTH, R; PLACHTA, B (eds.), Autor - Autorisation - Authentizität. Beihefte zu editio 21. Tübingen 2004. S. 31-38.

(31) Das erste Zitat stammt aus einem Brief Christoph Martin Wielands an den Schweizer Verleger, Schriftsteller und Maler Salomon Geßner vom 19. August 1767. In ihm heißt es: "Harscher gährt noch, wie Sie sagen - er muß noch jung seyn - er wird mit der Zeit sehr gut werden, insofern er nicht noch vorher ins Tollhaus komt, eh er ausgeraßt hat."

Das zweite Zitat ist einem Brief Wielands an den Verleger Georg Joachim Göschen vom 6. Juni 1805 entnommen. Es lautet: "Sie, mein Bester, haben mir durch das Geschenk von Romanus Neffen einen Genuß verschafft, wie mir selten zu Theil wird, und der mir in diesem Geist und Körper abspannenden Frühjahr um so wohlthätiger war."

(32) Durch die Erwähnung der Quelle, aus der Wielands Kenntnis dieser Person stammt, wird der Leser nun schon auf die richtige Spur geführt; sie wird bestätigt durch einen weiteren Brief, um den 23. Oktober 1767 geschrieben, an Salomon Geßner, in dem es heißt. "Dieser Tagemeldet man mir, daß der seltsame Mensch, der die Fragmente über die Neuere Literatur geschrieben hat, Härder heisse, und, mit Verlaub, Conrector der Schule zu Riga sey: Der Himmel sey seinen Schulknaben gnädig!" Gemeint ist also in diesen Fällen Johann Gottfried Herder, dessen richtigen Namen Wieland zunächst nicht kennt oder den er nur gehört, nicht gelesen hatte und deshalb nicht richtig wiedergab. Alle diese Briefe sind in Handschriften Wielands überliefert, ihre Texte sind also in der überlieferten Form autorisiert. Wie hat sich nun der Editor in diesem Fall zu verhalten? Er könnte zwar in den Text eingreifen und an allen Stellen statt Harscher den richtigen Namen Herder einsetzen, er würde damit aber den autorisierten Text verfälschen und in diesem Fall außerdem ein Wissen des Autors vortäuschen, das zu der Zeit, als diese Briefe geschrieben wurden, noch nicht gegeben war. Der autorisierte Text dieser Briefe, also auch die falsche Namensangabe (sowohl Harscher wie Härder), ist in der Ausgabe in der überlieferten Form wiederzugeben, die Richtigstellung hat in den Erläuterungen zu den betreffenden Briefen zu erfolgen, im Register sind außerdem entsprechende Hinweise zu geben, und gegebenenfalls könnte durch ein Zeichen am Rand der betreffenden Zeile der Leser darauf hingewiesen werden, daß er unbedingt im Apparat nachzuschlagen hat, um auf die richtige Spur zu kommen.

Das zweite Zitat, in dem von "Romanus Neffen" die Rede ist, stellt sich dagegen anders dar. Der originale Brief Wielands ist nicht überliefert. Er war 1880 bei der Firma Lepke zum Verkauf angeboten worden; sein Verbleib ist seitdem unbekannt. Offenbar der erste Käufer, Moritz Herz, veröffentlichte den Text des Briefes 1885 in der Zeitschrift *Die Gegenwart*, und auf dessen Textwiedergabe sind wir angewiesen, um diesen Brief in einer Edition abzurufen. Herz fügte der zitierten Stelle folgende Anmerkung hinzu: "'Der Neffe' ein Lustspiel von Romanus". Jedoch ist weder ein Komödienautor Romanus biographisch noch ein von diesem stammendes Stück "Der Neffe" bibliographisch nachzuweisen. ...

Zur Ostermesse 1805 erschien dagegen von Goethe (33) bei Cotta in Tübingen *Winkelmann und sein Jahrhundert. In Briefen und Aufsätzen herausgegeben von Goethe* und bei Göschen in Leipzig *Rameau's Nef-fe. Ein Dialog von Diderot. Aus dem Manuscript übersetzt und mit Anmerkungen begleitet von Goethe*, und, nun ist es klar, letzteres Buch hatte Göschen an Wieland geschickt. Herz hatte sich bei der Übertragung dieses Briefes verlesen und seine Verlesung noch durch eine unsinnige Anmerkung verschlimmert. Was hat der Editor in diesem Fall zu tun? Der ihm vorliegende Text ist unautorisiert (denn der originale Text des Briefes liegt ihm nicht vor), daß der mit Handschriften offenbar unerfahrene Herausgeber "Romanus" für "Rameau" verlesen hat, ist bei der Ähnlichkeit der Wörter wahrscheinlich, und deshalb hat der Editor das Recht und hier auch die Pflicht, den eindeutig falschen Text zu beseitigen und anstelle von "Romanus Neffen" den richtigen Titel, "Rameaus Neffen", einzusetzen.

zu Ziff. 3.112:

Textkritik: Strofenfolge

zu: vgt4.03

aus: Karl Konrad Polheim, Die scheinbare Autorisation oder Der Schutz des Autors vor sich selbst in: BEIN, T; NUTT-KOFOTH, R; PLACHTA, B (eds.), Autor - Autorisation - Authentizität. Beihefte zu editio 21. Tübingen 2004. S. 67-72.

(70)

Wem Gott will rechte Gunst erweisen,
Den schickt er in die weite Welt,
Dem will er seine Wunder weisen
In Fels und Wald und Strom und Feld.

Die Bächlein von den Bergen springen,
Die Lerchen schwirren hoch vor Lust,
Was sollt' ich nicht mit ihnen singen
Aus voller Kehl' und frischer Brust?

Die Trägen, die zu Hause liegen,
Erquicket nicht das Morgenroth,
Sie wissen nur vom Kinderwiegen,
Von Sorgen Last und Noth um Brodt.

Den lieben Gott laß ich nur walten;
Der Lerchen, Bächlein, Wald und Feld
Und Erd' und Himmel will erhalten,
Hat auch mein' Sach' auf's best' bestellt!

Das vierstrophige Gedicht ist in der Handschrift zweispaltig geschrieben. Die Strophen "Wem Gott will" und "Die Trägen" stehen in der ersten Spalte untereinander, ebenso die Strophen "Die Bächlein" und "Den lieben Gott laß ich" in der zweiten Spalte. Umgekehrt ausgedrückt stehen nebeneinander die Strophen "Wem Gott will" und "Die Bächlein" und darunter nebeneinander "Die Trägen" und "Den lieben Gott laß ich". Der Setzer hat - wie man das ja zunächst tun würde - spaltenweise gelesen, also I. Strophe "Wem Gott will", II. Strophe "Die Trägen", III. Strophe "Die Bächlein", IV. Strophe "Den lieben Gott laß ich". In dieser Gestalt ist das Gedicht erstmals, 1823, und in der letzten autorisierten Buchfassung 1826 veröffentlicht worden, und ebenso wurde es fortan gedruckt, gelesen und gedeutet.

In der Handschrift folgt wenige Seiten später das zweite Gedicht "Wohin ich geh' und schaue", ebenfalls vierstrophig und wiederum zweispaltig geschrieben. Aber hier sind die Strophen mit Ziffern versehen, und diese laufen waagrecht: die II. Strophe steht neben der I., darunter die III. und IV. Strophe ebenfalls nebeneinander. Der Befund ist eindeutig und legt den Schluß nahe, diese Anordnung auch im ersten Gedicht zu sehen. Das kodikologische Kriterium trägt uns also auf, das Gedicht "Wem Gott will rechte Gunst erweisen" waagrecht zu lesen, und das heißt entgegen der (71) bisherigen Überlieferung die II. und III. Strophe zu vertauschen. Diese Entscheidung ist vor kurzem bestätigt worden, nämlich durch die große, erst jetzt wieder aufgetauchte *Marmorbild*-Handschrift, die zwar etwas älter als die *Taugenichts*-Handschrift ist, jedoch in dieselbe Zeitspanne und Schaffensperiode Eichendorffs gehört: dort sind alle mehrstrophigen Gedichte - bis auf ein Sonett tatsächliche(!) alle - zweispaltig waagrecht angeordnet.

zu Ziff. 3.113:

"King Lear" - textlich 'Schwere Geburt' '

zu:

aus: Alexandra BRAUN-RAU, William Shakespeares "King Lear" in seinen Fassungen. Ein elektronisch-dialogisches Editionsmodell. Tübingen 2004.

vgl: <http://www.textkritik.uni-muenchen.de/abraun-rau/Lear/>

vgl: http://www.niemeyer.de/links/link_material.html/

(8) Nachdem im Dezember 1607 der Auftraggeber des Drucks, Nathaniel Butter, Druckvorlage und Papier geliefert hatte, wurde am 8. Dezember 1607 in der Londoner Druckerei von Nicholas Okes mit dem Druck der ersten Quartoausgabe begonnen. Unterbrochen durch eine einwöchige Weihnachtspause wurde Q1 binnen vier Wochen im Januar 1608 fertiggestellt. Okes, der den Betrieb erst 1607 als offizieller Juniorpartner von George Snowdon übernommen hatte, besaß selbst keine Erfahrung im Druck von Dramen. Wie Blayney an einigen Auffälligkeiten der Herstellung nachweisen konnte, war *King Lear* der erste, nach Übernahme der Druckerei hergestellte Dramentext im Quartformat. (9) Der Quartodruck von *King Lear* wurde in "Guyot Pica" von zwei Setzern (B und C) gesetzt, denen jeweils unterschiedliche Textabschnitte zugeschrieben werden können. So scheint zunächst nur ein Setzer (B) für den Satz der Bögen A-G verantwortlich gewesen zu sein. Bei Bogen H kam dann ein zweiter Setzer (C) mit eigenem Setzkasten hinzu, der im Verbund mit B die restlichen Bögen bearbeitete. Der hohen Quote an Fehlleistungen beim Satz nach zu urteilen, wurde die erste Quartoausgabe direkt von Shakespeares Entwurfshandschrift gesetzt. Obwohl die handschriftliche Vorlage nur schwer lesbar gewesen zu sein scheint, hielten wohl weder Okes noch Butter aus Kostengründen die Anfertigung einer Reinschrift für notwendig. Die Unerfahrenheit Okes beim Druck von Dramen und Schwierigkeiten der Setzer mit ihrer Druckvorlage mögen der Grund dafür gewesen sein, daß spätestens ab Bogen C nicht wie üblich nach Druckformen, sondern nach Textverlauf gesetzt wurde. Die Abkehr vom gängigen Verfahren bewirkte jedoch, daß Okes limitierter Vorrat an Drucktypen beim Satz des Dramas nicht ausreichte. Beide Aspekte, die unzureichende Menge an Drucktypen als auch die unleserliche Druckvorlage, mögen für einen Großteil der in der ersten Quartoausgabe auftretenden Druckfehler verantwortlich sein.

(11) Inwieweit die Preßkorrekturen der ersten Quartoausgabe von *King Lear* Manuskriptlesungen oder Konjekturen des Korrektors repräsentieren, ist nicht mehr zu klären. Bis heute liegen keine gesicherten Erkenntnisse über die Identität des Korrektors vor. Es dürfte jedoch auszuschließen sein, daß Shakespeare selbst Korrektur las. Er scheint nie am Druck seiner Stücke beteiligt gewesen zu sein. Im allgemeinen beschäftigten namhafte Druckereien einen *Proofreader*, der auf die Anforderungen des jeweiligen Druckprojekts spezialisiert war. Oft übernahm auch der Inhaber der Druckerei die Aufgabe des Korrektors. Da die flüchtige und ungenaue Korrektur der ersten Quartoausgabe nicht auf die Anstellung eines professionellen Korrekturlesers schließen läßt, ist bei *King Lear* eher damit zu rechnen, daß Nicholas Okes selbst für die Korrektur zuständig war.

zu Ziff. 3.114:

Textkritischer Alltag

zu:

aus: R. EGGER-WENZEL, Ein neues Sira-Fragment des MS C: Biblische Notizen 138 (2008) 107-114.

(107) Es kommt einer Sensation gleich, dass nach 26 Jahren, als damals Teile einer neuen Handschrift F aufgefunden wurden, nun ein kleines Stückchen mehr des hebräischen Sira das Licht der Welt wieder erblickt hat. Shulamit Elizur hat in ihrem Beitrag. "A New Hebrew Fragment of Ben Sira (Ecclesiasticus)" dankenswerter Weise die Photographie des Manuskriptblattes samt zeilenweiser Abschrift erstmals veröffentlicht. Ein Vergleich mit den bereits publizierten Photographien von Teilen des MS C bei Gaster, Schirmann und Scheiber zeigt, dass die Zuweisung zu MS C aufgrund des Formats des aufgefundenen Blattes und der Eigenart der Schreibweise mancher Buchstaben durchaus berechtigt ist. Zudem hatte Beentjes in seinem Artikel von 1988 bereits angemerkt, dass zwischen 36,24a sowie 6,18b ein Blatt fehlen müsse, das jetzt exakt diese Lücke schließt.

Fiktiv überhöht in UMBERTO ECO "Der Name der Rose" (Beginn des Vorworts). Beschrieben in: H. SCHWEIZER, Fantastische 'Opferung Isaaks'. Textanalyse in Theorie und Praxis. Lengerich 2006. S. 312f.

3.12 Kritische Edition (vs. Faksimile)**Literatur**

TERRAS, M; ROBERTSON, P: Downs and Acrosses. Textual Markup on a Stroke Level. Literary and Linguistic Computing 19 / 3 (2004) 397-414.

zu Ziff. 3.121:

Textkritik

GRASS, G, Danziger Trilogie. Die Blechtrommel. Katz und Maus. Hundejahre. Darmstadt 1980. [Die Blechtrommel]

zu:

(249) "Matzerath entschloß sich auf Gretchen Schefflers Zureden hin, meine Geliebte zu heiraten. Wenn ich also ihn, meinen mutmaßlichen Vater, als Vater bezeichne, muß ich feststellen: mein Vater heiratete meine zukünftige Frau, nannte später meinen Sohn Kurt seinen Sohn Kurt, verlangte also von mir, daß ich in seinem Enkelkind meinen Halbbruder anerkannte und meine geliebte, nach Vanille duftende Maria als Stiefmutter in seinem nach Fischlaich stinkenden Bett duldete."

(315) "Da trat Oskar näher. Da zog es ihn heran. Da wollte er auf den Teppich, wollt nicht mehr auf den Fliesen stehen. Eine Altarstufe gab ihn an die nächste weiter. So stieg ich hinauf und hätte ihn lieber hinabsteigen sehen. »Jesus«, kratzte ich einen Rest Stimme zusammen, »so haben wir nicht gewettet. Sofort gibst du mir meine Trommel wieder. Du hast dein Kreuz, das sollte dir reichen!« Ohne abrupt abzubrechen, beendete er die Trommelei, kreuzte die Stöcke übertrieben sorgfältig auf dem Blech und reichte mir ohne Widerrede, was Oskar im leichtsinnig gepumpt hatte. Schon wollte ich ohne Dank und hastig wie zehn Teufel die Stufen runter und raus aus dem Katholizismus, da berührte eine angenehme, wenn auch befehlerische Stimme meine Schulter: »Liebst du mich, Oskar?« Ohne mich zu drehen, antwortete ich: »Nicht daß ich wüßte.«"

(372) "Immer wenn die Jugendlichen den Waggon des Herrn Matzerath besetzten, erhoben sich die vier Nonnen und hielten ihre an den Kutten hängenden Kreuze hoch. Die vier Kurzfixe beeindruckten die jungen Burschen sehr."

(402) "Die aber schienen mir schön und geschickt gewesen zu sein, wie auch der Kopf der Frau, der schon in der Zinkkiste war, eine gewisse Ebenmäßigkeit durch den Nachkriegswinter siebenundvierzig-achtundvierzig, der ja bekanntlich schlimm war, hinübergerettet hatte, so daß abermals von Schönheit, wem auch verfallener, die Rede sein konnte."

(482) "Ich konnte nur zustimmend nicken, denn noch am Vortage hatte ich bei Ebbe zwischen Muscheln und leeren Krabbenschalen den deutlich sprechenden Knopf einer kanadischen Uniform gefunden."

(513) "Vittlar liebte es immer schon, mich zu verwirren. Ich soll mein Bett verlassen, Jünger sammeln, nur weil ich dreißig Jahre zähle."

aus: Johann Wolfgang Goethe, Die Wahlverwandtschaften. Ein Roman. Insel taschenbuch 2125. ¹1997.

(185) "... ihre alte Tätigkeit regt sich wieder; sie erblickt, wo sie auch hinsieht, im vergangenen Jahre vieles getan und empfindet Freude am Getanen."

aus: Niklas Luhmann, Die Realität der Massenmedien. Opladen ²1996: Westdeutscher Verlag.

(185) "Das Primärziel müßte sein, die Klienten zu lehren, die bei allen Unterscheidungen inhärente Paradoxie selbst zu sehen und zugleich zu sehen, daß Beobachtungen nur möglich sein, wenn man

die Paradoxien wieder in die Form einer Unterscheidung bringt, die im Moment zu überzeugen scheint."

aus: John UPDIKE, Terrorist, 2006: Rowohlt:

(257) "'Freunden helfen einander', sagt Charlie nachdrücklich."

aus: Günter GRASS, Die Rättin. Darmstadt 1986.

(327) denn der Raum Matarnia, Firoga, Zokovo, Kartury, im (!) dem - wo früher Bysewo lag - seit Jahren Flugpisten betoniert liegen, ist von zwei Vernichtungssystemen, der Hitze- und Druckwelle und dem radioaktiven Fallout sowie von beschleunigten Neutronen- und Gammastrahlen betroffen.

zu Ziff. 3.122:

Textkritik

zu:

aus: MARTENS, G, Methoden der Textkritik und Textedition: BRINKER, K; ANTOS, G; HEINEMANN, W; SAGER SF (eds.), Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischen Forschung. HSK 16.1. Berlin 2000: de Gruyter. S. 201-208 // I.2.4

(201) Erst die im 20. Jh. einsetzende Neubesinnung der Aufgaben und Ziele einer Edition brachte die Notwendigkeit mit sich, nun auch den editorischen Begriff *Text* im einzelnen zu klären. Die bis heute nicht abgeschlossene Diskussion ist allerdings überaus kontrovers. Danach zeichnen sich - entsprechend dem unterschiedlichen Selbstverständnis der Editionsphilologen - verschiedene Richtungen ab, den Begriff definitiv einzugrenzen. So stehen sich zum einen Auffassungen, die den Text (1) als eine intentionale Sprachgestalt sehen, die nicht überliefert ist oder in dieser Form niemals existiert hat und die daher erst vom Editor *hergestellt* ("konstituiert") werden muß, anderen Verständnisweisen gegenüber, die Text (2) allein als bereits bestehendes Faktum zulassen, das vom Textherausgeber *festzustellen* und zu dokumentieren ist. Verschiedenartige Textbegriffe ergeben sich vor allem durch die Frage, ob zu einem Werk überlieferte Varianten (3) als solche zum Text des Werkes gehören, in ihrer Gesamtheit also den Text des Werkes bilden, oder ob sie demgegenüber (4) als Abweichungen zum edierten Text jeweils einen eigenen gesonderten Text begründen...

(202) Werke der Antike und des Mittelalters sind nur in wenigen Ausnahmefällen im vom Autor stammenden Original oder in einer von ihm gebilligten Fassung überliefert. Zumeist sind es sehr viel später entstandene Abschriften von Abschriften, die unsere Kenntnis der antiken und mittelalterlichen Literatur begründen. So sahen es Editoren seit jeher als ihre Aufgabe an, neben der Sicherung der Überlieferung das Überlieferte selbst kritisch zu überprüfen und das "Echte" von "Unechtem" zu scheiden. Die Rückgewinnung der ursprünglichen, authentischen Textgestalt, Entstellungen des originalen Wortlauts zu erkennen, zu beseitigen oder doch zumindest zu bezeichnen, war für lange Zeit das vornehmste Ziel editorischer Tätigkeit.

zu Ziff. 3.123:

Editionsphilologie

SCHICKEL, J, Sappho. Strophen und Verse. it 309. Frankfurt/M 1978: Insel
Stichwort: Textkritik, freie Konjektur, Pragmatik in Konstituierung, Text(fragment) als "historisches Faktum" (17f)

XII

*Scheinen will mir, er komme gleich den Göttern,
 jener Mann, der dir gegenüber niedersitzen darf und nahe den süßen Stimmenzauber vernehmen*

und des Lachens lockenden Reiz. Das läßt mein Herz im Innern mutlos zusammenkauern. Blick ich dich ganz flüchtig nur an, die Stimme stirbt, eh sie laut ward,

ja, die Zunge liegt wie gelähmt, auf einmal läuft mir Fieber unter der Haut entlang, und meine Augen weigern die Sicht, es überauscht meine Ohren,

mir bricht Schweiß aus, rinnt mir herab, es beben alle Glieder, fahler als trockne Gräser bin ich, einer Toten beinahe gleicht mein Aussehn . . .

Aber alles trägt sich noch, da . . .

(86f) *Aber alles trägt sich noch, da . . .*

"Bruno Snell nimmt an, im uns verlorenen Schluß des Gedichtes habe »in irgendeiner Form der Gedanke des »Ausgleichs« gestanden... , entweder: »da du das Glück der Hochzeit besitzt«, oder: »da Leid und Freude wechseln««. Denys Page, der 1955 eine Einführung in das Studium der alten Lesbischen Poesie veröffentlicht und im gleichen Jahr (zusammen mit Edgar Lobel) die »Poetarum Lesbiorum Fragmenta« griechisch ediert hat, ist vorsichtiger. Er meint, es sei einmal »fashionable« gewesen, diese Worte - *ἀλλὰ παντολματον, επει* - so zu verdrehen, daß sie »den Anfang einer Kritik an den vorangehenden Versen abgeben können«. Aber »seit der Rest des Satzes, und damit des Poems, fehlt, gibt es kein Mittel zu bestimmen, ob er »alles ist ertragen worden« oder »alles muß ertragen werden bedeutet«. Ist schon das Vorhandene solcher Zurückhaltung wert, desto mehr das Verlorene. Gar nicht sicher scheint, daß Sappho ihre Gefühle, der andern sich freuend, ausgeglichen haben soll; sie hätte auch, an sich selber leidend, schließen können: Ich trag es, da ich die Situation zu erkennen, darüber zu sprechen vermag. Vermocht hat sie es sicherlich; doch ist zu fragen: um welchen Preis, für welchen Gewinn. Was hat sie bekommen, was hat sie bezahlt? Die erste Antwort, Sapphos lyrische Einkunft (wenn sie bloß lyrisch war!) betreffend, gibt Bruno Snell, der frühgriechisches Dichten geradezu mit dem »Erwachen der Persönlichkeit« identifiziert:

»Im vollen Sinn ihr Eigenes, Persönliches ist ... das Gefühl der Hilflosigkeit. Erst die gehemmte Liebe, die nicht zu ihrer Erfüllung kommt, bemächtigt sich mit besonderer Kraft des Bewußtseins. Wo der Strahl eines starken, lebendigen Strebens sich an einem Widerstand totläuft, leuchtet er auf. In diesem am Widerstand bewußt gewordenen Gefühl, das etwas durchaus Persönliches ist, verlieren die sonst gepriesenen Dinge ihren Wert: Hier bricht der

Gegensatz auf zwischen Schein und Sein, zwischen dem, was andere hochhalten, und dem, was man selbst als wesentlich begreift.«"

zu Ziff. 3.124:

Computerunterstützte Editionsphilologie

zu: vgt4.33

ROBINSON, P M W; GABLER, H W (eds.):

Making Texts for the Next Century.

A Special Issue of *Literary and Linguistic Computing*

Vol. 15 Number 1 April 2000

DEEGAN, M (ed.):

Electronic Scholarly Editing - Some Northern European Approaches

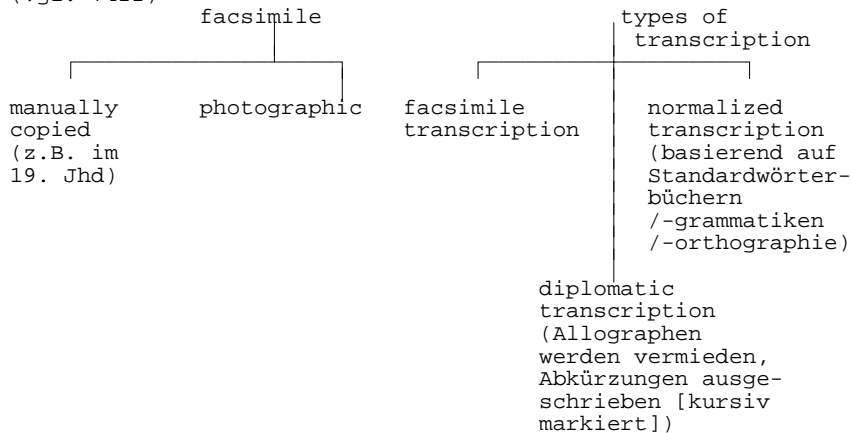
A Special Issue of *Literary and Linguistic Computing*

Vol. 19 Number 1 April 2004

Textkritische Aufbereitung alter Handschriften

aus: O. E. HAUGEN, Parallel Views: Multi-level Encoding of Medieval Nordic Primary Sources: Literary and Linguistic Computing 19/1 (2004) 73 - 91.

(vgl. 74ff)



(80)

Criterion	Facsimile	Diplomatic	Normalized
Word division	faithfully copied	faithfully copied	always regularized
Line division	faithfully copied	usually copied	always left out
Page division	faithfully copied	usually copied	always left out
Abbreviations	faithfully copied, no expansions	expanded and usually shown in italics	always expanded silently
Graphemic inventory	nearly complete	usually limited to phonemic oppositions	limited to the standard orthography
Punctuation	faithfully copied	usually copied	always regularized
Capitalization (persons and place names)	never regularized	usually not regularized	always regularized
Emendations by the transcriber /editor	never in the transcription	occasionally, always explicitly marked	often, not always explicitly marked

zu Ziff. 3.1241:

Älteste Bibel-Handschrift der Welt ist online

zu:

aus: SPIEGEL ONLINE 6. Juli 2009

CODEX SINAITICUS

1600 Jahre nach ihrer Entstehung kann man nun auch bequem am Bildschirm in ihr lesen: Die älteste erhaltene Bibel-Handschrift der Welt ist im Internet. Doch noch sind nicht alle der insgesamt 1450 Seiten online.

London - Der Codex Sinaiticus ist 1600 Jahre alt - und damit die älteste Bibel-Handschrift der Welt. Teile des Textes wurden nun ins Internet gestellt, wie die British Library in London und die Universitätsbibliothek Leipzig mitteilten. Forschungsteams aus Deutschland, Großbritannien, Russland und Ägypten hatten das alte griechische Manuskript aufgearbeitet. Nun sind rund 800 der 1460 Seiten auf

www.codex-sinaiticus.net

einsehbar.

Die Originalblätter befinden sich zu Teilen in der Universitätsbibliothek Leipzig, der British Library, dem St. Katharinenkloster auf Sinai in Ägypten und in der russischen Nationalbibliothek in St. Petersburg. Das Projekt hatte sich über Jahre hingestreckt, vergangenen Sommer wurde der erste Online-Teil vorgestellt.

Das Manuskript gebe einen Einblick in die frühe Entwicklung des Christentums und zeige, wie der Bibeltext von Generation zu Generation weitergegeben wurde, erklärte Scot McKendrick von der British Library. Der Codex Sinaiticus sei einer der "wertvollsten Schätze der Welt". "Bibel-Forscher können nun online weltweit zusammenarbeiten. Das war vor einigen Jahren noch gar nicht möglich."

Die digitale Fassung geht über die reine Abbildung des Textes, der ausschließlich in Großbuchstaben, ohne Leerzeichen und Bilder gesetzt ist, hinaus. Internet-Nutzer und Forscher können die Seiten wie einen digitalen Film verschieben, Textstellen vergrößern und sich die griechische Transkription anzeigen lassen. Von Teilen gibt es auch eine deutsche und englische Übersetzung. Die vier-spaltige Bibelabschrift in altgriechischer Sprache gilt als eines der größten Pergamentbücher der Antike. Sie war Jahrhunderte lang im Katharinenkloster aufbewahrt worden. 1844 hatte der Leipziger Theologe Konstantin von Tischendorf die 33,5 Zentimeter breiten und 37,5 Zentimeter hohen Pergamentblätter in einem Papierkorb des Klosters entdeckt. Die Mönche erlaubten ihm, 43 von insgesamt 129 gefundenen Blättern mit nach Leipzig zu nehmen.

© SPIEGEL ONLINE 2009

zu Ziff. 3.125:

Textprobleme und Erwartungen des Lesers

zu:

aus: Umberto ECO, Das Foucaultsche Pendel. München ¹⁷2005.

(689) »Und die Popelicans?«

»In Provins hatte es Katharer gegeben, die dann ordnungsgemäß verbrannt worden sind, und der Großinquisitor war ein reuiger Ketzer und wurde Robert le Bougre genannt. Also war's nichts Besonderes, wenn es da auch eine Straße oder eine Gegend gab, die weiterhin als Sitz der Ketzer bezeichnet wurde, auch als es die Katharer nicht mehr gab.« »Auch noch 1344...«

»Wer sagt dir denn, daß dieses Dokument von 1344 stammt? Dein Oberst hat hier > 36 Jahre nach dem Heuwagen< gelesen, und tatsächlich war damals ein *p*, das in einer bestimmten Weise gemacht war, mit einer Art Apostroph, die normale Abkürzung für *post*, aber ein anderes *p* ohne Apostroph bedeutete *pro*. Der Verfasser dieses Textes ist ein friedlicher Händler, der sich ein paar Notizen gemacht hat über seine Geschäfte in der Grange-aux-Dîmes, also in der Rue Saint-Jean, nicht in der Johannisnacht, und er hat sich einen Preis notiert: sechsunddreißig Sous oder Deniers oder was für Münzen die damals hatten, für eine oder für jede Fuhre Heu.« »Und die hundertzwanzig Jahre?«

»Wer spricht von Jahren? Ingolf hat etwas gesehen, das er als 120 a... abgeschrieben hat. Wer sagt dir, daß es ein a war? Ich habe nachgeschaut in einer Tabelle der damals üblichen Abkürzungen und hab gefunden, daß man für *denier* oder *denarius* seltsame Zeichen benutzte - eins, das wie ein Delta aussieht, und ein anderes wie ein Theta, eine Art links angeknabberter Kreis. Kritzel das eilig hin, als kleiner Händler, der sich Notizen macht, und schon kann es so ein Eiferer wie dein Oberst mit einem a verwechseln, weil er die Sache mit den 120 Jahren schon irgendwo gelesen hat, du weißt besser als ich, daß er sie in jeder beliebigen Rosenkreuzergeschichte lesen konnte, er wollte etwas finden, das so ähnlich klang wie *post* (690) *120 annos patebo!* Und was macht er dann? Er findet eine Reihe von *it* und liest sie als *iterum*. Aber *iterum* wurde mit *itm* abgekürzt, und *it* bedeutete *item* desgleichen, ebenso, ein Ausdruck, der speziell für repetitive Listen benutzt wurde. Unser Händler kalkuliert, was ihm bestimmte Aufträge einbringen werden, die er bekommen hat, und listet auf, was er wohin liefern muß. Er muß Rosen liefern, Kreuzritter-Rosen von Provins, das ist es, was r... s... *chevaliers de Pruins* heißt. Und da, wo der Oberst *vainjance* gelesen hat (weil ihm der Ritter Kadosch im Kopf herumspukte), muß man *jonchée* lesen, Blumenschmuck. Die Rosen wurden benutzt, um Kränze oder Teppiche daraus zu flechten, für die verschiedenen Feste. . . . Das ist eine Wäscheliste, ich wiederhole es dir.«

3.13 Vergleich verschiedener Fassungen → begründete Wahl

zu Ziff. 3.131:

AUS:
 "Ohren-vergnügendes und
 Gemüth-ergötzendes
 Tafel-Confect" 1733/46
 von Johann Valentin
 Rathgeber (1682-1750)
 in: Das europäische
 Kunstlied. Im Auftrag des
 Verbandes Bayerischer
 Schulmusikerzieher
 hg.v. M.
 JÜRGENSON-RUPPRECHT; K.
 SUTTNER. Mainz 1979. S.14f.

AUS:
 "Augsburger Tafelkonfekt"

 hg.v. W. PUDELKO.
 Kassel o.J.
 S.20f.

Textkritik, lectio difficilior, Editionsphilologie

Nase
 grobe kleine dicke dünne
 breite schmale lange kurtze
 alte junge andre mehr.
 1.Nun hört mich an ihr Herren
 ich will mein Maul aufsperrn
 in angelweit und breit
 da will ich produciren
 viel Nasen abcopiren
 für lange Zeit und Weil.
 Da seh ich ein kleine
 ein hübsche
 ein feine dort guckt eine für
 die ist zimlich dürr
 dort wieder ein stumpffe
 dort wieder ein krumbe
 es hab'n halt die Haas'n
 gar artliche Nas'n.
 Es gibt Nase Näsle
 Näslein
 muß doch ein jeder mit seiner
 zufrieden wohl seyn.
 2.Dort seh ich einen Lappen
 sein Nas gleicht eim
 Danzapffen
 lang dick groß starck darbey
 nicht weit darvon ist einer
 der komt heraus nicht feiner
 als wie ein Gänß-Arschlein.
 Auch gibts viel probirte
 proportionierte
 recht feine recht zart
 auf höflichste Art
 halt wie sie gerathen
 bald schön bald mit Schaden
 bald ita bald non
 bald Sonn und bald Monn.

Nase!
 Grobe, kleine, dicke, dünne,
 breite, schmale, lange, kurze,
 alte, junge, andre mehr!
 1.Nun hört mich an ihr Herren,
 tut euer Ohr aufsperrn
 recht angelweit und breit!
 Was will ich jetzo bringen?
 Viel Nasen euch besingen
 für lange Weil und Zeit.
 Da seh ich ein' kleine,
 ein hübsche
 ein feine, dort guckt eine für,
 die ist ziemlich dürr.
 Dort wieder ein stumpffe,
 dort wieder ein' krumme.
 Es habn halt die Hasn
 gar artliche Nasn!
 Es gibt: Nase, Näsle,
 Näslein,
 muß doch ein jeder mit seiner
 zufrieden wohl sein.
 2.Dort seh ich einen tappen,
 der hat ein
 Tannenzapfen,
 lang, dick groß, stark dabei.
 Nicht weit davon ist einer,
 der kommt heraus mit seiner
 als wie ein Äpfelein.
 Auch gibts viel gezierte,
 wohlproportionierte,
 recht feine, recht zart,
 auf höflichste Art.
 Halt wie sie geraten,
 bald schön bald mit Schaden
 bald grad und bald krumm,
 bald schlau und bald dumm.

Es gibt Nase Näsle ...

| Es gibt: Nase, Näsle, ...

Str.1

Ø

Str.1

2. Ich seh ein' lange, lange, bald wie ein Hopfenstange, nicht gar so lang ist sie. Dort steht ein ungeschickte, dort ein zusammengeflickte, wieviel sind Nasen hier! Die hat einen Griffel als wie ein Rheinschiffel, die hat einen Ast gar lustig verpaßt. Der hat einen Haken mit allerlei Zacken und bild sich doch ein, sein Nasen wär fein.

Str.2

Ø

Str.3

4. Ei, ei, seht an ihr Denker, dort seh ich was, zum Henker, ist aller Nasen Nest, als wie ein Faust gewaltig, verzieret mannigfaltig mit lauter Gartengmüs. Sie ist recht possierlich, schier übernatürlich, fein blau untermengt und gelblich gesprengt, fein knorrig und faltig und holprig und spaltig, ein Sorten ganz neu, ein Wunder dabei.

Ø

5. Schau, schau nur diese Zwiebel gleicht einem alten Giebel, verbogen schief und krumm. Die tut eim Rauchfang gleichen, das geben ihre Zeichen rund um die nas herum. Acht kein Fleisch, kein Suppen, wann sie nur zu schnupfen, grämt sie sich um nichts, wann voll nur die Büchs. Die schaut auf gen Himmel, die zeigt schon Schimmel, vom Bier und vom Wein so scheint es zu sein.

zu Ziff. 3.132:

Textkritik / GOETHE'S 'Faust'

zu:

aus: SCHÖNE, A (ed.) Johann Wolfgang Goethe. Faust. Texte. Frankfurt/M 1994. (= Bd.I) SCHÖNE, A Johann Wolfgang Goethe. Faust. Kommentare. Frankfurt/M 1994. (= Bd.II)

SCHÖNE, I,147 + II,321

Mein Busen drängt
sich nach ihm hin.

Ein glättend abschwächender Eingriff Goethescher Selbstzensur. In der *Frühen Fassung* hieß es, rücksichtslos direkt gegen den Verhaltenskodex des zeitgenössischen Publikums verstoßend: *Mein Schoos! Gott! drängt / Sich nach ihm hin* - ein tonbeugender Aufschrei hob dort das Wort *Gott* aus der metrischen Senkung, so dass dieses für Margarete ganz ungeheure Eingeständnis als einzig dreihebiger Vers das zweihebig gefaßte Regemaß ihrer Rede durchbrach.

zu Ziff. 3.133:

Editionsphilologie: Ziele/Kriterien?

zu:

aus: BRENNER, P J, Das Problem der Interpretation. Eine Einführung in die Grundlagen der Literaturwissenschaft. Konzepte der Sprach- und Literaturwissenschaft 58. Tübingen 1998: Niemeyer

(275) Diese Suche nach der philologischen Wahrheit erhält später eine andere Form die sich bis in die editionsphilologischen Diskussionen der Gegenwart erhalten hat. Je stärker sich die Philologie neben der Rekonstruktion mittelalterlicher Texte neuerer Literatur zuwandte, desto deutlicher trat der Autor als der Urheber des Textes in den Blick. Der 'Wille des Autors' wird zum Zauberwort, an dem sich die Editionsphilologie ausrichtet, die enge Bindung von Text und Autor erscheint unter dieser Prämisse konsequent: *"der Text ist die letztgültige sprachliche Gestalt, die dem Werk durch den Autor verliehen worden ist als Ergebnis des schöpferischen Prozesses; er drückt diejenige Realisierung der schöpferischen Intention aus, die unter den Bedingungen der Entstehung des Werkes und den schriftstellerischen Möglichkeiten des Autors erreicht werden konnte."* In der Mitte des 20. Jahrhunderts ließ sich auf der Grundlage derartiger Prämissen ganz unbefangenen Auffassung vertreten, daß mit mäßigem Aufwand ein "zuverlässiger Text" erreichbar sei: "Ein zuverlässiger Text, so läßt sich definieren, ist ein Text, der den Willen des Autors repräsentiert." Diese naive Vorstellung ist durch die Entwicklung der folgenden Jahrzehnte wieder in weite Ferne gerückt. Die Entfaltung der Editionsphilologie hat dazu geführt, daß sich die Hoffnung auf den 'gesicherten' Text als der Grundlage einer jeden Interpretation als trügerisch erwiesen hat. Die Editionstechnik hat sich mit zunehmendem Kenntnisstand und Aufwand bei der Herstellung von Texten damit abfinden müssen, daß jeder Text nicht nur das Werk seines Autors ist. An seiner Entstehung haben (276) viele Faktoren mitgewirkt, identifizierbare und anonyme; dazu gehören der Setzer und Verleger ebenso wie die Zensur und die Zufälle der Überlieferungsgeschichte, die dem Text jene Gestalt gegeben haben, in der er dem Interpreten schließlich gegenübertritt. Die Editionsphilologie hat seit ihren Anfängen versucht, diese autorfremden Einflüsse zu identifizieren und damit zu neutralisieren, aber sie hat sich seit je sehr schwer damit getan...

Denn die Literaturgeschichte zeigt, daß die Zurückführung der Überlieferung auf eine einmalige und eindeutige authentische Textgestalt, wie sie dem Willen des Autors entspricht, oft schon an der einfachen Tatsache scheitert, daß ein Autor mehrere Fassungen des gleichen Textes hinterlassen hat. Nicht die Überlieferungsgeschichte, sondern der Autor selbst wird dann verantwortlich dafür, daß die Eindeutigkeit des Textes verlorengegangen ist. Der Wunsch nach dem 'authentischen' Text in einer eindeutig vorliegenden oder rekonstruierbaren Fassung wird damit kaum erfüllbar.

zu Ziff. 3.134:

Varianten, aber kein Autograph

zu:

aus: E. VOSS, Fast die Quadratur des Kreises. Zur Edition der Suiten für Violoncello solo von Johann Sebastian Bach BWV 1007-1012, in: HENKES, C; HETTICHE, W u.a. (eds.), Schrift - Text - Edition (FS GABLER). editio Bn. 19. Tübingen 2003. S. 305-312.

(303) Die Suiten für Violoncello solo von Johann Sebastian Bach sind weder durch ein Autograph noch durch eine überprüfte Abschrift oder einen zeitgenössischen Druck überliefert. Es gibt also keine Quelle - so der in der Musikwissenschaft gebräuchliche Terminus für Textzeuge -, die nachweislich durch die Hände des Autors gegangen ist. Die Überlieferung ist ohne Autorisation - was jedoch nicht heißt, daß es auch nur den geringsten Zweifel an der Echtheit der Kompositionen gäbe.

Überliefert sind die Suiten durch vier Abschriften und einen Druck. Die Bachforschung belegt sie mit den Siglen A-E:

- A: Abschrift von Bachs zweiter Ehefrau Anna Magdalena, angefertigt nach allgemeiner Annahme zwischen 1727 und 1731;
- B: Abschrift von Johann Peter Kellner aus dem Jahre 1726 (Kellner war zu dieser Zeit Kantor in Frankenhain in Thüringen; er war mit Bach persönlich bekannt und hinterließ Abschriften auch anderer Bachscher Werke);
- C: Abschrift von zwei unbekanntem Kopisten etwa vom Beginn der 2. Hälfte des 18. Jahrhunderts (die Abschrift von Schreiber II setzt in Suite III Bourrée I Takt 12 4. Zählzeit ein);
- D: Abschrift von unbekannter Hand etwa vom Ende des 18. Jahrhunderts;
- E: Erstdruck, Paris 1824 (Janet et Cotelle). Diese Quelle steht C und D so nahe, daß sie in der Bachforschung allgemein als sekundär gilt oder gar als irrelevant.

Schon der große zeitliche Abstand zwischen den Quellen läßt vermuten, daß die Edition nicht dadurch zu bewerkstelligen ist, daß man gleichsam die Quersumme aus ihnen zieht. Diese Vermutung wird durch den inhaltlichen Befund der Quellen bestätigt. Sie unterscheiden sich in so starkem Maße, qualitativ wie quantitativ, daß die Annahme einer auch nur indirekten gemeinsamen Vorlage auszuschließen ist. Diese Unterschiede lassen sich nicht, wie man vielleicht (306) aufgrund der zeitlichen Distanzen zwischen den Quellen vermuten könnte, als Konsequenz von Korrumpierung oder Fehlverständnis und deren Zunahme betrachten. Sie erscheinen vielmehr als Ausweise eigenständiger Versionen. ...

Im übrigen (308) stellt sich die Frage, ob B überhaupt eine Abschrift im herkömmlichen Sinne darstellt; einige Eigenheiten nämlich, deren auffälligste einige auf dem Violoncello unausführbare fünftönige Akkorde sind, deuten darauf hin, daß weniger eine peinlich genaue Kopie beabsichtigt war, als vielmehr eine subjektive Aneignung der Komposition, was - wie selbstverständlich "alcune licenze" einschließt. ... Besonders auffällig sind die regelmäßig, in einigen Suitensätzen geradezu gehäuft auftretenden Staccatopunkte oder -striche. Bach verwendet Staccatozeichen im Unterschied oder Gegensatz zu den Artikulationsbögen selten. ...

(311) Die Reihe der Beispiele ließe sich fortsetzen, so daß zur Qualität der Indizien noch deren Quantität hinzukommt. Ob diese Indizien die Nicht-Authentizität der Quellen B-E und aller ihrer Varianten im Sinne von Endgültigkeit beweisen, bleibe dahingestellt und soll hier auch nicht behauptet werden. Zumindest jedoch legitimieren sie den grundsätzlichen Zweifel und, unabhängig von den Folgerungen, die man zieht, die Notwendigkeit der strikten Scheidung der Quellen. Die Konsequenz des Zweifels an der Authentizität der Quellen B-E aber ist notwendig die Überzeugung, daß Quelle A, die Abschrift von Bachs Ehefrau Anna Magdalena, diejenige Quelle mit dem größten Authentizitätsgrad ist. Eine Edition, die der Bachschen Niederschrift und Intention so weit wie nur möglich nahe kommen möchte, hätte also diese Quelle zugrunde zu legen.

Freilich - und hier zeigt sich die Crux der Angelegenheit - ist Quelle A alles andere als unproblematisch. Zum einen ist sie stellenweise unvollständig, was bis hin zum Fehlen ganzer Notenfolgen reicht. ... (312) Aber nicht nur der Versuch, über A hinaus- und näher ans Autograph heranzukommen, ist zum Scheitern verurteilt, auch eine korrekte Wiedergabe von A erscheint utopisch.

zu Ziff. 3.135:

Vorlage ist sakrosankt - bei der Edition

zu:

aus: W. STEPPE Aus Fehlern lernen. in: HENKES, C; HETTICHE, W u.a. (eds.), Schrift - Text - Edition (FS GABLER). editio Bh. 19. Tübingen 2003. S. 349-360.

(352) Ich möchte hier eine Anekdote einflechten, die uns vor Augen führt, daß man darüber auch anders denken kann. In einem seriösen deutschen Verlagshaus sollte nach dem Kriege eine ziemlich anspruchsvolle ontologische Abhandlung, die seit ihrer Erstveröffentlichung im Jahre 1927 schon etliche Auflagen erlebt hatte, noch einmal neu herausgebracht werden. Ein junger Mitarbeiter der Firma, der damals gerade erst am Anfang seiner beruflichen Laufbahn stand (und der mir die Umriss dieser Geschichte später erzählt hat), hatte die technischen Abläufe zu beaufsichtigen. Der Neusatz, welcher nötig geworden war, sollte sich bis in die Paginierung hinein möglichst eng an die vorangegangenen Auflagen anlehnen, damit Zitate, die sich auf jene bezögen, auch weiterhin leicht auffindbar wären. Nun hätte man doch wohl, wenn ohnehin alles beim alten bleiben würde, vollends auch die alten Seitenzahlen im Inhaltsverzeichnis unbesehen abdrucken können, nicht wahr? Unser Gewährsmann ließ sich freilich in seiner professionellen Sorgfalt nicht beirren: Er überprüfte die Angelegenheit und mußte dabei feststellen, daß die den Text gliedernden Überschriften nicht überall ganz sauber in das Inhaltsverzeichnis übertragen worden waren. Unverzüglich ordnete er eine Handvoll Korrekturen an, welche diese Peinlichkeit aus der Welt räumen würden. Als er jedoch seinem Chef davon berichtete, wurde er nicht etwa belobigt, sondern gerüffelt: Wie konnte er sich unterstehen, die vom Autor so und nicht anders vorgegebene Textur des Werkes anzutasten? Ob ihm denn der Name Heidegger keinen Respekt eingejagt habe? Der Gescholtene mußte seine Korrekturen zurückziehen, und so blieb die ruhmreiche Schrift über *Sein und Zeit* dazu verurteilt, bis auf weiteres mit einem leicht fehlerhaften Inhaltsverzeichnis vor die Öffentlichkeit zu treten.

zu Ziff. 3.136:

Textkritik / Präsuppositionen

zu: vprag7.341

aus: Heinrich PLETT, Textwissenschaft und Textanalyse. Semiotik, Linguistik, Rhetorik. UTB 328. Heidelberg 1975

(90f) Allgemein läßt sich über die kommunikative Bedeutung der Präsuppositionen behaupten: Je stärker ein Text von ihnen abhängig ist, desto eher ist die Möglichkeit seiner unterschiedlichen Dekodierung gegeben. Dies ist besonders der Fall bei Schrifttexten, welche der unmittelbaren Möglichkeit einer informativen Rückkopplung im Situationskontext entbehren. Noch schwieriger gestaltet sich das Problem, wenn zwischen Vertextung und Textentschlüsselung ein größerer historischer Abstand liegt, so daß sich das Reservoir möglicher Präsuppositionen inzwischen verändert (z. B. erweitert) hat. Aus der Diachronie dieser Voraussetzungen erklärt sich die historische Entwicklung und Relativität von Textanalysen. Sie ist nicht nur augenfällig bei der Deutung von semantischen Leerstellen, z.B. im politischen Vokabular, sondern gerade auch bei der Füllung von Kohärenzlücken. Kohärenzlücken pflegen in fast allen Texten zu erscheinen. Besonders evident sind sie dort, wo Texte durch mechanische Schäden (Störungen im Kommunikationskanal: z. B. Papierzerfall) fehlende Textglieder aufweisen. Diese müssen vom Editor ergänzt werden, wenn ein kohärenter Text vorgelegt werden soll. Die Disziplin, welche sich u. a. dieser Aufgabe widmet, trägt den Namen »Textkritik«. Die ältere (»klassische«) Textkritik deutete ihre Aufgabe vornehmlich als eine restaurative; sie versuchte auf dem Wege der historischen Rekonstruktion, d. h. durch Ergründung des Autorwillens, eine dem Urtext möglichst entsprechende Textform herzustellen. Alle davon abweichenden Lesarten (Varianten) gelten unter diesem Aspekt als Verfälschungen des Originals, hingegen läßt sich die neuere Textkritik von dem Gedanken leiten, daß überlieferungsbedingte Varianten jeweils als zeichenhafte Repräsentanten für bestimmte Stadien innerhalb des Textbildungsprozesses zu gelten haben. Diese Auffassung wird in verschiedenen Arbeiten einer von Martens/Zeller unter dem Titel Texte und Varianten (1971) herausgegebenen Aufsatzanthologie zu Problemen der Textedition vertreten. Dort heißt es an einer Stelle; »Text ist als solcher schon immer ein der statischen Fixierung sich entziehender Vorgang; er umfaßt die gesamte Sprachwerdung einer intellektiv oder sensorisch erfaßten außersprachlichen Wirklichkeit« (Martens 1971:169), definiert als eine Aufgabe der Textwissenschaft »die Analyse des gesamten Prozesses [der Textbildung] von der ersten Konstitution durch den Texturheber bis zu der jeweiligen Konkretisation durch den jeweiligen Leser unter Einschluß des gesamten Textherstellungsvorganges« (Schmidt, P. 1973: 122).

Mit Überlegungen dieser Art ist angedeutet, welche Bedeutung die Präsuppositionen für die Textbildung im allgemeinen und die Textkohärenz (91) im besonderen haben können. Da Präsuppositionen wie auch der gesamte Akt der Textkommunikation Kenntnisse über die »außersprachliche Wirklichkeit« (Martens), d. h. Referenzobjekte, voraussetzen, kann die Textpragmatik nicht ohne die Textsemantik auskommen. Daß die Umkehrung dieser Bedingungsrelation ebenfalls gilt, wird ein weiteres Kapitel nachweisen.

Definition: Textpragmatische Kohärenz ist begründet in der Person des emittierenden oder perzipierenden Kommunikationsteilnehmers. Dieser ergänzt (substituiert) auf Grund seines Vorwissens (Präsuppositionen) vorhandene Textlücken und schafft dadurch eine kohärente Textfolge. Diese pragmatische

Kohärenz ist vorhanden auf der Ebene einer Tiefenstruktur. Die Variabilität ihrer möglichen Interpretationen schafft voneinander abweichende Kohärenzen und Texte. Der pragmatische Kohärenzbegriff macht deutlich, daß der Text ein prozessuales Phänomen mit unterschiedlichen Entwicklungsstufen darstellt.

3.14 Textstammbaum, Textfamilien

Literatur

WINDRAM, H F; SHAW, P u.a.: Dante's *Monarchia* as a test case for the use of phylogenetic methods in stemmative analysis. *Literary & Linguistic Computing* 23/4 (2008) 443-463.

zu Ziff. 3.141:

Genealogie von Hss - Anleihen bei Bioinformatik

zu:

aus: *Literary & Linguistic Computing* 20/2 (2005)

- 1) WINDRAM, H f; HOWE C J; SPENCER, M: The Identification of Exemplar Chance in the *Wife of Bath's Prologue* Using the Maximum Chi-Squared Method: 189ff

Abstract: Chaucer's *Wife of Bath's Prologue* survives in both hand-written and early print witnesses dating from the 15th century. The introduction of material from more than one exemplar into a new copy results in contamination of a textual tradition. This contamination causes problems in standard phylogenetic analysis. We use an application of the maximum χ^2 method (developed for the detection of recombination in DNA sequences) to identify locations where scribes may have changed their exemplar whilst copying the tale. Our results are largely in agreement with other published sources, indicating that this method may prove useful in the analysis of a contaminated tradition.

Anmerkungen: Mit dem Ansatz kann - rechnerisch abstrakt - die Nähe/Ferne von Textvarianten zueinander bestimmt und in Baumform sichtbar gemacht werden. Man muss sich aber hüten, die Baumdarstellung zeitlich-genealogisch zu interpretieren: es ist möglich, dass Textpaare, die vom materialen Text her wenig miteinander zu tun haben, annähernd den gleichen Ähnlichkeitsgrad zu einander haben. Um historisch-genealogisch die Daten auszuwerten sind Methodenschritte nötig, die auch berücksichtigen, was *materialiter* in den Texten geschrieben ist.

- 2) OUSAKA, Y; YAMAZAKI, M: Genealogical Classification of Saddharmapundarika Manuscripts Based on Many-Variable Analysis: 193ff

Abstract: Many manuscripts of the Saddharmapundarika, which are among the most important manuscripts for the study of Buddhism, have been discovered in very different localities and are classified according to their place of discovery into the following three groups: Nepalese, Kashmirian, and Central Asian manuscripts. For the genealogical classification of these manuscripts, principal component analysis and cluster analysis, were applied to the data. As a result, we could successfully classify these manuscripts into two large groups and several smaller groups: one large group consists of ten paper manuscripts from Nepal and the other comprises nine palm-leaf plus two paper manuscripts. The Kashmir and Central Asian manuscripts and few from Nepal manuscripts belong to the small groups.

Anmerkungen: Methoden wie Cluster-Analyse berücksichtigen, was *materialiter* im Text steht (Wortformen in bestimmten Konstellationen). Diese Befunde in verschiedenen Versionen können verglichen werden. Dies könnte zumindest ein Schritt sein, um tatsächlich zu einer Rekonstruktion der Genealogie im Textbildungsprozess zu kommen. Vgl. Anleihen aus der Biologie: Zu hohem Prozentsatz stimmen die Genome von Affen und Menschen überein. Irgendwann hat sich diese Ausstattung in kleinen Bereichen differenziert. Ab da lassen sich diese Differenzen in allen weiteren Exemplaren nachweisen: der genealogische Strang hat sich verzweigt.

3.15 Echtheitsüberprüfung**Literatur**

SCHLITZ, S A: The TEI as luminol: Forensic philology in a digital age. *Literary & Linguistic Computing* 24/2 (2009) 173-185.

zu Ziff. 3.151:

Echtheitsüberprüfung

zu:

aus: FORSYTH, R S, HOLMES, D I, u.a.: Cicero, Sigonio, and Burrows: Investigating the Authenticity of the *Consolatio*. *Literary and Linguistic Computing* 14/3 (1999) 375-400.

(375) When his daughter Tullia died in 45 BC, the Roman orator Marcus Tullius Cicero (106-43 BC) was assailed by grief, which he attempted to assuage by writing a philosophical work now known as the *Consolatio*. Despite its high reputation in the classical world, only fragments of this text - in the form of quotations by subsequent authors - are known to have survived the fall of Rome. However, in 1583 a book was printed in Venice purporting to be a rediscovery of Cicero's *Consolatio*. Its editor was a prominent humanist scholar and Ciceronian stylist called Carlo Sigonio. Some of Sigonio's contemporaries, notably Antonio Riccoboni, voiced doubts about the authenticity of this work; and since that time scholarly opinion has differed over the genuineness of the 1583 *Consolatio*. The main aim of this study is to bring modern stylometric methods to bear on this question, to see whether internal linguistic evidence supports the belief that the *Consolatio* of 1583 is a fake, very probably perpetrated by Sigonio himself. A secondary objective is to test the application of methods previously used almost exclusively on English texts to a language with a different structure, namely Latin. Our findings show that the language of the 1583 *Consolatio* is extremely uncharacteristic of Cicero, and indeed that the text is much more likely to have been written during the Renaissance than in classical times. The evidence that Sigonio himself was the author is also strong, though not conclusive.

3.2 Aufbereitung des Textes

Literatur

- AEJEMELAEUS, A: Function and Interpretation of ׀ in Biblical Hebrew. JBL 105 (1986) 193-209.
- ALTMANN, H (ed./eds.): Intonationsforschungen. Tübingen 1988
- ANDERSEN, F I: The Hebrew verbless clause in the Pentateuch. Journal of Biblical literature monograph series, 14. Nashville 1970 // I.2.1
- BOLINGER, D: Intonation and its uses. Melody in grammar and discourse. London 1989
- BRINKER, K: Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin ³1992 // I.2.4
- BUCHER-GILLMAYR, S: Biblical words transmitted by lyric poetry. Actes du cinquième Colloque International 'Bible et Informatique: »Translation et Transmission«, Aix-en-Provence 1-4 septembre 1997. Genève 1998 357-368.: Honoré Champion Éditeur // EB
- CHAO, Y R: Language and Symbolic Systems. Cambridge 1968 // I.2.1
- FECHT, G: Metrik des Hebräischen und des Phönizischen. Ägypten und Altes Testament 19. Wiesbaden 1990: Harrassowitz // I.2.1
- FREEDMAN, D N; FORBES AD; ANDERSEN FI: Studies in Hebrew and Aramaic orthography. Biblical and Judaic Studies from the University of California, San Diego, 2. Winona Lake 1992 // I.2.0
- GOEBL, H: Dendrogramme im Dienst der Dialektometrie. Zwei hierarchisch-agglomerative Klassifikationen von Daten des Sprachatlasses AIS in: KLENK, U (ed./eds.): Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.73. Stuttgart 1992 55-73. // I.2.7
- GOEBL, H: Spannungsverhältnisse in dialektalen Netzen. Ein Hinweis zu disziplinübergreifender Diskussion in: KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 63-83. // I.2.7
- GRIMM, J; GRIMM, W: Deutsches Wörterbuch. Bd.5. Leipzig 1873
- GROSS, W: Die Pendenskonstruktion im Biblischen Hebräisch. Studien zum althebräischen Satz I. ATS 27. St. Ottilien 1987 // I.2.1
- HALLIDAY, M A K: An introduction to functional grammar. London ²1994: Arnold // I.2.4
- JENNI, E: Die hebräischen Präpositionen. Band 1: Die Präposition Beth. Stuttgart 1992: Kohlhammer // I.2.7
- JENNI, E: Die hebräischen Präpositionen. Band 2: Die Präposition Kaph. Stuttgart 1994: Kohlhammer // I.2.1
- JOÜON, P: Grammaire de l'Hébreu Biblique. Rom 1965 (reprint von 1923) // I.2.1
- KÖNIG, F E: Historisch-kritisches Lehrgebäude der hebräischen Sprache. Syntax der hebräischen Sprache Bd.III

- 2.Hälfte 2.Teil. Hildesheim 1979 (reprint von 1897) // I.2.1
- LEUNG, C-H; KAN, W-K: A Statistical Learning Approach to Improving the Accuracy of Chinese Word Segmentation. Literary and Linguistic Computing 11 / 2 (1996) 87-92. // EB
- MEISENBURG, T: Graphische und phonische Integration von Fremdwörtern am Beispiel des Spanischen. Zeitschrift für Sprachwissenschaft 11,1 (1992) 47-67.
- MEYER, R: Hebräische Grammatik Bd.I-IV. Berlin ³1966-1972 // I.2.1
- MURTONEN, A: Hebrew in its West semitic setting. A comparative survey of non-masoretic hebrew dialects and traditions. Part one: A comparative Lexicon. Section A: Proper Names. Leiden 1986
- RICHTER, W: Grundlagen einer althebräischen Grammatik. A. Grundlagen einer sprachwissenschaftlichen Grammatik. B. Die Beschreibungsebenen: I. Das Wort (Morphologie). ATS 8. St. Ottilien 1978 // I.2.1
- RICHTER, W: Grundlagen einer althebräischen Grammatik. B. Die Beschreibungsebenen: II. Die Wortfügung (Morphosyntax). ATS 10. St. Ottilien 1979 // I.2.1
- SAILHAMER, J H: The Translation Technique of the Greek Septuagint for the Hebrew Verbs and Participles in Psalms 2-41. Studies in biblical Greek 2. Frankfurt / M 1989 // I.2.1
- SCHWEIZER, H: Rezension zu: Walter Groß, Die Pendenskonstruktion im Biblischen Hebräisch. Studien zum althebräischen Satz I. ATS 27. St.Ottilien 1987. ThLZ 114 (1989) 506-508. Abk: (1989REZ)
- SIEBESMA, P A: The Function of the niph'al in Biblical Hebrew. Biblical Hebrew. In relationship to other passive-reflexive verbal systems and the pu'al and hoph'al in particular. Studia Semitica Neerlandica 29. Assen 1991 // I.2.1
- SMITH, N; McENERY, J u.a.: Issues in Transcribing a Corpus of Children's Handwritten Projects. Literary and Linguistic Computing 13 (4 / 1998) 217-225.
- SZEMERÉNYI, O: Einführung in die vergleichende Sprachwissenschaft. Darmstadt ⁴1990 // I.2.1
- WARDHAUGH, R: An Introduction to Sociolinguistics. Oxford ²1992: Blackwell
- WILLI-PLEIN, I: ׀. Ein Übersetzungsproblem. VT 23 (1973) 90-99.
- WONNEBERGER, R: Zur Syntax und Pragmatik der Partikel im Bereich der Bibel in: WEYDT, H (ed./eds.): Die Partikeln der deutschen Sprache. Berlin / New York 1979 488-499.
- ZIMA, P V; STRUTZ, J: Komparatistik. Einführung in die Vergleichende Literaturwissenschaft. UTB 1705 1992: A. Francke

3.21 Übersetzungsbegründung

Literatur

- ADAMZIK, K: Syntax und Textgliederung. Hypotaktischer Stil, Nominalstil, graphischer Stil in: HINDELANG, G; ROLF, E; ZILLIG, W (ed./eds.): Der Gebrauch der Sprache (FS F. HUNDSNURSCHER). Münster 1995 15-39.
- BANERJEE, B P, S; MULLICK: Word Sense Disambiguation and WordNet Technology. Literary and Linguistic Computing 22 / 1 (2007) 1-15.
- BARR, J: Comparative Philology and the Text of the Old Testament. Oxford 1983 (reprint von 1968) // I.2.1
- BARTON, E L: Nonsentential constituents. A theory of grammatical structure and pragmatic interpretation. Pragmatics and Beyond NS 2 1990: Benjamins
- KOLLER, W: Einführung in die Übersetzungswissenschaft. UTB 819. Heidelberg ⁴1992 // I.2.1
- KREUZER, H . Zeitschrift für Literaturwissenschaft und Linguistik 84 / 21 (ed./eds.): [Übersetzung]. Göttingen (1991): Vandenhoeck & Ruprecht // I.2.1
- OCHS, E: Transcription as Theory in: OCHS, E; SCHIEFFELIN, B B (ed./eds.): Developmental Pragmatics. London 1979 // I.2.4
- QUIROZ, G; BERTHOUD-PAPANDROPOULOU, I; THOMMEN, E; VOGEL, C (ed./eds.): Les unités discursives dans l'analyse sémiotique: la segmentation du discours. Textanalyse in Universität und Schule 12. Bern 1998: Peter Lang // I.2.1
- SCHINDELE, M: Textkonstituierung zu Daniel 8 in: BADER, W (ed./eds.): "Und die Wahrheit wurde hinweggefegt". Daniel 8 linguistisch interpretiert. THLI 9. Tübingen 1994 3-16.: Francke // Raum B106
- STOLZE, R: Hermeneutisches Übersetzen. Linguistische Kategorien des Verstehens und Formulierens beim Übersetzen. TBL 368. Tübingen 1992: Narr // I.2.4
- STOLZE, R: Übersetzungstheorien. Eine Einführung. Tübingen ²1997: Narr // I.2.0
- WONNEBERGER, R: Normaltext und Normalsynopse. Neue Wege bei der Darstellung alttestamentlicher Texte. Zeitschrift für Sprachwissenschaft 3,2 (1984) 203-233.

zu Ziff. 3.211:

Verstehensprüfung / GOETHE'S 'Faust' '

zu:

aus: SCHÖNE, A (ed.) Johann Wolfgang Goethe. Faust. Texte. Frankfurt/M 1994. (= Bd.I) SCHÖNE, A Johann Wolfgang Goethe. Faust. Kommentare. Frankfurt/M 1994. (= Bd.II)

SCHÖNE, I,83 + II,269

Da wird der Geist euch wohl dressiert,
In spanische Stiefeln eingeschnürt,
Daß er bedächtiger so fort an
Hinschleiche die Gedankenbahn,
Und nicht etwa, die Kreuz' und Quer
Irrlichteliere hin und her.

spanische Stiefeln] Die Schienbeine zusammenpressendes Marterinstrument beim zweiten Grad der Inquisitionstortur.

Irrlichteliere] Goethesche Wortbildung für eine unstete, sprunghaft-hüpfende Bewegung, wie das Irrlicht sie vortäuscht.

SCHÖNE, I,114 + II,290

MEPHISTOPHELES ...
Jetzt ohne Schimpf und ohne Spaß,
Ich sag' euch, mit dem schönen Kind
Geht's ein- für allemal nicht geschwind.

Schimpf] in der älteren Bedeutung: Scherz

SCHÖNE, I,127 + II,302

MEPHISTOPHELES ...
Ein schönes Fräulein nahm sich seiner an,
Als er in Neapel fremd umher spazierte;
Sie hat an ihm viel Lieb's und Treu's getan,
Daß er's bis an sein selig Ende spürte.

MARTHE
Der Schelm! der Dieb an seinen Kindern!

Über Frau Marthes Mann:]

Neapel] gängige Kurzform von Neapel

- Anspielung auf eine Dirne, von der er sich die Syphilis geholt hatte.

- *Schelm* nicht als witziger Mensch, sondern nach älterem Sprachgebrauch "verworfenener Mensch, Betrüger, treuloser Gatte oder Geliebter".

SCHÖNE, I,151 + II,326

MARGARETE: ...
Doch meine Mutter schläft nicht tief:
Und würden wir von ihr betroffen,

Ich wär gleich auf der Stelle tot!

betroffen] Für: angetroffen, ertappt.

SCHÖNE, I,173 + II,351

Das drängt und stößt, das ruscht und klappert!
Das zischt und quirlt, das zieht und plappert!

ruscht] nicht zu verbessern in "rutscht". Nach dem Grimmschen Wörterbuch: "sich mit Geräusch bewegen". Lautmalend.

zu Ziff. 3.212:

Übersetzungstheorie

zu: vprag8.3014

Vgl. im Internet:

<http://www-ct.informatik.uni-tuebingen.de/ct/manuskripte.htm>

Unter Ziff.1 drei Aufsätze zu den Problemen des genauen Übersetzens, der computerunterstützten Kontrolle der Qualität von Übersetzungen, der Verbindung von Wörtlichkeit der sprachlichen Wiedergabe und der Hinzunahme von notwendigen Zusatzinformationen.

Im SLANG-Projekt eigenes *tool*: aus einem Quelltext wird eine Eins-zu-Eins-Entsprechung in der Zielsprache erstellt. Da die Grammatiken differieren müssen Hilfskonventionen die Unterschiede ausgleichen. Diese Abbildung des Quelltextes in der Zielsprache wird *Arbeitsübersetzung* genannt.

Die *Arbeitsübersetzung* dokumentiert, dass man die Wortbedeutungen und grammatischen Konstruktionen des Quelltextes genau wahrgenommen hat.

Auf dieser Basis kann dann erst eine *literarische Übersetzung = Übertragung* erstellt werden. In sie fließen im Idealfall auch analytische Erkenntnisse zum Text und das nötige Wissen zu historischem Umfeld und Kultur ein.

So gesehen bildet sich eine Übertragung erst **im Kopf des Lesers** nachdem ihm die unterschiedlichen Informationsmöglichkeiten zur Verfügung gestellt worden waren.

Experiment dazu: Vgl.

<http://www-ct.informatik.uni-tuebingen.de/daten/jgrollen.pdf>

Aktualisierte und materialreichere Zusammenfassung (für Übersetzungstheorie vgl. "Einleitung"):

<http://www-ct.informatik.uni-tuebingen.de/daten/jguebers.pdf>

zu Ziff. 3.213:

Ilias Übersetzung oder Übertragung?

zu:

aus: HOMER, Ilias. Übertragen von Raoul Schrott. Kommentiert von Peter Mauritsch. München 2008.

(XXXIII)

Bei dieser Fassung handelt es sich deshalb nicht um eine philologische - das 'Wörtliche liebende' - Übersetzung. Sie versucht die entgegengesetzte Richtung einzuschlagen und Homer von seinem Ufer abzuholen, um ihn ins Heute zu bringen. Um dort auch anzukommen, verzichtet sie auf die Nachbildung der Konstituenten seiner Diktion - die durch den Hexameter bedingte Formelsprache -, um ihren Sinngehalt mit modernen Mitteln wiederzugeben.

Das Übersetzen gerät dadurch zu einer Übung im Ersetzen: bei dem die Art, wie etwas von Homer formuliert wurde, durch das substituiert wird, was er sagt und wie er es meint. ...

Daß die Poesie eine Suggestionkunst ist, die ihre Formen einsetzt, um jenseits notgedrungener sprachlicher Defizienzen auf eine Realität zu verweisen, ist eine Universalie. Schon die älteste uns überlieferte Poetik, die chinesische, definierte Dichtung als etwas, das unser eigentliches Denken in Worte zu fassen versucht - oder anders gesagt: Poesie verbalisiert Intention: Unter diesem Gesichtspunkt ist jedes dichterische Werk bereits eine erste Übersetzung dessen, was der Autor fühlt und denkt und dennoch nur durch die beschränkten Mittel einer Diktion wiedergeben kann - um es auf diese Weise dem Leser in den Kopf zu setzen.

Damit dies nach dreitausend Jahren wieder gelingen kann, stellt man sich am besten vor - so Wieland in einem Kommentar zur Voßschen Übersetzung -, wie Homer gesungen hätte, wenn unsere deutsche Sprache die seine (XXXIV) gewesen wäre. Zu berücksichtigen ist dabei, daß der Anspielungshorizont Homers ein völlig anderer als der gegenwärtige war. Das bedeutet zunächst, daß man hinter jedem seiner Verse das Bild sehen muß, das er skizziert - um es dann mit heutigen Mitteln nachzuzeichnen. Wo er Realien abbildet, die der damaligen Lebenswirklichkeit entnommen waren, sind diese dem Leser erst wieder zugänglich zu machen. Und was seinen damaligen Zuhörern etwa an mythischem Hintergrundwissen präsent war, muß heute paraphrasiert, was einmal zwischen den Zeilen mitschwang, aktualisiert werden.

Diese Fassung schreibt deshalb gewissermaßen als Metaversion mit, was wir jetzt über den Text und seine Hintergründe wissen.

(42) achilleús und odysseús - die er unaufhörlich anpöbelte - war er längst ein dorn im auge; jetzt aber warf er agammennon alles an den kopf, in empörenderem brustton beifallheischend - und brachte die achaier doch nur gegen sich auf und in harnisch:

agamennon!! was paßt dir denn jetzt schon wieder nicht??
was willst du noch?? ist deine hütte etwa nicht voll beute??
und voll der edelsten frauen, die du als allererster kriegst
wenn wir sie in irgendeiner stadt schnappen? ja -haben dir
diese reichen troianischen roßtäuscher denn immer noch
nicht genug an gold als lösegeld für ihre muttersöhnchen
die ich oder sonst wer als geisel nahm, hergeschleppt?
brauchst du noch ein mädchen, das du entjungfern kannst

um es nach lust und laune durchzuziehen? was ist das denn für ein führer, der uns bloß deswegen in den krieg schickt? und ihr? bloße weicheier seid ihr! langhaarige waschweiber!! kommt - segeln wir heim und lassen ihn samt seinem plunder und seinen frauchen am strand sitzen, damit er endlich merkt was er ohne uns, die wir nichts taugen sollen, zustandebringt!
der ehrabschneider hat doch dem achilleús - der von beiden noch der beßre ist - nur aus raffgier seine trophäe gestohlen!

3.22 Textsegmentierung**Literatur**

- GIL, A: Textadverbiale in den romanischen Sprachen. Eine integrale Studie zu Konnektion und Modalisatoren in Spanischen, Französischen und Italienischen. Bonner Romanistische Arbeiten 53. Frankfurt 1995: Peter Lang // I.2.1
- GROSS, W: Zum Problem der Satzgrenzen im Hebräischen. Beobachtungen an Pendenskonstruktionen. BN 35 (1986) 50-72.
- HALFORD, B K: Talk Units. The Structure of Spoken Canadian English. ScriptOra 87. Tübingen 1996 // n.b.
- KOLB, H-P: Monotonicity and Locality. Problems of Principle Based Parsing. Zeitschrift für Sprachwissenschaft 10,2 (1991) 161-190.
- LOHNSTEIN, H: Grammatik und Parsing. Zeitschrift für Sprachwissenschaft 13,1 (1994) 56-109.
- MIKHEEV, A: Text Segmentation in: MITKOV, R (ed./eds.): The Oxford Handbook of Computational Linguistics. Oxford Handbooks Series. Oxford 2003 201-218.: Oxford University Press // I.2.4
- NAUMANN, S: Adaptives Parsen in: KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 127-147. // I.2.7
- OSWALD, W: Text Segmentation and Pragmatics. Actes du Quatrième Colloque International "Bible et Informatique: «Matériel et Matière» L'impact de l'informatique sur les études bibliques.", Amsterdam, 15-18 August 1994. Paris-Genève 1995 140-152. // I.2.0
- PATERNOST, J: Rezension zu: Konrad EHLICH, Interjektionen. Tübingen 1986. General Linguistics 28 (1988) 228-236.
- RASOLOSON, J N: Interjektionen im Kontrast. Am Beispiel der deutschen, madagassischen und französischen Sprache. Arbeiten zur Sprachanalyse 22. Frankfurt 1994 // I.2.1
- RICHTER, W: Transliteration und Transkription. Objekt- und metasprachliche Metazeichensysteme zur Wiedergabe hebräischer Texte. ATS 19. St. Ottilien 1983 // I.2.1
- SCHMITZ, KD: Automatische Segmentierung natürlichsprachiger Sätze. Sprache und Computer 3. Hildesheim 1986: OLMS // I.2.1
- SCHWEIZER, H: Wovon reden die Exegeten?. Zum Verständnis der Exegese als verstehender und deskriptiver Wissenschaft. Theologische Quartalschrift 164 (1984) 161-185.
- SCHWEIZER, H: Textsegmentierung in Äußerungseinheiten. Sprache und Datenverarbeitung 18 / 2 (1994) 3-18.
- STEUBE, A; ZYBATOW, G (ed./eds.): Zur Satzwertigkeit von Infinitiven und Small Clauses. LA 315. Tübingen 1994: Niemeyer // I.2.4
- ZIFONUN, G: Kommunikative Einheiten in der Grammatik. Tübingen 1987

zu Ziff. 3.2201:**Satz**

zu:

aus: M. HEPPERLE, Gestammelte Werke. 1979.Satzzeichen setzen

Ich schreite wie ein Sämann durch die Zeilen
und streue Pünktchen, Doppelpünktchen, in die Welt.
Durch reinen Zufall trifft sich aber doch bisweilen,
daß eines dahin, wo es stehen müßte, fällt.

Geneigter Leser, mög er sich bequemen,
falls ihn ein fehlerhaft-gestreutes stört,
es zu versetzen, es an jene Stelle nehmen
dahin, wo es nach seiner Meinung hingehört.

Wann ist Satz ein Satz, wann ist er keiner?
Bei dieser Frage seh'ich überhaupt kein Land.
Mach ich ihn länger, zwei draus, oder keiner?
Was meist dagegensteht ist der Satzgegenstand.

Die Frage, wo ein Fragezeichen gar vonnöten
wird meist sogar von mir korrekt gelöst,
weil sich hier orthographisch kaum Gelegenheiten böten,
daß man als Deutscher gegen das Gesetz verstößt.

In Klammern setz ich selten meine Sachen,
ich sperre ugnern einen Satz in einen Zaun.
Vergeß ich einerseits die Klammer wieder zuzumachen,
so ist mir andererseits noch keiner abgehau.

Nun, sein Sie tolerant zu den Toleranzen
und scheinen sie auch manchmal reichlich groß.
Ich wäre froh, betrachte ich es mir im Ganzen,
es gäb bloß Zeichensetzungsdiskrepanzen.
Sie sagen jetzt mit Recht "Was heißt denn hier schon 'bloß'".

zu Ziff. 3.221:

Textbegrenzungssignale

zu:

aus: BRINKER, K, Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin ²1988.

(18) Die Bestimmung, daß Texte b e g r e n z t e Satzfolgen darstellen, verweist auf die sog. Textbegrenzungssignale. Es handelt sich dabei um bestimmte sprachliche und nicht-sprachliche Mittel. Zu den sprachlichen Signalen für Textanfang bzw. Textschluß gehören z.B. Überschriften, Buchtitel und bestimmte Einleitungs- und Schlußformeln; an nicht-sprachlichen Mitteln sind vor allem bestimmte Druckanordnungskonventionen zu nennen (Buchstabengröße bei Überschriften, Leerzeilenkontingent usw.) sowie die auf mündliche Texte beschränkte Sprechpause (die als Abgrenzungskriterium allerdings nicht immer eindeutig ist). Diese und andere Textbegrenzungssignale kennzeichnen also die Zeichen- bzw. Satzfolgen, die für den Emittenten den Charakter der Selbständigkeit und Abgeschlossenheit besitzen, kurz: die er als Texte verstanden wissen will.

zu Ziff. 3.222:

Textsegmentierung - unklare Theorie

zu:

aus: BRINKER, K, Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin ²1988.

(22) Wir wollen zunächst eine primär ausdrucksorientierte Satzdefinition vorstellen. Wir knüpfen dazu an die auf L. Tesnière zurückgehende Dependenz- bzw. Valenzgrammatik an. Auf der Grundlage des Valenzmodells kann S a t z als eine sprachliche Einheit definiert werden, die sich aus einem Verb (Prädikat) als dem strukturellen Zentrum und einer Reihe von Satzgliedpositionen (Subjekt, Objekte, Adverbialbestimmungen usw.) konstituiert, die jeweils in bestimmten Abhängigkeitsrelationen zum "tragenden" Verb stehen. Sätze können danach als einfache Sätze oder als Teilsätze (z.B. Haupt- und Gliedsätze in sog. Satzgefügen) realisiert sein. Der Unterschied zwischen einer solchen grammatischen Satzdefinition und einem auf der Interpunktion beruhenden Satzbegriff läßt sich recht gut an den folgenden Versen von M. Claudius verdeutlichen:

(1) Der Mond ist aufgegangen, die goldnen Sternlein prangen am Himmel hell und klar; der Wald steht schwarz und schweiget, und aus den Wiesen steigt der weiße Nebel wunderbar.

Von der Interpunktion her gesehen liegt nur ein Satz vor; unter grammatischem Aspekt können wir aber zumindest vier (Teil-)Sätze unterscheiden.

(24) Der valenzgrammatische Satzbegriff ist - wie gesagt - primär an der Ausdrucksstruktur des Satzes orientiert. Als (komplexe) sprachliche Zeichen haben Sätze aber auch eine Inhaltsseite; sie bezieht sich vor allem auf die Satzbedeutung (im engeren Sinn), d.h. auf den vom Satz ausgedrückten Sachverhalt, den wir als Proposition bezeichnen.

Zur Explikation des Propositionsbegriffs knüpfen wir an die Sprechakttheorie. J.R. Searles an, der zwischen der illokutiven Rolle und dem propositionalen Gehalt einer Äußerung differenziert. Eine Sprechhandlung wie "ein Versprechen geben" (z.B. *Ich verspreche dir, daß ich morgen komme*) läßt sich danach in zwei Teile zerlegen: Der erste Teil enthält den Indikator des Sprechhandlungstyps, bezeichnet also den Modus der Kommunikation, d.h. die Beziehung, die der Sprecher zum Angesprochenen herstellt (*ich verspreche dir*); für diesen Aspekt der Sprechhandlung verwendet Searle (im Anschluß an J.L. Austin) die Termini "illokutionärer Akt" bzw. "illokutionäre Rolle" (eine Äußerung). Der zweite (grammatisch abhängige) Teil (*daß ich morgen komme*) ist gegliedert in die Referenz, d.h. die Setzung eines Kommunikationsgegenstandes (*ich*), und die Prädikation, d.h. die Zuordnung von Eigenschaften zum gesetzten Gegenstand (*morgen kommen*). Searle nennt diese Komponente einer sprachlichen Handlung den "propositionalen Akt" bzw. "propositionalen Gehalt" (einer Äußerung). In syntaktischer Hinsicht wird die Referenz durch Eigennamen, Pronomen und andere Nominalgruppen, die Prädikation durch Prädikate realisiert.

(25) Aus der Unterscheidung zwischen Illokution und Proposition folgt, daß verschiedene illokutive Akte den gleichen propositionalen Gehalt haben können. Die Proposition *daß Hans den Raum verläßt* kann z.B. mit folgenden illokutiven Rollen verbunden werden: *Hans verläßt den Raum* (Feststellung oder Behauptung) - *Hans, verlaß den Raum!* (Aufforderung) - *Verläßt Hans den Raum?* (Frage) - *Würde Hans doch den Raum verlassen!* (Wunsch) usw. Der Sprecher referiert in den Äußerungen auf dieselben Gegenstände (Hans, ein bestimmter

Raum), zwischen denen eine bestimmte Relation ("verlassen") hergestellt wird (Prädikation). ...

So ist z.B. die Äußerung *Hans hat das Buch trotz seiner Krankheit beendet* ein Satz mit zwei Propositionen, da die konzessive Adverbialbestimmung *trotz seiner Krankheit* eine Proposition vertritt (explizit: *Hans hat das Buch beendet, obwohl er krank war/ist*). Demgegenüber besteht die Äußerung *Hans glaubt, daß der Urlaub schön wird* aus z w e i Sätzen, die aber nur eine Proposition realisieren. Bei dem *daß*-Satz handelt es sich um einen sog. Objektsatz (*Hans glaubt X*), dessen Inhalt in die Proposition des übergeordneten Satzes integriert ist (als zweiter Referenzteil).

zu Ziff. 3.223:

Satzdefinition

zu:

aus: EHLICH, K, Der Satz. Beiträge zu einer pragmatischen Rekonstruktion, in: REDDER, A; REHBEIN, J (eds.), Grammatik und mentale Prozesse. Tübingen 1999. S. 51-68.

(52) ... selbst dort, wo Linguistik sich wieder explizit als Satzlinguistik versteht, unterbleibt die Bestimmung der Grundkategorie "Satz". Genauer: Man verläßt sich auf das, was man immer schon zu wissen meint.

(53) Betrachten wir auch nur in einigen europäischen Sprachen das, was als Terminus für "Satz" aufzufinden ist, so kommen wir auf die in (1) zusammengestellte Liste.

(1)	Satz (dt.)	propositio (lat.)
	setning (norw.)	proposition (frz.)
	zin (ndl.)	proposition (engl.)
	mening (norw.)	hermeneia (gr.)
	sententia (lat.)	logos (gr.)
	sentence (engl.)	phrasis (gr.)
		phrase (frz.)
		phrase (engl.)

Die Liste (1) ließe sich durch weitere Beispiele, etwa aus dem Slawischen, ergänzen. Bereits in der vorliegenden Form macht sie sehr deutlich, wie divergent die Beziehungen sind. Weitere Ausdrücke spielen - mit einer nicht immer ganz klaren Abgrenzung - im Englischen eine Rolle, wo der Terminus

(2a) clause

eingeführt ist, der seinerseits auf

(2b) clausula

zurückgehen dürfte (vgl. z.B. Audax in Keil 1880, 324).

In den verschiedenen Sprachen konkurrieren also durchaus verschiedene Bezeichnungen, z.T. sogar in ein- und derselben Sprache: Im Französischen haben wir zwei Kandidaten für den Satz, nämlich die "proposition" und die "phrase", im Englischen neben "sentence" auch "phrase" und "clause". ...

(54) Da haben wir zunächst solche Bezeichnungen wie "phrasis" oder "logos", die sich auf die Äußerung selbst, das Äußern, beziehen. "phrasis" ist das Reden. "logos", das "Wort", ist eigentlich das Sprechen, das "legein", bzw. das, was sich als Resultat daraus ergibt (vgl. Benseler 1900, s.v.). Andere Bezeichnungen hingegen wie etwa "propositio", "proposition", "proposition" haben es im wesentlichen zu tun mit einem spezifischen Teilaspekt des *II-Bereichs*, des mentalen Bereichs, und zwar bezogen auf den *Sprecher*; mit dem nämlich, was der Sprecher "proponit", was er vorlegt, was er *mental sozusagen vorbereitet* (vgl. Hoffmann 1996), um es dann in der Verbalisierung zu äußern.

Die "sententia" hingegen hängt mit "sentire" zusammen, dem Fühlen oder "Wahrnehmen". Es bezieht sich auf einen ganz anderen Teilsektor des *mental Bereichs*, nämlich die Wahrnehmungstätigkeit insgesamt und deren mentale Umsetzung.

Anders liegt es bei der "hermeneia", die sich in der klassischen aristotelischen Untersuchung "peri hermeneias" auf das bezieht, was der "Satz" ist. "hermeneia" hat es ersichtlich mit dem zu tun, was im Deutschen als "Hermeneutik" präsent ist, also der Verste-

henslehre. "hermeneia" bezieht sich also gleichfalls auf den *mentalen Bereich*, aber den mentalen Bereich auf der Seite des *Hörers* einer Äußerung, auf dessen Verstehens-, dessen Umsetzungsprozesse.

Der mit dem deutschen Wort "Sinn" assoziierte niederländische Ausdruck "zin" hingegen bezieht sich auf eine Art *Abstraktion* aus diesen *mentalen Qualitäten* (vgl. norw. "mening", entsprechend dt. "Meinung" bzw. engl. "meaning").

Ganz anders liegt es beim "Satz", der, wie eben schon ausgeführt, als *Setzung* (vgl. norw. "setning") auf das *Äußerungsergebnis* bezogen ist, insbesondere auf den spezifischen Anwendungsbereich des Rechtssatzes. Das *Feld*, das hier aufgerufen wird, assoziiert zugleich einen anderen Ausdruck, der in der Satzdiskussion eine große Rolle spielt - insbesondere in der deutschen philosophischen Tradition -, nämlich das "Urteil". (Auch das lat. "sententia" hat als Äußerung der Meinung eine derartige Einschränkung auf den Urteilsspruch früh erfahren. Im engl. "sentence" als Urteilsspruch und im frz. "sentence" setzt sich dies fort bis dahin, daß im Französischen neben der Bedeutung "Sentenz" nur noch diese juristische Bedeutung verblieben ist.)

zu Ziff. 3.2231:

"Satz" - pragmatische Einheit

zu: vsem5.101

aus: R. B. BRANDOM, Begründen und Begreifen. Eine Einführung in den Inferentialismus. Frankfurt/M 2001.

(166) Wenn man statt dessen mit der semantischen Interpretation von Sätzen anfängt - denn schließlich besitzen sie ja pragmatische Priorität -, welchen Grund gibt es dann, diese Sätze so zu zerlegen, um auch subsententiale Ausdrücke zu interpretieren? Warum sollte man überhaupt das semantisch signifikante Vorkommen irgendeiner weiteren Kategorie in Betracht ziehen? Frege gibt zu Beginn einer seiner späteren Arbeiten diese Antwort: "Erstaunlich ist, was die Sprache leistet, indem sie mit wenigen Silben unübersehbar viele Gedanken ausdrückt, daß sie sogar für einen Gedanken, den nun zum ersten Mal ein Erdenbürger gefaßt hat, eine Einkleidung findet, in der ihn ein anderer erkennen kann, dem er ganz neu ist. Dies wäre nicht möglich, wenn wir in dem Gedanken nicht Teile unterscheiden könnten, denen Satzteile entsprächen."

Die Fähigkeit, eine unbestimmte Zahl neuer Sätze hervorzubringen und zu verstehen, stellt ein verblüffendes und wesentliches Merkmal der sprachlichen Praxis dar. Chomsky hat nachfolgend darauf hingewiesen, daß eine solche Kreativität eher die Regel als die Ausnahme ist. Praktisch jeder Satz, den ein erwachsener Muttersprachler äußert, wird zum ersten Mal geäußert - nicht bloß von diesem Sprecher, sondern zum ersten Mal in der Geschichte der Menschheit. Der hohe Anteil von sentential Neuem zeigt sich in den Untersuchungsberichten empirisch aufgezeichneter Gespräche und wird statistisch belegt, wenn man die Anzahl der Sätze von bis zu dreißig Wörtern mit der Anzahl vergleicht, die deutsche Sprecher bis zu diesem Zeitpunkt geäußert haben konnten, selbst wenn sie nichts anderes getan hätten als zu reden. "Bitte reiche mir das Salz" kommt wohl recht häufig zum Zug, aber es ist extrem unwahrscheinlich, (167) daß zum Beispiel ein zufällig ausgewählter Satz dieses Buches schon einmal niedergeschrieben oder sonstwie von jemand anderem geäußert worden ist. ...

(208) Eine von Kants epochemachenden Einsichten, die Frege und Wittgenstein für uns bestätigt und gesichert haben, besteht in seiner Anerkennung des *Vorrangs des Propositionalen*. Für die vorkantische Tradition stand fest, daß jede ordnungsgemäße semantische Erklärung mit einer Lehre von den Begriffen oder (singulären bzw. generellen) Termini anzufangen habe, deren Bedeutungshaltigkeit unabhängig von und vor der Bedeutungshaltigkeit von Urteilen erfaßt werden könne. Ausgehend von dieser grundlegenden Interpretationslehre erklärt dann eine Lehre von den Urteilen die Kombination von Begriffen zu Urteilen und wie die Richtigkeit der sich ergebenden Urteile davon abhängt, was wie kombiniert wurde. Gestützt auf diese abgeleitete Interpretation der Urteile erklärt schließlich eine Lehre der Konsequenzen die Kombination von Urteilen zu Inferenzen und wie die Richtigkeit der Inferenzen davon abhängt, was wie kombiniert wurde.

Kant lehnt das ab. Eine seiner tiefgreifendsten Neuerungen besteht in der Behauptung, daß die fundamentale Einheit des Bewußtseins oder der Erkenntnis, das kleinste Begreifbare, das *Urteil* sei. Die Urteile sind deswegen grundlegend, weil sie die kleinste Einheit darstellen, für die man auf der kognitiven Seite *Verantwortung* übernehmen kann, ebenso wie Handlungen die entsprechende Einheit der Verantwortung auf der praktischen Seite bilden. ...

(209) Daraus ist die Lehre zu ziehen, daß die Relation zwischen Bezeichnung und Bezeichnetem nur als ein Aspekt des Urteilens oder Behauptens verstanden werden kann, *daß* etwas (was durch einen Aussagesatz und nicht durch einen singulären Terminus oder ein Prädikat allein ausgedrückt wird) so der Fall, d.h. *wahr* ist. Damit urteilt, glaubt oder behauptet man, *daß* eine Proposition oder eine Behauptung wahr ist (eine Tatsache ausdrückt oder feststellt), *daß* etwas *die Wahrheit über* einen Gegenstand oder eine Reihe von Gegenständen ist, *daß* ein Prädikat *auf* etwas anderes *zutrifft*. Folglich muß man sich sowohl darum kümmern, was gesagt oder ausgedrückt wird, als auch darum, *worüber* damit gesprochen wird, *von* wessen Wahrheit die Rede ist. Wichtig ist also sowohl der Gedanke als auch das, *wovon* der Gedanke handelt.

zu Ziff. 3.2241:

Satz - Äußerung

zu:

aus: J. Lyons, Linguistic Semantics. An introduction. Cambridge 1995 (reprint 1997)

(37f)

I am taking for granted, for the time being, the reader's ability to identify the sentences of any language in which he or she is competent: i.e., to distinguish them from other combinations of words that are not sentences. I will now make the further assumption that some are non-sentences because they are grammatically incorrect and others because they are grammatically incomplete, or elliptical; and that, once again, those who are competent in the language, whether they are native speakers or not, can identify these two subclasses of non-sentences. As we shall see later, many everyday utterances are grammatically incomplete, but, in context, both acceptable and interpretable. ...

First, sentence-meaning is (to a high degree) **context-independent**, whereas utterance-meaning is not: that is to say, the meaning of an utterance is (to a greater or less degree) determined by the context in which it is uttered...

(135f) There are other, actual or potential, utterances which we can classify, no less readily, as grammatical, but meaningless. Among them, we can list, with their authors, such famous examples as

(7) Colourless green ideas sleep furiously (Noam Chomsky)

(8) Quadruplicity drinks procrastination (Bertrand Russell)

(9) Thursday is in bed with Friday (Gilbert Ryle).

Of course, none of these is uninterpretable, if it is appropriately contextualized and the meaning of one or more of its component expressions is extended beyond its normal, or literal, lexical meaning by means of such traditionally recognized rhetorical principles as metaphor, metonymy of synecdoche...

(138) Here I am concerned to emphasize, first, that meaningfulness, or semantic well-formedness (if we use that term and, for the present at least, accept what it implies), is readily distinguishable, in clear cases, from grammaticality, and, second, that not every utterance which is judged to be unacceptable on the grounds that it does not make sense is properly regarded as semantically ill-formed.

zu Ziff. 3.2242:

Äußerung = öffentlich verantwortete Rede

zu: vgt2.201 vprag6.0641

aus: SOTTONG, H; MÜLLER, M, Zwischen Sender und Empfänger. Eine Einführung in die Semiotik der Kommunikationsanalyse. Berlin 1998.

(33f) Wenn wir uns die Elemente der Semiosphäre nochmals vor Augen halten, so fällt sogleich auf, daß es sich bei all diesen Elementen um Hervorbringungen von Menschen handelt - um Hervorbringungen, die elementar sind für die Vermittlung, Verständigung und Organisation menschlicher Gesellschaften und Kulturen. Daß wir derart unterschiedliche Phänomene als Kommunikationsakte bzw. Äußerungen einstufen, beruht auf zwei Merkmalen, die diese Phänomene von anderen Dingen, die wir wahrnehmen, unterscheiden: Wir unterstellen erstens, daß es eine Person oder Instanz gibt, die dieses Phänomen produziert hat, und zweitens, daß uns durch dieses Phänomen irgend etwas mitgeteilt wird, daß es also eine "Botschaft" übermittelt. Die unterstellten Produzenten von Äußerungen können sowohl konkrete Personen sein - der Autofahrer, der sich an die Stirn tippt oder der Wirt, der die Speisekarte verfaßt hat - als auch Institutionen oder juristische Personen: der Gesetzgeber, der die Straßenverkehrsordnung beschlossen hat und die durch die Verkehrsschilder kommuniziert, oder die Bank, die als Unternehmen sich dafür entschieden hat, ihr Gebäude so und nicht anders zu gestalten. Der Produzent einer Äußerung ist also immer diejenige Instanz, von der wir annehmen, daß sie die Äußerung verantwortet. Dies muß nicht unbedingt der tatsächliche Urheber der Inhalte sein, die die Äußerung übermittelt. So nehmen wir zum Beispiel an den Produzenten einer Politikerrede den betreffenden Politiker an, auch wenn die Rede in Wirklichkeit ein Referent für ihn geschrieben hat - aber der Politiker ist derjenige, der diese Äußerung in der Semiosphäre verantwortet. Innerhalb eines Kommunikationszusammenhangs oder -mediums können auch verschiedene Produzenten angenommen werden. So wird man bei nicht namentlich gezeichneten Artikeln in einer Zeitung oder Zeitschrift "die Zeitung" oder "die Redaktion" als kollektive Produzenteninstanz betrachten, beim mit einem Namen versehenen Leitartikel dagegen den Träger dieses Namens. Es ist also in vielen Fällen nicht wesentlich, wer nun beispielsweise tatsächlich einen bestimmten Text geschrieben hat, sondern als wessen Text er in der Kultur, in der er geäußert wird, gilt. Ob wir annehmen, daß es für ein Phänomen einen Produzenten gibt, hängt sehr stark davon ab, was die jeweilige Kultur, in der dieses Phänomen auftaucht, für Realität hält. So könnte etwa eine Kultur - vielleicht trifft dies auf die vorklassische griechische Antike zu - für das Phänomen "Blitz" eine Produzenteninstanz, nämlich die Götter, annehmen; wenn diese Kultur dann auch noch annimmt, daß ihr durch das Schleudern von Blitzen durch die Götter etwas mitgeteilt werde, dann wäre der Blitz in dieser Kultur eine Äußerung. Ebenso ist es kulturrelativ, wer als Produzent einer Äußerung angenommen wird. Während wir denjenigen, den wir sprechen sehen, als den Produzenten einer sprachlichen Äußerung betrachten, auch wenn er äußerst wirres Zeug redet, nahm die mittelalterliche Kultur in einem solchen Fall wahlweise an, aus diesem Sprecher spreche der Teufel, oder aber Gott. Im ersten Fall reagierte sie mit Exorzismus, im zweiten mit der Kanonisierung des betreffenden Individuums.

(40) *Äußerungen*

Die Semiosphäre besteht aus miteinander vernetzten und interferierenden Äußerungen.

Äußerungen sind Kommunikationsakte, gleich welcher Art.

Der Unterschied zwischen Äußerungen und anderen Phänomenen liegt in den Kriterien der Wahl und der Arbitrarität.

Das Kriterium der Wahl bedeutet, daß der Produzent einer Äußerung die Wahl hat, eine Äußerung zu produzieren oder nicht, oder eine ganz andere Äußerung zu produzieren.

Das Kriterium der Arbitrarität bedeutet, daß die Form einer Äußerung in keinem ursächlich-notwendigen Zusammenhang mit ihrer Botschaft bzw. Inhalt steht. Die Formen von Äußerungen sind für jede Kultur frei wählbar.

Äußerungen lassen sich von anderen Äußerungen durch die Kriterien der Kohärenz des Produzenten, des Kontextes und des Adressaten abgrenzen.

zu Ziff. 3.2243:

Gegenstände der Pragmatik
Satz/Äußerung - pragmatische Kompetenz

zu: vprag1.01 vgt4.222

aus: MEIBAUER, J, Pragmatik. Eine Einführung. Tübingen 1999: Stauffenburg.

(1) Als am 10. Juli 1997 der damals amtierende Bundespräsident Roman Herzog von Jugendlichen in Kronach gefragt wurde, ob die Rechtschreibreform zurückgenommen werden sollte, antwortete er:

(1) Ich habe mich nie mit der Rechtschreibreform befaßt. Ich befaße mich nur mit wichtigen Dingen. (Schwäbisches Tagblatt, 11.7.97, S.2)

Wie die Jugendlichen auf diese Antwort reagiert haben, ist mir nicht bekannt. Es würde mich jedoch nicht wundern, wenn sie enttäuscht gewesen wären. Der Bundespräsident gibt ihnen ja deutlich zu verstehen, daß er etwas für unwichtig hält, was sie selbst für wichtig halten. Diese Meinung des Bundespräsidenten "Die Rechtschreibreform ist unwichtig" ist allerdings nicht wörtlich ausgedrückt, sie muß erschlossen werden. Wir nennen eine solche erschlossene Gesprächsandeutung eine **Implikatur**.

Durch seine Äußerung hat der Bundespräsident auch eine Behauptung aufgestellt. Würde man ihm nachweisen können, daß er sich in Wirklichkeit doch schon mit der Rechtschreibreform befaßt hat, könnte man ihn der Lüge überführen. Wir sagen, der Bundespräsident hat den **Sprechakt** der Behauptung ausgeführt.

Stellen Sie sich nun vor, Sie selbst hätten die Äußerung unter (1) gemacht. Der Ausdruck *ich* würde sich dann nicht auf den Bundespräsidenten Roman Herzog, sondern auf Sie beziehen. Der Ausdruck *ich* erlaubt diese Möglichkeit prinzipiell. Um zu wissen, worauf sich *ich* bezieht, muß man wissen, von wem die Äußerung stammt. Das Phänomen, daß sich ein Ausdruck wie *ich* auf verschiedene Personen beziehen kann, nennt man **Deixis**....

Andere Forscher nehmen noch die Gebiete der Präsupposition und der Konversationsanalyse hinzu. Unter einer **Präsupposition** ist eine Gesprächsvoraussetzung zu verstehen. Zum Beispiel setzt der Bundespräsident voraus, daß es eine bestimmte Rechtschreibreform gibt. Die Äußerung des Bundespräsidenten gilt auch als Antwort auf eine Frage. Solche Handlungsabläufe wie Frage - Antwort, die ein Gespräch strukturieren, werden von der **Konversationsanalyse** untersucht...

(8) Die pragmatische Kompetenz sagt uns, wie wir eine Grammatik in der Kommunikation zur Anwendung bringen, sei es in der Produktion oder der Interpretation von Äußerungen.

Dank unserer pragmatischen Kompetenz erkennen wir zum Beispiel, daß der Bundespräsident Herzog mit seiner Äußerung (1) zu verstehen gegeben hat, daß er die Rechtschreibreform für unwichtig hält.

Unsere grammatische Kompetenz sagt uns dagegen, daß die Gebilde unter (1) Sätze sind. **Sätze** sind Einheiten der Grammatik; sie sind durch die Regeln der Grammatik bestimmt und dadurch gewissermaßen abstrakte Objekte. In einer konkreten Kommunikationssituation werden Sätze ausgesprochen ("geäußert"), und zwar durch Sprecher, die damit etwas mitteilen wollen. Geäußerte Sätze bezeichnet man als **Äußerungen**. ...

Wenn eine Äußerung gemacht wird, geschieht das in einer bestimmten Äußerungssituation. All das, was in einer Äußerungssituation für die menschliche Verständigung relevant ist, hat man auch als **Kontext** bezeichnet. Dazu werden im allgemeinen Sprecher, Hörer, Zeitpunkt und Ort gerechnet. Bei unserem Beispiel ist Roman Herzog der Sprecher, die Jugendlichen sind die Hörer, der Ort ist Kronach, der Zeitpunkt ist der 10.7.1997. Stellen wir uns weiter vor, wir hätten zusätzlich auch noch Journalisten als Zuhörer, dann würden wir die Jugendlichen als **Adressaten** der Äußerung bezeichnen (obgleich entsprechende Statements natürlich oft "mit Blick auf die Presse" gemacht werden.)

zu Ziff. 3.2244:

aphrastisch/phrastisch; mono-/multiprozedural

zu:

aus: EHLICH, K, Der Satz. Beiträge zu einer pragmatischen Rekonstruktion, in: REDDER, A; REHBEIN, J (eds.), Grammatik und mentale Prozesse. Tübingen 1999. S. 51-68.

(64) Es gibt sprachliche Handlungsformen, die in sich vollständig suffizient, *selbstsuffizient* sind (s.Ehlich 1992). Ich nenne sie die *monoprozeduralen* sprachlichen Handlungsformen. Bei ihnen werden einzelne Prozeduren isoliert von den Interaktanten in einer solchen Weise eingesetzt, daß sie ihren jeweiligen Handlungszweck vollständig in dieser Isolation erreichen. Dies ist insbesondere der Fall bei der Nutzung der größten Gruppe der expeditiven Prozedur, nämlich bei denen, für die *Interjektionen* eingesetzt werden. Ein für sich geäußertes *ah* ist kommunikativ in sich suffizient, es erfüllt seinen Handlungszweck in sich...

(65) In einer spezifischen Verwendungsweise können auch deiktische Prozeduren monoprozeduralen Charakter tragen. Dies ist zum Beispiel der Fall bei einem isolierten "da". Seine Artikulation erfüllte die volle sprachliche Leistung: Der Fokus der Aufmerksamkeit der Hörer richtet sich auf den vom Sprecher gemeinten Punkt, die Wahrnehmungstätigkeit der Adressaten wird in einer vom Sprecher gewünschten Weise gesteuert.

Von den monoprozeduralen Formen sind die *multiprozeduralen* zu unterscheiden. Diese zerfallen wiederum in zwei Untergruppen, die mit einem präzisen Namen zu kennzeichnen nicht ganz einfach ist. Ich unterscheide sie als *Prozedurenkombination* und *Prozedurenintegration*.

Zu den prozedurenkombinierenden Verfahren gehören - am Beispiel des expeditiven Feldes - der *Imperativ* und der *Vokativ*. In beiden Fällen wird die expeditiv Qualität einem je spezifischen Element aus dem Symbolfeld appliziert, also etwa einem Verb oder einem Eigennamen...

Systematisch von der Prozedurenkombination zu unterscheiden ist die *Prozedurenintegration*. Der prozedurenintegrierende Bereich hat im wesentlichen seine Realisierung in dem, was wir die eigentliche Domäne des Satzes als einer eigenen Form sprachlichen Handelns nennen können. Die *Prozedurenintegration* erhält ihre Struktur aus den spezifischen Strukturentscheidungen, die im jeweiligen Sprachsystem hinsichtlich der Aufgabenverteilung für das Zustandebringen komplexer interaktiver Verständigung vorgenommen wurden.

zu Ziff. 3.2245:

aphrastisch

zu:

aus: HOFFMANN, L, Ellipse und Analepse, in: REDDER, A; REHBEIN, J (eds.), Grammatik und mentale Prozesse. Tübingen 1999. S. 69-90.

(69) Sätze wie die folgenden gelten Linguisten oft als "elliptisch", als unvollständig und damit ergänzungsbedürftig, damit ein grammatisch "wohlgeformter" und verstehbarer Satz entsteht.

- (1) da sin manche → die kumme entweder [] sin angetrunke → oder sie komme frisch vun der arweit [] hawwe irgendwas zu verzehle (...) un daß du die leut net beleidische duscht muscht du a: noch was druff antworte → mit ja und nein gehts das nadierlich net → also muscht [] irgenwas rauslosse. (R. Schmitt, Die Schwellensteher, 239)
- (2) Armut bringt außer Weisheit auch Verdruß und Kühnheit [] außer Ruhm auch bittre Mühn. (B. Brecht, Die Dreigroschenoper, 188)
- (3) Man muß deutlicher werden, das ist [] an Fälle erinnern. (E. Bloch, Spuren, 82)
- (4) Aber weil er so gut geschrieben ist, liest man auch leicht über Wichtiges hinweg. [] Meint, man hätte es verstanden. [] Hat es aber nicht []. (U. Wesel, Juristische Weltkunde, 204)
- (5) Das klingt paradox. [] Ist es aber deshalb nicht, weil die vielverschlungenen Wege von Organisationen die moralischen Hemmungsenergien schwächen... (G. Anders, Die Antiquiertheit des Menschen 1, 246)
- (6) Noch der armseligste Mensch ist fähig, die Schwächen des bedeutendsten [], noch der dümmste, die Denkfehler des Klügsten zu erkennen. (Th. W. Adorno, Minima Moralia, 57)
- (7) Schüler: Sie sin aber fleißig (...)
Lehrerin: [] Bin ich doch immer (Redder, Schulstunden, 10 (vereinfacht))
- (8) Daneben [] der Postwagen mit seinen Rappen, unter deren Hufen der romantische Funke hervorsprüht. (W. Benjamin, Das Passagenwerk 1, 212) ...

Ellipsen sind immer wieder im Blick auf den "vollständigen Satz" - was immer das sei (vgl. Hoffmann 1997, Kap. C 4.3) - und selten konsequent funktional betrachtet worden. Als Norm galt (und gilt) in der Grammatiktradition der Verbalsatz, zu dem eine Ellipse ergänzt oder aus dem sie durch Reduktion abzuleiten war. Tatsächlich ist auch im Deutschen der Verbalsatz nur eine (wenngleich frequent) Form neben anderen...

Die bisherige Diskussion um die Ellipse zeigen zwei Schwächen, auf die ich kurz eingehen möchte.

Schwäche 1: Ellipsen gelten als "unvollständige Sätze"...

(71) Schwäche 2: Ellipsen werden "kontextfrei" betrachtet, nicht im Rahmen eines Sprecher-Hörer-Modells bzw. der Dynamik der Verbalisierung und ihrer Ressourcen.

zu Ziff. 3.2246:

Textsegmentierung

zu:

M. STREERUWITZ, Verführungen. 3. Folge Frauenjahre. stb 2726. Frankfurt/M 1997.

(68f) "Die Großmutter wollte noch mit Helene sprechen. Sie solle zu ihr in die Wohnung kommen. Hinüber. Es ging dann um Geld. Die Telefonrechnung. In der Wohnung hatte es nur einen Anschluß gegeben. Nach der Teilung waren 2 Apparate installiert worden. Sie waren umzuschalten. Die alte Frau hatte gesagt, sie bräuchte ja kaum noch ein Telefon. Das würde schon gehen. So. Und sie hatte die Rechnung bezahlen wollen. Als ihren Beitrag. Jetzt plötzlich wollte sie von Gregor einen Anteil. Er solle zahlen. Aber sie könne ihn nicht erreichen. Dieser Aichenheim sage immer, er werde es ausrichten. Aber Gregor rief nie zurück. Ob sie, Helene. Das verstünde. »Na. Er wird nicht mit dir sprechen wollen«, meinte Helene. Die Schwiegermutter antwortete, es müsse Regelungen geben. So ginge das alles nicht. Und überhaupt. Gregor habe ja wahrscheinlich Gründe dafür, daß er das Haus meide. Von ihr. Seiner Mutter. Da sei er ja nie weggegangen. Da sei er geblieben. Bis Helene aufgetaucht war. Helene antwortete, sie habe diesen Mann nicht erzogen. Verzogen. Doch wohl eher. Da liege doch das Grundübel. In der Erziehung. Im Drill zum Egoisten. Mütter stellten sich eben zu viel vor. In ihren Söhnen. Besonders wenn sie Witwen wären. Der Sohn als Mann-Ersatz. Das könne ja nicht gutgehen. Die Schwiegermutter verwies Helene ihrer Wohnung. Helene sagte, sie hätte ohnehin nie in diese Wohnung gewollt."

zu Ziff. 3.2251:

Textsegmentierung, Äusserungseinheiten

zu:

GRASS, G, Danziger Trilogie. Die Blechtrommel. Katz und Maus. Hundejahre. Darmstadt 1980. [Die Blechtrommel]

Stichwort: ÄE/aphrastischer Stil [Übung/Kontrolle möglich: Textform nur mit Punkten, ohne ,;- usw. zugrundelegen]

(356) "Da fand ich ein mageres gußeisernes Kreuz mit mürben Schnörkeln und verkrusteten Buchstaben wie: Mathilde Kunkel - oder Runkel. Da fand ich - soll ich oder soll ich nicht - im Sand zwischen Disteln und Strandhafer - soll ich - drei oder vier - tellergroße, bröckelnd rostige Metallkränze, die vormals - soll ich - vielleicht Eichenlaub oder Lorbeer dargestellt hatten - soll ich etwa nicht - wog die in der Hand - soll ich etwa doch - zielte - soll ich - das überragende Kreuzende - oder nicht - hatte einen Durchmesser von - soll ich - vielleicht vier Zentimetern - nicht - einen Abstand von zwei Metern befahl ich mir - soll ich - und warf - nicht - daneben - soll ich abermals - zu schief stand das Eisenkreuz - soll ich - Mathilde Kunkel oder hieß sie Runkel - soll ich Runkel, soll ich Kunkel - das war der sechste Wurf und sieben gestand ich mir zu und sollte sechsmal nicht und warf sieben - sollte, hing ihn über - sollte - bekränzte Mathilde - sollte - Lorbeer für Fräulein Kunkel - soll ich? fragte ich die junge Frau Runkel - ja, sagte Mathilde; sie starb sehr früh, im Alter von siebenundzwanzig Jahren und achtundsechzig geboren. Ich aber stand im einundzwanzigsten Lebensjahr, als mir der Wurf beim siebenten Versuch glückte, als ich jenes - »Soll ich, soll ich nicht?« - in ein bewiesenes, bekränztes, gezieltes, gewonnenes »Ich soll!« vereinfachte."

(388f) "Und ich verließ mit dem Schmuck die Wohnung, sah in dem Schmuck die Vorstufe zum, macht mich auf den Weg zum, fuhr zum Hauptbahnhof, weil, dachte ich mir, wenn das klappt, dann, verhandelte lange über und war mir im klaren, daß... aber der Einarmige und der Sachse, den die anderen Assessor nannten, waren sich nur über den Sachwert im klaren, ahnten nicht, wie überreif sie mich fürs Glück machten..."

(389) "...Glück gibt es vielleicht nur ersatzweise, Glück immer Ersatz fürs Glück, das lagert sich ab: Marmorglück, Sandsteinglück, Elbsandstein, Mainsandstein, Deinsandstein, Unsersandstein, Glück Kirchheimer, Glück Grenzheimer. Hartes Glück: Blaubank. Wolkig brüchiges Glück: Alabaster. Widiastahl dringt glücklich in Diabas. Dolomit: grünes Glück. Sanftes Glück: Tuff. Buntes Glück von der Lahn. Poriges Glück: Basalt. Erkaltetes Glück aus der Eifel."

(505f) "Ich bete an. Was? Glas. Was Glas? Weckglas. Was weckt das Glas ein? Weckglas weckt Finger ein. Was Finger? Ringfinger. Wessen Finger? Blond. Wer blond? Mittelgroß. Mißt Mittelgroß einen Meter sechzig? Mittelgroß mißt einen Meter dreißig. Was besonderes? Leberfleck. Wo Fleck? Oberarm Innenseite. Links rechts? Rechts. Ringfinger wo? Links. Verlobt? Ja, doch ledig. Bekenntnis? Reformiert. Unberührt? Unberührt. Geboren wann? Weiß nicht. Wann? Bei Hannover. Wann? Im Dezember. Schütze oder Steinbock? Schütze. Und der Charakter? Ängstlich. Gutwillig? Fleißig, auch schwatzhaft. Besonnen? Sparsam, nüchtern, auch heiter. Schüchtern? Naschhaft, aufrichtig und bigott. Blaß, träumt meistens von Reisen, Menstruation unregelmäßig, träge, leidet gerne

und spricht darüber, selbst einfalllos, passiv, läßt es drauf ankommen, hört gut zu, nickt zustimmend, verschränkt die Arme, senkt beim Sprechen die Lider, schlägt, wenn angesprochen, die Augen groß auf, hellgrau mit braun nahe der Pupille, Ring vom Vorgesetzten geschenkt bekommen, der verheiratet, wollte zuerst nicht annehmen, nahm an, schreckliches Erlebnis, faserig, Satan, viel weiß, verreiste, zog um, kam wieder, konnte nicht ablassen, auch Eifersucht aber unbegründet, Krankheit aber nicht selbst, 'Tod aber nicht selbst, doch, nein, begleitet schon vorher, kann nicht mehr ... Amen? Amen."

(523) "[...] nur Oskar blieb übrig, sah Ameisen zu und wußte: das ist ihr Schatten, der sich vervielfältigt hat und der Süße nachgeht, und alle die Worte: Gebenedeite, Schmerzreiche, Seliggepriesene, Jungfrau der Jungfrauen ... und alle die Gesteine: Basalt, Tuff, Diabas, Nester im Muschelkalk, Alabaster so weich ... und all das zersungene Glas, durchsichtige Glas, hauchdünn geatmete Glas ... und Kolonialwaren: Mehl und Zucker in blauen Pfund- und Halbpfundtüten. Später vier Kater, deren einer Bismarck hieß, die Mauer, die frisch gekälkt werden mußte, ins Sterben verstiegene Polen, auch Sondermeldungen, wenn wer was versenkte, Kartoffeln, die von der Waage polterten, was sich zum Fußende hin verjüngt, Friedhöfe, auf denen ich stand, Fliesen, auf denen ich kniete, Kokosfasern, auf denen ich lag ... alles im Beton Eingestampfte, der Saft der Zwiebeln, der die Tränen zieht, der Ring am Finger und die Kuh, die michleckte ... Fragt Oskar nicht, wer sie ist!"

aus: Günter GRASS, Die Rättin. Darmstadt 1986.

(133) Immer noch nicht können wir euren Ausgang, diese allzu menschliche Dramaturgie begreifen: Ofen auf, Hexe rein, Klappe zu, Hexe tot! Vorhang, Schluß der Vorstellung.

(415) Nie habe ich sie aufgeregter gesehen. Mal auf dem Holm, dann die Lange Brücke rauf runter, im Werftgelände, ums Hafenbecken lief sie auf Kaimauern, sprang schließlich auf einen Poller, suchte Worte, brachte keinen Satz zu Ende. Herrchen, Lieberchen! rief sie. Sollten etwa. Könnte es sein, daß trotz Kälte, Eis, Finsternis. Und obgleich Staubstürme, dieser verdammte Strahlensegen, den selbst wir kaum. Und dennoch, weil damals? Nicht auszudenken. Einige Exemplare womöglich. Schau nur, wie erwartungsvoll wir. Volle Hoffnung. Aber auch ängstlich...

Aus einem Radwanderführer:

"Durch I. auf der Radwegbrücke über die Bahngleise ~ dann rechts in den Radweg einbiegen, der in den **Blumenweg** übergeht ~ rechts auf die **Donaustrasse** ~ geradeaus auf der Radwegbrücke über die Donau ~ danach geradeaus auf die Nebenstraße ~ den Bahngleisen entlang ~ nach der Kläranlage auf der Radwegbrücke die Donau nach links überqueren ~ dann vor ...

zu Ziff. 3.2261:

Äußerungen → Satz

zu: vprag2.02 vprag4.11

Bei Sprache in Texten haben wir es mit lebendiger, tatsächlicher Kommunikation zu tun, also mit Erscheinungsformen der *parole*. Folglich sind die Einheiten, aus denen sich Texte zusammensetzen, zunächst einmal - unverfänglich gesagt: *Äußerungen*.

Satzlinguisten sind an *Sätzen* interessiert. Sie untersuchen weniger die *parole* (die immer die Einbettung von Äußerungen in die Texte berücksichtigen muß), als das *System* der betr. Sprache (*langue*).

Daher ist in Abgrenzung von den Satzlinguisten der Hinweis wichtig: Texte bestehen nicht einfach aus Sätzen. Ein "Mhm", ein "Ja", ein "Oh jeh!" usw. kann man nicht einfach als Satz einstufen. Texte bestehen aus *Äußerungen* (bei uns: **Äußerungseinheit(en) = ÄE(en)**). Vielleicht ist je die Mehrzahl davon satzhaft. Immer bleibt aber ein mehr oder weniger großer Anteil von ÄEen übrig, der nicht-satzhaft ist, der aber je wichtige und beschreibbare Funktionen im Text hat.

LEVINSON (1990,[1983]) 18ff betont die Wichtigkeit, *Satz* und *Äußerung* zu unterscheiden. *Satz* sei eine theoretische Entität innerhalb einer Theorie der Grammatik, *Äußerung* dagegen das Hervorbringen eines Satzes, Satz-Analogs oder Satzfragments in einem wirklichen Kontext.

Demzufolge geht eine pragmatisch orientierte Sprachbeschreibung von den gegebenen *Äußerungen* eines Textes aus. Der Wunsch nach einer erschöpfenden, theoretisch konsistenten Beschreibung/Interpretation führt dazu, die Einzeläußerungen zunehmend vor dem Hintergrund des Sprachsystems, der *langue*, zu verstehen.

D.h. die Pragmatik versucht, hinter den individuellen Besonderheiten des Einzeltextes mehr und mehr *die abstrakte Bedeutungsstruktur* des Textes freizulegen.

Prämisse: Wenn ein Text vorliegt, dann müssen die vielfältigen, oft disparat erscheinenden ÄEen der *parole* verstehbar sein im Rahmen eines theoretischen Gesamtkonzepts (der *langue* zugehörig), das es folglich nur bewußt zu machen gilt. Das führt - in der Regel - zu zwei sich ergänzenden Ansätzen:

- Semantik: Das interne grammatische System, die kognitiven Basisstermini für das Bedeutungsverstehen, werden direkt und unmittelbar auf die einzelnen Äußerungen angewendet.¹⁵ Damit läßt sich alles, was *anscheinend* ein Satz ist, beschreiben. - Aber erstens gibt es "Reste" in Texten, deren Funktion auch beschrieben sein will. Zweitens geben auch Sätze und ihre Elemente verschiedenartigen Anlaß zu kritischer Rückfrage. Drittens steht aus, daß der Text als ganzer beschrieben wird, nicht lediglich als Addition zusammenhängender Atome. Viertens ist der Text in seinem Verhältnis zum Äußerungskontext noch nicht erfaßt. - Wegen dieser Defizite ist mit der Semantik die Frage, was der Text *sagt* - $f(a,b)$ - mit den wei-

¹⁵ Unter Einsatz von Simplifizierungen wie: Verb=Prädikat, Nomen=Dingbeschreibung=Aktanten-fähig.

teren Zusatzkategorien der Basisterminologie - noch längst nicht beantwortet. Ein zweiter Zugang ist erzwungen:

- Pragmatik: Erneuter, modifizierter, wesentlich stärker differenzierter Versuch, den Text mit seinen Äußerungen auf das interne grammatische System, die kognitiven Basistermini hin zu beschreiben und zu interpretieren, nun aber so, dass die genannten Defizite (aus der Semantik) sukzessive abgebaut werden. Daraus ergibt sich die entsprechende Differenzierung:

=Textgrammatik Die semantischen "Reste" werden in ihrer Funktion beschrieben. Die Prädikate (aus der Semantik) werden kritisch auf ihren Realitätsgehalt befragt: strengerer "Prädikat"-Begriff. - Die grammatisch faßbare Verknüpfung der ÄEen im Wortsinn wird beschrieben (aufgrund der Basisterminologie wie auch durch Pronominalisierung). - Auf der Basis dieser neuen Erkenntnisse schließlich: Nach dem Kriterium der Einheit von Ort/Zeit/Akteuren/Isotopie werden Textpassagen als zusammengehörig (=TGE ~ "Szene") definiert, wobei die Teile davon (=ÄEen) ausnahmslos im Dienst der jeweiligen TGE eine definierbare *Textfunktion* wahrnehmen. - Dem Erkenntnisfortschritt gegenüber der Semantik werden nun aber immer noch Info-Lücken, Widersprüchlichkeiten, Paradoxien, "Unsinn" (wenn man nur die Wortbedeutung berücksichtigt) entgegenstehen, die den bislang im Wortsinn beschriebenen Text noch nicht befriedigend im Rahmen des grammatischen Systems verstehen lassen. Das erzwingt eine weitere Beschreibungsanstrengung:

=Textlinguistik Versuch, das Gelesene mittels einer anderen, erschlossenen *gemeinten* Bedeutung konsistenter und umfassender zu verstehen. Das schließt weitere kritische Rückfragen ein: Kritische Analyse dessen, was als "Nomen" seit der Semantik begegnete. Aufdecken von Abstraktionen. Beschreibung und Bestimmung der Funktion von Stilfiguren. Auf der Basis dieser Erkenntnis eine weitere Kontextbeschreibung. Dialog-, Isotopie-, Thema-Rhemabeschreibung. - Auf der Basis dieser vielen neuen Erkenntnisse erneut die Aufgabe: Nach dem Kriterium der Einheit von Ort/Zeit/Akteuren/Isotopie werden Textpassagen als zusammengehörig (=TLE) definiert. Die Einheiten der vorigen Ebene (TGEen) nehmen ausnahmslos *Textfunktionen* bezüglich der betreffenden TLE wahr. Auf Bedeutungsebene wird der Text mehr und mehr als *Hierarchie* sichtbar; *alle* Textteile sind darin einbezogen.

=Textpragmatik Kein Text formuliert alle zu seinem Verständnis notwendigen Informationen (er wäre langweilig bzw. unlesbar), immer ist er beeinflusst von Implikationen und Präsuppositionen, d.h. von Wissen, das sich aus der Kommunikationssituation ergibt bzw. dort vorausgesetzt ist. Wie ist die Relation zwischen dem - inzwischen - aufwändig beschriebenen Text und der erschließbaren/anzunehmenden Kommunikationssituation (mit ihren Komponenten) zu beschreiben? Hilft letztere, um immer noch bestehende Unklarheiten auszuräumen? - Gesamttext = TPE = Textpragmatische Einheit, für den die Einheiten der Ebene zuvor ausnahmslos *Textfunktionen* wahrnehmen.

Die Bedeutungsanalyse strebt also ab der Semantik eine theoretisch schlüssige Gesamtinterpretation an. Der erste naive Zugang der Semantik führt dazu, dass der/die InterpretIn sich an vielen Stellen in einem Text "die Finger verbrennt", d.h. dass die Sprachbeschreibung scheitert. Die *Äußerungen* leisten einer theoretisch-systemhaften Sicht großen Widerstand, weil sie vielfältig den theoretischen Erfordernissen einer kompletten und verstehbaren Mitteilung *nicht* entsprechen. Dieser Widerstand wird auf den nachfolgenden pragmatischen Schritten sukzessiv abgebaut - durch Schärfstellung der "Linse". Das Ziel ist erreicht, wenn am Schluß eine konsistente, also plausible, und *alle* Elemente des Textes einschließende Interpretation gefunden wurde.

zu Ziff. 3.2262:

Textsegmentierung, Äusserungseinheiten

zu:

aus: H. SCHWEIZER, Textsegmentierung in Äußerungseinheiten: Sprache und Datenverarbeitung 18 (1994) 3-18.

Kriterien für Äußerungseinheiten⁶

Eine Äußerungseinheit (=ÄE) ist, was eine kommunikativ/inhaltlich beschreibbare, in sich abgerundete Funktion erfüllt.¹⁷ Nicht bloße Sachverhaltslogik/-semantik ist somit die Beurteilungsbasis - wie meist in der langen Geschichte versuchter Satzdefinitionen -, sondern die Frage der aktuellen Sprachverwendung, die Einbeziehung des Sprechers und seiner über die Formulierungen faßbar werdenden Interessen und Intentionen.¹⁸

Zugleich ist eine ÄE immer durch die Grenzen der umgebenden ÄEen mitdefiniert. Wichtiger pragmatischer Gesichtspunkt: Bei der Segmentierung eines Textes in ÄEen wird der Text sequentiell (linear) und entsprechend der unumkehrbaren Leserichtung (Textanfang → Textende) betrachtet. Immer wenn in dieser Bewegung - entsprechend den folgenden Kriterien - ein Einschnitt notwendig ist, beginnt eine neue ÄE. Eine ÄE kann in zwei Formen auftreten:

- satzhaft (phrastisch)
- nicht-satzhaft (aphrastisch)

¹⁶ Die Kriterien entspringen nicht positivistischen Setzungen, sondern sind Kristallisationen des Sprachverständnisses, wie es in SCHWEIZER, H, Metaphorische Grammatik. Wege zur Integration von Grammatik und Textinterpretation in der Exegese. ATS 15. St.Ottilien 1981.²1990. bzw. SCHWEIZER, H, Biblische Texte verstehen. Arbeitsbuch zur Hermeneutik und Methodik der Bibelinterpretation. Stuttgart 1986, erläutert ist.

¹⁷ Unter "Funktion" ist hier der Bereich "Sprechakt/Illokution" zu verstehen: Inwiefern gibt ein Sprecher zu erkennen, daß er eine neue Sprechhandlung beginnt, mit einer veränderten Einstellung zu Sachverhalt (das kann auch heißen: er redet von einem neuen Svh) und/oder Adressaten? Stellt er Fakten nüchtern dar? Versucht er, zum Handeln zu animieren? Drückt er eigene Gefühle aus? Sorgt er sich nur um den Sprechkontakt, die Aufmerksamkeit der Kommunikationspartner? Beginnt er einen Gedanken, bricht ihn aber vorzeitig ab zugunsten zunächst eines anderen, komplett formulierten?

¹⁸ Daher liegt es auf der Linie verschiedener in diesem Sinn vorgehender linguistischer Werke, wenn - dies als weitere Illustration - von der "langue"-Orientierung ab - und zur "parole"-Orientierung übergegangen wird. Auch wird nicht mehr nur von "Satz" sondern eben von "Äußerung" gesprochen - was immer diese speziell beinhaltet. Daher wurde auch als englischer Terminus "Illocution Unit" (=IU) vorgeschlagen, vgl. SCHWEIZER, H, The Predication-Model as a component of a Semantic and Pragmatic Content-Analysis: "Bible et Informatique: méthodes, outils, résultats", Jérusalem, 9-13 Juin 1988. Genève 1989. 537-562.

Der so aufgeteilte Text wird neu geschrieben und durchgezählt.¹⁹ - Es fungieren - wenn die Numerierung per Programm vollzogen wird - als Steuerzeichen: | für Kapitelanfang, @ für Anfang eines grafischen Abschnitts, # Beginn einer ÄE. Dabei besteht ein Inklusionsverhältnis:

(Kapitel (Abschnitt (ÄE))).²⁰

Vorschlag für die Syntax der Zählung: [Kapitel] und [Abschnitt] werden - durch Komma getrennt - numerisch gezählt, die einzelne ÄE alphabetisch angefügt.²¹ Es können auch andere Zählweisen entwickelt werden.

- (1) Ein Satz hat nur ein finites, d.h. konjugiertes Verb.²² - Das Kriterium gilt folglich für morphologisch faßbare konjugierte Verbformen. Letztere werden nicht - wie es oft geschieht - austauschbar mit dem semantischen Terminus "Prädikat" belegt.
- (2) Immer wenn - nun semantisch gedacht - eine "Prädikation" anzunehmen ist (eine Subjekt-Prädikat-, also im Kern eine zweipolige Struktur, s.u. Kriterium 13), wird eine eigene ÄE angenommen. In vielen Sprachen kann eine einen selbständigen Sinn ergebende Äußerung (= Satz) nominal, verblos vorliegen. Vgl. Reporterdeutsch: "der Ball von Müller zu Beckenbauer, ...".²³
- (3) Nach einer Redeeinleitung: eine neue ÄE beginnt. Die Formen des Übergangs von besprechender zu direkter Rede können sehr verschieden sein.²⁴
- (4) Relativsätze werden abgetrennt. Im Fall einer Einbettung die

¹⁹ Es kann dann, wenn die ÄEen zunächst am Endtext bestimmt wurden, nach literarkritischem Ausscheiden von Zusätzen an einer früheren synchronen Textschicht weitergearbeitet wird, zu einer Spannung kommen: Was beim Endtext getrennt werden mußte, kann bei der früheren Fassung zusammengehören. Beispiel: Gen 43,15a+c (vgl. SCHWEIZER, H, Josefsgeschichte. Konstituierung des Textes Teil I: Argumentation. THLI 4/1. 360 S. Teil II: Textband. THLI 4/2. 203 S. Tübingen 1991. THLI 4/2 S.171). Es ist eine arbeitspraktische Entscheidung, ob man mit dieser Zählung weiterarbeitet oder sie anpaßt (was dem Sprachbefund besser entspräche).

²⁰ Vgl. auch Regel (11).

²¹ Da manche Abschnitte sehr lang sind, kann es leicht sein, daß ein Durchgang des Alphabets nicht ausreicht. Daher ist in solchen Fällen neu mit dem Alphabet zu starten, aber zusätzlich anzugeben, in welchem Alphabet-Durchgang man sich befindet, z.B. **15,4h2** = Kapitel 15, Abschnitt 4, ÄE h im 2. alphabetischen Durchgang.

²² Im Hebräischen wird die Verbindung w=/wa= + konjugierte Verbform als sicheres Signal der Satzeröffnung genommen.

²³ Oder: **Gerhard Schröder** (SPD), Ex-Bundeskanzler, nach dem Rückzug Franz Münteferings von dessen politischen Ämtern: ARBEIT GUT, HILFE GROSS, MENSCH KLASSE [3 nominale, satzhafte ÄEen].

²⁴ u.U. ist der Übergang überhaupt nicht markiert, aber klar aus dem Kontext erschließbar (vorher dritte Person, jetzt erste oder zweite; Sprecherwechsel). Satzzeichen liefern Steuerungssignale (die gibt es nur nicht in allen Schriften). Ausformulierte Redeeinleitung ("er sprach:"). Oft treffen mehrere solcher Indizien zusammen.

Einzelelemente durchzählen in der geschilderten Weise. Keine Sondernotation!²⁵

- (5) Konjunktionale Nebensätze (logische Konjunktionen: "daß, damit, weil, obwohl, wenn" usw.) werden abgetrennt - seien sie nominal oder verbal. Vor die Konjunktion können noch *koordinierende* Wörter treten ("und weil...", oder daß...").
- (6) Verblöse ÄEn mit benennbarer eigener kommunikativer Funktion (z.B. als kurze Antwort auf eine Frage) werden abgetrennt.
- (7) Erkennbar parallelisierte Gedanken (z.B. in Poesie) werden abgetrennt; vgl. Parallelismus membrorum ("Fallen auch tausend zu deiner Rechten / Zehntausend zu deiner Linken").²⁶
- (8) Infinitiv-Konstruktionen ("um zu..." u.ä.) werden nicht abgetrennt, weil bei ebenfalls vorhandenem konjugiertem Verb der vom Autor gewollte Satzakzent auf diesem liegt und nicht gleich stark auf dem Infinitiv.
- (9) Setzung eines neuen Themas: es wird (oft aphrastisch) eine inhaltliche Wendemarke im Text, auf die erst anschließend wieder ein vollständiger Satz folgt, gesetzt ("aber bezüglich des Mose, so wissen wir ..."). Es können aber auch einzelne Glieder eines Satzes zur Betonung nicht-satzhaft vorangestellt werden.
- (10) Vokativ (u.U. eingebettet - s.o. (3)), Aufmerksamkeit-erregende Partikeln, Interjektionen haben aufgrund ihrer eigenständigen, nämlich die Kommunikation steuernden/sichernden Funktion als eigene ÄE zu gelten.
- (11) Jedes Glied der Prädikation kann zusätzlich beschrieben werden (vgl. Genitivverbindung, Apposition u.ä.; hier zusammengekommen unter: Adjunktion). "Zusatzbeschreibung" hat Affinität zu unmittelbarem Anschluß bzw. Kontaktstellung. Ist diese Verbindung von beschriebenem (=Signifikat) und beschreibendem (=Signifikant) Element aufgesprengt, so wächst die Wahrscheinlichkeit, daß der Autor die Zusatzbeschreibung als neue ÄE versteht. Dieses Kriterium läßt sich weiter präzisieren. Es sind folgende Fälle denkbar, wenn eine koordinierende Wortform vorliegt:
 - a. Die Koordination bezieht sich auf das unmittelbar vorhergehende Nomen. Hierbei besteht kein Grund zur Abtrennung ("Einigkeit und Recht und Freiheit").
 - b. Die Koordination bezieht sich auf ein weiter vorne liegendes Nomen, das inzwischen durch eine weitere Adjunktion (z.B. Genitiv) beschrieben war. Es besteht somit nicht nur

²⁵ "Der Hund, der grau war, ist tot." = a:Der Hund b:der grau war c:ist tot. Nicht: a:Der Hund b:(oder al:)der grau war a: ist tot. - Begründung: (1) Segmentierungen werden auf der Basis linearer Textwahrnehmung allein bestimmt (also nicht aufgrund der Erkenntnis von Einbettungen). (2) Ist (1) eingehalten, so bereitet es später bei der Analyse kein Problem, die vielfältigen Verschachtelungen (eben nicht nur bei Relativsätzen) in Texten herauszuarbeiten.

²⁶ "Parallelisierte Gedanken" sind mehr als nur die beschreibende Erweiterung etwa eines Nomens (s.u. Regel 10)! - Die Leserichtung (vorn→hinten) ist einzuhalten. Liest man bis Gen 50,20c - vgl. THLI 4/2 S.181 - käme man nicht auf den Gedanken, 20b+c zu trennen (Infinitive werden nicht abgetrennt). Erst nach der Lektüre von 20d entsteht rückwirkend der Wunsch, die Parallelität von 20c/20d herauszustellen. Ein derartiges Argumentieren entgegen der Leserichtung ist aber zu vermeiden.

eine lineare Abfolge von Koordinationen, sondern zusätzlich auch eine Hierarchie. Aber das Erstnomen und das durch Koordination angeschlossene Nomen sind in gleicher Funktion belegt, daher kein Grund zur Abtrennung. Beispiel: "Er kaufte Reis aus China und Honig aus Kanada."

- c. Der Koordinationsbezugspunkt ist nicht mit den Fällen a. und b. vergleichbar, sondern zielt auf die Ebene des Prädikats. Folgt auf Koordination ein finites Verb (z.B. "Er kam und sah.") wird ohnehin abgetrennt. Es gibt aber auch Fälle, in denen dieses zweite finite Verb ausbleibt und trotzdem klar ist, daß nicht lediglich eine nominale Wortgruppe gebildet werden soll, sondern eine neue Äußerungseinheit, z.B. "Er kaufte gestern den Reis und das Brot heute."²⁷
 - (12) Bei Texten, die eine verbürgte optische Gliederung (z.B. Abschnitte, bewußter Zeilenwechsel) aufweisen (in der Regel nicht bei antiken Handschriften), wird auch diese als vom Autor gewollte Textsegmentierung berücksichtigt:²⁸ Neuer optischer Abschnitt = neue ÄE.²⁹ - Analoges gilt für die Satzzeichen: Schwache Trenner (Komma) können nicht als Kriterium genommen werden, sehr wohl jedoch **;;?!"»** u.ä.
 - (13) Während eine Prädikation - also bei einer phrastischen ÄE - in ihrer Basis immer mindestens als 2-stellige Relation zu sehen ist,³⁰ ist eine aphrastische ÄE immer ein nominales Element ohne auch nur eine Relation aus dem Prädikationsbereich.³¹ Dieses eine Element kann in sich sehr komplex sein.³² Liegen somit irgendwo zwei derartige Sprachelemente vor, die - außerhalb der Prädikatsthematik - nicht in gewohn-
-
- ²⁷ In solchen Fällen hilft zusätzlich die Beobachtung der weiteren Elemente, im gen. Beispiel die inhaltliche Opposition und chiastische Stellung der Zeitangaben. Es wirkt also Kriterium (7) herein. Eine solche Argumentation ist immer textnäher als die bloße Behauptung einer "Tilgung" des Verbs.
- ²⁸ Optische Gliederungen bzw. Gliederungszeichen (z.B. Petucha und Setuma im Hebräischen), die lediglich auf nachträgliche Herausgeber des Textes zurückgehen, zählen nicht zu diesem Kriterium.
- ²⁹ Das gilt auch dann, wenn nur von den mitgeteilten Inhalten her sich eine ÄE-Grenze nicht aufdrängen würde. Vgl. auch die Zeilengliederung moderner Gedichte bzw. das Stilmittel "Enjambement".
- ³⁰ [1.Aktant=Inhalt a]←[Prädikat=Relationstyp]→[Inhalt b], das heißt bei dem Satz: "Ich ging": [1.Aktant=ich]←[Prädikat=dynamisch-initiativ]→[<<GEHEN>>]. - Verben mit obligatem 2. oder gar noch 3.Aktanten sind entsprechend mehrstellig. Es sollte hierbei jedoch nicht einfach hochgezählt werden (also ein traditionell als "transitiv" bezeichnetes Prädikat nun: 3-stellig, eines auch mit "indirektem Objekt" womöglich als 4-stellig). Denn die Valenzen der "Basis der Prädikation" sind den weiteren Aktanten hierarchisch vorgeordnet. - Orts- und Zeitangaben werden außerhalb dieser Relationen (=Valenzen) geführt.
- ³¹ Andere Relationen, die sich in weiterem Kontext von satzhaften ÄEn finden, sind möglich. z.B. eine "Kausal"-Verknüpfung: "denn unser Bruder" (Gen 37,27d).
- ³² Emphatisch verstärkt (Gen 37,13f), durch Adjunktionen unterschiedlichster Art erweitert (Gen 37,25g; 41,3c.10c; 42,15c).

ter Weise aufeinander bezogen werden können, so handelt es sich um zwei aphrastische ÄEn.³³

Diese Kriterien beanspruchen nicht vollständig zu sein. Auch haben nicht alle das gleiche Gewicht. In der Regel ergänzen sich mehrere Kriterien und führen zu einer Entscheidung. Unsicherheiten ergeben sich öfters bei der Frage der Parallelisierung und der Zusatzbeschreibung.³⁴

Die Kriterien sind weiterentwickelt im Gefolge von

SCHWEIZER, H, Biblische Texte verstehen. Arbeitsbuch zur Hermeneutik und Methodik der Bibelinterpretation. Stuttgart 1986. S.38f.

Fortschreibung des Textes im Internet:

<http://www-ct.informatik.uni-tuebingen.de/ct/ae.html>

zu Ziff. 3.2271:

zu: vgt2.251

Hermann Hesse, Legenden. stb 909. Frankfurt 1983. S. 127-139.

Die Belagerung von Kremna

Zur Zeit der Kaiser Aurelian, Tacitus und Probus hatte in den seit Jahrhunderten als Räubernester verrufenen kleinasiatischen Provinzen Isaurien, Pisidien und Lykien ein gewisser Lydius einen gefürchteten Namen. Er war ein Isaurier und unter Philippus Arabs geboren. Und fast alle seine Vorfahren waren Räuber gewesen. Sein Vater war auf einem Beutezug in Lykien umgekommen, sein Großvater und zwei Oheime am selben Tage am Galgen gestorben. Sein ursprünglicher Name ist unbekannt; von seinem zwanzigsten Jahr an nannte er sich Lydius und machte sich unter diesem Namen in jenen Ländern berühmt. Lydius war von Natur ein kluger und besonnener Mensch, kühn, doch maßvoll in seinen Unternehmungen. Er wußte die Menschen zu benutzen und ihre Liebe oder Furcht seinen Absichten dienstbar zu machen. So klomm er rasch von Erfolg zu Erfolg und kostete schon als Jüngling Ruhm und Macht, ohne dessen müde oder gesättigt zu werden. Erst in der Zeit nach seinem dreißigsten Jahre, als ihm immer kühnere Handstreich wie Wunder glückten, begann ihn der Rausch und Übermut der Unbesiegbaren zu erfassen, so daß er die von den Göttern gesetzten Grenzen überschritt und endlich jäh zu Fall kam. Bei einem Zuge durch Kilikien, den Lydius mit seiner zahlreichen Schar unternahm, stieß einst ein ionischer Grieche namens Hephaistion zu ihm, der bisher den kilikischen Seeräubern Dienste getan hatte, nun aber es vorzog, dieser berühmten Schar zu folgen. Von da an vermochte Lydius immer größere Dinge zu unternehmen, die ihm alle gelangen, denn dieser Hephaistion war ein [...]

³³ Das Problem taucht z.B. bei doppelten Vokativen auf: "JAKOB, JAKOB!" (vgl. Gen 46,2cd), wo das zweite Element eben nicht eine Adjunktion zum ersten ist, sondern eine gleichwertige Wiederholung. Anders bei Vokativ mit Adjunktion: "a:Aber, b:Herr Vorsitzender, c:ich behaupte..."

³⁴ Für die jeweiligen Sprachen lassen sich recht viele dieser Kriterien im Programm erfassen, so daß viele Standardsituationen automatisch analysiert werden. Diese Grob-Analyse verlangt aber eine Prüfung und Ergänzung um die Analyse der Fälle, die differenzierter zu beurteilen sind.

<http://www-ct.informatik.uni-tuebingen.de/ct/ae.html>

Hermann Hesse, // Legenden. // stb 909. // Frankfurt 1983. // S. 127-139. //

Die Belagerung von Kremna //

Zur Zeit der Kaiser Aurelian, Tacitus und Probus hatte in den seit Jahrhunderten als Räubernester verrufenen kleinasiatischen Provinzen Isaurien, Pisidien und Lykien ein gewisser Lydius einen gefürchteten Namen. // Er war ein Isaurier // und unter Philippus Arabs geboren. // Und fast alle seine Vorfahren waren Räuber gewesen. // Sein Vater war auf einem Beutezug in Lykien umgekommen, // sein Großvater und zwei Oheime am selben Tage am Galgen gestorben. // Sein ursprünglicher Name ist unbekannt; // von seinem zwanzigsten Jahr an nannte er sich Lydius // und machte sich unter diesem Namen in jenen Ländern berühmt. // Lydius war von Natur ein kluger und besonnener Mensch, kühn, doch maßvoll in seinen Unternehmungen. // Er wußte die Menschen zu benutzen und ihre Liebe oder Furcht seinen Absichten dienstbar zu machen. // So klomm er rasch von Erfolg zu Erfolg // und kostete schon als Jüngling Ruhm und Macht, // ohne dessen müde oder gesättigt zu werden. // Erst in der Zeit nach seinem dreißigsten Jahre, // als ihm immer kühnere Handstreich wie Wunder glückten, // begann ihn der Rausch und Übermut der Unbesiegbaren zu erfassen, // so daß er die von den Göttern gesetzten Grenzen überschritt // und endlich jäh zu Fall kam. // Bei einem Zuge durch Kilikien, // den Lydius mit seiner zahlreichen Schar unternahm, // stieß einst ein ionischer Grieche namens Hephaestion zu ihm, // der bisher den kilikischen Seeräubern Dienste getan hatte, // nun aber es vorzog, dieser berühmten Schar zu folgen. // Von da an vermochte Lydius immer größere Dinge zu unternehmen, // die ihm alle gelangen, // denn dieser Hephaestion war ein [...]

zu Ziff. 3.2272:

Segmentierung / Aufbereitung

zu: vgt3.342

AUSGANGSTEXT

Charles Dickens. David Copperfield. Oxford 1989. p.176-183.

CHAPTER XV

I MAKE ANOTHER BEGINNING

MR. DICK and I soon became the best of friends, and very often, when his day's work was done, went out together to fly the great kite. Every day of his life he had a long sitting at the Memorial, which never made the least progress, however hard he labored, for King Charles the First always strayed into it, sooner or later, and then it was thrown aside, and another one begun. The patience and hope with which he bore these perpetual disappointments, the mild perception he had that there was something wrong about King Charles the First, the feeble efforts he made to keep him out, and the certainty with which he came in, and tumbled the Memorial out of all shape, made a deep impression on me. What Mr. Dick supposed would come of the Memorial, if it were completed; where he thought it was to go, or what he thought it was to do; he knew no more than anybody else, I believe. Nor was it at all necessary that he should trouble himself with such questions, for if anything were certain under the sun, it was certain that the Memorial never would be finished.

It was quite an affecting sight, I used to think, to see him with the kite when it was up a great height in the air. What he had told me, in his room,

TRENNSTELLEN EINGESETZT [automatisch]

15,0
|Charles Dickens.
#David Copperfield.
#Oxford
#1989.
#p.176-183.
@CHAPTER XV
@I MAKE ANOTHER BEGINNING
@MR. DICK and I soon became the best of friends,
#and very often,
#when his day's work was done,
#went out together to fly the great kite.
#Every day of his life he had a long sitting at the Memorial,
#which never made the least progress,
#however hard he labored,
#for King Charles the First always strayed into it, sooner or later,
#and then it was thrown aside,

#and another one begun.
 #The patience and hope
 #with which he bore these perpetual disappointments,
 #the mild perception he had
 #that there was something wrong about King Charles the First,
 #the feeble efforts
 #he made to keep him out,
 #and the certainty
 #with which he came in,
 #and tumbled the Memorial out of all shape,
 #made a deep impression on me.
 #What Mr. Dick supposed
 #would come of the Memorial,
 #if it were completed;
 #where he thought
 #it was to go,
 #or what he thought
 #it was to do;
 #he knew no more than anybody else,
 #I believe.
 #Nor was it at all necessary
 #that he should trouble himself with such questions,
 #for
 #if anything were certain under the sun,
 #it was certain
 #that the Memorial never would be finished.
 @It was quite an affecting sight,
 #I used to think,
 #to see him with the kite

[Kontrolle durch verstehenden Sprachbenutzer]

ZÄHLUNG [automatisch]

15,1a : Charles Dickens.
 15,1b : David Copperfield.
 15,1c : Oxford
 15,1d : 1989.
 15,1e : p.176-183.
 15,2a : CHAPTER XV
 15,3a : I MAKE ANOTHER BEGINNING
 15,4a : MR. DICK and I soon became the best of friends,
 15,4b : and very often,
 15,4c : when his day's work was done,
 15,4d : went out together to fly the great kite.
 15,4e : Every day of his life he had a long sitting at the Memori-
 al,
 15,4f : which never made the least progress,
 15,4g : however hard he labored,
 15,4h : for King Charles the First always strayed into it, sooner
 or later,
 15,4i : and then it was thrown aside,
 15,4j : and another one begun.
 15,4k : The patience and hope
 15,4l : with which he bore these perpetual disappointments,
 15,4m : the mild perception he had
 15,4n : that there was something wrong about King Charles the
 First,
 15,4o : the feeble efforts
 15,4p : he made to keep him out,
 15,4q : and the certainty
 15,4r : with which he came in,

15,4s : and tumbled the Memorial out of all shape,
 15,4t : made a deep impression on me.
 15,4u : What Mr. Dick supposed
 15,4v : would come of the Memorial,
 15,4w : if it were completed;
 15,4x : where he thought
 15,4y : it was to go,
 15,4z : or what he thought
 15,4a2: it was to do;
 15,4b2: he knew no more than anybody else,
 15,4c2: I believe.
 15,4d2: Nor was it at all necessary
 15,4e2: that he should trouble himself with such questions,
 15,4f2: for
 15,4g2: if anything were certain under the sun,
 15,4h2: it was certain
 15,4i2: that the Memorial never would be finished.
 15,5a : It was quite an affecting sight,
 15,5b : I used to think,
 15,5c : to see him with the kite

zu Ziff. 3.2273:

Längster Satz im Gesamtwerk Thomas Manns

zu: vprag2.261 vgt3.3611

aus: Th. Mann, Joseph und seine Brüder. Bd.1 Die Geschichten Jaaköbs. Der junge Joseph. Frankfurt/M 1980.

Vgl. Mat. 3.341 (vgt)

- ein Makrosatz bestehend aus vielen einzelnen Äusserungseinheiten (satzhaft oder nicht-satzhaft)
- wodurch entsteht - nicht lediglich über Interpunktion - der Eindruck, dieser eine Satz bilde eine Einheit?
- welche Teilfunktion steuert die einzelne ÄE zur gesamten Sinn-einheit bei?
- Lassen sich die Kategorien der Semantik auch hier anwenden?
- Ist der Gesamtsatz als komplexe hierarchische Struktur darstellbar (mit Relationen und Knoten)?
- Nach Analyse und Beschreibung: Frage nach der Funktion eines solchen Satzes am Beginn eines Romans.

zu Ziff. 3.2274:

Gedicht: Zeile ↔ Strofe ↔ Äusserungseinheit

zu: vprag5.121

- 1 Das Kamel
- 2 Es stand ein Kamel in der Wüste am Nil,
- 3 fraß nur ganz wenig und nur wenig entfiel.
- 4 Es hat schon die Kindheit im Sande verbracht
- 5 und nie an den Wechsel des Wohnorts gedacht.
- 6 Den Leitsatz im Auge - egal was auch kam
- 7 achala machala - dahaaam ist dahaam.
- 8 Sein bescheidenes Dasein ist dadurch erklärlich:
- 9 Vitamine gibt's ganzjährig relativ spärlich,
- 10 nicht wie bei glücklichen Kühen im mannshohen Grase -
- 11 hier muß man schon froh sein an einer Oase.
- 12 Auch Steine müssen hier als Seltenheit gelten,
- 13 weil Pharaonen aus ihnen Pyramiden erstellten,
- 14 Die Gegend ansonsten - bis hinunter nach Theben -
- 15 ist, abgesehen von Ausgrabungen, relativ eben.
- 16 Da hat ein Kamel im Cirkus und Zoo
- 17 durch den Umgang mit Menschen ein höheres Niveau.
- 18 Doch von den vielen Kamelern, Zebras und Affen
- 19 kann ja nicht jeder beim Cirkus Hagenbeck schaffen.
- 20 Höhepunkt des Tages ist mittags um drei,
- 21 da kommt täglich ein Bus mit Touristen vorbei,
- 22 die ab und zu aus den Fensterchen gucken,
- 23 um Kaugummis ins alte Ägypten zu spucken.
- 24 Bald schon kommen sie in den länglichen spitzen
- 25 Schatten der hohen Pyramiden zu sitzen.
- 26 Der Rest will sich in den Grabkammern tummeln
- 27 und den Göttern an den Zehnägeln fummeln.
- 28 Gegen vier fährt der Omnibus meist wieder fort
- 29 aus der Wüste, das Kamel aber steht immer noch dort
- 30 auf allen vier Beinen im sengenden Sand,
- 31 da wo es anfänglich des Gedichtes schon stand.
- 32 (Manfred Hepperle, Es war einmal ein Uhu 2. Teil. Ravensburg 1979)

Gliederung in

Strofen

Zeilen

ÄÄen-Äusserungseinheiten

- 1.01. 01 Das Kamel
- 2.02. 02 Es stand ein Kamel in der Wüste am Nil,
2.03. 03 fraß nur ganz wenig
2.03. 04 und nur wenig entfiel.
2.04. 05 Es hat schon die Kindheit im Sande verbracht
2.05. 06 und nie an den Wechsel des Wohnorts gedacht.
2.06. 07 Den Leitsatz im Auge
2.06. 08 - egal was auch kam
2.07. 09 achala machala
2.07. 10 - dahaaam ist dahaam.
- 3.08. 11 Sein bescheidenes Dasein ist dadurch erklärlich:
3.09. 12 Vitamine gibt's ganzjährig relativ spärlich,
3.10. 13 nicht wie bei glücklichen Kühen im mannshohen Grase
3.11. 14 - hier muß man schon froh sein an einer Oase.
- 4.12. 15 Auch Steine müssen hier als Seltenheit gelten,
4.13. 16 weil Pharaonen aus ihnen Pyramiden erstellten,
4.14a. 17 Die Gegend ansonsten
4.14b. 18 - bis hinunter nach Theben -,
4.15. 19 ist, abgesehen von Ausgrabungen, relativ eben.
- 5.16-17. 20 Da hat ein Kamel im Cirkus und Zoo durch den Umgang
mit Menschen ein höheres Niveau.
5.18-19. 21 Doch von den vielen Kamelern, Zebras und Affen kann
ja nicht jeder beim Cirkus Hagenbeck schaffen.
- 6.20. 22 Höhepunkt des Tages ist mittags um drei,
6.21. 23 da kommt täglich ein Bus mit Touristen vorbei,
6.22-23. 24 die ab und zu aus den Fensterchen gucken, um Kaugum-
mis ins alte Ägypten zu spucken.
- 7.24-25. 25 Bald schon kommen sie in den länglichen spitzen
Schatten der hohen Pyramiden zu sitzen.
7.26-27. 26 Der Rest will sich in den Grabkammern tummeln und den
Göttern an den Zehnägeln fummeln.
- 8.28-29a.27 Gegen vier fährt der Omnibus meist wieder fort aus
der Wüste,
8.29b-30.28 das Kamel aber steht immer noch dort auf allen vier
Beinen im sengenden Sand,
8.31. 29 da wo es anfänglich des Gedichtes schon stand.
- 9.32. 30 (Manfred Hepperle, Es war einmal ein Uhu 2. Teil.
Ravensburg 1979)

zu Ziff. 3.2275:

Tokenisierung = Satzeinheit? "Satz" und Punkt?

zu:

aus: U. DEMSKE, Das Mercurius-Projekt. Eine Baumbank für das Frühneuhochdeutsche in: G. ZIFONUN, W. KALLMEYER (eds.), Sprachkorpora - Datenmengen und Erkenntnisfortschritt. IdS Jahrbuch 2006. Berlin 2007. S. 91-104.

(94) Der erste Schritt auf dem Weg zu einem syntaktisch annotierten Korpus umfasst die Zerlegung des Textes in Satzeinheiten, d.h. seine Tokenisierung. Bereits bei dieser Arbeitsaufgabe manifestiert sich ein entscheidender Unterschied zwischen den im Rahmen der Negra- und Tigerprojekte bearbeiteten Texten aus der Gegenwart einerseits und dem Text aus der Frühen Neuzeit andererseits: Während für die Texte aus dem heutigen Deutsch im Wesentlichen davon ausgegangen werden kann, dass Satzeinheiten durch Punkte abzugrenzen sind, markieren Satzzeichen im Frühneuhochdeutschen Satzgrenzen nur sehr unzuverlässig (Stoltz 1990). Dieses Problem möchte ich an 95) einem Textausschnitt illustrieren: Der folgende Ausschnitt enthält insgesamt drei Satzgefüge, von denen allein der Umfang des letzten Satzgefüges unstrittig ist. Der Textteil vor dem Absatz ist durch Punkte in zwei Satzeinheiten gegliedert: Der Beginn des ersten Satzes fällt mit dem Beginn des Textausschnitts zusammen (= *Es*), der zweite Satz beginnt laut Interpunktion und Majuskel mit dem Pronominalabverb *worüber*. Inhaltlich bilden beide Sätze eine Einheit. Unter syntaktischen Gesichtspunkten wäre eine Satzgrenze allerdings eher vor dem ebenfalls hervorgehobenen *Mons(ieur)* zu setzen, da hier mit einem V2-Satz eine syntaktisch unabhängige Einheit beginnt, wohingegen das mit einer Majuskel beginnende *worüber* einen V-letzt-Satz einleitet. Im Gegenwartsdeutschen sind diese weiterführenden Nebensätze abhängige Satzteile größerer Satzgefüge.

(2) **Es** sind viel/ welche vermeynen/ daß man allhier nicht die beste Opinion gegen Frankreich habe/ aus Ursachen/ weil sie in der Levante die unsrige nicht zum besten meynen sollen *Mons. d'Almeras* soll damahls/ als die Veränderung in Portugall vorgegangen/ mit vielen Frantzösischen Schiffen auff der Revier von Lisabona gewesen seyn. **Worüber** auch seltzame Gedanken entstehen.

Heute hat die Königl. Societät die Überbringung des Blutes aus einem Schafe in einen Menschen experimentirt, welches wol succedirte vor die quantität von 12 à 14 Untzen. (M 813.27)

Die Schlussfolgerung, die meines Erachtens aus Textbeispielen wie (2) zu ziehen ist, lautet, dass die Tokenisierung historischer Texte anders als im Gegenwartsdeutschen nicht automatisch zu leisten ist, sondern entweder manuell oder halbautomatisch zu erfolgen hat. Auf jeden Fall müssen vor der Tokenisierung die einschlägigen (syntaktischen) Kriterien für Satzwertigkeit genauestens festgeschrieben werden (Demske-Neumann 1990, Demske et al. 2004).

zu Ziff. 3.228:

"Koordination" im alten Grammatik-Konzept

zu: vsem4.3 vprag2.1331

aus: T. OSBORNE, Shared Material and Grammar: Toward A Dependency Grammar Theory of Non-Gapping Coordination for English and German: Zeitschrift für Sprachwissenschaft 25 (2006) 39-93.

[58] 6. String Coordination (SC)

The term "String Coordination" (SC) is borrowed from Heringer (1996: 198). He (1996: 199) characterizes SC in the following manner:

"Man kann sich das ganze so vorstellen, dass der Sprecher den String an einer Stelle unterbricht und einen Parallelweg geht, der natürlich on-line nicht parallel, sondern nur nach dem ersten Teilweg begangen werden kann. An seinem Ende kehrt er zur Abzweigstelle zurück." ...

This general understanding of SC is adopted here. SC is viewed as a mechanism that interrupts the production of a structure to repeat a part [59] of that structure that was just produced; when the coordination is finished, the original structure picks up again where it left off. The mechanism is represented schematically as follows:

(69) Fred sent a letter to Sue yesterday.
and a package to Jane

The original string is interrupted after *Sue*, the structure of the string corresponding to a *letter to Sue* is reproduced with other lexemes, i. e. a *package to Jane*, and the original string is then re-continued to its completion.

This coordination mechanism is quite flexible. It can coordinate most strings, seemingly regardless of their status as constituents or non-constituents:

(70) Fred ate lunch before work.
a. [Fred] and [Sue] ate lunch before work.
b. [Fred prepared] and [Sue ate] lunch before work
c. [Fred ate lunch] and [Sue watched TV] before work.
d. [Fred ate lunch before] and [Sue watched TV after] work.
e. Fred [prepared] and [ate] lunch before work.
f. Fred [ate lunch] and [prepared dinner] before work.
g. Fred [ate lunch before] and [watched TV after] work.
h. Fred [ate lunch before work] and [watched TV after work].
i. Fred ate [lunch] and [dinner] before work.
j. Fred ate [lunch before] and [dinner after] work.
k. Fred ate [lunch before work] and [dinner after work].
l. Fred ate lunch [before] and [after] work.
m. Fred ate lunch [before work] and [after work]
n. Fred ate lunch before [work] and [training].

The sentences (!) (70a-n) demonstrate just how flexible the mechanism is; it can coordinate every possible string in (70). Contrary to the impression generated by (70a-n) though, it will be made clear that SC is actually limited in the strings it can coordinate. ...

[62] 6.1. The Contiguity Requirement (CR)

Given Heringer's 'on-line' view of SC, it should be apparent that the conjuncts of coordinate structures are contiguous. This conti-

guity is expressed in the first constraint on SC, which is called the CR; it predicts where *shared material*, i. e. material common to each of the conjuncts, can and must appear in relation to the conjuncts.

Contiguity Requirement (CR)

(i) Within the coordinate structure, shared material may not follow the (first) root in the initial conjunct.

(ii) All material external to the conjuncts must be shared.

Part (i) prevents shared material from appearing after the first root in the initial conjunct and anywhere within the non-initial conjunct(s). The constraint makes concrete predictions about the strings that can and cannot be coordinated. Shared material may not appear between the conjunct roots. ... [63]

(73) a. *[Tom bought] an old car and [he fixed].
b. Tom [bought an old car] and [fixed it].
c. *Tom [bought] an old car and [fixed].
d. ?[Tom bought], and [hefixed] an old car.
e. Tom [Bought] and [fixed] an old car.
f. *[Bought an old car] and Tom [fixed it].
g. *[bought] and Tom [fixed] an old car.

The shared material in the acceptable (73b, d, e) appears before the first conjunct root and/or after the final conjunct bracket. (73d) is less acceptable than (73b, e); notice, however, that the marginality disappears if the subject changes, e. g. *Tom bought, and Sue fixed, an old car*. In contrast, the ungrammatical (73a, c, f, g) have shared material appearing between the conjunct roots, which is prohibited by the CR. The CR holds for German as well. (74) ...

a. *[Thomas kaufte] ein altes Auto und [er reparierte].
b. Thomas [kaufte ein altes Auto] und [reparierte es].
c. "Thomas [kaufte] ein altes Auto und [reparierte].
d. ?[Thomas kaufte] und [er reparierte] ein altes Auto.
e. Thomas [kaufte] und [reparierte] ein altes Auto.
f. *[Kaufte ein altes Auto] und Thomas [reparierte es].
g. *[Kaufte] und Thomas [reparierte] ein altes Auto.

[66] With respect to derivational theories, parts (i) and (ii) of the CR are analogous to the two components of Ross' (1967: 161) CSC:

Coordinate Structure Constraint (CSC):In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

By preventing shared material from appearing between the conjunct roots, part (i) of the CR covers the same data as the 'no conjunct movement' component of the CSC. And by requiring all conjunct external material to be shared, part (ii) of the CR addresses the same data as the prohibition on non-ATB extraction out of conjuncts. It should be noted that while the CR and the CSC largely address the same data, they differ significantly in their core views of the nature of syntax. The CR references surface strings only, whereas the CSC peers below the surface, acknowledging a movement procedure.

An important aspect of part (i) of the CR is that it does not specify the behavior of the material within the initial conjunct that precedes the first root. Indeed, the pre-root material in the initial conjunct exists in a grey zone that is not covered by part (i) or part (ii) of the CR. At times, German allows such material to be shared. Wilder (1994: 314, 1997: 75) discusses the following cases: ... [67] The conjunct-internal *ein Buch* is shared. Since it

appears before the roots *geholt* in (81a) and *hat* in (82a), its position does not violate either part of the CR. Note that if *ein Buch* is placed between or after the conjunct roots, ungrammaticality results:

- (81) b. * Ich habe [gestern *geholt*] und [heute ein Buch I have yesterday picked.up and today a book *zurückgegeben*].
given.back
'I picked up a book yesterday and today gave it back.'
- c. *Ich [*holte* gestern] und [*gab* ein Buch heute zu- rück].
I picked up yesterday and gave a book today back
'I picked up a book yesterday and returned it today.'
- (82) b. *dass Hans [mir gekauft hat] und [ihr ein Buch geben wird]
that Hans me bought has and her a book give will
'that Hans bought me a book and will give it to her'

[84] 7.3. Subjekt-lücken-Koordination (SLK)

The final (and most serious) challenge to the CR is called *Subjekt-lücken-Koordination* 'subject-gap coordination' (SLK) (Büring & Hartmann 1998) or Subject Lacking in F-structure coordination (SLF) (Höhle 1990). SLK occurs when a shared subject appears between the conjunct roots.

- (133) [In den Wald *ging* der Jäger] und [*fang* einen Hasen].
in the forest went the hunter and caught a rabbit
'The hunter went into the forest and caught a rabbit.'
(Wunderlich 1988: 289)
- (134) [Gestern *spielten* sie Klavier] und [*sangen* Lieder].
yesterday played they piano and sang songs
'Yesterday they played piano and song songs.'
(Hudson 1988: 335)
- (135) [In Italien *schätzt* man Rotwein] und [*hasst* die Franzosen].
In Italy values one red.wine and hates the French
In Italy one values red wine and hates the French.'
(Büring & Hartmann 1998: 173)

These sentences all have a shared subject appearing between the roots, so they violate the CR. The fact that they are acceptable challenges the current theory.

In order to address this challenge, four observations about SLK demonstrate that it actually involves subordination.

1. SLK occurs only with subjects; it does not occur with other constituents, e. g. objects or PPs.
2. SLK cannot occur if the pre-verbal position of the non-initial conjunct is occupied. [85]
3. SLK occurs best with subjects that are semantically typical, i. e. with agents or patients. If the subject involves some other semantic role, e. g. cause, then acceptability decreases markedly.
4. The non-initial conjunct of SLK typically adds 'adverbial' meaning to the initial conjunct.

Each of these points is illustrated in the following paragraphs.

Höhle (1990: 231) and Wilder (1996: 176) point out that SLK is limited to subjects, meaning that ungrammaticality results if some other constituent besides the subject is shared.

- (136) *[Vielleicht *sucht* sie den Hund] und [*findet* sie nicht].
maybe seeks she the dog and finds she not
'Perhaps she is looking for the dog and not finding it.'
(Höhle 1990: 230)
- (137) * [Der Jäger *hat* den Hasen gesucht] und [*hat* er auch the hunter has the rabbit sought and has he also gefunden].
found
'The hunter sought the rabbit and also found it.'
(Wilder 1996: 176)
- (136) - (137) attempt to share a direct object. As the CR predicts, these attempts fail because the shared material appears between the roots.
- Wilder (1996: 175) and Büring & Hartmann (1998: 177) point out that SLK fails if the preverbal position in the non-initial conjunct is occupied.
- (138) *[Gestern ist Maria einkaufen gegangen] und [heute hat Yesterday is Maria shopping gone and today has sich einen Pullover gekauft].
herself a sweater bought
'Yesterday Maria went shopping and she bought herself a sweater today.'
(Wilder 1996: 174)
- (139) *[In Italien *schätzt* man Rotwein] und [aus Frankreich in Italy regards one red.wine and out.of France *importiert* deshalb Trauben].
imports therefore grapes
'In Italy one regards red wine highly and therefore imports grapes from France.'
(Büring & Hartmann 1998: 177)

In diese (Beschreibungs-)Schwierigkeiten gerät, wer *nicht* mit "Äusserungseinheiten" operiert.

"und"-Anschlüsse darauf prüfen, ob sie *gleichartige* Informationen verbinden ("*Einigkeit und Recht und Freiheit*" - drei Nomina). Wenn ja: sie gehören zu *einer* ÄE.

Verbindet "und" *verschiedenartige* Informationen ("*Sie sucht den Hund und findet ihn nicht*" Nomen + konjugiertes Verb), so liegen unterschiedliche ÄEen vor. Folgerung: dieses "und" hat satzverknüpfende Funktion, darf nicht mit semantischem "und" verwechselt werden.³⁵

³⁵ Ohnehin dürfen nicht *zwei* konjugierte Verben in *einer* ÄE vorkommen. Aber auch: AE1 "Er sandte gestern Blumen" AE2 "und vorgestern Wein".

3.23 Transkription / Präsentation

Literatur

- BERRIE, P: A New Technique for Authenticating Content in Evolving Marked-up Documents. *Literary and Linguistic Computing* 22 / 1 (2007) 17-25.
- HUITFELDT, C; SPERBERG-McQUEEN, C M: What is transcription?. *Literary and Linguistic Computing* 3 (2008) 295-310.
- SCHMIDT, T: Computergestützte Transkription. Modellierung and Visualisierung gesprochener Sprache mit texttechnologischen Mitteln. *Sprache, Sprechen und Computer* 7. Frankfurt am Main 2005 ISBN 3 631 53514 7 br: Peter Lang // I.2.7
- SCHROTT, R: HOMER Ilias (Übertragung). München 2008
- SCHWEIZER, H: Der Computer und Übersetzungen unterschiedlich starker Wörtlichkeit. Erfahrungen mit dem Text der Josefsgeschichte. Actes du cinquième Colloque International 'Bible et Informatique: »Translation et Transmission«, Aix-en-Provence 1-4 septembre 1997. Paris 1998 95-112.: Honoré Champion Editeur Abk: (AIBI)
- SCHWEIZER, H: Qualitätsmessung bei Übersetzungen? Aspekte der Theorie und Werkstattbericht. *Literary and Linguistic Computing* 14 / 3 (1998) 436-444.
- SCHWEIZER, H: Übersetzung als Drama. Wie simuliert man komplexe Dekodierungsprozesse? in: GIL, A; HALLER, J, u.a. (ed./eds.): Modelle der Translation. Grundlagen für Methodik, Bewertung, Computermodellierung. SABEST 1. Frankfurt / M 1999 401-422. // EB
- VANHOUTTE, E; VAN DEN BRANDEN, R: Describing, transcribing, encoding and editing modern correspondence material. a textbase approach. *Literary and Linguistic Computing* 24 (2009) 77-98.

zu Ziff. 3.231:

Transkription gesprochener Sprache

zu:

aus: SCHWITALLA, J, *Gesprochenes Deutsch. Eine Einführung.* Berlin 1997: E. Schmidt

(21-23) Birgit und ihr Freund Gerd sind bei gemeinsamen Bekannten (Karin und Martin) eingeladen. Birgit erzählt von einem Besuch bei ihrem früheren Freund Bodo und dessen Frau Silvia in einer kleinen Stadt. Ihre Erzählung entwickelt sich immer mehr zu einer Klatschgeschichte. Alle Sprecher stammen aus Südbaden. Zum Verständnis der Verschriftlichung gesprochener Äußerungen gebe ich hier die wichtigsten Transkriptionszeichen wieder:...

\ fallende Intonation
/ steigende Intonation
* kurze Pause
** mittlere Pause
2,5 längere Pause mit Angabe der Dauer in Sekunden
< lauter gesprochen
> leiser gesprochen
→ schneller gesprochen
← langsamer gesprochen
" starker Akzent
: auffallende Dehnung
Jürgen Unterstrichenes wird simultan gesprochen
(eben) vermuteter Wortlaut
(...) Unverständliches
= Verschleifungen zwischen Wörtern
Grenzsymbole für das Gelten eines Kommentars
LACHEND Kommentare in Großbuchstaben

Text 1

01 BI: #kann=s dir gar nit sage\##* es war nu"r wichtig/
K #LACHEND-----#

02 BI: wer#het#im haus die meischde gru"ndfläche/ wer het noch
K #HAT#

03 BI: am längschde #siebbe be"# abzuschreibe/**
K #DER ABSCHREIBEPARAGRAPH 7b#

04 GE: aha

05 BI: und we"r klei"det sei"nen spe"ck am vo"rteilhafteschde\
06 BI: die Silvia /*het also so ein ein

07 MA: LACHT

08 BI: medikame"ntenköfferle noch dabei ghabt/ denn sie het

09 BI: ja au tau"send wehwehle\Silvia Silvia isch ni"cht

10 BI: berufstätig/ Silvia kümmert ich nu"r um das neue

11 BI: haus und um die aufzucht von diesem Jürgen\
12 KA: #aufzucht#
K #LACHEND#

13 BI: <aber de Jürgen isch noch ein liebliches

14 BI: wenn der Bodo mol mich über de disch
15 GE: wie reds du wie reds denn du

16 BI: <nein/ >so war=s\ →hab ich dir doch erzählt/
17 BI: no het der Bodo mol so wie er des früher so schön

18 BI: #kenne het so e weng zwingert un a weng gflördet#
K : #KÖNNEN HAT SO EIN WENIG GEZWINKERT UND EIN WENIG GEFLIR-
TET#

19 BI: het er gwagt/* sei matrone nebedran #verschdehsch/#*
K #VERSTEHST DU#

20 BI: de Bodo guckt mich an/ redet mit mir

21 GE: isch jetzt doch=ne komische zeitform wie er
 22 GE: so schön kenne het des gibt=s grammatikalisch
 23 MA: hm/
 24 GE: gar nich\ ne/ ja/ is des perfekt/
 25 MA: doch doch\ do gibt=s
 26 BI: im badische gibt=s des*
 27 MA: alles grammatikalisch\
 28 BI: #do her mol vor# vor so * er isch mir so gegenüber
 K #DA HÖR MAL ZU#
 29 BI: gsesse het mir dann was erzählt/* un het mir
 30 BI: dann so zugezwingert\ s=isch natürlich seiner Silvia
 31 BI: ni"cht entga"nge* un sofort het er sich
 32 BI: en schdrafende blick einfange/*
 33 BI: weil die Silvia nehm ich uff jeden fall amol an\
 34 BI: weiß/* daß des vor hundert johr mol mit
 35 BI: uns was war\ un sowas von einer* ←tra"nfunsligen
 36 KA: LACHT
 37 BI: a"lten frau\ ja des war die
 38 GE: die horrorparty war des\
 39 BI: rocky horror picture show war des\
 ("verunglimpfung")

zu Ziff. 3.2311:

SGML/XML - Merkmale der Markup-Sprachen

zu:

aus: SCHMIDT, T, Computergestützte Transkription. Modellierung und Visualisierung gesprochener Sprache mit texttechnologischen Mitteln. Sprache, Sprechen und Computer 7. Frankfurt/M 2005

(63)

(1) Das Prinzip der Trennung von Form und Inhalt

Wie Witt (2002: 7f) anmerkt, stellt das Prinzip der Trennung von Form und Inhalt ein Grundprinzip für die Verwendung von Auszeichnungssprachen dar:

In nahezu allen einführenden Texten zu SGML und XML wird darauf hingewiesen, dass das einzuhaltende Grundprinzip ihrer Verwendung in der Trennung von Form und Inhalt liegt. (...)

Es kann (...) davon gesprochen werden, dass das Ziel der Trennung von Inhalt und Form als Konsens anzusehen ist.

Die Dichotomie *Form vs. Inhalt* ist dabei nur eine von vielen möglichen Formulierungen. Sie findet sich an anderen Stellen z.B. unter den Begriffen *graphische vs. logische Struktur* (Wohlberg 1999), *Form vs. inhaltliche Strukturierung* (Lobin 1999b), *Format/Präsentation vs. Inhalt* (Ray 2001), *(visual) display vs. underlying representation* (Johansson 1995) oder *presentation vs. organization of the data* (Knowles 1995) wieder. Historisch gesehen ist dieses Prinzip aus dem Bedürfnis entstanden, digital gespeicherte Dokumente in verschiedenen Kontexten wieder verwenden zu können. Die inhaltliche Auszeichnung hat sich in dieser Hinsicht im Vergleich zur formbasierten als flexibler handhabbar erwiesen. Warum das Prinzip der Trennung von Form und Inhalt eines Dokumentes praktikabel ist, obwohl jedes Dokument die Summe aus seiner Form und seinem Inhalt ist, erklärt Michel (1999: 34) so: (64)

(...) obwohl Inhalt und Layout nur Aspekte von Dokumenten sind, d.h. nur begrifflich, nicht physisch, zu trennen sind, verlangt XML von uns die Abstraktion vom Layout eines Dokuments und damit die Abstraktion von dem, was seinen Austausch zwischen Plattformen, Anwendungen und Medien problematisch machen kann.

Der Vorteil eines inhaltsbasierten Markups liegt demnach in seinem höheren Abstraktionsgrad. ...

(2) Die OHCO-These

Ein weiterer Grund für die technische Praktikabilität von Auszeichnungssprachen liegt darin, dass sie von einer starken Grundannahme über die Struktur von Dokumenten ausgehen: So wie im relationalen Modell angenommen wird, dass sich Daten immer in Relationen bzw. Tabellen organisieren lassen, so liegt den Sprachen XML und SGML die These zugrunde, dass Dokumente immer eine geordnete Hierarchie darstellen, also baumförmig strukturiert sind. Darauf aufbauend kann dann die Arbeit mit XML- oder SGML-Daten als ein Manipulieren von Bäumen aufgefasst und bei der Implementierung der zugehörigen Technologien von entsprechenden Kenntnissen über solche Strukturen profitiert werden - beispielweise, um schnelle und effektive Algorithmen zu erhalten.

Die These ist unter dem Namen *OHCO-These* (DeRose et al. 1990: 6) bekannt. Sie besagt genauer, dass jedes Dokument eine "Ordered Hierarchy of Content Objects" darstellt, dass also alle in einem Dokument enthaltenen (inhaltlichen) Objekte sich erstens (linear)

anordnen lassen und zweitens als Bestandteile einer Hierarchie fungieren. Für das Gebiet der Auszeichnungssprachen heißt das konkret, dass durch (XML- oder SGML-)Elemente ausgezeichnete Umgebungen sich nicht überlappen dürfen. Wie Witt (2002: 41) ausführt, erweist sich diese These bei genauerer Betrachtung aber bereits für den ursprünglichen Anwendungsbereich von Auszeichnungssprachen - nämlich geschriebene Texte - als eine zu starke Einschränkung: (65)

In Texten können potentiell mehrere unabhängige Hierarchien existieren, deren Begrenzungen sich überlappen. Es kann vermutet werden, dass sich die Hierarchie-These nur entwickeln konnte, da ihr ein sehr enger Textbegriff zugrunde liegt. Texte bestehen aus Kapiteln, Abschnitten, Überschriften etc. Naheliegender war dies insbesondere deshalb, da die Wurzeln der Annotation im Buchdruck lagen (...).

(3) Dokumentzentrisches vs. Datenzentrisches Verständnis

Wie bereits mehrfach bemerkt, verdanken die Auszeichnungssprachen ihre Entstehung und auch ihre Namen vornehmlich dem Bedürfnis, die in *schriftlichen Dokumenten* enthaltene Information flexibel nutzbar zu halten, indem man sie mittels Markup explizit auszeichnet und so einer computergestützten Verarbeitung zugänglich macht. Wie z.B. Thompson (2001) deutlich macht, hat sich besonders für den Fall von XML aber herausgestellt, dass der Nutzen und damit der Einsatzbereich von Auszeichnungssprachen weit über die Anwendung auf Dokumente im engeren Sinne hinausgeht. Er stellt daher dem ursprünglich *dokumentzentrischen* Verständnis von XML ein *datenzentrisches* entgegen und charakterisiert letzteres durch die Aussage: "A document is just a container in which information or data is poured".

(66) Dass sich Auszeichnungssprachen auch in diesem ursprünglich nicht anvisierten Sinne als nützlich erweisen, kann man mit mehreren Tatsachen begründen, die hier nur stichwortartig wiedergegeben werden sollen (vgl. dazu wiederum Thompson 2001):

- XML und SGML sind offene Standards, deren Definition und Pflege in den Händen des W3C-Konsortiums liegen und damit langfristig gesichert sind.
- Speziell XML baut auf vorhandenen Technologien und Standards auf. Es erlaubt eine konsequente Verwendung des UNICODE-Standards zur Kodierung einzelner Zeichen, ist selbst eine Anwendung des SGML-Standards und weist nicht nur in dieser Hinsicht offensichtliche Parallelen zum weitverbreiteten HTML auf.
- Wie oben ausgeführt basieren XML und SGML auf der OHCO-These, deren zentrale Aussage zwar nur begrenzte Gültigkeit für schriftliche Dokumente besitzt, dafür aber auf viele weitere Datenstrukturen angewendet werden kann - denn auch nicht textförmige Daten sind oft hierarchisch strukturiert.

zu Ziff. 3.2312:

Transkription: geschrieben ↔ gesprochen

zu:

aus: SCHMIDT, T, Computergestützte Transkription. Modellierung und Visualisierung gesprochener Sprache mit texttechnologischen Mitteln. Sprache, Sprechen und Computer 7. Frankfurt/M 2005

(90) Im Prinzip lassen sich die obigen Ausführungen zunächst auf all diejenigen Sprachen, deren Standardorthographie auf dem lateinischen Alphabet (oder allgemeiner: auf alphabetischen Schriftsystemen, was z.B. auch das griechische und kyrillische Schriftsystem einschließt) basiert, übertragen. Beispielsweise existieren selbstverständlich auch für das Englische, das Französische oder das Russische Regelwerke, die die Schreibweise einzelner Wörter und die Verwendung von Interpunktionszeichen verbindlich festlegen, und diese können analog zum Deutschen als Referenz für die Transkription der Standardlautung verwendet werden und gleichzeitig die Basis für eine literarische Umschrift dieser Sprachen darstellen.

Dabei finden sich aber große graduelle Unterschiede im Hinblick auf die Systematizität vor allem der Phonem-Graphem-Korrespondenz. Zur Lauttreue ("faithfulness") der Orthographien verschiedener Sprachen sagt Sproat (2002: 35):

An obvious example of a fairly orthographic system is Finnish; other systems that are typically claimed to be faithful are Spanish and Serbo-Croatian. At the other extreme are orthographies that are relatively poor indicators of surface pronunciation. The archetypical instance of such an orthography is English, but other examples abound, including French, Danish, Hebrew and Russian.

Er unterscheidet bei den hinsichtlich des phonologischen Prinzips problematischen Orthographien weiter drei Typen. Englisch ist demnach ein Beispiel für eine Sprache, deren Orthographie schlicht *idiosynkratisch* ist, d.h. Unregelmäßigkeiten in der englischen Orthographie sind weder einheitlich, noch lassen sie sich ausreichend durch andere Prinzipien als das phonographische rechtfertigen, z.B.:

<i>tough, though, through</i>	(Graphemkette <ough> entspricht drei verschiedenen Lautungen)
<i>Worcester, Gloucester, Leicester</i>	(Graphemkette <ce> hat keine eindeutige lautliche Entsprechung)

Semitische Schriftsysteme wie das hebräische oder das arabische hingegen weisen eine *unvollständige Abdeckung* auf der phonologischen Ebene auf - sie bilden lediglich die Konsonanten und nicht die Vokale auf Grapheme ab.

Die französische Orthographie schließlich ist ein Beispiel für ein System, dessen Inkonsistenz auf der phonologischen Ebene teilweise *durch andere Prinzipien erklärbar* ist. Unterschiedliche Schreibweisen identisch lautender Wörter motivieren sich dort z.B. aus grammatischen Differenzierungen:

<i>mange, manges, mangent</i>	[mäʒ]
<i>mangeait, mangeais, mangeaient</i>	[mäʒɛ]
<i>manger, mangé, mangée, mangés, mangées</i>	[mäʒe]

zu Ziff. 3.2313:

Transkription: Partiturnotation

zu:

aus: SCHMIDT, T, Computergestützte Transkription. Modellierung und Visualisierung gesprochener Sprache mit texttechnologischen Mitteln. Sprache, Sprechen und Computer 7. Frankfurt/M 2005

(93) Das Grundprinzip dieser Notationsform beschreiben Rehbein et al. (1993:4) so:

Die Grundidee [...] ist, beim Verschriftlichen die Kommunikation nicht sukzessive in einzelnen Zeilen abzutragen, sondern für jeden Beteiligten eine eigene Handlungslinie zu konzipieren. [...] Die Beiträge der Aktanten werden entsprechend ihrer Handlungslinien graphisch separiert dargestellt; die Handlungslinien aller in derselben Kommunikation agierenden Aktanten zu einer *fortlaufenden Fläche* zusammengefasst.

Der grundlegende Unterschied zwischen der Notation schriftsprachlicher Texte und der Partiturnotation besteht demnach darin, dass bei der ersteren Symbole entlang einer *eindimensionalen* Linie, bei der letzteren hingegen auf einer *zweidimensionalen* Fläche angeordnet werden. Während schriftsprachliche Texte somit nur ein logisches oder sprachliches *Vorher und Nachher* abzubilden vermögen, wird es durch die zweite Dimension der Partiturnotation zusätzlich auch möglich, die in der gesprochenen Sprache so häufig auftretenden zeitlich *parallelen* Beziehungen graphisch zu repräsentieren.

Formal lässt sich diese Methode als Einordnung von symbolischen oder ikonischen Einheiten in ein zweidimensionales Koordinatensystem beschreiben. Auf der von links nach rechts verlaufenden Achse sind dabei einzelne Zeitpunkte geordnet abgetragen, während die von oben nach unten verlaufende Achse eine Abfolge der im obigen Zitat als *Handlungslinien* bezeichneten Elemente aufnimmt. Dabei kann es - wie Rehbein et al. (1993: 5) auch sagen ("Die singuläre Handlungslinie eines Aktanten kann [...] wieder in unterschiedliche Linien aufgefächert werden") - für einen Sprecher auch mehr als eine Handlungslinie geben z.B. wenn verschiedene Handlungstypen unterschieden werden. Ein Beispiel soll dies verdeutlichen:

MAX (v)	Du fällst mir immer ins Wort.	Siehst Du, Du hast es schon, wieder getan.
MAX (nv)	<i>fuchelt mit den Armen</i>	
TOM (v)		Stimmt ja gar nicht
TOM (nv)		<i>abfällige Geste</i>

Dieses nach dem Layoutprinzip der Partiturnotation organisierte Transkript kann als die Visualisierung eines Symbolischen Modells aufgefasst werden, das die in ihm enthaltenen Symbolketten jeweils einem Start- und Endpunkt sowie einem Sprecher und einem Handlungstyp zuordnet:

	Start	Ende	Sprecher	Typ	Symbolkette
1	t_0	t_1	MAX	v	Du fällst mir immer ins
2	t_1	t_2	MAX	v	Wort.
3	t_0	t_2	MAX	nv	fuchelt mit den Armen
4	t_3	t_4	MAX	v	Siehst Du, Du hast es schon
					wieder getan.
5	t_1	t_2	TOM	v	Stimmt
6	t_2	t_3	TOM	v	ja gar nicht.
7	t_1	t_3	TOM	nv	abfällige Geste

zu Ziff. 3.232:

Gliederung/Transkription/Arbeitsübersetzung

zu:

aus: 2 Sam 14

וַיַּעַן יוֹאָב בֶּן צִרוּיָהּ	wa=ye[n]da ^ˁ YW ^ˁ B bân ṢRYH	1a	und=er/erfuhr JOAB, (des)-ZERUJA Sohn
כִּי לֵב הַמֶּלֶךְ עַל אֲבָשָׁלוֹם	ki(y) leb ha=mäläk	1b	dass das-Herz des=Königs ()
וַיִּשְׁלַח יוֹאָב חֲקֹרָעַי	^ˁ al ^ˁ BŠLWM wa=yišlah YW ^ˁ B TQW ^ˁ =H	2a	über ABSCHALOM und=er/sandte JOAB
וַיִּקַּח מִשָּׁם אִשָּׁה חַכְמָה	wa=yi[l]qah mi[n]=šäm ^ˁ iššä(h) ḥkämä(h)	2b	und=er/nahm von=dort eine-Frau eine-wei- se
וַיֹּאמֶר אֵלָיו הֲחֹנְאֵלִי נָא	wa=yo(^ˁ)mār ^ˁ el-ä(y)=hä	2a	und=er/sprach zu=ihr
וְלִבְשֵׁי נָא בְגָדֶיךָ	w=libši(y) nā(^ˁ) bigde(y) ^ˁ ebäl	2e	Trauerritten-voll- ziehe doch und=ziehe-an doch Kleider (von-)Trau- er
וְאֵל חֲסֹרְכֵי שְׂמֹךְ	w= ^ˁ al täsü(w)ki(y)	2f	und=ja-nicht ver- wende
וְהִיָּיתָ כְּאִשָּׁה	šämän w=hayit k= ^ˁ iššä(h)	2g	Öl! und=sein-wirst\du wie eine-Frau
וְהַיָּמִים רַבִּים עָלַי מִחַמְאָלָתְךָ עַל מַח	zä(h) yâmi(y)m rabbi(y)m mit ^ˁ abälät	2h	welche () Tage viele eine-Trauerritten- vollziehende
וְרַבְחָא אֶל הַמֶּלֶךְ	^ˁ al met u(w)=bâ(^ˁ)t	3a	über einen-Verstor- benen und=hingehen-wir- st\du
וַיְדַבֵּר אֵלָיו כַּדְבָר הַזֶּה	^ˁ äl ha=mäläk w=dibbart ^ˁ el-a(y)=w k=[h]a=dâbâr ha=zä(h)	3b	zu dem=König und=reden-wirst\du zu=ihm entspre- chend=dem=Wort dem=folgenden
וַיִּשֶׂם יוֹאָב אֶת הַדְּבָרִים בְּפִי הָאִשָּׁה	wa=yäsäm YW ^ˁ B ^ˁ ät ha=dbâri(y)m	3c	und=er/legte JOAB → die=Worte
וַיֹּאמֶר הָאִשָּׁה חַכְמָלְעִית אֶל הַמֶּלֶךְ	b=pi(y)=hä wa=to(^ˁ)mār hä= ^ˁ iššä(h) ha=tqo ^ˁ i(y)t	4a	in=Mund=ihren und=sie/sprach die=Frau die=teko- itische
וַתַּפֵּל עַל אֶפְיָהּ אֶרְצָה	^ˁ äl ha=mäläk wa=ti[n]pol ^ˁ al ^ˁ app-ä(y)=hä	4b	zu dem=König und=sie/fiel auf Gesicht=ihres Erd=wärts
וַתִּשְׁתַּחֲוֶה	^ˁ ärs=ä(h) wa=tištâḥu(w)	4c	und=sie/verbeugte- sich

zu Ziff. 3.233:

001:::37,02b*	JOSEF gewesen-ist ein-Hüter mit Brüdern+seinen bei+dem-Kleinvieh
002:::37,02c	und+er () ein-Knabe
003:::37,03a	und+ISRAEL geliebt-hat → JOSEF aus+der-Gesamtheit Söhne+seiner
004:::37,03b	denn ein-Sohn des-Alters () er ↑+ihm
005:::37,03c	und+gemacht-hatte&er[0020] ↑+ihm einen-Leib- -rock
006:::37,05a	und+träumte JOSEF einen-Traum
007:::37,05b	und+er\erzählte ↑+Brüdern+seinen
008:::37,06a	und+er\sprach zu+ihnen
009:::37,06b	hört doch den+Traum den+selben
010:::37,06c	welchen geträumt-habe\ich
011:::37,07a	und+da
012:::37,07b	wir () Bindende Garben[0040] in+Mitten des+Feldes
013:::37,07c	und+da
014:::37,07d	aufgestanden-ist Garbe+meine
015:::37,07e	und+sogar stehenblieb\sie
016:::37,07f	und+da
017:::37,07g	(im-Kreis-)stellten-sich Garben+eure
018:::37,07h	und+sie\verneigten-sich vor+Garbe+meiner
019:::37,08a	und+sprachen zu+ihm Brüder+seine
020:::37,08b	?+als-König-Regieren du-möchtest\als-König-regieren über+uns
021:::37,08c	(?)-oder Herrschen[0060] du-willst\herrschen über+uns
022:::37,11a	und+eifersüchtig-waren auf+ihn Brüder+seine
023:::37,11b	und+Vater+sein bewahrt-hat → die+Rede/Sache
024:::37,12a	und+gingen Brüder+seine (um-)zu+weiden → das-Kleinvieh Vaters+ihres in+SICHEM
025:::37,13a	und+sprach ISRAEL zu JOSEF[0080]
026:::37,13b	?+nicht Brüder+deine () Weidende in+SICHEM
027:::37,13c	geh+doch
028:::37,13d	und+ich-werde/mod.\schicken+dich zu+ihnen
029:::37,13e	und+er\sprach zu+ihm
030:::37,13f	da+ich
031:::37,14a	und+er\sprach zu+ihm
032:::37,14b	geh also
033:::37,14c	sieh →/nach dem-Wohlbefinden Brüder+deiner und+→/nach dem-Wohlbefinden[0100] des+Kleinviehs
034:::37,14d	und+bring+mir Meldung
035:::37,14e	und+er\schickte+ihn aus+der-Ebene(-von) HEBRON
036:::37,14f	und+er\gelangte SICHEM+wärts
037:::37,15a	und+stieß-auf+ihn ein-Mann
038:::37,15b	und+da ein-Umherirrender auf+dem(-freien)-Feld
039:::37,15c	und+fragte+ihn der+Mann folgendermaßen
040:::37,15d	was du\suchst
041:::37,16a	und+er\sprach
042:::37,16b	→[0120] Brüder+meine ich () ein-Suchender
043:::37,16c	sage+doch doch ↑+mir
044:::37,16d	wo sie () Weidende
045:::37,17a	und+sprach der+Mann
046:::37,17b	weitergezogen-sind\sie von+hier
047:::37,17c	denn gehört-habe\ich Redende
048:::37,17d	wir\gehen+doch DOTAN+wärts
049:::37,17e	und+ging JOSEF[0140] nach Brüdern+seinen
050:::37,17f	und+er\ fand+sie in+DOTAN
051:::37,18a	und+sie\sahen →+ihn von+weitem
052:::37,18b	und+bevor er\gelangte zu+ihnen
053:::37,18c	und+sie\verschworen-sich gegen+ihn zu+töten+ihn
054:::37,19a	und+sie\sprachen jeder zu Bruder+seinem
055:::37,19b	da
056:::37,19c	der-Meister der+Träume[0160] der+genau+selbe angekommen-ist

057::37,20a und+jetzt
 058::37,20b geht
 059::37,20c und+wir\umbringen+ihn
 060::37,20d und+wir\werfen+ihn in+einen der+Brunnen
 061::37,20e und+sprechen-werden/mod.\wir
 062::37,20f ein-Tier ein-Böses gefressen-hat+ihn
 063::37,20g und+wir-werden/mod.\sehen
 064::37,20h was werden Träume+seine
 065::37,23a und+es\war
 066::37,23b als gekommen-war JOSEF[0180] zu Brüdern+seinen
 067::37,23c und+sie\ausziehen-ließen → JOSEF → Rock+seinen →
 den-Leib- -rock
 068::37,23d welcher () auf+ihm

zu Ziff. 3.234: vprag7.13

Der Text am Ende der Pragmatik

TLE1* (TGE1*)

Block: 1

001::37,02b* JOSEF gewesen-ist ein-Hüter mit Brüdern+seinen
 bei+dem-Kleinvieh
 002::37,02c und+er () ein-Knabe

Block: 2

003::37,03a und+ISRAEL geliebt-hat → JOSEF aus+der-Gesamtheit
 Söhne+seiner
 004::37,03b denn ein-Sohn des-Alters () er ↑+ihm
 005::37,03c und+gemacht-hatte&er[0020] ↑+ihm einen-Leib- -rock

Block: 3

006::37,05a und+träumte JOSEF einen-Traum
KHS: 37,05b-11b **Nr. 1**
 007::37,05b und+er\erzählte ↑+Brüdern+seinen
 008::37,06a und+er\sprach zu+ihnen
 009::37,06b hört doch den+Traum den+selben
 010::37,06c welchen geträumt-habe\ich
 011::37,07a und+da
 012::37,07b wir () Bindende Garben[0040] in+Mitten des+Feldes
 013::37,07c und+da
 014::37,07d aufgestanden-ist Garbe+meine
 015::37,07e und+sogar stehenblieb\sie
 016::37,07f und+da
 017::37,07g (im-Kreis-)stellten-sich Garben+eure
 018::37,07h und+sie\verneigten-sich vor+Garbe+meiner
 019::37,08a und+sprachen zu+ihm Brüder+seine
 020::37,08b ?+als-König-Regieren du-möchtest\als-König-regieren
 über+uns
 021::37,08c (?) -oder Herrschen[0060] du-willst\herrschen
 über+uns
 022::37,11a und+eifersüchtig-waren auf+ihn Brüder+seine
 023::37,11b und+Vater+sein bewahrt-hat → die+Rede/Sache
ENDE KHS: **Nr. 1**
 024::37,12a und+gingen Brüder+seine (um-)zu+weiden →
 das-Kleinvieh Vaters+ihres in+SICHEM

TLE2* (TGE2*)

KHS: 37,13a-14e

Nr. 2

Block: 4

025::37,13a und+sprach ISRAEL zu JOSEF[0080]
 026::37,13b ?+nicht Brüder+deine () Weidende in+SICHEM
 027::37,13c geh+doch
 028::37,13d und+ich-werde/mod.\schicken+dich zu+ihnen
 029::37,13e und+er\sprach zu+ihm
 030::37,13f da+ich
 031::37,14a und+er\sprach zu+ihm
 032::37,14b geh also
 033::37,14c sieh →/nach dem-Wohlbefinden Brüder+deiner
 und+→/nach dem-Wohlbefinden[0100] des+Kleinviehs

034::37,14d und+bring+mir Meldung
 035::37,14e und+er\schickte+ihn aus+der-Ebene(-von) HEBRON
 ENDE KHS: Nr.2
 036::37,14f und+er\gelangte SICHEM+wärts
 KHS: 37,15a-17d Nr.3
 037::37,15a und+stieß-auf+ihn ein-Mann
 038::37,15b und+da ein-Umherirrender auf+dem(-freien)-Feld
 039::37,15c und+fragte+ihn der+Mann folgendermaßen
 040::37,15d was du\suchst
 041::37,16a und+er\sprach
 042::37,16b →[0120] Brüder+meine ich () ein-Suchender
 043::37,16c sage+doch doch ↑+mir
 044::37,16d wo sie () Weidende
 045::37,17a und+sprach der+Mann
 046::37,17b weitergezogen-sind\sie von+hier
 047::37,17c denn gehört-habe\ich Redende
 048::37,17d wir\gehen+doch DOTAN+wärts

ENDE KHS:

Nr.3

Block: 5

049::37,17e und+ging JOSEF[0140] nach Brüdern+seinen
 050::37,17f und+er\ fand+sie in+DOTAN
 051::37,18a und+sie\sahen →+ihn von+weitem
 052::37,18b und+bevor er\gelangte zu+ihnen
 KHS: 37,18c-20h Nr.4
 053::37,18c und+sie\schworen-sich gegen+ihn zu+töten+ihn
 054::37,19a und+sie\sprachen jeder zu Bruder+seinem
 055::37,19b da
 056::37,19c der-Meister der+Träume[0160] der+genau+selbe
 angekommen-ist
 057::37,20a und+jetzt
 058::37,20b geht
 059::37,20c und+wir\umbringen+ihn
 060::37,20d und+wir\werfen+ihn in+einen der+Brunnen
 061::37,20e und+sprechen-werden/mod.\wir
 062::37,20f ein-Tier ein-Böses gefressen-hat+ihn
 063::37,20g und+wir-werden/mod.\sehen
 064::37,20h was werden Träume+seine

ENDE KHS:

Nr.4

TLE2* (TGE3*)

Block: 6

065::37,23a und+es\war
 066::37,23b als gekommen-war JOSEF[0180] zu Brüdern+seinen
 067::37,23c und+sie\ausziehen-ließen → JOSEF → Rock+seinen →
 den-Leib- -rock
 068::37,23d welcher () auf+ihn

zu Ziff. 3.235:

Textkritik: elektronische Edition

zu:

aus: Alexandra BRAUN-RAU, William Shakespeares "King Lear" in seinen Fassungen. Ein elektronisch-dialogisches Editionsmodell. Tübingen 2004.

vgl: <http://www.textkritik.uni-muenchen.de/abraun-rau/Lear/>

vgl: http://www.niemeyer.de/links/link_material.html/

(115) Um die aus textkritischer Sicht erforderlichen Prämissen in einer elektronischen Edition abdecken zu können, werden die im vorhergehenden Kapitel angesprochen Textebenen *Edierte Fassungen*, *Faksimiles* und *Transkriptionen* integriert. Zudem enthält die Edition *Erschließungshilfen*, etwa den *Textkritischen Kommentar* und *Verzeichnisse*, ebenso wie weiterführende *Materialien*. ...

(117) Um aber auch dem textkritisch interessierten Nutzer, der mit dem Sprachgebrauch Shakespeares nur wenig vertraut ist, den Zugang zu einer Auseinandersetzung mit der Überlieferungsproblematik der Fassungen zu erleichtern, werden in die Edition auch die *Modern Spelling*- Fassungen von Q1 und F1 integriert. Während die *Old Spelling*-Texte sprachgeschichtlich interessierten Nutzern dienlich sind, kann die Edition durch die Einbindung der *Modern Spelling*-Texte unter anderen die Funktion einer Studienausgabe erfüllen oder Schauspielern als Grundlage für eine fassungsbezogene Interpretation und Aufführung dienen. ...

Ein erhöhtes Maß an Transparenz wird durch die Beigabe von Faksimiles erreicht. Da der Nutzer mit Hilfe der Faksimiles in die Lage versetzt wird, die editorische Beurteilung einer Textstelle in den Überlieferungsträgern nachprüfen zu können, fungieren die Faksimiles vornehmlich als Instrument zur Texterschließung und zur Verdeutlichung des editorischen Vorgehens. ...

(118) Ergänzend zu den edierten Fassungen und den Faksimiles der Drucke werden Transkriptionen der ersten Quarto- und der ersten Folioausgabe zur Verfügung gestellt. Die Integration der Transkriptionen ist ein zusätzliches Angebot an den Nutzer, bei schlecht lesbaren Stellen in den elektronischen Faksimiles auf die editorischen Transkriptionen zurückzugreifen und so die Textgrundlage für die edierten Fassungen im Vergleich mit den Originaldokumenten prüfen zu können. ...

Aufgabe des textkritischen Kommentars ist es, das editorische Verfahren bei der Textkonstitution zu erläutern und Überlieferungsprobleme transparent zu machen. Kommentiert werden Eingriffe bei Textfehlern und diskursiv gestalteten Textstellen mit gleichwertigen Varianten. Fassungsvarianten werden nur kommentiert, wenn das editorische Verfahren von der Textkonstitution früherer Editoren der Fassungen abweicht. ...

(119) Während die Gründe für die Annahme und Emendation eines Textfehlers im Kommentar diskutiert werden, dokumentieren die Verzeichnisse *Textfehler Q1* und *Textfehler F1* die in den jeweiligen Textgrundlagen als fehlerhaft eingestufteten Lesungen. Daneben können im Verzeichnis *Substantielle Varianten* die sinntragenden, als Fassungsvarianten eingestufteten Lesungen, im Verzeichnis *Akzidentelle Varianten* alle Divergenzen in Orthographie und Interpretation eingesehen werden. ...

Hilfestellung zu textkritischen Untersuchungen geben weiterführende Materialien wie *Nachschlagewerke* sowie Auszüge aus

dem *OED* und aus den *Quellen und Vorlagen* des Dramas. Auch Spezialindices, Literaturverzeichnisse und ein Glossar sind in der Edition enthalten. Im Ordner *Internet* können zudem Links zu textkritisch relevanten Shakespeare-Seiten im World Wide Web (WWW) abgespeichert werden. ...

Um dem Nutzer einen Überblick über die Materialität der Drucke und deren Druckprozeß zu geben, enthält die Edition einen einführenden Essay zu den *Textzeugen*. Ebenfalls integriert ist ein Überblick über die *Editionsgeschichte* der Drucke und eine Erläuterung der im Rahmen der Edition verfolgten *Editionsprinzipien*. Auch ein Vorwort (*Zu dieser Ausgabe*) ist beigefügt.

zu Ziff. 3.236:

Formen der Textpräsentation

zu:

nach: H. SCHWEIZER, Wovon reden die Exegeten? Zum Verständnis der Exegese als verstehender und deskriptiver Wissenschaft: Theologische Quartalschrift 164 (1984) 161-185.

a) Ps 126: transkribiert; nach Äußerungseinheiten gegliedert; Arbeitsübersetzung

š(i)y(r) ha-ma ^c lo(w)t	1a	Lied des Hinaufgehens
b=su(w)b	1b-1	Beim-Wenden
YHWH	1b-2	JAHWES
ʾät	1b-3	→
šbu(w)t	1b-4	das-Geschick
šYYWN	1b-5	ZIONS
hayinu(w)	1b-6	waren-wir/werden-wir-sein
k=holmi(y)m	1b-7	wie=Träumende.
ʾaz	2a-1	Damals
yi[n]male(ʾ)	2a-2	erfüllte/wird-erfüllen
šhoq	2a-3	Lachen
pi(y)=nu(w)	2a-4	Mund=unsere
w=lšon-e=nu(w)	2b-1	und=Zunge=unsere
rinna(h)	2b-2	Jubel,
ʾaz	2c-1	dann
yo(ʾ)mru(w)	2c-2	werden-sie-sprechen/haben-sie-gesprochen
b=[h]a=goyim	2c-3	in=den-Völkern:
higdi(y)l	2d-1	»Groß-erwiesen-hat-sich
YHWH	2d-2	JAHWE
l-a= ^c šo(w)t	2d-3	im=Tun
^c im	2d-4	mit
ʾellä(h)	2d-5	ihnen.<
higdi(y)l	3a-1	Groß-erwiesen-hat-sich
YHWH	3a-2	JAHWE
l-a= ^c šo(w)t	3a-3	im=Tun
^c imm-a=nu(w)	3a-4	mit=uns.
hayi(y)=nu(w)	3b-1	Wir-werden-sein/wir-waren
šmeḥi(y)m	3b-2	Sich-Freude.
šu(w)b-a(h)	4a-1	Wende-doch,
YHWH	4b-1	JAHWE,
ʾät	4c-1	→
šbu(w)t-e=nu(w)	4c-2	Geschick=unser
k-a=ʾpi(y)qi(y)m	4c-3	wie=die-Bäche
b=[h]a=nägäb	4c-4	in=dem=Südland!
*ha=zor ^c i(y)	5a-1	Die Säenden
b=dim ^c a(h)	5a-2	in/mit=Träne(n),
b=rinna(h)	5b-1	in/mit=Jubel
yiqšoru(w)	5b-2	werden-sie-ernten;
haló(w)k	6a-1	Gehen
ye[h]lek	6a-2	er-geht,
w=bako(w)	6b-1	und=Weinen,
noše(ʾ)	6c-1	ein-Tragender
māšāk	6c-2	des-Wurfs
ha=zara ^c	6c-3	des=Samens,
bo(ʾ)	6d-1	Kommen
yabo(ʾ)	6d-2	er kommt
b=rinna(h)	6d-3	in/mit=Jubel

noúe(ʹ 6e-1 ein-Tragender
 ʹlummot-a(y)=w 6e-2 Garben=seine.

e) Lk 18, 9-14 nach Äußerungseinheiten gegliedert

9a ἔλεπεν δὲ καὶ πρὸς τινὰς τοὺς πεποιθότας ἐφ' ἑαυτοῖς
 9b ὅτι ἐλισὶν δίκαιοι
 9c καὶ ἔξουθενοῦντας τοὺς λοιποὺς τὴν παραβολὴν ταυτὴν
 10a ἄθροοποι δύο ἀνέβησαν εἰς τὸ ἱερὸν προσεῦξασται,
 10b ὁ εἰς φαρισαῖος
 10c καὶ ὁ ἕτερος τελῶνης.
 11a Ὁ φαρισαῖος σταθεῖς πρὸς ἑαυτὸν ταῦτα προσήχετο
 11b ὁ θεός,
 11c εὐχαριστῶ σοι
 11d ὅτι οὐκ εἰμὶ ὡσπερ οἱ λοιποὶ τῶν ἀνθρώπων
 11e ἄρπαγες,
 11f ἄδικοι,
 11g μοιχοί,
 11h ἢ καὶ ὡς οὗτος ὁ τελῶνης.
 12a νηστεύω δις τοῦ σαββάτου,
 12b ἀποδεκατῶ πάντα
 12c ὅσα κτῶμαι.
 13a ὁ δὲ τελῶνης μακρόθεν ἐστῶς οὐκ ἤθελεν οὐδὲ τοὺς ὀφθαλμοὺς
 ἐπάραι εἰς τὸν οὐρανόν,
 13b ἀλλ' ἔτυπεν τὸ στῆθος αὐτοῦ λέγων·
 13c ὁ θεός,
 13d ἰλάστητί μοι τῷ ἁμαρτωλῷ.
 14a λέγω ὑμῖν,
 14b κατέβη οὗτος δεδικαιωμένος εἰς τὸν οἶκον αὐτοῦ
 παρ' ἐκεῖνον·
 14c ὅτι πᾶς ὁ ὑψῶν ἑαυτὸν ταπεινωθήσεται,
 14d ὁ δὲ ταπεινῶν ἑαυτὸν ὑψωθήσεται.

3.3 Textbildungsprozess: Welche synchrone Schicht - jede inzwischen textkritisch überprüft - soll der späteren Beschreibung / Interpretation zugrundegelegt werden?

Literatur

- COXON: The Presentation of Authorship in Medieval German Narrative Literature 1220 - 1290. Oxford Modern Languages and Literatures Monographs. Oxford 2001 ISBN 0-19-8160178: Oxford University Press
- DIXON, D, P; MANNION: Goldsmith and the British Magazine. Literary and Linguistic Computing 13 (1998) 37-49.
- DONNER, H: Die literarische Gestalt der alttestamentlichen Josefsgeschichte. SHAW. PH 2. Heidelberg 1976 // I.2.7
- EISSFELDT, O: Hexateuch-Synopse. Die Erzählung der fünf Bücher Mose und des Buches Josua mit dem Anfange des Richterbuches in ihre vier Quellen zerlegt und in deutscher Übersetzung dargeboten samt einer in Einleitung und Anmerkungen gegebenen Begründung. Darmstadt. Leipzig 1962 (reprint von 1922)
- FERRER, D: Hypertextual Representation of Literary Working Papers. Literary & Linguistic Computing 10,2 (1995) 143-145.
- GREENWOOD, HH: Common Word Frequencies and Authorship in Luke's Gospel and Acts. Literary and Linguistic Computing 10 / 3 (1995) 183-187.
- HOLMES, D I; FORSYTH, R S: The Federalist Revisited: New Directions in Authorship Attribution. Literary & Linguistic Computing 10,2 (1995) 111-127.
- LEDGER, G; MERRIAM, T: Shakespeare, Fletcher, and the Two Noble Kinsmen. Literary & Linguistic Computing 9,3 (1994) 235-248.
- LEDGER, G: An Exploration of Differences in the Pauline Epistles using Multivariate Statistical Analysis. Literary & Linguistic Computing 10,2 (1995) 85-97.
- MEALAND, DL: Correspondence Analysis of Luke. Literary and Linguistic Computing 10 / 3 (1995) 171-182.
- MERRIAM, T: Heterogeneous Authorship in Early Shakespeare and the Problem of Henry V. Literary and Linguistic Computing 13 (1998) 15-28.

3.31 Probleme auch noch in textkritisch geklärten Texten?

zu Ziff. 3.311:

Literarkritik: Vollständigkeit / Zusätze / Auslassungen

DETTMERING, P (ed.), Die Kinder- und Hausmärchen der Brüder Grimm. Urfassung 1812/1814. Lindau o.J. S.319.	Kind- und Hausmärchen gesammelt durch die Brüder Grimm mit Zeichnungen von Otto Ubbelohde. Lizenzausgabe von 1922. it 112. Frankfurt/m 1975. ³⁶
--	--

=====	
Vorrede	Vorrede
1. Der Froschkönig oder der eiserne Heinrich	1. Der Froschkönig oder der eiserne Heinrich
2. Katz und Maus in Gesellschaft	2. Katze und Maus in Gesellschaft
3. Marienkind	3. Marienkind
4. Gut Kegel- und Kartenspiel	4. ?Märchen von einem der auszog, das Fürchten zu lernen
5. Der Wolf und die sieben jungen Geislein	5. Der Wolf und die sieben jungen Geißlein
6. Von der Nachtigall und der Blindschleiche	6. ???
7. Von dem gestohlenen Heller	7. ???
8. Die Hand mit dem Messer	8. ???
9. Die zwölf Brüder	9. Die zwölf Brüder
10. Das Lumpengesindel	10. Das Lumpengesindel
11. Brüderchen und Schwesterchen	11. Brüderchen und Schwesterchen
12. Rapunzel	12. Rapunzel
13. Die drei Männlein im Walde	13. Die drei Männlein im Walde
14. Von dem bösen Flachsspinnen	14. ?Die drei Spinnerinnen
15. Hänsel und Gretel	15. Hänsel und Gretel
16. Herr Fix und Fertig	16. ???
17. Die weiße Schlange	17. Die weiße Schlange
18. Strohalm, Kohle und Bohne auf der Reise	18. Strohalm, Kohle und Bohne ???
19. Von den Fischer und siine Fru	19. Von dem Fischer un? syner Fru
20. Von einem tapfern Schneider	20. Das tapfere Schneiderlein
21. Aschenputtel	21. Aschenputtel
22. Wie Kinder Schlachtens mit einander gespielt haben	22. ???
23. Von dem Mäuschen, Vögelchen und der Bratwurst	23. Von dem Mäuschen, Vögelchen und der Bratwurst
24. Frau Holle	24. Frau Holle
25. Die drei Raben	25. Die sieben Raben
26. Rothkäppchen	26. Rot?käppchen

³⁶ In Fettdruck sind die Abweichungen gegenüber der Urfassung markiert.

zu Ziff. 3.312:

Literarkritik: Vollständigkeit / Zusätze / Auslassungen

Federico García Lorca, Trece canciones españolas antiguas.

Lied: Los pelegrinitos.

zu: lectio brevis? Redaktion \triangleq Kürzungen (die unterstrichenen Strophen werden von der Notenausgabe ausgelassen)I) Notenausgabe
(gekürzt)II) DG Stereo 435 848-2
(komplett: I+II)

Hacia Roma caminan dos pelegrinos a que los case el Papa, mamita, porque son primos, niña bonita, porque son primos, niña.	Es wandern nach Rom zwei Pilger, damit der Papst sie verheirate, Mutter, denn sie sind Vettern, hübsches Mädchen, denn sie sind Vettern, Mädchen.
---	--

<u>Einen Hut aus Wachstuch</u> <u>trägt der Junge,</u> <u>und die kleine Pilgerin,</u> <u>Mutter,</u> <u>einen aus Samt,</u> <u>hübsches Mädchen,</u> <u>einen aus Samt,</u> <u>Mädchen.</u>	<u>Sombrerito de hule</u> <u>lleva el mozuelo</u> <u>y la pelegrinita,</u> <u>mamita,</u> <u>de terciopelo,</u> <u>niña bonita,</u> <u>de terciopelo,</u> <u>niña.</u>
---	---

<u>Als sie über die Brücke,</u> <u>die Brücke Victoria gingen,</u> <u>stolperte die Patin,</u> <u>Mutter,</u> <u>es fiel die Braut,</u> <u>hübsches Mädchen,</u> <u>es fiel die Braut,</u> <u>Mädchen.</u>	<u>Al pasar por el puente</u> <u>de la Victoria,</u> <u>tropezó la madrina,</u> <u>mamita,</u> <u>cayó la novia,</u> <u>niña bonita,</u> <u>cayó la novia,</u> <u>niña.</u>
---	--

Han llegado a Palacio suben arriba y en la sala del Papa, mamita, los desaminan, niña bonita, los desaminan, niña.	Sie stehn vorm Palast und gehen hinauf in den Saal des Papstes, Mutter, werden geprüft, hübsches Mädchen, werden geprüft, Mädchen.
---	---

<u>Der Papst hat sie gefragt</u> <u>wie sie hießen,</u> <u>er sagte: Peter,</u> <u>Mutter,</u> <u>und sie sagte: Anna,</u> <u>hübsches Mädchen,</u> <u>und sie sagte: Anna,</u> <u>Mädchen.</u>	<u>Le ha preguntado el Papa</u> <u>cómo se llaman.</u> <u>El le dice que Pedro,</u> <u>mamita,</u> <u>y ella que Ana,</u> <u>niña bonita,</u> <u>y ella que Ana,</u> <u>niña.</u>
--	--

<u>Der Papst hat sie gefragt,</u> <u>wie alt sie wären.</u> <u>Sie gibt zur Antwort: Fünfzehn,</u> <u>Mutter,</u> <u>er gibt zur Antwort: Siebzehn,</u> <u>hübsches Mädchen,</u> <u>er gibt zur Antwort: Siebzehn,</u>	<u>Le ha preguntado el Papa</u> <u>que qué edad tienen.</u> <u>Ella dice que quince,</u> <u>mamita,</u> <u>y él diecisiete,</u> <u>niña bonita,</u> <u>y él diecisiete,</u>
--	---

Mädchen.niña.

Le ha preguntado el Papa
de dónde eran.
Ella dice de Cabra,
mamita,
y él de Antequera,
niña bonita,
y él de Antequera,
niña.

Der Papst hat sie gefragt
woher sie kämen.
Sie sagte: aus Cabra,
Mutter,
und er aus Antequera,
hübsches Mädchen,
und er aus Antequera,
Mädchen.

Le ha preguntado Papa
que si han pecado.
El le dice que un beso,
mamita,
que le había dado,
niña bonita,
que le había dado,
niña.

Der Papst hat sie gefragt
ob sie gesündigt.
Er antwortet: Einen Kuß,
Mutter,
habe er ihr gegeben,
hübsches Mädchen,
habe er ihr gegeben,
Mädchen.

Y la pelegrinita
que es vergonzosa.
Se le ha puesto la Cara,
mamita,
como una rosa,
niña bonita,
como una rosa,
niña.

Und das Pilgermädchen
Schämig sei sie,
und im Gesicht sei sie,
Mutter,
wie eine Rose geworden,
hübsches Mädchen,
wie eine Rose geworden,
Mädchen.

<u>Und der Papst gab zur Antwort</u> <u>aus seiner Kammer:</u> <u>Ach, wär ich doch ein Pilger,</u> <u>Mutter,</u> <u>wie hier die beiden,</u> <u>hübsches Mädchen,</u> <u>wie hier die beiden,</u> <u>Mädchen.</u>	<u>Y ha respondido el Papa</u> <u>desde su cuarto:</u> <u>¡Quién fuera pelegrino,</u> <u>mamita,</u> <u>para otro tanto,</u> <u>niña bonita,</u> <u>para otro tanto,</u> <u>niña!</u>
--	--

Las campanas de Roma
ya repicaron,
porque los Pelegrinos,
mamita,
ya se casaron,
niña bonita,
ya se casaron,
niña.

Die Glocken von Rom
läuteten nun,
denn die Pilger,
Mutter,
sind schon verheiratet,
hübsches Mädchen,
sind schon verheiratet,
Mädchen.

zu Ziff. 3.313:

Mehrere Texte in einem?

zu:

aus: Sintfluterzählung, Gen 8

(1) Da dachte Gott an Noach und an alle Tiere und an alles Vieh, das bei ihm in der Arche war. Gott ließ einen Wind über die Erde wehen, und das Wasser sank. (2) Die Quellen der Urflut und die Schleusen des Himmels schlossen sich; der Regen vom Himmel ließ nach, (3) und das Wasser verlief sich allmählich von der Erde. So nahm das Wasser nach hundertfünfzig Tagen ab. (4) Am siebzehnten Tag des siebten Monats setzte die Arche im Gebirge Ararat auf. (5) Das Wasser nahm immer mehr ab, bis zum zehnten Monat. Am ersten Tag des zehnten Monats wurden die Berggipfel sichtbar. (6) Nach vierzig Tagen öffnete Noach das Fenster der Arche, das er gemacht hatte, (7) und ließ einen Raben hinaus. Der flog aus und ein, bis das Wasser auf der Erde vertrocknet war. (8) Dann ließ er eine Taube hinaus, um zu sehen, ob das Wasser auf der Erde abgenommen habe. (9) Die Taube fand keinen Halt für ihre Füße und kehrte zu ihm in die Arche zurück, weil über der ganzen Erde noch Wasser stand. Er streckte seine Hand aus und nahm die Taube wieder zu sich in die Arche. (10) Dann wartete er noch weitere sieben Tage und ließ wieder die Taube aus der Arche. (11) Gegen Abend kam die Taube zu ihm zurück, und siehe da: In ihrem Schnabel hatte sie einen frischen Olivenzweig. Jetzt wußte Noach, daß nur noch wenig Wasser auf der Erde stand. (12) Er wartete weitere sieben Tage und ließ die Taube noch einmal hinaus. Nun kehrte sie nicht mehr zu ihm zurück. (13) Im sechshundertersten Jahr Noachs, am ersten Tag des ersten Monats, hatte sich das Wasser verlaufen. Da entfernte Noach das Verdeck der Arche, blickte hinaus, und siehe: Die Erdoberfläche war trocken. Am siebenundzwanzigsten Tag des zweiten Monats war die Erde trocken. ...

3.32 Soziolinguistische Informationen

zu Ziff. 3.321:

Forensische Linguistik - Autorenerkennung

zu: vprag7.5122

aus: HÄNLEIN, H, Studies in authorship recognition - a corpus-based approach. EHS XIV,352. Frankfurt/M 1999.

(17f) Another equally notorious case in which linguistic expertise was involved is that of Derek Bentley in 1952. Coulthard describes how linguists came to the conclusion that the record of the witness's statement had been forged.

Cf. Coulthard 1992b:245, 251f. and Coulthard 1993:88f: Derek Bentley, a mentally retarded 19-year-old, was sentenced to death and executed for the murder of a policeman who had been shot by Chris Craig, an underage friend of Bentley's. Bentley was the last man to be hanged in Britain. The verdict turned out to be wrong, but no posthumous pardon was granted. Before his death, Derek Bentley had repeatedly insisted that the police officers had "helped" him with his statement, but the policemen unanimously denied this. The expert witnesses involved in the case after Derek Bentley's death first looked at two corpora: one consisted of the policemen's statements on how the murder had come about and the other contained ordinary statements by other witnesses in other cases. High frequency of temporal *then*, especially the postponement of temporal *then* behind the subject (*I then...*) turned out to be typical feature of "policeseek", whereas it occurred extremely rarely in the reference corpus. Since temporal and postponed *then* was very frequent in the record of Derek Bentley's confession the expert witnesses concluded that it had been forged. Another surprising feature in Bentley's confession is the high frequency of negatives, which is untypical in normal narrative. The expert testimony yielded the following explanation for this: the policemen had asked Bentley questions, which he answered in the negative. These exchanges had then been incorporated into the record of the statement, which should have been a verbatim transcription of an unprompted monologue. Furthermore, the case involved the interpretation of an ambiguous utterance made by Bentley and addressed to his friend Chris Craig immediately before the latter shot the policeman: the phrase *let him have it* was interpreted by the prosecution as 'shoot him' while the defence alleged that it was meant to be 'give it (the gun) to him'.

zu Ziff. 3.322:

Strukturiertes Text-Korpus: Identifikation der Einzeltexte

zu: vprag7.5112

aus: MEURMAN-SOLIN, A, Structured Text Corpora in the Study of Language Variation and Change: Literary and Linguistic Computing, Vol. 16, No.1 (2001) 5-27.

(11f) In fact, under their leadership, a research team has now completed the *Corpus of Early English Correspondence (CEEC)*, which comprises 2.7 million words of running text. In the CEEC there are almost 800 informants over the period 1417-1681. This single-genre corpus is structured by sociolinguistically relevant variables such as the writer's social rank, gender, mobility, age, etc. The corpus provides evidence for what can be considered a relatively new field of study: historical sociolinguistics. The following set of 27 variables coded into the CEEC structures the language-external information about each letter and its writer (Table 3.2 of Nevalainen and Raumolin-Brunberg, 1996, p.50):

1. Last name
2. First name
3. Title
4. Year of birth
5. Year of death
6. First letter
7. Last letter
8. Sex
9. Rank
10. Father's rank
11. Social mobility
12. Place of birth
13. Main domicile
14. Migrant
15. Education
16. Religion
17. Number of letters
18. Number of recipients
19. Kind of recipients
20. Number of words
21. Letter of contents
22. Letter quality
23. Collection
24. Career
25. Migration history
26. Extra
27. Complete

(12) In addition to this method of providing detailed sociolinguistically relevant information in a single-genre corpus, the compilers have also introduced a system that explicitly states the extent to which and how the data included can be considered reliable. The letters and their writers are categorized into the following groups (Nevalainen and Raumolin-Brunberg, 1996, p.43):

A = autograph letter in a good original-spelling edition; writer's social background recoverable;

B = autograph letter in a good original-spelling edition; part of the writer's background information missing;

C = non-autograph letter (secretarial work or copy) in a good original-spelling edition; writer's social background recoverable;

D = doubtful or uncertain authorship; problems with the edition, the writer's background information, or both.

3.4 Theorie und Praxis der Methode zur Identifizierung von Zusätzen (Literarkritik)

Literatur

- HÄNLEIN, H: Studies in Authorship Recognition. A Corpus-based Approach. EHS 14: Anglo-Saxon Language and Literatur 352. Frankfurt / M 1999: Peter Lang // I.2.5
- HARK, H: Der Gevatter Tod. Ein Pate fürs Leben. Stuttgart 1986
- KUHL, C: Die »Wiederaufnahme« ein literarkritisches Prinzip?. ZAW 64 (1952) 1-11.
- SCHWEIZER, H: Literarkritischer Versuch zur Erzählung von Michaben Jimla (1 Kön 22). BZ NF (1979) 1-19.
- SCHWEIZER, H: Der Sturz des Weltenbaumes (Ez 31) - literarkritisch betrachtet. Theologische Quartalschrift 165 (1985) 197-213.
- SCHWEIZER, H: Fragen zur Literarkritik von Gen 50. Diskussionsbeitrag zu R. Bartelmus BN 29 (1985) 35-53. BN 36 (1987) 64-68.
- SCHWEIZER, H: Literarkritik. Theologische Quartalschrift 168 (1988) 23-43. Abk: (1988 Lkr)
- SCHWEIZER, H: Leckere Häppchen oder dicke Mehlsoße?. Replik zu: Burkhard R. Knipping, Textwahrnehmung 'häppchenweise'. Bemerkungen zu Harald Schweizers "Die Josefsgeschichte" und zu seiner Literarkritik: BN 62 (1992) 61-95: BN 63 (1992) 52-57.
- SCHWEIZER, H: Angst vor Wahrnehmung solo?. Zu: Objektive Ergebnisse bei textinterner Literarkritik. Einige Anmerkungen zur Subjektivität literarkritischer Beobachtungen in Harald Schweizers Studie "Die Josefsgeschichte" von Bernd Willmes BN 67 (1993) 54-86: BN 69 (1993) 24-28.
- SCHWEIZER, H: Weitere Impulse zur Literarkritik. Biblische Notizen 80 (1995) 73-99.
- SCHWEIZER, H: Verwässerung der Texte (Redaktoren) in: SCHWEIZER, H (ed./eds.): "... deine Sprache verrät dich!" Grundkurs Religiosität. Essays zur Sprachkritik. Forum Religionskritik 1. Münster 2002 147-163. 3-8258-5869-3: LIT Verlag // I.2.7
- THEOBALD, M: Der Primat der Synchronie vor der Diachronie als Grundaxiom der Literarkritik. Methodische Erwägungen an Hand von Mk 2,13-17 / Mt 9,9-13. BZ NF 22 (1978) 161-186.
- WAHL, H: Empathie und Text. Das selbstpsychologische Modell interaktiver Texthermeneutik. Theologische Quartalschrift 169 (1989) 201-222.
- WERLITZ, J: Studien zur literarkritischen Methode. Gericht und Heil in Jesaja 7,1-17 und 29,1-8. BZAW 204. Berlin 1992: de Gruyter // I.2.1

zu Ziff. 3.41:

Literarkritische Problemstellung in essayhafter Form

zu:

Harald Schweizer

Vgl.

<http://www-ct.informatik.uni-tuebingen.de/daten/pattl.ps> bzw.
[../pattl.pdf](http://www-ct.informatik.uni-tuebingen.de/daten/pattl.pdf)

In dem Essay wird auf S. 1+2 in bildhafter Form (Kuchen + Additionen) erläutert, worum es in der Literarkritik (= "Textarchäologie") geht.

Vgl. für Literarkritik, aber auch alle weiteren Punkte im Kapitel "Konstituierung des Textes" Materialien / Illustrationen auf der homepage:

<http://www-ct.informatik.uni-tuebingen.de/ct/syntax.htm>

zu Ziff. 3.42:

GRASS, G, Danziger Trilogie. Die Blechtrommel. Katz und Maus. Hundejahre. Darmstadt 1980. [Die Blechtrommel]

zu:

Stichwort: LkrB/lkr Bruch vs. stilist. Erscheinung

(366f) "Als Ende Mai meine Großmutter Anna Koljaiczek zu Fuß aus Bissau über Brenntau nach Langfuhr kam, uns besuchte und sich schwer atmend auf die Chaiselongue warf, lobte der Herr Fajngold das Kurtchen sehr und fand auch für Maria lobende Worte. Als er meiner Großmutter lang und breit die Geschichte meiner Krankheit erzählte, dabei immer wieder auf die Nützlichkeit seiner Desinfektionsmittel hinwies, fand er auch Oskar lobenswert, weil ich so still und brav gewesen, während der ganzen Krankheit nie geschrien habe."

zu Ziff. 3.421:

Aufhebung fixierter Erzählperspektive

zu: vprag8.244

aus: J. SEELHORST, Autoreferentialität und Transformation. Zur Funktion mystischen Sprechens bei Mechthild von Magdeburg, Meister Eckhart und Heinrich Seuse. Tübingen 2003.

(132) Mit der Ich-Form rekurriert Mechthild zum einen auf die Ebene personaler Erfahrung, wodurch sie formal den Eindruck von Authentizität und biographischer Saturiertheit vermittelt, zum andern eröffnet sie damit die Möglichkeit einer identifikatorischen Lektüre im Sinne eines allgemein menschlichen Ichs. Auch die unterschiedlichen Figurationen der dritten Person bieten Gelegenheit zur Identifikation: ein Mensch, eine Seele; das heißt, der Perspektivenwechsel intendiert eine Öffnung der Textwelt auf die Lesenden hin, denen verschiedene Rollen angeboten werden, eigenes Verstehen und eigene Erfahrungswirklichkeit mit dem Text zu verknüpfen.

Auf der anderen Seite stört die Aufhebung einer einheitlichen Perspektivierung den Lektüreprozeß, zumal das Changieren zwischen erster und dritter Person im einzelnen sich weder inhaltlich noch formal aus dem Kontext motivieren läßt. In ihrer Grundlosigkeit wirken die Perspektivenwechsel verunklarend, indem sie die Position der Erzählerin in der Schwebelage halten. Ebenso wie der Wechsel der Sprechweisen lenken sie die Aufmerksamkeit auf die Konstitution des Textes und unterminieren so die durch sie eröffnete Möglichkeit einer identifikatorischen Lektüre. Die Auflösung fixierter Rollenidentitäten fungiert somit zugleich als ein Medium der Vermittlung und als eine Weise, die Differenz zwischen Textwelt und Leser durch die autoreferentielle Betonung der Medialität bewußt zu halten.

zu Ziff. 3.43:

Revidierte Literarkritik

zu:

SCHWEIZER, H, Literarkritik. ThQ 168 (1988) 23-43.

RABE, N, Vom Gerücht zum Gericht. Revidierte Text- und Literarkritik der Kundschaftererzählung Numeri 13.14 als Neuansatz in der Pentateuchforschung. THLI 8. Tübingen 1994.

Lkr 1. Stufe: Sammlung von Leseauffälligkeiten

RABE 224: "Sämtliche tatsächlichen oder möglichen Leseauffälligkeiten sind aufzulisten, jedoch noch nicht zu bewerten. Jede Leseauffälligkeit benötigt zwei Pole, deren Zueinander sich als problematisch erweist. Diese Pole sind inhaltlich zu beschreiben und mit den Bezeichnungen der betreffenden Äe im Text zu verankern - ein Postulat der Nachvollziehbarkeit und Präzision. Ferner erweist es sich als hilfreich, die Leseauffälligkeiten zu klassifizieren. Zum einen lassen sich Auffälligkeiten und Störungen im Lesevorgang konstatieren:

- a. 'Syntaktischer Bruch': Ein morphologisch vorhandenes Subjekt, ein Pronomen läßt sich nicht identifizieren oder auf genannte Größen beziehen.
- b. 'Inhaltliche Spannung': Zwei Daten widersprechen sich.
- c. 'Terminologische Differenz': Ein und dieselbe Sache oder Person wird verschieden bezeichnet.
- d. 'Terminologische Indifferenz': Zu trennende Sachen, Sachverhalte oder Personen werden begrifflich zusammengefaßt.
- e. 'Mehrfachnennung': Wiederholung von Wörtern, Sätzen usw.
- f. 'Informationsdefizit': Angaben des Textes überfordern (im Augenblick) den Leser.

Zum andern sollen 'stilistische Wechsel' notiert werden, wie der Wechsel vom plastisch-beschreibenden zum abstrakt-wertenden Stil, vom knappen zum ausladenden Stil, von wörtlicher Rede zur Handlungsdarstellung, von Prosa zu Poesie, von der Wiederholung eines Gedankens zu seiner Weiterführung, Tempusänderungen usw. Diese Variationen sind in Texten üblich, können aber im Zusammentreffen mit den Leseauffälligkeiten a.-f. als untergeordnete literarkritische Indizien bisweilen Textstellen markieren, an denen sich leicht Sekundäres anschließen läßt."

Lkr 2. Stufe: Minimale Leseeinheiten (MLEen)

RABE 224: "Es werden diejenigen Textbereiche angegeben, in denen keine Leseauffälligkeit (mit *beiden* Polen! - H.S.) im unter 1. beschriebenen Sinn vorliegt. Sobald eine Leseauffälligkeit (mit ihrem zweiten Pol - H.S.) auftritt, muß diese fraglos verstehbare 'minimale Leseinheit' (vorher, H.S.) abgebrochen werden."

Lkr 3. Stufe: Diskussion der Leseauffälligkeiten / Erarbeitung der Teiltexthe (TT)

RABE 225: "Der Text wird von vorn gelesen, indem das Verhältnis der ersten usw. minimalen Leseeinheit zur jeweils folgenden überprüft wird. Dabei sind alle unter 1. gesammelten Leseauffälligkeiten zu gewichten - als stilistisch-pointierte, als

im Kontext funktionierende bzw. wenigstens akzeptable Textphänomene oder als Verstehensprobleme. Auf diese Weise lassen sich literarkritische Brüche zwischen minimalen Leseeinheiten identifizieren und 'Teiltex-te' ... bestimmen ... Zu unterstellen ist, daß jemand, der in ein so vielschichtiges Gebilde, wie es ein Text darstellt, noch erkennbar eingegriffen hat, in der Regel nicht nur ein einziges Indiz dafür hinterlassen wird, sondern mehrere Ungereimtheiten, Unpäßlichkeiten der Grammatik und der Sprachstruktur. Daher sollten zur Konstatierung eines literarkritischen Bruches mindestens zwei nach der Diskussion etablierte Verstehensprobleme angeführt werden."

Lkr 4. Stufe: Verhältnisbestimmung von Teiltex-ten

nach SCHWEIZER: Bis jetzt ist nachgewiesen, daß unmittelbar benachbarte Teiltex-te sich ausschließen. Man nimmt nun - immer vom Textanfang her - weiter entfernte TTe und prüft sie ebenfalls nach den Kriterien von 1. durch: wirken zwischen ihnen ebenfalls Auffälligkeiten, die einen textlichen Anschluß verhindern? Oder entsteht durch das Aneinanderkoppeln ein problemlos lesbarer, störungsfreier Text?

Lkr 5. Stufe: Gegenkontrolle

nach SCHWEIZER: Man vergewissert sich, ob alle unter 1. gesammelten literarkritischen Beobachtungen auf einer der Ebenen zuvor aufgegriffen wurden und eine zufriedenstellende Erklärung gefunden haben.

zu Ziff. 3.44:

Literarkritik in 5 Stufen³⁷

zu:

(I) Beobachtung möglicher Verstehensprobleme

T e x t in Äußerungseinheiten

(II) Minimale Leseeinheiten

AAAABBBBCCCCCCCCCCCCDDEEEEEEEEEEEEEFFFFFGHHHHHHIIIIIIJJJKKKKLL

(III) Teiltex-te

AAAABBBB	CCCCCCCCCCCC	DDEEEEEEEEEEEEEFFFF	GHHHH	IIIIIII...
----------	--------------	---------------------	-------	------------

(IV) Schichten

AAAABBBB DDEEEEEEEEEEEEEFFFF

CCCCCCCCCCCC

GHHHH IIIIIII

(V) Gegenkontrolle

Sind alle Beobachtungen aus (I) berücksichtigt?

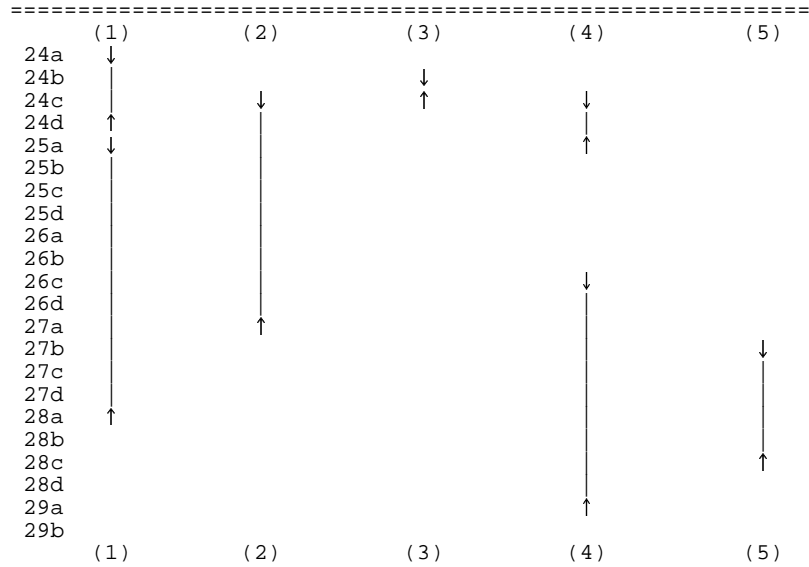
³⁷ Vgl. Schweizer, H, Literarkritik. Theologische Quartalschrift 168 (1988) 23-43. - Ders., Weitere Impulse zur Literarkritik: Biblische Notizen 80 (1995) 73-99.

zu Ziff. 3.45:

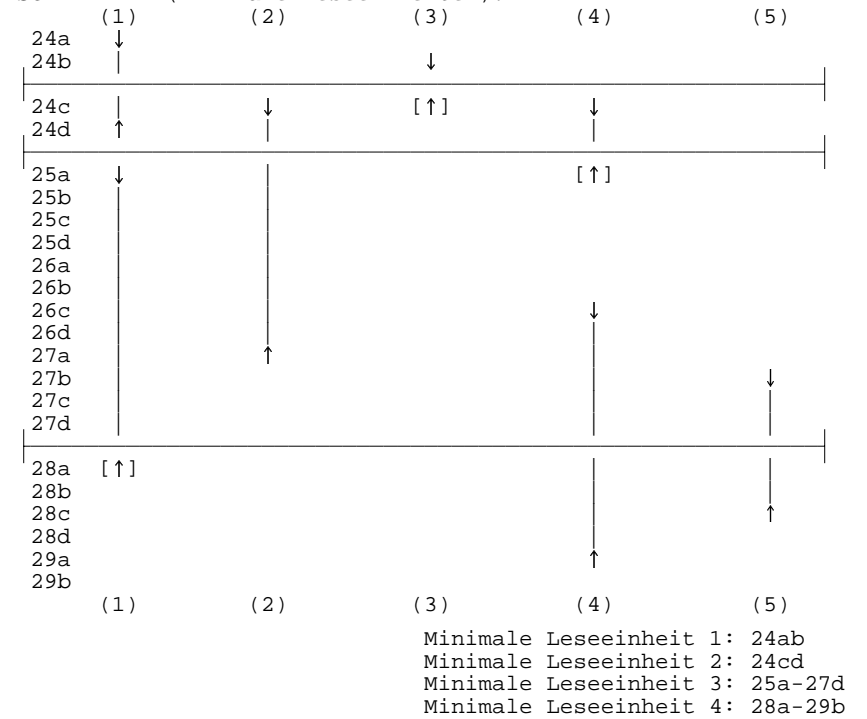
Literarkritik: Operationalisierung

zu:

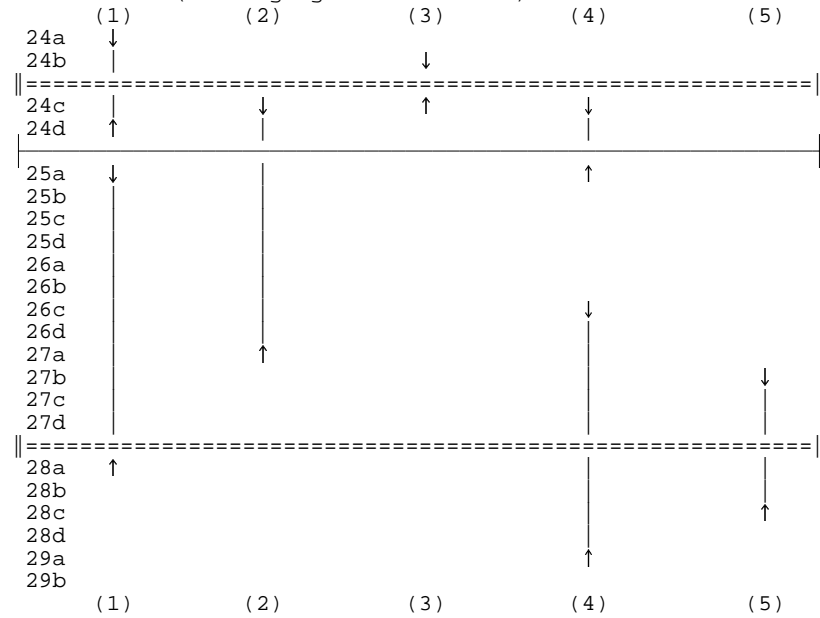
SCHRITT I (Sammeln von Beobachtungen)



SCHRITT II (Minimale Leseeinheiten).



SCHRITT III (Festlegung der Teiltexte).



Teiltext 1: 24ab =MLE1
 Teiltext 2: 24c-27d=MLE2+3
 Teiltext 3: 28a-29b=MLE4

SCHRITT IV (Suche nach Schichten bei nicht-benachbarten Teiltexten).

24a
 24b
 28a
 28b
 28c
 28d
 29a
 29b

SCHRITT V (Kontrolle: Sind alle Beobachtungen von Schritt I verarbeitet?).

41::lo2:									↑
42::lp2:									↑
43::lq2:									↑
44::lr2:									↑
45::ls2:									↑
46::lt2:									↑
47::lu2:									↑
48::lv2:									↑
49::lw2:									↑
50::lx2:									↑
51::ly2:									↑
52::lz2:									↑
53::la3:									↑
54::lb3:									↑
55::lc3:									↑
56::ld3:									↑
57::le3:									↑
58::lf3:									↑
59::lg3:									↑
60::lh3:									↑
61::li3:									↑
62::lj3:									↑
63::lk3:									↑
64::ll3:									↑
65::lm3:									↑
66::ln3:									↑
LkrB:	1	2	3	4	5	6	7	8	9

LkrB

1. Inh. Spannung: Im Dialog kommt der Mann zu einem negativen Gottesbild. Dieser Wertung wird anschließend eine positive entgegengesetzt: lw-la2 vs. ld2.

2. Inh. Spannung: Die positive Wertung hat im bisherigen Textverlauf keinen Rückhalt, hängt in der Luft. Die negative dagegen ist narrativ vorbereitet: la-la2 vs. ld2.

3. Spannung in der Wortwahl: Abstraktionen ("Armut", "Reichtum") sind neu vor dem Hintergrund bisheriger anschaulicher Sprache: la-la2 vs. ld2.

4. Spannung bei den Akteuren: Angegriffen muß sich der Dialogpartner "Gott" fühlen. Von dem aber wird keine Reaktion berichtet. In diese Erzähllücke springt dagegen ein Kommentator ein, der damit - wäre er mit dem Erzähler identisch - die Kühnheit des Mannes unterläuft und entwertet: lj-la2 vs. lb2-ld2.

5. Terminol. Differenz: Von "Gott" ist zunächst immer mit Artikel und Adjektiv "der liebe" die Rede. Im Kommentar: ohne Artikel und Adjektiv: ll+lw vs. ld2.

- - - - -

6. Syntaktischer Bruch: "Also wendete er sich..." Das zuletzt genannte Subjekt war "Gott". Es muß jetzt aber wieder "der Mann" als Subjekt gedacht werden: ld2 vs. le2.

7. Inh. Spannung: "Also..." - woran schließt dies an? Sicher nicht an den Kommentar. Zu dem ist keine gedankliche Brücke formuliert (z.B. "in seiner Verblendung" oder: "verstockt"): lb2-ld2 vs. le2.

8. Terminol. Differenz: In Dialog und Kommentar war von "Gott" die Rede (vgl. oben die LkrB 5); in der Fortführung: "von dem Herrn": ll+lw+ld2 vs. le2.

9. Inh. Spannung: Nach der Abstraktion nun wieder Rückkehr in die konkret-anschauliche Erzählfiktion: ld2 vs. le2-ln3.

zu Ziff. 3.47:

Inhaltsfreie Autoridentifikation + Literarkritik

zu:

aus: T. MERRIAM, *Kind John* Divided: Literary and Linguistic Computing 19/2 (2004) 181-195.

Verfahren, die in Ziff. 4.4(Ausdrucks-Syntax) behandelt werden, können die vorgestellte Literarkritik nicht ersetzen, aber unterstützen. Vgl. aus gen. Aufsatz die Versuchsanordnung:

(181) A particular set of function words may be selected for their ability to discriminate between two suspected authors. The relative frequencies of *and, I, is, it, of, their, with* and *you*, words constituting the 2nd, 3rd, 10th, 14th, 11th, 5th, 35th, 12th, and 7th most frequently used words by Shakespeare respectively, provide one such set. From a previous study, eight of the nine (all excepting *I*) were shown to discriminate Shakespeare and Marlowe ...

Die Statistik teilt *Kind John* in 2 Teile, der eine zweifelsfrei von Shakespeare, der andere von anderer Hand, höchstwahrscheinlich von Marlowe. Aber das müsste durch weitere und andersartige Untersuchungen erhärtet werden.

3.5 Übersicht über gängige Motive und Techniken von Interpolatoren/Redaktoren

Literatur

GROEBEN, N; CHRISTMANN, U: Textoptimierung unter Verständlichkeitsperspektive in: ANTOS, G; KRINGS, HP (ed./eds.): Textproduktion. Ein interdisziplinärer Forschungsüberblick. Konzepte der Sprach- und Literaturwissenschaft 48. Tübingen 1989 165-196.: Niemeyer // I.2.0

SCHWEIZER, H: Weitere Impulse zur Literarkritik. Biblische Notizen 80 (1995) 73-99.

WONNEBERGER, R: Redaktion. Studien zur Textfortschreibung im Alten Testament, entwickelt am Beispiel der Samuel-Überlieferung. FRLANT 156. Göttingen 1992: Vandenhoeck & Ruprecht // I.2.1

zu Ziff. 3.51:

Motive für redaktionelle Eingriffe / Zusammenfassung

zu:

1. Tötung der Erzählspannung durch Vorgriffe.
2. Zerdehnung durch Auffüllen von Erzähllücken.
3. Schmarotzender Mißbrauch durch Zusatzakzente und Korrekturen.
4. Der Fluch der bösen Tat: redaktioneller Zugzwang.
5. Emotionale Nivellierung. Provozierende Aussagen werden umgebogen, sind damit weniger kreativ, weniger anstößig.
6. Gewaltsame Integrierung der religiösen Ebene = Mißachtung der Erzählfiktion.
7. Oberlehrerhafte Verstärkungen, falsche Dramatisierungen.
8. Nicht-Aushalten poetischer Bildhaftigkeit und Provokation.
9. Ideologische Korrekturen (wenn z.B. das Gute nicht siegt; wenn sozial bestimmende Institutionen - Kirche, Staat, Familie - attackiert werden).
10. Vermeidung von angsterregenden Schilderungen (physische Grausamkeit, Rache usw.) zur vorgeblichen "Schonung" der LeserInnen.

zu Ziff. 3.52:

Motive und Effekte der sekundären Bearbeitungen
der ursprünglichen Josefsgeschichte / Beispiele

zu:

[mit "TT" sind nachfolgend literarkritisch herauspräparierte "Teiltex-te" gemeint. Was sich konkret als Text hinter den Etiketten verbirgt, kann bequem nachgeschlagen werden:

H. Schweizer, Die Josefsgeschichte. Konstitutierung des Textes. Teil II: Textband. THLI 4/ii. Tübingen 1991. - Vgl. die Textversion ab S. 83ff]

1. Tötung der Erzählspannung durch Vorgriffe.

TT7 + TT9: Erst aufgrund von TT8 + TT10 (träumen, Bericht vom Traum) ergibt sich erzählerisch die Konfrontation.

TT23 + TT25 + TT27: Dreimal wird massiv die Unterstützung Josefs durch Jahwe genannt, bevor die Verführungsszene beginnt: der Leser ist zur Genüge gewappnet, um die Befürchtung, Josef könne nicht widerstehen, klein zu halten.

TT89: Nimmt bewertend vorweg, was in TT90A + TT91 erst allmählich entwickelt wird. Zudem Pathos: "der Mächtige", "s. verneigen zur Erde", "Hartes".

TT101: Laut urspr. JG (TT102) sind sowohl Akteur im Text und Leser überrascht über das Silber im Sack. Der, der TT101 einfügt, nimmt dem Leser die Überraschung, stattet ihn mit überlegenem Wissen aus. Das Leseerlebnis wird seiner urspr. Dramatik beraubt.

TT117: Der gute Ausgang der anvisierten Wiederbegegnung mit Josef wird vorab schon angedeutet. Dadurch ist der Risikocharakter, das Ungewisse, weggeblasen.

TT128: wie TT117

TT236: Bevor Josef und die Brüder in TT238A selbst das klärende Gespräch führen, wird in Form einer neu geschaffenen Botensituation und durch Abstrakta ("Schuld, Sünde, Böses") die Tendenz des folgenden Gesprächs vorweggenommen.

2. Zerdehnung durch Auffüllen von Erzähllücken.

TT42: Die urspr. JG sagt nur, daß der Pharao zornig war auf seine beiden Hofbeamten. Ein Grund wird nicht genannt. Diese Lücke füllt TT42 aus, kann es aber auch nur sehr abstrakt ("verfehlten sich").

TT69: Die Amtseinsetzung Josefs (TT68) war einem Bearbeiter zu wenig feierlich und prunkvoll. Die entsprechende Ausmalung holt er nach.

TT77: Die männerzentrierte Grundschrift wird um Frau + Kinder ergänzt. Zugleich wird die weitgehend profane Grundschrift um fromme Danknamen ergänzt. Josef zieht anhand der Kinder ein religiöses Zwischenfazit.

TT98: , der sich urspr. nur auf TT96 beziehen konnte, stellt sicher, daß Josef das Schuldeingeständnis wahrnimmt. Der Bearbeiter will auf dieser Grundlage die weitere Vorgehensweise Josefs verständlich machen. Dies ist aber völlig überflüssig, da i.S.d. urspr. JG Josef reichlich Gelgenheit hatte, sich ein Urteil zu den Brüdern zu bilden.

TT107: In TT106A waren 3 der Brüder namenlos erwähnt worden (vgl. 42,32cd.33d). TT107 präzisiert und trägt nach: Josef, Simeon, Benjamin.

TT124-129: Obwohl nicht aus einem Guß, so wird die Situation von 43,17c (a) durch Gefühl und Selbstreflexion, (b) durch ein Gespräch mit diesem Mann aufgefüllt. 43,24a greift wieder den Faden von 43,17c auf: Ende der Zusätze.

TT197: Kommt dem Bedürfnis nach Zeitdauer-Angaben nach.

TT200: Die verhaltene Andeutung des Sterbens Israels (47,31d) wird ausgenutzt, ihn kurzfristig wiederaufleben zu lassen: dadurch lassen sich noch einige nachträgliche Akzente einbauen.

TT201: Analog zu TT200 (aber von anderer Hand).

TT209: Die ohnehin im Text entwickelte Vorstellung von der Gebrechlichkeit Jakob/Israels wird - im übrigen Text funktionslos - konkretisiert durch Hinweis auf seine Beinahe-Blindheit.

TT231B: Eine im Text funktionslose Namensätiologie wird eingefügt.

TT232+234: Da im Grundtext nur Josef den Vater in Kanaan begräbt, stellen die beiden TT sicher, daß auch die anderen Söhne (und andere Begleiter) am Begräbnis teilnehmen.

TT239-244: An den Schluß der urspr. JG in TT238A, wo die innere Aussöhnung betont wird, hängen mehrere Bearbeiter noch einen äußeren Ausklang der Erzählung, weitergeführt bis zum Tod Josefs.

3. Schmarotzender Mißbrauch durch Zusatzakzente und Korrekturen.

TT193 + TT195: Ausnahmeregelung der Priester über Landbesitz.

TT220: War z.B. durch TT201 Israel kurzfristig wieder zu Kräften gekommen, so muß wieder, nun definitiv, zu seinem Sterben übergeleitet werden. Dies geschieht durch Umbiegen der Aussageabsicht der urspr. JG: dort wird für die neue Lebensmöglichkeit in Ägypten geworben. Josef und die Brüder haben keinen Drang mehr zurück nach Kanaan. Darauf reagiert TT220 mit der gegenläufigen Tendenz, unterstrichen durch die Nennung Gottes.

TT200+205+212+215+218+219: Gemeinsam ist den TT, daß Efraim und Manasse eine Sonderstellung eingeräumt wird, sie werden besonders gesegnet und vor Gott ausgezeichnet. Die beiden haben aber keine einheitliche Lobby, weil z.T. Efraim und Manasse als gleichberechtigt angesehen werden, z.T. wird Efraim ein Vorrang eingeräumt (TT216.219).

TT226: Während im Grundtext (TT225) Josef pauschal zum "Haus des Pharao" spricht, macht TT226 einen Botenauftrag daraus. Damit wird Josefs Souveränität reduziert. Er wird stärker als Untergebener gesehen.

4. Der Fluch der bösen Tat: redaktioneller Zugzwang.

TT55: Wenn schon durch TT51 nachträglich dafür gesorgt worden war, daß der Pharao ein zweites Mal träumt, so muß natürlich auch ein zweiter Traumbericht Josef gegenüber eingefügt werden.

TT57: Wenn zuvor das Träumen Pharaos verdoppelt worden war, so ist der Hinweis sinnvoll und notwendig, daß es sich letztlich doch nur um einen Traum handelt. Nachdem TT57 zuvor isoliert worden war, kann nun gefolgert werden, daß durch TT57 der Bearbeiter implizit zugibt, daß das einmalige Träumen ursprünglich ist. Es fehlt ihm auch der lange Atem, seinen Zusatz (Traum von den Ähren) erzählerisch weitgreifend in den Text zu integrieren.

TT209: Läßt man Israel in 47,31d nicht sterben, sondern belebt ihn kurzfristig wieder (TT200+201), dann ergibt sich die Notwendigkeit, zwischendurch auf seine Gebrechlichkeit hinzuweisen.

5. Emotionale Nivellierung.

TT155: "seid nicht erregt auf dem Weg". Warum nicht? Und wodurch ist diese Mahnung veranlaßt?

6. Gewaltsame Integrierung der religiösen Ebene = Mißachtung der Erzählfiktion.

TT23: a) die Offenheit und das Risiko in Josefs Schicksal (TT22B/TT22A) werden nicht ausgehalten sondern beschwichtigend kommentiert ("Mann des Gelingens"). b) Josefs Geschichte wird in religiösen Hintergrund hineingestellt. Das "Mitsein Jahwes" befreit ebenfalls von Angst.

TT38: Josef am Tiefpunkt seiner Karriere. Der wird vom Bearbeiter emotional nicht ausgehalten. Stattdessen wird ein positiv-freundlicher Akzent beigefügt. Motto: Alles nicht so schlimm.

TT94: Völlig unmotiviert im Rahmen der Handlungsanweisungen an die Brüder beteuert Josef seine Gottesfurcht. Ein Bearbeiter wollte offenbar sicherstellen, daß Gott bei der positiven Problembewältigung mitwirkt. Die urspr. JG kommt an dieser Stelle ohne ihn aus.

TT117: Angesichts der angstbesetzten Rückreise nach Ägypten führt ein Bearbeiter den Vertrauen vermittelnden Lenker der Geschichte ein: EL SCHADDAI. Der Rückgriff auf das Übernatürliche mindert die aktuelle Unsicherheit. Gott als Beruhigungsspielle?³⁸

TT131: Josef mit frommem Gruß an Benjamin. Dagegen kommt im Kontakt mit den Brüdern (TT130) Josef ganz ohne Verweis auf Gott aus.

TT143: Was geschehen ist, entspringt alles einem weisen Plan Gottes. Wieder wird dadurch die Erzählfiktion zerstört.

TT160: Was geschehen wird/soll, entspricht einem Plan/einer Verheißung Gottes.

TT202: EL SCHADDAI als aus der Vergangenheit heraus das Schicksal von Jakob und seiner Sippe auch in Zukunft zum Guten lenkender Faktor wird eingeführt.

TT242: Mit Hilfe Gottes wird redaktionell über die aktuelle Erzählung hinaus auf den nächsten thematischen Erzählkomplex geschaut. Das Thema "Heraufführung aus Ägypten" hat in der urspr. JG keinen Anhaltspunkt.

7. Oberlehrerhafte Verstärkungen, falsche Dramatisierungen.

TT20: Die Trauer in 37,34c reichte anscheinend noch nicht. TT20 addiert eine Kitsch-Dramatik. Erzählerisch-emotional: Holzhammer.

TT51: Der eine Traum Pharaos wird um einen zweiten ergänzt. Dem Bild vom Nil wird ein rein agrarisches zur Seite gestellt, willkürlich, wohl deshalb, um eine Brücke zum späteren

³⁸ Umberto ECO, Das Foucaultsche Pendel. München ¹⁷2005. S. 412: "Die Menschheit kann den Gedanken nicht ertragen, daß die Welt per Zufall entstanden ist, durch einen Irrtum, bloß weil vier unvernünftige Atome auf der nassen Autobahn ineinandergerast sind. Also muß sie eine kosmische Verschwörung suchen. Gott, die Engel oder die Teufel."

Getreide-/Hunger-Problem herzustellen. TT51 = mildert die Kluft zwischen dem Bild von den Kühen und dem später erwähnten Getreide. Nivellierung.

TT53: Die urspr. JG betont (in 41,13) nachhaltig die Entsprechung von Traumdeutung und Geschehnissen. Darauf kommt es auch an (zur Legitimierung Josefs). Demgegenüber bringt TT53 überflüssigerweise in Erinnerung, was faktisch passierte. Erstens weiß man das längst, zweitens ist dieser Aspekt aktuell unbedeutend.

TT60: drückt pathetisch auf die Tränendrüse verglichen mit dem schlanken Erzählstil von TT59. Musterbeispiel für falsche Dramatisierung. Zweimalige Erwähnung Gottes hat ebenfalls die Funktion der Einschüchterung.

TT76 will durch den Hinweis auf Unzählbarkeit den Überfluß an Getreide herausstreichen, gerät dabei aber ins Stottern.

TT96: Emphatisch-moralisierende Verbreiterung mit vielen Abstrakta ("Bedrängnis, Seele, Anflehen") und Wertungen ("Schuldige").

TT97: Nochmals nachträgliche Rechthaberei Rubens.

TT109: Abstrakt, nicht in die aktuelle Erzählfiktion eingebunden, wird die negative Situation dramatisierend in Erinnerung gerufen: Pathos.

TT114: Gegenüber TT115 wird das zurückzubringende Silber verdoppelt.

TT123: Hinzufügung eines Gefühls (Furcht).

TT133: 43,32b wird durch rabiate Wertung ("Greuel") verstärkt.

TT136-TT139: An deplazierter Stelle wird die Silber-Episode (vgl. TT102) nochmals durchgespielt. Nun aber sind Planung und Durchführung anschaulich beschrieben. Außerdem werden die Brüder nun auch wegen des Silbers im Verhör in die Zange genommen, während urspr. nur ihre Ungereimtheiten in der Rede Anlaß für Verdacht waren (TT91) und das Silber nur eine heimliche Rückerstattung des Preises war.

TT141: Die Kulisse für Josefs Auftritt wird ausgeweitet.

TT146: Zu "Israel" wird ausweitend die ganze Großfamilie samt Vieh hinzugenommen.

TT148: Aspekt "Eile" wird eingebracht.

TT151: Feierliche Sprechweise (explizit performativ)

TT153: Die Erwähnung der "Verpflegung" (45,21c) genügt anscheinend nicht: Ein Bearbeiter stellt die Öppigkeit dessen heraus, was Josef den Brüdern mitgab.

TT163-TT178: Obwohl in sich nicht einheitlich liegt doch eine Orgie von Zusatzbeschreibungen/Spezifizierungen der Großfamilie Jakob/Israels vor.

TT191: Pathetisch

TT196: Quantitative Ausweitung

TT203: Es wird zum Zusatz eine noch überwältigendere Aussage der positiven Zukunft hinzugesetzt.

TT223: Nichtssagender Hinweis auf einen Brauch, den man sich im Hintergrund zu denken habe. Damit soll wohl dem Eindruck gewehrt werden, die Einbalsamierung Israels sei lediglich eine Privatinitiative Josefs.

TT224: Zeitliche und bez. der Personengruppe zahlenmäßig dramatische Ausweitung der Trauer um Israel. Sehr unglaubwürdig,

da laut Grundtext nicht einsichtig ist, wieso Israel eine so zentrale Bedeutung für Ägypten gewonnen hätte.

TT228-231A: Verschiedene Strategien, um den Aufwand der Teilnehmer, die Dramatik der Trauer zu steigern.

TT237: Unpassend von der K.-Situation her (vgl. TT236) wird nochmals das Thema "Schuld" gebracht, GOTT neu erwähnt, eine heftige Emotion Josefs berichtet.

8. Nicht-Aushalten poetischer Bildhaftigkeit und Provokation.

TT51: Das Bild von den Kühen (=TT50) wird ergänzt durch das von den Ähren, wohl deshalb, weil (a) auf Bildebene die "Ähren" vertrauter sind (=palästinisches Kolorit?) als die Nil-Kühe und - (b) - weil sich damit eine Brücke von der Bild- zur Sach-Ebene schlagen läßt: Später wird der Hunger, also das Getreide-/Ährenproblem, dominierendes Thema sein. Ein Ergänzter konnte also die große Spanne zwischen Bild- und Sachebene nicht aushalten, glaubte auch, die Textrezipienten seien bei der Lektüre überfordert, fühlte sich deshalb zu nachhelfend-belehrendem Eingriff veranlaßt.

TT55: ... im Verhältnis zu TT54 analog wie TT51/TT50.

zu Ziff. 3.53:

Textstrukturen // Textverständlichkeit

Typisierung nach GROEBEN/CHRISTMANN (1989), SCHWEIZER (1988,ThQ).

zu:

I.

Kognitive klare Gliederung -thematisch kontinuierl. -Gebrauch kurzer, geläufiger, konkreter und anschaulicher Wörter -Bild-Text-Kombinationen -Advance Organizer -sequentielles Arrang. -semantische Redundanz -Zusammenfassungen -Hervorhebung -Überschrift/Glossier. -Fragen/Lernzielangaben	Geringe motivierende und stimulierende Kraft. Hohe Verständlichkeit
--	--

II.

Höhere Komplexität des Textmaterials -them. diskontinuierl. -nicht geläufige Wörter und Wortkombinationen -Satzschachtelungen -Nominalisierungen -große Satzlänge -hoher Info-Wert -konfliktgenerierende Fragen -Bezug auf überraschend neue Information -gleichwahrscheinliche Problemalternativen -Überraschungsmomente	Reduktion von Verständlichkeit und Behalten. Optimale Motivierung.
--	---

Ein Text nach TYP II kann mit Motivation und Techniken nach TYP I bearbeitet worden sein. Resultat:

III.

Unklare Gliederung, widersprüchliche Informationen, stilistische Irritationen	Abbruch der Motivation. Bewußter oder unbewußter Ausstieg aus der Textrezeption
--	---

zu Ziff. 3.54:

Redaktion

zu:

aus: WONNEBERGER, R, Redaktion. Studien zur Textfortschreibung im Alten Testament, entwickelt am Beispiel der Samuel-Überlieferung. FRLANT 156. Göttingen 1992. S.32f

These 1: *Redaktion kann nur erforscht und beschrieben werden, wenn dahinter regelgeleitetes Handeln der Redaktoren vorauszusetzen ist.* Obwohl schon lange Redaktionsgeschichte betrieben wird, ist die Frage nach Regelmäßigkeit bisher so gut wie gar nicht gestellt worden.

These 2: *Eine Redaktionstheorie kann nicht rein induktiv aus der Exegese entwickelt werden.* Es bedarf vielmehr eines Wechselspiels zwischen zwei Polen: Der induktiven Textexegese auf der einen Seite und einem Modell von Redaktion auf der anderen. Der größere Nachholbedarf liegt derzeit auf Seiten des Modells.

These 3: *Als Ausgangspunkt für ein Modell bietet sich jener Ansatz der sprachanalytischen Philosophie an, der als Pragmatik bezeichnet wird und den Handlungscharakter sprachlicher Äußerungen untersucht.* Redaktionelles Handeln ist uns ausschließlich in Gestalt von Texten überkommen und muß daher in seinem Handlungscharakter erst rekonstruiert werden. Es ist zugleich sprachliches Handeln, insofern Redaktoren selbst formulieren, und nicht-sprachliches Handeln, insofern Redaktoren Texte einfügen, tilgen oder umstellen. Durch den Ansatz bei der Pragmatik sollen vor allem die redaktionellen Handlungen in ihrer Regelmäßigkeit und nach ihren Intentionen erfassbar gemacht werden.

These 4: *Beim Aufbau eines Modells ist die Frage nach der Einbettung in den Lebenskontext einzubeziehen ("Sitz im Leben" der Redaktion).* Obwohl der Lebenskontext meist nur indirekt erschließbar ist, liegen in ihm wesentliche Voraussetzungen für Redaktion, also zum einen Texte von besonderer Geltung, die nicht einfach zu vergessen oder durch andere zu ersetzen sind, und zum anderen Anlässe, also Motive, diese Texte zu verändern und damit zu erreichen, daß sie bestimmten Anforderungen der Gemeinschaft besser gerecht werden.

These 5: *Ein solches Modell wird mindestens drei Stufen haben: zuoberst die Stufe der globalen Intentionen; sodann ihre Umsetzung in Stoffe und Argumentationsketten; schließlich die konkrete Veränderung der bestehenden Texte.* Wir möchten natürlich vor allem die Intentionen auf der höchsten Ebene kennen... Wir können sie aber nur induktiv von unten ermitteln und sind damit allen exegetischen Problemen und Unsicherheiten bis hin zur Textkritik ausgeliefert.

These 6: *Ein literarkritischer Ansatz reicht für ein solches Modell nicht aus. ...*

These 7: *Es muß versucht werden, Kriterien für die Rekonstruierbarkeit von Redaktion und den Grad ihrer Sicherheit zu erarbeiten.* Es ist gut denkbar, daß bestimmte redaktionelle Änderungen keine oder nicht genügend sicher beobachtbare Spuren hinterlassen haben. Rekonstruierbar ist Redaktion nur dort, wo wir entweder die Quellen haben. (Chr; Evangelienharmonien) oder wo sie Spuren in Form von Kohärenz (Dtr) oder Textstörungen (1.S) hinterlassen hat.

These 8: *Hier kommt dem Erkennen redaktioneller Muster eine ähnliche Bedeutung zu wie dem Erkennen von Gattungen in der Formgeschichte.* Wenn in der Formgeschichte eine Gattung aufgrund eindeutiger Fälle erst einmal etabliert ist, dann können ihre Spuren auch dort entdeckt werden, wo dies ohne solches Mustererkennen nicht möglich wäre. Es besteht Hoffnung, daß sich bestimmte redaktionelle Techniken an klaren Beispielen erforschen lassen und dadurch auch an weniger klaren Beispielen erkennbar werden.

These 9: *Durch redaktionelle Eingriffe werden in vielen Fällen hybride Gattungen im Sinne der Formgeschichte entstehen.* In der Formgeschichte wird zwischen Rahmen- und Gliedgattung unterschieden. In beiden Fällen handelt es sich jedoch um integere Gattungen. Hybride Gattungen entstehen durch redaktionelle Einfügungen, die weder in das Textmuster der Vorlage passen noch auch selbst eine Gattung bilden.

These 10: *Es gibt sehr unterschiedliche Typen von Redaktion.* (z.B. Kompilationsredaktion [Pentateuch], Adaptionsredaktion [Dtr], vgl. S.93f).

3.6 Präsentation des Ergebnisses für die folgenden Interpretationsschritte

B. INTERPRETATION

Literatur

- RICHTER, W (ed./eds.): Biblia Hebraica transcripta Band 1ff. St. Ottilien 1991ff: EOS // I.2.1
- SCHWEIZER, H: Die vorhergesehene Katastrophe. Der Sturz des Weltenbaumes (Ez 31) in: SCHWEIZER, H (ed./eds.): "... Bäume braucht man doch!" Das Symbol des Baumes zwischen Hoffnung und Zerstörung. Sigmaringen 1986 89-108. // I.2.7
- SCHWEIZER, H: JOSEPH. Mit Photocollagen von Jonas Balena. Urfassung der alttestamentlichen Erzählung (Gen 37-50). Tübingen 1993: Klöpfer&Meyer
- SCHWEIZER, H: Revidierte, morph-analoge Arbeitsübersetzung in: SCHWEIZER, H (ed./eds.): Computerunterstützte Textinterpretation. Die Josefsgeschichte beschrieben und interpretiert im Dreischritt: Syntax-Semantik-Pragmatik. THLI 7. Tübingen 1995 1-39 (ii).: Francke // I.2.7

zu Ziff. 3.61:

Segmentierung 1: Wörter

zu:

aus: Studienarbeit Volker Wysk "Portierung der JOSEF-Datenbank (Turbo-PASCAL nach SQL)".

Nr.	Wortform	Kap	Vers	ÄE	Wort	s/o
1	V.Y*.W.B	37	1	1	1	s
2	Y.J.Q.B	37	1	1	2	s
3	B*.%R.3	37	1	1	3	s
4	M.GV.R.Y	37	1	1	4	s
5	%.B.YV	37	1	1	5	s
6	B*.%R.3	37	1	1	6	s
7	K*.N.J.N	37	1	1	7	s
8	%.L*.H	37	2	1	1	s
9	T*.L.DV.T	37	2	1	2	s
10	Y.J.Q.B	37	2	1	3	s
11	YV.C.P	37	2	2	1	o
12	B*.N	37	2	2	2	s
13	W.B.J	37	2	2	3	s
14	J.S.R.H	37	2	2	4	s
15	W.N.H	37	2	2	5	s
16	H.Y.H	37	2	2	6	o
17	R.J.H	37	2	2	7	o
18	%.T	37	2	2	8	o
19	%.X.YV	37	2	2	9	o
20	B*.3*.%N	37	2	2	10	o
21	V.HV.%	37	2	3	1	o
22	N.J.R	37	2	3	2	o
23	%.T	37	2	4	1	s
24	B*.N.Y	37	2	4	2	s
25	B.L.H.H	37	2	4	3	s
26	V.%.T	37	2	4	4	s
27	B*.N.Y	37	2	4	5	s
28	Z.L.P*.H	37	2	4	6	s
29	N.W.Y	37	2	4	7	s
30	%.B.YV	37	2	4	8	s
31	V.Y*.B.%	37	2	5	1	s
32	YV.C.P	37	2	5	2	s
33	%.T	37	2	5	3	s
34	D*.B*.T.M	37	2	5	4	s
35	R.J.H	37	2	5	5	s
36	%.L	37	2	5	6	s
37	%.B.YH.M	37	2	5	7	s
38	V.Y.S.R.%.L	37	3	1	1	o
39	%.H.B	37	3	1	2	o
40	%.T	37	3	1	3	o
41	YV.C.P	37	3	1	4	o
42	M.K*.L	37	3	1	5	o
43	B*.N.YV	37	3	1	6	o

.....

zu Ziff. 3.62:

Segmentierung 2: Äußerungseinheiten

zu:

aus: Studienarbeit Volker Wysk "Portierung der JOSEF-Datenbank (Turbo-PASCAL nach SQL)".

ÄE-Nr.	von Wortnr.	bis Wortnr.
1	1	7
2	8	10
3	11	20
4	21	22
5	23	30
6	31	37
7	38	43
8	44	48
9	49	52
10	53	54
11	55	60
12	61	62
13	63	66
14	67	69
15	70	71
16	72	75
17	76	77
18	78	81
19	82	83
20	84	84
21	85	89
22	90	90
23	91	92
24	93	94
25	95	95
26	96	97
27	98	99
28	100	102
29	103	105
30	106	109
.....		

zu Ziff. 3.63:

Textpräsentation nach diachroner Analyse

zu:

nach: H. Schweizer, Computerunterstützte Textinterpretation. THLI 7. Tübingen 1995.

Äußerungseinheiten	innerhalb ÄEen	Zählung pro ÄE	KEY
Zählung pro ÄE	absol. Zählung	an Endtext orient.	absol.
Endtext, incl.	ursprüngliche	ursprüngliche	Zählg
Überarbeitungen	Textschicht	Textschicht	End-
Transkription A	Transkription B	Arbeitsübers.	text
		morph-analog	

37,1a

WF:01 V.Y*.W.B			01
WF:02 Y.J.Q.B			02
WF:03 B*.*.R.3			03
WF:04 M.GV.R.Y			04
WF:05 *.B.YV			05
WF:06 B*.*.R.3			06
WF:07 K*.N.J.N			07

37,2a

WF:01 *.L*.H			08
WF:02 T*.L.DV.T			09
WF:03 Y.J.Q.B			10

37,2b*

WF:01 YV.C.P	(1)iu+s+p [יְרוּסָפ]	W:1JOSEF	11
WF:02 B*.N			12
WF:03 W.B.J			13
WF:04 J.S.R.H			14
WF:05 W.N.H			15
WF:06 H.Y.H	(2)h+y+h [העה]	W:6gewesen-ist	16
WF:07 R.J.H	(3)r+y+h [רעה]	W:7ein-Hüter	17
WF:08 *.T	(4)a+t [אא]	W:8mit	18
WF:09 *.X.YV	(5)a+x+iu [אחיו]	W:9Brüdern+seinen	19
WF:10 B*.3*.*N	(6)B+C+an [בצאן]	W:10bei+dem-Kleinvieh	20

37,2c

WF:01 V.HV.*	(7)u+hUa [והוא]	W:lund+er ()	21
WF:02 N.J.R	(8)n+y+r [נער]	W:2ein-Knabe	22

37,2d

WF:01 *.T			23
WF:02 B*.N.Y			24
WF:03 B.L.H.H			25
WF:04 V.*.T			26
WF:05 B*.N.Y			27
WF:06 Z.L.P*.H			28
WF:07 N.W.Y			29
WF:08 *.B.YV			30

37,2e

WF:01 V.Y*.B.*			31
WF:02 YV.C.P			32
WF:03 *.T			33
WF:04 D*.B*.T.M			34
WF:05 R.J.H			35
WF:06 *.L			36
WF:07 *.B.YH.M			37

37,3a

WF:01 V.Y.S.R.*.L	(9)u+i+w+r+a+l [וישראל]	W:lund+ISRAEL	38
WF:02 *.H.B	(10)a+h+b [אהב]	W:2liebte	39

WF:03 *.T	(11)a+t [אא]	W:3→	40
WF:04 YV.C.P	(12)iu+s+p [יְרוּסָפ]	W:4JOSEF	41
WF:05 M.K*.L	(13)m+K+l [מכל]	W:5aus+der-Gesamtheit	42
WF:06 B*.N.YV	(14)B+n+iu [בנין]	W:6Söhne+seiner	43
37,3b			
WF:01 K*.Y	(15)K+i [כי]	W:1denn	44
WF:02 B.N	(16)b+n [בן]	W:2ein-Sohn	45
WF:03 Z.Q.N.YM	(17)z+q+n+i^m [זקנים]	W:3des-Alters ()	46
WF:04 HV.*	(18)hUa [הוא]	W:4er	47
WF:05 LV.	(19)lu+ [לו]	W:5↑ihm	48
37,3c			
WF:01 V.J.S.H	(20)u+y+w+h [ועשה]	W:lund+er-hatte\gemacht	49
WF:02 LV.	(21)lu+ [לו]	W:2↑ihm	50
WF:03 K*.T.N.T	(22)K+t+n+t [כחנת]	W:3einen-Leib-	51
WF:04 P*.C*.YM	(23)P+S+im [פטים]	W:4-rock	52

37,4a

WF:01 V.Y*.R.*V*			53
WF:02 *.X.YV			54

37,4b

WF:01 K*.Y			55
WF:02 *.TV.			56
WF:03 *.H.B			57
WF:04 *.B.YH.M			58
WF:05 M.K*.L			59
WF:06 *.X.YV			60

37,4c

WF:01 V.Y*.S.N.*V.			61
WF:02 *.TV.			62

37,4d

WF:01 V.L.*			63
WF:02 Y.K.LV.			64
WF:03 D*.B*.RV.			65
WF:04 L.W.L.M			66

37,5a

WF:01 V.Y*.X.L.M	(24)u+I+x+l+m [ויחלם]	W:lund+träumte	67
WF:02 YV.C.P	(25)iu+s+p [יְרוּסָפ]	W:2JOSEF	68
WF:03 X.LV.M	(26)x+lu+m [חלום]	W:3einen-Traum	69

37,5b

WF:01 V.Y*.G*.D	(27)u+I+G+d [ויגיד]	W:lund+er\erzählte	70
WF:02 L.*.X.YV	(28)l+a+x+iu [לאחיו]	W:2↑+Brüdern+seinen	71

37,5c

Äußerungseinheiten	innerhalb ÄEen	Zählung pro ÄE	KEY
Zählung pro ÄE	absol. Zählung	an Endtext orient.	absol.
Endtext, incl.	ursprüngliche	ursprüngliche	Zählg
Überarbeitungen	Textschicht	Textschicht	End-
Transkription A	Transkription B	Arbeitsübers.	text
		morph-analog	

zu Ziff. 3.64:

Textpräsentation im HTML-Format

zu:

aus: Computerunterstützte Analyse der Josefsgeschichte. Hier: Portierung der Datenbank (Turbo-Pascal → SQL). Studienarbeit Wysk.

```
<table border=2>
```

```
<tr><td>Jos 37,1/1 W:1 <td>V.Y*.W.B <td>s
<tr><td>Jos 37,1/1 W:2 <td>Y.J.Q.B <td>s
<tr><td>Jos 37,1/1 W:3 <td>B*.%R.3 <td>s
<tr><td>Jos 37,1/1 W:4 <td>M.GV.R.Y <td>s
<tr><td>Jos 37,1/1 W:5 <td>%B.YV <td>s
<tr><td>Jos 37,1/1 W:6 <td>B*.%R.3 <td>s
<tr><td>Jos 37,1/1 W:7 <td>K*.N.J.N <td>s
<tr><td>Jos 37,2/1 W:1 <td>%L*.H <td>s
<tr><td>Jos 37,2/1 W:2 <td>T*.L.DV.T <td>s
<tr><td>Jos 37,2/1 W:3 <td>Y.J.Q.B <td>s
<tr><td>Jos 37,2/2 W:1 <td>YV.C.P <td>o
<tr><td>Jos 37,2/2 W:2 <td>B*.N <td>s
<tr><td>Jos 37,2/2 W:3 <td>W.B.J <td>s
<tr><td>Jos 37,2/2 W:4 <td>J.S.R.H <td>s
<tr><td>Jos 37,2/2 W:5 <td>W.N.H <td>s
<tr><td>Jos 37,2/2 W:6 <td>H.Y.H <td>o
<tr><td>Jos 37,2/2 W:7 <td>R.J.H <td>o
<tr><td>Jos 37,2/2 W:8 <td>%T <td>o
<tr><td>Jos 37,2/2 W:9 <td>%X.YV <td>o
<tr><td>Jos 37,2/2 W:10 <td>B*.3*.%N <td>o
<tr><td>Jos 37,2/3 W:1 <td>V.HV.% <td>o
<tr><td>Jos 37,2/3 W:2 <td>N.J.R <td>o
<tr><td>Jos 37,2/4 W:1 <td>%T <td>s
<tr><td>Jos 37,2/4 W:2 <td>B*.N.Y <td>s
<tr><td>Jos 37,2/4 W:3 <td>B.L.H.H <td>s
<tr><td>Jos 37,2/4 W:4 <td>V.%T <td>s
<tr><td>Jos 37,2/4 W:5 <td>B*.N.Y <td>s
<tr><td>Jos 37,2/4 W:6 <td>Z.L.P*.H <td>s
<tr><td>Jos 37,2/4 W:7 <td>N.W.Y <td>s
<tr><td>Jos 37,2/4 W:8 <td>%B.YV <td>s
<tr><td>Jos 37,2/5 W:1 <td>V.Y*.B.% <td>s
<tr><td>Jos 37,2/5 W:2 <td>YV.C.P <td>s
<tr><td>Jos 37,2/5 W:3 <td>%T <td>s
<tr><td>Jos 37,2/5 W:4 <td>D*.B*.T.M <td>s
<tr><td>Jos 37,2/5 W:5 <td>R.J.H <td>s
<tr><td>Jos 37,2/5 W:6 <td>%L <td>s
<tr><td>Jos 37,2/5 W:7 <td>%B.YH.M <td>s
<tr><td>Jos 37,3/1 W:1 <td>V.Y.S.R.%L <td>o
<tr><td>Jos 37,3/1 W:2 <td>%H.B <td>o
<tr><td>Jos 37,3/1 W:3 <td>%T <td>o
<tr><td>Jos 37,3/1 W:4 <td>YV.C.P <td>o
<tr><td>Jos 37,3/1 W:5 <td>M.K*.L <td>o
<tr><td>Jos 37,3/1 W:6 <td>B*.N.YV <td>o
</table>
```

zu Ziff. 3.65:

Gedicht / Segmentierung

zu: vprag5.24

aus: MAZZA, C, Unités discursives dans le poème "Aux arbres" d'Yves Bonnefoy: QUIROZ, G, u. a. (eds.), Les unités discursives dans l'analyse sémiotique. TAUSCH 12. Frankfurt 1998. S.143-154.

(144) C'est à partir de ces considérations sur les unités discursives que nous allons à présent analyser le poème "Aux Arbres" d'Yves Bonnefoy.

Le texte de référence

Aux Arbres

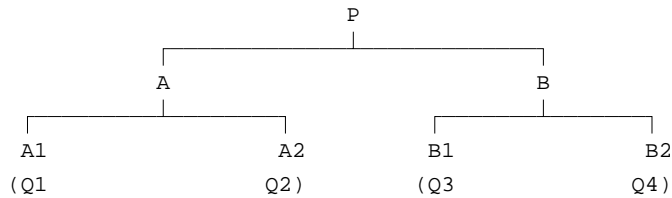
		syllabes
1	Vous qui vous êtes effacés sur son passage,	12
2	Qui avez refermé sur elle vos chemins,	12
3	Impassibles garants que Douve même morte	12
4	Sera lumière encore n'étant rien.	10
5	Vous fibreuse matière et densité,	10
6	Arbres, proches de moi quand elle s'est jetée	12
7	Dans la barque des morts et la bouche serrée	12
8	Sur l'obole de faim, de froid et de silence.	12
9	J'entends à travers vous quel dialogue elle tente	12
10	Avec les chiens, avec l'informe nautonier,	12
11	Et je vous appartiens par son cheminement	12
12	A travers tant de nuit et malgré tout ce fleuve.	12
13	Le tonnerre profond qui roule sur vos branches,	12
14	Les fêtes qu'il enflamme au sommet de l'été	12
15	Signifient qu'elle lie sa fortune à la mienne	12
16	Dans la médiation de votre austérité.	12

Segmentation et unités discursives

"Aux Arbres" se présente typographiquement comme une succession de quatre quatrains monographiques. Le poème ne comporte que des vers au nombre de syllabes paires. La distribution syllabique en miroir d'alexandrins et de décasyllabes sur les deux premiers quatrains (12 12 12 10/10 12 12 12) nous fournit un premier signal démarcateur, au niveau métrique, qui nous permettra d'envisager un découpage du poème en une partie A, constituée par les deux premiers quatrains, et une partie B embrassant les quatrains finaux. Au niveau lexical nous pouvons unir les deux premiers quatrains par le couplage du pronom personnel à la deuxième personne du pluriel "vous" en position initiale de la première et deuxième strophe (v.1/5). L'itération du pronom personnel "vous" n'a pas lieu dans le troisième quatrain, l'attente d'une troisième reprise est donc déçue. L'apostrophe pronominale "vous" ne sera pas suivie de (145) verbes principaux dans les deux premières strophes, contrairement aux quatrains 3 ou 4 où les phrases sont complètes. La série ternaire du vers 8 "Sur l'obole de faim, de froid, et de silence" renforce encore le caractère d'autonomie relative de la partie A par son effet de clôture.

Pour distinguer les deux parties A et B, faisons appel à une opposition lexicale et taxique entre le lexème "silence" en position finale du vers 8 et "j'entends" en position initiale du vers 9. Signalons enfin la rupture temporelle passé/présent qui sépare les deux premiers quatrains des deux derniers.

La division strophique est à elle seule un critère de découpage assez fort qui nous permet de subdiviser à leur tour les unités principales A et B en sous-unités hiérarchiquement inférieures (A1, A2 et B1, B2). En désignant par P le poème considéré comme un système clos de relations, nous proposons le modèle d'organisation suivant:



Nous aurons à construire entre les différentes unités discursives (A (A1; A2)-B (B1; B2) des relations d'équivalence sémantique et de transformation en vertu desquelles tout contenu de la première unité (A) se retrouve dans la seconde (B), de manière identique, transformée ou déniée.

zu Ziff. 3.66:

Textsegmentierung - komplexes Thema

zu: vprag4.12

nach: H. SCHWEIZER, Die Josefsgeschichte... THLI 4, 1991. bzw.

H. SCHWEIZER (Hg.), Computerunterstützte Textinterpretation... THLI 7, 1995

=====Zunächst: Konstituierung des Textes=====

0.1 UNVOKALISIERTER HEBRÄISCHER TEXT

ישבו לאכל לחם וישאו עיניהם ויראו והנה ארחח ישמעאלים באה מגלעד וגמליהם נשאים נכאת וצרי ולט הולכים להוריד מצרימה ויאמר יהודה אל אחיו מה בצע כי נהרג את אחינו וכסינו את דמו לכו ונמכרו לישמעאלים וידנו אל תהי בו כי אחינו בשרנו הוא וישמעו אחיו ויעברו אנשים מדינים סחרים וימשכו ויעלו את יוסף מן הבור וימכרו את יוסף לישמעאלים בעשרים כסף ויביאו את יוסף מצרימה ויקנהו איש מצרי מיד הישמעאלים ויהי בבית אדניו המצרי וימצא יוסף חן בעיניו וישרת אתו ויפקדוהו על ביתו וכל יש לו נתן בידו ולא ידע אתו מאומה כי אם הלחם אשר הוא אוכל ויהי יוסף יפה תאר ויפה מראה ומשא אשת אדניו את עיניה אל יוסף ותאמר שכבה עמי ויאמר ויאמר אל אשת אדניו הו אדני לא ידע אתי מה בבית וכל אשר יש לו חשך ממני מאומה כי אם ארוח נתן בידי איננו גדול בבית הזה ממני ולא באשר את אשתו

0.2 VOKALISIERTER HEBRÄISCHER TEXT

ohne mas. Akzente

נִשְׁבּוּ לֶאֱכֹל-לֶחֶם וַיִּשְׂאוּ עֵינֵיהֶם וַיִּרְאוּ וְהִנֵּה אֶרְחַח יִשְׁמַעְאֵלִים בָּאָה מִגְלֵעַד וַיִּגְמְלֵיהֶם נִשְׂאִים נִכְאֹת וְצָרִי וְלוֹט הוֹלְכִים לְהוֹרִיד מִצְרֵימָה וַיֹּאמֶר יְהוּדָה אֶל-אָחִיו מַה-בְּצַע כִּי נִהְרַג אֶת-אֲחִינוּ וְכָסִינוּ אֶת-דַּמּוֹ לָכוּ וְנִמְכְּרוּ לִישְׁמַעְאֵלִים וַיִּדְנוּ אֶל-תְּהֵי-בּוֹ כִּי-אֲחִינוּ בְּשָׂרֵנוּ הוּא וַיִּשְׁמְעוּ אָחִיו וַיַּעֲבְרוּ אַנְשֵׁים מִדִּינִים סַחְרִים וַיִּמְשְׁכוּ וַיַּעֲלוּ אֶת-יוֹסֵף מִן-הַבּוֹר וַיִּמְכְּרוּ אֶת-יוֹסֵף לִישְׁמַעְאֵלִים בְּעֶשְׂרִים כֶּסֶף וַיְבִיאוּ אֶת-יוֹסֵף מִצְרַיִם וַיִּקְנֶהוּ אִישׁ מִצְרִי מִיַּד הַיִּשְׁמַעְאֵלִים וַיְהִי בְּבֵית אֲדֹנָיו הַמִּצְרִי וַיִּמְצָא יוֹסֵף חֵן בְּעֵינָיו וַיִּשְׂרַח אֹתוֹ וַיִּפְקְדֵהוּ עַל-בֵּיתוֹ וְכָל-יֶשֶׁר לּוֹ נָתַן בְּיָדוֹ וְלֹא יָדָע אֹתוֹ מֵאוֹמֶה כִּי אִם הַלֶּחֶם אֲשֶׁר הוּא אוֹכֵל וַיְהִי יוֹסֵף יָפֵה תֹאֵר וְיָפֵה מְרֹאֵה וּמִשָּׂא אִשְׁתֹּ אֲדֹנָיו אֶת עֵינֶיהָ אֶל יוֹסֵף וַתֹּאמֶר שְׁכַבָה עִמִּי וַיֹּאמֶר אֵל אִשְׁתֹּ אֲדֹנָיו הֲוֹ אֲדֹנִי לֹא-יָדָע אֵמֶי חֹשֶׁךְ מִמֶּנִּי מֵאוֹמֶה כִּי אִם אֲרוּחַ נָתַן בְּיָדִי אֵינְנוּ גָדוֹל בְּבֵית הַזֶּה מִמֶּנִּי וְלֹא-בְּאִשְׁרֵי אֶתְּ-אִשְׁתִּי

1. EINS-ZU-EINS-ÜBERSETZUNG

ENDFASSUNG/sequentuell

und+sie\setzten-sich (um-)zu+essen Brot und+sie\hoben Augen+ihre
 und+sie\sahen und+da eine-Karawane(-von) ISMAELITERN eine-Kommende
 aus+GILEAD und+Kamele+ihre Tragende Tragakant und+Mastix
 und+Ladanum Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts
 und+sprach JUDA zu Brüdern+seinen was ein-Gewinn falls
 wir-werden/mod.\umbringen → Bruder+unseren
 und+verdecken-werden/mod.\wir → Blut+sein geht
 und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN und+Hand+unsere
 nicht sei gegen+ihn denn Bruder+unser Fleisch+unser er und+hörten

Brüder+seine und+vorbeikamen Männer MIDIANITISCHE Kaufleute und+sie\zogen und+sie\herausholten → den-JOSEF aus dem+Brunnen und+sie\verkauften → JOSEF an+die-ISMAELITER für+zwanzig Silber(stücke) und+sie\brachten → JOSEF ÄGYPTEN+wärts und+zurückkehrte RUBEN zu dem+Brunnen und+da die-Nichtexistenz des-Josef in+dem-Brunnen und+er\zerriß → Kleider+seine und+er\zurückkehrte zu Brüdern+seinen und+er\sprach das+Kind die-Nicht-Existenz-von+ihm und+ich wohin ich ein-Gehender und+sie\nahmen → den-Rock des-JOSEF und+sie\schlachteten einen-Bock (von-)Ziegen und+sie\eintauchten den-Rock in+das-Blut und+sie\zuspielten den+Leib- rock und+sie\brachten zu Vater+ihrem und+sie\sprachen dieses gefunden-haben>wir schau-genau-hin doch ob+der-Rock Sohnes+deines dieser oder nicht und+er\genau-anschaute+ihn und+er\sprach der-Rock Sohnes+meines ein-Tier ein-böses hat-gefressen+ihn zerfleischt zerfleischt-ist JOSEF und+JAKOB zerriß Gewänder+seine und+er\legte einen+Sack auf+Hüften+seine und+er\trauerte über Sohn+seinen Tage viele und+aufstanden die-Gesamtheit Söhne+seiner und+die-Gesamtheit Töchter+seiner um+zu+trösten+ihn und+er\weigerte-sich zu+getröstet-werden und+er\sprach fürwahr ich-will\hinuntersteigen zu Sohn+meinem (als-)Trauernder die-Unterwelt+zu und+beweinte ihn Vater+sein und+die-MIDIANITER verkauften ihn nach ÄGYPTEN an+POTIFAR Höfling des-PHARAO Oberster der+Schutzwache [ganz Gen 38] und+JOSEF wurde ÄGYPTEN+zu hinabgebracht und+kaufte+ihn POTIFAR Höfling des-PHARAO Oberster der+Schutzwache ein-Mann ein-ÄGYPTISCHER aus+der-Hand der+ISMAELITER die hinabgebracht-hatten+ihn dort+hin und+war JAHWE mit JOSEF und er\war ein-Mann des-Gelingens und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN und+sah Herr+seiner daß JAHWE mit+ihm und+die-Gesamtheit was er ein-Tuender JAHWE ein-Vollendender durch+Hand+seine und+fand JOSEF Gefallen in+Augen+seinen und+er\bediente →+ihn und+er\Aufsicht-führen-ließ+ihn über Haus+sein und+die-Gesamtheit Existenz ↑+ihm und+es\war von+da Aufsicht-führen-ließ\er ihn in+Haus+seinem und+über die-Gesamtheit dessen Existenz ↑+ihm und+JAHWE segnete → das-Haus des+ÄGYPTERS wegen JOSEF und+war der-Segen JAHWES in+der-Gesamtheit dessen Existenz ↑+ihm im+Haus und+auf+dem-Feld und+er\überließ die-Gesamtheit was ↑+ihm in+der-Hand des-JOSEF und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart) Etwas nämlich außer dem+Brot welches jener ein-Essender und+war JOSEF ein-Schöner(-an) Gestalt und+ein-Schöner(-an) Aussehen und+es\war nach den+Sachverhalten diesen und+erhob die-Frau Herrn+seines → Augen+ihre (hin-)zu JOSEF und+sie\sprach leg-dich+doch mit+mir und+er\weigerte-sich und+er\sprach zu der-Frau Herrn+seines da Herr+mein nicht wahrgenommen-hat in+meiner(-Gegenwart) was in+dem-Haus und+die-Gesamtheit dessen Existenz ↑+ihm gegeben-hat&er in+Hand+meine Nicht-Existenz(-von)+ihm ein-Großer in+dem-Haus dem+selben von-her+mir und+nicht zurückgehalten-hat&er vor+mir Etwas nämlich außer →+dir denn du Frau+seine

2. MIT LITERARKRITIK

sekundäre Teile markiert

und+sie\setzten-sich (um-)zu+essen Brot und+sie\hoben Augen+ihre und+sie\sahen und+da eine-Karawane(-von) ISMAELITERN eine-Kommende aus+GILEAD und+Kamele+ihre Tragende Tragakant und+Mastix und+Ladanum Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts und+sprach JUDA zu Brüdern+seinen was ein-Gewinn falls wir-werden/mod.\umbringen → Bruder+unseren und+verdecken-werden/mod.\wir → Blut+sein geht

und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN und+Hand+unsere nicht sei gegen+ihn denn Bruder+unser Fleisch+unser er und+hörten Brüder+seine und+vorbeikamen Männer MIDIANITISCHE Kaufleute und+sie\zogen und+sie\herausholten → den-JOSEF aus dem+Brunnen und+sie\verkauften → JOSEF an+die-ISMAELITER für+zwanzig Silber(stücke) und+sie\brachten → JOSEF ÄGYPTEN+wärts

und+zurückkehrte RUBEN zu dem+Brunnen und+da die-Nichtexistenz des-Josef in+dem-Brunnen und+er\zerriß → Kleider+seine und+er\zurückkehrte zu Brüdern+seinen und+er\sprach das+Kind die-Nicht-Existenz-von+ihm und+ich wohin ich ein-Gehender und+sie\nahmen → den-Rock des-JOSEF und+sie\schlachteten einen-Bock (von-)Ziegen und+sie\eintauchten den-Rock in+das-Blut und+sie\zuspielten den+Leib- rock und+sie\brachten zu Vater+ihrem und+sie\sprachen dieses gefunden-haben>wir schau-genau-hin doch ob+der-Rock Sohnes+deines dieser oder nicht und+er\genau-anschaute+ihn und+er\sprach der-Rock Sohnes+meines ein-Tier ein-böses hat-gefressen+ihn zerfleischt zerfleischt-ist JOSEF und+JAKOB zerriß Gewänder+seine und+er\legte einen+Sack auf+Hüften+seine und+er\trauerte über Sohn+seinen Tage viele und+aufstanden die-Gesamtheit Söhne+seiner und+die-Gesamtheit Töchter+seiner um+zu+trösten+ihn und+er\weigerte-sich zu+getröstet-werden und+er\sprach fürwahr ich-will\hinuntersteigen zu Sohn+meinem (als-)Trauernder die-Unterwelt+zu und+beweinte ihn Vater+sein und+die-MIDIANITER verkauften ihn nach ÄGYPTEN an+POTIFAR Höfling des-PHARAO Oberster der+Schutzwache [ganz Gen 38] und+JOSEF wurde ÄGYPTEN+zu hinabgebracht

und+kaufte+ihn

POTIFAR Höfling des-PHARAO Oberster der+Schutzwache

ein-Mann ein-ÄGYPTISCHER aus+der-Hand der+ISMAELITER

die hinabgebracht-hatten+ihn dort+hin und+war JAHWE mit JOSEF und er\war ein-Mann des-Gelingens

und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN

und+sah Herr+seiner daß JAHWE mit+ihm und+die-Gesamtheit was er ein-Tuender JAHWE ein-Vollendender durch+Hand+seine

und+fand JOSEF Gefallen in+Augen+seinen und+er\bediente →+ihn und+er\Aufsicht-führen-ließ+ihn über Haus+sein und+die-Gesamtheit Existenz ↑+ihm

und+es\war von+da Aufsicht-führen-ließ\er ihn in+Haus+seinem und+über die-Gesamtheit dessen Existenz ↑+ihm und+JAHWE segnete → das-Haus des+ÄGYPTERS wegen JOSEF und+war der-Segen JAHWES in+der-Gesamtheit dessen Existenz ↑+ihm im+Haus und+auf+dem-Feld und+er\überließ die-Gesamtheit was ↑+ihm in+der-Hand des-JOSEF

und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart) Etwas nämlich außer dem+Brot welches jener ein-Essender und+war JOSEF ein-Schöner(-an) Gestalt und+ein-Schöner(-an) Aussehen

und+es\war nach den+Sachverhalten diesen

und+erhob die-Frau Herrn+seines → Augen+ihre (hin-)zu JOSEF und+sie\sprach leg-dich+doch mit+mir und+er\weigerte-sich und+er\sprach zu der-Frau Herrn+seines da Herr+mein nicht wahrgenommen-hat in+meiner(-Gegenwart) was in+dem-Haus und+die-Gesamtheit dessen Existenz ↑+ihm gegeben-hat&er in+Hand+meine Nicht-Existenz(-von)+ihm ein-Großer in+dem-Haus dem+selben von-her+mir und+nicht zurückgehalten-hat&er vor+mir Etwas nämlich außer →+dir denn du Frau+seine

3. URSPRÜNGLICHE SCHICHT/sequentiell

und+sie\setzten-sich (um-)zu+essen Brot und+sie\hoben Augen+ihre
 und+sie\sahen und+da eine-Karawane(-von) ISMAELITERN eine-Kommende
 aus+GILEAD und+Kamele+ihre Tragende Tragakant und+Mastix
 und+Ladanum Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts
 und+sprach JUDA zu Brüdern+seinen was ein-Gewinn falls
 wir-werden/mod.\umbringen → Bruder+unseren
 und+verdecken-werden/mod.\wir → Blut+sein geht
 und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN und+Hand+unsere
 nicht sei gegen+ihn denn Bruder+unser Fleisch+unser er und+hörten
 Brüder+seine und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 und+sie\zogen und+sie\herausholten → den-JOSEF aus dem+Brunnen
 und+sie\verkauften → JOSEF an+die-ISMAELITER für+zwanzig
 Silber(stücke) und+sie\brachten → JOSEF ÄGYPTEN+wärts
 und+kaufte+ihn ein-Mann ein-ÄGYPTISCHER aus+der-Hand
 der+ISMAELITER und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 und+fand JOSEF Gefallen in+Augen+seinen und+er\bediente →+ihn
 und+er\Aufsicht-führen-ließ+ihn über Haus+sein und+die-Gesamtheit
 Existenz ↑+ihm gegeben-hat&er in+Hand+seine und+nicht
 wahrgenommen-hat&er in+seiner(-Gegenwart) Etwas nämlich außer
 dem+Brot welches jener ein-Essender und+war JOSEF ein-Schöner(-an)
 Gestalt und+ein-Schöner(-an) Aussehen und+erhob die-Frau
 Herrn+seines → Augen+ihre (hin-)zu JOSEF und+sie\sprach
 leg-dich+doch mit+mir und+er\weigerte-sich und+er\sprach zu
 der-Frau Herrn+seines da Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart) was in+dem-Haus und+die-Gesamtheit dessen
 Existenz ↑+ihm gegeben-hat&er in+Hand+meine
 Nicht-Existenz(-von)+ihm ein-Großer in+dem-Haus dem+selben
 von-her+mir und+nicht zurückgehalten-hat&er vor+mir Etwas nämlich
 außer →+dir denn du Frau+seine

4. PLUS WORTZAHL

(Ausdrucks-)SYNTAX

und+sie\setzten-sich (um-)zu+essen Brot und+sie\hoben Augen+ihre
 und+sie\sahen und+da eine-Karawane(-von) ISMAELITERN eine-Kommende
 aus+GILEAD und+Kamele+ihre Tragende Tragakant und+Mastix
 und+Ladanum Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
 und+sprach JUDA zu Brüdern+seinen was ein-Gewinn falls
 wir-werden/mod.\umbringen → Bruder+unseren
 und+verdecken-werden/mod.\wir → Blut+sein geht
 und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN und+Hand+unsere
 nicht sei gegen+ihn[0240] denn Bruder+unser Fleisch+unser er
 und+hörten Brüder+seine und+vorbeikamen Männer MIDIANITISCHE
 Kaufleute und+sie\zogen und+sie\herausholten → den-JOSEF aus
 dem+Brunnen und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke) und+sie\brachten → JOSEF ÄGYPTEN+wärts
 und+kaufte+ihn ein-Mann ein-ÄGYPTISCHER aus+der-Hand
 der+ISMAELITER und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 und+fand JOSEF Gefallen in+Augen+seinen und+er\bediente[0280]
 →+ihn und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 und+die-Gesamtheit Existenz ↑+ihm gegeben-hat&er in+Hand+seine
 und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart) Etwas nämlich
 außer dem+Brot welches jener ein-Essender und+war[0300] JOSEF
 ein-Schöner(-an) Gestalt und+ein-Schöner(-an) Aussehen und+erhob
 die-Frau Herrn+seines → Augen+ihre (hin-)zu JOSEF und+sie\sprach
 leg-dich+doch mit+mir und+er\weigerte-sich und+er\sprach zu
 der-Frau Herrn+seines[0320] da Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart) was in+dem-Haus und+die-Gesamtheit dessen
 Existenz ↑+ihm gegeben-hat&er in+Hand+meine

Nicht-Existenz(-von)+ihm ein-Großer in+dem-Haus dem+selben
 von-her+mir und+nicht zurückgehalten-hat&er[0340] vor+mir Etwas
 nämlich außer →+dir denn du Frau+seine

5. PLUS KAPITEL/VERS Traditionelles Referenzsystem
- 37,25 und+sie\setzten-sich (um-)zu+essen Brot und+sie\hoben
 Augen+ihre und+sie\sahen und+da eine-Karawane(-von)
 ISMAELITERN eine-Kommende aus+GILEAD und+Kamele+ihre
 Tragende Tragakant und+Mastix und+Ladanum Gehende
 (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
- 37,26 und+sprach JUDA zu Brüdern+seinen was ein-Gewinn falls
 wir-werden/mod.\umbringen → Bruder+unseren
 und+verdecken-werden/mod.\wir → Blut+sein
 geht und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN
 und+Hand+unsere nicht sei gegen+ihn[0240] denn Bruder+unser
 Fleisch+unser er und+hörten Brüder+seine
- 37,27 und+vorbeikamen Männer MIDIANITISCHE Kaufleute und+sie\zogen
 und+sie\herausholten → den-JOSEF aus dem+Brunnen
 und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke) und+sie\brachten → JOSEF
 ÄGYPTEN+wärts
- 39,01 ... und+kaufte+ihn ... ein-Mann ein-ÄGYPTISCHER aus+der-Hand
 der+ISMAELITER ...
- 39,02 ... und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
- 39,04 und+fand JOSEF Gefallen in+Augen+seinen
 und+er\bediente[0280] →+ihn und+er\Aufsicht-führen-ließ+ihn
 über Haus+sein und+die-Gesamtheit Existenz ↑+ihm
 gegeben-hat&er in+Hand+seine
- 39,06 und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart) Etwas
 nämlich außer dem+Brot welches jener ein-Essender
 und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
- 39,07 und+erhob die-Frau Herrn+seines → Augen+ihre (hin-)zu JOSEF
 und+sie\sprach leg-dich+doch mit+mir
- 39,08 und+er\weigerte-sich und+er\sprach zu der-Frau
 Herrn+seines[0320] da Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart) was in+dem-Haus und+die-Gesamtheit
 dessen Existenz ↑+ihm gegeben-hat&er in+Hand+meine
- 39,09 Nicht-Existenz(-von)+ihm ein-Großer in+dem-Haus dem+selben
 von-her+mir und+nicht zurückgehalten-hat&er[0340] vor+mir
 Etwas nämlich außer →+dir denn du Frau+seine

6. PLUS ÄUSSERUNGSEINHEIT (ÄE) Bedeutungswissen
- 37,25a und+sie\setzten-sich (um-)zu+essen Brot
- 37,25b und+sie\hoben Augen+ihre
- 37,25c und+sie\sahen
- 37,25d und+da
- 37,25e eine-Karawane(-von) ISMAELITERN () eine-Kommende
 aus+GILEAD
- 37,25f und+Kamele+ihre () Tragende Tragakant und+Mastix
 und+Ladanum
- 37,25g Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
- 37,26a und+sprach JUDA zu Brüdern+seinen
- 37,26b was () ein-Gewinn

37,26c falls wir-werden/mod.\umbringen → Bruder+unseren
 37,26d und+verdecken-werden/mod.\wir → Blut+sein
 37,27a geht
 37,27b und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN
 37,27c und+Hand+unsere nicht sei gegen+ihn[0240]
 37,27d denn Bruder+unser
 37,27e Fleisch+unser () er
 37,27f und+hörten Brüder+seine
 37,28a und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 37,28b und+sie\zogen
 37,28c und+sie\herausholten → den-JOSEF aus dem+Brunnen
 37,28d und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke)
 37,28e und+sie\brachten → JOSEF ÄGYPTEN+wärts
 39,01b* und+kauften+ihn ein-Mann ein-ÄGYPTISCHER aus+der-Hand
 der+ISMAELITER
 39,02c und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 39,04a und+fand JOSEF Gefallen in+Augen+seinen
 39,04b und+er\bediente[0280] →+ihn
 39,04c und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 39,04d und+die-Gesamtheit
 39,04e Existenz () ↑+ihm
 39,04f gegeben-hat&er in+Hand+seine
 39,06d und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart)
 Etwas nämlich außer dem+Brot
 39,06e welches jener () ein-Essender
 39,06f und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
 39,07b und+erhob die-Frau Herrn+seines → Augen+ihre
 (hin-)zu JOSEF
 39,07c und+sie\sprach
 39,07d leg-dich+doch mit+mir
 39,08a und+er\weigerte-sich
 39,08b und+er\sprach zu der-Frau Herrn+seines[0320]
 39,08c da
 39,08d Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart)
 39,08e was () in+dem-Haus
 39,08f und+die-Gesamtheit
 39,08g dessen Existenz () ↑+ihm
 39,08h gegeben-hat&er in+Hand+meine

086::37,27c und+Hand+unsere nicht sei gegen+ihn[0240]
 087::37,27d denn Bruder+unser
 088::37,27e Fleisch+unser () er
 089::37,27f und+hörten Brüder+seine
 090::37,28a und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 091::37,28b und+sie\zogen
 092::37,28c und+sie\herausholten → den-JOSEF aus dem+Brunnen
 093::37,28d und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke)
 094::37,28e und+sie\brachten → JOSEF ÄGYPTEN+wärts
 095::39,01b* und+kauften+ihn ein-Mann ein-ÄGYPTISCHER aus+der-Hand
 der+ISMAELITER
 096::39,02c und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 097::39,04a und+fand JOSEF Gefallen in+Augen+seinen
 098::39,04b und+er\bediente[0280] →+ihn
 099::39,04c und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 100::39,04d und+die-Gesamtheit
 101::39,04e Existenz () ↑+ihm
 102::39,04f gegeben-hat&er in+Hand+seine
 103::39,06d und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart)
 Etwas nämlich außer dem+Brot
 104::39,06e welches jener () ein-Essender
 105::39,06f und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
 106::39,07b und+erhob die-Frau Herrn+seines → Augen+ihre
 (hin-)zu JOSEF
 107::39,07c und+sie\sprach
 108::39,07d leg-dich+doch mit+mir
 109::39,08a und+er\weigerte-sich
 110::39,08b und+er\sprach zu der-Frau Herrn+seines[0320]
 111::39,08c da
 112::39,08d Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart)
 113::39,08e was () in+dem-Haus
 114::39,08f und+die-Gesamtheit
 115::39,08g dessen Existenz () ↑+ihm
 116::39,08h gegeben-hat&er in+Hand+meine

=====Ab hier: Deskription und Interpretation=====

7. PLUS FORTLAUFENDE NUMMER Berechenbarkeit
 073::37,25a und+sie\setzten-sich (um-)zu+essen Brot
 074::37,25b und+sie\hoben Augen+ihre
 075::37,25c und+sie\sahen
 076::37,25d und+da
 077::37,25e eine-Karawane(-von) ISMAELITERN () eine-Kommende
 aus+GILEAD
 078::37,25f und+Kamele+ihre () Tragende Tragakant und+Mastix
 und+Ladanum
 079::37,25g Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
 080::37,26a und+sprach JUDA zu Brüdern+seinen
 081::37,26b was () ein-Gewinn
 082::37,26c falls wir-werden/mod.\umbringen → Bruder+unseren
 083::37,26d und+verdecken-werden/mod.\wir → Blut+sein
 084::37,27a geht
 085::37,27b und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN

8. PLUS TEXTDEIXIS PRAGMATIK: Textgrammatik
 073::37,25a und+sie\setzten-sich (um-)zu+essen Brot
 074::37,25b und+sie\hoben Augen+ihre
 075::37,25c und+sie\sahen
 076::37,25d und+da
 077::37,25e eine-Karawane(-von) ISMAELITERN () eine-Kommende
 aus+GILEAD
 078::37,25f und+Kamele+ihre () Tragende Tragakant und+Mastix
 und+Ladanum
 079::37,25g Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
 080::37,26a und+sprach JUDA zu Brüdern+seinen
 081::37,26b was () ein-Gewinn
 082::37,26c falls wir-werden/mod.\umbringen → Bruder+unseren
 083::37,26d und+verdecken-werden/mod.\wir → Blut+sein

084::37,27a geht
 085::37,27b und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN
 086::37,27c und+Hand+unsere nicht sei gegen+ihn[0240]
 087::37,27d denn Bruder+unser
 088::37,27e Fleisch+unser () er
 089::37,27f und+hörten Brüder+seine

Block: 7

090::37,28a und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 091::37,28b und+sie\zogen
 092::37,28c und+sie\herausholten → den-JOSEF aus dem+Brunnen
 093::37,28d und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke)
 094::37,28e und+sie\brachten → JOSEF ÄGYPTEN+wärts
 095::39,01b* und+kaufte+ihn ein-Mann ein-ÄGYPTISCHER
 aus+der-Hand der+ISMAELITER
 096::39,02c und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 097::39,04a und+fand JOSEF Gefallen in+Augen+seinen
 098::39,04b und+er\bediente[0280] →+ihn
 099::39,04c und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 100::39,04d und+die-Gesamtheit
 101::39,04e Existenz () ↑+ihm
 102::39,04f gegeben-hat&er in+Hand+seine
 103::39,06d und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart)
 Etwas nämlich außer dem+Brot
 104::39,06e welches jener () ein-Essender

Block: 8

105::39,06f und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
 106::39,07b und+erhob die-Frau Herrn+seines → Augen+ihre
 (hin-)zu JOSEF
 107::39,07c und+sie\sprach
 108::39,07d leg-dich+doch mit+mir
 109::39,08a und+er\weigerte-sich

9. PLUS TG EINHEITEN

PRAGMATIK: Textgrammatik

TGE4*

073::37,25a und+sie\setzten-sich (um-)zu+essen Brot
 074::37,25b und+sie\hoben Augen+ihre
 075::37,25c und+sie\sahen
 076::37,25d und+da
 077::37,25e eine-Karawane(-von) ISMAELITERN () eine-Kommende
 aus+GILEAD
 078::37,25f und+Kamele+ihre () Tragende Tragakant und+Mastix
 und+Ladanum
 079::37,25g Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
 080::37,26a und+sprach JUDA zu Brüdern+seinen
 081::37,26b was () ein-Gewinn
 082::37,26c falls wir-werden/mod.\umbringen → Bruder+unseren
 083::37,26d und+verdecken-werden/mod.\wir → Blut+sein
 084::37,27a geht
 085::37,27b und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN
 086::37,27c und+Hand+unsere nicht sei gegen+ihn[0240]

087::37,27d denn Bruder+unser
 088::37,27e Fleisch+unser () er
 089::37,27f und+hörten Brüder+seine

TGE5*

Block: 7

090::37,28a und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 091::37,28b und+sie\zogen
 092::37,28c und+sie\herausholten → den-JOSEF aus dem+Brunnen
 093::37,28d und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke)
 094::37,28e und+sie\brachten → JOSEF ÄGYPTEN+wärts
 TGE6*
 095::39,01b* und+kaufte+ihn ein-Mann ein-ÄGYPTISCHER
 aus+der-Hand der+ISMAELITER
 096::39,02c und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 097::39,04a und+fand JOSEF Gefallen in+Augen+seinen
 098::39,04b und+er\bediente[0280] →+ihn
 099::39,04c und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 100::39,04d und+die-Gesamtheit
 101::39,04e Existenz () ↑+ihm
 102::39,04f gegeben-hat&er in+Hand+seine
 103::39,06d und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart)
 Etwas nämlich außer dem+Brot
 104::39,06e welches jener () ein-Essender

TGE7*

Block: 8

105::39,06f und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
 106::39,07b und+erhob die-Frau Herrn+seines → Augen+ihre
 (hin-)zu JOSEF
 107::39,07c und+sie\sprach
 108::39,07d leg-dich+doch mit+mir
 109::39,08a und+er\weigerte-sich
 110::39,08b und+er\sprach zu der-Frau Herrn+seines[0320]
 111::39,08c da
 112::39,08d Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart)
 113::39,08e was () in+dem-Haus
 114::39,08f und+die-Gesamtheit
 115::39,08g dessen Existenz () ↑+ihm
 116::39,08h gegeben-hat&er in+Hand+meine
 117::39,09a Nicht-Existenz(-von)+ihm () ein-Großer in+dem-Haus
 dem+selben von-her+mir
 118::39,09b und+nicht zurückgehalten-hat&er[0340] vor+mir Etwas
 nämlich außer →+dir
 119::39,09c denn du () Frau+seine

10. PLUS TL EINHEITEN

PRAGMATIK: Textlinguistik

TLE2* (TGE4*)

073::37,25a und+sie\setzten-sich (um-)zu+essen Brot
 074::37,25b und+sie\hoben Augen+ihre

075::37,25c und+sie\sahen
 076::37,25d und+da
 077::37,25e eine-Karawane(-von) ISMAELITERN () eine-Kommende
 aus+GILEAD
 078::37,25f und+Kamele+ihre () Tragende Tragakant und+Mastix
 und+Ladanum
 079::37,25g Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
 080::37,26a und+sprach JUDA zu Brüdern+seinen
 081::37,26b was () ein-Gewinn
 082::37,26c falls wir-werden/mod.\umbringen → Bruder+unseren
 083::37,26d und+verdecken-werden/mod.\wir → Blut+sein
 084::37,27a geht
 085::37,27b und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN
 086::37,27c und+Hand+unsere nicht sei gegen+ihn[0240]
 087::37,27d denn Bruder+unser
 088::37,27e Fleisch+unser () er
 089::37,27f und+hörten Brüder+seine

TLE3* (TGE5*)**Block: 7**

090::37,28a und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 091::37,28b und+sie\zogen
 092::37,28c und+sie\herausholten → den-JOSEF aus dem+Brunnen
 093::37,28d und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke)
 094::37,28e und+sie\brachten → JOSEF ÄGYPTEN+wärts

TLE3* (TGE6*)

095::39,01b* und+kaufte+ihn ein-Mann ein-ÄGYPTISCHER
 aus+der-Hand der+ISMAELITER
 096::39,02c und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 097::39,04a und+fand JOSEF Gefallen in+Augen+seinen
 098::39,04b und+er\bediente[0280] →+ihn
 099::39,04c und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 100::39,04d und+die-Gesamtheit
 101::39,04e Existenz () ↑+ihm
 102::39,04f gegeben-hat&er in+Hand+seine
 103::39,06d und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart)
 Etwas nämlich außer dem+Brot
 104::39,06e welches jener () ein-Essender

TLE3* (TGE7*)**Block: 8**

105::39,06f und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
 106::39,07b und+erhob die-Frau Herrn+seines → Augen+ihre
 (hin-)zu JOSEF
 107::39,07c und+sie\sprach
 108::39,07d leg-dich+doch mit+mir
 109::39,08a und+er\weigerte-sich
 110::39,08b und+er\sprach zu der-Frau Herrn+seines[0320]
 111::39,08c da
 112::39,08d Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart)
 113::39,08e was () in+dem-Haus
 114::39,08f und+die-Gesamtheit
 115::39,08g dessen Existenz () ↑+ihm
 116::39,08h gegeben-hat&er in+Hand+meine

117::39,09a Nicht-Existenz(-von)+ihm () ein-Großer in+dem-Haus
 dem+selben von-her+mir
 118::39,09b und+nicht zurückgehalten-hat&er[0340] vor+mir Etwas
 nämlich außer →+dir
 119::39,09c denn du () Frau+seine

11. PLUS KHS

PRAGMATIK: Textlinguistik

TLE2* (TGE4*)

073::37,25a und+sie\setzten-sich (um-)zu+essen Brot
 074::37,25b und+sie\hoben Augen+ihre
 075::37,25c und+sie\sahen
 076::37,25d und+da
 077::37,25e eine-Karawane(-von) ISMAELITERN () eine-Kommende
 aus+GILEAD
 078::37,25f und+Kamele+ihre () Tragende Tragakant und+Mastix
 und+Ladanum
 079::37,25g Gehende (um-)zu+hinunterbringen ÄGYPTEN+wärts[0220]
KHS: 37,26a-27f **Nr.5**
 080::37,26a und+sprach JUDA zu Brüdern+seinen
 081::37,26b was () ein-Gewinn
 082::37,26c falls wir-werden/mod.\umbringen → Bruder+unseren
 083::37,26d und+verdecken-werden/mod.\wir → Blut+sein
 084::37,27a geht
 085::37,27b und+wir-werden/mod.\verkaufen+ihn den-ISMAELITERN
 086::37,27c und+Hand+unsere nicht sei gegen+ihn[0240]
 087::37,27d denn Bruder+unser
 088::37,27e Fleisch+unser () er
 089::37,27f und+hörten Brüder+seine

ENDE KHS:

Nr.5

TLE3* (TGE5*)**Block: 7**

090::37,28a und+vorbeikamen Männer MIDIANITISCHE Kaufleute
 091::37,28b und+sie\zogen
 092::37,28c und+sie\herausholten → den-JOSEF aus dem+Brunnen
 093::37,28d und+sie\verkauften → JOSEF an+die-ISMAELITER[0260]
 für+zwanzig Silber(stücke)
 094::37,28e und+sie\brachten → JOSEF ÄGYPTEN+wärts

TLE3* (TGE6*)

095::39,01b* und+kaufte+ihn ein-Mann ein-ÄGYPTISCHER
 aus+der-Hand der+ISMAELITER
 096::39,02c und+er\war in+dem-Haus Herrn+seines des+ÄGYPTISCHEN
 097::39,04a und+fand JOSEF Gefallen in+Augen+seinen
 098::39,04b und+er\bediente[0280] →+ihn
 099::39,04c und+er\Aufsicht-führen-ließ+ihn über Haus+sein
 100::39,04d und+die-Gesamtheit
 101::39,04e Existenz () ↑+ihm
 102::39,04f gegeben-hat&er in+Hand+seine
 103::39,06d und+nicht wahrgenommen-hat&er in+seiner(-Gegenwart)
 Etwas nämlich außer dem+Brot
 104::39,06e welches jener () ein-Essender

TLE3* (TGE7*)

Block: 8

105::39,06f und+war[0300] JOSEF ein-Schöner(-an) Gestalt
 und+ein-Schöner(-an) Aussehen
 KHS: 39,07b-09c Nr.6

106::39,07b und+erhob die-Frau Herrn+seines → Augen+ihrē
 (hin-)zu JOSEF

107::39,07c und+sie\sprach

108::39,07d leg-dich+doch mit+mir

109::39,08a und+er\weigerte-sich

110::39,08b und+er\sprach zu der-Frau Herrn+seines[0320]

111::39,08c da

112::39,08d Herr+mein nicht wahrgenommen-hat
 in+meiner(-Gegenwart)

113::39,08e was () in+dem-Haus

114::39,08f und+die-Gesamtheit

115::39,08g dessen Existenz () ↑+ihm

116::39,08h gegeben-hat&er in+Hand+meine

117::39,09a Nicht-Existenz(-von)+ihm () ein-Großer in+dem-Haus
 dem+selben von-her+mir

118::39,09b und+nicht zurückgehalten-hat&er[0340] vor+mir Etwas
 nämlich außer →+dir

119::39,09c denn du () Frau+seine
 ENDE KHS: Nr.6

4. (Ausdrucks-) SYNTAX**Literatur**

- BAAYEN, H; HALTEREN, H van: Outside the Cave of Shadows: Using Syntactic Annotation to Enhance Authorship Attribution. *Literary and Linguistic Computing* 11 / 3 (1996) 121-131. // EB
- Brauch, W: Programmierung mit FORTRAN. Eine Einführung in Basic FORTRAN IV. Stuttgart 1972: B.G. Teubner // I.2.5
- DAUSES, A: Systemcharakter und Relativität der Sprache. Stuttgart 1996: Steiner // EB
- Faculté des Lettres et Sciences Humaines (ed./eds.): *Statistique et Analyse Linguistique. Colloque de Strasbourg (20-22 avril 1964)*. Paris 1966: Presses universitaires de France // I.2.5
- HAN, Y S; PARK, H R (et.al.): An Upper Bound Estimate for the Entropy of Korean Texts. *Literary and Linguistic Computing* 11 / 3 (1996) 141-146. // EB
- HOLMES, D I; SINGH, S: A Stylometric Analysis of Conversational Speech of Aphasic Patients. *Literary and Linguistic Computing* 11 / 3 (1996) 133-140. // EB
- HUG, M: Das Menzerath-Gesetz in der *Vulgata in*: Grzybek, P; Köhler, R (ed./eds.): *Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday*. *Quantitative Linguistics* 62. Berlin 2007 245-257.: Mouton de Gruyter // I.2.5
- JANDL, E: für alle. Darmstadt 1974
- KREUZER, H; GUNZENHÄUSER, R (ed./eds.): *Mathematik und Dichtung. Versuche zur Frage einer exakten Literaturwissenschaft*. München 1969: Nymphenburger // I.2.0
- MARTINDALE, C; TUFFIN, P: If Homer is the Poet of the *Iliad*, then he may not be the Poet of the *Odyssey*. *Literary and Linguistic Computing* 11 / 3 (1996) 109-120. // EB
- McKENNA, W; ANTONIA, A: 'A Few Simple Words' of Interior Monologue in *Ulysses: Reconfiguring the Evidence*. *Literary and Linguistic Computing* 11 / 2 (1996) 55-66. // EB
- MERGENTHALER, V: Sehen schreiben - Schreiben sehen. *Literatur und visuelle Wahrnehmung im Zusammenspiel*. Hermaea 96. Tübingen 2002 ISBN 3-484-15096-3: Niemeyer // I.2.7
- VOß, W: *Statistische Methoden und PC-Einsatz*. UTB 1445. Opladen 1988 // I.2.5

zu Ziff. 4.01:

SYNTAX: Wahrnehmung von Objekten ↔ Einbildungskraft

zu: vsyst4.01 vsem7.221

aus: V. MERGENTHALER, Sehen schreiben - Schreiben sehen. Literatur und visuelle Wahrnehmung im Zusammenspiel. Tübingen 2002

(298) Das im Leser sich Vollziehende hat einen anderen, "objektiveren" ontologischen Status als allein durch Zeichen repräsentierte oder durch Unbestimmtheitsstellen zwar angeregte, doch vom Leser hervorgebrachte Vorstellungen, denn die Verknüpfungsleistung wird nicht durch die produktive (299) Lesetätigkeit des Rezipienten erbracht, sondern ist im Text als positives Faktum angelegt. Mit der hermeneutischen Komplettierung eines Textes, wie ihn die Wirkungsästhetik im Konzept der "Appellstruktur des Textes" anvisiert, hat dieses Verfahren nur noch strukturell zu tun, denn nicht eine sinnbildende Verknüpfungsleistung zwischen zwei "schematisierten Ansichten", sondern eine neuronal-mechanische Vorstellungssteuerung vollzieht sich hier. Wolfgang Iser unterscheidet (mit Jean-Paul Sartre)

zwischen Wahrnehmung und Vorstellung als zwei getrennten Vermögen (...), da für die Wahrnehmung immer ein Objekt vorgegeben sein muß, während die konstitutive Bedingung für die Vorstellung gerade darin besteht, daß sie sich auf ein Nicht-Gegebenes oder Abwesendes bezieht. In der Lektüre literarischer Texte müssen wir uns deshalb immer Vorstellungen bilden, weil die "schematisierten Ansichten" uns nur ein Wissen davon bieten, über welche Voraussetzungen der imaginäre Gegenstand vorgestellt werden soll. Es sind die sprachlich nicht manifesten Implikationen des Textes sowie seine Unbestimmtheitsgrade und seine Leerstellen, die die Einbildungskraft mobilisieren, damit der imaginäre Gegenstand als Korrelat des Vorstellungsbewußtseins erzeugt werden kann.

zu Ziff. 4.011:

Linguistik u. a. empirisch: Daten

zu:

aus: Ch. Lehmann, Daten - Korpora - Dokumentation in: G. ZIFONUN, W. KALLMEYER (eds.), Sprachkorpora - Datenmengen und Erkenntnisfortschritt. IdS Jahrbuch 2006. Berlin 2007. S. 9-27.

(11) Während die jüngeren Fortschritte der Hermeneutik in Richtung einer objektiven Hermeneutik (vgl. z.B. Oevermann 2002) unbestritten sind, ergibt das Kriterium der Objektivität doch eine Hierarchie der Wissenschaftlichkeit unter den bestehenden Disziplinen: Ein unangefochtenes Prestige von Wissenschaftlichkeit haben nur die logischen Disziplinen. Die empirischen Disziplinen werden seit Hempel 1966, soweit eben möglich, nach dem Modell der logischen Disziplinen aufgezogen - z.B. die Physik nach dem Modell der Mathematik. Die hermeneutischen Disziplinen hinwiederum gelten überhaupt nicht als *sciences*, sondern lediglich als *humanities*. Manche von ihnen, darunter die Psychologie, die Soziologie und die Linguistik, wollen am liebsten keine hermeneutischen Disziplinen sein. Sie behandeln die Unterscheidung zwischen Erzeugnissen der Mitmenschen und anderen erfahrbaren Gegenständen als irrelevant und gehen nach Möglichkeit empirisch oder gar logisch vor.

Im Falle der Linguistik ist dies bis zu einem gewissen Grade durch die Natur des Gegenstandes gerechtfertigt. Sprachzeichen haben eine wahrnehmbare Seite, Sprechakte sind beobachtbare raumzeitliche Ereignisse. Man kann sie empirisch angeben. Ferner ist Sprache, in ihrer kognitiven Dimension, die äußere Form des Denkens und insofern logischer Erkenntnis zugänglich. Beides ändert freilich nichts daran, dass Sprachtätigkeit, in ihrer sozialen Dimension, Verständigung ist und dass diese hermeneutischer Methodik bedarf. Der Gegenstand Sprache hat wesentlich diese drei Aspekte, den hermeneutischen, den empirischen und den logischen. Jeglicher Versuch, einen davon zu eskamotieren oder zu verabsolutieren, ist steriler Reduktionismus.

Dies vorausgeschickt, wird die Linguistik im folgenden überwiegend in ihrer Eigenschaft als empirische Disziplin behandelt. Dies ist einfach eine Folge der Themenstellung: Daten gibt es nur in empirischen Disziplinen.

1. Daten

1.1 Funktion von Daten

Ein Datum ist eine Repräsentation eines Phänomens aus dem Gegenstandsbereich einer Wissenschaft, die als gegeben angenommen wird. Die Repräsentationsbeziehung ist nicht frei manipulierbar, sondern durch die wissenschaftliche Methodik geregelt. Das ist eben die Basis dafür, dass das Datum als gegeben angenommen wird. Es vertritt für die Zwecke wissenschaftlicher Argumentation das Phänomen selbst, welches i.a. nicht zur Hand ist. Daten spielen in einer Disziplin eine methodische Rolle in dem (12) Maße, in dem sie eine empirische Disziplin ist. In induktiver Methodik dient das Datum als Indiz und empirische Evidenz, in deduktiver Richtung als Prüfstein im Text einer Theorie.

Eine Wissenschaft konstituiert sich gemäß ihrem Erkenntnisinteresse. Es gibt legitime oder jedenfalls in der Gesellschaft unangefochtene Erkenntnisinteressen, die nichts mit in der uns umgebenden Welt vorfindlichen und öffentlich beobachtbaren Phänomenen zu tun haben. Für die Linguistik aber gilt in dieser Hinsicht folgendes:

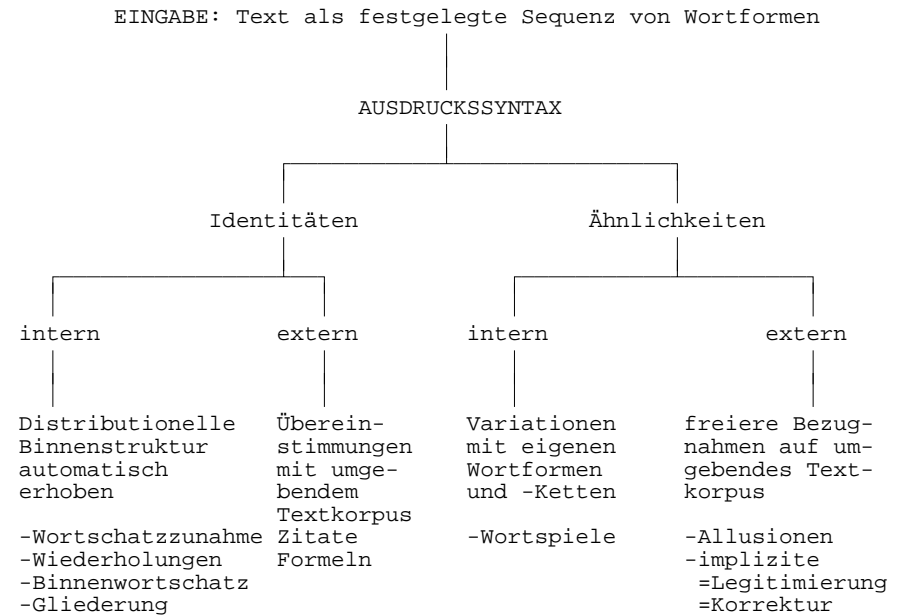
1. Eine Disziplin, die sich für eine empirische erklärt, muss ihren Gegenstandsbereich mit Bezug auf unabhängig von ihr selbst in der Welt vorfindliche und öffentlich beobachtbare Phänomene konstituieren. In dem Maße, in dem sie das nicht tut, ist es vielleicht eine logische oder hermeneutische Disziplin oder vielleicht auch Esoterik.
2. Es gibt in der uns umgebenden Welt vorfindliche und öffentlich beobachtbare Phänomene, nämlich Sprechakte im weitesten Sinne, an deren wissenschaftlicher Erforschung ein öffentliches Interesse und geradezu ein Auftrag besteht und für die keine andere Disziplin zuständig ist.
3. Die Linguistik selbst hat sich, in Gestalt herausragender Vertreter wie Leonard Bloomfield und Noam Chomsky, für eine empirische Disziplin (und gelegentlich gar, in rhetorischer Polarisierung, für eine Naturwissenschaft) erklärt. Das erfordert, dass sie ihren Gegenstandsbereich auf die in Nr. 1 beschriebene Weise konstituiert. Dessen vorgegebene Basis sind gerade die in Nr. 2 genannten Phänomene.
4. Jegliche Wissenschaft, die über reine Idiographie hinausgehen will, arbeitet mit Abstraktionen. Das gilt jedenfalls für die Linguistik und a fortiori für die Systemlinguistik und ihrem Gegenstand, das Sprachsystem. Von den notwendigen und angestrebten Abstraktionen sind zu allererst die Daten betroffen. Sie werden gewisser Eigenschaften entkleidet, denen das Erkenntnisinteresse nicht gilt. So entstehen sekundäre Daten, die oft eine sehr vermittelte Beziehung zu den Phänomenen haben.
5. Wie jegliche empirische Wissenschaft kann auch die Linguistik ihren Anspruch, eine solche zu sein, nur aufrecht erhalten, wenn sie die Beziehung ihrer Theorien und a fortiori ihrer Daten auf die Phänomene jederzeit methodisch kontrolliert und objektiv nachvollziehbar macht. In dem Maße, in dem sie das nicht tut, ist ihr Anspruch, eine empirische Disziplin zu sein, wissenschaftspolitische Rhetorik.
6. Gemäß dem in § 1 Gesagten ist die Linguistik sowohl eine empirische als auch logische als auch eine hermeneutische Disziplin. Sie hängt also (13) nicht ausschließlich von Daten ab, sondern gewinnt Erkenntnis auch auf andere Weise. Desto komplexer ist ihre Methodologie und desto größer die Gefahr der Manipulation von Ergebnissen.

zu Ziff. 4.02:

B. Deskription und Interpretation des Einzeltextes

zu: vsyst4.32

(1) Ausdrucks-Syntax oder SYNTAKTIK



4.1 Zeichentheorie - Distributionsanalyse - Computer**Literatur**

- ANDERSEN, F I; FORBES, A D: The vocabulary of the Old Testament. Rome 1989 // I.2.7
- BARR, G K: The Use of Cumulative Sum Graphs in Literary Scalometry. Literary and Linguistic Computing 12 (1997) 103-111.
- BEST, K-H; JINYANG, Z: Zur Häufigkeit von Wortlängen in Texten deutscher Kurzprosa (mit einem Ausblick auf das Chinesische) in: KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 19-30. // I.2.7
- EVEN-SHOSHAN, A: A new Concordance of the Bible. Thesaurus of the Language of the Bible. Hebrew and Aramaic. Roots Words, Proper Names, Phrases and Synonyms. Jerusalem 1981 // I.2.7
- GELB, I J: Sequential reconstruction of Proto-Akkadian. Assyriological studies 18. Chicago 1969
- GELB, I J; BARTELS, J u.a: Computer-aided analysis of Amorite. Assyriological studies 21. Chicago 1980 // I.2.7
- GROSS, S: Lese-Zeichen. Kognition, Medium und Materialität im Lese-prozeß. Darmstadt 1994: Wiss. Buchgesellschaft // I.2.0
- HARDMEIER, C; TALSTRA, E: Sprachgestalt und Sinngehalt. Wege zu neuen Instrumenten der computergestützten Textwahrnehmung. ZAW 101 (1989) 408-428.
- HELBIG, G: Entwicklung der Sprachwissenschaft seit 1970. Unter dem besonderen Aspekt der Grammatik-Theorie. WV studium Bd. 161 1990: Westdeutscher Verlag
- HO LEE, D: Rückläufiges Wörterbuch der deutschen Sprache. Berlin 2005 ISBN 3 11 018197 5: de Gruyter // I.2.7
- LEHMANN, C: Daten - Korpora - Dokumentation in: KALLMEYER,W; ZIFONUN,G (ed./eds.): Sprachkorpora - Datenmengen und Erkenntnisfortschritt. Jahrbuch des IdS 2006. Berlin 2007 9-27. 978-3-11-019273-5: de Gruyter // I.2.7
- LEVISON, M; LESSARD, G: A Multi-Level Approach to the Detection of Second Language Learner Errors. Literary and Linguistic Computing 15 / 3 (2000) 313-322.
- LISOWSKY, G: Konkordanz zum hebräischen Alten Testament. Stuttgart 1958
- MARTEN, S; SPERFELD, S: Kommunikationsdesign. Zur Sinnhaftigkeit der Materialität von Kommunikaten in: POHL,I (ed./eds.): Semantik und Pragmatik-Schnittstellen. Sprache, System und Tätigkeit 59. Frankfurt/M 2008 109-141. 978-3-631-57061-6 br: P. Lang // I.2.4
- MUTHMANN, G: Rückläufiges deutsches Wörterbuch. RGL 78. Tübingen 1988 // I.2.7
- PAPROTTÉ, W: Theorie und Empirie in der Sprachwissenschaft am Beispiel neuerer korpuslinguistischer Forschung in: KÖNIG, P P; WIEGERS, H (ed./eds.): Satz - Text - Diskurs. Akten des 27. linguistischen Kolloquiums,

Münster 1992. Band 2. Linguistische Arbeiten 313. Tübingen 1994 19-26.: Niemeyer // I.2.4

- RADTKE, P: Die Kategorien des deutschen Verbs: zur Semantik grammatischer Kategorien. TBL 438. Tübingen 1998: Narr // I.2.4
- SCHWEIZER, H: Die Sprache der Zeichenkörper. Textinterne (Ausdrucks-)Syntax zu Dan 8 in: BADER, W (ed./eds.): "Und die Wahrheit wurde hinweggefegt". Linguistische Analysen von Daniel 8. THLI 9. Tübingen 1994 17-30.: Francke // Raum B106
- SEARLE, J R: Ist der menschliche Geist ein Computerprogramm?. Nein. Ein Programm manipuliert Symbole nur, während das Gehirn ihnen Bedeutung beilegt. Wie also soll ein Programm denken können, wenn es nicht versteht, womit es eigentlich hantiert?. Spektrum der Wissenschaft (1990) / März 40-47.
- WETTLER, M: Sprache, Gedächtnis, Verstehen. Berlin 1981 // I.2.0

4.11 Zeichentheorie

Literatur

SCHNEIDER, J G . Zeitschrift für Semiotik 30/1-2 (ed./eds.):
Medialität und Sozialität sprachlicher Zeichen
(2008)

4.12 Stichwort: "Syntax"

zu Ziff. 4.121:

Zunehmende Trennung: Ausdruck vs. Inhalt

zu:

aus: RADTKE, P, Die Kategorien des deutschen Verbs. Zur Semantik grammatischer Kategorien. TBL 438. Tübingen 1998.

(42f) Bevor ich mich der inhaltlichen Seite der Verbalkategorien zuwende, möchte ich festhalten, welche Konsequenzen all diese Überlegungen für den weiteren Verlauf dieser Arbeit haben. Dazu nehme ich die folgenden vier terminologischen Festlegungen vor.

Erstens: Sämtliche Verbformen, die sich aus einem Basismorphem und einem grammatischen Marker zusammensetzen, bezeichne ich als flektiert. Infinitive Formen wie beispielsweise der Infinitiv *schreiben*, das Partizip I *schreibend* gelten ebenso als flektiert wie die finiten Formen (*ich*) *schreibe* oder (*ich*) *habe geschrieben*. Der grammatische Marker ist - im Gegensatz zum Derivationsmorphem - ein wortartenspezifischer Markierungstyp. Er wird über Regeln, die ausschließlich für Verben gelten, die Flexionsregeln, realisiert.

Zweitens: Die Begriffe "Flexion" und "Finitheit" werden voneinander getrennt; flektierte Verben sind nicht notwendigerweise finit. Finit sind Verbformen genau dann, wenn sie dazu dienen, das Bündel der Verbalkategorien auszudrücken, d.h. jeweils genau eine Kategorie aus den Kategorisierungen "Person", "Numerus", "Tempus", "Modus" und "Genus verbi". Eine Form wie (*ich*) *schreibe* gilt als finit; eine Form wie der Infinitiv Perfekt Aktiv *geschrieben haben* gilt deshalb als infinit, weil sie nicht das gesamte Kategorienbündel, sondern nur einen Teil davon ausdrückt.

Drittens: Als Verbalkategorien bezeichne ich nur die Kategorien, die für die Finitheit einer Verbform sorgen, d.h. im Kategorienbündel auftreten können. So sind beispielsweise "1. Person", "Singular", "Präsens", "Indikativ", "Aktiv", Verbalkategorien. Andere Kategorien, die ebenfalls spezifischerweise am Verb markiert werden, gehören nicht dazu. So ist "Infinitiv" zwar eine Kategorie des Verbs, aber eben keine Verbalkategorie.

Viertens: Da die Verbalkategorien im Deutschen nicht nur auf morphologischem, sondern auch auf syntaktischem Wege symbolisiert werden, muß man auch zusammengesetzte Formen berücksichtigen. Eine Form wie *wird geschrieben* drückt die Kategorie "Passiv" genau durch das Zusammenspiel von morphologischen und syntaktischen Markern aus. Aus eben diesem Grund gilt für mich die gesamte Form als finit und nicht nur das Hilfsverb, das üblicherweise als der finite Bestandteil der Konstruktion angesehen wird...

(43) Wir haben gesehen, daß nur mit dem Verweis auf inhaltliche Gegebenheiten die formal verschieden symbolisierten Verbalkategorien homogene Klassen bilden.

zu Ziff. 4.122:

Ausdruck vs. Inhalt//Plerem vs. Morphem

zu:

aus: RADTKE, P, Die Kategorien des deutschen Verbs. Zur Semantik grammatischer Kategorien. TBL 438. Tübingen 1998.

(18ff) Das Bergenholt & Mugdanche Verbparadigma umfaßt für das Deutsche lediglich die Formen, die bezüglich der Kategorien "Präsens" und "Präteritum", "Singular" und "Plural" sowie "1. Person", "2. Person" und "3. Person" markiert sind. Hinzu kommt die Unterscheidung bezüglich des Flexionsklassentyps, d.h. die Unterscheidung zwischen starken und schwachen Verben. Es ergibt sich bei binären Merkmalskontrasten - [-Prät] entspricht "Präsens", [-Plural] entspricht "Singular" - folgendes Bild:

Personalsuffix	-stark		+stark	
	-Prät	+Prät	-Prät	+Prät
-Plural	1 Person 2 Person 3 Person	e st t	e est e	e st t O st O
+Plural	1 Person 2 Person 3 Person	en t en	en et en	en t en en t en

Ergänzen müßte man dieses Schema noch um die Formen des Imperativs. Schon hier wird die morphologische Argumentationslinie der Autoren deutlich. Entscheidendes Kriterium ist, ob etwas - eine Kategorie, ein Flexionsklassentyp - sich am Ausdruck belegen läßt...

(19) Bergenholz & Mugdan setzen nicht - wie weit verbreitet - das Morphem als die kleinste bedeutungstragende Einheit an. Bei ihnen sind es, der strukturalistischen Tradition der Kopenhagener Schule folgend, die sogenannten Plereme: "Das Minimalzeichen, das im folgenden kurz *Plerem* genannt werden soll, hat einen Ausdruck, der im allgemeinen aus mehreren Phonemen bzw. Graphemen besteht, und einen Inhalt, der sich in mehrere Bestandteile aufgliedern läßt." Exemplarisch analysieren sie das Verb *wollten*: Den Ausdrücken *wollt*, sowie *en* entspricht der Inhalt "woll", "Präteritum" sowie "3. Person" und "Plural". Ausdrucksgleiche Plereme mit verschiedenem Inhalt werden als Morph bezeichnet, wohingegen ausdrucksverschiedene Plereme mit gleichem Inhalt als Morphem bezeichnet werden...

Für die Zugehörigkeit zu einem Morph ist die Ausdrucksseite (20) des Plerems maßgeblich, für die Zugehörigkeit zu einem Morphem die Inhaltsseite. Dem Morph *en* entsprechen beispielsweise die Plereme *en* ("3. Person"/"Plural") und *en* ("1. Person"/"Plural"). Dem Morphem "Präteritum" nach der oben wiedergegebenen Übersicht die Plereme *e*, *O*, *est*, *st* usw. mit dem Inhalt "Präteritum". Soll die Zugehörigkeit eines Plerems zu einem Morphem ausgedrückt werden, so wird es Allomorph genannt: "Diejenigen Plereme, die zu demselben Morphen gehören, werden als dessen *Allomorphe* bezeichnet." Bei Allomorphen handelt es sich also um nichts anderes als um bereits bezüglich ihres Inhalts klassifizierte Plereme...

Das Morphem selbst kann nie allein ausdrucksseitig identifiziert werden. Es ist eine abstrakte Einheit, die sich zwar inhaltlich bestimmen läßt, ausdrucksseitig aber immer durch ein oder mehrere Allomorphe repräsentiert wird. Alle Morpheme sind also via Analyse der Plereme zugänglich.

zu Ziff. 4.123:

Schwierige "Konjugation"

zu: vsem10.1

aus: RADTKE, P, Die Kategorien des deutschen Verbs. Zur Semantik grammatischer Kategorien. TBL 438. Tübingen 1998.

(13-15) Unklar ist bereits der Begriff der Verbflexion. Was soll darunter verstanden werden? ... Man geht also im Deutschen - versteht man diesen Begriff rein formal - wie im Lateinischen von einer "conjugatio", einer Verbindung zwischen Verbalstamm und entsprechender Endung, aus.

Doch eine solche Anlehnung an traditionelle Konzepte birgt Schwierigkeiten. Problematisch sind alle drei hier verwendeten Begriffe...

(14) Erstens: Schwierigkeiten bereitet bereits die Vorstellung einer Verbindung von Stamm und Endung. Für Formen wie *schreib-e* ist sie fraglos zutreffend: *schreib* ist der Stamm, *-e* die Endung. Aber was sollte bei einer Form wie *schrieb* der Stamm und was die Endung sein? ...

Zweitens: Nimmt man nun als Konsequenz an, der Ablaut übernehme in einer Verbform wie *schrieb* Rolle und Funktion einer Endung, so stößt man zudem auf die Schwierigkeit, einen festgefügtten Stamm identifizieren zu können. Der sollte im Rahmen eines Konjugationsmodells eigentlich *schreib* heißen. Erlaubt man nun Veränderungen am Stamm, so müßte man annehmen, es gäbe hier feste und variable Bestandteile, was die Redeweise von "dem" Stamm verkompliziert und letztlich zu Aussagen führen kann wie: der Stamm von *schreiben* besteht aus *schr b*. ...

Drittens: Der Konjugationsbegriff ist darüber hinaus auch deshalb für das Deutsche unpassend, weil er - rein formal betrachtet - Phänomene mitabdeckt, die nach genereller Auffassung nicht darunter gefaßt werden sollen. Der Infinitiv (*schreiben*), das Partizip I (*schreibend*) und das Partizip II (*geschrieben*) weisen nämlich sehr wohl Endungen auf (letzteres zusätzlich noch den Ablaut und ein Präfix - oder handelt es sich um ein Zirkumfix *ge-en?*) und entsprechen so der oben genannten Regel "Stamm + Endung". Unter Konjugation wird jedoch nur das Bilden von finiten Verbformen verstanden, d.h. von Formen, die bezüglich der Verbalkategorien markiert sind...

(15) Viertens: Problematisch ist der Konjugationsbegriff nicht nur deshalb, weil er auf der einen Seite möglicherweise zu weit gefaßt ist - und so infinite Verbformen einschließt -, sondern auch, weil er auf der anderen Seite zu eng ist. Er schließt nämlich Formen aus, die man gerne als finit bezeichnen würde, Formen wie *habe geschrieben* beispielsweise, die wie die synthetische Form *schreibe* imstande sind, Tempus auszudrücken. Nur bei synthetischen Verbformen könnte man einigermaßen gerechtfertigt von einem morphologischen Verfahren (der Konjugation) sprechen, bei analytischen hingegen verläßt man bereits den Bereich der Morphologie und gelangt in den Zuständigkeitsbereich der Syntax. ... Da allein die Konjugation zu finiten Verbformen führt, kann konsequenterweise nur der Bestandteil *habe* als finit gelten, nicht aber die Form *habe geschrieben* insgesamt.

zu Ziff. 4.124:

Lexik: Formseite + Inhaltsseite

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(6)

a) Die Formseite eines sprachlichen Zeichens ist die jeweilige Ausprägung auf der Formebene.

b) Die Inhaltsseite eines sprachlichen Zeichens ist die jeweilige Ausprägung auf der Inhaltsebene.

Hiernach stellen sich Morphologie und (lexikalische) Semantik als Grundlagendisziplinen für die Lexikologie heraus, denn von der Morphologie dürfen wir formorientierte Kriterien für die lexikalischen Elemente und von der (lexikalischen) Semantik inhaltsorientierte Kriterien, am besten in Form einer Bedeutungskonzeption, für die lexikalischen Elemente erwarten. Die Aussonderung der Morphologie für die formorientierten Kriterien erklärt sich unter dem Aspekt der systematischen Beschreibung für den Wortschatz und das Lexikon. Denken wir jedoch an das mentale Lexikon, dann haben natürlich Phonologie und Graphematik mindestens den gleichen Stellenwert, da hier die Formanalyse häufig unmittelbar auf die konkrete Realisierung im Medium übergeht.

zu Ziff. 4.125:

Lexikalische Einheit

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(31) Das Wort als sprachliches Zeichen verlangt nach einer Formseite und einer Inhaltsseite. Die ideale Vorstellung dabei ist, daß die Verbindung zwischen Form und Inhalt die Klammer für die Einheit des Wortes darstellt.

a. Die Bedeutungslehre oder Semasiologie kümmert sich um die Ausprägungen auf der Inhaltsebene für eine vorgegebene Formseite.

b. Die Bezeichnungslehre oder Onomasiologie kümmert sich um die Ausprägungen auf der Formebene für eine vorgegebene Inhaltsseite.

(33) Die Inhaltsseite wird für die Klammerbildung entscheidend, sobald es sich um eine Form mit mehreren Inhalten handelt. Können wir hierbei noch einen gewissen Zusammenhang zwischen den Inhalten erkennen, gilt dies als deutlicher Hinweis auf die Einheit des Elementes - man spricht in diesem Fall von einem **polysemen** Element, während im Falle von Inhalten ohne einen erkennbaren Zusammenhang wir eher von verschiedenen Einheiten mit zufällig identischer Form reden - man spricht in diesem Fall von **homonymen Elementen**.

zu Ziff. 4.126:

Lemma / Lexem

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(39) Unter einem Lemma verstehen wir die repräsentative Wortform im Wörterbuch - auch Nennform genannt - oder die repräsentative Wortform in einem Register/Index als Stichwort bzw. Eintrag.

Lat. *lemma* "Überschrift" kommt aus dem Griechischen mit der Bedeutung "Aufgegriffenes". In der mittelalterlichen Emblemforschung taucht deshalb der Begriff "Lemma" auch als Element des dreiteiligen Emblems (Bild, Lemma als Motto [Inscriptio], Subscriptio) auf.

Unter einem Lexem verstehen manche Autoren/Autorinnen eine abstrakte Einheit des Wortschatzes, im Gegensatz zu den konkreten Wortformen, dort dann häufig "Wort" genannt. Die Angabe des Lexems erfolgt meist über das Basismorphem bzw. den "Stamm"; z.B. schreib-.

(40) Bei der im nächsten Kapitel vertretenen Paradigmaauffassung entsteht eigentlich kein Bedarf für den Begriff "Lexem", zumal meist unklar bleibt, ob es sich um eine Einheit mit Form- und Inhaltsausprägung oder nur mit Formausprägung handelt.

zu Ziff. 4.127:

€ - Zeichen für Tauschwert

zu:

aus: R. KRÜGER, Der EURO: Zur historischen Semiologie eines Namens und Münzzeichens: Zeitschrift für Semiotik 30/1-2 (2008) 115-136.

(116) In zwölf Staaten Europas vollzog sich im Kontext der seit dem 1. Januar 2002 neu eingeführten Euro-Währung einer der größten Prozesse der gemeinsamen Einübung eines neuen Zeichensystems. Lediglich die Hellenisierung und die anschließende Latinisierung der mediterranen Welt sowie die Hispanisierung Amerikas dürften - soweit wir den europäischen Kulturraum überblicken - semiologische Revolutionen oder Umstrukturierungen noch größeren Ausmaßes gewesen sein.

Im Zuge der Euro-Einführung vollzogen sich komplexe Prozesse der Einübung des Umgangs mit neuen Zeichen. Münzen sind allgemeine Äquivalente und damit abstrakte Zeichen für den Tauschwert von Waren, die für den Menschen als Konsumenten jedoch vor allem einen Gebrauchswert darstellen. (Nur sie sind der Gegenstand der nachfolgenden Studie). Damit Münzen als derartige Zeichen anerkannt werden, müssen sie standardisierte materielle, graphische und taktile Merkmale aufweisen, die allen an den Ware-Geld-Transaktionen beteiligten Personen bekannt sind.

Das Gepräge einer Münze ist ein Versprechen. Es sichert dem Nutzer dieser Münze zu, dass sie erstens tatsächlich von autorisierter Stelle erzeugt wurde (117) und dass sie zweitens auch zu dem eingprägten Wert in Zahlung genommen wird. Gepräge und sinnliche Beschaffenheit einer Münze sind also Eigenschaften, welche die Annahme begründen, dass der Gebrauchswert des Objekts Münze in seinem Tauschwertversprechen liegt, welches darin besteht, dass sie zu einem vereinbarten Kurs für andere Tauschwerte in Zahlung genommen wird. Nur unter diesen Bedingungen kann eine weitere Eigenschaft der Geldzeichen gesichert werden: ihre universelle Gültigkeit in dem vom Münzmeister definierten geographischen Raum (sowie in dem geographischen und sozialen Raum, in dem diese Münze gegebenenfalls stillschweigend Anerkennung findet). Diese wird durch die Definition von Geldzeichen als gesetzlichen Zahlungsmitteln garantiert: Es wird per Gesetz festgelegt, dass jeder verpflichtet ist, diese Geldzeichen als Zahlung entgegenzunehmen (Annahmewang). Der Annahmewang ist faktisch die Durchsetzung des staatlichen Münzmonopols bis hin zur kleinsten Transaktion. Dazu jedoch muss das Geldzeichen als solches erkennbar und jedem bekannt sein. Eine internationale Währung benötigt demnach auch eine international bekannte Standardisierung der Merkmale, an denen sie von allen als gültiges Zahlungsmittel erkannt werden kann. Zudem müssen Verfahren gefunden werden, mit denen die Einübung des Gebrauchs der Geldzeichen vollzogen werden kann. Wird der Gebrauch von Geldzeichen von jedem einzelnen üblicherweise im Kontext der Akkulturation in einer konkreten Gesellschaft erlernt, so stellt die Neueinführung von Geldzeichen die gesamte Gesellschaft vor eine komplexe Lernaufgabe.

4.13 Stichwort: "Distribution"

Literatur

- DIMPEL, F M: Computergestützte textstatistische Untersuchungen - an mittelhochdeutschen Texten. Tübingen 2004 ISBN 3-7720-8041-3: A. Francke // I.2.7
- LUONG, X; JUILLARD, M u.a.: Trees and After: The Concept of Text Topology. Some Applications to Verb-Form Distributions in Language Corpora. Literary and Linguistic Computing 22 / 2 (2007) 167-186.

zu Ziff. 4.131:

Rekursion

zu: vsyst4.11

aus: Utz Maas, Grundkurs Sprachwissenschaft Teil I. Die herrschende Lehre. München 21974.

(184f) Wir haben syntaktische Konstruktionen vor uns. Diese reduzieren wir auf ein nukleares Element, indem wir sie durch endlich viele Anwendungen einer Operation aus ihm entwickeln und dabei die Anzahl der Anwendungen speichern. Ein solches Verfahren heißt eine *Rekursion*; das beste Beispiel dafür ist die Entwicklung der natürlichen Zahlen, wie wir sie in Kap. 4.2. vorgeführt haben: Sie erinnern sich, daß jede Zahl durch den Rückgriff (die "Rekursion") auf das Anfangselement 0 reduziert werden konnte. Die Zahl 4 ist so erklärt als der vierte Nachfolger (die Rekursion geschieht bei den Zahlen ja mit Hilfe der Nachfolge-Operation) von 0: $4 = \text{der } 4^{\text{te}} \text{ O}'$. Die Zahlen sind aber auch das beste Beispiel für einen Algorithmus: der Ausdruck "berechnen" kommt ja aus dem Operieren mit Zahlen. Wenn man die Syntax als einen Algorithmus versteht, dann heißt das, daß man sie ebenso berechenbar macht, wie das Probleme des Operierens mit natürlichen Zahlen sind...

Harris stellt hier also einen syntaktischen Algorithmus vor. Die Operation, die ihm dabei die Rekursion definiert (entsprechend zur Nachfolgeoperation bei den Zahlen), ist das Anfügen eines Komplementes (wir erinnern uns aus Kap. 7.2 an die mengentheoretische Bedeutung von "Komplement": Sei N ein ausgezeichnetes Element ["N" wie "Nukleus"], dann zerfällt die Folge bzw. Konstruktion, die N enthält, in N und \bar{N}). Die Komplementierung ist nun rekursiv; zunächst definieren wir:

- (1) N ist eine Konstruktion
- (2) Mit jeder Konstruktion K ist auch $K + \bar{K}$ eine Konstruktion.

Jetzt läßt sich aber jede Konstruktion in endlich vielen Schritten so zerlegen, daß wir sie auf ihren Nukleus N reduzieren können. Für jede Konstruktion K gilt nämlich, daß sie entweder nur aus Nukleus besteht ($K = N$), oder aus einer Konstruktion K' plus ihrem Komplement ($K = K' + \bar{K}'$); für dieses K' wiederum gilt, daß es entweder ein Nukleus ist ($K' = N$), oder es besteht wieder aus einer Konstruktion K'' plus ihrem Komplement ($K' = K'' + \bar{K}''$) usw., bis wir beim Nukleus angekommen sind.

Beispiel (HS):

"seitenwind" = $K (\neq N)$

"seiten" + "wind" : $K = K' + \bar{K}'$

"seit" + "en" : $K' = K'' + \bar{K}''$

"sei" + "t" : $K'' = K''' + \bar{K}'''$

"t"	: $\bar{K}''' = N_3$	B_3
"sei"	: $K'''' = \bar{K}''''' + K''''$	
"s"	: $\bar{K}''''' = N_1$	B_1
"ei"	: $K'''' = N_2$	
"e" + "n"	: $\bar{K}'' = N_4 + N_5$	$B_4 + B_5$
"wind"	: $\bar{K}' = N_6$	

Unterschiede:

- N_2 und N_6 sind im Deutschen als eigene, selbstständige Wortformen nachweisbar: sie sind nicht weiter zerlegbar, so dass - u.z. oberhalb der Einzelbuchstaben - ein Nukleus nachzuweisen wäre (in der Liste des deutschen Wortschatzes).
- N_1 , N_3 , N_4 und N_5 sind Einzelbuchstaben, als selbstständige Wortformen im Deutschen nicht gebräuchlich, aber als unselbstständige, elementare Bildelemente. Bildelemente sind aber keine Konstruktion (=Nukleus). Folglich andere Etikettierung: B_1 , B_3 , B_4 und B_5 .
- Nur unter Einbeziehung von Bedeutungswissen - was hier aber verboten ist - könnte man differenzieren:
 - Die Hinzufügung von "s" zu "ei" kreierte ein neues Wort: "sei". Ähnlich "t" angefügt an "sei": "seit".
 - Dagegen erweitern "e" + "n" nur K'' , ermöglichen seine Verbindbarkeit, schaffen aber kein neues Wort.]

zu Ziff. 4.132:

Distributionsanalyse/Ebenentrennung/Textrezeption

zu: vprag1.71 vsyst4.03

aus: Utz Maas, Grundkurs Sprachwissenschaft Teil I. Die herrschende Lehre. München ²1974.

(151-153) Voraussetzung für die strikte Durchführung des distributionalistischen Programms und damit Bestandteil des Distributionalismus ist die *Homogenität* der Daten. Wir haben schon im vorigen Kapitel gesehen, daß die Homogenitätsforderung ein Korrektiv für die strikt empirische Vorgehensweise nötig gemacht hat: die Verfahren können nicht auf ein beliebiges Corpus angewandt werden, sondern nur auf ein Corpus, das ausschließlich Daten aus einem Idiolekt enthält. Ein Idiolekt war aber nichts anderes, als die Datenmenge, die in Hinblick auf die gegebenen Verfahren äquivalent ist. Die Legitimation einer solchen Konstruktion mit außermethodischen Überlegungen ist ja, wie ich mit Hinweis auf Kritiker wie Pike angemerkt habe, nicht berechtigt: im Verfahren wie im Bewußtsein der Sprecher wie der Sprachgemeinschaft gibt es keine Entsprechung zu Idiolekten.

Eine ähnliche Konstruktion, die als Korrektiv zur Rettung der strikten Verfahren eingeführt wird und damit zugleich eine Abweichung vom empiristischen Wissenschaftspostulat notwendig macht, ist die strikte Trennung der grammatischen Ebenen (Phonologie, Morphologie, Syntax usw.) - schon bei Bloomfield haben wir ja gesehen, daß jede Ebene unabhängig mit eigenen Axiomen und Definitionen aufgebaut wird, wobei der Aufbau aller Ebenen aber strikt parallel vor sich geht. Auch hier gilt wieder, daß diese Unabhängigkeit der Ebenen eine notwendige Annahme für die angestrebte Formalisierung der Theorie ist: die distributionellen Relationen sind ja immer nur auf einer gegebenen Menge definiert und können also immer nur solchen Eigenschaften (bzw. Prädikaten) Rechnung tragen, die mit dieser Menge definiert sind. Die Verfahren der Phonologie operieren aber auf einer Menge von phonetisch klassifizierten Daten, die der Morphologie auf einer Menge von phonologisch definierten Daten und schließlich die der Syntax auf einer Menge von morphologisch definierten Daten. Es versteht sich also, daß dabei nicht auf Eigenschaften zurückgegriffen werden kann, die bei einem gegebenen Stand der Analyse (einer gegebenen Ebene also) noch nicht definiert sind - insbesondere ist ein Rückgriff auf die Bedeutung der Äußerung nicht möglich, wenn diese noch nicht in ihrer Gesamtheit auf allen "niedrigeren" Ebenen analysiert ist. Wir können also für den Distributionalismus noch ein grundlegendes Axiom formulieren, das insbesondere von dem Hauptsatz vorausgesetzt wird: "Jede Sprache ist beschreibbar auf verschiedenen unabhängigen Ebenen, von denen jede darstellbar ist als eine endliche Menge von verschiedenen Elementen."

Nun leuchtet aber ein, daß diese Trennung bzw. Unabhängigkeit der Ebenen keinen Sinn macht, wenn die grammatische Analyse auch nur annähernd etwas mit der Analyse der Äußerungen durch den Hörer (und Sprecher) beim Sprechakt zu tun haben soll. Für diese gilt sicherlich (das haben alle psychologischen Untersuchungen in dieser Richtung bestätigt), daß dem Verstehens- und auch dem Sprechprozeß eine Art Strategie zugrundeliegt, die erst eine Art Hypothese über den Inhalt (die Bedeutung) der folgenden Äußerungen macht, und die dann anhand dieser Hypothese punktuell aufgrund lautlicher und grammatischer Kriterien ihre Richtigkeit überprüft - falls man das in einem Grammatikmodell mit verschiedenen Ebenen

darstellt: eine Überprüfung auf allen Ebenen gleichzeitig also. Es scheint auch so zu sein, daß dabei eine Strategie für das Durchführen der Stichproben zugrundeliegt, die auf die Überprüfung eines großen Teils der Äußerung verzichten kann, diesen also "überhört" bzw. "verschluckt"; das erklärt auch das hohe Maß an Routinisierung beim Sprechen bzw. Hören - und die allgemein zu beobachtende Angewohnheit von Hörern, bei einer lebhaften Diskussion dem Sprecher "ins Wort" zu fallen, d.h. ihn zu unterbrechen, bevor er seine Äußerung im grammatischen Sinne zuende gesprochen hat. Die Annahme, daß Sprecher und Hörer erst vollständig die phonologische, dann die morphologische, dann die syntaktische und dann die semantische Struktur hervorbringen, ist schlicht absurd.

(214-216) Jedem Element x einer Sprache ist dabei seine Klasse zugeordnet, d.h. seine Distribution $D(x) = \{P_1, P_2, \dots, P_n\}$. Jeder Position P_i entspricht aber eine Folge $S_i = (\alpha_{s_i}, P_i, \omega_{s_i})$. Bezeichnen wir die Anfangsfolge (also alle α_{s_i}) aus $D(x)$ mit A_x und die Endfolge (also alle ω_{s_i}) mit Ω_x , so ergibt sich für ein gegebenes Element a an einer gegebenen Stelle in einem Text, daß jedes Element y , das auf a in diesem Text folgen kann, ein Element aus Ω_a sein muß. Ein solches Element sei b . Für alle auf b folgenden Elemente z gilt nun, daß sie sowohl ein Element aus Ω_a wie aus Ω_b sein müssen, also $z \in \Omega_a \cap \Omega_b$ usw. Das folgende Beispiel kann das illustrieren.

Der Anfang einer Äußerung sei #*der* + x , wobei x als Variable für das folgende Element steht. *der* ist ein Element der Morphemklasse "Art(ikel)", deren Distribution ich hier in ein paar Beispielen andeute:

D (Art) = Nomen: der Mann
 Verb: der hat es ihm gegeben
 Proform: der da!
 Adjektiv: der diesjährige Showmaster
 Adverb: der äußerst gerissene Showmaster
 #: das war der

Das folgende Element sei *diesjährige*, also ein "Adj(ektiv)": #*der* + *diesjährige* + x .

D(Adj) = Nomen: der diesjährige Showmaster
 Verb: der diesjährige ist besser als der vom letzten Jahr
 Proform: der diesjährige da!
 Adjektiv: der diesjährige gerissene Showmaster
 Adverb: der diesjährige äußerst gerissene Showmaster
 Konjunktion: der diesjährige, wie wir alle wissen, äußerst gerissene Showmaster

#: das war der diesjährige
 Jetzt gilt für $x: x \in D(\text{Art}) \cap D(\text{Adj})$ bzw. $x \in D(\text{Art}+\text{Adj})$.
 Das folgende Element sei *Showmaster*, also ein "Nomen": #*der* + *diesjährige* + *Showmaster* + x .

D(Nomen) = Nomen: der diesjährige Showmaster Franz-Josef
 Verb: der diesjährige Showmaster spinnt
 Proform: der diesjährige Showmaster dahinten
 Artikel: der diesjährige Showmaster, ein prächtiger Mensch, ...
 Konjunktion: der diesjährige Showmaster, wie wir alle wissen, ein prächtiger Mensch, ...

#: das war der diesjährige Showmaster
 Jetzt gilt, daß $x \in D(\text{Art}) \cap D(\text{Adj}) \cap D(\text{Nomen})$.
 Das folgende Element sei *ist*, also ein Verb:
 #*der* + *diesjährige* + *Showmaster* + *ist* + x .
 D (Verb) = Nomen: der diesjährige Showmaster ist Franz-Josef
 Proform: der diesjährige Showmaster ist da
 Adjektiv: der diesjährige Showmaster ist gerissen
 Adverb: der diesjährige Showmaster ist äußerst gerissen
 Artikel: der diesjährige Showmaster ist ein prächtiger Mensch
 Konjunktion: der diesjährige Showmaster ist, wie wir alle wissen, ein prächtiger Mensch

-
- DISTRIBUTIONALISMUS: - erste Hälfte der 1950er Jahre
- richtig/wichtig: Verteilung der Sprachphänomene "sagt" etwas. Betonung statistischer Betrachtung.
 - Problem: Was ist ein *Sprachphänomen*? Wessen Verteilung soll gemessen werden? Wörter oder Bedeutungsfunktionen? - Man blieb bei traditionellen Grammatikkategorien (= Ausdruck + Bedeutung), also bei einem Mix aus Wort und Bedeutungsfunktion.
 - Systemtheoretisch (und semiotisch) muss getrennt werden: nur so kann die Gleichartigkeit der Elemente (Ausdrücke bzw. Bedeutungsfunktionen) zum Zug kommen.
 - Methodische Zwickmühle im traditionellen Konzept: Berechnet man Verteilung der Ausdrücke (Wörter), kann man von *Bedeutungsseite* her widerlegt werden (weil eine Bedeutungsfunktion immer auch noch anders realisiert werden kann), und umgekehrt.
 - Deswegen bürgerte sich die Missachtung des DISTRIBUTIONALISMUS ein.
 - Der Fehler liegt nicht im statistischen Ansatz, sondern im Grammatikkonzept.
 - Sobald Ausdrucksseite und Bedeutungsseite klar getrennt werden, ist auf *beiden Seiten* die Beachtung der Distributionen sehr ergiebig.

In H. Schweizer, Computergestützte Textinterpretation, THLI 7, Tübingen 1995, finden sich Distributionsanalysen auf *beiden Seiten*: (a) der Ausdrucksebene (Beitrag BADER), (b) der semantisch-pragmatischen Kategorien (SCHWEIZER - vgl. die (Farb-)Grafiken am Ende von Band II). // I.2.7

zu Ziff. 4.133:

Unschärfe Semantik oder Distribution von Ausdrücken?

zu: vsem1.552

aus: RIEGER, B.B., Unschärfe Semantik: zur numerischen Modellierung vager Bedeutungen von Wörtern als fuzzy Mengen: FRIEMEL, H-J; MÜLLER-SCHÖNBERGER, G; SCHÜTT, A (eds.), Forum '90. Wissenschaft und Technik. Neue Anwendungen mit Hilfe aktueller Computer-Technologien. Trier 8./9. Oktober 1990. Proceedings. Berlin 1991. S.80ff.

(82) Im Unterschied zur klassischen oder *scharfen* (*crisp*) Mengentheorie, in der ein Individuum *alternativ* im Hinblick auf eine Menge entweder Element ist oder nicht, kann man in der neuen Theorie die Zugehörigkeit eines Individuums zu einer deswegen *unscharf* genannten Menge *graduell* angeben. Das geschieht vermöge der charakteristischen Funktion $\mu_A(x)$, die für ein Element x der Menge A nicht nur - wie im Sinne der klassischen Mengen - die Werte 0 (für *nichtzugehörig*) oder 1 (für *zugehörig*) annehmen kann, sondern auch jeden beliebigen anderen Wert zwischen 0 und 1, wobei etwa $\mu_A(x) = 0.2$ eine geringere Zugehörigkeit des Elements x zur Menge A anzeigt, als $\mu_A(x) = 0.8$.

Allgemein wird eine unscharfe Teilmenge A von X charakterisiert durch die Zugehörigkeitsfunktion

$$(1) \quad \mu_A: X \rightarrow [0,1]$$

die jedem $x \in X$ einen (und nur einen) Zugehörigkeitswert $\mu_A(x)$ aus dem Intervall $[0,1]$ zuordnet, der den Grad angibt, mit dem das Individuum x als Element der unscharfen Menge A zu gelten hat. Die unscharfe Menge A besteht also aus der Menge der geordneten Paare

$$(2) \quad A := \{(x, \mu_A(x))\} \text{ für alle } x \in X$$

aus: RIEGER, B.B., Relevance of Meaning, Semantic Disposition, and Text Coherence. Modelling Reader Expectations from Natural Language Discourse: CONTE, M-E; PETÖFI, J S; SÖZER, E (eds.), Text and Discourse Connectedness. Amsterdam 1989. S.154-173.

the present approach has been based on the algorithmic analysis of discourse that real speakers/writers produce in actual situations of performed or intended communication on a certain subject domain. The approach makes essential use of algorithmic means to detect within a set of input discourses the differences of usage regularities of lexical items. Specified numerically, these data serve to map each lexical onto a fuzzy subset of the vocabulary employed that serves to describe an item's conceptual stereotype meaning. The set of such fuzzy subsets (Zadeh 1965) forms a system of stereotypical concept representations which constitutes a topological space structure whose abstract elements are related by mutual distances.

(161) Those items which co-occur frequently in a number of texts will positively be correlated and hence called *affined*, those of which only one (and not the other) frequently occurs in a number of texts will be negatively correlated and hence called *repugnant*. Different degrees of *word-repugnancy* and *word-affinity* may thus be ascertained without recourse to an investigator's or his test-person's word and/or world knowledge (*semantic competence*), but can instead be based solely upon the usage regularities of lexical items observed in a corpus of pragmatically homogeneous discourses, spoken or written by real authors in actual or intended acts of communication (*communicative performance*)...

Table 1. Topological environment E (*ARBEIT*)

ARBEIT 0.000

ALLGEMEIN 8.332	ANBIET 8.756	AUSGAB 10.392
STADT 10.711	PERSON 11.075	LEHR 11.811
GEBIET 11.831	VERBAND 12.041	UNTERNEHMEN 12.130
VERKEHR 12.312	HERRSCH 12.362	VERANTWORT 12.543
EINSATZ 13.980	STELLE 14.120	WERB 15.561
ORGANIS 16.146	VERWALT 16.340	MODE 16.842
GESCHAEFT 16.873	UNTERRICHT 18.275	BITT 19.614
...		

The resulting system of sets of fuzzy subsets constitutes the *semantic space*. As a distance-relational data structure of stereotypically formatted meaning representations, it may be interpreted topologically as a hyperspace with a natural metric. Its linguistically labelled elements represent *meaning points*, and their mutual distances represent *meaning differences*.

(169) Thus, reducing the three sentences to a string of lexical items (lexemes) only, and italicizing those of them (Tab. 2) whose occurrence had been predicted by the LD-algorithm, will reveal the capacity of *semantic dispositions* to provide the knowledge based relevant *expectations* a cognitive system has to derive in order to decide whether or not a piece of discourse may be called *coherent*.

Table 2. String of (*expected*) lexemes (*in italics*)

...*arbeit* - markt - herrsch - allgemein - anbiet - person - lehr - ausbild - verantwort - wirtschaft - verwalt - verb - einsetz - industri - such - beruf - grupp - kenntnis - erfahr - gebiet - computer - techn - diplom - verfueg - wunsch - faehig - leit - stell - ausueb - organis - unterricht - schul - verband - gebiet - ebene - stadt - unternehm - versuch - allgemein - inform - elektron - geb...

zu Ziff. 4.1341:

Distributive Semantik

zu: vsem1.553

aus: HERINGER, H-J, Das höchste der Gefühle. Empirische Studien zur distributiven Semantik. Tübingen 1999.

(10) Die distributive Semantik kann man als Spielart oder als Ausführung einer Gebrauchstheorie der Bedeutung ansehen. Zentraler Slogan der Gebrauchstheorie - er wird meistens Wittgenstein zugeschrieben - ist: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.

Dieser Slogan wehrt bestimmte Auffassungen ab. Er richtet sich gegen die Idee,

- es gebe eine sprachfreie Welt, die für Bedeutungen bestimmend sei,
- es gebe ein Reich der Bedeutungen in Unabhängigkeit von den Zeichen,
- es gebe Bedeutungen als identifizierbare Gegenstände,
- es gebe Bedeutungen ohne die Sprecher, die Zeichen verwenden,

Das alles sollten Sie vergessen, wenn Sie sich mit dieser Theorie befassen...

(13) Traditionelle Beschreibungen von Bedeutung stellen das zu beschreibende Wort in ein mehr oder minder kompliziertes Netz von Ausdrücken einer Beschreibungssprache, von Merkmalen usw. Die einfachste Art ist die Bedeutungsangabe im Wörterbuch: Wasser "klare, farblose, durchsichtige Flüssigkeit".

Diese Angabe ist nicht einmal eine definitorische Äquivalenz, da nicht jede klare, durchsichtige Flüssigkeit Wasser ist; sie ist sozusagen explikativ gedacht, als erster Hinweis auf die Bedeutung. Wichtiger aber: Sie kann die Variabilität und Dynamik der Bedeutung nicht erfassen. Wie wird sie damit fertig, dass unter den Belegbeispielen dann *blaues Meer* angeführt wird, das ja nach der Definition eine Kontradiktion enthielte?

Analoges gilt für die Darstellung der Bedeutung in Netzwerken, die meistens eine direkte Erweiterung der traditionellen Definition bleiben.

(15) Die Rückführung auf Sätze ist auch ein Reflex der Überzeugung: "Nur im Zusammenhang eines Satzes bedeuten die Wörter etwas." (Frege 1884:§64) oder "Nur im Zusammenhang eines Satzes hat ein Name Bedeutung" (Wittgenstein 1969:20). Die Bedeutung eines Wortes ermitteln wir als Beitrag zur Satzbedeutung.

Die Bedeutung eines Satzes ergibt sich aus der Bedeutung seiner Wörter. Und umgekehrt!

Die Bedeutung ist überhaupt nur im Gebrauch entstanden. Und sie entsteht ständig neu.

Mal anders gesagt: Sie verändert sich. Noch anders gesagt: Sie ist nicht etwas, sondern eine Feststellung des Gebrauchs, ein Kondensat der Verwendung.

(22) Nach dem Saussureschen Zeichenmodell unterscheidet man bei sprachlichen Zeichen ihren Ausdruck und ihren Inhalt. Beide sind konventionell miteinander verbunden, beide existieren aber nicht

unabhängig von einer Sprache. Sie existieren nur qua Zeichen, also in einer Sprache und durch eine Sprache.

Und beide Seiten wie das ganze Zeichen sind Schematisierungen. Der Ausdruck ist nicht das physikalische Lautereignis, der Inhalt ist nicht das je Gemeinte (Saussure 1984:98; Jäger 1983).

Die Schemata sind konventionell, sie haben ihr individuelles Pendant, insofern sie im sprachlichen Wissen einzelner Individuen sind. Das semantische Wissen eines Individuums besteht also in der Kenntnis der jeweiligen Konvention - oder besser in Hypothesen über die Konvention. Oder noch besser: Die Konventionen sind solche Hypothesen.

(29) Die Unterscheidung in paradigmatische und syntagmatische Assoziation ist ein Artefakt. Jedenfalls ist die Alternative nicht entscheidbar. Sind *fünf* → *sechs*, *fragen* → *antworten* syntagmatisch oder paradigmatisch? Und wie stünde es mit *sechs* → *fünf*?

Für den graduellen Übergang von Syntagmatik zur Paradigmatik sprechen auch die psycholinguistischen Experimente, in denen festgestellt wurde, daß Kinder anfänglich viel stärker syntagmatisch assoziieren und erst im Laufe der Zeit zu mehr Assoziationen in der gleichen Kategorie kommen (Ervin 1961, Entwistle 1966; Palermo/Jenskins 1965).

Semantisches Lernen beginnt mit Kontiguität und Kontext und bringt seine Strukturen ins episodische Wissen. Erst langsam und später wird episodisches Wissen immer mehr stehendes Wissen, fester und langsamer upgedatet.

(30) Linguistischer und psycholinguistischer main stream entwickelten jedenfalls den propositionalen Ansatz weiter. Beispiele für linguistische Weiterentwicklungen sind der gesamte Komponentialismus und die Semanalyse. Sie sind geprägt durch folgende Grundannahmen:

- Es gibt eine universale Menge semantischer Merkmale (*figurae* oder *Seme*).
- Die *Seme* sind außersprachlich oder außereinzelsprachlich, wie es heißt.
- Die *Sem*sprache hat und repräsentiert eine propositionale Struktur; sie hat eine logische Syntax.

(31) Pottiers Behauptung, das *Semem* sei eine Menge von *Semen* (Pottier 1964:122) würde allerdings zu einem ärmlichen Verständnis führen, das tatsächlich der Semanalyse nicht zugrunde liegt. Vielmehr wird eher stillschweigend vorausgesetzt, das *Semem* habe eine innere logisch-syntaktische Struktur.

58. Demgegenüber geht die distributive Semantik von folgenden Grundannahmen aus:

- Grundmenge der Semantik ist V, die Menge der Wörter oder *Lexeme* der Objektsprache.
- Die Semantik ermittelt die Struktur dieser Menge V; sie braucht kein zusätzliches Beschreibungsvokabular und ist streng sprachspezifisch.
- Die Struktur von V ist nicht rein propositional.

59. Das Semantische Wissen eines Individuums wird im Spracherwerb aufgebaut (und wahrscheinlich nie abgeschlossen). Das Individuum muß sein Wissen gewinnen aus Kommunikationen, also aus Untermengen verwendeter Zeichen und aus den entsprechenden Situationen. Das Wissen wird kaum darin bestehen, dass alle Äußerungen und

Situationen im Gedächtnis bleiben. Es findet eine Schematisierung oder Verdichtung statt.

Wie dies aussieht, wissen wir nicht. Wir sollten aber in Betracht ziehen,

- daß es keine Grenze zwischen dem sogenannten sprachlichen und dem sogenannten enzyklopädischen Wissen zieht;
- daß das semantische Wissen Wörter in Beziehung setzt, in Affinitäten bringt;
- daß (in einer lokalen Metaphorik) die Wörter im Wissen einander unterschiedlich fern stehen.

(34) Der Sündenfall des Distributionalismus, der sich hier andeutet, war die Betrachtung von Mengen, Klassen und Kategorien an Stelle der realen Einheiten. Das erklärte sich durch traditionelle Zielsetzung und Aufsprung auf den Merkmalszug, es war sicherlich auch ein Problem des Handlings. Schnelle Abstraktion mußte her, weil man eben mit den riesigen Datenmengen sonst nicht zurechtgekommen wäre:

Therefore in discourse analysis as in descriptive linguistics, we collect those elements which have like distributions into one class, and thereafter speak of the distribution of the class as a whole rather than of each element individually. (Harris 1952:6)

(38) **76.** Wenn unsere Grundüberlegungen zur Konstitution und zum Erlernen von Bedeutung stimmen, dann muß die Bedeutung in Texten stecken. Wir brauchen dann für die semantische Untersuchung nur:

- (i) gute Methoden der Kondensierung von Distributionen,
- (ii) plausible Formate der Bedeutungsdarstellung,
- (iii) Konvertierungsmethoden, die uns aus Kondensaten Bedeutungsdarstellungen erstellen lassen.

zu Ziff. 4.1342:

Distribution

zu:

aus: HERINGER, H-J, Das höchste der Gefühle. Empirische Studien zur distributiven Semantik. Tübingen 1999.

(50) **112.** Ein zu untersuchendes Wort nennen wir eine Wurzel. Sei w_0 eine Wurzel, u_i eine Äußerung, die w_0 enthält:

(23) $u_i - w_0$ heißt der Kotext von w_0 in u_i .

Die Menge aller Kotexte von w_0 heißt ihre Distribution. Die Distribution D_0 von w_0 ist also die Menge aller Umgebungen von w_0 in einem Korpus. Allgemein:

(24) $D_0 = \{u_i - w_0 \mid u_i - w_0 \subset C_i\}$

Die Länge eines u_i kann bestimmt werden durch grammatische Kriterien oder rein formal mit einem Radius um w_0 , der die Anzahl symmetrisch einbezogener Wörter w_i angibt.

113. Die syntagmatischen Relationen zwischen w_i sind zweistellig. Sie können allgemein als direkte oder indirekte Nachbarschaftsrelationen verstanden werden. Da ein Korpus linear ist, kann man es abbilden auf die geordnete Menge der natürlichen Zahlen N . Es bedeuten dann:

(25) $w_i < w_j = i < j$

w_i steht links von w_j , wenn es eine kleinere Indexzahl hat.

(26) $w_i \ll w_j = (i+1=j)$

w_i steht direkt links von w_j , wenn sein Index um 1 kleiner ist.

Entsprechend für die rechte Nachbarschaft:

(27) $w_i > w_j = i > j$

(28) $w_i \gg w_j = (i-1=j)$

114. Für die Untersuchung eines Wortes w_0 im Korpus C_i ermitteln wir die Umgebung aller Vorkommen von w_0 . Da die Umgebung eines w_i im Prinzip jeweils das gesamte Korpus ist, müssen wir sinnvolle Ausschnitte ausschneiden. Diese Ausschnitte nennen wir Fenster, sie entsprechen in etwa einem Eintrag in einem KWIC-Index, in dem Zeilen mit dem Suchwort in der Mitte aus dem Korpus herausgeschnitten werden.

Das Fenster F_0 von w_0 ist definiert durch:

(29) $F_0 = w_1 w_2 \dots w_r w_0 \dots w_{2r}$

(51) Ein Fenster ist definiert durch seinen Durchmesser, das ist die Anzahl der Vorgänger und Nachfolger von w_0 . In der Regel wird man das Fenster symmetrisch wählen. F_r ist entsprechend w_0 mit seiner Umgebung im Radius r , also eine Kette der Länge $2r+1$, die besteht aus w_0 zusammen mit r Wörtern vor w_0 und r Wörtern danach.

Wie groß die angemessenen Fenster für bestimmte Fragestellungen sein sollten, ist eine empirische Frage.

115. Ein Sonderfall des Fensters ist der sogenannte Beleg. Ein Beleg für w_0 ist üblicherweise eine Kette von w_i , die nach grammatischen Gesichtspunkten begrenzt ist und w_0 mit umfaßt, also

$B_1 = w_1 w_2 \dots w_0 \dots w_m$

Der Beleg ist in der Regel darum auch nicht symmetrisch.

116. Ein Belegkorpus für W_0 ist eine Untermenge des Untersuchungskorpus aus n Fenstern und somit n Vorkommen von W_0 :

$B_0^n C_i$

Das Belegkorpus ist eine Menge von Fenstern, die aus dem Untersuchungskorpus herausgeschnitten sind.

(32) $B_0^n = \{F_0^x \mid F_0^x C_i\}$

Traditionell würde man allerdings Belege ausschneiden, so dass B_0^n aus n Belegen für W_0 bestünde.

Wie groß muß n sein, damit man Aussagen über die Bedeutung von W_0 gewinnen kann? Ab welcher Größe von n können diese Aussagen als repräsentativ angesehen werden? Wo liegt der Übergang von Eigenschaften des Korpus zu Eigenschaften des Deutschen?

zu Ziff. 4.1343:

Nachbarschaft = Semantik ?

zu:

aus: HERINGER, H-J, Das höchste der Gefühle. Empirische Studien zur distributiven Semantik. Tübingen 1999.

(56) **128.** Aufgabe einer semantischen Untersuchung des Wortes W_0 könnte sein, festzustellen, mit welchen Nachbarn Äußerungen von W_0 üblicherweise vorkommen und mit welchen Ereignissen aus E die Vorkommen verbunden sind. Jede semantische Untersuchung endet in einer Beschreibung oder Darstellung. In dieser Beschreibung gehen auch die Ereignisse aus E nur als sprachliche Äußerungen ein. Darum besteht jede semantische Beschreibung letztlich nur aus einem Text, aus einer Kette von Wörtern. Methodisch interessant ist es, diese Erkenntnis ernst zu nehmen. Ist also eine Semantik möglich, die selbst keine Beschreibungstexte produziert, keine Formulierungen für die Ereignisse aus E ? Eine entsprechende These ist:

- (i) Die Bedeutung eines Wortes W_0 kann charakterisiert werden in der Struktur eines Korpus.
- (ii) Das Korpus muss keine beschreibenden Äußerungen des Linguisten enthalten.

Dies ist die Grundthese der Distributiven Semantik. Sie zu widerlegen ist keine triviale Aufgabe. Denn immerhin steht auch stets zur Disposition, was denn die Bedeutung eines Wortes sei. Einen fixen Eichpunkt haben wir da nicht.

129. Eine speziellere These ist:

- (i) Die Bedeutung von W_0 findet ihren Niederschlag in der technisch eruierten Distribution von W_0 .

Entscheidende Fragen sind nun:

- (ii) Wie findet die Bedeutung von W_0 in D_0 ihren Niederschlag?
- (iii) In welchen Aspekten der eruierten Distribution findet die Bedeutung von W_0 ihren Niederschlag?

Auf jeden Fall gewinnt die Grundthese eine gewisse Plausibilität dadurch, daß auch ein Lerner in der Ontogenese keine Bedeutungserklärungen bekommt, sondern seine Kenntnisse aus den Äußerungen entnehmen muß. ...

(57) **129.** Ein noch immer naheliegender Einwand gegen die Distributive Semantik bestünde darin, die Elemente aus D^* zu reinen, äußerlichen Zeichen zu deklarieren, sozusagen zu reinen Lautfolgen. Ein solcher Reflex der Ideologie "Hier das Zeichen, hier die Bedeutung" wäre kurzsichtig. Die Elemente aus D^* sind Okkurrenzen vollständiger Zeichen, mit Ausdruck und Inhalt. Ein Zeichen ist nicht ohne Bedeutung; weder Ausdruck noch Inhalt sind eigenständig vorgegeben. Die Identität des Zeichens, die Abgrenzung und Kategorisierung von Ausdruck und Inhalt entsteht in einem Gesamtzusammenhang.

Wenn die Bedeutungen irgendwo sind, dann sind sie auch im Korpus!

zu Ziff. 4.1344:

Wortdistanz / Bedeutung

zu:

aus: HERINGER, H-J, Das höchste der Gefühle. Empirische Studien zur distributiven Semantik. Tübingen 1999.

(67) **152.** Fassen wir zusammen: Ein Kondensat für W_0 gewinnt man aus der Distribution dadurch, daß man errechnet,

- mit welcher relativen Häufigkeit bestimmte Nachbarn W_i mit W_0 im Belegkorpus vorkommen;

- in welchen Distanzen die jeweiligen Nachbarn von den W_0 vorkommen. Grundlage dafür ist das gesäuberte und abgemachte Belegkorpus.

130. Im Kondensat lagern sich bestimmte W_i um die untersuchte Wurzel W_0 . Diese W_i sind dem W_0 affin. Affinität drückt sich aus in der jeweiligen Distanz zu W_0 . Die Affinitäten können auch einzeln wiedergegeben werden:

$$W_0 \rightarrow W_i = a$$

besagt, daß die Affinität von W_0 und W_i den Wert a hat.

Je kleiner a , umso relevanter ist ein W_i für die Bedeutung von W_0 , so die Grundthese.

Das Ensemble der Affinitäten wird dargestellt durch die Kanten in einem radialen Graphen.

Die Lage der W_i im radialen Graphen ist ein Artefakt; sie besagt nichts.

154. Ein Kondensat können wir uns vorstellen als Blick in das lexikalische Gesamtnetzwerk von einem bestimmten Punkt W_0 aus. Ein gutes Bild ist: Man greift W_0 in dem Gesamtnetzwerk und lupft das Netz an dieser Stelle. Die verknüpften W_i werden sich mitheben je nachdem, wie weit sie von W_0 entfernt sind und wie hoch man das Netz hebt. ...

(70)

157. Weitere Fragen im Zusammenhang mit der textuellen Basis sind:

- Wieviel Belege braucht man für verlässliche distributive Aussagen?

- Wie groß sollte der Radius der Fenster sein?

- Wie sollte die Affinität berechnet werden?

Solche Fragen, die auf den ersten Blick eher technischer Natur scheinen, erweisen sich als tiefer reichende sprachtheoretische Probleme. Wir wissen nicht, wie weit die semantische Domäne eines Lexems im Kontext reicht oder wie sie langsam fadet. Jedenfalls gibt es nach textlinguistischer Lehre kein Argument, die Domäne mit Satzgrenzen abzuschließen. Wir wissen ebensowenig, ob für die Bedeutung eines Lexems mehr die Frequenz des kotextuellen Miteinanders zählt oder die kotextuelle Nähe, und erst recht nicht, wie beides zu verrechnen wäre. Wie so oft wird eine Antwort auch vom Zweck der Untersuchung abhängen.

zu Ziff. 4.1345:

Syntax ↔ Semantik

zu:

aus: HERINGER, H-J, Das höchste der Gefühle. Empirische Studien zur distributiven Semantik. Tübingen 1999.

(98)

197. Halten wir fest: Affinitäten sind keine Prädikationen. Wenngleich das Argument-Prädikator-Verhältnis vielleicht letztendlich zu Grunde liegt, ist die Prädikation nicht recoverable.

Direkte Prädikationen liegen aus mehreren Gründen nicht vor:

(i) Die Richtung ist unbekannt. Man weiß nicht, ob W_0 Prädikator zu W_i ist oder umgekehrt.

(ii) Die Prädikation könnte indirekt sein, möglicherweise über mehrere Kettenglieder, so daß ein W_j eine Prädikation zu W_i wäre, das seinerseits W_0 affin ist.

(iii) Die Affinität ist neutral gegenüber der Negation. Darum ist zu W_0 sowohl W_i als auch dessen Antonym W_i affin.

198. Affinitäten sind nur thematische Zusammenhänge. Sie charakterisieren semantische Dimensionen, in denen die Prädikation sinnvoll ist.

Der thematische Zusammenhang von W_0 und W_i hat also doch mit dem P-A-Verhältnis zu tun, und er ist abhängig von der syntaktischen und semantischen Kategorie der W_i

(206) **206.** Die distributive Darstellung enthält tiefere semantische Aussagen nur versteckt und für den Normalverbraucher unsichtbar.

zu Ziff. 4.135:

Stilstatistik - ohne Theorie?

zu:

aus: F. M. DIMPEL, Computergestützte textstatistische Untersuchungen - an mittelhochdeutschen Texten. Tübingen 2004.

(17) ... folgende Vorannahmen zu benennen, die freilich kaum bewiesen werden können:

A) Bei der quantifizierenden Textanalyse wird ein Teilbereich dessen untersucht, was - wie auch immer - als "Stil" bezeichnet werden kann. "Stil" kann auf unterschiedlichem Niveau stabil sein: innerhalb eines einzelnen Werkes, innerhalb von größeren Abschnitten eines Werkes.

B) Es gibt Textmerkmale, bei denen Differenzen zwischen Werken unterschiedlicher Autoren, unterschiedlichen Werken eines Autors und unterschiedlichen Textabschnitten eines Werkes gemessen werden können.

C) Es ist plausibel, daß die meßbaren Unterschiede zwischen unterschiedlichen Autoren größer sind als die zwischen den Werken eines Autors, sowie daß die Unterschiede zwischen den Werken eines Autors größer sind als die zwischen den Abschnitten eines Werkes.

Doch bleibt als Fazit: Eine allgemeine Theorie der quantifizierenden Stilanalyse gibt es nicht.

zu Ziff. 4.15:

SHANNON: Entropy in Kommunikationstheorie / "Information"

zu: vprag9.5611 vgt3.430

aus: V. K. BALASUBRAHMANYAN; S. NARANAN, Entropy, information and Complexity in: KÖHLER, R u.a. (eds.), Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005

(879) Shannon is to be credited with a clear formulation of a mathematical theory of communication where the concept of entropy was related closely with a quantitative definition of the amount of information conveyed by a message in a communication process. Shannon's information theory is schematically shown in Figure 61.1. All the essential elements of a communication system can be seen in the figure.

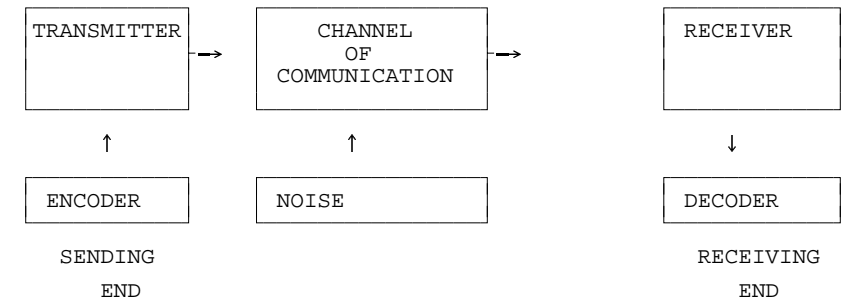


Fig. 61.1: Shannon's model of a communication system for messages

Shannon's theory is a mathematical theory and gives theoretical limits for the performance requirements needed and achievable by the different units constituting a communication system. ...

Shannon recognized the role of information as reducing the uncertainty in a situation of receiving the message. Shannon postulated the amount of information as being equal to the uncertainty reduced by receiving the message. The message for this purpose is considered to be a statistical collection of symbols transmitted by the system. The other aspect of a message, for example the meaning conveyed by it or its significance was not considered and was out of the purview of the theory. In this sense the theory is an abstract theory with its own axioms, rules etc. Its correspondence to the real world of efficient codes for transmission of messages (Huffman, 1952, 1098) and error correcting codes by many workers and in motivating others to use these ideas to study problems in fields other than communication engineering.

3.1 Shannon entropy

Kapur (1980) has summarized all the characteristic essentials of Shannon entropy, which can be considered as a very useful measure of the quantity of information for a message. It equates the quantity of information in a message to the uncertainty decreased by the receipt of the message. A message is treated as a random collection of symbols, which can be transmitted and received by the communication system. Treating the collection of symbols as purely governed by the rules of probability, the aim of the

communication process is to transmit messages efficiently and without error so that the receiver of the message can get the message faithfully and reconstruct what was sent by the sender.

4.2 Was Grammatiken unter "Morphologie" verstehen

Literatur

- ALEGRIA, I, M (et. al.); URKIA: Automatic Morphological Analysis of Basque. *Literary and Linguistic Computing* 11 (1996 / 4) 193-203.
- ARONOFF, M: Morphology by Itself. Stems and Inflectional Classes. *Linguistic Inquiry Monograph* 22. Cambridge / Mass. 1994: The MIT Press // I.2.7
- BAUER, L: *Introducing Linguistic Morphology*. Edinburgh 1988: Edinburgh University Press // I.2.7
- BECKER, T: Morphologische Ersetzungsbildungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 12,2 (1993) 185-217.
- BELLE, W van; LANGENDONCK, W van (ed./eds.): *The Dative. Volume 1: Descriptive Studies. Case and grammatical Relations across languages 2*. Amsterdam 1995: John Benjamins // I.2.7
- CHENGYU FANG, A; NELSON, G: Tagging the Survey Corpus: a LOB to ICE. Experiment using AUTASYS. *Literary & Linguistic Computing* 9,3 (1994) 189-194.
- CORBETT, G G: *Gender. Cambridge Textbooks in Linguistics*. Cambridge 1991: Cambridge University Press // I.2.4
- GETTERT, H: *Syntax zwischen Hierarchie und Linearität* 1995: UB Tübingen (Sign: 10 G 7994; 5 Mikrofiches)
- GLINZ, H: Grammatiken im Vergleich. *Deutsch-Französisch-Englisch-Latein. Formen-Bedeutungen-Verstehen. Germanistische Linguistik Band 136*. Tübingen 1994: Niemeyer // I.2.4
- HAUSSER, R (ed./eds.): *Linguistische Verifikation. Dokumentation zur Ersten Morpholympics*. Sprache und Information 34. Tübingen 1996: Niemeyer // I.2.7
- HOFFMANN, L (ed./eds.): *Deutsche Syntax. Ansichten und Aussichten*. Berlin 1992: Walter de Gruyter // I.2.4
- JANSSEN, A: Segmentierung französischer Wortformen in Morpheme ohne Verwendung eines Lexikons in: KLENK, U (ed./eds.): *Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache*. ZDL BH.73. Stuttgart 1992 74-95. // I.2.7
- JÄPPINEN, H: Finite State Computational Morphology in: KLENK, U (ed./eds.): *Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache*. ZDL BH.73. Stuttgart 1992 96-109. // I.2.7
- KÖPCKE, K-M: Die Klassifikation der schwachen Maskulina in der deutschen Gegenwartssprache. Ein Beispiel für die Leistungsfähigkeit der Prototypentheorie. *Zeitschrift für Sprachwissenschaft* 14.2 (1995) 159-180.
- KOSKENNIEMI, K: *Two-Level Morphology: A General Computational Model for Word-form Recognition and Production*. Publication no.11. Helsinki 1983: University of Helsinki

- LUTZEIER, PR: Lexikologie. Ein Arbeitsbuch. Tübingen 1995: Stauffenburg // I.2.5
- MILLER, J: Semantics and syntax. Parallels and connections. Cambridge 1985: Cambridge University Press // EB
- PALMER, F R: Grammatical Roles and Relations. Cambridge Textbooks in Linguistics. Cambridge 1994: Cambridge University Press // I.2.4
- RICKHEIT, M: Wortbildung. Grundlagen einer kognitiven Wortsemantik. Psycholinguistische Studien. Opladen 1993 // I.2.4
- RIEGER, B: Unscharfe Semantik. Die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung vager Wortbedeutungen in Texten. Frankfurt / M 1989 // I.2.4
- ROTHWEILER, M (ed./eds.): Spracherwerb und Grammatik. Linguistische Untersuchungen zum Erwerb von Syntax und Morphologie. Linguistische Berichte, Sonderheft 3 / 90 1991
- SCHMITZ, U: Computerlinguistik. Eine Einführung 1992: Westdeutscher Verlag
- SCHULTE-SASSE, J; WERNER, R: Einführung in die Literaturwissenschaft. UTB 640. München 1990: Wilhelm Fink // I.2.4
- SGARBAS, K; FAKOTAKIS, N: A PC-KIMMO-Based Morphological Description of Modern Greek. Literary and Linguistic Computing 10 / 3 (1995) 189-201.
- SPECHT, G: Wissensbasierte Analyse althebräischer Morphosyntax. Das Expertensystem AMOS. ATS 35. St. Ottilien 1990 // I.2.5
- SPENCER, A; ZWICKY, A M (ed./eds.): The Handbook of Morphology. Oxford 1998: Blackwell // I.2.7
- STEIN, S: Affixoide. Überlegungen aus semantischer und pragmatischer Perspektive zu einer umstrittenen Einheit der Wortbildung in: POHL, I (ed./eds.): Semantik und Pragmatik-Schnittstellen. Sprache, System und Tätigkeit 59. Frankfurt/M 2008 179-216. 978-3-631-57061-6 br: P. Lang // I.2.4
- TROST, H: Morphology in: MITKOV, R (ed./eds.): The Oxford Handbook of Computational Linguistics. Oxford Handbooks Series. Oxford 2003 25-47.: Oxford University Press // I.2.4
- VITAS, D; KRSTEV, C: Local grammars in word counting in: Grzybek, P; Köhler, R (ed./eds.): Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 665-678.: Mouton de Gruyter // I.2.5
- WUNDERLICH, D; FABRI, R: Minimalist Morphology: An Approach to Inflection. Zeitschrift für Sprachwissenschaft 14.2 (1995) 236-294.
- XIAOLING, H; NIGEL, W: Syntactic Positions of Prepositional Phrases in the History of Chinese. Using the Developing Sheffield Corpus of Chinese for Diachronic Linguistic Studies. Literary and Linguistic Computing 22/4 (2007) 419-434.

4.21 Inkonsistenzen**Literatur**

- DROR, J; DUDU,, Sh; TALMON, R; WINTNER, Sh: Morphological Analysis of the Qur'an. Literary and Linguistic Computing 19 / 4 (2004) 431-452.

zu Ziff. 4.211:

Form ↔ Inhalt

zu:

aus: POURADIER DUTEIL, F, Studien zur Prädikation. Eine Analyse der Verbalperiphrasen der Phase im Französischen. Frankfurt 1999.

(28) Es geht mir im folgenden nicht um eine Theorie der grammatischen Kategorien, sondern um eine zwangsweise unvollständige Auflistung der syntaktisch-semantischen Funktionen im französischen Verbalbereich. Um bei solchen Funktionen von grammatischen Kategorien sprechen zu können, muß laut Guchmann die Einheit sowohl des Inhalts (semantische Funktion) als auch des Ausdrucks (morphosyntaktische Ausdrucksweise) vorhanden sein:

Die Zweiseitigkeit (Inhalt und Form) dieser Elemente [die Grundeinheiten des grammatischen Systems und ihre distinktiven Merkmale (Verf.)], die sich in der Einheit der Funktion (grammatischer Inhalt) und der diesen Inhalt signalisierenden sprachlichen Mittel ausdrückt, ist die natureigene Qualität der grammatischen Elemente, die folglich stets zweidimensional sind. Deshalb ist jedes grammatische Element durch die distinktiven Merkmale auf zweierlei Art charakterisiert: 1. durch die funktionalen oder Bedeutungsmerkmale und 2. durch die strukturell-formalen Merkmale (GUCHMANN 1966:265)

(29) Die Bedingung der Einheitlichkeit wird nicht immer zufriedenstellend erfüllt. Die Entsprechung von Ausdruck zu Inhalt ist nicht immer eine Entsprechung eins zu eins. Einer Klasse von Bedeutungsmerkmalen können mehrere Klassen von Ausdrucksmitteln entsprechen, wie auch umgekehrt...

Unter Berücksichtigung der Tatsache, daß bestimmte Merkmale der lexikalischen Bedeutung unter Zuhilfenahme von analytischen Mitteln der grammatischen Bedeutung präzisiert bzw. modifiziert werden können, werde ich im folgenden drei Gruppen von Kategorien bzw. semantischen Funktionsklassen der grammatischen wie auch der lexikalischen Bedeutung berücksichtigen, ohne den Anspruch der Vollständigkeit erheben zu wollen:

1. Synthetische (flexionelle) Kategorien der grammatischen Bedeutung:
Numerus / Person
Tempus (absolute Zeit)
Modus (Indikativ, Konjunktiv, etc.)
Aspekt (Perspektiv vs Imperfektiv, nur beschränkt vorhanden).
2. (31) Analytische (nicht-flexionelle) Kategorien / Funktionsklassen der grammatischen Bedeutung:
Tempus (relative Zeit)
Diathese (aktive, passive, rückbezügliche, reziproke, kausative, rezessive)
Modalität (als Ergänzung zu flexionellem Modus)
Aktionart.
3. Modifizierbare Kategorien der lexikalischen Bedeutung:
Zeitcharakter
prädikative Kategorie (Zustand, Prozeß, Handlung)
Valenz bzw. Rollenrahmen.

zu Ziff. 4.212:

Unausrottbare Bedeutungsaspekte

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(42) DIE FORMSEITE DES WORTES

In diesem Kapitel wird versucht zu klären, was wir uns unter der Formseite eines Wortes vorstellen können. Vorgestellt wird die Paradigma-Konzeption.

Im Medium der gesprochenen oder geschriebenen Sprache begegnen uns keine Wörter sondern Wortformen. Dies zeigt sich am deutlichsten bei den sog. flektierbaren Wörtern:

der Verkäufer redet auf die unschlüssige Kundin ein
die Verkäufer reden auf die unschlüssigen Kundinnen ein.

Die jeweils direkt übereinander stehenden Wortformen können wir als Realisationen ein und desselben Wortes ansehen. Die in der ersten und fünften Spalte auftretenden Wortformen sind damit Formen des bestimmten Artikels. Solche Namen für die abstrakten Wörter parat zu haben - ein anderer wäre "Personalpronomen", ist allerdings eher die Ausnahme. Bereits bei der zweiten Spalte können wir außer der Angabe, daß es sich um ein Substantiv mit der von uns formal vorgenommenen Bezeichnung W handelt, nichts Konkretes sagen. In dieser Situation liegt nun folgendes Vorgehen nahe: Die Zusammenfassung aller derjenigen Wortformen, die man intuitiv als Wortformen desselben Wortes auffassen würde, soll als möglicher Repräsentant für das ansonsten nicht ansprechbare Wort gelten. Der Fall der zweiten Spalte ergäbe damit eine Menge aus drei Elementen: (*verkäufer, verkäufers, verkäufern*). Dieses Vorgehen kann allerdings noch nicht befriedigen. Bei einer Menge kann jedes Element nur einmal auftreten, insofern hätten wir z.B. mit der vorgeschlagenen Mengendarstellung nicht die Möglichkeit, die bereits in unseren beiden Beispielsätzen auftretenden Unterschiede für eine Form wie *verkäufer* zu erfassen. Im ersten Beispielsatz ist *verkäufer* im Singular markiert, im zweiten dagegen im Plural. Erfasst werden muß also neben der reinen (!) Form auch die Menge der morpho-syntaktischen Eigenschaften bezüglich aller relevanten Kategorien, wie z.B. "Person", "Kasus", "Genus", "Numerus", "Modus", "Genus verbum", "Tempus", "Aspekt", "Komparation", wobei deren Relevanz natürlich von Wortart zu Wortart variiert.

Unser ominöses Substantiv W, genauer, dessen Formseite, läßt sich somit folgendermaßen repräsentieren:

```
{<verkäufer, {Nom., Sing., Mask.}>, <verkäufer, {Dat., Sing., Mask.}>, <verkäufers, {Gen., Sing., Mask.}>, ..., <verkäufern, {Dat., Plur., Mask.}>, ..., <verkäufer, {Akk., Plur., Mask.}>}
```

(43) Wie man unschwer erkennt, ist, "Genus" innerhalb dieser Menge keine relevante Dimension. Substantive im Deutschen sind im Hinblick auf ihre Wortformen genusfest. Aus diesem Grund ist *verkäuferin* eine Form eines anderen Wortes!

Man nennt eine solche Menge von Paaren wie in unserem Beispiel auch ein **Paradigma**, genauer ein **Einzelparadigma**. Gr. *paradeigma* "Beispiel" bezieht sich dagegen auf ein Paradigma als Repräsentant eines Flexionstyps; eine Verwendung, die wir aus jeder traditionellen Grammatik her gewohnt sind und bei der wir besser von einem **Gruppenparadigma** reden sollten.

Wir können nun feststellen:

Die Formseite eines Wortes besteht aus einer Menge von geordneten Paaren, wobei jedes Paar an der ersten Stelle eine Wortform des fraglichen Wortes und an der zweiten Stelle die Menge von morpho-syntaktischen Eigenschaften dieser Wortform enthält.

Das für unsere Sprechweise eingeführte Verständnis der Formseite eines Wortes hat zwar etwas Abstraktes ergeben, insofern eben Mengen abstrakte Gegenstände darstellen, aber immerhin werden diese Mengen aus konkreten Realisationen, den Wortformen, konstruiert.

zu Ziff. 4.213:

Morphologie: Unsicherheit in Quantitativer Linguistik

zu:

aus: C. PRÜN, P. STEINER, Quantitative Morphologie: Eigenschaften der morphologischen Einheiten und Systeme in: KÖHLER, R u.a. (eds.), Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005

(227) Die Morphologie als Lehre von der Struktur der Wörter (vgl. Bauer 2003, 3ff.; Flohr/Pfingsten 2002; Matthews 1991, 9) befasst sich mit der Form und der Bildung von Wörtern. Sie untersucht sprachliche Einheiten, die unter den vagen Begriff "Wort" fallen (*Wortformen, Wortformtokens, Wortformtypes, Lexeme, Lemmata*) und kategorisiert sie (*Inhaltswörter vs. Funktionswörter, offene vs. geschlossene Wortklassen, Wortarten*). Für die Untersuchung der inneren Struktur werden die Bestandteile von Wortformen, die *Morpheme*, beschrieben und klassifiziert, und es wird erfasst, wie diese syntagmatisch kombiniert werden können. Prinzipiell wird bei der Kombination von Morphemen unterschieden zwischen den morphologischen Prozessen der *Flexion* und der *Wortbildung*, die wiederum u.a. *Derivation* und *Komposition* umfasst (Aronoff 2000).

Die Quantitative Morphologie befasst sich jedoch nicht nur mit den morphologischen Einheiten, ihren Eigenschaften und den morphologischen Prozessen, sondern auch mit Erklärungen - zum Beispiel, warum bestimmte Häufigkeitsverteilungen auftreten - und mit Zusammenhängen zwischen verschiedenen Spracheigenschaften.

Da die Quantitative Morphologie sich exakter Methoden bedient, müssen ihre Einheiten hinreichend klar operationalisiert sein. Die oben erwähnten Begriffe der morphologischen Einheiten sind aber teilweise auch innerhalb der Qualitativen Linguistik nicht eindeutig definiert. Insbesondere der Morphembegriff und die Einteilung der Morpheme in Unterklassen werden unterschiedlich gefasst. Daher wird im folgenden Abschnitt näher auf Probleme und Fragen der Definition von morphologischen Einheiten eingegangen.

1.1. Morphologische Einheiten

Morpheme werden von *Morphen* repräsentiert, die die gleiche Bedeutung besitzen. Das für die deutsche Sprache bekannteste Beispiel sind die Morphe *-e, -en, -er, -s*, die *Allomorphe* des Pluralmorphems sind. Linke/Nussbaumer/Portmann (2001, 69f) postulieren allerdings den Verzicht auf die signifié-Seite des Morphems, so dass es sich hier um synonyme Morpheme handeln würde.

Es ist problematisch, Morphe(me) zu identifizieren bzw. sie im Text zu segmentieren, denn hält man sich an die Form, so gerät man an Erscheinungen wie *Sandhi, liaison, Assimilationen, Portmanteu-Morphe(me)* u.ä. Hält man sich an die Bedeutung, dann verliert man die Möglichkeit, Länge zu messen, die im synergetischen Regelkreis (vgl. Köhler 1986) eine wichtige Rolle spielt. ...

Morpheme bzw. Morphe werden danach unterscheiden, ob sie als (potentiell) freie Morphe(me), wie *Haus* oder *die*, auftreten können oder ob es sich um (obligatorisch) (228) gebundene Morphe(me) handelt (Bauer 2003, 13). Gebundene Morpheme können Morphe mit lexikalischer Bedeutungsfunktion (wie *Schorn* in *Schornstein* oder *Kirch* in *Kirchplatz*) oder Morphe mit grammatischer Bedeutung, also *Affixe*, sein.

Die Unterscheidung zwischen Affixen und freien Morphemen ist indes nicht immer einfach, und damit wird auch die Definition von

Komposition und Derivation problematisch, denn für Komposita als Ergebnisse des Prozesses der Komposition gilt im Allgemeinen als *differentia specifica*, dass sie "aus mindestens zwei frei vorkommenden Morphemen oder Morphemkonstruktionen zusammengesetzt" sind (Bußmann 2002, 362, 402).

Hier ergibt sich allerdings das Problem, dass sich die Bedeutung einer Form in einem Kompositum von der Bedeutung der freien Form unterscheidet. Zum Beispiel betrachten Altmann/Best/Kind (1987, 132) Morphe(me) wie *-freundlich*, *-fest* und *-frei* als Suffixe, während Höhle (1985; 328ff.) sie in Verbindung mit freien Morphemen bringt. Der Übergang zwischen freien und gebundenen Morphe(me)n ist von gradueller Natur bzw. die Frage, wann die Bedeutung einer Form so stark modifiziert ist, dass von einer neuen Bedeutung gesprochen werden kann, kann nicht geklärt werden.

zu Ziff. 4.214:

Verfrühtes Bedeutungswissen: Zusammensetzung + Ableitung

zu:

aus: KIM, Gyung-Uk, Ist *trinkfreudig* eine Zusammensetzung oder eine Ableitung? Ergebnisse aus Recherchen in gegenwartssprachlichen Korpora der deutschen Standardsprache in: Schwitalla, J; Wegstein, W, Korpuslinguistik deutsch: synchron - diachron - kontrastiv. Würzburger Kolloquium 2003. Tübingen 2005. S. 155-162.

(155) Man könnte darüber streiten, ob *trinkfreudig* zum Wortbildungstyp einer Zusammensetzung oder einer Ableitung gehört. *Ein trinkfreudiger Mann* kann nicht mehr durch "ein Mann, der beim Trinken freudig ist" paraphrasiert werden, sondern durch "ein Mann, der gern und häufig trinkt". *Ein sangesfroher Mensch* bedeutet nicht einen "Menschen, der beim Singen froh ist", sondern "einen, der gerne singt". Zur Erzeugung neuer Benennungseinheiten hat eine Umsemantisierung des Morphems "freudig" oder "froh" stattgefunden, sodass neue Inhaltsmuster und damit neue Wortbildungsmodelle (Komposition → Derivation) entstehen. Man könnte sogar meinen, dass solche Prozesse ständig stattfinden, weil einerseits immer neue Benennungen benötigt werden und andererseits eine Wortbildungskonstruktion, sowie sie von der Sprachgemeinschaft angenommen ist, sowohl zu Demotivierung als auch zu Ummotivierung tendiert.

Dagegen kann *vitaminarme Kost* durch "Kost, die arm am Vitaminen ist" paraphrasiert werden. Wenn eine WBK (Wortbildungskonstruktion) durch eine Paraphrase, die beide Konstituenten in sich enthält, paraphrasiert werden kann, gehört diese Konstruktion noch zu einer Zusammensetzung, die aus Bestimmungswort und Grundwort besteht. Wenn aber eine Wortbildungskonstruktion durch eine Paraphrase, die beide Konstituenten in sich enthält, nicht mehr oder nur unter großen Schwierigkeiten paraphrasiert werden kann, gehört diese Konstruktion zu einer Ableitung. Das bedeutet, die ursprüngliche zweite Konstituente einer Zusammensetzung hat die eigentliche Bedeutung verloren und ist zu einem Suffix geworden. Wenn man einmal *Mordshitze* vergleicht, kann man sofort feststellen, dass *Mordlust* durch "Lust, zu morden" und *Mordshitze* nicht mehr durch "Hitze des Mordes", sondern durch "eine sehr große Hitze" paraphrasiert wird. *Mordlust* ist noch eine Zusammensetzung, während *Mordshitze* schon zu einer Präfixbildung geworden ist (Vgl. Kim 1996).

zu Ziff. 4.215:

Morphologie in Computerlinguistik

zu:

aus: M. KLENNER, Robuste Auszeichnung Grammatischer Funktionen: Zeitschrift für Sprachwissenschaft 26 (2007) 187-211.

(188) Die Aufgabe besteht in der Auszeichnung von einfachen Phrasen (d.h. Phrasen, die keine weiteren Phrasen enthalten - sogenannte Chunks) mit grammatischen Funktionen (GF) (von nun an *GF-Etikettierung*). Es wird dabei gefordert, dass jeder dieser Chunks inklusive der Verben (außer dem Wurzelverb) eine grammatische Funktion zugewiesen bekommt. Auf diese Weise entsteht eine Baumstruktur, die sich von einem regulären Dependenzbaum lediglich darin unterscheidet, dass sie nicht auf die Ebene der lexikalischen Köpfe hinunterreicht, sondern ganze Phrasen relationiert.

Die Auswahl der grammatischen Funktionen ist dabei von den Verben im Satz abhängig. Betrachten wir ein Beispiel (vgl. Bsp. 1).

- (1) a. [Jeder]_S glaubt [Sorbas]_I
 b. [Jeder]_S glaubt, [dass [Sorbas]_S kommt]_C
 c. [Sorbas]_S glaubt [an ein gutes Ende]_P
 d. [Sorbas]_S glaubt [alles]_D

Das Verb "glauben" hat mindestens vier Subkategorisierungsrahmen. Da es uns nicht auf den Phrasentyp (NP, PP usw.), sondern die grammatische Funktion ankommt, spezifizieren wir den Rahmen (wie die LFG) mittels grammatischer Funktionen. Wir verwenden im Artikel folgende Abkürzungen: *S* = Subjekt, *D* = direktes Objekt, *I* = indirektes Objekt, *P* = pra(!)positionales Objekt (Komplement, kein Adjunkt), *C* = Komplementsatz, *J* = Adjunkt, *T* = attributiver Modifikator (z.B. attributiv gebrauchte PP).

(189) Die Etikettierungsaufgabe kann ohne ein Lexikon mit verbspezifischen Subkategorisierungsrahmen nicht zufriedenstellend gelöst werden. Mit Hilfe dieses Lexikons können Etikettierungen ausgeschlossen werden, die vom jeweiligen Verb nicht subkategorisiert werden (z.B. "Sorbas" als nicht erlaubtes Genitivobjekt von "glauben"). ...

(196) Die Merkmale ersetzen eine vollständige Syntaxanalyse; die Behauptung ist, dass sie zur Bestimmung der grammatischen Funktionen ausreichen. Wie bereits erwähnt sind die morphologischen Analysen oft mehrdeutig, z.B. bei Eigennamen, wo bis auf den Genitiv alle Kasus möglich sind. Die Dimensionen der Vektoren sind so aufgebaut, dass alle Gertwol-Analysen repräsentiert werden können. Generell

- das syntaktische Label der Chunks
- die Wortklasse der Köpfe der Chunks
- das Lemma der Köpfe der Chunks
- Person, Kasus, Numerus und Genus bei Nomen
- Kasus bei Präpositionen
- Numerus und Person bei Verben
- die Distanz der Kandidaten in Chunks
- der syntaktische Typ der intervenierenden Chunks
- die Anzahl der intervenierenden Verben zwischen den Kandidaten
- die Anzahl von intervenierenden Satzzeichen zwischen den Kandidaten
- die Richtung der Dependenzrelation

- ein Aktiv-Passiv-Flag

(197) erhält der Maximum-Entropie-Klassifikator im Fall von ambigen Daten alle Alternativen.

4.22 Suche nach Grundlagenzu Ziff. 4.221:**Thesen zur Morphologie**

zu: vsyst4.101 vgt5.211

Vgl. im *Internet*:<http://www-ct.informatik.uni-tuebingen.de/ct/morph.html>

Thesen und Anmerkungen sind als Hypertext organisiert.

Es wird sowohl die gängige Morphologieauffassung dargestellt und kritisiert, wie auch in Grundzügen ein neuer ausdrückssyntaktischer Zugang vorgestellt.

Kann als Spezialthema genommen werden.

zu Ziff. 4.222:

Morphologie: Kritik an traditioneller Position

zu: vsyst3.11

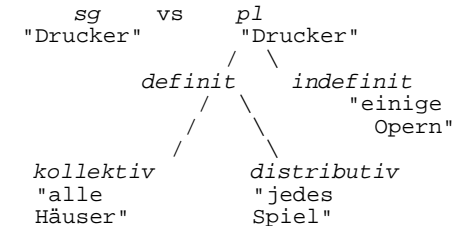
aus: H. SCHWEIZER, Internes Papier zur Morphologie.

- (61) Morphem = kleinste bedeutungstragende Einheit?
 Unterschiedliche Elemente: BEDEUTUNG + EINHEIT (=wohl Ausdrücke)
 Bsp.: Verpflichtung in: Hamburg + Zürich. Wenn ich beiden Verpflichtungen gleichzeitig gerecht werden will, komme ich etwa bis KASSEL (geographische Mitte), komme aber weder nach Hamburg noch nach Zürich. Kassel war aber nie mein Ziel.
 Der traditionelle Morphembegriff hinterläßt also Defizite bei der gleichzeitigen Analyse von SEMANTIK + MORPHOLOGIE³⁹
- (62) Bsp. "Pluralmorphem" Von einer semantischen Position aus (*Plural*) wird in Richtung *Ausdrucksebene* geschaut. Die semantische Einfach-/Einheitlichkeit hat auf Ausdrucksebene völlige Verschiedenheit zur Folge:
Opern
Häuser
Spiele
Drucker (?)
poratu
- (63) Neudefinition von Morphem: Gegeben sind - gedruckt oder handschriftlich - irgendwelche tatsächlichen Realisierungen von Ausdrücken (=Morphe). Meine Aufgabe ist nun, zunächst diese Morphe zu lesen und - durch alle individuelle Realisierung hindurch - zu prüfen, ob mir das Einzelmorph bekannt vorkommt, weil es einem Ausdrucksmuster, das ich gespeichert habe, entspricht. Kommt es zu einem Wiedererkennen, so kann ich weitergehen und fragen, welche Bedeutung ich mit diesem Ausdruck verbinde. Bleibt das Wiedererkennen aus, bleibt der Ausdruck für mich fremd, unleserlich. Die Frage nach der Bedeutung ist obsolet. Als Beispiel für letzteres dient das obige Phantasiewort: *poratu*, das zumindest im Ausdrucksrepertoire des Deutschen keinen Platz hat.⁴⁰
- (64) explizite Ausdrucksanalyse + explizite Semantik
 Nur kurz sei illustriert, wie inkonsequent und verkürzt Grammatiken im Gefolge des traditionellen Morphembegriffs vorgehen, daß nicht nur eine schlüssigere, weil homogenere Ausdrucksanalyse gefordert ist, sondern - im gleichen Zug - eine ausführlichere Semantik.

³⁹ Gestützt auf die ausführliche Rezension von N. MÜLLER [ZS 14.1 (1995) 136-141] kann diese Kritik auch bezogen werden auf: KÖPCKE, K-M, Schemata bei der Pluralbildung im Deutschen. Versuch einer kognitiven Morphologie. Studien zur deutschen Grammatik 47. Tübingen 1993.

⁴⁰ Zu dieser Neudefinition vgl. SCHWEIZER, H, Computerunterstützte Textinterpretation. Die Josefsgeschichte beschrieben und interpretiert im Dreischritt: Syntax - Semantik - Pragmatik. Band III. THLI 7. Tübingen 1995. S.5-8.

In *Opern*, *Häuser*, *Spiele* je das "Pluralmorphem" realisiert zu sehen besagt - semantisch - viel zu wenig. Es müßte gleichzeitig hervorgehoben werden, was für ein "Plural" vorliegt, es müßte gesehen werden, daß das Numerus-Problem mit dem der Determination verknüpft ist, daß ein kognitives System im Hintergrund steht und daß die zufriedenstellende *semantische* Beschreibung nicht auf einzelne, unselbständige Morpheme beschränkt bleiben darf, sondern - bei Bedarf - auch Wortgruppen einbeziehen muß.



Im Fall von "jedes Spiel" haben wir sogar *kein* Pluralmorphem (im Sinn der traditionellen Grammatik) am Nomen, dennoch liegt ein definitiver Plural vor, der alle Einzelelemente der Klasse einbezieht. Das Beispiel zeige exemplarisch die Inkonsistenz der traditionellen Sichtweise oder wie man damit allenfalls nach Kassel kommt, nicht aber nach Hamburg und nach Zürich, wo man eigentlich hinwollte.

Die Diskrepanz rührt daher, dass "Numerus" - obwohl eine genuin semantische Kategorie - in der trad. Grammatik nur behandelt wird, insofern sie *grammatisch*, d.h. über Wortveränderung Spuren hinterläßt. Man sieht zwar, dass es auch *lexikalische* Quantoren gibt. Diese werden aber nicht unter das 'morphologische' Numerusverständnis genommen. Man verleugnet also die Einsicht in den semantischen Charakter aufgrund von Restriktionen, die von der Ausdrucksseite her getroffen werden, u. z. ziemlich willkürlich. Beleg für diese Sicht: J. L. ITURRIOZ-LEZA; S. SKOPETEAS, 100. Numerus, S. 1053-1066 in: BOOIJ, G; LEHMANN, C; MUGDAN, J ; SKOPETEAS, S, Morphologie/Morphology: Ein internationales Handbuch zur Flexion und Wortbildung. 2. Halbband. Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science. 17/2. Berlin 2004: de Gruyter.

zu Ziff. 4.223:

Stilstatistik / Stilometrie

zu:

aus: HÄNLEIN, H, Studies in authorship recognition - a corpus-based approach. EHS XIV,352. Frankfurt/M 1999.

(71) Stylostatistics, which is also called stylometry and which is a relatively recent branch of stylistics, mainly deals with the quantification of style markers. It is defined as "(...) the use of statistical methods to study literary style, authorship, or some other characteristic of an author or a text." (Asher/Simpson 1994 Vol.8:4383) Stylostatistics is a mean of studying empirically a phenomenon that has been approached from a merely intuitive perspective for a long time (Sanders 1973:17). Among the features that are typically investigated in stylometry are sentence length, word-frequency distributions, frequency and distribution of certain parts of speech, grammatical forms and syntactic constructions (Sanders 1973:17). The main instruments of stylostatistics are concordances, word-frequency lists, alphabetical lists and other statistical representations of the run-on text (Bolz 1984:197f)...

(75) Ultimately, even though statistics can help to gain important insights into language despite these idealisations and inaccuracies, to put rather bluntly, the method is incompatible with the object under investigation, language. The quantitative analyses in the empirical part of this study exhibit this paradox to a great extent: due to the fuzziness of linguistic data, the decisions about the assignment of these data to certain classes were sometimes very difficult to make. Many examples did not fit into any of the categories (hence categories such as "various" or "miscellaneous"), others seemed to fit into more than one category at the same time. Therefore, quantitative linguistic analysis is not purely empirical in most cases; most quantitative analyses involve some amount of interpretation.

(79) It is important to be aware of the fact that the numerical data yielded by concordance programs are merely approximations and that most results are flawed by inaccuracies at least to some extent. For instance, the mere statistical data (number of tokens, types, type-token ratio, etc.) elicited by the individual software tools deviate from each other. Likewise, none of the software would identify items like *contain*, *contained* and *containing*, *flag* and *flags*, *Disney* and *Disney's*, etc. as tokens that belongs to the same type. These lexical items, therefore, are counted as separate word-forms. *U.N.* is analysed as two "words" *U* and *N*. Numbers are also broken up: *1,000* gets two entries *1* and *000* in the word lists. The tagged version of the corpus, too, contains a certain number of inaccuracies: the number of modals, for instance, measured by counting the tags for *modal verb* (MD) deviates slightly from the number of modals that results from a cumulative search of the text files for individual modal verb tokens (*would*, *could*, *may*, *might*, etc.). Similarly, the number of proper nouns as analysed manually does not equal the number of *proper noun* tags (PN) in the corpus. It is because of these inaccuracies that the raw corpus was more widely used than the tagged version.

zu Ziff. 4.224:

Rückläufige Wortliste

zu:

aus: Rezension von Martin NEEF zu: Gustav Muthmann, Reverse English Dictionary. Based on Phonological and Morphological Principles. Berlin 1999: Zeitschrift für Sprachwissenschaft 20.2 (2001)318-319.

(318) Rückläufige Wörterbücher sind eine segensreiche Erfindung, nicht nur für Morphologen. Da für die englische Sprache seit einer strikt alphabetischen Wortliste von Martin Lehnert (1971) kein derartiges Werk mehr auf dem Markt erschienen ist, hat sich Gustav Muthmann, ausgewiesen durch ein rückläufiges (1988) und ein phonologisches Wörterbuch zum Deutschen (1996), an diese Aufgabe gesetzt und nun ein Buch publiziert, das durch eine reiche Binnenstruktur wissenschaftlichen Fortschritt dokumentiert. Wie der Untertitel verkündet, folgt Muthmanns Wörterbuch nicht nur alphabetischen, sondern auch phonologischen und morphologischen Prinzipien. Morphologische Prinzipien sind bereits aus Muthmann (1988) bekannt und bewährt; sie ermöglichen es, Wörter, die mit dem gleichen Suffix enden oder dasselbe Kompositumshinterglied besitzen, auf einen Blick zu überschauen. So sind Abteilungen mit dem Suffix *-able* wie *passable* in der Wortliste räumlich getrennt von Wörtern mit dem freien Morphem *able* wie *disable* und von Wörtern, die mit der Buchstabenfolge *<able>* enden, der kein Morphemstatus zukommt, wie *sable*. Hilfreich ist dabei die Angabe der Wortart für jedes Wort (wie auch bei Lehnert 1971).

Wesentlich erweitert gegenüber Muthmanns rückläufigem Wörterbuch zum Deutschen (1988) ist der Bezug auf phonologische Prinzipien. Die Wörterliste führt die einzelnen Lemmata nämlich nicht nur in orthografischer Form auf, sondern zugleich in einer phonetischen Transkription, die sich aus pragmatischen Gründen an der britischen Aussprache orientiert (Abweichungen des amerikanischen Englisch und Varianten werden aber berücksichtigt). Diese Erweiterung des Konzepts eines rückläufigen Wörterbuchs ist gerade für die englische Sprache sinnvoll, weil identische Buchstabenfolgen oft unterschiedlich ausgesprochen werden. Dies illustriert die obige Folge *<able>*, die mal [-bl], mal [erble] ausgesprochen wird. Da die Wortliste optisch sehr ansprechend und übersichtlich gestaltet ist, genügt eine kurze Einarbeitungszeit, um Zusammenhänge zwischen Wörtern erkennen zu können, die ein vorwärts-alphabetisch geordnetes Wörterbuch nicht zu geben vermag.

4.3 Neuansatz: Inhaltsfreie Morphologie -

Literatur

- BORGES, Jorge Luis: Die Bibliothek von Babel in: BORGES, J L (ed./eds.): Die Bibliothek von Babel. Erzählungen. Universal-Bibliothek Nr. 9497. Stuttgart 1974 47-57.: Reclam
- DRESSLER, W U; LUSCHÜTZKY, H C; PFEIFFER, O E; RENNISON, J R (ed./eds.): Contemporary Morphology. Trends in Linguistics. Studies and Monographs 49. Berlin 1990 // I.2.7
- DREWS, E: Die Bedeutung von Morphemen für die Sprachanalyse. Zur mentalen Verarbeitung lexikalischer und grammatischer Morpheme. Psycholinguistische Studien. Opladen 1989 // I.2.7
- FLENNER, G: Ein Quantitatives Morphsegmentierungssystem für spanische Wortformen in: KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 31-62. // DI.2.7
- FREEDMAN, D N; FORBES AD; ANDERSEN FI: Studies in Hebrew and Aramaic orthography. Biblical and Judaic Studies from the University of California, San Diego, 2. Winona Lake 1992 // I.2.0
- HALLE, M; MARANTZ, A: Distributed Morphology and the Pieces of Inflection in: HALE, K; KEYSER, S J (ed./eds.): The View from Building. 20 Essays in Linguistics in Honour of Sylvain Bromberger. Cambridge, MA 1993 111-176.: MIT-Press
- HAN, Y S; PARK, H R (et.al.): An Upper Bound Estimate for the Entropy of Korean Texts. Literary and Linguistic Computing 11 / 3 (1996) 141-146. // EB
- HARRIS, Z S: Distributional Structure. WORD X (1954) 146-162.
- HEIN, D: UNIX gestützte maschinelle morphologische Untersuchung zur Komposition am Beispiel des DUDEN (Deutsches Universal Wörterbuch). Holos Reihe Linguistik 6. Bonn 1995: Holos // I.2.7
- JANSSEN, A: Segmentierung französischer Wortformen in Morphe ohne Verwendung eines Lexikons in: KLENK, U (ed./eds.): Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.73. Stuttgart 1992 74-95. // I.2.7
- KLENK, U (ed./eds.): Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.73. Stuttgart 1992 // I.2.7
- KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 // I.2.7
- KLENK, U; LANGER, H: Morphological Segmentation Without a Lexicon. Literary and Linguistic Computing 4 (1989) 247-253.
- KLENK, U: Verfahren morphologischer Segmentierung und die Wortstruktur des Spanischen in: KLENK, U (ed./eds.): Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.73. Stuttgart 1992 110-124. // I.2.7
- KLENK, U: Automatische morphologische Analyse arabischer Verbformen in: KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 84-101. // I.2.7
- KRÄMER, S: Sprache und Schrift oder: Ist Schrift verschriftete Sprache?. Zeitschrift für Sprachwissenschaft 15,1 (1996) 92-112.
- LANGER, H; NAUMANN, S: Syntaktische Hierarchie und lineare Abfolge in: KLENK, U (ed./eds.): Computatio Linguae. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.73. Stuttgart 1992 125-145. // I.2.7
- NEEF, M: Wortdesign: Das Lexembildungsmuster *Gehopse* und die Kopfllosigkeit von 'Ableitungen'. Zeitschrift für Sprachwissenschaft 15,1 (1996) 61-91.
- RENOUF, A: What do you think of that: A pilot study of the phraseology of the core words of English in: LEITNER, G (ed./eds.): New Directions in English Language Corpora. Methodology, Results, Software Developments. Topics in English Linguistics 9. Berlin, New York 1992 301-317. // I.2.1
- SCHINDELE, M: Darstellung morphologischer Zerlegungen hebräischer Wörter. Biblische Notizen 75 (1994) 22-25.
- SCHWAB; NUSSBAUM (ed./eds.): Pattern recognition by humans and machines. Vol.1: Speech perception. Vol.2: Visual perception. Academic Press Series in cognition and perception 21987
- SCHWEIZER, H: Sprachkritik als Ideologiekritik. Zur Grammatikrevision am Beispiel von ORB. THLI 1. Tübingen 1991: Francke // I.2.7
- WEBER, N: Wortanalytisches Wörterbuch. Strukturelle Bauelemente deutscher Wörter als maschinenlesbares lexikographisches Material. Sprache und Datenverarbeitung 12 (1988) 3-18.

4.31 Cursorische Beobachtungenzu Ziff. 4.311:**Zeichenkörper - Konvention - Bedeutung**

zu: vgt2.121

aus: TAWADA, Y, Übersetzungen. Tübingen 2002.

(33f) Eine Sprache, die man nicht gelernt hat, ist eine durchsichtige Wand. Man kann bis in die Ferne hindurchschauen, weil einem keine Bedeutung im Weg steht. Jedes Wort ist unendlich offen, es kann alles bedeuten.

Ich sehe das Wort "du". Es ist schwierig zu glauben, daß es gar nichts mit dem deutschen Wort "du" zu tun hat. Ein "du", das man nicht kennt, kann alles bedeuten: ein Getreidesack, eine Anziehpuppe, eine Taube oder eine Tür. Egal, was ich mir darunter vorstelle, die beiden Buchstaben. "d" und "u" bleiben so, wie sie sind. Die Schriftzeichen interessiert es vielleicht gar nicht, was sie in einem Land bedeuten. In Deutschland bedeuten sie das, in Frankreich jenes. Sie sind Reisende, sie werden unterwegs immer wieder anders verstanden, je nachdem, in welcher Sprache sie übernachten. Ihre Körper bleiben aber dieselben, nämlich ein "d", ein Halbkreis mit einer erhobenen Hand, und ein "u", ein leeres ... Gefäß.

Eine Sprache, die man nicht versteht, liest man äußerlich. Man nimmt ihr Aussehen ernst. Das Gesicht eines französischen Textes sieht runder aus als das eines deutschen. Es fehlen die eckigen Schultern der großen Buchstaben, die im Deutschen jeder Zeile einen architektonischen Charakter geben.

zu Ziff. 4.312:

Ausdrucksformale Morphologie / Hebräisch

zu:

Aus: SCHWEIZER, H, Sprachkritik als Ideologiekritik. Zur Grammatikrevision am Beispiel von QRB. THLI 1. Tübingen 1991. Vgl. S. 51-55.

Separat sind folgende Wörter belegt:

Gen 10,20	...LMŠRHTM LLŠNTM	MŠRH LŠN
Gen 23,18	...L'BRHM LMQNH L'YNY	'BRHM LMQNH 'YN
Dtn 28,26	...LM'KL LKL	M'KL KL
Gen 32,1	...LBN LMQMW	BN MQMW
Gen 36,40	...LMŠPHTM LMQMTM BŠMTM	MŠPH MQM BŠM
Gen 13,2	...BMQNH BKSP WBZHB	MQNH KSP ZHB
Gen 21,14	...BMDBR B'R	MDBR B'R 'R
Dtn 1,1	...BMDBR B'RBH	MDBR B'RBH
Gen 50,26	...B'RWN BMSRYM	'RWN MŠRYM
Dtn 2,26	...ML'KYM MMDBR	ML'KYM MDBR

Folgende weitere Wörter sind isoliert belegt: (sofern nicht aus den Belegen erschließbar)

Gen 5,15	...HMŠ ŠNYM WŠŠYM ŠNH	ŠŠYM
Gen 5,25	...ŠB' WŠMNYM ŠNH WM'T ŠNH	ŠMNYM M'T
Gen 11,5	...LR'T 'T H'YR W'T HMGDL	'YR MGDL
Gen 13,7	...BYN R'Y MQNH 'BRM W+	
	BYN R'Y MQNH LWT	
Gen 31,51	...T HGL HZH W'T HMSBH HZ'T	
Dtn 28,20	...T HM'RH 'T HMHWMH W'T HMG'RT	M'RH MHWMH MG'RT

Andere Formen wenn isoliert:

Gen 23,9	...M'RT HMKPLH	M'RH MKPLH
Gen 23,19	...L PNY MMR'	PNYM
Gen 1,16	...ŠNY HM'RT HGDLYM	ŠNYM GDLYM
Gen 27,7	...LPNY MWTY	MWT
Gen 29,22	...NŠY HMQWM	'NŠYM MQWM
Gen 30,3	...N'KY MMNH	

zu Ziff. 4.313:

Wortschatz

zu:

aus: B. GLOOR, staat - sex - amen. 81 Sprachbeobachtungen. Zürich 2004.

Nr. 42

Wie viele Wörter hat der Mensch?

Adenauer soll nur 800 gehabt haben. Und Kohl? Und Schröder?

Der Wortfrequenzforscher J.A. Pfeffer weist nach, dass 1300 Wörter 90 Prozent eines einfachen Gesprächs ausmachen. Lässt man als Alltagsthemen Bereiche wie Wohnen, Essen, Verkehr, Gesundheit etc. gelten und rundet noch ein wenig auf, sind es 2000. So gross ist ein Grundwortschatz, den man zum Erlernen einer neuen Sprache überall kaufen kann. Darunter geht nichts.

Kafkas "Prozess" ist aus 6500 Wörtern gemacht, Melvilles "Moby Dick" aus 12 000. Den grössten Wortschatz dürfte Joyce gehabt haben: Allein der "Ulysses" enthält 32 000 verschiedene Wörter.

Die Lexika der deutschen Standardsprache ("Duden Wörterbuch" oder "Brockhaus Wahrig", je 6 Bd.) enthalten rund 250 000 Wörter. (Die Zahl der Verknüpfungsregeln dieser Wörter wird auf rund 1000 geschätzt.) Der nicht besonders ausführliche Duden-Begleitband "Medizinische Fachausdrücke" verzeichnet 30 000 Begriffe. Die Kartothek der Wörterbuchkanzlei der Akademie der Wissenschaften in Wien enthält über 5 Millionen Wörter deutscher Mundarten. Die Sammlung der Schweizer Mundartwörter, das so genannte Idiotikon, wird bei seiner Fertigstellung (falls es jemals fertig wird) über 500 000 Einträge enthalten. Die technische Terminologiebank "Team" verzeichnet 3 Millionen technische Fachausdrücke. Die Anzahl chemischer Verbindungen beläuft sich auf 5 Millionen, wobei jeden Tag etwa 1000 dazukommen. Auch in der Biologie hat jede der rund 6 Millionen Arten ihren eigenen Namen. Dann die Steine, die Sterne, die Firmen...

Wie viele Wörter gibt es? Schwer zu sagen. 50 Millionen? 80? Und wie viele Wörter hat ein Mensch? Ein gebildeter verwendet 10 000 bis 15 000 und versteht vielleicht das Vierfache davon. Aber Benns gesamte Lyrik ist aus bloss 2700 Wörtern gemacht. Und Rilkes Gedichte umfassen gerade mal 5000. Der Wortschatz ist nur das eine. Die Beherrschung der Verknüpfungsregeln bestimmt die sprachliche Kompetenz viel mehr als blosser Wortreichtum.

Buchstabenstatistik

Der Franzose Georges Perec hat einen Roman geschrieben. Er heisst "La disparition" und enthält kein einziges e. Die Übersetzung ins Deutsche besorgte (ausgerechnet!) ein Mann mit ebenfalls vier e im Namen: Eugen Helmlé. Die deutsche Ausgabe heisst "Anton Voyls Fortgang" (1986) und ist ein Kunststück für sich. Ein Auszug:

"Mobiliar und Luxus sagt ihm nichts, darum ist Antons Wohnung schlicht und schmucklos. Kalkwand, Tisch, Stuhl und Sofa, und dazu stinkt's furchtbar nach Knoblauch. Damit hat sich's. Für Bad und so was hat Anton Voyl nichts übrig, hält's für nutzlos, das ist für ihn Klimbim und Hokuspokus."

In diesen rund 300 Zeichen müsste statistisch 37-mal der Buchstabe e vorkommen. Sein Fehlen bewirkt bei fortgesetztem Lesen etwas Dumpfes und Lichtloses in der Sprache. Immerhin ist das e im Deutschen (wie auch im Französischen) mit Abstand der häufigste

Buchstabe. Sein Prozentanteil liegt bei 12,9, gefolgt von n (7,9), i (6,4), r (5,9) und s (5,7). Diese Reihenfolge e-n-i-r-s ist nicht in allen Sprachen gleich. Sie lautet im Französischen e-s-a-n-i, im Englischen e-o-t-h-a. Erwartungsgemäss am seltensten sind im Deutschen die Buchstaben q, j, x, y und ö (jeweils unter 0,2).

4.32 Ausdrucksbasierte Formenlehre (Morphologie)

Literatur

FUHRHOP, N: Das graphematische Wort (im Deutschen): Eine erste Annäherung-. Zeitschrift für Sprachwissenschaft 27 (2008) 189-228.

zu Ziff. 4.321:

Hjelmslev: Texte / Segmentierung

zu: vprag4.301 vgt2.123

aus: A. MEHLER, Textbedeutung. Zur prozeduralen Analyse und Repräsentation struktureller Ähnlichkeiten von Texten. Sprache, Sprechen und Computer 5. Frankfurt/M 2001.

(31) Die primären Erfahrungsdaten der Sprachwissenschaft bestehen somit in Mengen unanalysierter Texte und nicht in isolierten Sätzen. Texte stellen die einzigen sprachlichen Größen dar, die dem Sprachwissenschaftler vor jeder Analyse gegeben sind. Hieraus folgt, daß die Einheiten der Sprache (ob Wörter oder Sätze) dem Sprachwissenschaftler nicht unmittelbar vorliegen. Ihre Abgrenzung resultiert vielmehr aus der *Textanalyse*: (32)

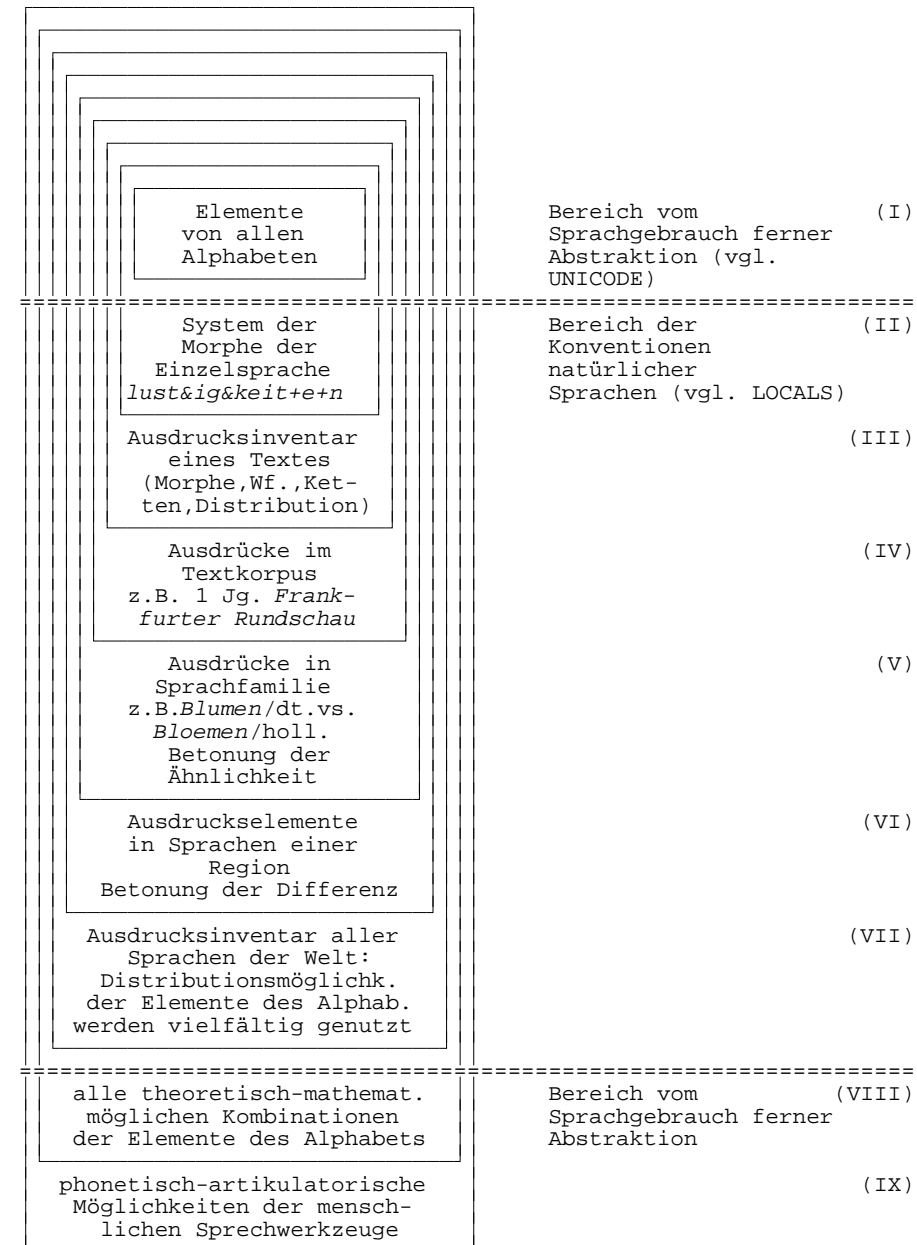
"Die Gegenstände, für die sich die Sprachtheorie interessiert, sind Texte. Das Ziel der Sprachtheorie ist, eine Verfahrensweise zu entwickeln, mittels derer ein vorliegender Text durch eine widerspruchsfreie und erschöpfende Beschreibung begriffen werden kann. Aber die Sprachtheorie muß ebenso angeben, wie man jeden beliebigen anderen Text von der gleichen angenommenen Beschaffenheit in derselben Weise begreifen kann, und dies geschieht, indem uns die Sprachtheorie Werkzeuge an die Hand gibt, die für jedweden Text dieser Art benutzt werden können." (Hjelmslev 1974:21).

Die Zugrundelegung des Strukturprinzips hat zur Folge, daß Texte *unabhängig* von ihrer pragmatischen bzw. kommunikativen Einbettung analysiert werden, und zwar ausschließlich mit Hilfe der *deduktiven* Textanalyseprozedur. Dabei bedeutet *Widerspruchsfreiheit*, daß kein Text *T* vor dem Hintergrund desselben Analyse Kriteriums auf verschiedene Weisen zerlegt werden darf. Ziel ist es, widersprüchliche Aussagen der Art "*A ist eine Komponente von T*" und "*A ist keine Komponente von T*" auszuschließen. Die Forderung nach *erschöpfender Analyse* bezieht sich darauf, daß jeder Text (von bestimmter vorausgesetzter allgemeiner Beschaffenheit) vollständig zu zerlegen ist, ohne daß komplexe Komponenten unzerlegt bleiben. Insbesondere das Kriterium der *Widerspruchsfreiheit* orientiert sich an einem Binarismus (entweder ist *A* eine Komponente von *T*, oder nicht) dessen Geltung im Rahmen der Glossematik vorausgesetzt wird.

zu Ziff. 4.322:

Ausdrucks-SYNTAX in systemtheoretischer Sicht

vsyst4.02



(I) + (IX): Das Alphabet gibt es nicht. Es gibt viele ähnliche, aber in Teilen eben doch auch abweichende Alphabete. Ein Kernbereich dieser Schrift Elemente stimmt mit Unschärfen je überein. Er hängt damit zusammen, daß es phonetisch eine begrenzte Menge artikulatorischer Möglichkeiten gibt - durch alle Sprachen und Kulturen hindurch ("Der Kehlkopf Davids war prinzipiell kein anderer als unserer"). Aus diesen phonetischen Gesamtmöglichkeiten wählt die einzelne Sprache eine Untermenge aus (das Arabische hat Kehllaute, die im Deutschen nicht vorkommen); daraus wiederum wählt die Schrift der Einzelsprache die Elemente aus, die sie repräsentieren will (Kehlkopfverschluss bei mit Vokal beginnenden Wörtern - 'aber im Gegensatz zu a in *Tal* - wird im Deutschen nicht geschrieben,

im Arabischen sehr wohl *أَبْرَاهِيمُ*.

- Piktographische u.a. Schriften müßten erst entsprechend transkribiert werden.

(I) + (II): Aus den wenigen Elementen der in der Einzelsprache gebräuchlichen Schrift (bedeutend komplexer: Japanisch, Chinesisch) werden durch Kombinatorik sehr viele und unterschiedliche *Morphe* gebildet (=selbständige Wortformen; unselbständige Wortkerne, die sich mit weiteren unselbständigen Bildelementen verbinden; Compounds).

Es kann dann *synchron* sein, dass sich z.B. die gleichen Wortformen in mehreren Sprachen finden. Bisweilen ist dies durch *diachrone* Verbindung erklärbar. Die Ausdrücke:

i vitelli dei romani sono belli

sind auf Latein und auf Italienisch lesbar, bedeuten dann zwar völlig Verschiedenes. Aber das Wortbildungsrepertoire beider Sprache hat aufgrund ihres geschichtlichen Zusammenhangs Überlappungen.

Andere Identitäten kommen aufgrund von Zufällen zustande. So wohl auch in der Zeitungsoberschrift: "Mies ist hier jeder Mann". Der Reporter versteht *Mies* dabei vom Finnischen her, wo *Mies* "Mann" heißt.

(II)+ (III): Ein Einzeltext wählt aus dem gesamten Formen- und Wortschatz der Einzelsprache eine eigene Menge in eigener Verteilung aus: Der Einzeltext wird schon auf Ausdrucksebene unverwechselbar, erhält seine eigene Struktur, vgl. *Bader* in THLI 7.

(III) + (IV): Der Einzeltext steht bezüglich seiner Ausdrucksgestalt in Relation zum umgebenden Sprachgeschehen. Untersuchbar: Anspielung, Formelverwendung, Klischees, Zitat, Nachahmung eines vorgegebenen Stils, vgl. *Schindeler* in THLI 7.

ad (V): Komparatistik; romanische, anglistische, germanistische, semitische ... Philologie.

ad (VI): In Europa gibt es Texte in Deutsch, Spanisch, Baskisch, Finnisch - um nur Vertreter nicht-zusammengehörender Sprachfamilien zu nennen.

(VII) + (VIII): Es gibt keine Berechtigung anzunehmen, beide Mengen wären identisch. (VII) ist eine nicht-überschaubare Gesamtmenge, die auch durch historische Zufälligkeiten bestimmt ist. - (VIII) ist im Prinzip errechenbar.

[Vgl. die literarische Gestaltung des gleichen Gedankens: JORGE LUIS BORGES, *Die Bibliothek von Babylon*. Reclam

Und die Besprechung davon in: V. A. FERRETTI, *Boreale Geltung*. Zu Nördlichkeit, Raum und Imaginärem im Werk von Jorge Luis Borges. Frankfurt/m 2007. S. 241ff //I.2.7]

zu Ziff. 4.3221:

Différance: endlose Kette der Signifikanten (DERRIDA)

zu: vsyst4.031

aus: C. EMIG, Gewagte Spiele. Derrida und die ästhetischen Folgen. in: M. FICK; S. GÖßL (eds.), Der Schein der Dinge. Einführung in die Ästhetik. Tübingen 2002: Attempto. S.187-205. //I.2.0

(196) Es gibt also keine ein-deutige Bedeutungszuweisung mehr, jedes materielle Wortzeichen verweist, einer festen und unerschütterlichen sinngebenden Instanz ledig, auf eine schier unendliche Menge von Bedeutungsmöglichkeiten. Der Sinn kommt so in Bewegung, hat innerhalb des sprachlichen Systems keinen sicheren unveränderlichen Halt. Man braucht Mut, dies weiterzudenken, führt es doch, wie Derrida angekündigt hat, letztendlich zu einer Bedrohung des Subjekts, dessen Identität selbst im Verschwinden begriffen ist. Auch es ohne zentrierte Struktur, ein Netzwerk unterschiedlichster Meinungen und Einschreibungen, substantiell durch das Spiel von Texten bestimmt... Soll dieses subjektlose Subjekt nicht die Sinneffekte kontrollieren können?

Derrida kreiert für das "Spiel der Differenzen", in dem in einer endlosen Kette Signifikanten aufeinander verweisen, ein Kunstwort, die "différance". Es soll darauf hinweisen, daß der Sinn keinen Halt mehr findet, sondern immer aufgeschoben wird, nie "präsent" ist. Die "différance" ist ein Spiel ohne Zentrum und festen Grund, das aber gleichwohl selbst die einzige Grundlage von Sprache und Bedeutung darstellt. Derrida formuliert das so:

Sich zu erlauben, nichts sagen zu wollen, bedeutet zu spielen anfangen, und zuallererst, in das Spiel der différance einzutreten, das bewirkt, daß kein Wort, kein Begriff, keine Hauptaussage sich anschicken, von der theologischen Präsenz eines Zentrums aus die Bewegung und die textuelle Ausbreitung der Differenzen zusammenzufassen und zu lenken.

Das heißt aber auch: Alle Zeichen stehen zueinander in einem kontextuellen Zusammenhang, da jedes angewiesen ist auf die semantische Abgrenzung von anderen Zeichen. Diesen kontextuellen Zusammenhang nennt Derrida den "allgemeinen Text". Es gibt nichts außerhalb des Textes.

zu Ziff. 4.323:

Verschmelzen von Lesarten

zu:

aus: M. SCHULZE, Inkrementelle grammatische Induktion (116ff)

Segmentierung
in Morpheme

	Kategorie	Bedeutung
Ab/treib/ung/s/gesetz	(NOM: N-G)	Kompositum aus <u>ab, treiben, Gesetz</u>
Abt/reib/ung/s/gesetz	(NOM: N-G)	Kompositum aus <u>Abt, reiben, Gesetz</u>
Mehrzweck/saal	(NOM: M-G)	Kompositum aus <u>Mehrzweck, Saal</u>
Mehrzweck/s/aal	(NOM: M-G)	Kompositum aus <u>Mehrzweck, Aal</u>
...		
gehabt	(PII:PII haben)	Simplex: Partizip II
Geh/abt	(NOM:M-G)	Kompositum: <u>gehen, Abt</u>
Aktion	(NOM:F)	Simplex: Nomen
Akt/ion	(NOM:N-G)	Kompositum: <u>Akt, Ion</u>
Abteilungen	(NOM: P)	Simplex: Nomen Plural
Abtei/lungen	(NOM: P)	Kompositum: <u>Abtei, Lunge</u>
Abt/ei/lunge/n	(NOM: P)	Kompositum: <u>Abt, Ei, Lunge</u>

zu Ziff. 4.3231:

ausdruckssyntaktische Morphologie - Zerlegungen

zu:

aus: B. GLOOR, staat - sex - amen. 81 Sprachbeobachtungen. Zürich
2004.

Siebenundsiebzig falsche Trennungen Nr. 23

schrei	asyl	bet
bau	such	reibung
tomaten	ende	samt
lach		
sauf	politik	stromer
lauf	ersatz	zeugung
verkauf		
shit	miss	staub
	b	ecken
	rauch	
wahns	plange	kohl
innig	echt	enkel
lerchen		
massa	tram	europ
geigel	polin	apolitik
bluto	puff	all
range	erz	einer
	one	zieh
		ende
konsum	ego	to
enten	ist	get
	in	her
	/	
	ego	
	ist	
	innen	
delikat	arbeit	multi
essen	sauf	tal
	wand	ente
schwer		recht
starb	t	sex
eiter	eee	trem
innen	i	
pol	gen	
armee	italien	vorm
trem		arsch
be	na	land
steh	omi	es
ende		meist
/		er
best		schaft
eh		
ende		
ball	weit	au
ade	erb	tot

haut	ringen	ran
schon		sport
ende	komm	fing
	union	er
	/	ring
	komm	/
	unikator	ring
		fing
		er

. . .

zu Ziff. 4.3232:

Bedeutungsfreie Morphologie

zu:

aus: Duk Ho Lee, Rückläufiges Wörterbuch der deutschen Sprache. Berlin 2005.

(IX) 1.1 Terminologie

Im RWDS finden folgende Terminologien und Abkürzungen Verwendung:

(1) Stichwörter (SW)

Damit werden alle unspezifischen normalen Stichwörter bezeichnet.

(2) Kernstichwörter (KSW)

KSW bilden Wortreihen durch Wortzusammensetzung oder Kompositabildung. Die KSW stehen am Anfang der jeweiligen Wortreihe. Sie sind zur besseren Erkennbarkeit und wegen ihrer hierarchisch wichtigen Leitfunktion fett gedruckt.

(3) Ersatzkernstichwörter (EKSW)

EKSW üben die gleiche Funktion wie KSW aus, sind aber morphologisch unselbständige Worbildungselemente. Hierzu gehören alle Suffixe, die erstens gebundene Morpheme sind und zweitens nicht als selbständige Wörter vorhanden sind, wie z.B. "(ig)", "(lich)", "(haft)", "(keit)", "(ung)", sowie die Wortbestandteile, die in ihrer Stammform nicht gebräuchlich sind, aber Reihenbildungsfähigkeit besitzen, z.B. "(Trag)", mit den Komposita "Eintrag" (ZSW), "Auftrag" (ZSW) usw. oder "(Handlungstisch)" mit den Komposita "Behandlungstisch" (ZSW), "Verhandlungstisch" (ZSW) usw. Die EKSW sind in runde Klammern gesetzt und fett gedruckt.

(4) Grundkernstichwörter (GKSW)

GKSW üben die gleiche Funktion wie KSW aus, besitzen aber eine höhere hierarchische Funktion als die KSW, indem es in den betreffenden Reihen mehrfache KSW-Relationen in Form von Subgruppen gibt (z.B. "Weg" - "Wanderweg" - "Rundwanderweg"). Sie sind zusätzlich zum Fettdruck der KSW und EKSW einen Schriftgrad größer gedruckt.

(5) Zugehörige Stichwörter (ZSW)

ZSW sind die Stichwörter, die einer KSW/EKSW/GKSW-Reihe zugehörig sind und die daher innere Mitglieder der Reihen darstellen.

(6) Schlusstichwörter (SSW)

SSW sind die jeweils letzten Mitgliedswörter der Wortreihen. Sie sind zur besseren Erkennbarkeit der Grenzfunktion kursiv und fett gedruckt.

(7) Großschlusstichwörter (GSSW)

GSSW sind die jeweils letzten Mitgliedswörter der Wortreihen eines GKSW. Sie sind zur besseren Erkennbarkeit der Grenzfunktion auch kursiv und fett und zusätzlich zu den SSW einen Schriftgrad größer gedruckt.

(8) Wortreihen

Als Wortreihen werden die KSW/EKSW/GKSW mit der Gesamtheit ihrer ZSW bezeichnet. Es gibt somit drei Reihen: KSW-Reihen, EKSW-Reihen und GKSW-Reihen. Die Wortreihen haben im RWDS zentrale Funktionen, indem sie alle deutschen Abteilungen und Zusammensetzungen, die Reihen bildende Funktionen haben, zur Bildung der Wortfamilien und Inventare integrieren.

1.2 Makro- und Mikrostrukturebene

1.2.1 Makrostrukturebene

Die Stichwörter folgen auf der Makrostrukturebene grundsätzlich der linksläufigen, d.h. finalalphabetischen Anordnung. Hierzu gehören:

(1) Alle isoliert vorkommenden Stichwörter

(2) Alle Reihen bildenden Stichwörter: KSW, EKSW, GKSW

Die Finalität der SW-Anordnung wird auf der Makrostrukturebene durchgehend beibehalten, und die Grenzen aller Reihen, die durch KSW, EKSW, GKSW eingeleitet werden, sind mittels verschiedener graphischer Auszeichnungen wie Fettsatz und Kursivdruck deutlich markiert, wodurch Anschaulichkeit und Anordnungsstruktur immer bestens erkennbar bleiben.

1.2.2 Mikrostrukturebene

Die Stichwörter auf der Mikrostrukturebene sind rechtsläufig, d.h. initialalphabetisch angeordnet. Zur Mikrostrukturebene gehören alle ZSW, also die Stichwörter, die sich in einer Wortreihe befinden und daher innere Mitglieder der Subgruppen sind. Diese Initialalphabetisierung der Mikrostrukturebene dient vor allem der übersichtlichen Gestaltung der Wortfamilien, indem alle ZSW, die jeweils die Mitglieder der Reihen sind, von oben nach unten rechtsbündig alphabetisch angeordnet werden. Dieses neue Anordnungsprinzip des RWDS hat vor allem deshalb eine große Bedeutung, weil das Deutsche eine Sprache ist, die eine besondere Präferenz für Wortzusammensetzungen aufweist. Dies zeigt die große SW-Zahl einer KSW-Reihe, die nicht selten viele hunderte SW umfasst. Eine strikt finalalphabetische Anordnung der SW wäre deshalb für die Gestaltung der Mikrostrukturebene ungeeignet.

1.3 Beispiel der Anordnungsprinzipien

Nachfolgend werden die Vor- bzw. Nachteile der Anordnungsprinzipien an Hand der Wortgruppe "Kauf" dargestellt.

Die erste Liste (I. RWDS) stammt aus dem RWDS (Lee 2005), die zweite (II. Muthmann) aus Muthmann 2001.

I- RWDS 2005

Kauf
 Abkauf
 Abzahlungskauf
 Aktienkauf
 Ämterkauf
 Angstkauf
Ankauf
 Autoankauf
 Barankauf
 Großankauf
 Kunstankauf
Neuankauf
 Aufkauf
 Auskauf
 Autokauf
 Barkauf
 Bestimmungskauf
 Beziehungskauf
 Blankokauf
 Brautkauf
 Buchkauf
 Computerkauf
 Deckungskauf
 Detailkauf
 Distanzkauf

Domainkauf
 Ebay-Kauf
 Effektivkauf
Einkauf
 Bareinkauf
 Computereinkauf
 Direkeinkauf
 Fehleinkauf
 Gemeinschaftseinkauf
 Geschenkeinkauf
 Großeinkauf
 Gruppeneinkauf
 Handyeinkauf
 Interneteinkauf
 Internet-Einkauf
Lebensmitteleinkauf
Online-Lebensmittel-Einkauf
 Markteinkauf
 Masseneinkauf
 Materialeinkauf
 Neueinkauf
 Notebookeinkauf
 Offline-Einkauf
 Onlineeinkauf
 PC-Einkauf
 Radioeinkauf
 Samstagseinkauf
 Urlaubseinkauf
 Wareneinkauf
 Weihnachtseinkauf
 Wocheneinkauf
Wochenendeinkauf
 Einzelkauf
 Erbschaftskauf
 Fehlkauf
 Fernkauf
 Fernsehkauf
 Finanzkauf
 Fixkauf
 Frauenkauf
 Gattungskauf
 Gelegenheitskauf
 Genusskauf
 Gesamtkauf
 Grundkauf
 Grundstückskauf
 Gutskauf
 Hamsterkauf
 Handelskauf
 Handkauf
 Handykauf
 Hauskauf
 Hoffnungskauf
 Impulskauf
 Internetkauf
Kartenkauf
 Bahnkartenkauf
 Eintrittskartenkauf
Fahrkartenkauf
 Kreditkauf
 Kunstkauf
 Leerkauf

Leihkauf
 Leitkauf
 Loskauf
 Mantelkauf
 Maschinenkauf
 Meinungskauf
 Mietkauf
 Nachkauf
 Neukauf
 Notebookkauf
 Online-Kauf
 Panikkauf
 PC-Kauf
 Prämienkauf
 Ramschkauf
 Ratenkauf
 Reukauf
 . . .

II. Muthmann 2001

Kauf
 Abkauf
 Aufkauf
 Leihkauf
 Leikauf
 Rückkauf
 Stückkauf
 Zielkauf
 Fehlkauf
 Ankauf
 Aktienkauf
 Stimmenkauf
 Ratenkauf
 Einkauf
 Fehleinkauf
 Weihnachtseinkauf
 Großeinkauf
 Direkteinkauf
 Neueinkauf
 Scheinkauf
 Weinkauf
 Barkauf
 Wiederkauf
 Ämterkauf
 Hamsterkauf
 Verkauf
 Abverkauf
 Pfandverkauf
 Abendverkauf
 Handverkauf
 Resteverkauf
 Ramschverkauf
 Einzelverkauf
 Ladenverkauf
 Straßenverkauf
 Kartenverkauf
 Kleinverkauf
 Alleinverkauf
 Wiederverkauf
 Sonderverkauf
 Weiterverkauf
 Leerverkauf

Vorverkauf
 Kartenvorverkauf
 Inventurverkauf
 Räumungsverkauf
 Ausverkauf
 Totalausverkauf
 Saisonausverkauf
 Inventurausverkauf
 Räumungsausverkauf
 Schlussverkauf
 Saisonschlussverkauf
 Sommerschlussverkauf
 Winterschlussverkauf
 Flußverkauf
 Direktverkauf
 Notverkauf
 Billetverkauf
 Vorkauf
 Spezieskauf
 Beziehungskauf
 Deckungskauf
 Abzahlungskauf
 Meinungskauf
 Gattungskauf
 Stützungskauf
 Handelskauf
 Spezifikationskauf
 Loskauf
 Gelegenheitskauf
 Genusskauf
 Mietkauf
 Kreditkauf
 Leitkauf
 Angstkauf
 Zukauf
 Neukauf
 Reukauf

zu Ziff. 4.324:

Lexikographie - Lexikologie

zu: vsem3.121

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(2) Beginnen wir mit den knappen Erläuterungen zur Lexikographie:

1. Ihr allgemeinsten Gegenstandsbereich sind Wörterbücher. Wörterbücher sind nichts Vorgegebenes, sie sind bewußte Schöpfungen. Sie sollen ihren Benützern/Benutzerinnen als Ratgeber zu Fragen über Wörter und den Wortschatz dienen. Insofern müssen sich die Lexikographen/Lexikographinnen mit den Bedürfnissen potentieller Benutzer/Benutzerinnen auseinandersetzen und sich fragen, wie diese Bedürfnisse in geeignete und realisierbare Wörterbuchformate umgesetzt werden können...

2. Die Theorie der Lexikographie beschäftigt sich mit den möglichen und wünschenswerten Formaten von Wörterbüchern, mit den Voraussetzungen ihrer Erstellung und mit den Kriterien zur Bewertung der angesprochenen Abbildfunktion zwischen Wortschatz und Wörterbuchformat...

Kommen wir zu den Erläuterungen zur Lexikologie:

(3) 2. Allgemeinsten Gegenstandsbereich der Lexikologie sind Wortschätze. Wenn wir uns einig sind, daß Wortschätze irgendwie aus Wörtern bestehen, dann scheint klar zu sein, daß Wortschätze einen gewissen Status der Realität genießen, uns also in gewisser Weise vorgegeben sind. Dies ist ein wichtiger Unterschied zum Status des Gegenstandsbereiches der Lexikographie.

3. Die Theorie der Lexikologie versucht u.a. die Frage zu beantworten, was ein Wortschatz ist. Dabei ergibt sich sofort die Notwendigkeit, drei Auffassungen von 'Wortschatz' voneinander zu unterscheiden:

a. Wortschatz als die mentale Speicherung von Wörtern beim Individuum, also der Wortschatz bezogen auf ein Individuum. Als identifizierende Bezeichnung verwenden wir hierfür " **mentales Lexikon**".

b. Wortschatz als der Grundstock eines beliebigen Ausschnitts einer natürlichen Sprache, also der Wortschatz bezogen auf eine natürliche Sprache. Hierfür wollen wir die vertraute Bezeichnung "**Wortschatz**" weiterhin verwenden. Lexikographisches Arbeiten bezieht sich normalerweise genau auf dieses Verständnis von "Wortschatz".

c. Wortschatz im Verhältnis zur Grammatik, wofür sich die Bezeichnung "**Lexikon**" sinnvollerweise eingebürgert hat.

zu Ziff. 4.325:

Wortschatz

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(16) Schätzungen für den Umfang des individuellen Wortschatzes einer erwachsenen Person zeigen immense Schwankungen: Bei Bußmann (1990) werden 6000 bis 10 000 Wörter angeführt, während Aitchison (1994) von 50 000 bis 250 000 Wörtern spricht und der Umfang des Allgemeinwortschatzes in der Duden-Grammatik mit 300 000 bis 500 000 Wörtern angegeben wird. Nun sind Bußmanns Angaben auf das Deutsche bezogen und Aitchisons Angaben auf das Englische. Selbst wenn man den starken romanischen und germanischen Einfluß beim Englischen berücksichtigt, kann damit diese Diskrepanz nicht erklärt werden. Abgesehen von der Unklarheit, was jeweils als "Wort" gezählt wird, vermute ich einmal, die Wahrheit wird nicht einfach irgendwo in der Mitte liegen, sondern eher in Richtung auf die untere Grenze der Angaben von Aitchison. Ferner darf natürlich auch die Unterscheidung zwischen einem "aktiven", also für die Sprachproduktion zur Verfügung stehenden und einem "passiven", also für die Sprachrezeption zur Verfügung stehenden Wortschatz nicht vergessen werden. Erinnern wir uns an die Situation mit *oheim*, ein Wort, das man, wenn überhaupt, eher dem passiven Wortschatz der jüngeren Generation zuordnen möchte. Ein anderer Bereich, wo wir den Unterschied zwischen aktivem und passivem Wortschatz meist schmerzhaft erfahren müssen, sind Fremdsprachen.

(18) 1. Versprecher, die beim gesunden Sprecher/bei der gesunden Sprecherin auftreten, sind häufig nicht willkürlich. Dies ist eine in der Fehlerlinguistik bekannte Erscheinung. Typisch z.B. ist folgender Versprecher aus einem im Rundfunkstudio vorgelesenen Wetterbericht: *Im Süden noch Frost, sonst mehrere Grade unter Null*, wobei statt *unter* über zu lesen war, oder der Versprecher aus einer Rundfunkreportage eines Fußballspiels: *es mißlingt ihm ein fehlpaß*, bei dem die Rekonstruktion des Intendierten natürlich nicht selbstverständlich ist. Man darf aber vermuten, daß *paß* anstatt *fehlpaß* bei angefangener Konstruktion mit *mißlingen* anvisiert war. Sehen wir die beiden Präpositionen *unter*, *über* und die beiden Substantive *paß*, *fehlpaß* als Elemente an, die auf der Inhaltsebene einen Gegensatz untereinander ausdrücken - im Falle der Präpositionen sprechen wir später von einem Beispiel der Konversen-Relation und im Falle der Substantive sprechen wir später von einem Beispiel der Inkompatibilitäts-Relation, dann würde dies auf eine semantisch orientierte Strukturierung deuten. Interessant ist ferner bei beiden Beispielen, daß die Fehler offensichtlich durch solche vorangehenden Formen (*frost*, *mißlingen*) ausgelöst werden, die mit den Fehlern inhaltlich aufs engste zusammenhängen. Auch dies scheint auf eine semantisch orientierte Strukturierung hinzudeuten.

2. Die Psycholinguistik berichtet seit den 50er Jahren folgendes: Die Qualität der sog. "freien" Wiedergabe von visuell oder lautlich präsentierten Wortlisten von etwa 50 Wörtern steigt, falls die Präsentation nach begrifflichen Kategorien erfolgt. (19) Meist wird in der Wiedergabe der vorgegebenen Ordnung gefolgt. Bei einer willkürlichen Präsentation erfolgen die Antworten nach mehrmaliger Wiederholung trotzdem meist geordnet nach begrifflichen Kategorien. Psychologen/Psychologinnen sprechen hier von der Tendenz zur kategorialen Gruppenbildung...

Neben paradigmatischen Beziehungen, die untereinander alternativ auftretende Elemente involvieren, findet man bei den Wiedergaben zahlreiche Hinweise auf die Möglichkeit syntagmatischer Beziehungen, die miteinander auftretende Elemente involvieren: So werden gerne assoziativ verknüpfte Elemente, wie *beißen* und *zähne*, zusammen produziert. Assoziationen, worauf wir später noch genauer zu sprechen kommen, sind in diesen Kontexten übrigens keineswegs so beliebig, wie Laien sich dies meistens vorstellen. Als Assoziationen treten folgende Möglichkeiten auf: a) Koordination: *messer, gabel* b) Gegensätze: *links, rechts* c) Kollokationen: *beißen, zähne* oder *blond, haar* d) Über-/Unterordnung: *pferd, tier* e) Synonyme: *erhalten, bekommen*. Diese Möglichkeiten können ebenfalls alle mehr oder weniger der Inhaltsebene zugeordnet werden.

(23) Mit der psychologischen Realität der Gegliedertheit des mentalen Lexikons meinen wir, daß die der Sprachfähigkeit zugrundeliegende Organisation im Gehirn Gegliedertheiten der besprochenen Art auf den für die Sprachproduktion oder Sprachrezeption relevanten Repräsentationsebenen zuläßt, wobei diese Gegliedertheiten wohl eher von temporär beschränkter Natur sind.

(38) Die lexikalischen Elemente im Wortschatz bestehen (zumindest) aus den abstrakt gedachten Wörtern und den abstrakt gedachten phraseologischen Einheiten.

Es muß ausdrücklich betont werden, daß nicht nur bei den phraseologischen Einheiten Abgrenzungsschwierigkeiten auftreten, sondern eben auch im Bereich der Wörter. Als Besonderheiten seien erwähnt:

a. **Kurzwörter** wie *AEG, AIDS, ampel, bafög, BMW, bus, DJ, HOMER, jwd, rolly, Saba, spusi, Stasi, taxi, uni. ...*

b. **Internationalismen** wie *bretzel, kimono, poncho, tango*. Hier handelt es sich um in zahlreichen Sprachen in praktisch gleicher Form auftretende Wörter für (39) *Eßwaren, Kleidungsstücke, Tänze* usw., wobei die betroffenen "Sachen" zwar ganz bestimmten Kulturgemeinschaften entstammen, aber weite Verbreitung darüber hinaus gefunden haben.

c. **Unsinnswörter** wie *verdaustig*, die von der Form her im Deutschen durchaus denkbare Wortformen für Wörter darstellen, aber offensichtlich keinen Inhalt aufweisen. Daß sie in einem geeigneten Kontext dennoch interpretierbar sein mögen, demonstriert Lewis Carroll anhand der ersten Verszeile des Gedichtes "Jabberwocky", in der deutschen Übersetzung "Der Zipferlake": *Verdaustig wars, und glasse Wieben...*

d. Als Realisationserscheinungen in einem Medium mögen bestimmte Formen auftreten. So im Lautmedium sog. **phonologische Wörter** wie das schwäbische Element *hof* für die Kette *hast du in der Frage hof* 'n *flysæl* für, *Hast du einen Schlüssel?* oder im Schriftmedium sog. **graphematische Wörter** wie *str.* für *straße*.

zu Ziff. 4.3251:

Mentales Lexikon

(Ausdrucks-)Syntax + Semantik + Pragmatik Wortnachbarschaften + Bedeutungsstruktur + Isotopien

zu: vsem3.122 vprag6.0451

aus: K. BEIßNER, *I see what you mean* - Metaphorische Konzepte in der (fremdsprachlichen) Bedeutungskonstruktion. EHS 14, 391. Frankfurt/M 2002: P. LANG

(126)

The mental lexicon is not a fixed dictionary with a set amount of information about each word, but an active system in which new links are perpetually being formed. (Aitchison, 1994, S.167; vgl. auch Raupach, 1994, S.37)

Das mentale Lexikon ist gemäß einer kognitiv-konstruktivistischen Konzeption und einer damit einhergehenden konnektionistischen Modellierung von Wissenskonstruktionsprozessen nicht, wie der Name suggeriert, als eine Zusammenstellung von Lexemen zu verstehen, in der jedem Eintrag eine bestimmte festgelegte Bedeutung zugeordnet ist. Es handelt sich dabei vielmehr um ein netzwerkartiges Gebilde, in dem sich die Bedeutung der einzelnen Bausteine sowohl aus den für die jeweilige Einheit gespeicherten sprachlichen und kognitiven Informationen als auch aus den zum Teil sehr zahlreichen und komplexen Verknüpfungen zu anderen Einträgen ergeben. ...

(127) Aitchison (1994) demonstriert, daß die Aktivierung bestimmter Lexeme bei der Produktion von Sprache nicht so sehr von logischen Beziehungen innerhalb der Sprache, wie z.B. Synonymie und Antonymie, abhängig ist, als vielmehr von der konventionellen Gebundenheit einzelner Lexeme und Lexemkombinationen an bestimmte semantische Felder, "different words tend to get attached to different semantic fields" (S.91). Die Autorin demonstriert dies sehr anschaulich am Beispiel der möglichen Antonyme des Adjektivs *short* (S.96)

A *short* life versus a *long* life

A *short* man versus a *tall* man

A *short* temper versus a *good* temper

In *short* supply versus a *plentiful* supply

und der Verwendung der Synonyme *broad* und *wide*, deren Wahl nicht konzeptuell sondern konventionell bedingt ist:

The choice of synonym is primarily dictated by the surrounding context. *Wide* may be more closely linked to *road*, and *broad* to *mind*, than *wide* and *broad* to each other. Similarly, *gallop* may be more closely attached to *horses*, and *sprint* to *human*, than *gallop* and *sprint* to each other. (Aitchison, 1994, S.97; vgl. auch Levelt, 1991, S.8-9)

(128) ... die von der Autorin hervorgehobene Erkenntnis, daß es zwischen dem semantisch-syntaktischen Modul, in dem die Bedeutungen einzelner Lexeme gespeichert sind, und den allgemeinen kognitiven Wissensstrukturen eines Individuums keine klare Trennung gibt; "it is therefore impossible to say where the "meaning" of a word ends and general knowledge begins" (S.226). Erst durch diese Verbindung von Sprache und Weltwissen ist es möglich, die unterschiedliche Bedeutung von Lexemen in unterschiedlichen Kontexten zu erschließen bzw. eine adäquate

Wortwahl zur Beschreibung einer bestimmten Situation oder zum Ausdrücken einer Intention vorzunehmen. Denn welche Teilbedeutung eines Lexems und damit verbundene Konzeptstrukturen in einer konkret wahrgenommenen Situation als relevant empfunden und daher aus dem mentalen Lexikon abgerufen werden, hängt vom jeweiligen Kontext ab, also davon, welche anderen kognitiven Strukturen bei dem wahrnehmenden Individuum aktiviert werden und somit den semantischen Rahmen (*frame*) für die Produktion und Rezeption der jeweiligen Äußerungen bilden; "the mental lexicon is therefore concerned above all with links, not locations" (S.230).

(140) Eine Vielzahl stabiler Lexemkombinationen wie Idiome, Kollokationen und die von Nattinger und DeCarrico (1992) als "lexical phrases" bezeichneten Lexemkombinationen liefern zudem konventionalisierte Fertigteile bzw. Rahmenstrukturen, die sowohl die Produktion als auch die Rezeption von Sprache erleichtern, "permitting speakers (and hearers) to direct their attention to the larger structure of the discourse, rather than keeping it focused narrowly on individual words as they are produced" (Nattinger & DeCarrico, 1992, S.32; vgl. auch Fernando, 1996, S.215f). ...

(141) Das soeben Dargelegte stimmt mit den in den vorangegangenen Kapiteln herausgestellten Erkenntnissen überein, daß fremdsprachliche Lexeme und deren Bedeutung nicht in Isolation, sondern in möglichst vielen unterschiedlichen Kollokationen, d.h. als Teil unterschiedlicher *multi-word chunks* gelernt werden sollten.

zu Ziff. 4.326:

Wortkern + Affixe / Suffixe

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(99) Beispiel einer **Wortfamilie** ist die Klasse {*rüsten, aufrüsten, entrüsten, wettrüsten, rüstung, gerüst, rüstig,...*}. Da Kompositionen aus Elementen der Wortfamilie mit anderen Elementen nicht ausgeschlossen sind, sind die Wortfamilien in ihrem Umfang nicht so leicht überschaubar. So würden wir bei unserem Beispiel als weitere Elemente auch Linkserweiterungen wie *ritterrüstung* und Rechtserweiterungen wie *rüstzeug, rüstungswettlauf, entrüstungssturm* finden. Ferner ergibt sich dadurch die Möglichkeit für einzelne Wörter, an mehreren Wortfamilien beteiligt zu sein. So würde etwa *rüstungswettlauf* zumindest auch an der Wortfamilie beteiligt sein, zu der *laufen* gehört.

Zur Bezeichnung der jeweiligen Wortfamilie wird ein elementarer Repräsentant aus der Klasse - im Beispiel: *rüsten* - oder ein morphologisches Grundelement - im Beispiel: *rüst-* - gewählt.

Eine Wortfamilie ist eine Klasse von Wörtern, die allesamt mittels morphologischer Prinzipien aus einem Grundelement herleitbar sind.

(100) Wortfamilien sollen ihrem Namen nach aus Wörtern bestehen, flektierte Formen des jeweiligen Wortparadigmas werden auf dieser Ebene nicht erfaßt. Die nach unserer Festlegung entscheidenden Kriterien betreffen die Formebene von Wörtern und greifen offensichtlich über Wortartgrenzen hinweg, finden wir doch in unserem Beispiel Elemente der Wortarten "Verb", "Substantiv" und "Adjektiv". Diese Flexibilität ist bei einem allgemeinen Forminteresse sicher von Vorteil. Erwartungen über inhaltliche Zusammenhänge müssen aber schon allein aus diesem Grunde gedämpft werden. Immerhin könnte man sich denken, daß die jeweiligen Elemente wenigstens sich die auf der Inhaltsebene angesprochenen Domänen teilen. Dem ist jedoch nicht so, wie wir bei unserem Beispiel rasch erkennen. Das Adjektiv *rüstig* dient als Charakterisierung älterer Menschen, ist also auf die Domäne "Personen" bezogen. *gerüst* im Sinne von "Baugerüst" zählt zu den konkreten Kollektiva und erlaubt zahlreiche metaphorische Erweiterungen in andere Bereiche. Die Domäne "Militär" faßt wenigstens die Elemente *rüstung, rüsten, aufrüsten* zusammen, wobei einige der Elemente dennoch Lesarten in anderen, nicht unbedingt geteilten Domänen aufweisen. Bei *entrüsten* ist eigentlich erst über den Slogan der Friedensbewegung "Entrüstet Euch!" wieder eine Anbindung an die Domäne "Militär" erfolgt. Die auf der Formebene angesiedelten morphologischen Zügel sind nicht straff genug, um über einen längeren Zeitraum hinweg eine inhaltliche Klammer zu garantieren.

zu Ziff. 4.3261:

Kasus - Türkisch: schön unterschieden

zu:

aus: B. BLAKE, 102. Case, S. 1073-1090 in: BOOIJ, G; LEHMANN, C; MUGDAN, J ; SKOPETEAS, S, Morphologie/Morphology: Ein internationales Handbuch zur Flexion und Wortbildung. 2. Halbband. Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science. 17/2. Berlin 2004: de Gruyter.

(1073) **Case** in its most central manifestation is an inflectional system of marking noun phrases for their relationship as dependents to their heads. Typically case marks the relationship of a noun phrase to a verb as in the following example from Turkish.

(1) Mehmet	adam-a	elma-lar-ı
Mehmet(NOM)	man-DAT	apple-PL-ACC

ver-di
give-PAST.3.SG

"Mehmet gave the apples to the man."

Mehmet is in the nominative case which indicates the subject. *Adam* is in the dative case which indicates that it is the recipient of *vermek* "give". *Elmalar* is in the accusative which indicates that it is the direct object. The accusative also indicates that it is definite or at least specific, since in Turkish only specific direct objects are marked by the accusative. Non-specific direct objects are put in the nominative.

Turkish has a system of six cases as in Tab. 102.1. The genitive marks an adnominal, dependent noun phrase as in *adam-ın ev-i* "man-GEN house-3.SG.POSS (of the man his house)", i.e. "the man's house". In Turkish the possessor, if definite, is cross-referenced by a possessive pronominal suffix on the possessed noun, hence the *-i*. The locative marks location as in *Istanbul-da* "in Istanbul", and the ablative indicates "from" or "out of" as in *Ankara-dan* "from Ankara".

nominative	<i>adam</i>
accusative	<i>adamı</i>
genitive	<i>adamın</i>
dative	<i>adama</i>
locative	<i>adamda</i>
ablative	<i>adamdan</i>

Tabl 102.1: Turkish case system

... (1074) We need to make a distinction between **cases**, (of which there are six in a system of oppositions) **case markers** or **case forms** (the affixes or word forms respectively that realise the oppositions) and **case relations** (the syntactic (!) or semantic relations that the cases express). Case relations, also referred to by other labels such as grammatical relations or functions, need not be in a one-for-one correspondence with cases.

zu Ziff. 4.327:

Linearität / (Ausdrucks-) SYNTAX

zu: vsyst4.12 vsem10.001

aus: RIEGER, B, Wissenserwerb aus natürlichsprachlichen Texten. Zur Interdisziplinarität eines computerlinguistischen Forschungsprogramms, in: KÜHLWEIN, W, RAASCH, A (eds.), Angewandte Linguistik heute. *forum* Angewandte Linguistik Bd. 20. 1990. S.101-124.

(112) Für die meisten Anwendungszusammenhänge ist die *Linearität* der textuellen Darbietungsweise dessen, was wir als *Bedeutung* oder *Inhalt* von sprachlichen Ausdrücken zu bezeichnen uns angewöhnt haben, ein nicht nur ausreichendes Darstellungsprinzip, sondern in seiner linguistischen Funktion sogar eine Bedingung semiotischer Struktur- und Systembildung. Ohne den Zwang zur Linearisierung ist etwa die Konstitution *syntagmatischer* und *paradigmatischer* Relationen zwischen sprachlichen Elementen nicht denkbar, die schon F. DE SAUSSURE als eine strukturbildend-funktionelle Beschreibung sehr fundamentaler Ordnungsprinzipien der kommunikativen Verwendung und Aggregation von Zeichen unterschied. Darüber hinaus stellt Linearisierung aber auch ein gewissermaßen erzieherisches Prinzip dar bei der "allmählichen Verfertigung der Gedanken beim Sprechen" und/oder Schreiben zur Überführung vieldimensionaler, konzeptueller Zusammenhänge in Texte, wie ebenso auch umgekehrt beim allmählichen Aufbau und Vollzug gedanklicher und konzeptueller Zusammenhänge aus der Linearität sprachlicher Texte. Diese Prozesse, bei deren Analyse und Rekonstruktion sich das strukturbildende Moment der Linearisierung wie eine Art Leitlinie nutzen lässt zur Modellierung von Bedeutung und Wissen in Texten, können übergreifend - sowohl als *Verstehen* in der Anknüpfung an schon vorhandene wie auch als *Lernen* im Zustandekommen neuer Zusammenhänge - als *Bedeutungskonstitution* bezeichnet werden, deren (simulativ/kreative) Nachbildung im Computer heute die zentrale Herausforderung der einschlägigen Forschung bildet.

zu Ziff. 4.328:

Morphologie: Sonderrolle von Fremdwörtern (und Onomatopoetika, Eigennamen)

zu:

aus: H. WEGENER, *Pizzas* und *Pizzen* - die Pluralformen (un)assimilierter Fremdwörter im Deutschen: Zeitschrift für Sprachwissenschaft 23 (2004) 47-112.

(50)

(1) zweistufige Entwicklung

fremder Plural	s-Plural	assimilierter Plural
a. Themata Conti Mensae Pizze	Themas Kontos Mensas Pizzas	Themen Konten Mensen Pizzen
b. (généraux) Couches	Generals Couchs	Generäle Couchen

(2) einstufige Entwicklung

fremder Plural	s-Plural	assimilierter Plural
	a. Lifts Fracks Balkons b. Cakes/Keks Straps Pumps c. Saunas Datschas Albums	Lifte Fräcke Balkone Kekse Strapse Pumpse Saunen Datschen Alben

(3) direkte Integration

fremder Plural	s-Plural	assimilierter
(boxes) (bosses) (chances) (gangsters)	- - - -	Boxen Bosse Chancen Gangster

(51) Die zweistufige Entwicklung der Klassizismen und Italianismen in (1) und (2c) wirft die Frage auf, warum z.B. für das Fremdwort *Pizza*, für das ... zunächst der fremde, italienische Plural *Pizze* (mit finale[m] [e]) und dann "eindeutschend" *Pizzas* gebraucht wurde, nicht direkt die phonetisch näherliegende Form *Pizzen* gebildet wurde, warum das Deutsche also vor der Bildnerung der assimilierten Form die Übergangsform *Pizzas* ausbildet - und warum es diese wieder aufgibt. Warum verläuft die Assimilation von *Pizze* zu *Pizzen* nicht direkt von -e zu -en, was phonologisch nahe liegt, sondern von -e über -as zu -en, also -e > -en > -en, warum dieser Umweg über die s-Form?

(106) Fremdwörter sind weder lautmalerische, ikonische Gebilde noch haben sie einen besonderen semiotischen Status, aber Fremdwörter sind zu Beginn ihrer Existenz im fremden Sprachraum noch nicht allgemein bekannt. Eine Pluralform, die die Laut-, Fuß- oder Silbenstruktur des fremden Wortes verändern würde, könnte die Verbreitung des neuen Wortes gefährden: *Pizzen* lässt die Basis *Pizza* schwerer erkennen als *Pizzas* ... Für fremde Wörter ist es wichtig, zunächst einmal in der Zielsprache etabliert zu werden. Strukturbewahrende Pluralformen stellen daher einen Ausgleich für mangelnde Etabliertheit dar. Das Korrespondenzgebot erklärt, warum "eindeutschend" ... an die Stelle der italienischen substitutiven Pluralform *Pizze* zunächst die "konservative" Pluralform *Pizzas* tritt. Erst wenn das Fremdwort einen gewissen Bekanntheitsgrad erreicht hat, erlaubt es die Ausbildung assimilierter Flexionsformen. Es ist also funktional, Fremdwörter (zunächst) morphologisch gesondert, "schonend" zu behandeln, genau so wie Onomatopoeika und Eigennamen.

zu Ziff. 4.329:

Materialität der Medien

zu:

aus: J. G. SCHNEIDER, Medialität und Sozialität sprachlicher Zeichen: Zeitschrift für Semiotik 30/1-2 (2008) 3-10.

(4) Will man diese Materialität und ihre "sinnmitemzeugende" Funktion systematisch untersuchen, so sieht man sich allerdings ... von vornherein mit einem Problem konfrontiert. Die Annahme, Medien seien mehr als bloße Hilfsmittel, wird nämlich durch unsere Alltagserfahrung zunächst keineswegs gestützt:

"Wir hören nicht Luftschwingungen, sondern den Klang der Glocke; wir lesen nicht Buchstaben, sondern eine Geschichte; wir tauschen im Gespräch nicht Laute aus, sondern Meinungen und Überzeugungen, und der Kinofilm lässt gewöhnlich die Projektionsfläche vergessen" (Krämer 1998: 74).

Das Medium gerät a l s s o l c h e s gar nicht in den Blick, es entfaltet seinen Einfluss gewissermaßen im Verborgenen. Zum Beispiel wirken sich die Lautlichkeit und "Fluktuanz" der gesprochenen Sprache auf die mündlichen Performanzen aus, die Visualität und Starrheit der geschriebenen Sprache auf die schriftlichen: Die gesprochene Sprache weist eine deutlich andere Syntax auf die geschriebene. Typisch mündliche Strukturen wie "Die CD da, kannst du mir die mal leihen?" oder "weil das widerspricht sich ja" werden in den letzten Jahren auch von Grammatikern zunehmend einbezogen und reflektiert. Medien hinterlassen, in der Regel unbemerkt, ihre "Spuren" an der jeweiligen "Botschaft" (vgl. Krämer 1998 und Stetter 2005: Kap. 2). Gleichzeitig sind sie notwendige Bedingung für Humankommunikation: Geistige Inhalte müssen sich - so ein Grundgedanke der kulturwissenschaftlichen Medientheorie - immer in Medien manifestieren; es gibt keine medienunabhängigen Inhalte. In diesem Sinne ist ein Medium deutlich mehr als ein Transportmittel: Es gibt dem jeweils mediatisierten Inhalt seine spezifische Form (vgl. Stetter 2005: 79f).

4.33 Automatische Analyse

Literatur

XANTHOS, A: Apprentissage automatique de la morphologie. Le cas des structures racine-schéme. Sciences pour la communication 88. Bern 2008 978-3-03911-756-7br: Peter Lang // I.2.7

4.331 Überlegungen zum Algorithmus

Literatur

KLEIN, H: Computer gestützte Inhaltsanalyse mit INTEXT. Aktuelle Medien- und Kommunikationsforschung 8. Münster 1996 // I.2.7

MENGAL P: Statistique descriptive. Exploration. Recherches en sciences de l'éducation 26. Bern 2004 ISBN 3-03910-237-0br: Peter Lang // I.2.5

SAMUELSSON, C: Statistical Methods in: MITKOV, R (ed./eds.): The Oxford Handbook of Computational Linguistics. Oxford Handbooks Series. Oxford 2003 358-375.: Oxford University Press // I.2.4

SCHMIDT-THIEME, L: Assoziationsregel-Algorithmen für Daten mit komplexen Strukturen. Mit Anwendungen im Web Mining. Informationstechnologie und Ökonomie 24. Frankfurt / M 2003 ISBN 3-631-52213-4br: Peter Lang // I.2.5

WAGNER, R: Korpusbasierte Konstruktion unscharfer Einheiten. Revision und Modifikation Harris'scher Verfahren in: HITZENBERGER, L (ed./eds.): Angewandte Computerlinguistik: Vorträge im Rahmen der Jahrestagung 1995 der Gesellschaft für Linguistische Datenverarbeitung (GLDV) e.V., Regensburg, 30.-31.3.1995. Sprache und Computer 15. Hildesheim 1995 154-158.: Olms // I.2.7

zu Ziff. 4.3311:

Übergangswahrscheinlichkeiten

zu:

aus: R. WAGNER, Korpusbasierte Konstruktion unscharfer Einheiten (157f)

Paralleler Übergangsansatz: Bei diesem Verfahren wird jedem Buchstabenzwischenraum das zugehörige Bigramm zugeordnet. Silbengrenzen werden an den Minima dieser Werte gesetzt.

Beispiel: Im Wort "Kalender" finden sich folgende Übergangshäufigkeiten:

K --->a --->l --->e --->n --->d --->e --->r
0.14 0.10 0.17 0.23 0.12 0.45 0.23

Daraus ergibt sich die Segmentierung "Ka-len-der".

Unterschiede zu klassischen Silben sind hauptsächlich in der Nähe von Doppelvokalen ("ei" oder "aa") und "sch" und "ch" zu finden. Möglicherweise sollten andere Ausgangseinheiten als Grapheme gewählt werden.

Die wichtigsten Unterschiede des Übergangsansatzes zum Harris'schen Verfahren sind:

- I. Statt der Anzahl der potentiellen Nachfolger wird die Übergangshäufigkeit zum tatsächlichen Nachfolger benutzt. Statt eines kombinatorischen wird also ein numerisches Verfahren benutzt. Dabei ist zu beachten, daß diese beiden Werte zusammenhängen: Sind die Häufigkeiten der Nachfolger gleichverteilt, ist die Übergangshäufigkeit zur Anzahl der Nachfolger reziprok.
- II. Statt der gesamten vorangehenden Zeichen werden nur die beiden direkt vorangehenden Zeichen berücksichtigt. Es wird also eine Art Markov-Annahme getroffen. Dadurch wird das Verfahren wesentlich effizienter.

Aufgrund der beobachteten Phänomene läßt sich grob die Hypothese aufstellen, daß sich

Silben durch Bigramme und
Morpheme durch Ngramme

ermitteln lassen.

zu Ziff. 4.3312:

Parsing-Definition

zu: vsem10.22

aus: LANGER, H: Parsing-Experimente. Praxisorientierte Untersuchungen zur automatischen Analyse des Deutschen. Frankfurt/M 2001. Sprache, Sprechen und Computer 4

(35) Der Ausdruck "Parsing" leitet sich aus den *partes orationis*, den Teilen der Rede, her. Im weitesten Sinne umfaßt Parsing alle Formen der (automatischen) Analyse sprachlicher Ausdrücke, aber gerade für diesen ursprünglichen Sinn (grammatische Analyse als Feststellung der Wortarten) hat sich in den vergangenen Jahren der Begriff "Tagging" (Wortarten-Tagging, "Part-of-speech tagging") etabliert; unter Parsing versteht man heute eher solche Analyseprozesse, die substantiell über das bloße Annotieren eines Textes mit Wortarten hinausgehen und die die grammatische "Struktur" einer Äußerung aufdecken. Die Grenzen sind jedoch fließend, so wird der Begriff "Tagging" z.B. auch für die Annotation von Texten mit syntaktischen Funktionen wie "Subjekt" und "Objekt" verwendet (Brants et al. 1997).

Der Standardanwendungsbereich von Parsing-Methoden in der Computerlinguistik ist die (Satz-)Syntax, es gibt aber auch zahlreiche Anwendungen in anderen Bereichen, z.B. prosodisches und phonologisches Parsing, morphologisches Parsing, semantisches Parsing usw.

Ebenso vielfältig wie die Phänomenbereiche, in denen Parsingmethoden zum Einsatz kommen, sind die Grammatikformalismen, für die Parsing-Algorithmen neu entwickelt oder angepaßt wurden. Im Zentrum stehen jedoch nach wie vor und immer wieder Parsingalgorithmen für kontextfreie Grammatiken, die allerdings praktisch nie in reiner Form, sondern fast immer in Verbindung mit zusätzlichen Mechanismen (z.B. Merkmalsstrukturen oder probabilistischen Bewertungen) verwendet werden. Auch bei Formalismen, bei denen das kontextfreie Skelett relativ schwach ausgeprägt ist, z.B. in HPSG-orientierten Grammatiken, wird für Zwecke des Parsings zumeist auf Grundalgorithmen für kontextfreie Syntaxen zurückgegriffen.

zu Ziff. 4.3313:

Wortkombinationen

zu:

aus: H. KLEIN, Computerunterstützte Inhaltsanalyse mit INTEXT. Münster 1996.//I.2.7

(58ff) *Wortkombinationen* sind die Möglichkeit, nicht nur eine Zeichenkette in der Zeile einer Wörterliste auszugeben, sondern immer mehrere, die in einer Texteinheit direkt hintereinander stehen. Die Anzahl der Zeichenketten ist frei wählbar ...

Wortpermutationen bestehen grundsätzlich immer nur aus zwei Zeichenketten. Dazu wird in jeder Texteinheit die erste Zeichenkette mit jeder folgenden kombiniert, danach die zweite Zeichenkette mit jeder folgenden, und das bis zur vorletzten Zeichenkette der Texteinheit. Besonders bei der Konstruktion von Wortstammfolgen, mit denen mehrere Zeichenketten mit definierbarem Abstand und Reihenfolge innerhalb einer Texteinheit gesucht werden können, ist das sehr nützlich...

Wortkombinationen und -permutationen sind zur Überwindung der Grenzen der kontextfreien Einwortanalyse geeignet, da sie den Kontext berücksichtigen und so Mehrwortanalysen ermöglichen.

Beispiele für Wortkombinationen

attraktiv,spontan	auf alles	aus, mittelblond
attraktiv, sucht	auf bloße Oberfläche	aus 20jähriger Ehe
attraktiv so sagt	auf deine Zuschrift	aus dem Raum
attraktive, etwas	auf dem Besen	aus dem Weg
...	auf den Arm	aus guter Familie
auch nur	auf der Erde	aus Parität gesicherten
auch albern	auf der Suche	ausgeglichenes Wesen
auch anders	auf der Welt	ausgeschlossen. Chiffre
...	...	ausgestattet, wünscht

Beispiele für Wortpermutationen

ab alles	Abend am	aber
ab Alltag	Abend auf	aber Abend
ab anderswo	Abend im	aber aber
ab Anmutig	Abend Kaminfeuer	aber allein
ab auch	Abend Sommer	aber als
ab ausgeprägter	Abend Terrasse	aber alten
ab beheimatende	Abend unserer	aber am
ab bemüht	Abend Winter	aber an
...	...	aber anderen

4.332 Eigener Ansatz (Rekursion)

zu Ziff. 4.3321: vsyst4.13

Paradigmatisch-distributionelle Morph-Analyse

ILLUSTRATION

- Stand: 1989 -- Stand: 1989 -- Stand: 1989 -- Stand: 1989 -- Stan

PARADIGMA N.

Ebene der Einzel-Zeichenformen (=Alphabet)

Morph-Atom (M)				
P1.1	∅=	//	//	+∅ A1.1
P	=			+

PARADIGMA I.

Ebene der Einzel-Zeichenformen (=Alphabet)

Morph-Atom (M1)					
Präfix		//	//		Affix
P1.1	∅=			+∅	A1.1
P2.1	e=			+s	A2.1
P3.2	w==			+e	A3.1
P4.3	sch==			++s	A2.2
P5.2	s==			++m	A4.2
P6.2	d==			++r	A5.2
P7.2	r==			++n	A6.2
P8.2	f==			+n	A6.1
P9.2	b==				
P10.2	k==				
P10.1	k=				

PARADIGMA II.

Ebene der Einzel-Zeichenformen (+Alphabet)

Morph-Atom (M2)				
P1.1	∅=	//	//	+∅ A1.1
P11.1	h=			+l A7.1
P12.1	m=			++l A7.2
P10.1	k=			+++s A2.3

P8.1	z=	++t	A8.2
P9.1	b=		

PARADIGMA III.

Ebene der Einzel-Zeichenformen (+Alphabet)

Morph-Atom (M3)				
		//	//	
	∅=			+∅ A1.1
P3.1	w=			+t A8.1
P4.1	sch=			+d A9.1
P4.2	sch==			++e A3.2
P13.2	l==			

Beispiele:

ein	/P2.1/M1/A1.1/
eine	/P2.1/M1/A3.1/
eines	/P2.1/M1/A3.1/A2.2/
wein	/P3.2/P2.1/M1/A1.1/
schweinen	/P4.3/P3.2/P2.1/M1/A3.1/A6.2/
keiner	/P10.2/P2.1/M1/A3.1/A5.2/
kind	/P10.1/M1/A9.1/
schwind	/P4.2/P3.1/M1/A9.1/
rein	/P7.2/P2.1/M1/A1.1/
alt	/P1.1/M2/A8.1/
halt	/P11.1/M2/A8.1/
halle	/P11.1/M2/A7.1/A3.2/
bald	/P9.1/M2/A9.1/
wert	/P3.1/M3/A8.1/
schwer	/P4.2/P3.1/M3/A1.1/
leer	/P13.2/P2.1/M3/A1.1/
leine	/P13.2/P2.1/M1/A3.1/
kert	/P10.1/M3/A8.1/

Definitionen: -"Wortform" ist eine grafische Einheit von Buchstaben, die zwischen anerkannten Trennzeichen steht (Blanks, Satzzeichen, Klammerungen).

- "Morph-Atom" ist die Verbindung von mindestens 2 Buchstaben, aufgrund von Distributionsanalyse erarbeitet. Die Zahl der an sich möglichen Buchstabenkombinationen und die Zahl der in der betr. Einzelsprache tatsächlich gebräuchlichen Buchstabenkombinationen ergibt eine Differenz: Die nicht-genutzten Kombinationen werden oft in anderen Sprachen genutzt, so daß sich von hier her die ausdrucksformale Differenz der Einzelsprachen erläutern läßt.

- "Buchstaben" sind die Elemente des Ausdrucksrepertoires, aus denen die Ausdrücke der jeweiligen Sprache konstruiert werden. In den alphabetischen Schriften ist - mit Unschärfen - dieses Grundinventar an Basiselementen des Ausdrucks gleich. Der Grund, daß die Schriften/Sprachen doch gerade auf Ausdrucksebene verschieden sind, ist im Bereich "Morph-Atom" zu suchen.

Wortform \rightarrow Morph-Atom₁...Morph-Atom_n

Morph-Atom \rightarrow Buchstabe₁...Buchstabe_n

Zugang von automatisierter Distributionsanalyse her:

Ziel: Ein vorgegebenes Textkorpus wird durch differenzierte Distributionsanalyse automatisch immer mehr auf die ausdrucksformalen Basiselemente reduziert, die insgesamt die Variablen darstellen, mit denen letztlich der gesamte Text erstellt ist: eine sehr begrenzte Zahl von Variablen wird durch Häufigkeit und unterschiedliche Kombinatorik zu vielfältiger Verwendbarkeit gebracht. Kleinste hierbei zu erzielende Einheiten sind nicht nur Wortformen (ob zusammengesetzt oder nicht) sondern auch Bestandteile von Wortformen (z.B. regelhafte Prä-, In-, Affixe), so daß auch auf diesem Weg zumindest eine große Nähe zu den "Morph-Atomen" erreicht werden kann.

- "AUTOSYNT" reduziert - nach der Definition der in der betr. Sprache geltenden Trennzeichen die im vorgegebenen Textkorpus begegnenden Wortformen per REGISTER-Programm auf je einmaliges Vorkommen. - Dann wird jede Wortform zusätzlich danach befragt, ob sie mit PRÄFIX, mit AFFIX oder mit PRÄ- und AFFIX begegnet. Zusammen mit den EINZELLEXEMEN liegen somit 4 Befunde vor.

- AUTOIII geht von der reinen (und u.U. willkürlichen) Kombinatorik weg und betrachtet die EINZELLEXEME als in der betr. Sprache legitimierte Ausdruckselemente. Nach dem SORTIEREN wird gefragt, welche einzelne Wortform in dieser Liste mit Affix begegnet. Anschließend: ist dieses Affix in der gleichen Liste ebenfalls als Einzelwortform belegt (z.B. |Baum|, |Baumhaus| = |Baum + haus|, in der Liste ist auch |Haus| belegt)? Wenn ja, dann gilt die zusammengesetzte Wortform als zureichend erklärt und wird aus den weiterhin erklärungsbedürftigen Befunden ausgeschieden.

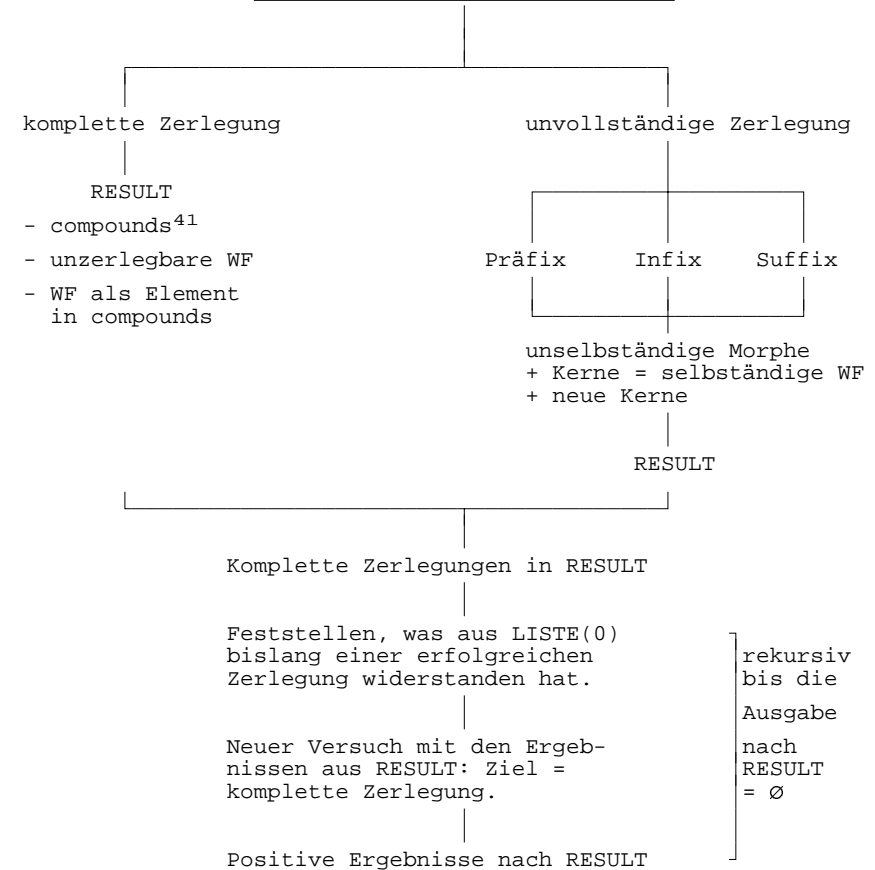
- In der Datei THES sind als in der betr. Sprache legitimierte Befunde abgespeichert: die Liste der Einzel-Wortformen |xxx|, die Liste der Wortformen, die um ein Affix erweitert sein können |xxx+|, die Wortformen, die als Affix fungieren können |+xxx|. In der Datei UNKLAR sind die noch zu erklärenden Befunde abgespeichert.

zu Ziff. 4.3322:

Morphologie - ausdrucksyntaktisch

zu: vsyst4.15 vgt5.221

Liste von Wortformen - LISTE(0)



⁴¹ Juristensprache als Übungsfeld: "Verkehrsinfrastrukturfinanzierungsgesellschaftsgesetz", oder: "Vermögenszuordnungszuständigkeitsübertragungsverordnung".

zu Ziff. 4.3323:

Morphologie: Neuansatz

zu: vsyst4.14

aus: H. SCHWEIZER, Internes Papier zur Morphologie. (reduzierte Anmerkungen)

(29) Denksportaufgabe

Wer entwickelt einen Algorithmus, der beim Input: *Leserlebnis* bei gegebenem Korpus folgenden Output liefert:

```

/les+
  +e+
    +er+
      +leb+
        +nis/

```

Weiterhin wären - als Etappen auf dem Weg zur vollständigen Zerlegung - */lese+* und *+erlebnis/* zusätzlich als Ergebnis festgehalten.

Das wäre in etwa ein Ergebnis, wie es von unserer Ausdruckssyntax angezielt wird.

(30) Weiterarbeit auf der Basis des erweiterten Beispielkorpus

(a) *be / antrag / en* - die Häufigkeit von *be* gegenüber *bea* als Präfix ist signifikant hoch; ebenso die von *en* gegenüber *gen* als Suffix.

(b) den Rest */antrag/* gibt es als selbständige Wortform.

(c) dieser Rest kann aber seinerseits nochmals analysiert werden: *an / trag*.

Dieses Vorgehen entspricht einer Kombination beider bisher getrennten Vorgehensweisen, der Kombination verschiedenartiger Plausibilitäten: (i) es wird nach Stellen gesucht, die rechnerisch auffallen, d.h. bevorzugt für eine Nahtstelle in Frage kommen; (ii) die hierbei erzielten Ergebnisse werden überprüft, indem - wie oben - getestet wird, ob der herausgearbeitete Kern tatsächlich in dieser Einzelsprache belegt ist.

...

(35) Zwischenreflexion: Man kann unsere Fragestellung systemtheoretisch betrachten. Das soll anhand einiger wesentlicher Merkmale, die ein System kennzeichnen, erläutert werden:

- Nimmt man die auf der Basis des Alphabets (unter Einbeziehung der Sonderzeichen, die in verschiedenen europäischen Schriften gebräuchlich sind)⁴² möglichen

⁴² Literaturtip: vgl. die Erzählung: Die Bibliothek von Babel, in: BORGES, JL, Die Bibliothek von Babel. Erzählungen. Stuttgart 1974: RECLAM, S.47-57. Auf der Basis von 25 orthographischen Symbolen sind in jener Bibliothek ungeheuer viele verschiedene Bücher gesammelt. Angesichts eines Buches, das ein reines Buchstabenlabyrinth bietet, aber am Schluß die Zeile: *O Zeit, deine Pyramiden:* "Man ersieht hieraus: auf eine einzige verständliche Zeile oder eine richtige Bemerkung entfallen Meilen sinnloser Kakophonien, sprachlichen Kauderwelschs, zu-

Kombinationen (zudem bei variabler Wortlänge), so repräsentiert diese Menge sowohl die in all diesen Sprachen/Schriften tatsächlich gebräuchlichen Wortformen, als auch - darüberhinaus - Wortformen, die statistisch möglich wären, aber faktisch nirgend im Gebrauch sind. Auf der Basis reiner Statistik/Quantifizierung ist es also nie möglich, die Ebene der Einzelsprache zu erreichen.

- Das Ausdrucksrepertoire der Einzelsprache entsteht dadurch, daß sie (d.h. die Gemeinschaft ihrer Sprecher) aus der Gesamtmenge der an sich möglichen Ausdruckskombinationen einen spezifischen Teil seligiert, aktiv benutzt, den großen Rest aber unberücksichtigt läßt. Dadurch gewinnt die Einzelsprache Individualität, hebt sich von anderen Einzelsprachen, denen eine andere Selektion zugrundeliegt, ab. Zwar gibt es Schnittmengen (produktive Wortformen, die in mehreren Sprachen gebräuchlich sind).

- Diese Selektion einer Einzelsprache stellt zugleich das Ausdrucksrepertoire dar, das in der betreffenden Einzelsprache akzeptiert ist, was sich so auswirkt: ein Sprachbenutzer kann davon ausgehen, daß er bei Verwendung dieser Ausdrücke verstanden wird.

- Ein Text seligiert im Rahmen dieser Einzelsprache nochmals, da er nie alle an sich in der Einzelsprache möglichen Ausdruckskombinationen verwendet. Dadurch gewinnt auf Ausdrucksebene jeder Einzeltext seine Unverwechselbarkeit. Sie beruht auf zwei Faktoren: (a) Auf der Auswahl der Ausdruckskombinationen; (b) auf deren spezifischer linearen Anordnung - was entsprechende Wiederholungen der seligierten Ausdrücke einschließt.

- Die Elemente, die hierbei im Spiel sind, sind homogen: Es sind Kombinationen unterschiedlicher Dimension auf der Basis einer gegebenen Grundmenge (Alphabet). - Typen von Elementen im Einzeltext - rekursiv dargestellt: *Kette von Wortformen* → *selbständige Wortform* → *Kern+Klitikon/a* *oder* *selbständige Wortform+selbständige Wortform* *oder* *selbständige Wortform+Klitikon+selbständige Wortform* *oder* ...

- Die Ausdruckselemente z.B. in einem Text erfüllen ihre Funktion, den Rezipienten in seinem Verhalten zu steuern, ihn zu veranlassen, unter Einbeziehung seiner Welt-, Situations- und Sprachkenntnis eine bestimmte Bedeutung zu konstruieren, nur, wenn sie unverwechselbar sich voneinander unterscheiden. Die

sammenhanglosen Zeugs." (50) Etwas später: "Diese Beispiele setzten einen genialen Bibliothekar instand, das Fundamentalgesetz der Bibliothek zu entdecken. Und zwar stellte dieser Denker fest, daß sämtliche Bücher, wie verschieden sie auch sein mögen, aus den gleichen Elementen bestehen: dem Raum, dem Punkt, dem Komma, den zweiundzwanzig Lettern des Alphabets. Auch führte er einen Umstand an, den alle Reisenden bestätigt haben: *In der ungeheuer weiträumigen Bibliothek gibt es nicht zwei identische Bücher.* Aus diesen unwiderleglichen Prämissen folgerte er, daß die Bibliothek total ist und daß ihre Regale alle irgend möglichen Kombinationen der zwanzig und soviel orthographischen Zeichen (deren Zahl, wenn auch außerordentlich groß, nicht unendlich ist) verzeichnen, mithin alles, was sich irgend ausdrücken läßt: in sämtlichen Sprachen." (51)

Ausdruckselemente - läßt man mal Wiederholungen beiseite - sind durch Differenz zueinander charakterisiert. Diese ist natürlich nur durch Bezug aufeinander, Vergleich miteinander erkennbar. Der Ausdruck *Buch* erbringt seine Orientierungsleistung, seine Verwendbarkeit in der Kommunikation dadurch, daß er - und alle Sprachteilnehmer in dieser Einzelsprache wissen es - konstant nicht mit *Bach* ausgetauscht werden kann, auch nicht mit *Bub*, *Bus* u.ä. - Darin zeigt sich die "Selbstreferentialität" des Systems: Die Elemente bekommen ihre Individualität durch Differenz von weiteren Elementen desselben Systems.

- Anders ausgedrückt: bezüglich des Ausdrucksrepertoires einer Einzelsprache gibt es keinen Fixpunkt, archimedischen Punkt, von dem aus man die Analyse starten müßte bzw. wo die Analyse nur dann erfolgversprechend verläuft, wenn man von diesem Punkt aus startet. Da alles mit allem in der beschriebenen Weise zusammenhängt, kann, ja *muß* man einen beliebigen Ausgangspunkt wählen. Unter Verarbeitung eines ersten Korpus und sukzessiver Erweiterung des Untersuchungskorpus werden sich die Ergebnisse bei Untersuchungen von verschiedenen Ausgangspunkten her - sehr bald erkennbar - immer mehr einander annähern.
- Ein solches System ist auch "autopoietisch", d.h. auf Ausdrucksebene gibt es nur Ausdrücke. Ich kann - diachron - mein Ausdrucksrepertoire erweitern (durch Lernen). Es gibt aber - etwa durch weit vorangetriebene Ausdrucksanalyse keine Chance, diese Ebene zu verlassen und etwa in den Bereich der Bedeutungen zu gelangen.
- Bei systemtheoretischer Beschreibung gibt es zwei verschiedene Betrachtungsweisen: (a) *statisch* orientiert ist die Beschreibung des Systems aller Wortformen einer Einzelsprache durch das Stichwort der Differenz, s.o. - (b) *dynamisch* dagegen ist die Perspektive, wenn ich - bezogen auf konkrete Äußerungen (z.B. Text) - die lineare Verkettung von Ausdrücken untersuche.

Aus all dem folgt für die praktische Analyse:

- Bei klug eingesetzter rekursiver Analyse müßte der Computer das ideale Arbeitsinstrument einer Ausdrucksanalyse sein.
- Verfahren, die auf bloße Statistik zielen, sind von vornherein fehlorientiert.
- Stattdessen ist immer die kontrollierende Rückkoppelung zu integrieren: Welche Wissenssicherheiten (d.h. in dieser Einzelsprache akzeptierte Ausdrücke bzw. Ausdruckskombinationen) habe ich bereits erkannt, so daß ich darauf bei der Analyse noch unerforschter Bereiche einsetzen kann?
- Jeder neue Erkenntnissschritt enthält die Möglichkeit/Wahrscheinlichkeit, früher erzielte Ergebnisse zu differenzieren (nicht: korrigieren!).
- Folglich kann die Analyse nicht lediglich in einem linearen Fortschreiten bestehen, sondern immer auch in einer neuen kritischen Durchsicht früher erzielter Ergebnisse.
- Effekt dieser Prozedur wird sein: Wir sind nicht in einem vergleichsweise frühen Stadium mit der Analyse der Ausdrücke dieses Einzeltextes fertig und müssen nach weiteren Texten - zur Erweiterung des Ausdrucks-Repertoires - greifen. Vielmehr lassen sich

Ergebnisse - gewonnen an Ausdruck_y - übertragen auf Ausdruck_x, obwohl dessen Bezeugung^x (einschließlich Varianten_y) für sich genommen nicht so gut ist, daß sie die Aktion rechtfertigen würde. Dadurch werden z.B. neue Wortkerne sichtbar gemacht, mit denen das Korpus neu durchgegangen werden kann.

- (36) Weiterführung der früheren Auflösung einer Wortform in: selbständige Wortform + Endung.

BEURTEILUNG DER ENDUNGEN

Bsp. *wo+hnung*

Zusätzliche Bedingungen notwendig, u.z. aus dem gegebenen Korpus heraus. Ausgang von den Einmalwörtern in sortierter Form und von den Zerlegungen in selbständige Wortformen.

- (a) Wo ist ein Wort in - mindestens - dreifacher Form belegt: (aa) als selbständige Wortform, (bb) mit angehängter selbständiger Wortform (cc) mit angehängtem, noch ungeklärtem Element: ?
- (b) An welchen anderen Wortformen findet sich ebenfalls dieses noch erklärungsbedürftige Element?
- (c) Subtraktion von ? an diesen weiteren, vielleicht nur einmal belegten Wortformen, so daß, streicht man es als inzwischen wahrscheinlich gemachte legitime Endung ab, ein neuer Wortkern sichtbar wird.

Gegeben sei die Reihe: *musik - musiker - musikschiule - musikant*. Akzeptiert bzw. fraglich: *musik - musik+er - musik+schule - musik?ant*. Dieser Befund erlaubt aber die Hypothese, *+ant* als legitime, unselbständige Endung einer Wortform zu betrachten. - Die Endung begegnet auch in *lieferant*. Die Wortform kann nun vollständig zerlegt werden: *liefer+ant*.

Bedingung: Diese noch unerklärte Endung muß an mindestens zwei Wörtern hängen um weiterbehandelt zu werden.

Mit dieser Prozedur haben *wo+hnung* und *wo+lke* keine Chance mehr, sie sind als "Schrott" entlarvt. Damit verringert sich die Menge der ungeklärten Endungen.

- (37) In Analogie zu (36) z.B. folgende Ergebnisse:

blei+b
blei+ch
schreib
man+ch

Konvention:

ch zählt im Deutschen nur als *ein* Graph. Die Begründung läuft nicht phonetisch, sondern distributionell: Es gibt kein Wort, das mit *c* aufhört. Wenn *c* dann auch *ch*.

sch Ist anders gelagert. Zunächst muß die Trennung von Komposita (*rauch+schwaden*) und gängigen Wortendungen (*häus+chen*) durchgeführt sein, bevor man feststellt, daß *sch* immer eine Einheit bildet.

st, sp analog zu *sch* zu behandeln

Es ist aber zweifelhaft, ob mit derartigen Konventionen gearbeitet werden soll, ob dadurch nicht - unnötigerweise - mögliche Ergebnisse bereits zur Prämisse erhoben werden.

(38) Zwischenreflexion: Eine Einzelsprache bildet ihr Repertoire an Ausdrücken (100.000 - 300.000?) unter starken Restriktionen:

1. In Alphabet-basierten Sprachen stehen nur ca. 25 Bildelemente zur Verfügung.
2. Aus der Vielzahl der an sich möglichen Kombinationen wird in der Einzelsprache nur ein sehr begrenztes Korpus ausgewählt (ansonsten gäbe es die Verschiedenheit der Einzelsprachen nicht).
3. Sprachgebrauch und kognitive Kompetenz der Sprachbenutzer verhindern, daß die Wortformen allzu lang sind. Das Gros der Wortkerne liegt bei einer Länge von 2-8 Buchstaben.

Es ist also trivial, daß die Schaffung neuer Wortformen/Kerne dadurch geschehen muß, daß entweder "kreativ" eine Wortform geschaffen wird, deren Einzelelemente (bei beibehaltener Abfolge) keine andere Wortform ergeben (*eben*), oder dadurch, daß eine bestehende Wortform an einer Stelle erweitert wird (*fort* = *f+ort*).

Im letzteren Fall liegt also noch nicht das Problem: Kern + Prä-/Suffix vor, sondern zunächst nur die Schaffung weiterer Wortformen. Das Problem wird sein, diesen normalen und legitimen Vorgang von der Modifikation einer Wortform durch Hinzufügung gängiger Prä-/In-/Suffixe zu unterscheiden.

(39) Vgl. vor dem Hintergrund von (37) und (38) die Befunde:

grün
grün+de
grün+de+n
grün+de+r
grün+de+r+jahre
grün+de+te
grün+dlich
grün+dlich+e
grün+dlich+en
grün+dlich+es
grün+dlich+es
grün+dlich+keit
grün+donnerstag
grün+dung
grün+e
grün+es
grün+gekachelter
grün+gestrichenen
grün+kohl
grün+liche
grün+liche+m
grün+liche+n
grün+roten
grün+silbrig
grün+straße
grün+t

(40) Auf der Basis von (38) und (39) läßt sich gut eine Prozedur entwickeln, die weiterführt:

1. Es wird eine selbständige Wortform + Endung zugrundegelegt, die mindestens ein weiteres Mal nochmals belegt ist. z.B. *grün+dlich*. Bei unserer

Suche war *+dlich* 3x belegt (z.B. in *nie+dlich*). Es wird gesucht in der Liste selbständiger Wortformen.

2. Probehalter wird der Kern um einen Buchstaben erweitert; dementsprechend reduziert sich die Endung um einen Buchstaben. Auf dieser Basis wird erneut gesucht, ob sich nun diese Endung anderweitig findet und ob zugleich dieser neue Kern anderswo nachweisbar ist. *+lich* findet sich weitere 3x, so daß wir für die reduzierte Endung *+lich* 6 Belege haben.
3. Der neue Kern *gründ* wird als selbständig nicht nachweisbar sein. Sobald die Prozedur aber mit anderen Endungen durchgespielt ist, haben sich dort sowohl neue Endungen wie auch neue Kerne finden lassen.

Das Durchspielen dieser Prozedur ausgehend von mehreren Einzelbefunden wird sehr schnell die Menge der Wortkerne erweitern, vor allem, weil durch das Erkennen von akzeptierten Endungen die gegebenen Wortformen weiter analysiert werden können und dadurch weitere Kerne zum Vorschein kommen. Die Analyse müßte damit einen sich beschleunigenden Charakter annehmen.

...

(44) KERN / KERNERWEITERUNG / ENDUNG:

Hier der Versuch einer terminologischen Vereinbarung. Auf dieser Basis soll anschließend zusammenfassend skizziert werden - soweit es derzeit möglich ist - die unterschiedlichen TYPEN von erreichbaren Ergebnissen zusammenzustellen.

KERN : Darunter wird eine Einheit von mindestens zwei Buchstaben verstanden, die distributionell nachgewiesen ist als Anker für mehrere Arten von *erweiternden* Umformungen.⁴³

KERNERWEITERUNG: Es handelt sich um ein Bildelement, das nicht als selbständige Wortform nachweisbar ist, das aber in Verbindung mit unterschiedlichen Kernen offenbar neue Kerne schafft. Weiteres Kriterium: Diese Kernerweiterungen sind nachweisbar mit Endungen (s.d.), bilden dann also nicht den Schluß der Wortform. Zu denken ist an: *&heit*, *&keit*, *&ung*, *&ling*, *&end*, *&lein*, *&chen*. SCHREIBKONVENTION: *&* soll in Zukunft den Beginn einer Kernerweiterung anzeigen. u.ä.

ENDUNG : Unselbständiges Bildelement oder selbständige Wortform (*/er/*), die eben auch in der Position einer Endung begegnet, das Ende der Wortform

⁴³ Nachtrag Juli 96: Es wurde bei der Definition nicht berücksichtigt, daß KERN ein zweites Merkmal hat: KERN ist zunächst eine Einheit, die in sich nicht weiter vollständig zerlegbar ist. Eine solche Einheit ist dann Ausgangspunkt für verschiedenartige Erweiterungen (=ANKER). KERN ist auch nicht identisch mit WORTFORM; KERN kann - s.u. - zusammenfallen mit WORTFORM, kann aber auch nur einen Teil einer WORTFORM darstellen = unselbständiger Kern.

markierend; angehängt entweder an KERN, wenn dieser = selbständige WORTFORM, oder an KERNERWEITERUNG. SCHREIBKONVENTION: + soll in Zukunft den Beginn einer Endung anzeigen.

: Zwei Gesichtspunkte müssen beachtet werden: (1) Es ist mit mehrstufigen Endungen zu rechnen, z.B. *ver=schied+en+er+lei*. - (2) Noch nicht gesichert ist im Moment, ob "Endung" und "Kernerweiterung" distributiv sicher unterscheidbar sein werden.

INFIX

: Unselbständiges Bildelement, das bei COMPOUNDS (s.d.) den Übergang zwischen den separat nachweisbaren Wortformen markiert. Es ist zu überlegen, ob bei COMPOUNDS nicht grundsätzlich der Übergang markiert werden soll. Wenn explizit die Schnittstelle morphologisch nichts aufweist, dann eben durch Leerstelle, z.B. *Mäd&chen-Ø-handel-s-schul+e*. SCHREIBKONVENTION: -x- soll in Zukunft die Schnittstelle bei Compounds anzeigen.

PRÄFIX

: Analog zu ENDUNG. Auch Präfixe können addiert auftreten. SCHREIBKONVENTION: = soll in Zukunft das Ende eines Präfixes anzeigen, z.B. *vor=arbeit*, *un=ver=hohl+en*, *un=ver=käuf+lich*.

WORTFORM

: = Wf. ist das, was in der Einzelsprache bzw. dem zugrundeliegenden Untersuchungstext zwischen blanks begegnet. SCHREIBKONVENTION: | soll in Zukunft Beginn und Ende einer Wortform anzeigen, z.B. *|deshalb|*.

COMPOUND

: ist aus mindestens zwei Wortformen zusammengesetzt, die durch Infix verbunden sind. Das Infix kann dabei Ø sein:

|donau-Ø-dampf-Ø-schiff-Ø-fahrt-s-ge=sell&schaft-s-kapit-än|

zu Ziff. 4.3324:

Morphologie - (ausdrucks-)syntaktisch

zu:

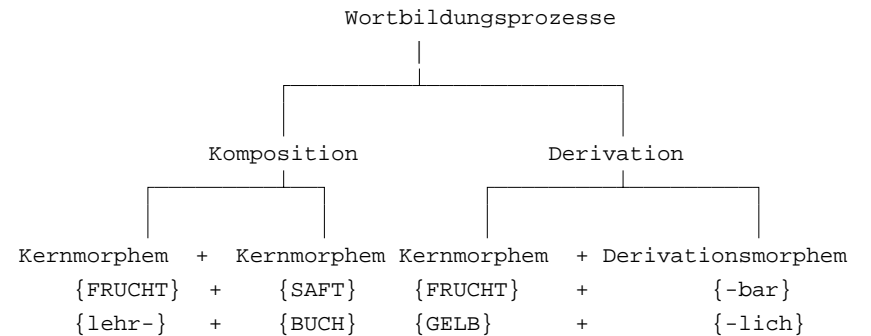
nach: R. PÖRINGS; U. SCHMITZ (eds.), Sprache und Sprachwissenschaft. Eine kognitiv orientierte Einführung. Tübingen 2003.

(55)

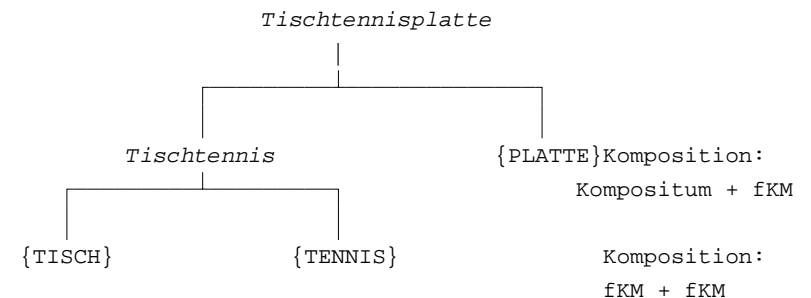
Viele Wörter wie *Traum*, *Elefant*, *zwei*, *weiß* etc. bestehen jeweils nur aus einem einzigen freien Kernmorphem. Man nennt sie deshalb auch **einfache Wörter** oder **Simplizia**.

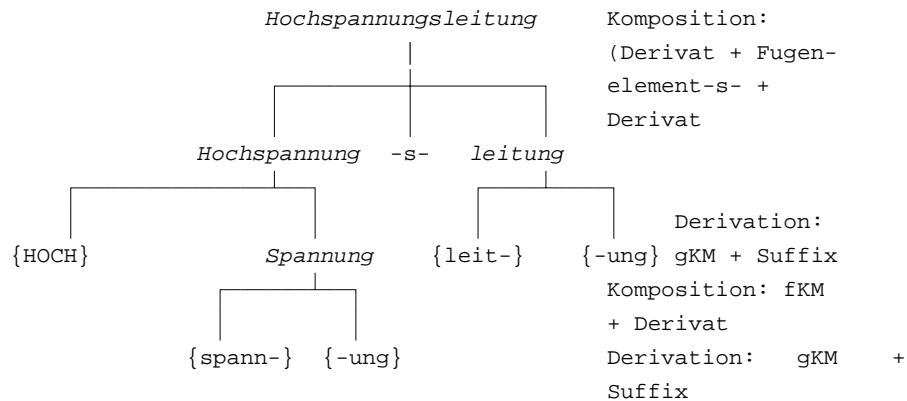
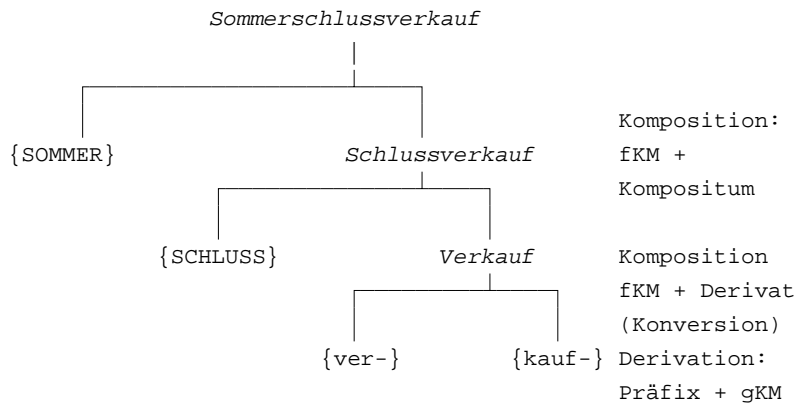
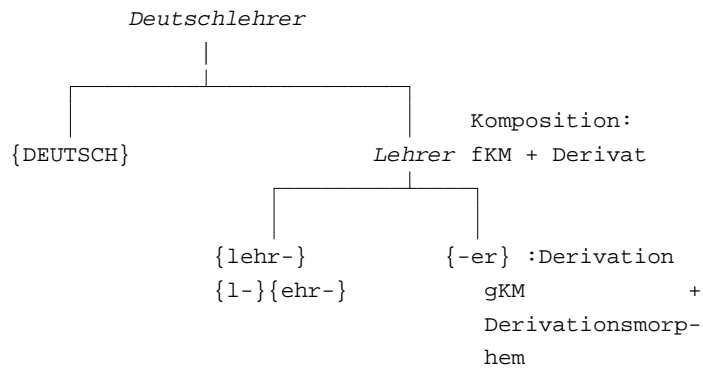
Freie und gebundene Kernmorpheme wie {FRUCHT}, {SAFT}, {lauf-}, und {BAND} lassen sich zu **komplexen Wörtern** zusammensetzen, nämlich *Fruchtsaft* bzw. *Laufband*. Das Ergebnis einer solchen **Zusammensetzung** oder **Komposition** ist ein **Kompositum**.

(56) Komplexe Wörter können auch gebildet werden, indem im einfachsten Fall ein gebundenes Morphem an ein freies oder gebundenes Kernmorphem angehängt wird, z.B. {FRUCHT} + {bar} → *fruchtbar*, {GELB} + {-lich} → *gelblich* oder {spring-} + {-er} → *Springer*. ... man bezeichnet diesen Prozess zur Bildung komplexer Wörter als **Ableitung** oder **Derivation**.



Nicht nur Kernmorpheme, sondern auch bereits zusammengesetzte oder abgeleitete komplexe Wörter können die Basis für weitere Prozesse der Zusammensetzung bzw. Ableitung sein, wie folgende Beispiele zeigen (KM = Kernmorphem, fKM = freies Kernmorphem, gKM = gebundenes Kernmorphem):





(57) Derivate können zum einen durch Voranstellung eines Derivationsmorphems vor ein freies oder ein gebundenes Morphem gebildet werden wie in {un-} + {KLUG} → *unklug* bzw. in {be-} + {schreib-} → *beschreib(en)*. Solche vorangestellten gebundenen Morpheme bezeichnet man als **Präfixe**, den Wortbildungsprozess als **Präfigierung**. Zum ändern können gebundene grammatische Morpheme aber auch an ein Kernmorphem angehängt werden; in diesem Fall

bezeichnet man sie als **Suffixe** und den Prozess als **Suffigierung**, z.B. {FURCHT} + {-bar} → *furchtbar*. **Zirkumfixe** umgeben ein Kernmorphem von beiden Seiten her, wie {ge- + -et} in *gearbeitet*. In manchen Sprachen gibt es darüber hinaus noch **Infixe**, die in ein Kernmorphem eingefügt werden (z.B. lat *vinco, vici*). Im Deutschen spielen Infixe keine Rolle. Präfixe, Suffixe, Zirkumfixe und Infixe werden zusammengefasst als **Affixe** (von lat. *affigere* 'anheften') bezeichnet.

HS: Eigentliche methodische/algorithmische Herausforderung ist sicher unterscheiden zu können zwischen

- wo tritt ein gängiges, in der betr. Einzelsprache gebräuchliches **Prä/Affix** an eine Wortform, wobei - trotz Erweiterung - die Wortform die selbe bleibt;
- oder: wo tritt/treten an eine Wortform ein/mehrere Buchstabe/n, so dass dadurch aber eine **neue Wortform** entsteht (an die dann wieder gängige Affixe treten können)?

4.333 Häufigkeit und Wortlänge

zu Ziff. 4.3331:

ZIPF: Relative Häufigkeit und Wortlänge

zu:

aus: C. PRÜN, Das Werk von G. K. ZIPF in: KÖHLER, R u.a. (eds.), Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005

(143) In seiner ersten Veröffentlichung, *Relative Frequency as a Determinant of Phonetic Change* (Zipf 1929), postuliert er ein Prinzip, demzufolge die Häufigkeit einer sprachlichen Einheit ihre Form bestimmt. ...

[Prinzip der Häufigkeit. Das Gewicht bzw. der Grad der Auffälligkeit eines Wortes, einer Silbe oder eines Lautes ist umgekehrt proportional zur relativen Häufigkeit dieses Wortes, dieser Silbe oder dieses Lautes in der Rede. Wenn der Gebrauch häufiger wird, verliert die Form an Gewicht und wird einfacher aussprechbar und umgekehrt.] ...

Die Gültigkeit dieses Satzes leitet Zipf von drei kommunikativen Notwendigkeiten ab: erstens muß die Unterscheidbarkeit sprachlicher Einheiten oder Merkmale gewährleistet sein, und zwar je seltener eine Einheit ist, desto besser muß sie von sich ähnlichen Einheiten und von den umgebenden Einheiten derselben Ebene abheben. Häufige Einheiten werden stärker erwartet; ihr Informationsgehalt ist geringer und es ist deshalb nicht so zwingend notwendig, daß sie genau verstanden bzw. aufwendig (144) artikuliert werden. Zweitens muß der Aufwand beim Sprechen ökonomisch verteilt werden, das heißt, wenn ein Element häufig, aber aufwendig auszusprechen ist, dann tendieren die Sprecher dazu, die Aussprache zu vereinfachen. Diese beiden Tendenzen, die seit Köhler (1986) als Systembedürfnisse "Minimierung des Dekodierungsaufwandes" und "Minimierung des Produktionsaufwandes" umschrieben werden, bilden einen Spannungsbereich, in dem geringer Sprecheraufwand mit hoher Häufigkeit und Redundanz mit geringer Häufigkeit funktional verbunden sind. ...

Für Wörter mißt Zipf den Aufwand bzw. die Deutlichkeit als Wortlänge. In (Zipf 1932, 8ff.) zeigt er, wie die Deutlichkeit seltener gebrauchter Wörter dadurch gewährleistet wird, daß sie im Durchschnitt länger sind als die häufig gebrauchten Wörter. Dieses Verhältnis ist das Resultat von Abkürzungsprozessen, die alle überflüssigen Elemente eines Wortes wegfallen lassen, solange sich die Kürzung durch große Häufigkeit des Wortes vertreten läßt (Zipf 1932, 18f). Wie bei den Lauten gilt auch hier, daß häufige, stärker erwartete Elemente weniger Information tragen und deshalb unwichtiger sind bzw. leichter geraten werden können. Eine Folge davon ist, daß es wenige häufige, aber mit zunehmender Gebrauchshäufigkeit immer mehr verschiedene Wörter gibt. Auch diesen Zusammenhang drückt Zipf (1932, 24) mit einer Formel aus, nämlich (a = Worthäufigkeit; b = Anzahl Wörter mit Häufigkeit a)

$$a b^2 = \text{const.},$$

(145) was später als eins der "Zipfschen Gesetze" bezeichnet wird (auch als Frequenzklassenbelegungsgesetz und als Häufigkeitsverteilung eines Textes bezeichnet).

zu Ziff. 4.3332:

ZIPF: Sprachökonomie / Distribution

zu:

aus: C. PRÜN, Das Werk von G. K. ZIPF in: KÖHLER, R u.a. (eds.), Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005

(148) Die Elemente der Sprache unterliegen dem Druck, als effektive Kommunikationsmittel zu funktionieren, aber dabei Sprachverarbeitung und Gedächtnis möglichst minimal zu belasten. Unifikation entlastet den Sprecher, indem Wörter mehrere Bedeutungen tragen können und das Vokabular kleiner wird. Sie kann deshalb als die individualistische Tendenz, verstanden werden. Diversifikation bewirkt, daß Wort und Bedeutung im 1 : 1-Verhältnis stehen und sichert damit die Informationsübertragung, weshalb ihre Funktion eine soziale ist. Zipf (1945a, 27ff.; 75ff.) kann allerdings zeigen, daß die Häufigkeit der Wörter nicht linear mit der Anzahl ihrer Bedeutungen abnimmt, sondern Wörter mit mehreren Bedeutungen systematisch häufiger sind als solche mit wenigen und die Häufigkeiten der Zipfschen Kurve folgen, d.h. wiederum wirkt Unifikation stärker im Bereich großer Häufigkeit, Diversifikation dagegen im Bereich geringer Häufigkeit.

Bestimmte Charakteristika der Rang-Frequenz-Verteilung lassen Schlußfolgerungen über den seelischen Zustand des Sprechers zu. Während Kinder ein verkleinertes Lexikon haben, das aber im Prinzip ökonomisch strukturiert ist (Zipf 1942a), haben Texte von Menschen in Stresssituationen oder mit psychischen Störungen deutlich abweichende Häufigkeitsstrukturen. Die Abweichung geschieht in der Regel in Richtung einer verstärkten Unifikation, was auf eine starke Ich-Orientierung der Probanden hinweist (Whitehorn/Zipf 1943; Zipf 1972, 288ff.).

Auch die Verwendung von gleich häufigen Wörtern (Zipf 1945b; 1972, 45ff.) scheint ökonomisch zu geschehen, denn ihre Abstände nähern sich einer Gleichverteilung an. Zipf (1945b, 47ff.) bringt dies mit einem Konzept der Zeit-Perspektive in Verbindung, die die Gestaltung von Texten bestimmt. In wie weit die Zeit-Perspektive mit dem Text als abgeschlossenem Ganzen (Orlov 1982) zusammenhängt, ist bis heute nicht wissenschaftlich bearbeitet worden.

4.34 Neologismen

4.35 Dialekte

Literatur

HYVÖNEN, S; LEINO, A u.a.: Multivariate Analysis of Finnish Dialect Data - An Overview of Lexical Variation. Literary and Linguistic Computing 22 / 3 (2007) 271-290.

4.4 Einzeltextbezogen: Algorithmen, Darstellung und Auswertung einer textinternen Befunderhebung

Literatur

- BADER, W: Ausdruckssyntax und Textgrammatik. Statistik im Methodendreischnitt Syntax-Semantik-Pragmatik. Actes du Premier Colloque International Bible et Informatique: Le Texte. Paris-Genève 1986 51-79. // I.2.0
- BADER, W: Ausdruckssyntaktische Untersuchungen biblischer Texte. Literary and Linguistic Computing 4 (1989) 282-285.
- BADER, W: Algorithmen zur internen Ausdruckssyntax in: BADER, W (ed./eds.): "Und die Wahrheit wurde hinweggefegt". Daniel 8 linguistisch interpretiert. THLI 9. Tübingen 1994 149-159.: Francke // Raum B106
- BADER, W: Interne Syntax: Der Befund an identischen Wortformen in: SCHWEIZER, H (ed./eds.): Computerunterstützte Textinterpretation. Die Josefsgeschichte beschrieben und interpretiert im Dreischnitt: Syntax-Semantik-Pragmatik. THLI 7. Tübingen 1995 17-41 (i); 40-48 (ii).: Francke // I.2.7
- CACCIARI, C (ed./eds.): Similarity in Language, Thought and Perception (Papers presented at the workshop held in 1991 at the International Center for Semiotic and Cognitive Studies of San Marino). Semiotic and cognitive studies 1. Turnhout 1995: Brepols // I.2.0
- CLEMENT, T E: 'A thing not beginning and not ending'. using digital tools to distant-read Getrude Stein's *The Making of Americans*. Literary and Linguistic Computing 3 (2008) 361-381.
- HŘEBIČEK, L: Hurst's indicators and Text. Some properties of word-frequency series in: ALTMANN, G; KOCH, W A (ed./eds.): Systems. New Paradigms for the Human Sciences. Berlin, New York 1998 572-588.: de Gruyter // EB
- HŘEBIČEK, L: Contextual word prominence in: Grzybek, P; Köhler, R (ed./eds.): Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 238-243.: Mouton de Gruyter // I.2.5
- HOOVER, D L: Frequent Word Sequences and Statistical Stylistics. Literary and Linguistic Computing 17 / 2 (2002) 157-180.
- JUILLARD, M, X; LUONG: Words in the Hood: A New Look at the Distribution of Words in Texts. Literary and Linguistic Computing 12 (1997) 71-78.
- KOSAREV, R, YU; PIOTROWSKI: Synergetics and 'Insight' Strategy for Speech Processing. Literary and Linguistic Computing 12 (1997) 71-78.
- LEUNG, C-H; KAN, W-K: A Statistical Learning Approach to Improving the Accuracy of Chinese Word Segmentation. Literary and Linguistic Computing 11 / 2 (1996) 87-92. // EB

- LINMANS, A J M: Correspondence Analysis of the Synoptic Gospels e. Literary and Linguistic Computing 13 (1998) 1-13.
- McKENNA, W; ANTONIA, A: 'A Few Simple Words' of Interior Monologue in *Ulysses: Reconfiguring the Evidence*. Literary and Linguistic Computing 11 / 2 (1996) 55-66. // EB
- MEALAND, D: Measuring Genre Differences in Mark with Correspondence Analyses. Literary and Linguistic Computing 12 (1997) 209-217.
- MEHLER, A: Textbedeutung. Zur prozeduralen Analyse und Repräsentation struktureller Ähnlichkeiten von Texten. Sprache, Sprechen und Computer 5. Frankfurt / M 2001 ISBN 3-631-38648-6 br: Peter Lang // I.2.7
- NAKAMURA, JH; SINCLAIR, J: The World of Woman in the Bank of English: Internal Criteria for the Classification of Corpora. Literary & Linguistic Computing 10,2 (1995) 99-110.
- NIDA, E A; LOUW, J P et al.: Style and Discourse. With special reference to the text of the Greek New Testament. Roggebai 1983
- PANCHANAN, M: On script complexity and the Oriya script in: Grzybek, P; Köhler, R (ed./eds.): Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 473-484.: Mouton de Gruyter // I.2.5
- RAPP, R: Are word senses reflected in the distribution of words in text? in: Grzybek, P; Köhler, R (ed./eds.): Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 575-586.: Mouton de Gruyter // I.2.5
- RIEGER, B: Unscharfe Semantik. Die empirische Analyse, quantitative Beschreibung, formale Repräsentation und prozedurale Modellierung vager Wortbedeutungen in Texten. Frankfurt / M 1989 // I.2.4
- SCHWEIZER, H: Die Sprache der Zeichenkörper. Textinterne (Ausdrucks-)Syntax zu Dan 8 in: BADER, W (ed./eds.): "Und die Wahrheit wurde hinweggefegt". Linguistische Analysen von Daniel 8. THLI 9. Tübingen 1994 17-30.: Francke // Raum B106
- SCHWEIZER, H: Preparations for the Redemption of the World. Distribution of Words and Modalities in Chapter I of Don Quixote in: KIEGELMANN, M; GÜRTLER, L (ed./eds.): Research Questions and Matching Methods of Analysis. Qualitative Psychology Nexus: Vol. 3. Tübingen 2003:

4.41 Beispiele, Motivation

zu Ziff. 4.411:

PLANKENHORN, G: AWA! Kleine Sprengsätze und andere Gedichte. o.J.
Tübingen. 2. Aufl.

zu: vsyst4.301

Stichwort: Ausdruckssyntax, Konkrete Poesie,

VOM ÄRBA

Dees
Dees dô
on dees dô
on dees dô noh
on dees dô nôh noh
on dees dô nôh noh ao
on dees dô nôh noh ao noh
on dees dô nôh noh ao nôh noh
on sell
on sell dô
on sell dô noh
on sell dô nôh noh
on sell dô nôh noh ao
on sell dô nôh noh ao noh
on sell dô nôh noh ao nôh noh
braucha kenne mir aegendlich alles
wôesch, mir kenne aegendlich alles braucha

zu Ziff. 4.412:

PLANKENHORN, G: AWA! Kleine Sprengsätze und andere Gedichte. o.J.
Tübingen. 2. Aufl.

zu: vsyst4.302 vprag7.22

Stichwort: Ausdruckssyntax, Implikationen

Haelix Blächle

Audo
Au dô
au dô a Audo
dô a Audo
ao dô a Audo
dô ao a Audo
au ao dô a Audo
au dô mae Audo
mae Audo !
Dôlao !
Dôlao !!
Mae Audoooo !!

zu Ziff. 4.413:

kurt marti, republikanische gedichte. neuwied 1971 (st.gallen 1959). s. 22

zu: vprag1.601

Stichwort: Ausdruckssyntax vs. Textgrammatik

- | | |
|-----------------------------------|-----|
| coué unser bestalozzi | (1) |
| die besten schulen | (2) |
| haben | (3) |
| die sich selbst am besten schulen | (4) |
| in der überzeugung | (5) |
| daß nur sie | (6) |
| die besten schulen | (7) |
| haben | (8) |

zu Ziff. 4.414:

Textinterne Ausdrucks-SYNTAX

zu:

aus: Besprechung der ADORNO-Biografie von Detlev Claussen (2003) durch Andreas Bernard in FAZ.

... Das Verhängnisvolle ist, dass sie (= die Biografie, H.S.) sich immer wieder um die Gunst ihrer Leser bringt - und zwar durch einen erstaunlichen Mangel an stilistischer Sorgfalt. Im Laufe der Lektüre fallen zahllose Wiederholungen und Überschneidungen ins Auge, so als hätte die Zeit gefehlt, die umfangreichen Einzelkapitel noch einmal aufeinander abzustimmen. So wird etwa in dem Abschnitt "Der Identische", der die chronologische Ordnung unterbricht und einen Exkurs über die Beziehung zu Horkheimer einschleibt, zu großen Teilen noch einmal dasselbe erzählt wie in einem früheren Kapitel. Mehr als alles andere befremdet jedoch die Wiederkehr bestimmter versatzstückartiger Formulierungen, so etwa die vom "älteren Mentor" Kracauer, die nicht weniger als 25 Mal auftaucht. Auch dass sich Horkheimer in den fünfziger Jahren "nach Montagnola zurückzog", erfährt der Leser in einem knappen Dutzend Nebensätzen. Mit welcher erstaunlicher Nachlässigkeit dieses so ambitionierte Buch letztendlich veröffentlicht wurde, zeigt schließlich auch, dass mehr als einmal exakt dieselbe Aussage auf zwei aufeinanderfolgenden Seiten zu finden ist. (Seite 328 etwa: Den von Bloch, "angeforderten wieder aufgelegten 'Kierkegaard' relativiert Adorno als eine höchstens *traumhafte Antezipation*"; Seite 330: "Adorno schickte Bloch den wieder aufgelegten 'Kierkegaard', dessen Bedeutung er herunterspielt und höchstens im Sinne einer *traumhaften Antezipation* gelten lassen will.") Das Gedächtnis dieser Biographie ist auffallend kurz, sie vergisst ihre Sätze zuweilen von Seite zu Seite. Das ist um so problematischer, als Claussen von seinem Lehrer einmal mit fast beängstigender Bewunderung sagt: "Das Verschlingen einer Unmenge von Literatur bei einer die menschlichen Dimensionen übersteigenden Erinnerungsfähigkeit charakterisiert zweifellos Adorno". Auch darin hätte der Biograph seinem Gegenstand ein wenig gerechter werden müssen.

zu Ziff. 4.415:

nonsens englisch ausgesprochen = deutsch?

zu: vprag8.122

aus: Redemanuskript von Präsident Kennedy bei seinem Berlin-Besuch.

ish Froy-er mish in bear-LEAN sur zine.

ish Froy-er mish in DOICH-lont sue zine.

ish bin DONK-bar fear dane HAIRTS-lishen emn-FONG.

dane zee mere eebe-OLL ba-WRY-tet hobben.

ish HAUFER doss mine ba-ZUKE dot-sue BUY-traket,

dee FROYNT-shofts vishen DOICH-lont oont Ah-MAY-ree-cah

VIE-tar sue FEST-iggen.

zu Ziff. 4.416:

Gleichheit der Ausdrücke ↔ Verschiedenheit der Bedeutungen

zu: vprag3.101

aus: M. HEPERLE, Gestammelte Werke. 1979.

Anlagen

Im Sommer, wo bevorzugt an wärmenden Tagen Anlagenanlieger in den Anlagen lagen, unangelegte Anlagen aber meist überwiegen, wird das Anlegen im Anliegen der Anlieger liegen.

Ein Gärtner wird also der Anlieger wegen seine Veranlagung ins Anlegen der Anlagen legen. Da sieht das Finanzamt mit großem Vergnügen im Anlagenanlegen auch eine Veranlagung liegen.

Nach Lage der Anlage liegt die Sache nun klar, daß das Anlegen zwar Anliegeranliegen war. Unangelegte Anlagen füllen zwar dem Gärtner die Kassen, während sich erst angelegte vom Fiskus veranlagen lassen.

4.42 Worthäufigkeit / Distribution

Literatur

SOTOV, A: Lexical Diversity in a Literary Genre. A Corpus Study of the *R̥gveda*. *Literary and Linguistic Computing* 24 (2009) 435-447.

zu Ziff. 4.421:

Wortdistribution im Einzeltext

zu:

aus: McKEE, G; LAVER, D, u.a., *Measuring Vocabulary Diversity Using Dedicated Software: Literary & Linguistic Computing* 15/3 (2000) 323-337.

(323) Researchers in various fields of linguistic enquiry have calculated vocabulary diversity using the ratio of different words (Types) to total words (Tokens) - the Type - Token Ratio (TTR) - or measures derived from it. Such measures are flawed, however, because the values obtained are related to the number of words in the sample...

(324)... as longer and longer samples of language are produced, more and more of the active vocabulary is likely to be included and the available pool of new word types that can be introduced steadily diminishes. It is obvious that once a sample is large enough to have included all of the subject's active vocabulary, any further sampling of tokens can only result in a hyperbolic decline in the values for TTR. But, it is also the case that however small the sample is, as more and more tokens are taken, the likelihood is that (because of repetition of previously included types) the cumulative number of types will increase at a slower rate than the number of tokens and the TTR values inevitably fall...

The *vocd* program was developed to overcome these problems as part of the project "A new research tool: mathematical modelling in the measurement of vocabulary diversity". The approach is based on an analysis of the probability of new vocabulary being introduced into longer and longer samples of speech or writing. This yields a mathematical model of how TTR varies with token size. By comparing the mathematical model with empirical data in a transcript, it provides a new measure of vocabulary diversity that we refer to as *D*. The measure has three advantages: (1) it is not a function of the number of words in the sample; (2) it uses all the data available; (3) it is more informative because, as opposed to a single value of TTR, it represents how the TTR varies over a range of token size for each speaker or writer (i.e. it is based on the TTR versus token curve calculated from data for the transcript as whole rather than a particular TTR value on it).

zu Ziff. 4.422:

Interne Syntax, Binnenwortschatz und seine Distribution

vsyst4.31

I)

[JOSEPH] [war] [als] [Hirte] [mit] [seinen] [Brüdern] [beim]
[1] [2] [3] [4] [5] [6] [7] [8]

Klammert man zunächst *Buchstaben*, *Silben* als
Beschreibungs-/Analyseeinheiten aus, dann kann man einen Text
verstehen als lineare Abfolge von *Wörtern*, besser: *Wortformen*:

$Wortform_1 \dots Wortform_n$

II)

Die hebräische, ursprüngliche Josefsgeschichte erstreckt sich von

$Wortform_1 \dots Wortform_{2512}$

III)

Zu jedem Wort kann man dann, wenn es zum erstenmal auftritt,
notieren, an welcher Position im Text es noch belegt ist:

[1] וְיִשְׂרָאֵל 1 12 25 80 140 180 185 254 259 265 277 301 312 352 467
478 514 533 562 610 674 699 765 771 912 925 940 953
1021 1091 1117 1159 1170 1176 1189 1207 1269 1289 1531
1533 1557 1563 1576 1585 1632 1685 1700 1708 1712 1725
1734 1759 1814 1823 1855 1877 1896 1905 1915 1947 1966
1977 2034 2062 2098 2121 2135 2150 2162 2175 2178 2246
2275 2365 2374 2394 2434 2444 2452 2459 2479

token = jedes Vorkommen einer Wortform

type = Klasse von *n*-mal belegten gleichen Wortformen (*tokens*)

Damit ist die *Häufigkeit* gegeben (hier: 81x). Vergleich mit der
Häufigkeit anderer Wortformen führt auf *Leitwörter* (s.u.
zusätzlich das Thema "Erstreckung").

IV)

Aus den Daten geht hervor: Ist eine *Wortform*, die an *n*-ter Stelle
im Text auftritt, *neu* oder ist sie früher schon genannt worden?
z.B. kennen wir das *token* 1966 schon sehr lange, weil es zum *type*
1 gehört.

Damit kann man nach Textstellen suchen, wo auffallend viele *neue*
Wortformen in unmittelbarer Folge auftreten: Starkes Signal für
Abschnittsgliederung.

Wortnummer	Zahl neuer Wörter
1	10
26	17
101	10
157	17
209	12
.....	

V)

Welche Textbereiche sind durch starke *Verwendung neuer Wortformen*
gekennzeichnet, welche dagegen begnügen sich, auffallend breit
schon bekannte Wörter *wiederzuverwenden*? - Bereiche hoher
Wortschatzzunahme

Beginn mit Position	Länge	Neue Wörter	y'
1	405	277	0,68
463	79	43	0,54
579	120	69	0,57
714	25	17	0,63
.....			

VI)

Wo sind unmittelbar aufeinanderfolgend besonders viele Wortformen
wiederholt? **Wortschatzstagnation** = stark retardierendes Element,
der Einprägsamkeit dienend.

Wortnummer	Länge (10 und mehr wiederholte Wörter)
864	12
956	10
1161	11
1240	10
1814	16
.....	

VII)

Binnenwortschatz In welchem Textbereich ist die Anzahl von
Wortformen (*types*), die *nur hier* vorkommen, auffallend hoch? - In
der Josefsgeschichte kann der Bereich 220-530 als besonders
abgeschlossen gelten.

VIII)

Ergänzung zu **Leitwörtern**: Neben auffallend hoher Häufigkeit geht
es auch darum, daß diese Wörter eine weite Erstreckung haben, ihr
Vorkommen nicht nur lokal begrenzt ist.

Vgl.: WINFRIED BADER, *Interne Syntax*. in: HARALD SCHWEIZER,
Computerunterstützte Textinterpretation. Die
Josefsgeschichte beschrieben und interpretiert im
Dreischritt: Syntax - Semantik - Pragmatik. Bd. I:
Textbeschreibung und -interpretation. Tübingen 1995.
S.17-41; Bd. II: Anhänge zu den Textanalysen. S.40-48.

IX)

Cluster: Ergänzender Gesichtspunkt. Erstreckung vorzugeben (z. B.
Wortabstand $A = 20$). Frage: In welchen Textbereichen sind
auffallend viele *tokens* mit einem Wortabstand A miteinander
verzahnt? Wo bricht diese Verzahnung ab?

T₁ ++
 T₂ ++
 T₃ ++
 T₄ ++
 T₅ ++

+.....+

zu Ziff. 4.423:

Textkohärenz auf verschiedenen Analyseebenen

zu: vprag2.4422 vprag7.001

aus: BRINKER, K, Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin ²1988.

(27)

(1) **Ein Mann** war zu Rad unterwegs und wollte auf einen Berg steigen; er sah ein Anwesen liegen und stellte dort ein. *Der Mann* hieß Oberstelehn und hielt von sich nicht mehr viel; er konnte auch mit seinem Namen nicht Staat machen, der die Amtsstuben verdroß...

(aus: Gerd Gaiser, Eine Stimme hebt an. Roman. 1950 - Textanfang)

(2) Eines der ekelhaftesten Verbrechen wird **dem 47jährigen Düsseldorfer Rechtsanwalt H.J.O.** vorgeworfen. *Der Jurist* soll die Entführung des Millionärs Th.A. inszeniert und dessen Familie um sieben Millionen Mark erpreßt haben...

(aus: Die Zeit vom 7.1.1972 unter der Schlagzeile "Schwarzes Schaf")

(3) Ein 79jähriger Rentner wurde in der Nacht zum Dienstag in Harburg von **einem Auto** tödlich verletzt. *Das beschädigte Fahrzeug* und seine drei Insassen wurden im Laufe des Tages gefunden.

(aus: Die Welt vom 4.6.1980 unter der Schlagzeile "Todesfahrer gefaßt")

(4) Auf ungewöhnliche Weise wollte **ein 43 Jahre alter Mann aus Pforzheim** in der Nacht zum Donnerstag Selbstmord begehen. Wie die Polizei mitteilte, war *der Facharbeiter* nach Streitigkeiten in seiner Wohnung in Notarrest gebracht worden. Dort leerten die Beamten *dem Betrunkenen* vorschriftsmäßig die Taschen, um "Dummheiten" *des Gefangenen* zu verhindern. Eine halbe Stunde später fanden die Beamten *den Mann* mit aufgeschnittenen Pulsadern in seiner Zelle vor...

(aus: Die Welt vom 22.7.1977 unter der Schlagzeile "Gefährliches Glasauge")...

(29) Wenn wir von der unproblematischen Wiederaufnahme durch dasselbe Wort, der sog. Repetition, hier einmal absehen, sind - wie oben bereits angedeutet wurde - mindestens zwei Fälle zu unterscheiden:

- Wiederaufnahme durch andere Substantive
- Wiederaufnahme durch Pronomen

Wir gehen zuerst auf die Wiederaufnahme des Bezugsausdrucks durch andere Substantive bzw. substantivische Wortgruppe ein.

Eine genauere Betrachtung der Textbeispiele 2 und 3 zeigt, daß zwischen den hervorgehobenen Ausdrücken bestimmte Bedeutungsbeziehungen bestehen. Die Substantive *Rechtsanwalt* und *Jurist* bzw. *Auto* und *Fahrzeug* sind zwar nicht bedeutungsgleich (synonym) wie die Wörter *Briefträger* und *Postbote* oder *Fahrstuhl* und *Lift*, aber sie sind doch bedeutungsähnlich. Wir können sagen: die Wörter *Jurist* bzw. *Fahrzeug* stellen gewissermaßen Oberbegriffe zu (30) *Rechtsanwalt* bzw. *Auto* dar, d.h., sie besitzen einen größeren Bedeutungsumfang als die Wörter *Rechtsanwalt* bzw. *Auto*, während diese aufgrund spezifischerer Bedeutungsmerkmale einen größeren Bedeutungsinhalt aufweisen. Diese Bedeutungsbeziehungen bestehen nun unabhängig vom konkreten Text im Lexikon unserer Sprache, sie sind sozusagen im sprachlichen System selbst verankert.

Das ist anders in Textbeispiel 4. Die Substantive *Mann*, *Facharbeiter*, *Betrunkenener*, *Gefangener* stehen auch hier in der Relation der Wiederaufnahme, d.h., sie beziehen sich auf dieselbe Person. Zwischen diesen Wörtern bestehen aber keine besonderen, im Sprachsystem vorgegebenen Bedeutungsbeziehungen (wenn man von dem allgemeinen semantischen Merkmal "männlich", das alle Wörter verbindet, einmal absieht). Die referenzidentische Verknüpfung von *Mann*, *Facharbeiter*, *Betrunkenener*, *Gefangener* wird erst in diesem Text und durch diesen Text aufgebaut und hat somit nicht für alle möglichen Texte Gültigkeit, sondern nur für diesen Text, in dem sie bestimmt ist...

Wenn wir nun die **A b f o l g e** der Ausdrücke in den Textbeispielen (2) und (3) betrachten, dann läßt sich die folgende Regularität feststellen. Die Ausdrücke mit dem größeren Bedeutungsumfang, die sog. Oberbegriffe (*Jurist*, *Fahrzeug*) sind die wiederaufnehmenden Ausdrücke, während die Wörter mit der spezifischeren Bedeutung als Bezugsausdrücke fungieren - oder vereinfacht formuliert: In der Wiederaufnahmerelation folgt der Oberbegriff auf den Unterbegriff und nicht umgekehrt.

zu Ziff. 4.424:

Computergestützte Wortanalyse

zu:

aus: J. WALIŃSKI, Acquisition of terminology at the phrase level from comparable corpora: LEWANDOWSKA-TOMASZCZYK, B (ed.), Practical Applications in Language and Computers. Łódź Studies in Language 9. Frankfurt/M 2004. S. 33-47.

(35) **Keyness**

Keyness is a statistical method of extracting lexical items with high relevance to the topic of text. The term *keyness*, though it is in common use, is not well defined in Linguistics (Scott 1998: 65). It was implemented into computational algorithms of language processing by Scott, whose WordSmith Tools (OUP 1998) identifies *keyness* of words on a mechanical basis by comparing patterns of frequency existing between two corpora.

The *Keyness* score (or K-score) of word *x* is calculated by comparing the frequency of that word in the target corpus with that of the same word in a much larger reference corpus, taking also into consideration the total numbers of running words (tokens) in both corpora. The "key words" are calculated by comparing the frequency of each word in the target corpus with the frequency of the same word in the reference corpus. Therefore, "key words" are not the most frequent words, but the words, which are most *unusually* frequent in target corpus in comparison to the reference corpus. ...

(36) **Stop lists**

In order to refine the scope of acquisition to contain only relevant terminological items, the results obtained from the *keyness* algorithm can be filtered out with the technique of *stop lists*. The stop list (Scott: 1998: 49) is a list of words that should be excluded from the computational analysis of language data. For example, the key words analysis may omit common function words like *the*, *of*, *was*, *is*, *it*, etc., which are highly unlikely to constitute terminology. The stop list file must be prepared first specifying all lexical items to be ignored in processing. If the stop list is activated, the words in it will be filtered out of the *keyness* analysis. ...

(37) In the outcome stop lists containing separate items for English and for Polish were obtained including clusters composed mostly of function words, pronouns and modals. For Polish the list contained 341 clusters and 274 for English. Examples of top items in these lists for English are: *of the*, *in the*, *to the*, *and the*, *it is*, *to be*, *of a*; ...

Lemmatization

Lemma is "a set of graphic words having the same stem and/or meaning and belonging to the same grammatical word class, differing only in inflection and/or spelling." Thus, *go*, *goes*, *going*, *went* and *gone* from (!) a lemma or a lemma group.

Lemmatization is based on assigning each of morphosyntactic forms of the word the same tag, in order to indicate that they all belong to one lemma. In English, lemmatization may be considered somewhat redundant, since English is a language with simple inflectional morphology (Garside, Leech and McEnery 1997). In the process of lemmatization all the inflected forms of verbs and plural forms of nouns are lemmatized to their base forms, but

adjectives and adverbs like *interesting* and *interestingly* are not lemmatized to the noun *interest*, because they belong to different word classes. The present and past forms of modals are considered to constitute separate word types. In Polish lemmatization plays much more important role because of its rich inflection. The process is similar - nouns are lemmatized to their singular nominative forms, verb forms are reduced to infinitives.

Lemmatization can be performed automatically using the automatic lemmatization function of the WordSmith Tools (Scott 1998, <http://www.liv.ac.uk/~ms2928/>). To perform this operation a lemma dictionary has to be provided by the user. Fortunately the above tools gives (!) a possibility of performing a dynamic lemmatization procedure based the on-screen point-and-drag procedure in the WordList module. Using this procedure, the lemma list for terminological items can be generated dynamically and saved in a standard text file which groups required words together (*celeron* → *celerona*, *celerony*, etc.). Using this list WordSmith lemmatizes the items automatically. The list can be later re-used for projects concerning similare domains of language, which automates the matching and joining process. ...

(39) Automatic alignment of equivalent phrases

After the word clusters had been extracted from the corpus, the key question was how effective the comparable corpus can be in producing accurate bilingual equivalents lists with no human intervention at the phrase level. The alignment algorithm is based on the assumption that the comparability of texts from a restricted sub-language implies the existence of common terminology. A natural extension of this hypothesis is that not only the same terms are present, but also their distribution in both languages is similar, i.e. they occupy similar position on the keyness lists. Observation of the empirical data obtained in the terminology extraction procedure discussed above confirms the above claim.

The procedure of automatic acquisition of bilingual equivalents follows the algorithm developed by McEnery et al. (1997) for alignment of corresponding terms in parallel corpora. The algorithm (they used to perform terminology acquisition from the parallel corpus) is based on a naive window matching algorithm, which "takes all n -sized sequences from one text, and compared them to all potential m -sized regions in an aligned region of the parallel text." The string similarity between the lexical items existing in these regions is then computed to determine equivalent terms. The n and m regions, i.e. *windows of comparison*, are selected on the structural similarity existing potentially between the texts in the corpus. The windows with the highest similarity are deemed to be the best translations. ...

(40) String similarity

Evaluation of string similarity can be used for different tasks in computational linguistics. One application is the identification cognates from bilingual texts in the case of similar character sets between the languages under consideration. As stated by Tiedemann (1998): "Applications to technical texts are especially profitable because of internationalisation and similarities in the origin of technical terminology. String matching algorithms are used to compare word pairs and in practice even simple algorithms represent sufficient measures for string similarity."

zu Ziff. 4.425:

Type - Token - Ratio (TTR)

zu:

aus: J. TULDAVA, Stylistics, author identification in: KÖHLER, R u.a. (eds.), Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005

(362) Up to now types of TTR-measurement were proposed.

- (i) Ascribing to the position of each token (x) the number of types (y) up to this position. This is most usual practice yielding a monotone increasing series <token, type>.
- (ii) Dividing the number of types up to position x by the position itself yielding a not monotone decreasing function <token, type/token>. There are some variants of this method considering not each token separately but pooling them e.g. by hundreds in order to achieve monotonicity.
- (iii) The Köhler-Galle method (Köhler, Galle 1993) using the formula:

$$TTR_x = \frac{t_x + T - \frac{xT}{N}}{N}$$

where x = position in text (number of tokens up to position x), t_x = number of types up to position x (inclusively), T = number of types in the whole text, N = text length (number of tokens in the whole text), yielding rather a fractal which is even optically very characteristic of the text. Here we have < x, TTR_x >.

Let us exemplify these methods using a letter written by F. Schiller in Figure 27.1.

[FOLIE]

The words are in the first column, the tokens (positions) (x) in the second. In the third column the number of types up to position x given (Method (i)); in the fourth column the data according to Method (ii), i.e. types/tokens is given; in the last column one finds the numbers according to Köhler-Galle Method (iii); since there are $T = 80$ types and $N = 100$ tokens in the text, we obtain here e.g. $TTR_1 = [1 + 80 - 1(80)/100]/100 = 0.802$, $TTR_2 = [2 + 80 - 2(80)/100]/100 = 0.804$ etc. The series are displayed in Figure 27.2 to 27.4.

For modelling purposes evidently Method (i) is the most feasible but it does not mean that the other ones, especially Method (iii), reminding us of renewal processes, of avalanches in sand piles and of theory of runs, would not bring another special insight. But up to now, nobody tried to follow this way.

4.43 Wortketten / *distant-reading*zu Ziff. 4.431:Häufige Wörter - häufige Wortketten

zu:

aus: HOOVER, D L, Frequent Word Sequences and Statistical Stylistics: Literary and Linguistic Computing 17/2 (2002) 157-180.

(157) The search for a stylistic wordprint that characterizes the texts written by a given author has often focused on the frequencies of very frequent words. The ground-breaking work of John F. Burrows on Jane Austen (1987) convincingly demonstrates that the frequencies of words such as *the*, *and*, *of*, *a*, and *to*, which intuitively seem insignificant both semantically and stylistically, can nevertheless be used to distinguish authors, novels, and even characters within a single novel from each other, and be shown to have interesting and significant stylistic nuances. This is remarkable, especially when we remember that these five little words typically account for nearly 20% of all the word tokens in a novel. Precisely because such words are so frequent in English and so unlikely to be consciously regulated by authors, their frequencies may reflect authorial habits that remain constant in spite of differences in subject matter, point of view, or theme...

(158) Frequent word sequences are, of course, almost always less frequent than their constituent words, nevertheless, they, like frequent words, also seem likely to escape the conscious regulation of authors.

(160) Word sequences must be distinguished from phrases, which are complete linguistic units (prepositional phrases, noun phrases). As defined here, word sequences form a subset of collocations: those with a distance of zero words between the words under consideration.

[Beispiele aus einer Analyse, Häufigkeit in Klammern; 162] *of the* (5930), *in the* (4246), *to the* (2457), *it was* (2195), *he was* (2071), *and the* (2065), *on the* (2059), *he had* (1935), *at the* (1745), *in a* (1432) ...

zu Ziff. 4.432:

Textinterne (Ausdrucks-)Syntax: Alliterationen

zu: vsyst4.33 vgt5.231

aus: Alliterationen in G. Grass "Die Blechtrommel" (Studienarbeit Simone Capliuk) - Einstellung: ein nicht-zutreffender Befund ist erlaubt.

1 Buchstabe/Wortanfang - Wortkette: 3

---bett bunt bemalt
---durch dick [und] dünn
---ziegelsteinrückstände zwischen [den] zähnen

1 Buchstabe/Wortanfang - Wortkette: 4

---glücklich genug gewählt gewesen
---verhältnissen vertraut [zum] verpacken [der]
viertelpfundkleckse

1 Buchstabe/Wortanfang - Wortkette: 5

---schlagsahne so süß so süß
---sauerkraut seine schweinenieren [in] senfsoße seine

1 Buchstabe/Wortanfang - Wortkette: 6

---der dreijährige [gnom] der däumling der
---sieben sieht [meine] strukturellen schräginformationen sagt sich

1 Buchstabe/Wortanfang - Wortkette: 7

---schreiben spielen schönschreiben sonnabend [rechnen] singen spielen spielen
---matzerath mühte [sich] meine mama mit mir meiner

2 Buchstaben/Wortanfang - Wortkette: 3

---glas glas gläschen
---invaliden instrumente interessieren
---geleckt gewaschen [und] gekämmt

2 Buchstaben/Wortanfang - Wortkette: 4

---setzen se [wenn] se setzen
---ein eimer [ist] ein eimer

2 Buchstaben/Wortanfang - Wortkette: 5

---gehäkeltes gestricktes [besticktes] geflochtenes geknotetes geklöppeltes
---der deibel dem [beißt] der [aff] denn

3 Buchstaben/Wortanfang - Wortkette: 3

---schluß schon [morgen] schluß
---verdunkelte vergrößerte [und] verkleinerte

3 Buchstaben/Wortanfang - Wortkette: 4

---schwester scholastika scholastika [agneta] schwester
---einmal eine [tat] eine einzigartige

4 Buchstaben/Wortanfang - Wortkette: 3

---zerschrie zersang zerscherbte
---hoch hoch [wie] hochwürden

5 Buchstaben/Wortanfang - Wortkette: 3

---kreuzundquer kreuzverhör kreuzworträtsel

6 Buchstaben/Wortanfang - Wortkette: 3

---ziegel ziegelbrenner [und] ziegelstreicher
---kinderjäckchen kinderlätzchen kinderhöschen

6 Buchstaben/Wortanfang - Wortkette: 4

---gleichgroßartiger gleichkühner gleichwertiger [oftmals]
gleichformatiger

1 Buchstabe/Wortende - Wortkette: 3

---oft spielt [er] mit
---grünlichem Schleim kam

1 Buchstabe/Wortende - Wortkette: 4

---polizeipräsidium am [wiebenwall] nahm
---rhythmus dieses liedchens

1 Buchstabe/Wortende - Wortkette: 5

---glühenden steinen an allen fingern
---mea culpa [ruft] mea maxima culpa
---leicht erotisiert [und] schmackhaft macht soweit

1 Buchstabe/Wortende - Wortkette: 6

---gucker hinter der tür [sobald] er [mein] zimmer
---nicht lügt [mir] nicht grünt [mir] nicht [grün] ist

1 Buchstabe/Wortende - Wortkette: 7

---verschieden großen glocken von verschieden hohen kirchtürmen
---des schiffahrtsmuseums [ein] altes gleichfalls museales patrizierhaus das

1 Buchstabe/Wortende - Wortkette: 8

---in den folgenden jahren einen braven sparsamen [leicht] stotternden
---wer hier der vater [ist] jener herr [matzerath] oder [ich] oskar

1 Buchstabe/Wortende - Wortkette: 9

---hätte eine ganze herrliche schönfenstrige lichtfangende lichtbrechende gotische kathedrale
---legten den ohnehin morschen lattenzaun [flach] banden jenen halblinden den

1 Buchstabe/Wortende - Wortkette: 10

---daumensprünge weiblichen proportionen nachsetzten [und] in den klassischen [acht] kopflängen ein [maß] sahen
---augen gehörten den erleuchteten vierecken an dann [und] wann [glitt] ein schatten

1 Buchstabe/Wortende - Wortkette: 11

---rechnen schreiben religion [freitag] rechnen schreiben spielen schönschreiben [sonnabend] rechnen singen spielen spielen
---schweigen von meinen [für] einen lächerlichen [preis] in vierzehntägigen abständen [über] den [ladentisch] gereichten blechtrommeln

1 Buchstabe/Wortende - Wortkette: 12

---war der vikar [rasczeia] hinter mir [mama] hinter mir alter [mann] hinter mir [und] der vikar
---denn schon nach wenigen minuten wenn man in einem stockdunklen [zimmer] von minuten sprechen kann
---sonnengebräunten kindern flauschigen bademänteln bunten strandbällen [und] segelschiffen stiegen [aus] straßenbahnen [die] von den seebädern

1 Buchstabe/Wortende - Wortkette: 13

---gemälden standuhren badewannen radioapparaten [und] garderobeständern [aufgebaut] radelten dazwischen achten schnecken spiralen wichen gegenständen [wie] kinderwagen

1 Buchstabe/Wortende - Wortkette: 14

---zwischen den gleisen hackten [und] schaufelten ostarbeiterinnen [mit] weißen kopftüchern [wir] standen [und] guckten in den schattigen kühlattenden

1 Buchstabe/Wortende - Wortkette: 16

---zwischen den flößern zwischen den heizern [und] flößern
zwischen steuermann heizern [und] kapitän zwischen [dem] kapitän
[und] den [ständig] wechselnden lotsen

2 Buchstaben/Wortende - Wortkette: 3

---kauend und [gerstenkaffee] schlürfend
---geht nicht [wirklich] nicht
---raus aus [dem] katholizismus

2 Buchstaben/Wortende - Wortkette: 4

---hustete weinte lachte [und] konnte
---tagtäglich rauch ich mich

2 Buchstaben/Wortende - Wortkette: 5

---stöckchen bekamen [doch] mitten [im] kauen besannen
---immer wieder [symmetrisch] angeordneter gemischter [chor]
entweder

2 Buchstaben/Wortende - Wortkette: 6

---männeraugen gesehen hätten einen scheuen [aber] schwieligen
---der forster [mit] übler bayrischer [aussprache] immer wieder

2 Buchstaben/Wortende - Wortkette: 7

---aufkommen lassen [und] spielen [wir] knickten falteten
zerschnitten [mit] scheren
---hecken büschen obstbäumen aufwölkten den [mercedes]
beschatteten gegen

2 Buchstaben/Wortende - Wortkette: 8

---ausgesuchtesten qualen darin bestehen den nackten menschen
[mit] den gerahmten
---marken fettmarken nahrungsmittelmarken reise marken
schwerarbeitermarken schwerstarbeitermarken [und] zwischen [die]
marken

2 Buchstaben/Wortende - Wortkette: 9

---den ganzen langen [an] einigen stellen [schon] knötchen
zeigenden fadenscheinigen faden
---zwischen den häßlichen [und] begabten den hübschen [aber]
schwatzhaften mädchen anzustellen

2 Buchstaben/Wortende - Wortkette: 10

---den trommelnden [permanent] dreijährigen [oskar] neben den
buckligen tränen [und] trommellosen [oskar] stellen

2 Buchstaben/Wortende - Wortkette: 11

---betonten vornamen [der] heiligen [unter] den einfalllosen
fragen [und] rauchgetrübten blicken [zweier] landgendarmen [vom]
kleinen [aber] breiten

3 Buchstaben/Wortende - Wortkette: 3

---bremsten drehten stiefelten
---verdienst kunst [und] verdienst

3 Buchstaben/Wortende - Wortkette: 4

---gewässerten gekochten gewürzten und abgeschmeckten
---mich wirklich ich [liebe] mich

3 Buchstaben/Wortende - Wortkette: 5

---gehäkelt gestricktes besticktes geflochtenes geknotetes
geklöppeltes

3 Buchstaben/Wortende - Wortkette: 6

---schnatterten brüllten lachten weinten [und] tobten [meine]
sogenannten

4 Buchstaben/Wortende - Wortkette: 3

---hungriger dreieckiger puppiger
---erregender schlafvertreibender [und] gedankenfördernder

4 Buchstaben/Wortende - Wortkette: 4

---wirsingkohl rosenkohl rotkohl [und] weißkohl
---abgegriffenen abgesehenen durchgelegenen [und] angestoßenen

5 Buchstaben/Wortende - Wortkette: 3

---quietschender stockender verrutschender
---rechtstadt altstadt pfefferstadt

5 Buchstaben/Wortende - Wortkette: 4

---liegenden sitzenden kriechenden laufenden
---gattinnen freundinnen sekretärinnen kunstgewerblerinnen

5 Buchstaben/Wortende - Wortkette: 5

---finnischen schwedischen polnischen freistädtischen [und]
reichsdeutschen

6 Buchstaben/Wortende - Wortkette: 3

---bücher abenteuerbücher liederbücher
---wiederkreuze krückenkreuze [und] stufenkreuze

6 Buchstaben/Wortende - Wortkette: 4

---finnischen schwedischen polnischen freistädtischen

zu Ziff. 4.4331:

(Ausdrucks-)Syntax/textintern: Schüttelreime

zu:

aus: M. HEPERLE, Gestammelte Werke. 1979.

Schlachtfest *Schüttelreime*

Der Bauer sagte: Meine Schwester
ist doch die Frau vom Schweinemäster.
Gleich hinter einem Steineschwall
befindet sich der Schweinestall,
wo Kinder oft mit rauhen Sitten
auf meines Schwagers Sauen ritten.

Nach einem Jahr, da macht er Schluß,
weil da die Sau zum Schlachter muß.
Zuvor kommt sie ins Bade meist,
damit sie keine Made beißt.

Das Tier dann voll Verachtung schlendert,
was auch nichts an der Schlachtung ändert.

Ob diese fette Lose dachte,
daß sie mal aus der Dose lachte?

Was kann man bei den Tieren nun
mit Lebern und mit Nieren tun?
Das Fleisch bis hin zur Bauchfleischrinde
wird sicherlich zur Rauchfleischbinde.
Und im Kamin am Speichereck,
da hängt bereits der Räucherspeck.

Die Bäuerin, die am Magen leidet,
das Fett von allen Lagen meidet.
Der Bauer ißt die Lagen mit,
weil er noch nie am Magen litt.

Der Metzger, der die Schürze wendet,
weil sie ein Fleck aus Würze schändet.
Bald wird man von den festen Rinden
nur Teile noch von Resten finden.

Die Frau in einer Schüssel rührt
und Feuer für den Rüssel schürt.
Die Kinder schauen sauer drein
sie werden voller Trauer sein.
Der Bauer sagt: O Schwester mein,
manchmal hat auch ein Mäster Schwein.

4.44 Autoridentifikation

Literatur

- BURROWS, J: All the Way Through: Testing for Authorship in Different Frequency Strata. *Literary and Linguistic Computing* 22 / 1 (2007) 27-47.
- GRIEVE, J: Quantitative Authorship Attribution: An Evaluation of Techniques. *Literary and Linguistic Computing* 22 / 3 (2007) 251-270.
- HIRST, G; FEIGUINA, OL'GA: Bigrams of Syntactic Labels for Authorship Discrimination of Short Texts. *Literary and Linguistic Computing* 22/4 (2007) 405-417.
- HOOVER, D L: Testing Borrows's Delta. *Literary and Linguistic Computing* 19 / 4 (2004) 453-475.
- HOOVER, D L: Delta Prime?. *Literary and Linguistic Computing* 19 / 4 (2004) 477-495.
- HOOVER, D l; HESS, S: An exercise in non-ideal authorship attribution. the mysterious Maria Ward. *Literary and Linguistic Computing* 24 (2009) 467-489.
- JOCKERS, M L; WITTEN, D M u.a.: Reassessing authorship of the *Book of Mormon* using delta and nearest shrunken centroid classification. *Literary & Linguistic Computing* 23/4 (2008) 465-491.
- JUOLA, P; BAAYEN, R H: A Controlled-corpus Experiment in Authorship Identification by Cross-entropy. *Literary and Linguistic Computing* 20 (2005) 59-67.
- LOVE, H: *Attributing Authorship: An Introduction*. Cambridge 2002
- MERRIAM, T: Untangling the derivatives. points for clarification in the findings of the Shakespeare Clinic. *Literary and Linguistic Computing* 24 (2009) 403-416.
- MIRANDA GARCÍA, A; CALLE MARTÍN, J: Function Words in Authorship Attribution Studies. *Literary and Linguistic Computing* 22 / 1 (2007) 49-66.
- TAMBOURATZIS, G; VASSILIOU, M: Employing Thematic Variables for Enhancing Classification Accuracy Within Author Discrimination Experiments. *Literary and Linguistic Computing* 22 / 2 (2007) 207-224.
- TEARLE, M; TAYLOR, K u.a.: An algorithm for automated authorship attribution using neural networks. *Literary & Linguistic Computing* 23/4 (2008) 425-442.

zu Ziff. 4.441:

Inhaltsfreie Autoridentifikation

zu:

aus: HOORN, J F; FRANK, S L u.a.: Neural Network Identification of Poets Using Letter Sequences: Literary and Linguistic Computing 14/3 (1999) 311-338.

(311) Stylistic differences among poets are usually sought in sound and semantics. In human analysis, the criteria for recognizing stylistic differences are manifold and intermingled. This study demonstrates that successful identification of poets based on their work is possible using one criterion: letter sequences. Poets show preferences for certain letter combinations, which are unique to their writing style. Using this criterion in machine computation demonstrates that semantics are not needed to identify poets correctly, and that, as a concession to utter parsimony, one minimal criterion of unique letter sequences is enough to fingerprint an author. A small sample of the work of three Dutch poets was used: Bloem (1887-1966), Slauerhoff (1898-1936), and Lucebert (1924-94). This sample formed the training set for the neural network program to analyse the unique letter patterns for each poet. Next, the program was fed a set of new poems, for which the author was to be identified. In choosing between two poets, the program succeeded in identifying the poet correctly for 80-90 per cent of the new poems. When the choice was between three poets, the score was ? -70 per cent correct. Since raw ASCII files are sufficient as input, and human pre-coding is unnecessary, neural network analysis of letter sequences may turn out to be a powerful tool in categorization and identification problems, such as genre, stylistics, and plagiarism.

zu Ziff. 4.442:

Wortstatistik: Stilanalysen

zu:

aus: Literary & Linguistic Computing 14/4 (1999)

- 1) BINONGO, J N G u.a.: The Application of Principal Component Analysis to Stylometry. S. 445-465.
→ Beschränkung auf das Vorkommen von "two marker words": *and*, *they*
- 2) FORSYTH, R S: Stylochronometry with Substrings, or: a Poet Young and Old. S. 467-477.
→ Am Bsp. von Gedichten (Yeats) werden mit "a quasi-random search algorithm" "marker substrings" erhoben und den Lebensperioden zugeordnet.
- 3) MEALAND, D L: Style, Genre, and Authorship in Acts, the Septuagint, and Hellenistic Historians. S. 479-505.
→ Partikeln (alla, kai, gar, de, oun / "aber, und, aber, denn, deshalb"), Konjunktionen (ei, ean, 'ews, 'ina, 'opou, 'opws / "if, if (fut.), while / until, in order that, where, so that"), Präpositionen (apo, eis, ek, en, dia, kata, meta, perim epi, pros / "from, into, out of, through / on account of, according to / against, with / after, around /concerning, on, to"). - Zweck: den Stil unterschiedlicher Autoren unterscheiden.
- 4) CONSTABLE, J u.a.: Word Length Frequency and Distribution in English: Part II. An Empirical and Mathematical Examination of the Character and Consequences of Isometruic Lineation. S. 507-535.
→ Zweck: Unterscheidung von Poesie und Prosa.

zu Ziff. 4.443:

Rückschluss: Text → Autor

zu: vprag7.5125

aus: E. Kleinschmidt, Autor und Autorschaft im Diskurs in: BEIN, T; NUTT-KOFOOTH, R; PLACHTA, B (eds.), Autor - Autorisation - Authentizität. Beihefte zu editio 21. Tübingen 2004. S. 5-16.

(5) In einer der letzten Erzählungen E.T.A. Hoffmanns, *Des Vettters Eckfenster*, findet sich die Episode vom Blumenmädchen, das der Schriftsteller-Vetter auf dem Markt als Leserin eines seiner Bücher antrifft. Er erbittet ihr Urteil darüber. Es mündet darin, man fühle sich bei der Lektüre des Werkes so, "als wenn man mitten darin säße". Stolz gibt der Vetter sich als Verfasser zu erkennen: "Hier, mein süßer Engel, hier steht der Autor des Buchs, welches Sie mit solchem Vergnügen erfüllt hat, vor Ihnen in leibhaftiger Person". Seine Erwartung zielt auf des Mädchens Interesse am "sublime(n) Genie, dessen schaffende Kraft solch ein Werk erzeugt" habe. Statt dessen geschieht etwas ganz anderes. Jegliche Reaktion bleibt aus, so daß dem Autor als "Schmach" bewußt wird, "daß das Mädchen niemals daran gedacht, daß die Bücher, welche sie lese, vorher gedichtet werden müßten. Der Begriff eines Schriftstellers, eines Dichters war ihr gänzlich fremd". Ironisch merkt der derart Verkannte nur noch an, er habe nicht mehr weiter nachgefragt, denn dann wäre vermutlich "der kindliche Glaube ans Licht gekommen, daß der Liebe Gott die Bücher wachsen ließe wie die Pilze". ...

(7) Die kommunikative Schnittstelle zwischen "Autor" und "Hörer/Leser" ist der Text, dessen Herausforderung, ihn zu verstehen, eine wie auch immer verortete Autordimension evoziert. Nicht die abstrakt gesetzte Idee eines Autors generiert die Vorstellung von Textualität, sondern die konkrete Existenz von Texten bringt die Frage nach dem Urheber als Möglichkeit, nicht einmal als Notwendigkeit hervor. Insofern ist der Autor ein spezifisches Phantasma im anthropologischen Kontext menschlicher (8) Kommunikations- und Formulierungsfähigkeit. Erst das Vorhandensein von Rede bringt auch den Redner ins Spiel, den, wenn er schwiege, niemand auch als Autor identifizierte. Die Entdeckung des Autors beruht primär auf der Frage nach dem Ursprung, nach den Regeln der Textgenese, von der zu wissen, gegebene Textualität legitimiert und strukturell erläutert...

(11) Was einem derart gedacht entgegentritt, subvertiert das traditionelle auktoriale Subjekt und das zentrierende Prinzip seiner Repräsentation im Text, so daß es nicht mehr zurücklesbar ist. Der Angriff macht aber zugleich den Sachverhalt stark, daß die Wirkungsrealität von Texten keinen Autorgaranten braucht, um zu funktionieren. Sie tritt als agierend auf, ohne daß auf einen personalen Agenten zurückgegriffen werden muß. Der Text erscheint so als sein eigener "Autor".

(16) Autorschaft hingegen wäre die unabweisbare Begriffssphäre von ermöglichender Produktivität. Sie ließe sich theoretisch aufspalten in zwei Strukturfelder der organisierenden Betrachtung. Das eine stellt sich als der rahmende Bereich des Prädispositiven dar, der sich mit dem Eintritt in die konkrete und überlieferungsfähige Textualität berührt, aber gerade nicht mit ihr identisch ist und nicht in ihr aufgeht. Er umfaßt vertraute formative Vorgabefaktoren wie Sprache, Textgenre, Stoff, Motivik usw., die als Bereitstellungspotentiale transsubjektiv fungieren.

Der andere Erfassungsbereich betrifft die spezifisch aktivierten Spielräume innerhalb dieser Vorgaben, die affirmativ ausgefüllt oder aber auch durchbrochen werden können. Dieser Bereich hängt sowohl von unbewußten als bewußten Eingriffen ab, bei denen ein gestaltendes Subjekt beteiligt, aber eben nicht allein dominant ins Spiel kommt. Diese Sphäre stellt eine deskriptive Irritation dar, da sie in eine theoretische Systematik von Autorschaft ein personales Substrat einführt. Dieses gilt jedoch nur scheinbar und ist von bedingter (15) Präsenz, da sowohl unbewußt wie bewußt konstituierte Schreibentscheidungen in ein allgemeines, überindividuelles Formulierungssystem eingebettet bleiben.

(16) Das Berliner Blumenmädchen E.T.A. Hoffmanns muß in der Tat nichts wissen vom Autor der Texte, die sie liest. Seine Autorschaft, die sie wahrnimmt und mit der sie umzugehen weiß, genügt, mag dies für den Autor auch eine schmerzliche Erfahrung sein.

zu Ziff. 4.444:

Worddistanzen, -erwartungen: Autoridentifikation

zu:

aus: JUOLA, P; BAAZEN, R H, A Controlled-corpus Experiment in Authorship Identification by Cross-entropy: Literary and Linguistic Computing, Vol 20, Suppl Issue, (2005) 59-67.

(59) The basic approach is to identify, by close inspection, a stylistic "fingerprint" characteristic of the author, and then determine whether this fingerprint is also present in a disputed work. Recent scholarship has tended to focus on specifically distributional "fingerprints" to be identified by computerized statistical analysis of the input texts.

(60) With this wide variety of techniques available, it is important and yet very difficult to compare the power and accuracy of different techniques. A fingerprint appropriate to distinguish between Jack London and Rudyard Kipling, for example, may not work to distinguish between Jane Austin and George Eliot.

(61) In information theory (Shannon, 1948, 1951), "entropy" is simply a measure of the unpredictability of a given event, given all relevant background information that could be brought to bear. "Cross-entropy" is a measure of the unpredictability of a given event, given a specific (but not necessarily) model of events and expectations. A person completely familiar with 20th century English may still find Shakespeare somewhat daunting, an effect of three centuries of language drift, but will be more comfortable than a German speaker with no English knowledge whatsoever. This difference can be quantified and measured as a "distance" between two samples.

As a general linguistic difference, this general technique has several advantages over other available work such as those of Burrows (1992a); Holmes (1998); Baayen et al. (1996). First, it seems to be widely applicable to a variety of linguistic and text-analysis problems. Second, the word "distance" is here used in its exact sense as a numerical measure that can be compared with other similarly scaled "distances" measured from unrelated documents. Third, the method is relatively parsimonious of input text, enough so that measurements of useful precision can be made from small samples; as will be discussed, the authorship of a disputed document can be determined using less than a page of data. Fourth, this technique is sensitive to all levels and aspects of language variation...

(62) As a quick summary, the first document is used as a sample from which one can make informed guesses about the next letter, word, grammatical construct, topic, etc. in the second document. The more closely linked the two documents are, the more accurate the guesses will be - just as someone who knows English well can predict that for any document written in English, the word "the" will be more common than the word "gryphon". This notion of accurate guessing can be accurately computed as a linguistic distance, where a low number implies accurate guessing, and therefore two documents close in all aspects. By contrast, a high number implies inaccurate guesses, implying in turn a substantial and significant difference in some way, be it language, authorship, topic, style, genre, date, or any other aspect.

zu Ziff. 4.445:

Erwartbare Worthäufigkeit: Autoridentifikation

zu:

aus: HOOVER, D L, Testing Burrows's Delta: Literary and Linguistic Computing: 19/4 (2004) 453-475.

(453) Burrows uses the frequencies of the 150 most frequent words of the entire set of texts his exposition of the method. After the frequencies of all these words in all of the texts are collected, he calculates the mean frequency for each word in the entire set and compares it with the word's frequency in a test (454) text and in the selection by one of the authors in the primary set. The result is two differences that are then compared with each other.

Delta is a relatively simple measure of difference, but its calculation and interpretation are not very transparent. In the interest of clarity, it seems worthwhile to trace through an example. Consider the word *of*, which is the third most frequent word in the texts that Burrows analyzes, with a mean frequency of 1.821 (presumably, this is its percentage of all the tokens) (2002a, p.272). For *Paradise Lost*, its frequency is 2.769, and for the selection by Behn from the main set, its frequency is 1.783. Delta compares how different the two texts are from the mean of the corpus, and here those differences are .948 for *Paradise Lost* and -.038 for Behn's selection, showing that Milton uses *of* much more frequently and Behn slightly less frequently than the mean for the corpus. Given how quickly word frequencies drop from the most frequent words to the hapax legomena, it is important to convert these absolute differences in frequency to z-scores (by subtracting the mean frequency of the word in the corpus from its frequency in the test text and dividing this difference by the standard deviation of the word in the corpus). This transforms the raw word-frequency information into a measure of the distance (in standard deviations) of each frequency from the mean frequency for the corpus, and shows that *Paradise Lost* is 3.015 standard deviations above the mean, and Behn's selection is .121 standard deviations below the mean. When Behn's z-score is subtracted from the z-score for *Paradise Lost* to determine the difference between the differences from the mean, the result is 3.136, showing that *of* is used very differently in *Paradise Lost* than it is in the works of Behn. When this procedure is followed for all 150 words, the result is a list of differences between the differences of the two texts from the mean, and the stage is set for the final step in the calculation of Delta. Because Burrows is interested in the pure differences between the differences, he eliminates their signs before calculating their mean. This result is Delta: "the mean of the absolute differences between the z-scores for a set of word-variables in a given text-group and the z-scores for the same set word-variables in a target text" (2002a, p.271). After Delta is calculated for each pair of texts, the primary author who shows the smallest mean differences from the test text, the smallest Delta, is the likeliest author of the text.

In spite of the simplicity of Delta and the fact that its calculation systematically removes information about the direction of the differences, it is remarkably effective in identifying authors in a difficult "open" test. When texts by 16 authors who are members of the original set and texts by 16 other authors are tested with Delta to determine likely authorship, "Of thirty-two long poems... fifteen are correctly identified and another 15 yield scores that correctly place them outside the main set"

(Burrows, 2003, p.15). Although these results are not completely accurate, they are very encouraging, suggesting that Delta can be a powerful tool in the early stages of authorship studies in which there are many possible claimants.

4.45 Sprachbarrieren

zu Ziff. 4.451:

Soziolinguistik / Sprachbarrieren

zu:

aus: DeFOREST, M; JOHNSON, E: The Density of Latinate Words in the Speeches of Jane Austen's Characters: Literary and Linguistic Computing 16/4 (2001) 389-401.

(389) The English language is largely made up of two different language groups, Latinate and Germanic. The choice of a word from either vocabulary has rhetorical implications. The gulf between these two vocabularies falls along with an ancient fault line, which, in classical antiquity, divided the Greek from the barbarian; the aristocrat from the plebeian; reason and self-control from passion; culture from nature, male from female. This fault line is preserved in "diglossia", a phenomenon of language described by C.A. Ferguson (1959). Ferguson showed that certain languages, such as Arabic and Modern Greek, are divided by grammar and vocabulary into two languages; one written, one spoken; one elevated, one popular. Latinate words in English resemble the dialects of the higher status. Its polysyllables are suitable for elevated discourse, concerns, and social class. Germanic words go with the popular dialect, spoken by people closer to the soil, with their earthier behaviour and talk.

This study describes how the diction of Austen's characters conforms to the expectations about Latinate vocabulary set out in the previous paragraph. Latinate densities indicate the components of character: social class, gender, education, mind, and mood. With Austen's judicious blending of Latinate and Germanic words, her characters reveal who they are.

(392) After the Norman Conquest in 1066, Britain was inhabited by two societies, each speaking a different language. The aristocrats spoke French and the peasants Anglo-Saxon. This linguistic apartheid lasted until the fourteenth century, when the two languages coalesced to form English. French was a Romance language, which means that it is largely composed of words derived from Latin.

Anglo-Saxon contributed prepositions, conjunctions, pronouns, and demonstratives - the "function words" - as well as words for things and actions in the real world. Latinate words tend to be either abstractions or modifications of natural things. Long ago, Sir Walter Scott showed, in *Ivanhoe*, that English had separate words for the animal and the meat of the animal (Scott, 1941: 36). The peasants, who spoke Anglo-Saxon, took care of the animals, which the speakers of French ate. *Beef*, *mutton*, and *pork* are derived from the French words for cow, sheep, and pig. In Modern English, the animal in its natural state goes back to an Anglo-Saxon word, whereas the word for an animal prepared for human consumption goes back to a French word. Nature and culture are perfectly opposed in the contrasting vocabularies, one Germanic, the other Latinate.

(393) Austen uses Latinate and Germanic words in *Mansfield Park* to contrast the life in Sir Thomas Bertram's estate and the one in Fanny Price's parents' home in Portsmouth. Mansfield is described in Latinate language:

The elegance, propriety, regularity, harmony - and perhaps, above all, the peace and tranquillity of Mansfield, were

brought to her remembrance every hour of the day, by the prevalence of everything opposite to them here (MP 391).

The Latinate words convey the stateliness of the house (no doubt a fine, neo-classical structure) and the civilized behaviour of its inhabitants. Although Julia and Maria Bertram compete bitterly for Henry Crawford, they never raise their voices or make a rude remark. In Portsmouth,

"[E]very body was noisy, every voice was loud. ... Whatever was wanted, was halloo'd for, and the servants halloo'd out their excuses from the kitchen. The doors were in constant banging, the stairs were never at rest, nothing was done without a clatter ..." (MP 392)

4.46 Bewertung der Daten / Ausblick / Ähnlichkeit

zu Ziff. 4.461:

Quantitäten

zu:

aus: SCHENDERA, CFG, Die Erforschung der Verständlichkeit von normativen Texten. Eine kritische Darstellung von Modellen, Methoden und Ergebnissen. Zeitschrift für Sprachwissenschaft 19.1 (2000) 3-33.

(9f) Zählen der Häufigkeit bestimmter Merkmale

Sollte das Verstehen ermöglichende oder erschwerende Potential eines bestimmten Textmerkmals bereits bekannt sein, dann würde, rein theoretisch, die Auszählung dieser Faktoren genügen, um die jeweilige Verständlichkeit oder Unverständlichkeit eines Gesetzestextes abschätzen zu können.

Zu den bekanntesten Verfahren zählen u.a.:

- *Zählen ohne Vergleiche*: z.B. das Zählen bestimmter Modalformen (Matzke 1988)
- *Querschnittsvergleiche*: z.B. Vergleiche zwischen verschiedenen Arten von Gesetzen (Wagner-Döbler & Philipps 1993)
- *Längsschnittvergleiche*: z.B. Vergleiche mehrerer Gesetzesarten über die Zeit hinweg (z.B. Brandt 1996, 1988).

Für all diese Verfahren gilt: Das theoretische Auszählen von möglicherweise verstehenserschwerenden Text-Merkmalen ist kein Beweis, sondern nur die Auskunft, daß eben in diesem oder jenem Textsample bestimmte Merkmale in dieser oder jener Häufigkeit vorgefunden wurden. Weitere: Aussagen sind nicht zulässig. ...

Die Auftretenshäufigkeit eines Elements ist ein rein quantitatives Merkmal, das mit seiner qualitativen Dimension der "Verständlichkeit" in keinerlei direkten Kausalbeziehung zu stehen braucht. Die Analyse nur des einen Elements, nämlich Text, klammert das zweite Element, den Leser, aus. Um von der Auftretenshäufigkeit auf das Interaktionsresultat "Verständlichkeit" schließen zu können, sind empirische Untersuchungen mit Lesern notwendig. Faktisch fehlt in jedem Falle der Einzel-, Quer- und Längsschnittanalysen noch der Beweis, daß das bloße Vorhandensein oder eine bestimmte Anzahl besonderer Textmerkmale die Verständlichkeit eines Gesetzestextes tatsächlich erschwert.

4.47 Übersetzung ohne Bedeutungswissen?

Literatur

BRANTS, T: Statistisch basierte Sprachmodelle und maschinelle Übersetzung in: KALLMEYER,W; ZIFONUN,G (ed./eds.): Sprachkorpora - Datenmengen und Erkenntnisfortschritt. Jahrbuch des IdS 2006. Berlin 2007 235-248. 978-3-11-019273-5: de Gruyter // I.2.7

zu Ziff. 4.471:

(Ausdrucks-)Syntax: Übersetzen ohne Bedeutungswissen

zu: vgt1.11

aus: Der SPIEGEL 44/2005: Parlieren auf Inuktitut. Eine Karlsruher Forschungsgruppe entwickelt Computer, die als Simultandolmetscher für Vorträge und TV-Nachrichten dienen. Das Fernziel: ein tragbares Sprachgenie für jedermann. S.184-6.

[nachfolgend werden die Passagen, die uns algorithmisch interessieren mit größerem rechtem Rand gesetzt; die anderen, die mit Spracherkennung zu tun haben, mit normaler Breite]

(186) Mehr und mehr Forscher verzichten darauf, dem Computer etwas über Sprache beizubringen; sie zapfen lieber das Internet an. Das Netz weiß gewissermaßen am besten, wie man treffsicher übersetzt; es ist schließlich voll von Texten, die bereits in mehreren Sprachen zugleich vorliegen.

Abertausende Dokumente der EU und der Vereinten Nationen lagern auf frei zugänglichen Rechnern; es gibt mehrsprachige Meldungen von Nachrichtenagenturen und nicht zuletzt internationale Klassiker wie etwa die Bibel. Mit anderen Worten: Fast jeder Satz ist - so oder ähnlich - schon einmal übersetzt worden. Der Computer muss sich den Erfahrungsschatz nur noch aneignen.

Sprachverstand ist dafür nicht notwendig. Der Computer vergleicht einfach ein Textpaar nach dem anderen, Satz für Satz. Am Ende weiß er, wie oft das englische "bank" mit "Bank" übersetzt wurde und wie oft mit "Ufer"; er hat sich aber auch gemerkt, um wie viel die Chancen für "Ufer" steigen, wenn im Umfeld von Gewässern die Rede war; und er weiß dass dem oft ein "das" vorausgeht, fast nie ein "der".

So entsteht mit der Zeit aus dem sprachblinden Wörterzählen etwas, das aussieht wie ein Wissen um Zusammenhang und angemessene Rede. Der Computer hat quasi eine neue Sprache gelernt.

"Der Witz daran ist, dass alles automatisch ablaufen kann, ohne viel menschliches Zutun", sagt Tanja Schultz, eine Mitarbeiterin von Waibel in Pittsburgh. Kein Vergleich mit den bisher üblichen Verfahren. Linguisten mussten dem Rechner möglichst viele Regeln, Ausnahmen. Mehrdeutigkeiten mühsam von Hand beibringen. Das führte zu keinem Ende, und bei jeder neuen Sprache begann das Eintrichtern wieder von vorn.

Die Statistik dagegen braucht sich um Regeln und Weltwissen nicht zu kümmern. Hier zählt vor allem die Rechenleistung und die Masse des vorübersetzten Materials. Je mehr der Computer sich einverleibt, desto trefflicher werden, jedenfalls theoretisch, seine Sätze.

Noch hat die Karlsruher Software - wegen der vielen EU-Dokumente - eine gewisse Vorliebe für bürokratische Wendungen. Wörter wie "Subventionen"

oder gar "Verkehrswegebeschleunigungsgesetz" sind in ihrem Lexikon eindeutig überrepräsentiert. Aber das ist nur eine Frage der Zeit. "Wir haben die steigende Rechenkraft und wachsenden Datenmassen im Web auf unserer Seite", sagt Institutsleiter Waibel. "Ich glaube, in zehn Jahren werden es alle so machen."

Der Informatiker Franz Josef Och zum Beispiel, ein Pionier der statistischen Übersetzung, arbeitet inzwischen bei der Suchmaschinenfirma Google. Vor zwei Jahren war Och, gebürtiger Franke, bei einem Wettbewerb des US-Militär aufgefallen. Seine Software bewältigte schon nach kurzem Training Texte aus der Hindu-Sprache besser als alle Konkurrenten. Bei Google tut sich dem Statistiker nun ein Himmereich von Daten auf: 16 Milliarden englische Wörter in den Speichern der Suchmaschine stehen ihm zur Verfügung. Och kann seine Software damit in unzähligen Wendungen und Feinheiten des Gebrauchs trainieren. Das Ziel ist eine möglichst geschmeidige Übersetzung ins Englische.

Darauf darf Waibel in Karlsruhe noch nicht so bald hoffen. Sein Simultandolmetscher hat es mit gesprochener Sprache in all ihrer Schludrigkeit zu tun. Er muss sich mit stotternden Sprechern herumschlagen, mit Füllwörtern, abgebrochenen Sätzen und endlosen Wiederholungen: "Wir wollten uns doch, hören Sie mal, hatten wir nicht, äh, zehn vereinbart?" Dazu womöglich Husten und Türenschlagen im Hintergrund.

Das ist nicht eben die gepflegte Diktion, mit der die Computer heute schon erstaunlich gut zurechtkommen. Bei vielen Ärzten und Anwälten sind Diktiersysteme mit automatischer Spracherkennung im Einsatz. Das Programm Dragon Naturally Speaking vom Marktführer Nuance zum Beispiel verspricht, bis zu 99 von 100 Wörtern fehlerfrei zu erkennen. Das gilt aber nur in ruhiger Umgebung; auch sollte der Sprecher stets gleichen Abstand zum Mikrofon halten.

In der akustischen Wildnis eines Parlaments dagegen muss der Computer sehen, wie er zurechtkommt. Dort ist das Erkennen, geschweige denn das Übersetzen "entsetzlich schwer", gesteht Waibel. Die Software muss erst lernen, Unwichtiges auszufiltern und Satzbrüche zu erkennen. Übersetzt wird dann möglichst nur der Kern der bezweckten Aussage.

Auf anderen Feldern geht es schneller voran. In Erprobung sind bereits tragbare Geräte für Mediziner in Notstandsgebieten, die den beschränkten Wortschatz des Leidens und Verarztes ("Wo tut es weh?" - "Wir müssen das röntgen") beherrschen.

Die Forscher in Karlsruhe und Pittsburgh haben aber auch allerhand Apparate im Sortiment, die eher für Geheimagenten gedacht scheinen. Darunter ist eine Spezialbrille, in die die Übersetzung als Laufschrift eingeblendet wird. Ein anderes Gerät kann Stimmen mitten im Raum wie aus dem Nichts erzeugen. Es handelt sich um eine Platte, bestückt mit Dutzenden winziger Lautsprecher, die Ultraschall aussenden.

Die Wellen sind so gegeneinander versetzt, dass sie erst an einer vorausberechneten Stelle spukhaft hörbar werden. So ließe sich beispielsweise eine Übersetzung punktgenau auf den Hörer richten, ohne dass der Nachbar etwas mitbekommt - eine technische Meisterleistung, die nebenher die Frage aufwirft, ob es wirklich so schlimm ist, einen Kopfhörer zu tragen.

Auch eine erste Anwendung für den elektronischen Bauchredner ist angepeilt. Eines Tages, meint Forscherin Schultz, werde man sich die Sensoren ja vielleicht implantieren lassen, statt sie umständlich aufzukleben. Wer das wagt, verfügt nicht nur über eingebaute Vielsprachigkeit. Er ist auch imstande, ungeniert zu telefonieren. Meldet sich zum Beispiel im Theaterpl das Handy, kann der sensorbestückte Mensch das Gespräch einfach annehmen, und niemand ringsum hört einen Laut. DER SPIEGEL 44/2005 S. 186 (Manfred Dworschak)

zu Ziff. 4.472:

Statistische maschinelle Übersetzung / Phrasen

zu:

aus: Th. BRANTS, Statistisch basierte Sprachmodelle und maschinelle Übersetzung in: G. ZIFONUN, W. KALLMEYER (eds.), Sprachkorpora - Datenmengen und Erkenntnisfortschritt. IdS Jahrbuch 2006. Berlin 2007. S.235-248.

(236) Statistische maschinelle Übersetzungssysteme kombinieren in der Regel mindestens zwei statistische Modelle: das Übersetzungsmodell und das Sprachmodell (Brown et al. 1990). Das Übersetzungsmodell liefert Wahrscheinlichkeiten $p(f/e)$ dafür, dass einzelne Wörter oder relative kurze Sequenzen von Wörtern in der Quellsprache f Übersetzungen von Wörtern oder Sequenzen in der Zielsprache e sind. Aufgrund dessen werden Satzbruchstücke als Kandidaten für Übersetzungen ausgewählt. Das Sprachmodell liefert Wahrscheinlichkeiten $p(e)$ für Sequenzen von Wörtern in der Zielsprache e . Diese sorgen dafür, dass die Kandidatenbruchstücke in die richtige Reihenfolge gebracht werden. Die meisten Systeme verwenden zusätzlich weitere Informationsquellen (*Features*) wie z.B. die Satzlänge, wie viele verschiedene Bruchstücke verwendet werden, wie stark die Reihenfolge in der Zielsprache von der Reihenfolge in der Ausgangssprache abweicht usw. Alle verwendeten statistischen Modelle und weiteren Features werden dann in einem log-linearen Model miteinander kombiniert:

$$\log(\text{score}) = \lambda_1 \log((p(f/e))) + \lambda_2 \log(p(e)) + \lambda_3$$

(237) Die Übersetzungswahrscheinlichkeiten $p(f/e)$ und die meisten anderen Features sowie deren Gewichtungparameter λ werden anhand von parallelen Trainingsdaten gelernt, d.h. anhand von Dokumenten, die Übersetzungen voneinander sind. Im phrasenbasierten Ansatz wird versucht, innerhalb von Sätzen möglichst lange Phrasen zu identifizieren, die häufig Übersetzungen voneinander sind (Och et al. 1999; Koehn et al. 2003). Dabei sind mit Phrasen in der Regel keine Phrasen im linguistischen Sinne gemeint, sondern beliebige Sequenzen von Wörtern, Zahlen und Satzzeichen, die häufig gemeinsam mit einer bestimmten Sequenz in der anderen Sprache auftreten.

4.5 Einzeltextbezogen: Automatische Erhebungstexter Befunde

Literatur

- BEHAR, H: Hubert de Phalèse's Method. Literary & Linguistic Computing 10,2 (1995) 129-134.
- BUCHER-GILLMAYR, S: A Computer-aided Quest for Allusions to Biblical Texts within Lyric Poetry. Literary and Linguistic Computing 11 / 1 (1996) 1-8.
- BUCHER-GILLMAYR, S: Biblical words transmitted by lyric poetry. Actes du cinquième Colloque International 'Bible et Informatique: »Translation et Transmission«, Aix-en-Provence 1-4 septembre 1997. Genève 1998 357-368.: Honoré Champion Éditeur // EB
- CHEON, M A: Zur Konzeption eines phraseologischen Wörterbuchs für den Fremdsprachler: am Beispiel Deutsch-Koreanisch. Lexicographica. Series maior 89. Tübingen 1998: Niemeyer // I.2.4
- DROSDOWSKI, G; SCHOLZE-STUBENRECHT, W: DUDEN - Redewendungen und sprichwörtliche Redensarten. Wörterbuch der deutschen Idiomatik. Band 11. Mannheim 1992 // A.2
- FÖLDES, C: Deutsche Phraseologie kontrastiv. Intra- und interlinguale Zugänge. Deutsch im Kontrast 15. Heidelberg 1996: J. Groos // I.2.7
- FORTIER, K J, P A; KEEN: Change Points: Ageing and Content Words in a Large Database. Literary and Linguistic Computing 12 (1997) 15-22.
- FRANKENBERG, H: Sprichwort und Slogan. Zur Funktion des Sprichwortes in der Konsumwerbung in: TSCHAUDER, G; WEIGAND, E (ed./eds.): Perspektive: textextern. Akten des 14. Linguistischen Kolloquiums Bochum 1979 Band 2. Tübingen 1980 73-84.: Niemeyer // I.2.4
- HOFFMANN, S; EVERT, S u.a.: Corpus Linguistics with <ka>BNCweb<ke> - a Practical guide. English corpus linguistics 6. Frankfurt/M 2008 978-3-631-56315-1 pb.: Peter Lang // I.2.7
- HÜMMER, C: Synonymie bei phraseologischen Einheiten. Eine korpusbasierte Untersuchung. Postdamer Linguistische Untersuchungen 3. Frankfurt/M 2009 978-3-631-58242-8 br.: Peter Lang // I.2.7
- KALLMEYER, W; ZIFONUN, G (ed./eds.): Sprachkorpora - Datenmengen und Erkenntnisfortschritt. Jahrbuch des IdS 2006. Berlin 2007 978-3-11-019273-5: de Gruyter // I.2.7
- KAWAGUCHI, Y; ZAIMA, S; TAKAGAKI, T; SHIBANO, K; USAMI, M (ed./eds.): Linguistic Informatics - State of the Art and the Future - The first international conference on Linguistic Informatics. Usage-Based Linguistic Informatics 1. Amsterdam 2005 Hb 90 272 3313 6 // I.2.7
- KIRK, J M: Corpus-Concordance-Database-VARBRUL. Literary & Linguistic Computing 9,4 (1994) 259-266.
- KJELL, BW: Authorship Determination Using Letter Pair Frequency Features with Neural Network Classifiers. Literary and Linguistic Computing 9 (1994) 119-124.

- LEITNER, G (ed./eds.): New Directions in English Language Corpora. Methodology, Results, Software Developments. Topics in English Linguistics 9. Berlin, New York 1992: de Gruyter // I.2.1
- LEWANDOWSKA-TOMASZCZYK, B (ed./eds.): Practical Applications in Language and Computers PALC 2003. Studies in Language 9. Frankfurt / M 2004 ISBN 0-8204-6593-7: Peter Lang // I.2.7
- LÜDELING, A; KYTÖ, M (ed./eds.): Corpus Linguistics. An International Handbook. Volume 1. Handbücher zur Sprach- und Kommunikationswissenschaft (HSK) 29/1. Berlin 2008 978-3-11-018043-5: Mouton de Gruyter // I.2.7
- MALMGREN, S G: Rez. zu: A. LEHR: Kollokationen und maschinenlesbare Korpora. Ein operationales Analysemodell zum Aufbau lexikalischer Netze. Tübingen 1996.. Literary and Linguistic Computing 12 (1997) 144-146.
- McENERY, T: Corpus Linguistics in: MITKOV, R (ed./eds.): The Oxford Handbook of Computational Linguistics. Oxford Handbooks Series. Oxford 2003 448-463.: Oxford University Press // I.2.4
- PILZ, K D: Phraseologie. Redensartenforschung. Stuttgart 1981 // I.2.7
- RIFFATERRE, M: Strukturelle Stilistik. München 1973
- RÖHRICH, L (ed./eds.): Lexikon der sprichwörtlichen Redensarten, 5 Bde.. Freiburg 1995: Herder (Taschenbuchausgabe) // A.2
- ROTHKEGEL, F: Feste Syntagmen. Grundlagen, Strukturbeschreibung und automatische Analyse. Tübingen 1973 // I.2.7
- SCHINDELE, M: Computerunterstützte Suche formelhafter Rede. Protokoll des 57. Kolloquiums über die Anwendung der EDV in den Geisteswissenschaften am 6. Februar 1993 (= Literary and Linguistic Computing 9 (1994) 169-172) 7-12.
- SCHINDELE, M: Möglichkeiten und Grenzen maschineller Befunderhebung zur Untersuchung von Formeln und geprägten Wendungen mit Beispielen aus Daniel 8 in: BADER, W (ed./eds.): "Und die Wahrheit wurde hinweggefegt". Daniel 8 linguistisch interpretiert. THLI 9. Tübingen 1994 31-38.: Francke // Raum B106
- SCHINDELE, M: Externe Syntax: Verbindung zu weiteren Texten in: SCHWEIZER, H (ed./eds.): Computerunterstützte Textinterpretation. Die Josefsgeschichte beschrieben und interpretiert im Dreischritt: Syntax-Semantik-Pragmatik. THLI 7 / i. Tübingen 1995 42-88.: Francke // I.2.7
- SCHOLZE-STUBENRECHT, W: DUDEN-Zitate und Aussprüche. Band 12. Mannheim 1993 // A.2
- SCHWITALLA, J; WEGSTEIN, N (ed./eds.): Korpuslinguistik deutsch: synchron - diachron - kontrastiv. Würzburger Kolloquium 2003. Tübingen 2006 978-3-484-73064-9: Niemeyer // I.2.7
- SMITH, N; HOFFMANN, S u.a.: Corpus Tools and Methods, Today and Tomorrow: Incorporating Linguist's Manual

- Annotations. Literary and Linguistic Computing 23/2 (2008) 163-180.
- STEIN, S: Formelhafte Sprache. Untersuchungen zu ihren pragmatischen und kognitiven Funktionen im gegenwärtigen Deutsch. Sprache in der Gesellschaft, Beiträge zur Sprachwissenschaft 22. Frankfurt / M 1995 // I.2.7
- STOLZ, T; STROH, C u.a.: Nicht ganz ohne. . . in: Grzybek, P; Köhler, R (ed./eds.): Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 633-646.: Mouton de Gruyter // I.2.5
- SVARTVIK, J (ed./eds.): Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82. Stockholm, 4-8 August, 1991. Trends in Linguistics: Studies and Monographs 65 1992: de Gruyter
- TSCHAUDER, G; WEIGAND, E (ed./eds.): Perspektive: textextern. Akten des 14. Linguistischen Kolloquiums Bochum 1979 Band 2. LA 89. Tübingen 1980: Niemeyer // I.2.4
- WIERINGEN, A L H M van: Parallel Clauses Between Third and Second Isaiah. A new kind of computer-concordance. BN 82 (1996) 21-26.

4.51 Phrasendreschmaschine

Literatur

WANDL-VOGT, E: Zitate per Mausclick? Das Textkorpus zum WÖRTERBUCH DER BAIRISCHEN MUNDARTEN IN ÖSTERREICH (WBÖ) als leistungsstarkes Werkzeug für die lexikographische Praxis. Literary and Linguistic Computing 23/2 (2008) 201-217.

zu Ziff. 4.511:

Formel, Klischee

zu: vprag7.5116

SANDERS, W, Sprachkritikastereien, und was der "Fachler" dazu sagt. Darmstadt 1992. S.44f.

Formelhaftigkeit läßt sich, allgemein beschrieben, als eine Art "Konservensprache" verstehen, die uns sprachliche Fertigkeit feilbietet. Eine wahre Flut - Sprachkritiker würden es sich nicht nehmen lassen, strafend von "Sündflut" zu reden - zeitgemäß-praktischer Wörter und Wendungen, die uns als Modevokabeln und ewig wiederholte Klischees auf der Zunge wie im Ohr liegen: "Manche Menschen schwimmen mit zwei, drei Dutzend Redensarten durchs Leben und erregen bei leidlich geschickter Verteilung der einzelnen Stücke den Eindruck, schlagfertig und mütterwitzig zu sein", spottet Hans Reimann. Aber es geht ja nicht einmal nur um derartige "flotte Sprüche". Vielmehr können sprachliche Durchschnittsköstler mit Hilfe solcher fix und fertigen Versatzstücke allgemeinsten Art mühelos runde zwei Drittel ihres täglichen Sprachgebrauchs bestreiten - was den Konsum an eßbaren Fertiggerichten bei weitem übersteigt: Sprachlich ist der Mensch offensichtlich anders, als er ist.

Sprachmoden ändern sich zwar, aber gegeben hat es sie seit eh und je. Wilhelm Schneider, namhafter Stilkundler der ersten Jahrhunderthälfte, beschreibt ein junges, sportliches, gutaussehendes Mädchen, welches jedoch, "ein so reiches Füllhorn von tönlichen Modeworten und -wendungen" ausgeschüttet habe, "daß man denkt: Schönes Kind, hättest du doch den Mund nicht aufgetan!" Wer fühlt sich da nicht an jenen Lessingschen Aphorismus erinnert (über eine Dame, deren Deutsch nicht das beste war), wo es in unübertrefflicher Kürze heißt: "Solange sie mich nicht ansprach, sprach sie mich an. Als sie mich aber dann ansprach, sprach sie mich nicht mehr an." In beiden Fällen geht es um schlechtes Deutsch; im letzten offensichtlich um falsches, im ersten um richtiges, das indes den stilistischen Ansprüchen nicht genügt: Es ist eben "formelhaft" und widerspricht damit dem Stil als innerer Form des Ausdrucks. Man sollte kaum glauben, daß Formel und Form, obwohl desselben Wortstammes, derart krasse Gegensätze bilden.

Sind Sie neugierig, was unser "schönes Kind" wohl gesagt haben könnte, als es den Mund auftrat? Hier ein paar Proben seiner Modewortkost: prima, ganz groß, hundertprozentig, irgendwie, bin im Bilde, kommt nicht in die Tüte. Da zeigt sich, daß aus dem damals jungen Mädchen mittlerweile eine ältere Dame geworden sein muß. Sicher weiß man auch heute noch alles "mit hundertprozentiger Sicherheit", und das klassische Verlegenheitswort irgendwie wird niemals aussterben (beiläufiger Hinweis für Kenner: ein noch stärkerer Ausdruck von Unsicherheit läßt sich, moderner, allenfalls erzielen durch ein locker angehängtes oder so). Aber statt "prima" hört man heute super, statt "ganz groß" vielleicht echt stark; um "im Bilde" zu sein, muß man jetzt voll im Trend liegen, oder wenn etwas nicht in Frage kommt, heißt es zeitgemäß: Mit mir nicht!

Zu Tausenden bevölkern sie unsere Sprachwelt: einerseits als umgangssprachliche Floskeln, die "in aller Munde" sind, andererseits als glatte "Phrasen" auch auf Zeitungs- und anderem Papier. Der Schriftsteller Sigismund von Radecki spottet über Leute, die so sprechen und schreiben: "Ihre Sprachtasten schlagen ganze Worte, ja ganze Sätze vorrätiger Einheitsfaktor an: sie sprechen Schreibmaschine." Benutzt wird in solcher Art des

Sprachgebrauchs ein routiniertes Repertoire sprachlicher Schablonen, deren allgemeine Beliebtheit sich gerade durch die Geläufigkeit "modischer" Wörter und Wendungen in den eingeschliffenen Bahnen unseres Sprachdenkens erklären läßt. Bequemes Hinschreiben, müheloses Verstehen - die Formelsprache macht es möglich: Sie erleichtert das Denken nicht nur, sie erspart es weitgehend. Man kann die Phrasen einfach dreschen, wie sie fallen - sicher haben Sie schon von der letzten Errungenschaft sprachtechnischen Erfindungsgeistes gehört, der "Phrasendreschmaschine"?

zu Ziff. 4.512:

Phraseologie: Definition

zu: vprag7.5113

In: CHEON, Mi-AE, Zur Konzeption eines phraseologischen Wörterbuchs für den Fremdsprachler. Lexicographica, Series Maior 89. Tübingen 1998.

(9f) Die Phraseologie im Sinne eines wissenschaftlichen Gegenstandsbereichs ist "eine relativ junge linguistische Teildisziplin". Ihr Gegenstand ist eine Gruppe von sprachlichen Erscheinungen, die "eine Bereicherung des Wortschatzes einer Sprache" darstellen, und sich dadurch auszeichnen, daß "(...) Wortgruppen, in speziellen Bedeutungen 'fest' und damit zu Bestandteilen des Wortschatzes werden können". Beispiele wie *jmdm. eins/etwas aufs Dach geben*, *'jmdn. zurechtweisen, tadeln'* oder *jmdn. auf dem Korb haben*, 'es auf jmdn. abgesehen haben, sich für jmdn. interessieren' machen deutlich, daß es sich um eine "merkwürdige und komplizierte Erscheinung" handelt.

Die Bezeichnungen des Gegenstandsbereichs in verschiedenen Sprachen (beispielsweise in Englisch "phraseology", in Russisch "frazеологija", in Französisch "phraseologie") leiten sich ursprünglich von "phrasis" (griech.-lat. 'rednerischer Ausdruck') bzw. "idioma" (gr. 'Eigentümlichkeit', 'Besonderheit') ab. In der aktuellen Sprachwissenschaft gilt, daß Phraseologie als Bezeichnung des Oberbegriffs fungiert, während Idiomatizität eine bestimmte Eigenschaft (unter mehreren) bezeichnet, was auf die Verwendung dieser Begriffe in der russisch- und englischsprachigen Forschung zurückgeführt wird. Heutzutage versteht man unter Phraseologie einerseits:

eine "sprachwissenschaftliche Teildisziplin, die sich mit der Erforschung der Phraseologismen beschäftigt".

Andererseits bezeichnet der Terminus aber auch den

"Bestand (Inventar) von Phraseologismen in einer bestimmten Einzelsprache"

wie z.B. Redewendungen, Kollokationen, Funktionsverbgefüge, Sprichwörter, "kommunikative Formeln" u.a. Für das einzelne Untersuchungsobjekt der Phraseologie existiert eine Vielfalt an autoren- bzw. modellspezifischen Termini wie "Phraseolexem" (z.B. Pilz, Wotjak), "Phraseologismus" (z.B. Fleischer, Burger), "feste Wortkomplexe" (Černyševa), "Idiom" (z.B. Dobrovolskij), u.a. Im folgenden werden feste Wortverbindungen der obigen Art in Anlehnung an die Vereinbarung des 1. Internationalen Symposiums über "Phraseologie und ihre Aufgaben" Phraseme genannt. Die hier getroffene Definition stellt einen Konsens innerhalb der Phraseologieforschung dar:

(10) "Phraseme sind bedeutungstragende Einheiten der Sprache, die als Ganzes im Verlauf der Rede reproduziert werden, über mindestens zwei Autosemantika verfügen, von denen wenigstens eines eine Umdeutung erfahren hat, und die als Satzglieder fungieren oder sich an den Satz anschließen, ohne alleine einen eigenen Text zu bilden".

zu Ziff. 4.513:

Unikale Komponenten

zu: vprag7.5115

In: CHEON, Mi-AE, Zur Konzeption eines phraseologischen Wörterbuchs für den Fremdsprachler. Lexicographica, Series Maior 89. Tübingen 1998.

(14) Phraseme können Komponenten enthalten, die nur in diesen auftreten. Sie werden deshalb in der Forschung "unikale Komponenten" genannt. Sie kommen nur in einer phraseologischen Verbindung vor, wie das Nomen *Anhieb* in *auf Anhieb*, das Verb *auswetzen* in *eine Scharte/Niederlage auswetzen* 'einen Schaden wieder gutmachen', oder das Abjektiv *abhanden* in *abhanden gehen/kommen* 'verlorengehen'. Zu den unikalen Komponenten zählt man nach W. Fleischer "auch eine Reihe von Fremdwörtern und sogar fremdsprachige Kombinationen mehrerer Wörter" wie *down* in *ganz down sein*, 'bedrückt, niedergeschlagen, *ad acta* in *ad acta legen*, 'als erledigt ansehen'.

zu Ziff. 4.514:

Spezielle feste Wortverbindungen

zu: vprag7.5114

In: CHEON, Mi-AE, Zur Konzeption eines phraseologischen Wörterbuchs für den Fremdsprachler. Lexicographica, Series Maior 89. Tübingen 1998.

(36f) Funktionsverbgefüge

Innerhalb der Phraseologie werden auch Funktionsverbgefüge (FVG) behandelt. Als solche werden prädikative Ausdrücke bezeichnet, die aus Verb (Funktionsverb) und einer Nominalphrase (NP) bestehen. Bürger betrachtet unter diesem Aspekt speziell FVG, bei denen insbesondere ein synonymes Verb für den Gesamtausdruck existiert, das "zur gleichen Wurzel wie das Nomen der Streckform" gehört. Für FVG insgesamt läßt sich feststellen, daß die Funktionsverben meist in einer "abgeleiteten Bedeutung" vorliegen bzw. eine häufig "vom freien Gebrauch deutlich geschiedene Bedeutung" aufweisen. Das Nomen hingegen zeigt weitgehend einen wendungsexternen Gebrauch.

Sprichwörter

Sprichwörter sind "Ausdrücke aus literarischen Quellen oder historischen Dokumenten (Zitate), die im allgemeinen sprachlichen Umgang eine verallgemeinerte Bedeutung erhalten haben", denen also Verallgemeinerung als semantische Funktion zugeordnet werden kann, bzw. "eine metaphorische, verallgemeinerte Bedeutung". Untersuchungen zum Sprichwort differenzieren den verallgemeinerten Charakter u.a. in die Funktionen:

1) "Depersonalisierung":

Eine Mitteilung kann indirekt und unpersönlich übermittelt werden z.B. *Morgenstund hat Gold im Mund* 'am Morgen läßt es sich gut arbeiten; wer früh mit der Arbeit anfängt, erreicht viel'.

2) "Betonung"

Der im Sprichwort referierte Sachverhalt bzw. das Verhalten soll dem Hörer ganz besonders deutlich vor Augen geführt werden, besonders beispielsweise durch Metaphorisierung wie in *der Apfel fällt nicht weit vom Stamm* 'jmd. ist in den negativen Anlagen den Eltern sehr ähnlich'.

3) Pädagogische Funktion:

Es wird auf gesellschaftliche Normen und Werte in appellativer Weise verwiesen, wie z.B. in *was du heute kannst besorgen, das verschiebe nicht auf morgen*.

Sprichwörter weisen überdies Satzwertigkeit auf, dienen also nicht als Satzbausteine, sondern fungieren als "Mikrotexte". Dies bedeutet, "dass sie keinerlei formale Möglichkeiten des Anschlusses an den Kontext aufweisen". Auf mögliche Kontextabhängigkeit der Sprichwort-Semantik weist Permjakov hin, indem er feststellt, daß die kontextuelle Einbettung dem Sprichwort ein semantisches Element, das ihm ursprünglich fehlt, hinzufügen kann.

Gemeinplätze

Semantisches Merkmal dieser festen und wie Sprichwörter auch satzwertigen Wortverbindungen ist die "Reduktion von Komplexität", was sich lexikalisch im Auftreten von Komponenten wie "man, jeder, alle, alles" u.a. ausdrückt, wie z.B. "wir sind alle nur Menschen", "was sein muß, muß sein" usw. "Charakteristisch für die hier auftretenden indexikalischen Ausdrücke ist, daß sie nur eine unspezifische Referenzleistung aufweisen, die sich letztlich in

einem Verweis auf den pragmatischen Kontext insgesamt erschöpft. Diesen Typ von vorgeformten Sätzen gibt es - soweit ich sehe - bei den Sprichwörtern nicht.

Geflügelte Worte

Notwendiges und hinreichendes Abgrenzungskriterium der geflügelten Worte ist die Aufzeigbarkeit der Urheberschaft. Es handelt sich deshalb um eine strukturell, semantisch und bezüglich der Idiomatizität heterogene Gruppe fester Wortverbindungen. Bei Beschränkung auf synchrone Sprachbeschreibung ergibt sich deshalb eine Verteilung auf entsprechende, durch synchron definierbare Kriterien konstituierte Gruppen fester Wortverbindungen.

"Kommunikative Formeln"

"Kommunikative Formeln" (*Guten Tag, wohl bekomm's* usw.) sind pragmatisch durch ihre Verwendung in der (geschriebenen oder gesprochenen direkten) Rede abgegrenzt und bezeichnen semantisch mehr oder weniger komplexe feste dialogstrukturierende und damit auch situationsgebundene sprachliche Einheiten unterschiedlicher Idiomatizität.

zu Ziff. 4.515:

Formel = Phraseologische Einheit

zu:

aus: LUTZEIER, R, P, Lexikologie. Ein Arbeitsbuch. Tübingen 1995.

(36) Zu den **idiomatischen Wendungen** zählen wir Verkettungen wie *ein großes theater machen, ins gras beißen, ruhig blut* und *jemandem den buckel runterrutschen*. Da *zwei große theater machen* als idiomatische Wendung nicht zu gebrauchen ist, sieht man sehr schnell, daß wir eventuell von eingeschränkter Flexion beim internen Aspekt auf der Formebene reden müssen. Die Markierung beim internen Aspekt für die Inhaltsebene stellt auf die isolierte Relevanz des durch den globalen Zugriff gewonnenen, "idiomatischen" Inhalts ab. Eine synchrone Remotivation, die von Sprechern/Sprecherinnen immer wieder gern vorgenommen wird, ist im allgemeinen als Remetaphorisierung zu sehen. Man wird schließlich für den externen Aspekt die globale Kombinierbarkeit mit anderen Elementen haben. Selbstverständlich muß man geringfügige interne Veränderungen bei den Wendungen zulassen, vgl. *er macht ein wahnsinnig großes theater* und *im vergleich zu andern macht er ein verhältnismäßig großes theater*.

zu Ziff. 4.516:

Formeln, geprägte Sprache, Klischee

zu:

aus: Wolfgang RICHTER, Exegese als Literaturwissenschaft. Göttingen 1971.

(99) "Die Fragestellung wendet sich einzelnen Gliedern der Einheit zu, kann aber auch zu dem Ergebnis kommen, daß eine Einheit eine Reihe von geprägten Elementen aufweist oder gar nur aus solchen besteht. Eigentümlich für diese Fragestellung ist es, daß sie über die untersuchte kleine Einheit hinausführt auf weitere kleine Einheiten...

Man hat es sich oft im Auffinden von 'Formeln' sehr leicht gemacht und solche nur gefühlsmäßig oder auf Mutmaßung hin angesetzt. Die Art der Untersuchung ist deshalb etwas in Mißkredit geraten. Sie muß indes geleistet werden, will man den Charakter der Einheit richtig bewerten und einordnen. Wichtig ist sie aus zwei Gründen: Einmal liefert sie Kriterien für die Beurteilung, ob die Einheit literarisch verfaßt oder schon vorliterarisch überliefert ist; zum andern führt sie über die gegebene Einheit hinaus auf weitere Einheiten und liefert Kriterien für die Frage, ob die untersuchte Einheit aus sich heraus verständlich oder Glied in einer Reihe von Einheiten, somit einer größeren Komposition ist.

Die Fragestellung ist durchaus nicht auf bloße Mutmaßungen angewiesen; sondern stützt sich auf nachprüfbar Kriterien, ist also ausweisbarer methodischer Schritt. Sie besteht in der Feststellung, ob formal Wortverbindungen oder Wortgruppen wiederholt belegt sind. Der Ausgang von einem einzelnen Wort oder Begriff erlaubt solche Feststellungen nicht; die Wortverbindung läßt aber den für ein Wort oder einen Begriff typischen Kontext erkennen und zeigt seine sprachliche Eingeflochtenheit und stilistische Besonderheiten des Gebrauchs. Eine Häufung solcher Besonderheiten führt so zum Schluß auf Eigentümlichkeiten von Autoren".

aus: Michael RIFFATERRE, Strukturele Stilistik. München 1973:

(139f.147) "Als Klischee betrachtet man eine Gruppe von Wörtern, die Urteile hervorrufen wie: nichts Neues, banal, abgedroschen, falsche Eleganz, abgenutzt, erstarrt usw. Wir können aus diesen Reaktionen auf die Existenz einer linguistischen Einheit schließen (einem zusammengesetzten Wort analog), da die Gruppe en bloc lexikalische oder syntaktische Einheiten ersetzen kann und da ihre Komponenten getrennt betrachtet, nicht mehr als Klischee empfunden werden. Diese linguistische Einheit ist expressiv, da sie ästhetische, ethische oder affektive Reaktionen hervorruft. Sie ist strukturele und nicht semantischer Art, da ein Ersatz durch ein Synonym das Klischee beseitigt. Sie läßt keine Varianten zu...

zwei wesentliche Merkmale:

Erstens: das Klischee weist eine starke und stabile Ausdruckskraft auf. Zunächst deshalb, weil die stereotype Wendung ein vollständiges Stilfaktum enthält, d.h., eine binäre, aus einem (Mikrokontext und einem mit diesem Kontext kontrastierenden Element gebildete Gruppe; dieser Kontrast zweier entgegengesetzter, aber untrennbarer Pole ist erstarrt: seine Wirkung ist also, wenn ich so sagen darf, konserviert... Diese Wirkung wird verstärkt durch den Makrokontext, in dem er steht und mit dem er umso mehr kontrastiert, als ein erstarrtes Stilfaktum definitionsgemäß keine lexikalischen Ersatzmöglichkeiten zuläßt:

folglich neigt er dazu, sich von dem dem Makrokontext eigenen Vokabular zu entfernen. Dieser Mangel an Übereinstimmung führt dazu, das Klischee als geschraubt, künstlich usw. zu bewerten, besonders dann, wenn die Ästhetik des Lesers auf Prinzipien wie Harmonie oder Einheitlichkeit der Sprache beruht ...

ein zweites Kennzeichen des Klischees: seine Wirksamkeit als Stilfaktum ist richtungsorientiert. Während das ursprüngliche Stilfaktum in seinem Kontext etwas Neues ist, ist das Klischee an den Makrokontext gebunden, in dem es zum ersten Mal auffiel; es funktioniert also als Zitat, als Verweis auf eine bestimmte soziale Ebene, auf bestimmte kulturelle Äußerungen."

zu Ziff. 4.517:

(Ausdrucks-)SYNTAX: Paar-/Zwillingsformeln

zu: vgt5.232

aus: G. MÜLLER, Beschränkungen für Binomialbildung im Deutschen. Ein Beitrag zur Interaktion von Phraseologie und Grammatik: Zeitschrift für Sprachwissenschaft 16.1/2 (1997) 5-51.

(6) Hierbei handelt es sich um formelhafte *und*-Verbindungen aus (im Regelfall) zwei Komponenten. Einige Beispiele für verschiedene Wortarten finden sich in (1):

- (1) a. *Koordinative Adjektiv/Adverb-Binomiale*:
fix und fertig, kurz und gut, gut und gern, klipp und klar, ab und zu, ganz und gar, angst und bange
- b. *Koordinative Substantiv-Binomiale*:
Katz und Maus, Kind und Kegel, Haus und Hof, Soll und Haben, Knall und Fall, Tag und Nacht, Sack und Pack, Grund und Boden, Drum und Dran, Biegen und Brechen, Nacht und Nebel, Schloß und Riegel
- c. *Koordinative Verb-Binomiale*:
hegen und pflegen, schalten und walten, kommen und gehen, zittern und zagen, (sich) recken und strecken

Koordinative Binomiale weisen die für Phraseologismen typischen Merkmale auf: Sie sind (mehr oder weniger) *semantisch opak*, und sie werden als *formelhaft* empfunden. Darüber hinaus haben koordinative Binomiale eine auf den ersten Blick merkwürdige und unerklärte Eigenschaft: Wie Malkiel (1959) ausführlich darlegt, zeigen Binomiale eine starke Tendenz zur *Irreversibilität*, d.h., eine Änderung der Abfolge führt entweder zu Ungrammatikalität (**klar und klipp*, **Pack und Sack*) oder zum Verlust der Formelhaftigkeit, mit einhergehender semantischer Transparenz der Einzelglieder (*gern und gut*, *Nacht und Tag*) ...

Man vergleiche etwa die folgenden Beispiele aus dem Englischen, Spanischen und Russischen:

- (2) a. *Englische Binomiale*:
salt and pepper, birds and bees, husband and wife, ladies and gentlemen, boys and girls, odds and ends, words and music, law and order, (a) rough and tough (speech), tried and true, hale and hearty, sweet and sour, to have and to hold, to live and learn, bought and paid for, safe and sane, rock and roll
- b. *Spanische Binomiale*: (8)
pan y agua ('Brot und Wasser'), corriente y mollente ('regulär, in Ordnung', 'laufend und mahlend'), punto y coma ('Punkt und Komma', 'Semikolon'), (echar a) cara y cruz ('eine Münze werfen', 'Kopf oder Zahl'), carne y huesos ('Fleisch und Knochen'), liso y llano ('einfach', 'glatt und eben'), blanco y negro ('weiß und schwarz'), cuerpo y alma ('Körper und Seele')
- c) *Russische Binomiale*:
den' i noč' ('Tag und Nacht'), grom i molnija ('Donner und Blitz'), muž i žena ('Mann und Frau'), šutki i pribautki ('Scherze und Redensarten'), vdol' i poperëk ('kreuz und quer'), vstrečnyj i poperečnyj ('jeder', 'Hinz und Kunz'), styd i sram ('Schimpf und Schande'), (razrjadit'sja v) puch i prach ('Daunen und Staub', 'sich mit aller Pracht herausputzen')

Zweitens ist die Binomialbildung im Deutschen tatsächlich zweifelsohne ein produktiver Prozeß, der in den verschiedensten Bereichen Anwendung findet, wie etwa in der Werbung (*flieg und spar*, (*in*) *Bad und WC*, *wisch und weg*), in Zeitschriften (wo Kolumnen Namen tragen können wie *gemein und geistreich*, *schnell und vergänglich*, *Kunst und Gewerbe* oder *Buch und Deckel*), in Comics (man vergleiche etwa *Plisch und Plum*, *Fix und Foxi*), und so weiter. In allen diesen Fällen liegt Irreversibilität vor (vgl. etwa **weg und wisch* oder **Foxi und Fix*).

(12) Vom semantischen Standpunkt aus gesehen hat daher ein koordinatives Binomial häufig denselben Status wie ein einfaches Wort: Es muß im mentalen Lexikon aufgelistet sein und dort eine Bedeutung zugeordnet bekommen...

(13) Unter dieser Perspektive ist das häufige Auftreten von Endreimen in Binomialen nicht verwunderlich - wie z.B. aus Forschungen zur *oral poetry* bekannt ist, erleichtern Endreime die mentale Speicherung. Typische Beispiele hierfür sind die unter (7-a) angegebenen. Eine ähnliche Funktion kann man aber auch den weit verbreiteten Alliterationen und Assonanzen attribuieren; vgl. (7-b) und (7-c):

- (7) a. *Endreime*:
Rat und Tat, sang- und klang(los), Sack und Pack, schalten und walten, Handel und Wandel, rank und schlank, dann und wann, (mit) Ach und Krach
- b. *Alliterationen*:
frank und frei, hoch und heilig, Biegen und Brechen, klipp und klar, Haut und Haar, gang und gäbe, Glanz und Gloria, Kind und Kegel, rauf und runter, Stumpf und Stiel, Drum und Dran, (im) großen und ganzen, hin und her, Land und Leute, Lust und Laune, Mann und Maus, Nacht und Nebel, Wind und Wetter
- c. *Assonanzen*:
angst und bange, Brief und Siegel, ganz und gar.

Darüber hinaus scheint mir eine interessante Hypothese, daß die strikten Reihenfolgerestriktionen in koordinativen Binomialen letzten Endes dieselbe Ursache haben, nämlich den durch die semantische Opazität verursachten Aufwand auszugleichen und so die mentale Abspeicherung zu erleichtern.

4.52 Intertextuelle Befunde

Literatur

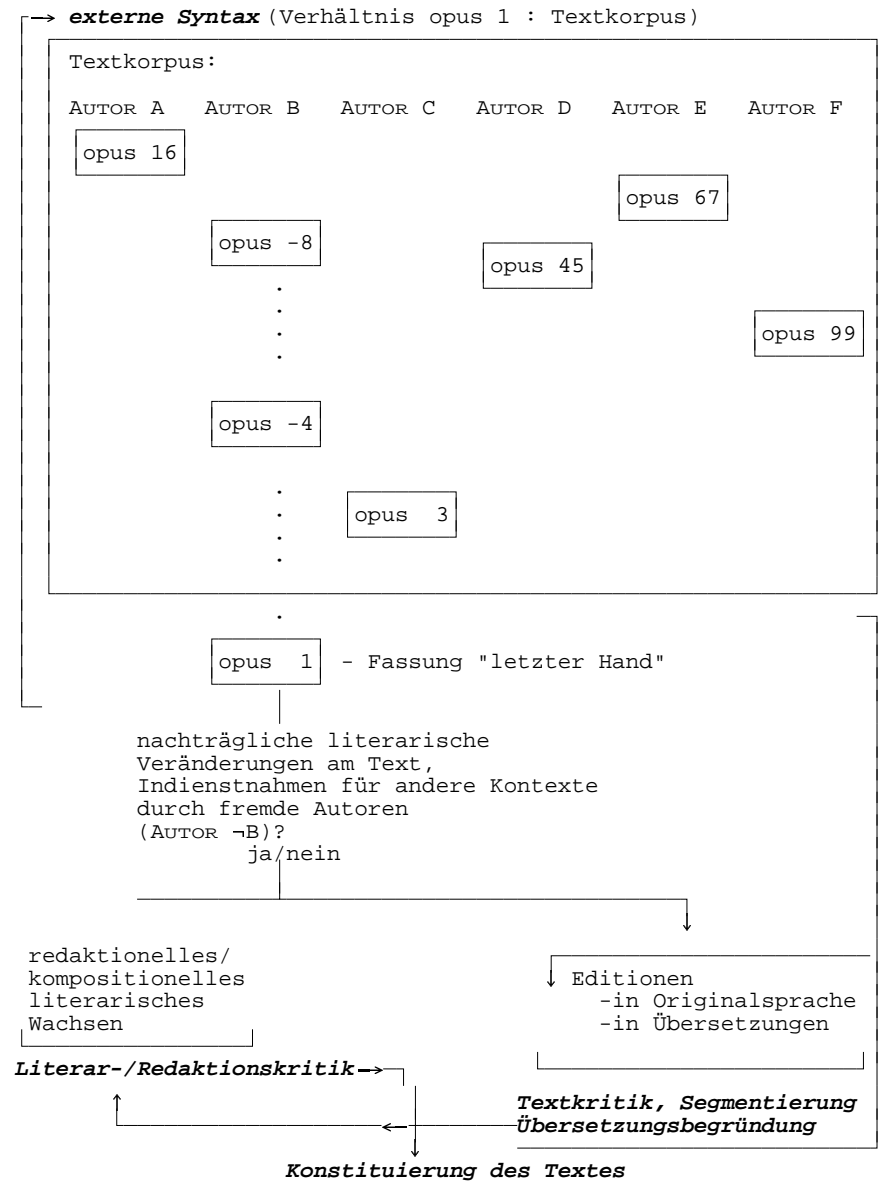
CLAVER ZOUOGOBO, J P: Le proverbe entre langues et cultures. Une étude de linguistique confrontative allemand/français/bêté. Etudes contrastives 10. Bern 2009 978-3-03911-667-6 br: P. Lang // I.2.7

OAKES, M P; FARROW, M: Use of the Chi-Squared Test to Examine Vocabulary Differences in English Language Corpora Representing Seven Different Countries. Literary and Linguistic Computing 22 / 1 (2007) 85-99.

zu Ziff. 4.521:

Entstehung - Überarbeitung - Überlieferung von Texten

Zu: vsyst7.51 vgt3.431



Hier angenommen: *Opus 1* ist der Text eines Autors, auf den man sich bezieht, für den der Autor bekannt ist, z. B. GOETHES *Faust*, oder das Gedicht von BRECHT *An die Nachgeborenen* usw.

Wichtig: *Opus 1* meint die Textfassung, die ein Autor selbst geschaffen hat (geschrieben bzw. autorisiert). "Opus 1" ist nicht identisch mit einem frühen Druck, denn auf dem Weg von der Abfassung zur Drucklegung können bereits Veränderungen am Text stattfinden.

Bei einem gegebenen Werk - *Opus 1* als existierende Handschrift oder nur noch als erschließbare Größe (weil das Original verloren ist) - einsetzend, gilt zugleich: dieser Text hat auf jeden Fall eine **Vorgeschichte** (denn er fiel nicht vom Himmel); und er hat eine **Nachgeschichte**.

Alle drei Faktoren - Vorgeschichte/Verfassen des Textes/Nachgeschichte - werden als *Textbildungsprozess* bezeichnet und sind Gegenstand verschiedener wissenschaftlicher Disziplinen (mit je spezifischen Methoden).

Alle Stadien eines solchen *Textbildungsprozesses* können durch Zettelsammlungen, Briefe, Handschriften, Typoskripte, Drucke bezeugt sein, oder aber nur durch einige späte Editionen - was die Rekonstruktion von *opus 1* entsprechend erschwert.

Ist durch die Schritte der **Konstituierung** ein Text ausreichend kritisch gesichtet, bereitgestellt, aufgearbeitet, steht also für weitere Deskription und Interpretation zur Verfügung, kann die (**Ausdrucks-**) **Syntax** anschließen - und zwar unter extensivem Rechneinsatz für die Datenerhebung :

- die **interne Syntax** analysiert die *string*-Kette des gegebenen Textes; die zuvor eingetragenen Segmentierungen des Textes, werden dabei übergangen;
- die **externe Syntax** macht sichtbar, welche *substrings* (z.B. Wortfolgen > 3) exakt/ähnlich in anderen Texten des (zuvor zu definierenden) Textkorpus vorkommen.

Die oben eingetragene **externe Syntax** kennzeichnet in der Grafik den Zuständigkeitsbereich, nicht eine Position im Ablauf der Methodenschritte. Zunächst muss die **Konstituierung** abgearbeitet werden, bevor es zur **Ausdrucks-Syntax** kommt.

zu Ziff. 4.522:

Externe Syntax, Intertextualität, Korpusanalyse

zu: vsyst4.201

I)

H e b r ä i s c h e B i b e l		Gen
JOSEFSGESCHICHTE Gen 37-50		
H e b r ä i s c h e	B i b e l	Ex
H e b r ä i s c h e	B i b e l	Num
H e b r ä i s c h e	B i b e l	Lev
H e b r ä i s c h e	B i b e l	Dtn
H e b r ä i s c h e	B i b e l	Jos
H e b r ä i s c h e	B i b e l	Ri
H e b r ä i s c h e	B i b e l	Rut
H e b r ä i s c h e	B i b e l	1Sam
H e b r ä i s c h e	B i b e l	2Sam
H e b r ä i s c h e	B i b e l	1Kön
H e b r ä i s c h e	B i b e l	2Kön
H e b r ä i s c h e	B i b e l	1Chr
H e b r ä i s c h e	B i b e l	2Chr
H e b r ä i s c h e	B i b e l	Esr
H e b r ä i s c h e	B i b e l	Neh
H e b r ä i s c h e	B i b e l	Ijob
H e b r ä i s c h e	B i b e l	Pss
H e b r ä i s c h e	B i b e l	Spr
H e b r ä i s c h e	B i b e l	Koh
H e b r ä i s c h e	B i b e l	Hld
H e b r ä i s c h e	B i b e l	Jes
H e b r ä i s c h e	B i b e l	Jer
H e b r ä i s c h e	B i b e l	Klgl
H e b r ä i s c h e	B i b e l	Ez
H e b r ä i s c h e	B i b e l	Dan
H e b r ä i s c h e	B i b e l	Hos
H e b r ä i s c h e	B i b e l	Joel
H e b r ä i s c h e	B i b e l	Am
H e b r ä i s c h e	B i b e l	Ob
H e b r ä i s c h e	B i b e l	Jona
H e b r ä i s c h e	B i b e l	Micha
H e b r ä i s c h e	B i b e l	Nah
H e b r ä i s c h e	B i b e l	Hab
H e b r ä i s c h e	B i b e l	Zef
H e b r ä i s c h e	B i b e l	Hag
H e b r ä i s c h e	B i b e l	Sach
H e b r ä i s c h e	B i b e l	Mal

II)

Wortform _n + Wf _{n+1} + Wf _{n+2}	+ Wf _{n+3}
wenn	Gnade ich-gefunden-habe in-deinen-Augen
Ex	
Num	
Lev	
Dtn	
Jos	
Ri	
Rut	
1Sam	
2Sam	
1Kön	
2Kön	
1Chr	
2Chr	
Esr	
Neh	
Ijob	
Pss	
Spr	
Koh	
Hld	
Jes	
Jer	
Klgl	
Ez	
Dan	
Hos	
Joel	
Am	
Ob	
Jona	
Micha	
Nah	
Hab	
Zef	
Hag	
Sach	
Mal	

III)

Erkennbar:

- Synchron: Welche Kombinationen von Wortformen im Ausgangstext finden sich auch im umgebenden Vergleichskorpus? Wer den Ausgangstext liest und das Vergleichskorpus kennt, der fühlt sich an diesen Stellen an die anderen Texte erinnert. Kann vom Autor auch gezielt eingesetzt werden: Anspielungen, Zitate, Formeln. Die anderen Texte verstärken so die Aussage des Ausgangstextes. - Dominierende Formeln/Idiome zeigen an, daß der Autor sich gesellschaftlich üblichem/akzeptiertem Sprachgebrauch anpaßt. Das weitgehende Fehlen von Formeln/Idiomen zeigt dagegen kreativen und eigenständigen Sprachgebrauch. - Interessant kann die Beobachtung der Ähnlichkeit sein: Ein Autor verwendet Formeln/Idiome, wandelt sie aber eigenständig ab, in der Regel, um wirkungsvolle Aussagekorrekturen anzubringen.
- Diachron: Zusammen mit weiteren Überlegungen können die Ergebnisse verwendet werden zur Erstellung einer relativen Chronologie der Texte: Setzt mein Ausgangstext das Vorhandensein anderer Texte erkennbar voraus? Der Ausgangstext müßte demnach später verfaßt sein.

IV)

WN 1 - WN 123

- (2,3) JG 37,2b* *hâyâh ro'âh 'ät*
JG 37,2b*; Ex 3,1;
- (2,2) JG 37,2b *hâyâh ro'âh*
JG 37,2b; Ex 3,1;
- (3,2) JG 37,2b *ro'âh 'ät*
JG 37,2b; Gen 30,36; Ex 3,1;
- (72,3) JG 37,12a *lir'owt 'ät so'n*
JG 37,12a; 1 Sam 17,15;
- (72,2) JG 37,12a *lir'owt 'ät*
JG 37,12a; 1 Sam 17,15; 2 Sam 7,7; 1 Chron 17,6;
- (73,2) JG 37,12a *'ät so'n*
JG 37,12a; Gen 29,10; Gen 30,36; Ex 3,1; 1 Sam 17,15;
Jer 23,1; Sach 11,4; Sach 11,7;

Bezieht man die Umgebungen der Fundstellen mit ein, so liest man an den V-Fundstellen des ersten Verweisbüschels:

Gen 37,2b* *yosep hâyâh ro'âh 'ät 'ähâyw baš so'n*
Ex 3,1 *uwmošâh hâyâh ro'âh 'ät so'n yitrow*

und an den V-Fundstellen des vierten Verweisbüschels:
Gen 37,12a *wayye1kuw 'ähâyw lir'owt 'ät so'n*
'biyhâm biškâm
1 Sam 17,15 *wdâwid holek wâšâb me'al ša'uwl lir'owt 'ät so'n 'biyw beyt lâhâm*

In beiden Fällen gibt es also eine zusätzliche, freier beschreibbare Gemeinsamkeit (-so'n bzw. -'biy)-, beide Vergleiche enthalten eine von r'h G (weiden, hüten) gebildete Wortform. Kontext von Ex 3,1 ist die Berufung des Mose, Kontext von 1 Sam 17 ist die Geschichte von David und Goliath, im Anschluß an die Berufungsgeschichte des David. Im ersten Fall verweisen auch alle in *hâyâh ro'âh 'ät* enthaltenen zweigliedrigen Wortketten lediglich nach Ex 3,1 (2. und 3. Verweisbüschel). Im zweiten Fall verweist die Wortkette *lir'owt 'ät* außerdem nach 2 Sam 7,7 und 1 Chron 17,6, beidesmal werden in der Reaktion des Natan auf das

Ansinnen Davids, einen Tempel zu bauen, die Richter als *lir'owt* 'ät *'ammij* (um zu hüten mein Volk) bezeichnet, beidesmal wird unmittelbar danach auf die Vergangenheit Davids als Hüter verwiesen. Die Wortfolge 'ät so'n verweist ebenfalls im zweiten Fall außer auf 1 Sam 17,6 noch auf Gen 29,10 (Jakob begegnet Rahel am Brunnen), Gen 30,36 (Jakob weidet das Vieh Labans), Ex 3,1 (s. o.), Jer 23,1 (Spruch über "die Hirten, die mein Volk weiden"), Sach 11,4.7 (Drohwort gegen schlechte Herrscher).

Die beiden Wortketten lassen sich als Anspielung auf die Hirtentätigkeit von Mose und David in deren Jugend (sachliche Parallele) verstehen, erst in zweiter Linie kommt deren spätere "Tätigkeit" als "Hüter des Volkes" in den Blick. Zum Ausdruck kommt die Bedeutung Josefs in einem allgemeinen Sinn (er wird im Zusammenhang mit Mose und David gesehen), als auch im Vorgriff auf die folgende Geschichte (Josef als Hüter von Volk = Familie).

Beide Befunde (Ex 3,1; 1 Sam 17,15) erzeugen zudem einen gewissen "Druck" auf Gen 37,2b*: Sie belegen die auch bedeutungsmäßig erwartbare Konstruktion *r'h G + 'ät* i. S. v. <<HÜTEN>> + [2. Aktant]. Diese Konstruktion bietet Gen 37,2b zunächst und legt damit die Deutung von Josef als dem Hüter seiner Brüder nahe. Erst durch das nachfolgende *bässa'n* ist diese Bedeutung wieder zu verwerfen, ist die Funktion der Brüder als [komitativ] zu bestimmen. Am Anfang der Josefsgeschichte steht also Irritation und Humor zu Lasten der Brüder.

Vgl.: MARTIN SCHINDELE, Externe Syntax. in: HARALD SCHWEIZER, Computerunterstützte Textinterpretation. Die Josefsgeschichte beschrieben und interpretiert im Dreischritt: Syntax - Semantik - Pragmatik. Bd. I: Textbeschreibung und -interpretation. Tübingen 1995. S.42-88.

zu Ziff. 4.523:

Externe Syntax: Bezugnahmen Josefsgeschichte → AT

Der Akteur JOSEF selbst wird durch die Referenztexte ungemein aufgewertet. Mehrfach wird er im Licht des Mose gezeigt,⁴⁴ aber auch in dem Davids,⁴⁵ mehrfach in dem Salomos,⁴⁶ Josuas,⁴⁷ ja auch Sanheribs.⁴⁸ Josef wird auch mit Jahwe/Gott,⁴⁹ mit Elischa in Verbindung gebracht.⁵⁰ Es wird durch die Texte im Hintergrund betont, wie großzügig Josef ist. Aber manche Anspielungen betonen auch - auf dem entsprechenden Stand der JG - die mögliche Ambivalenz der anstehenden Erfahrung: Josef also auch angsterregend.⁵¹ Darin liegt ein großer dramaturgischer Effekt der betr. Assoziationen.

Der PHARAO wird verschiedentlich als gut geschildert, er erfüllt die göttliche Vorsehung, er dient.⁵² Aber auch hier kann es ambivalent sein, zum mächtigen Pharao gerufen zu werden.⁵³

Mehrfach wird kontrastiv gegenüber den Referenztexten, die etwa die Botschaft enthalten, das Heil liege in einer Bewegung "weg von Ägypten", ausgesagt, gerade die Tendenz "hin zu Ägypten" sei die richtige.⁵⁴ Flankierend dazu: Kritik an Israel, denn es ist kulturell unterlegen.⁵⁵ Die Figur "Israel" ist gegenüber der Figur "Abraham" kleinlicher, ängstlicher im Denken.⁵⁶

Wenn die Referenztexte z.B. hochtheologische Theophanien enthalten,⁵⁷ so macht sich diese der Autor der JG dramatisierend zunutze. Sein expliziter Text enthält aber nichts von derartigen theologischen Denkmustern und Bildern.⁵⁸

⁴⁴ JG 37,2b//Ex 3,1; JG 41,3d//Ex 2,3; 7,15; JG 41,14ff//Ex 7,7f; 10,8; JG 42,28g f//Ex 13,14; JG 43,7e//Ex 34,27; JG 43,11c f + 45,19b f//Num 16,6; JG 45,16c//Ex 5,21; JG 45,19b f//Num 16,6; JG 45,21a //Ex 16,17; Num 5,4; Jos 4,8; JG 47,3c//Ex 5,1; JG 50,19b//Ex 20,20.

⁴⁵ JG 37,2b//1 Sam 17,5; JG 37,23ab//2 Sam 16,16; JG 40,20a//1 Sam 17,46; 2 Sam 1,2; JG 41,57b//2 Sam 24,8; JG 43,27a//1 Sam 30,21; JG 45,26a//1 Sam 24,2; JG 47,31a f//1 Sam 30,15.

⁴⁶ JG 40,20a//1 Kön 3,18; JG 40,20b//1 Kön 3,15; JG 41,8d//1 Kön 1,9; JG 41,37a//1 Kön 3,10; JG 45,27a //1 Kön 10,2.

⁴⁷ JG 41,38b ff//Num 27,18.

⁴⁸ JG 42,29b//Jes 36,22; JG 47,14a//2 Kön 18,15.

⁴⁹ JG 46,30b//Ps 21,7f; JG 47,29c f//Num 22,20.

⁵⁰ JG 37,19ab//2 Kön 7,3.6.9; JG 37,25b ff//2 Kön 6,20; JG 43,26e//2 Kön 2,15.

⁵¹ JG 40,14ef//Jer 22,5f; JG 40,20a//Gen 34,25; Est 5,1; JG 43,16b ff//2 Kön 10,22; JG 43,29a//Ri 19,17; JG 45,1c//2 Sam 13,9.

⁵² JG 41,1c f//Gen 24,30; kontrastiv: JG 45,21b//2 Kön 23,35.

⁵³ JG 41,14a f//Ex 9,27.

⁵⁴ JG 42,5a//Ex 14,22; Num 20,1.22; JG 46,5b//Ex 14,10.

⁵⁵ JG 41,57a//1 Sam 14,25; JG 46,34f//Ex 8,22.

⁵⁶ JG 47,29e//Gen 24,2.

⁵⁷ JG 40,20a//Ex 19,16; JG 42,14b f//Ez 38,17; JG 50,19b//Ex 20,20; Hag 2,5f.

⁵⁸ Berufungstexte: JG 37,16c//Ri 16,6.10; 1 Sam 9,18; 10,5.

zu Ziff. 4.524:

Einzeltext & Textkorpus

zu:

1. Wortlaut eines alttestamentlichen Textes

Jer 2,1 Das Wort des Herrn erging an mich: 2 "Geh, rufe Jerusalem laut ins Ohr: 'So spricht der Herr: Ich lohnte dir deine junge Liebe, deine bräutliche Minne, da du in der Wüste mir folgst, im Land ohne Aussaat. Israel galt dem Herrn als heiliges Gut, als Erstling seines Ertrages. Wer davon aß, der mußte es büßen, Unheil kam über ihn'" - Spruch des Herrn.

4 Höret das Wort des Herrn, Haus Jakobs und alle Geschlechter des Hauses Israel! 5 So spricht der Herr: "Was fanden Unrecht an mir eure Väter, daß sie mich verließen, daß sie herliefen hinter nichtigen Götzen und selber zunichte wurden?"

6 Sie fragten nicht: 'Wo ist der Herr, der uns aus Ägypten geführt, der uns durch die Wüste geleitet, durch Steppen und Schluchten, durch dürres und düsteres Land, das kein Wanderer durchzieht und kein Mensch bewohnt.'

7 Ich brachte euch in ein fruchtbares Land, damit ihr dessen Güter und Früchte genießen solltet. Doch ihr kamt und entweihet mein Land und machtet zur Greuelstätte mein Erbe".

2. Hinweise auf einbettendes literarisches Korpus:

1. Jes 28,13
Jer 7,1
Ez 21,1; 25,1
2. Jes 66,1. 12
Am 7,17
Ez 36,2
3. Hos 2,15-19
Hos 11,1-2
Jer 3,6-7
4. Jes 30,1; 54,17
Jer 7,11.13.19.30
5. Ez 13,1
Jer 11,6
Ez 34,7
2 Kön 20,16
6. Jes 28,16
Jer 19,1.3
Ez 36,3.4
7. 2 Kön 17,15
Jer 2,23; 9,13
8. Jer 16,14
Am 2,10
2 Kön 17,7
Ri 6,13
9. Am 2,10
Ps 106,9
Ps 136,16

10. Jer 51,43
Ez 29,11

11. Jes 32,15.18
Neh 9,36

12. Dtn 21,23
Num 35,34

13. Ez 5,11
Jer 16,18

14. vgl. insgesamt: Ps 81,6-13;
78,5-8.56-64;
106,34-43

- (1) Es gilt zuerst herauszufinden, worin die genannten Stellen mit Jer 2 übereinstimmen. Das Ergebnis können Sie zusammenfassend herauschreiben (es braucht nicht eine streng wörtliche Übereinstimmung zu sein).
- (2) Welches Bild bietet nach dieser Untersuchung Jer 2? (Also: Welche Elemente kommen in gleicher oder ähnlicher Form auch anderswo vor? - Durch Unterstreichung hervorheben)
- (3) Was will und was bewirkt ein Schriftsteller, wenn er einen solchen Text verfaßt?

3. Hinweise auf einbettendes literarisches Korpus:

Jer 2,1 Das Wort des Herrn erging an mich: 2 "Geh, rufe Jerusalem laut ins Ohr: 'So spricht der Herr: Ich lohnte dir deine junge Liebe, deine bräutliche Minne, da du in der Wüste mir folgst, im Land ohne Aussaat. Israel galt dem Herrn als heiliges Gut, als Erstling seines Ertrages. Wer davon aß, der mußte es büßen, Unheil kam über ihn'" - Spruch des Herrn.

4 Höret das Wort des Herrn, Haus Jakobs und alle Geschlechter des Hauses Israel! 5 So spricht der Herr: "Was fanden Unrecht an mir eure Väter, daß sie mich verließen, daß sie herliefen hinter nichtigen Götzen und selber zunichte wurden?"

6 Sie fragten nicht: 'Wo ist der Herr, der uns aus Ägypten geführt, der uns durch die Wüste geleitet, durch Steppen und Schluchten, durch dürres und düsteres Land, das kein Wanderer durchzieht und kein Mensch bewohnt.'

7 Ich brachte euch in ein fruchtbares Land, damit ihr dessen Güter und Früchte genießen solltet. Doch ihr kamt und entweihet mein Land und machtet zur Greuelstätte mein Erbe".

zu Ziff. 4.525:

Computergestützt: geprägte Wortketten

zu:

aus: A. PARTINGTON; J. MORLEY, 'At the heart of ideology: Word and Cluster/bundle frequency: LEWANDOWSKA-TOMASZCZYK, B (ed.), Practical Applications in Language and Computers. Łódź Studies in Language 9. Frankfurt/M 2004. S. 179-192.

(179) If we look at the string of words:

...at the heart of...

it is clear that it does not correspond to any structural unit described in classical pre-phraseological grammar. The string is made up of a prepositional phrase *at the heart* plus an embedded *of*-phrase fragment. Neither is the string an example of that class of fixed expression which Moon (1994: 117) calls "highly colourful metaphorical idioms and proverbs": *unlike, say, kick the bucket* it cannot be substituted by a single word.

Yet the sequence of words occurs 29 times in a corpus of 500,000 words taken from editorials in seven English popular and quality newspapers. The newspapers are the *Express, Guardian, Independent, Mirror, Sun, Telegraph* and *The Times*. The 500,000 words represent a module of the *Newspool Corpus* of media reporting and commenting on political matters, compiled with the help of a grant from the Italian Ministry for the University. This frequency makes it the fifth most common four-word sequence in the corpus.

Examining the contexts in which the sequence appears we discover that it is not the linguistic tick of a particular editorial writer or even the house style of (179) a single newspaper, nor can it be due to the burning question of a particular moment, for the expression occurs in 29 different editorials in all seven newspapers over the period between July 2000 and January 2002.

These combinations were garnered using the Wordlist function of *WordSmith Tools*. Scott, the author of the software, calls them *clusters* and defines them as "words which are found repeatedly in each others" company. They represent a tighter relationship than collocates, more like groups or phrases..." (1998), but Scott calls them clusters because the other terms already have uses in grammar. Effectively, they are a kind of extended collocation.

The most common 4-word clusters found in the editorial module are:

N.	Cluster	Frequency
1	the house of common	49
2	the house of lords	47
3	the end of the	43
4	the rest of the	38
5	at the heart of	29
6	of the house of	29
7	per cent of the	29
8	at a time when	27
9	at the end of	27
10	it is hard to	26

Biber refers to something very similar as *lexical bundles* (Biber et al. 1999; Biber and Conrad 1999). What differentiates these from the more primitive clusters are the following characteristics. They appear:

1. at least 10 times per million words,
2. in at least five different texts. ... (186)

N	cluster (4) FORWARD (WHB)	Frequ.
1	as we move forward	9
2	to move forward with	7
3	we look forward to	7
4	moving forward with this	6
5	president has put forward	5
6	continue to move forward	4
7	forward with this military	4
8	going to move forward	4

zu Ziff. 4.526:

Wiederverwendung / Umdeutung / Anspielungen / Ähnlichkeiten**politischer Kampf ↔ unpolitische Idylle**

zu:

nach: H. WIEGMANN, Und wieder lächelt die Thrakerin. Zur Geschichte des literarischen Humors. Frankfurt/M 2006. S. 268f:

<p>Osterspaziergang / Tucholsky Vom Eise befreit sind Strom und Bäche durch des Frühlings holden, belebenden Blick; das deutsche Volk zahlt des Krieges Zeche, und keiner bringt das Verlore- rene zurück. Die alten Monarchen, in ih- rer Schwäche, zogen sich in die Versenkung zurück. Von dorthier senden sie, fliehend nur, ohnmächtige Schauer körniger Reden (...) Hier fühl ich wahrhaft mich erhoben: "Was kümmert uns ein verlore- rener Krieg! Amerikanisches Mehl wird verschoben - nur der Schieber rettet den Sieg! Hätten wir genug zu essen wäre das Alte mit Gunst ver- gessen; Ludendorffen entbieten wir Huld Keiner ist Schuld! Keiner ist Schuld! Ich höre schon des Dorfes Getümmel, hier ist des Volkes wahrer Himmel. Zufrieden jauchzt die Reak- tion Keine Angst! sie vergessen schon! WAGNER: Mit euch, Herr Dok- tor, zu spazieren, ist ehrenvoll und ist Ge- winn. Doch würd ich nicht allein mich her verlieren, weil ich ein Feind von allem Rohen bin. Das Schreien und Sozialisie- ren</p>	<p>Osterspaziergang / Goethe Vom Eise befreit sind Strom und Bäche Durch des Frühlings holden, belebenden Blick; Im Tale grünet Hoffnungs- Glück; Der alte Winter, in seiner Schwäche, Zog sich in rauhe Berge zu- rück. Von dorthier sendet er, flie- hend, nur Ohnmächtige Schauer körnigen Eises In Streifen über die grünen- de Flur; Aber die Sonne duldet kein Weißes, Überall regt sich Bildung und Streben, Alles will sie mit Farben beleben; Doch an Blumen fehlts im Revier, Sie nimmt geputzte Menschen dafür. Kehre dich um, von diesen Höhen Nach der Stadt zurück zu sehen. Aus dem hohlen finstern Tor Dringt ein buntes Gewimmel hervor. Jeder sonnt sich heute so gern. Sie feiern die Auferstehung des Herrn, Denn sie sind selber aufer- standen, Aus niedriger Häuser dumpfen Gemächern, Aus Handwerks- und Gewerbes- Banden, Aus dem Druck von Giebeln und Dächern, Aus der Straßen quetschender Enge, Aus der Kirchen ehrwürdiger Nacht</p>
---	--

<p>ist mir ein gar verhaßter Klang das will ja nur das Volk verführen - uns Reichen wird ganz angst und bang. Wir wollen wieder die alten Zeiten, wie wollen wieder die Menge leiten Zufrieden jauchzt dann Groß und Klein Ich bin kein Mensch! Ich darfs nicht sein!</p>	<p>Sind sie alle an's Licht gebracht. Sieh nur sieh! wie behend sich die Menge Durch die Gärten und Felder zerschlägt, Wie der Fluß, in Breit' und Länge, So manchen lustigen Nachen bewegt, Und, bis zum Sinken überla- den, Entfernt sich dieser letzte Kahn. Selbst von des Berges fernen Pfadern Blinken uns farbige Kleider an. Ich höre schon des Dorfs Getümmel, Hier ist des Volkes wahrer Himmel, Zufrieden jauchzet groß und klein: Hier bin ich Mensch, hier darf ich's sein. WAGNER Mit euch, Herr Doktor, zu spazieren Ist ehrenvoll und ist Ge- winn; Doch würd' ich nicht allein mich her verlieren, Weil ich ein Feind von allem Rohen bin. Das Fiedeln, Schreien, Ke- gelschieben, Ist mir ein gar verhaßter Klang; Sie toben wie vom bösen Geist getrieben Und nennen's Freude, nen- nen's Gesang.</p>
--	--

4.53 Algorithmus

zu Ziff. 4.531:

Automatisierte Suche nach Formeln

zu:

aus: M. SCHINDELE, Computerunterstützte Suche formelhafter Rede: Protokoll des 57. Kolloquiums über die Anwendung der EDV in den Geisteswissenschaften am 6. Februar 1993: ZDV Universität Tübingen. (= LLC 9 (1994) 169-172).

(7ff)

0.1 Arbeitsbegriff. Unter formelhafter Rede verstehen wir die Verwendung "mehr oder monder starrer" Wortfolgen, denen neben ihrer unmittelbaren Wortbedeutung eine textpragmatische Funktion zukommt.

0.2 Beispiel. - Die Wortfolge "Im Namen des Volkes" kennzeichnet Gerichtsurteile.

0.3 Aufgabenstellung. - Es wird eine Konzeption und ein Programm vorgestellt, mit der zu einem gegebenen Text alle Belegstellen für "gewisse Wortfolgen" bis zu einer gewissen "Streuung" in einem anderen, grösseren Text gefunden werden. Als Belegstelle gilt zunächst eine Stelle, an der die Wortfolge vorkommt, wobei "gewisse Abstandsverhältnisse" gewahrt sind. ...

1. Formale Beschreibung des Problems

Wo nichts anderes angegeben, ist W ein Text, V ein Textkorpus mit den Bezeichnungen der Definitionen 1.1 und 1.2.

1.1 Definition. - Sei Σ ein Alphabet, Σ^+ die Menge der Wörter über Σ positiver Länge. Unter einem Text W verstehen wir eine endliche Folge $\bar{W} = (w_1, \dots, w_k)$ von Wörtern aus Σ^+ . Die Menge $M_W = \{1, \dots, k\}$ ist die Indexmenge von W . $k = |M_W|$ ist die Mächtigkeit von M_W , das heisst die Länge von W .

1.2 Definition. - Unter einem Textkorpus $V = (V_1, \dots, V_m)$ verstehen wir eine endliche Folge von Texten $V_i = (V_{i1}, \dots, V_{ik_i})$. $M_{V_i} = \{1, \dots, k_i\}$ ist die Indexmenge von V_i .

1.3 Definition. - Sei $U = (u_1, \dots, u_1)$ eine Wortfolge, $W = (w_1, \dots, w_k)$ ein Text mit den Indexmengen M_U und M_W . $f : M_U \rightarrow M_W$ sei eine streng monoton wachsende Funktion, derart, dass für alle $i \in M_U$ gilt, dass $u_i = w_{f(i)}$. Dann sagen wir, dass U in W vorkommt, und die Menge $F_W = \{f(i) \mid i \in M_U\}$ heisst eine Fundstelle von U in W . f heisst die Verweisfunktion von U nach W , $\sigma_{F_W} = f(l) - f(1) + 1$ heisst die Streuung von F_W .

1.4 Bemerkung. - Während es zu jeder Wortfolge U keine, genau eine oder mehrere Fundstellen in W geben kann, gibt es umgekehrt zu jeder Teilmenge F_W der Indexmenge M_W genau eine Wortfolge U , von der F_W Fundstelle ist.

1.5 Satz. - Sei W ein Text mit Indexmenge M_W . Dann sind die Teilmengen von M_W bezüglich der Relation \subseteq partiell geordnet.

1.6 Satz. - Sei W ein Text mit Indexmenge M_W . $P(M_W)$ sei die Menge der Teilmengen von M_W . $<_*$ sei eine totale (lineare) Ordnung auf $P(M_W)$ derart, dass für alle Teilmengen F_{W_i} und F_{W_j} von M_W gilt, dass aus $|F_{W_i}| < |F_{W_j}|$ folgt, dass $F_{W_i} <_* F_{W_j}$ ist. Dann ist $<_*$ eine totale (lineare) Ordnung auf $P(M_W)$, die die partielle Ordnung aus 1.5 respektiert.

1.7 Definition. - Als Standardordnung auf einer Fundstellenmenge bezeichnen wir im folgenden die totale Ordnung, die die Fundstellen als aufsteigend geordnete Mengen erstens der Mächtigkeit nach absteigend und zweitens lexikalisch aufsteigend anordnet.

1.8 Definition. - Unter einem Verweis Q von einem Text W in ein Textkorpora V verstehen wir ein geordnetes Tripel (F_W, U, F_{V_i}) bestehend aus einer Wortfolge U und Fundstellen F_W, F_{V_i} von U in W und in einem V_i in V .

1.9 Definition. - Als Standardordnung auf einer Verweismenge bezeichnen wir im folgenden die totale Ordnung, die die Verweise erstens aufsteigend ordnet bzgl. der Standardordnung auf den W -Fundstellen, zweitens aufsteigend anordnet bzgl. der Textnummer der V -Fundstellen und drittens aufsteigend anordnet bzgl. den Standardordnungen auf den Fundstellen der Texte von V .

1.10 Definition. - Unter einem einpunktigen Verweis verstehen wir einen Verweis $Q = (F_W = \{j_W\}, U, F_{V_i} = \{j_{V_i}\})$, also mit $|F_W| = 1$. Q ist durch j_{V_i}, i und $Q_d := j_{V_i} - j_W$ eindeutig bestimmt.

1.11 Definition. - Als Standardordnung auf einer Menge einpunktiger Verweise bezeichnen wir im folgenden die totale Ordnung, die die Verweise erstens aufsteigend anordnet bzgl. der Textnummer der V -Fundstellen, zweitens aufsteigend anordnet bzgl. Q_d und drittens aufsteigend anordnet bzgl. den Standardordnungen auf den Fundstellen der Texte von V .

1.12 Definition. - Sei $Q' = (F_W, U = (u_1, \dots, u_k), F_{V_i})$ ein Verweis. f_W, f_{V_i} seien die zugehörigen Verweisfunktionen von U nach W bzw. V_i . Q heisst homogen, wenn für alle m, n in M_U (der Indexmenge von U) gilt $f_{V_i}(m) - f_W(m) = f_{V_i}(n) - f_W(n)$, das heisst, wenn für die einpunktigen Verweise $Q_1 = (\{w_{f_W(1)}\}, u_1, \{v_{f_{V_i}(1)}\}), \dots, Q_k = (\{w_{f_W(k)}\}, u_k, \{v_{f_{V_i}(k)}\})$ gilt $Q_{1,d} = \dots = Q_{k,d}$.

1.13 Definition. - Sei $Q = (F_W, U, F_{V_i})$ ein Verweis. Q heisst maximal bzgl. seiner Streuung oder σ -vollständig, wenn für jeden Verweis $Q' = (G_W, U', G_{V_i})$ mit $F_W \subset G_W$ und $F_{V_i} \subset G_{V_i}$ gilt $\sigma_{G_W} > \sigma_{F_W}$.

1.14 Bemerkung. - Homogene Verweise sind nicht notwendig σ -vollständige, σ -vollständige Verweise nicht notwendig homogen.

2. Maschinelle Lösung des Problems: Die Spezifikation

Eine Lösung des Problems wurde in der Abteilung für Methodik computerunterstützter Textinterpretation der Fakultät für Informatik in TUSTEP implementiert, und liefert zufriedenstellende Ergebnisse für alttestamentliche (hebräische) Textkorpora. Sie ist derzeit folgendermassen spezifiziert:

2.1 Eingabedaten. - Ein Text W , ein Textkorpora V ; diese müssen "gewissen" formalen Kriterien genügen. Ferner ein $n \in \mathbb{N}$ mit $2 \leq n \leq 99$. W muss in V enthalten sein, d.h. es gibt einen Text V_i in V mit einer Fundstelle von W der Streuung k .

2.2. Ausgabedaten. -

$\{Q = (F_W, U, F_{V_i}) \mid Q \text{ homogen} \wedge (Q \text{ } \sigma\text{-vollständig} \vee \exists Q' = (F'_W, U, F'_{V_j}) Q' \text{ } \sigma\text{-vollständig und homogen}) \wedge 1 < \sigma_{F'_W} \leq n\}$ als Liste mit "gewissem" Format.

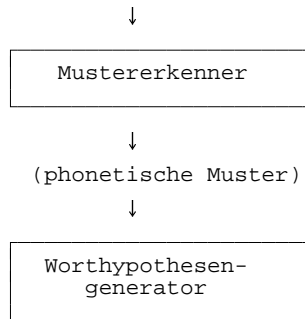
zu Ziff. 4.533:

Dialogsystem Speedis

zu:

aus: ... implementiert bei Daimler-Benz. Aus einem Papier von Anke Kölzer: "Methoden zur Spezifikation von Wissensquellen für ein natürlichsprachliches Dialogsystem".

(Laute) Die akustischen Signale, die über Telefon oder Mikrofon in das Dialogsystem eingehen, werden von diesem Modul vorverarbeitet. Mit Hilfe von statistisch basierten Sprachmodellen und Aussprachelexika wird ein Worthypothesengraph mit bewerteten Kanten erzeugt, die widerspiegeln, wie gut die jeweilige Hypothese ist.



↓

(Worthypothesen)

- (1) ich
- (2) moechte | moechten
- (3) morgen | morgens
- (4) noch | nach
- (5) hamburg | hamm | haben
- (6) uhr
- (4) + (6) hamburg
- (5) + (6) hamburg | hamm
- (5) + (6) homburg
- (7) #pause#
- (5) - (8) hamburg

.....

zu Ziff. 4.534:

"Syntax" - Programmierung

zu:

- im AB "Textwissenschaft" erstes funktionsfähiges Suchprogramm (Anfang der 1990er Jahre) von MARTIN SCHINDELE. Lief noch am Großrechner (PC waren noch nicht leistungsfähig genug) und unter TUSTEP. Ausführlich erhoben: Hebräische Josefsgeschichte im Verhältnis zum gesamten Alten Testament. Publikation in H. SCHWEIZER (ed.), Computerunterstützte Textinterpretation, 1995. 3 wesentliche Merkmale des Programms: (a) Auf *identische* Entsprechungen ausgerichtet, (b) bei der Länge der Wortkette im Suchtext wird nur eine Mindestzahl (z.B. 3) eingegeben. Der Algorithmus erkennt selbst, bis zu welcher Maximallänge Entsprechungen im Korpus zu finden sind. (c) Was man als "Einzeltext = Suchtext" eingibt, wird sequenziell automatisch im "Korpus" auf Entsprechungen hin geprüft.
- Neuprogrammierung unter JAVA durch MICHAEL HÜTTER, 2002. Lauffähiges Programm, auf *identische* Entsprechungen ausgerichtet.
- Neuprogrammierung als Diplomarbeit unter JAVA durch MICHAEL PACH (2005). Viele Variablen für die Suche einstellbar: Thema "Ähnlichkeit" (Maß an Ähnlichkeit, Permutationen, Affixe/Suffixe). Dadurch aufwändige Speicherverwaltung notwendig. In dieser Form - nach weiteren Verbesserungen (S. LAUCKNER) - Bestandteil des SLANG-Projekts.
- Weitere Informationen und Materialien unter: <http://www-ct.informatik.uni-tuebingen.de/ct/syntax>
- Aus einer Buchankündigung, März 2009: Chris Greaves, **ConcGram 1.0** A phraseological search engine. Studies in Corpus Linguistics Software, 1. CD-ROM. "ConcGram 1.0 is a corpus linguistics software package which is specifically designed to find all the co-occurrences of words in a text or corpus irrespective of variation. The software finds the cooccurences fully automatically; in other words, the user inputs no prior search commands. These co-occurrences are termed 'concgrams' and include instances of both constituency variation (*increase in expenditure, increase in the share of public expenditure*) and positional variation (*expenditure would inevitably increase*). ConcGram 1.0 is therefore uniquely suited to uncovering the full phraseological profile of a text or corpus. In addition to this principal function, ConcGram 1.0 allows the user to nominate concgram searches, and it also has the more traditional functions associated with corpus linguistics software."
- Die Publikation entspricht offenbar dem genau gleichen Konzept, das wir in den letzten 2 Jahrzehnten entwickelt haben.

4.54 Relative Chronologie**Literatur**

- HELLWIG, O: A chronometric approach to Indian alchemical literature. *Literary and Linguistic Computing* 24 (2009) 373-383.
- HILPERT, M; GRIES, S T: Assessing frequency changes in multistage diachronic corpora. Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing* 24 (2009) 385-401.
- ROOS, T; HEIKKILÄ, T: Evaluating methods for computer-assisted stemmatology using artificial benchmark data sets. *Literary and Linguistic Computing* 24 (2009) 417-433.
- STAMOU, C: Stylochronometry: Stylistic Development, Sequence of Composition, and Relative Dating. *Literary and Linguistic Computing* 23/2 (2008) 181-199.

zu Ziff. 4.541:**Intertextualität, Anspielung**

zu:

aus: Benjamin D. Sommer, *Exegesis, Allusion and intertextuality in The Hebrew Bible: A response to Lyle Eslinger: Vetus Testamentum XLVI (1996/4) 479-489.*

(484) If two texts share vocabulary items that are commonplace in Biblical Hebrew, the parallel between them is most likely coincidental. If they share terms that often appear together in biblical or ancient Near Eastern texts, then there is a strong likelihood that they independently draw on traditional vocabulary clusters. If (485) the vocabulary is neither common nor part of a known vocabulary cluster, then the possibility of genuine borrowing is strong. If a text repeatedly alters the wording or ideas of earlier texts in certain ways, or if it displays a particular preference for certain texts, then examples of shared vocabulary which display those tendencies are likely to represent genuine cases of borrowing.

The argument that an author alludes, then, is a cumulative one: assertions that allusions occur in certain passages become stronger as patterns emerge from those allusions. In any one passage that may rely on an older text, the critic must weigh evidence including the number of shared terms and their distinctiveness, the presence of stylistic or thematic patterns that typify the author's allusions, and the likelihood that the author would allude to the alleged source. Weighing such (486) evidence and hence identifying allusions constitutes an art, not a science, but the recognition that the study of literary history is no science should provoke neither surprise nor despair...

The terms allusion and intertextuality are not entirely compatible. "Allusion", as used by literary critics, does posit an earlier and a later text, so that the study of allusion necessarily involves a diachronic component. "Intertextuality", as most literary critics use the term, focuses on manifold linkages among texts or on connections between a text and commonplace phrases from the cultural systems in which the text exists. These connections do not arise exclusively from an intentional and signalled use of an earlier text, such as citation (which might be studied under the rubric of influence or allusion). Rather, they also result from the way a text reflects its linguistic, aesthetic, or ideological contexts; other texts may share those contexts, and hence readers may notice links among many texts, whether the authors of the texts (487) knew each other or not. Thus an intertextual approach differs markedly from approaches concerned with "influence" and "allusion". Intertextuality is concerned with the reader or with the text as a thing independent of its author, while an approach oriented towards allusion is concerned with the author as well as the text and reader, the study of intertextuality is synchronic, the analysis of allusion diachronic or even historicist; intertextuality is interested in a wide range of correspondences among texts, allusion with a more narrow set; intertextuality examines the relations among many texts, while allusions represent specific connections between a limited number of texts. The distinction between intertextuality on the one hand and allusion and influence on the other is basic to contemporary theoretical discussions of the relations between texts, though many readers (including Eslinger) continue to confuse them.

aus: Lyle Eslinger, Inner-biblical exegesis and inner-biblical Allusions: The Question of category: Vetus Testamentum XLIII (1992/1) 47-58.

(57) The study of the Bible has moved through at least two stages. Just now it seems caught between the second and a new third stage. First, it was read as history and its plot was taken for the real sequence of events that it describes. Second, in a reactionary movement still dominated by concern for history but now (58) suspicious of the history of the plot it portrays, the Bible was read as a reflection, both of the history that its plot lays out and, more clearly, of the period and society in which it was written. It is out of this second stage of reading that the study of i.b.e. has sprung. Lastly, the Bible is being read without regard for the issues of history and historicity. This shift in focus should not, as it so often is, be taken as a rejection of historical study: it is not. Rather, it is a conscious decision to focus on a given, biblical literature, and a rejection of an appropriation of this given for inappropriate purposes - the wringing of history from a literature whose historiographical purpose, if it has one, is unstated and, so far, not demonstrated.

zu Ziff. 4.542:

Feste Wortketten ⇒ Chronologie / Assoziationshintergrund

zu:

Im Fall der alttestamentlichen Josefsgeschichte (Gen 37 - 50) ist der automatisch erhobene Befund folgender:

- nahezu keine Formeln / feste Wortketten, die häufig sind und quer durch die hebräische Bibel nachgewiesen werden können;
- regelmäßig finden sich Wortketten - z. T. Länge 7 -, die so nur *einmal* in einem anderen Teil des Alten Testaments nachweisbar sind

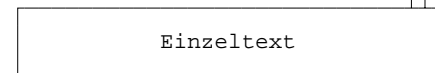
Vergleichskorpus

Buch 1 (um 750 v.Chr.)

Buch 2 (um 650 v.Chr.)

Buch 3 (um 550 v.Chr.)

Buch 4 (um 450 v.Chr.)



Buch 5 (um 350 v.Chr.)

Buch 6 (um 300 v.Chr.)

Auswertung:

- es liegt das Gegenteil allgemein verfügbarer Floskeln vor: **gezielte Anspielungen**
- den jeweiligen zweiten Text im Hintergrund muss man zur Interpretation heranziehen; er liefert beabsichtigte Assoziationen
- Streuungsrichtung beachten: (a) es ist nicht so, dass die Sprache der Josefsgeschichte sprachprägend gewirkt hat, also viele andere Texte Wortketten nachgeahmt hätten (eben das ist nicht nachweisbar); (b) der Befund ist umgekehrt: die Josefsgeschichte "pflückt" immer wieder gezielt aus schon vorliegenden Texten
- dadurch ist zunächst eine *relative Chronologie* möglich: die Josefsgeschichte muss später als die benutzten Texte entstanden sein
- kennt man die Entstehungszeit der anderen Texte, ergibt sich daraus - als Annäherung - eine *absolute Chronologie*
- durch Bereitstellung solcher Befunde kann der Rechner das Lesen von Texten unterstützen, deren Umfeld uns heute nicht mehr so präsent ist (hier: Wortlaut der hebräischen Bibel)
- Vgl. <http://www-ct.informatik.uni-tuebingen.de/daten/pattl.pdf> - in den Essay (zur Josefsgeschichte) sind immer wieder Beispiele solcher Verweise eingebaut

zu Ziff. 4.543:

Textkritik // Bioinformatik>>Verwandtschaften

zu:

aus: Vgl.

http://www.eurekalert.org/pub_releases/2009-02/dbnl-ftw021109.php

"feature frequency profiles" (FFP), that makes it possible to compare, classify, index and catalog just about any type of linear information that can be electronically stored. The kinds of information that can be analyzed with the FFP technique include nucleotide base and amino acid sequences, books, documents and possibly images. It could even prove to be the ultimate music organizer. ...

For example, Kim and his colleagues used the FFP technique to create a book tree composed of more than two dozen selected works under the categories of philosophy, mythology, religion, 19th Century fiction, science fiction and children's fiction. Their FFP-based book tree correctly grouped all books by category and author including some, such as the Koran, that were misplaced in a book tree based on a standard word frequency profile analysis. In the case of the Koran, the FFP-based tree placed it in the religion category on the same branch as the King James Bible and the Book of Mormon, whereas the word frequency book tree grouped it in the philosophy category, on the same branch as Plato's 'The Republic' and Socrates' 'The Apology'.

...

First current standard text analysis is based on the frequency at which different words appear, but genomic data consist of long strings of letters not words. Second, analysis based on the frequency of words does not provide local syntax - the relationship between adjoining words, a point that is critical in comparative genomics and turned out to be important in text comparisons as well. ...

"Text features longer than eight or nine letters do not occur frequently enough for frequency profile comparisons, and text features shorter in length do not give us enough information to distinguish one book from another", Kim says.

To apply their FFP technique to comparative analysis of books, they "delimiter-stripped" each book - meaning they stripped the text of all punctuation and spaces - then transformed the text into a single long string of letters. A "window" of eight letters in length was then advanced across this string one letter at a time, yielding a frequency profile of the features in which overlapping sequences of text reveal relationships between individual features. Comparing the feature frequency profile for each book analyzed produced astonishingly accurate trees that grouped books by author, genre or historical era.

4.55 Interpretation

zu Ziff. 4.551:

Situativer Kontext / kommunikativ-sprachlicher Kotext

zu: vprag7.01

aus: SOTTONG, H; MÜLLER, M, Zwischen Sender und Empfänger. Eine Einführung in die Semiotik der Kommunikationsanalyse. Berlin 1998.

(81f) Wenden wir uns zunächst dem einfacheren der beiden Begriffe zu, dem Kontext. In der Linguistik wird Kontext allgemein als die Umgebung definiert, in der eine bestimmte Äußerung vorkommt, und es wird zwischen dem situativen und dem normativen Kontext unterschieden. Es spielt also bei der Interpretation einer Äußerung zunächst eine Rolle, wer zu wem spricht, in welcher sozialen und/oder räumlichen Umgebung die Äußerung stattfindet, mit welchen Handlungen die Äußerung verknüpft ist. Die Anrede "Na du kleines Scheißerchen" von einer Mutter gebraucht, die ihrem Baby gerade die Windeln wechselt, hat offenkundig eine erheblich andere Bedeutung, als wenn die gleiche Äußerung von einem Schüler gegenüber einem Lehrer gemacht wird. Während das doppelte Diminutiv im ersten Falle nicht nur als liebevoll betrachtet werden kann, sondern auch die Relation zwischen der Sprecherin und dem Angesprochenen adäquat wiedergibt, ist es im zweiten Falle genau umgekehrt. Hier besteht die Botschaft der Äußerung primär darin, daß durch den Gebrauch dieser Äußerung in diesem Kontext eine soziale Regel massiv verletzt und die bestehende soziale Struktur verbal in ihr Gegenteil verkehrt wird: Der sozial rangniedere, jüngere spricht den ranghohen, erwachsenen Adressaten so an wie eine Mutter ihr Baby. Zudem werden hier dem Lehrer, dessen soziale Rolle durch ein Mehr an Macht, Wissen und Autorität definiert ist, all diese Eigenschaft implizit abgesprochen und unterstellt, er verfüge - gleich einem Baby - noch nicht einmal über die grundlegende zivilisatorische Fähigkeit der Reinlichkeit. Diese wesentlichen interpretatorischen Folgerungen aus dieser Äußerung sind nur aufgrund der Kenntnis des Kontextes und darüber hinaus des Wissens über soziale Codes, institutionelle Regeln von Schule etc. zu ziehen. Um die in diesem Falle provokative Bedeutung der Anrede zu verstehen, genügt in unserem Beispiel das Alltagswissen und die Alltagserfahrung...

(82) Kotexte sind selbstverständlich auch jenseits und außerhalb von Gesprächen relevant, wenn sich die Äußerung implizit oder explizit auf eine andere Äußerung bezieht. Der offenkundigste Fall einer solchen Referenzbeziehung ist das Zitat, wobei durch die Einbindung einer Teiläußerung in die aktuelle Äußerung potentiell die gesamte zitierte Äußerung als Kotext für diese relevant gesetzt wird. Zitiert werden kann in diesem Zusammenhang potentiell jede vorhergehende Äußerung, sei sie nun sprachlich oder nicht, künstlerisch oder nicht künstlerisch, Bestandteil des kulturellen Allgemeinwissens oder subkultureller bzw. privater Natur. Ob der Rezipient der Äußerung das Zitat als solches erkennt oder nicht und ob er über das dazugehörige kotextuelle Wissen verfügt oder nicht, ist dabei insofern relevant, als dieses Wissen eine Bedingung dafür ist, inwieweit er die Äußerung seines Gegenüber vollständig interpretieren kann.

Neben dem direkten Zitieren, also der Reproduktion einer bereits existierenden Teiläußerung innerhalb einer aktuellen Äußerung können Kotexte auch durch Erwähnungen relevant gesetzt werden: Beispielsweise durch die Nennung von Buch- oder Filmtiteln oder Figuren, die in Büchern oder Filmen vorkommen. Oder durch Nennung von Eigennamen, die mit der Produktion bestimmter Äußerungen verknüpft werden können: dabei kann es sich um Autorinnen und

Autoren handeln oder um eine Person, die für ganz bestimmte Äußerungen bekannt ist.

Schließlich kann es aber auch der Kontext selbst sein, der ganz bestimmte Kotexte aktiviert. Nehmen wir an, in den Fernsehnachrichten wird ein Auszug aus einer Politikerrede mit folgenden Wortlaut gesendet: "Wir Deutsche blicken nicht ohne Stolz auf die Tradition, für die dieser Ort steht, und wir wollen diese Tradition auch in Zukunft mit all unserer Kraft aufrechterhalten." Offenkundig ist für die Bedeutung dieser Äußerung von entscheidender Relevanz, wer sie an welchem Ort macht. Wenn wir beispielsweise aufgrund kotextuellen Wissens den Sprecher als den amtierenden Präsidenten des Bundestages und den Ort als den Plenarsaal des Parlaments identifizieren, dann können wir die angesprochenen Traditionen mit ziemlicher Sicherheit als die der immer noch jungen Demokratie in Deutschland identifizieren. Wenn aber derselbe Satz, von demselben Politiker vor der Münchner Feldherrnhalle gesprochen wird, dann wird mit dieser Äußerung offenbar etwas völlig anderes behauptet, das weitgehend oppositionell zur Bedeutung des zuerst geschilderten Falls ist - denn der situative Kontext "Feldherrnhalle" evoziert Kotexte über die präfaschistische und faschistische Geschichte dieser Lokalität. Hier wird deutlich, daß Kontext und Kotexte in einer engen Wechselbeziehung zueinander stehen.

zu Ziff. 4.552:

Kontrast zum Korpus

zu: vprag7.224

aus: W. GRIMM, Er nicht! Der Gottesknecht als Verweigerer des Heiligen Kriegs. Ein neues Verständnis von Jes 42,1-4: Biblische Notizen 138 (2008) 47-61.

(47)

Siehe, mein Knecht - ich stütze ihn,
mein Erwählter - Gefallen habe ich an ihm gefunden.
Ich gab meine Geist auf ihn,
Recht bringt er zu den Völkern hinaus.

Er schreit nicht herbei und er nimmt nicht hoch.
und er lässt draußen seine Stimme nicht hören.
Das geknickte Rohr zerbricht er nicht.
und den glimmenden Docht löscht er nicht aus. ...

(48) Dies provoziert sogleich die Frage: In welchem Sinne schreit der Knecht nicht? ...

Das Verb bezeichnet in einer Spezialbedeutung im politisch-militärischen Gebiet "das Aufbieten zur Heeresnachfolge (zum Heerbann) von wehrhaften Männern (und Streitwagen) von Beginn des Kampfes ..., sowie das dem Aufgebot folgende "sich Zusammenscharen" des Heerbanns ... Bezeichnend ist, dass ... eine konkrete Not durch Bedrohung übermächtiger äußerer Feinde die Ursache für den Aufgebotsruf ist, der das ganze Volk oder eine bzw. mehrere Stämme zur bewaffneten Hilfeleistung herbeiruft." (635f) Die Belegstellen, die diesem Summarium zugrunde liegen, sind in der biblischen Reihenfolge: Jos 8,16; Ri 4,10.13; 6,34.35; 7,23.24; 10,17; 12,1.2; 18,22.23; 1Sam 13,4; 14,20; 2Sam 20,4-5; 2Kön 3,21. ...

(49) "Er schreit nicht herbei", bricht mit einer Praxis der Führer im Heiligen Krieg, es verweigert sich der den JHWH-Krieg vorbereitenden Einberufung der Krieger. Ich halte es für möglich, dass der Prophet dabei speziell Gideon im Heiligen Krieg gegen die Midianiter vor Augen hatte. ...

Das zweite "Er nicht"-Sätzchen lautet: $w=lo(?) yi[n]šša(?)$ = "und er nimmt nicht hoch" bzw. "und er trägt nicht". (Das hier stehende Verb kann sowohl den Moment des Hochhebens als auch das andauernde Tragen eines Gegenstandes bezeichnen.) Aber was nimmt der Gottesknecht partout nicht hoch? ...

Im Zusammenhang mit dem Fall der Mauern von Jericho geht dem Ertönen der Widderhörner jeweils voraus, dass die Priester sie "hochnehmen" bzw. "tragen": "Sieben Priester sollen sieben Widderhörner vor der Lade her tragen, und am siebten Tag sollt ihr siebenmal die Stadt umschreiten, und die Priester sollen in die Widderhörner stoßen. (Jos 6,4) ... Sieben Priester sollen sieben Widderhörner vor der Lade JHWHs her tragen (6,6) ... Es geschah, wie Josua dem Volk gesagt hatte: Die sieben Priester, die die sieben Widderhörner vor der Lade JHWHs her trugen, zogen hinüber, stießen in die Widderhörner (6,8) ... Die sieben Priester trugen die sieben Widderhörner vor der Lade JHWHs, indem sie dauernd in die Widderhörner stießen, und die Kriegsmannschaft ging vor ihnen her, und die Nachhut ging hinter der Lade JHWHs, indem sie dauernd in die Widderhörner stieß (6,13)". ...

(51) "und nicht lässt er draußen seine Stimme hören". Es bezieht sich ganz offensichtlich auf das traditionelle Kriegsgeschrei des

JHWH-Volks. "Seine Stimme hören lassen" ist im Heiligen Krieg gegen Jericho ein synonym paralleler Ausdruck zu HRY' "das Kriegsgeschrei erheben". Josua befiehlt, unmittelbar nach dem Blasen der Widderhörner (6,9), dem Kriegsvolk: "Ihr sollt (vorläufig noch) nicht das Kriegsgeschrei erheben und eure Stimme (noch) nicht hören lassen und kein Wort soll aus eurem Munde hergehen, bis zu dem Tag, an dem ich euch sage: "Erhebt das Kriegsgeschrei!" (6,10; vgl. auch 2Kön 7,6) ...

Der Gottesknecht verweigert sich in lapidarer Kürze dem dreiteiligen Eröffnungszeremoniell des Heiligen Krieges: dem Herbeirufen der Krieger (vgl. Ri 6,34f; 7,23f); dem Fanfarenstoß des Schofar (Jos 6; Ri 7,19-22); dem Kriegsgeschrei der Krieger des JHWH-Volks (Jos 6,5.10; Ri 7,20b-21). ...

Der Anfang des Heiligen Krieges was das dreiteilige Eröffnungsritual - der Gottesknecht sagt "Nein!" dazu. Das Ende und Ziel des Heiligen Krieges war Vernichtung der Besiegten. Für diesen finalen Akt finden sich im Jesajabuch die Metaphern vom "angebrochenen Rohr" und vom (52) "glimmenden Docht". Da wird etwa ein im Krieg stark angeschlagener Pharao als "angebrochener Rohrstab" bezeichnet (Jes 36,6). Der Aramöerkönig, vor dem sich der jüdische König Ahas im syrisch-ephraimitischen Krieg so fürchtet, wird als "nur noch rauchender Brandscheitstummel" kleingeredet (Jes 7,4). ...

(53) Der Knecht JHWHs indessen verweigert sich - mit seinem vierten und fünften Nein - auch diesem finalen Aspekt des Heiligen Kriegs. Er verbindet die Absage mit zwei Metaphern, die sich, wie wir sahen, im Jesajabuch auf die restlose Auslöschung der Feinde beziehen: "Das geknickte Rohr zerbricht er nicht und den glimmenden Docht löscht er nicht aus." Er nicht: Vor dem Hintergrund der Praxis Josuas und Gideons präsentiert der Prophet den Minister JHWHs als radikalen Verweigerer, als die personifizierte Kritik jedes im Namen Gottes geführten Krieges.

4.6 Bewußtheitsgrade bei der Textrezeption und ausdrucksyntaktische Realisierungen (Zitate, Anspielungen, Gruppensprache, Formeln, vergleichbare Konstruktionen, relative Chronologie usw.)

Literatur

- BLUHM, L; RÖLLEKE, H: "Redensarten des Volks, auf die ich immer horche": Märchen - Sprichwort - Redensart: zur volkspoetischen Ausgestaltung der Kinder- und Hausmärchen durch die Brüder Grimm. Stuttgart 1997: Hirzel // I.2.7
- BURGER, H; DOBROVOL'SKIJ, P; KÜHN, P; NORRICK, N R (ed./eds.): Phraseologie / Phraseology - Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of contemporary Research. Handbücher zur Sprach- und Kommunikationswissenschaft 28 / 1. Berlin 2007 3-11-017101-5: deGruyter // I.2.7
- BURGER, H; DOBROVOL'SKIJ, D; KÜHN, P; NORRICK, N R (ed./eds.): Phraseologie / Phraseology - Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science. Handbücher zur Sprach- und Kommunikationswissenschaft 28 / 2. Berlin 2007 3-11-019076-1: deGruyter // I.2.7
- HORNERO, A M; LUZÓN, M J; MURILLO, S (ed./eds.): Corpus Linguistics. Applications for the Study of English. Linguistic Insights. Studies in Language and Communication 25. Bern 2008 978-1-4331-0667-5ph: Peter Lang // I.2.7
- IRIZARRY, E: Tampering with the Text to Increase Awareness of Poetry's Art. Theory and Practice with a Hispanic Perspective. Literary and Linguistic Computing 11 (1996 / 4) 155-162.
- SÁNCHEZ, M A: From Words to lexical Units. A Corpus-Driven Account of Collocation and Idiomatic Patterning in English and English-Spanish. Studien zu romanischen Sprachwissenschaft und interkulturellen Kommunikation 35. Frankfurt / M 2006 3-631-55759-0: Peter Lang // I.2.4
- TWEEDIE, F J; HOLMES, P D u.a.: The Provenance of 'De Doctrina Christianana' attributed to John Milton: A statistical Investigation. Literary and Linguistic Computing 13 (1998) 77-87.

4.61 Begriffsbildung

zu Ziff. 4.611:

Intertextualität in traditioneller Sicht

zu: vprag7.51101

aus: HERWIG, H, Literaturwissenschaftliche Intertextualitätsforschung im Spannungsfeld konkurrierender Intertextualitätsbegriffe: Zeitschrift für Semiotik, 24/2-3 (2002) 163-176

(164f) Das Phänomen der Intertextualität ist so alt wie die Textlichkeit. Von ihm sprechen bereits die antiken Poetiken. Texte existieren nicht in einem "Vakuum" (Ascel 2001: 287), sie sind immer auf andere Texte bezogen, "in einem schon vorhandenen Universum der Texte" (Stierle 1983: 7) situiert. "Jeder Text wird vom Produzenten wie vom Rezipienten mit Bezug auf Textwissen und Texterfahrung, d.h. vor dem Hintergrund zuvor produzierter und rezipierter Texte [...], wahrgenommen" (Fix 2000: 449). Auch die Konzeptualisierung von literarischen Gattungen oder Textsorten ist "ohne die Annahme intertextueller Bezüge undenkbar" (Aczel 2001: 287). Diesen Sachverhalten hat die literaturwissenschaftliche Quellen- und Einflussforschung immer schon Rechnung getragen. Sie hat ein reichhaltiges Begriffsrepertoire für die Beziehungen von Texten auf Vorläufertexte entwickelt, das von "Adaption", "Anspielung", "Cento", "Collage", "Epikrise", "Glosse", "Imitation", "Kommentar", "Kontradiktio", "Kontrafaktur", "Montage", "Nachbildung", "Palimpsest", "Palinodie", "Parodie", "Pastiche", "Persiflage", "Plagiat", "Quelle", "Reminiszenz", "Rezension", "Satire", "Transposition", "Travestie" bis zu "Übersetzung", "Zitat", und "Zusammenfassung" reicht. Die Einflussforschung fragt danach, was ein literarischer Text mit Vorläufertexten teilt, von diesen übernimmt. Diese Übernahmen können von einzelnen Wörtern über Sätze, syntaktische Merkmale, Bilder und (169) Motive bis zu Textsortenmustern, stilistischen Mitteln und allgemeinen Texteigenschaften reichen, vom Einzelwort bis zum ganzen Text. Die Aufmerksamkeit richtet sich dabei auch auf die Art der Adaption, auf die Frage, ob der spätere Text den Prätext oder das Textmuster bestätigt, verneint, ironisiert, modifiziert, collagiert oder verfremdet. Auch motivgeschichtliche Untersuchungen, Gattungstheorie, Stilistik und Toposforschung kommen nicht ohne die Berücksichtigung intertextueller Beziehungen aus.

zu Ziff. 4.612:

imitatio / Intertextualität

zu: vsyst7.52

aus: FIX, U, Aspekte der Intertextualität in: BRINKER, K; ANTOS, G; HEINEMANN, W; SAGER SF (eds.), Text- und Gesprächslinguistik. Ein internationales Handbuch der zeitgenössischen Forschung. HSK 16.1. Berlin 2000. S. 449-457 //I.2.4

(449) Der Gedanke, daß Texte sich immer auf Texte beziehen (müssen), liegt allen Auffassungen von Intertextualität zugrunde. Jeder Text wird vom Produzenten wie vom Rezipienten mit Bezug auf Textwissen und Texterfahrung, d.h. vor dem Hintergrund zuvor produzierter und rezipierter Texte, in der Kontinuität des jeweiligen Umgangs mit Texten wahrgenommen. Streng genommen hat man nur den allerersten Text, den man in seinem Leben als Text erfahren hat, ohne intertextuellen Bezug erlebt. Generell gilt, daß es unmöglich ist, außerhalb der Welt der Texte und unabhängig von ihr zu kommunizieren.

Hinzu kommt die speziellere Erfahrung des Sprachteilnehmers, daß es bei diesem generellen Textbezug nicht bleiben muß. Ein konkreter Text, ein Textexemplar kann sich auch - formal oder inhaltlich - auf andere Textexemplare beziehen, indem er Inhalts- oder Formelemente von Vortexten oder auch beides aufgreift bzw. auf Folgetexte verweist. Und unbewußt oder bewußt greift jeder Sprachteilnehmer, wenn er es mit Texten zu tun hat, auch auf sein Wissen über Textmuster zurück, bezieht sich also auf kulturell geregelte Konventionen des Schreibens und Sprechens und stellt jeweils intertextuelle Text-Textmuster-Beziehungen her. Selbst wenn er die Muster bricht, stellt er, sofern er dies nicht aus Unkenntnis, sondern bewußt tut, eine Text-Textmusterbeziehung her. Mit drei Typen intertextueller Beziehungen haben wir es also zu tun: Text-Textwelt-Beziehung, Text-Text-Beziehung und Text-Textmuster-Beziehung. ...

(451) Die Sprachwissenschaft kann mit dem Intertextualitätsbegriff der Literaturwissenschaft, wie sie meint, nichts anfangen. Vornehmlich stößt sie sich, wie allenthalben zu lesen ist, an der Weite, an dem "Schillernden" des Begriffs. Diese Vorstellung ist schon zum alltagsweltlichen Topos geworden, auf den sich beziehend man den Sinn der Beschäftigung mit Intertextualität in Frage stellt. Einen eigenen, auf ihre Bedürfnisse hin zugeschnittenen Begriff hat die Sprachwissenschaft lange Zeit nicht im Blick gehabt und nicht zu etablieren versucht. ...

(452) Genettes (1982) viel beachteter Versuch, Textbeziehungen zu differenzieren, setzt an beim Begriff der *Transtextualität* = erkennbare Präsenz eines Textes in einem anderen; *Paratextualität* = den Text begleitende Textsorten; *Metatextualität* = kommentierende Bezugnahme eines Textes auf einen Prätext; *Hypertextualität* = ein Text als die Folie eines anderen; *Architextualität* = Bezug auf Texttraditionen. (Auf die daran anschließenden Subkategorisierungen kann hier nicht näher eingegangen werden.) ...

Neben *Intertextualität* begegnen verwandte Begriffe wie *Intermedialität*, *Hypertext*, *Interaktivität* - Folge der Möglichkeit digitaler Codierung von Texten -, alle in inhaltlicher Nähe zum Begriff der Intertextualität, alle aber Ausdruck neuer Sehweisen auf Texte (neuer Art), auf Textbeziehungen (neuer Art) und neuer Möglichkeiten des Umgangs damit. *Intermedialität* bezieht sich auf multimediale Kommunikation. *Hypertext* meint das Ergebnis dieser

Art intermedialen Handelns, den holistischen Text, der ein Verbund sein kann von digital gespeichertem Sprachtext mit Bildern, Tonmaterial, Filmen, Graphiken u.ä., also mit Texten anderer semiotischer Kodes (vgl. Hess-Lüttich 1997). *Interaktivität* wird verstanden als praktikierbare Intertextualität: Auf Abruf bereitgehaltene Texte und Verknüpfungsstrukturen können jederzeit realisiert und jederzeit neu verknüpft werden (vgl. Sager 1997). Abzugrenzen von Intertextualität ist der Begriff der *Interdiskursivität* (vgl. Link 1986), der das Zusammenwirken institutionell geregelter, an Macht gebundener Redeweisen einer Gesellschaft meint.

zu Ziff. 4.613:

Intertextualität

zu: vprag7.51102

aus: HERWIG, H, Intertextualität als Mittel der Assimilations- und Orthodoxiekritik in Joseph Roths *Hiob. Roman eines einfachen Mannes*: Zeitschrift für Semiotik, 24/2-3 (2002) 261-277.

(262f) Eine solche Erwartungshaltung kann allenfalls vom Untertitel (263) *Roman eines einfachen Mannes* abgeleitet werden, nicht aber vom Haupttitel *Hiob*. Mit dem Haupttitel signalisiert Joseph Roth, dass sein Roman eines einfachen Mannes noch eine andere Textstrategie verfolgt: die Aktualisierung und sozialpsychologische Konkretisierung eines biblischen Stoffes. Und dieser Stoff verlangt am Schluss das im realistischen Kontext unwahrscheinliche "Wunder" (Roth 1982: 208). Bei genauer Betrachtung zeigt sich, dass der Text sogar drei kanonische Erzählungen des "Alten Testaments" übereinanderlegt:

- die Hiobsgeschichte: Hiobs Hadern mit Gott, weil er sich schuldlos gestraft fühlt und ihm vor dem Hintergrund der dogmatischen Vergeltungslehre seiner Zeit der Sinn seines Leidens uneinsehbar bleibt (*Hiob* 1-42),
- die Josephslegende, die vom Schicksal eines verlassenen Sohnes und Bruders berichtet, der es im Exilland zu Macht und Einfluss bringt und so zum Retter der Familie werden kann, die ihn in der Jugend ertränken wollte und verstoßen hat (*1 Mose* 37-50, Schweizer 1993, Shaked 1986: 289f),
- und die Geschichte der alttestamentlichen Richterin, Prophetin und siegreichen Heerführerin Debora (*Richter* 4f)...

(265) Schon auf der Ebene der textuellen Referenzillusion führt Joseph Roth in *Hiob* einen ideologiekritischen Zweifrontenkrieg:

- gegen die falsche Hoffnung, die sogenannte "Judenfrage" durch sexuelle oder nationale Assimilation an die Normen des östlichen oder westlichen Wirtschaftslandes lösen zu können,
- wie gegen den orthodoxen Versuch, den jüdischen Glauben durch Buchstabengehorsam gegenüber dem Gesetz vor jedem fremden Einfluss zu schützen.

Der Preis für die vollständige Assimilation ist der Verlust der ethnischen Identität (Hamerow 1984); der Preis für die Gesetzestreue das verpasste Leben.

zu Ziff. 4.614: vsyst4.21

Intertextualität

Themenheft: Zeitschrift für Semiotik, 24/2-3 (2002)

Literaturwissenschaftliche Intertextualitätsforschung
im Spannungsfeld konkurrierender
Intertextualitätsbegriffe (Henriette Herwig)

Intertextualität als Realisation von Weltliteratur:
Literarische Landschaften in Goethes *Faust*
(Anne Bohnenkamp)

Markierte Zitate und Kultur als Intertext: Varianten
der Intertextualität in Thomas Manns Roman
Königliche Hoheit (Franziska Schößler)

Intertextualität und Rezeptionsgeschichte bei Alfred
Döblin: "Goethe dämmerte mir sehr spät"
(Werner Stauffacher)

Intertextualität und Lektüresemiotik: Der *Räuber*-Roman
von Robert Walser (J. Ulrich Binggeli)

Intertextualität als Mittel der Assimilations- und
Orthodoxiekritik in Joseph Roths *Hiob. Roman
eines einfachen Mannes* (Henriette Herwig)

Schiffbruch des Erzählens: Eine intertextuelle Lektüre
der *Rotkäppchen*-Version in Ilse Aichingers
Roman *Die größere Hoffnung* (Brigitta Oesch)

Vom 'Werk' zur Intertextualität der *Stoffe*: Friedrich
Dürrenmatts Wandlung (Peter Rusterholz)

Vgl. am Institut für Deutsche Sprache, Mannheim, das "Projekt
Methoden der Korpusanalyse und -erschließung"

<http://www.ids-mannheim.de/kt/projekte/methoden/ka.html>

Stichwort: "Kookkurrenzanalyse" zur Aufdeckung von signifikanten
Regelmäßigkeiten bei der Verwendung von Wortkombinationen in den
Korpora.

zu Ziff. 4.615:

Lexem/Metapher: Ausdruck + Bedeutung

Unzulänglichkeit von Korpusanalysen

zu: vprag3.272

aus: K. BEIßNER, *I see what you mean* - Metaphorische Konzepte in
der (fremdsprachlichen) Bedeutungskonstruktion. EHS 14, 391.
Frankfurt/M 2002: P. LANG

(138)

The retrieval systems[with computer-held banks of text],
unlike human beings, miss nothing if properly instructed - no
usage can be overlooked because it is too ordinary or too
familial (...)

The human being, contrary to popular belief, is not well
organized for isolating consciously what is central and
typical in the language; anything unusual is sharply
perceived, but the humdrum everyday events are appreciated
subliminally. (Sinclair & Renouf, 1988, S.151; vgl. auch
Sinclair, 1991)

So konnte mittlerweile ... durch umfangreiche korpusbasierte
Analysen gezeigt werden, daß eine Vielzahl frequenter Lexeme
häufiger in metaphorisch erweiterten Bedeutungen als in ihrer
wörtlichen Bedeutung verwendet werden, diese Tatsache jedoch
bisher in Lehrbüchern für den Fremdsprachenunterricht nicht
berücksichtigt wird; "metaphorical uses of many common words are
more frequent in reality than the literal senses which dominate
textbooks" (Aston, 1995, S.258; vgl. auch Urdang, 1979). Gerade mit
Blick auf die Vermittlung von *phrasal verbs* und *idioms* im
Fremdsprachenunterricht verweist Sinclair (1991) auf die
Wichtigkeit von Korpusanalysen als Grundlage für die Entwicklung
entsprechender Lehrmaterialien, die die tatsächliche Verwendung
dieser Lexemkombinationen reflektieren.

Deignan (1999) macht jedoch auch auf die Grenzen der Möglichkeiten
aufmerksam, die korpusbasierte Analysen bieten, da sie sich
hauptsächlich zur Analyse frequenter und typischer Sprachmuster
und nur bedingt zur Untersuchung innovativer Sprachverwendungen
eignen; "corpus studies help to provide ways of determining what
is usual, not what is inventive" (S.196). Eine weitere
Unzulänglichkeit von Korpusanalysen besteht darin, daß auf Grund
der Notwendigkeit *bottom-up* zu arbeiten, indem bestimmte Lexeme
herausgesucht und deren kollokationales Profil sowie deren
mögliche metaphorisch motivierte Bedeutungen anhand von
Konkordanzen untersucht werden, metaphorische
Bedeutungserweiterungen anderer Lexeme des gleichen semantischen
Feldes leicht übersehen werden können. Die Autorin empfiehlt daher
eine systematische Analyse semantischer Felder unter Zuhilfenahme
von Thesauri.

zu Ziff. 4.62:

Phrasenkomposita

zu:

aus: J. MEIBAUER, Phrasenkomposita zwischen Wortsyntax und Lexikon: Zeitschrift für Sprachwissenschaft 22 (2/2003) 153-188.

(170) Eine weitere Auffassung liegt vor, wenn man unter Lexikalisierung die permanente Speicherung von fixierten Ausdrücken ('fixed expressions') im mentalen Lexikon versteht, wie Jackendoff (1995, 1997) es tut. Es muss dabei keine Abweichung in phonologischer oder semantischer Hinsicht vorliegen, es genügt die Annahme, dass der entsprechende Ausdruck als Ganzer gespeichert und abrufbar ist. In (171) Anlehnung an Jackendoff (1995, 1997) unterscheide ich zwischen Idiomen, Klischees, Titeln, Zitaten, Paarformeln, entlehnten Phrasen und Mengenangaben. In der folgenden Tabelle finden sich dazu passende Belege mit entsprechenden Erstgliedern von Phrasenkomposita:

(26) Typen und Belege lexikalisierten Erstglieder in Phrasenkomposita

1. Idiome	Die-Katze-im-Sack-kaufen-Effekt, Leck-mich-am-Arsch-Bullen, Schwarze-Kassen-Affäre, Affe-auf-Schleifstein-Position, Rote-Socken-Kampagne, Kopf-durch-die-Wand-Strategie, Oben-ohne-Show, Gott-sei-bei-uns-Vokabel, Ton-in-Ton-Kombination
2. Klischees	Eigentlich-geht's-uns-doch-ganz-gut-Menschen, Es-war-einmal-Atmosphären, Hier-kriegt-man-alles-was-man-braucht-Seminar, Genau-so-ist-es-Effekt, Abgerechnet-wird-am-Schluss-Taktik, Lass-uns-Freunde-bleiben-Spruch, Ich-trau-mich-nicht-Hommage
3. Titel	Romeo-und-Julia-Gefühl, Denn-sie-wissen-nicht-was-sie-tun-Jugenddrama, Mensch-ärger-dich-nicht-Spielfeld, Wer-wird-Millionär-Deutschland
4. Zitate	"Keine-Macht-den-Drogen"-Schmarrn, "Ich-bin-ein-Berliner"-Auftritt, "Buddeln-ist-rum"-Fest, Boris-"Ich-bin-drin"-Becker, Dieter-"Ich-hab-der-Tusse-keine-gefeuert"-Bohlen
5. Paarformeln	Messer-und-Gabel-Handhabung, Zwei-plus-Vier-Gespräche, Pro-und-contra-Debatte, Pfeffer-und-Salz-Haare, Freund-oder-Feind-Philosophie
6. Entlehnte Phrasen	Grand-Slam-Turnier, Fast-Food-Kino, Queen-of-England-Qualitätssiegel
7. Mengenangaben	630-Mark-Gesetz, Erste-Person-Perspektive, 180-Grad-Wende

Der wesentliche Unterschied zwischen Idiomen und Klischees besteht darin, dass nur die Bedeutung der Idiome nicht-kompositional ist.

4.62 Feste Wortketten

Literatur

- LUDEWIG, P: Korpusbasiertes Kollokationslernen. Computer-Assisted Language Learning als prototypisches Anwendungsszenario der Computerlinguistik. Sprache, Sprechen und Computer 9. Frankfurt am Main 2005 3-631-53958-4br: Peter Lang // I.2.5
- MIEDER, W: "Yes We Can". Barack Obama's Proverbial Rhetoric. Bern 2009 978-1-4331-0667-5ph: Peter Lang // I.2.7

zu Ziff. 4.621:

Feste Redewendungen

zu:

aus: BLUHM, L; RÖLLECKE, H, "Redensarten des Volks, auf die ich immer horche". Märchen - Sprichwort - Redensart. Zur volkspoetischen Ausgestaltung der *Kinder- und Hausmärchen* der Brüder Grimm. Stuttgart/Leipzig 1997: S. Hirzel

(19) Bei der folgenden Dokumentation ist der Begriff im Sinn Wilhelm Grimms weit gefaßt. Es sind einbezogen das formelhafte Sprichwort ("Wer A sagt, muß auch B sagen"), die gangbare Redensart ("jemanden abhobeln"), der redensartige Vergleich ("ehrlieh wie der Dieb am Galgen"), Lebensweisheiten ("man soll Bäume ziehen, solange sie jung sind"), volkssprachlich gewordene (Bibel-)Zitate ("unschuldiges Blut vergießen"), sprachlicher Niederschlag des Volksglaubens ("der wilde Jäger"), tradierte (meist alliterierende) Sprachformeln ("Geld und Gut"), volkstümlich typisierende Namengebung ("Packan"), volkssprachliche, bildkräftige Verunlimpfungen ("Bärenhäuter"), märchengerechte (nicht die stereotypen) Schlußformeln ("Da kam eine Maus, das Märchen war aus"). Nur ein derart weiträumiger Angang gewährleistet die umfassende und angemessene Darstellung des Grimmschen Konzepts volkspoetischer Ausgestaltung in den KHM...

(42) *KHM 4 Märchen von einem, der auszog, das Fürchten zu lernen* ...

Der Vater bemerkt zu seinem tumben Sohn:

1. *an dir ist Hopfen und Malz verloren* (⁷1857, I, 42; zuerst ²1819). Die Wendung ist vielfältig belegt; so gebraucht sie etwa Jacob Daniel Ernst 1682 in seinen 'Gemueths-ergetzlichkeiten': "es ist doch hopf und malz an ihm verloren". Als Lebensregel vermittelt der älteste Bruder dem einfältigen Märchenhelden:

2. *was ein Häkchen werden will, muß sich beizeiten krümmen* (⁷1857, I, 42; zuerst ²1819). Bereits seit dem Mittelalter ist diese sicherlich ein wenig zweifelhafte Lebensregel in der Literatur als gängige redensartige Weisheit belegt (vgl. auch um 1530 Luthers Sprichwörterammlung, Nr. 402). Noch 1860 verzeichnet sie L. Curtze als Sprichwort...

6. *So haben wir nicht gewettet* (⁷1857, I, 48; zuerst ²1819). In KHM 10 findet sich diese Redensart schon seit ¹1812. Sie ist in Literatur und Volksliteratur vielfältig belegt; etwa wiederum in Hebels 'Schatzkästlein des Rheinischen Hausfreundes': "Als das die anderen hörten, machten sie zwar große Augen und meinten, so sei's nicht gewettet".

Bei der Begegnung mit einer Leiche erinnert sich der Märchenheld:

7. *wenn zwei zusammen im Bett liegen, so wärmen sie sich* (⁷1857, I, 49; zuerst ²1819) - eine Weisheit, die bereits im Alten Testament zu finden ist: "wenn zwei beieinander liegen, wärmen sie sich" (Pred 4, 11). ...

(52f) KHM 15 Hänsel und Gretel ...

2. *wie sollt' ich's übers Herz bringen* (⁷1857, I, 100; zuerst ¹1812). Die Redensart wiederholt sich in den KHM noch mehrmals: in KHM 60 (seit ²1819), in KHM 97 (seit ³1837) und in KHM 135 (seit ¹1815). Wie viele der Herzmotaphern ist sie als gängige Sprachformel weit verbreitet und häufig belegt. Luther etwa verwendet die Junktur: "hat gott seinen einigen son für uns alle

dahin gegeben, wie künde ers denn über das Herz bringen, uns in geringerm Anliegen zu verlassen?" Als Gretel vom Plan der Eltern erfährt, sie gemeinsam mit Hänsel im Wald auszusetzen, fürchtet sie:

3. *nun ist's um uns geschehen* (⁷1857, I, 100; zuerst ¹1812). Die Wendung begegnet vergleichbar später in einigen anderen Märchen: in KHM 40 (seit ²1819), in KHM 56 (seit ³1837) und in KHM 93 (seit ⁶1850), hier allerdings mit der Erweiterung "so ist es leicht um dein Leben geschehen". Die Redensart ist in der Literatur häufig zu finden; sehr prägnant verwendet sie etwa Gryphius: "geschehn! es ist geschehn! mein König nicht um dich: nein! nein! ach leider nein! es ist geschehn um mich!" Für die Verbreitung sprechen auch Hebel und Goethe, der eine im 'Rheinischen Hausfreund': "Da war es um den armen Jüngling geschehn", der andere in seiner Ballade 'Der Fischer': "Da war's um ihn geschehen".

zu Ziff. 4.622:

Formelhaftes Sprechen

zu: vprag6.0831

aus: SCHWITALLA, J, Gesprochenes Deutsch. Eine Einführung. Berlin 1997: E. Schmidt

(118f) Kritiker der gesprochenen Sprache halten diese für stereotyp, formelhaft und vorgestanzt. Gewiß greifen wir zur Durchführung bestimmter Sprechakte auf gesellschaftlich vorgefertigte Formeln zurück (Begrüßungen, Verabschiedungen, Sich-Entschuldigen usw.). Aber Gesprochenes ist insgesamt nicht formelhafter als manche geschriebenen Textsorten auch (Politiker-, Festreden, offizielle Mitteilungen, Urteilsbegründungen, Zeugnisse, Anzeigentexte).

Coulmas schlägt folgende Kategorien vor:

1. Redewendungen (syntaktisch unvollständig, übertragene Bedeutung): *jemanden auf=s glatteis führen; einen strich drunter machen;*
2. Sprichwörter (satzförmig, übertragene Bedeutung, Kondensierung kulturellen Wissens): *wes brot ich eß, des lied ich sing; gleich und gleich gesellt sich gern;*
3. Gemeinplätze (satzförmig, Truismen, Tautologien): *wir sind alle nur menschen; was zuviel ist, ist zuviel;*
4. Routineformeln für bestimmte Sprechakte und dialogische Aktivitäten: *guten tag; darf ich vorstellen? wie bitte? hörst du mir überhaupt zu?*

Aus ethnographischen Forschungen zur Kommunikation geht hervor, daß, je traditionaler eine Kultur ist, um so häufiger auf einen Schatz an Sprichwörtern zurückgegriffen wird, mit denen Einzelfälle des täglichen Lebens unter allgemeine Gesetze subsumiert werden. Sprichwörter haben in weiten Teilen unserer Gesellschaft ausgedient, allseits akzeptierte Wahrheiten und Normen festzuklopfen. Viel häufiger werden sie zu witzigen Verdrehungen verwendet: *kommt zeit, kommt unrat; wer die moral hat, hat die qual.* Das gilt auch für geflügelte Worte (*wer nie sein brot im bette aß, weiß nicht, wie krümel pieksen*) und Sprüche. In bäuerlichen und Arbeitermilieus werden Sprichwörter, Maximen und Gemeinplätze dagegen durchaus zu konsensstiftenden Zwecken formuliert.

Routineformeln haben sehr viele Funktionen für die Herstellung, Durchführung und Beendigung von Gesprächen. Ich zähle einige auf:

- Begrüßungs- und Verabschiedungsformeln
- phatische Kommunikation (*schönes wetter heute*)
- Redeeinleitungsformeln (*wenn ich dazu etwas sagen darf; ich habe eine frage,*) und bekräftigende Beendigungsformeln (*ich bleibe dabei; tut mir schrecklich leid; so ist das*)
- metakommunikative Floskeln zur eigenen Rede (*meiner meinung nach; soweit ich weiß; ob du=s glaubst oder nicht;*)
- Formeln für bestimmte Sprechakte: Sich-Entschuldigen (*entschuldigung; tut mir leid*) und Ratifizieren (*macht nix*); Danken, Bitten, Loben, Drohen usw.

Vertreter von Institutionen bedienen sich formelhafter amtssprachlicher Ausdrücke, um ihre offizielle Rolle, z.T. auch ihre institutionelle Macht zu dokumentieren (*hiermit eröffne ich*

die sitzung;) Die Sprecher greifen dabei nur an einzelnen Stellen, nicht für längere Redeabschnitte, auf vorgefertigte Formulierungen zurück. Und selbst wenn wie in Gottesdiensten der Wortlaut vorgeschrieben ist, hat der Durchführende einen gewissen Spielraum der individuellen Gestaltung.

zu Ziff. 4.623:

Kollokationen | Phraseologie

zu:

aus: P. LUDEWIG, Korpusbasiertes Kollokationslernen. Computer-Assisted Language Learning als prototypisches Anwendungsszenario der Computerlinguistik. Sprache, Sprechen und Computer 9. Frankfurt/M 2005.

(11) Gegenstand der vorliegenden Schrift sind Kollokationen. Dabei handelt es sich um usuelle bzw. habitualisierte Wortverbindungen wie z.B. *Geld abheben, Zorn verraucht, schwarz sehen, tief beeindruckt* und *eingefleischter Junggeselle*. Heid charakterisiert kollokatorische Wortverbindungen als

[...] frequent, recurrent, conventionalized building blocks of the lexicon [...] often not predictable; but native speakers are quite good at identifying non-collocational combinations in other people texts, and they feel non-collocational texts are not fluent, not elegant or just not the "usual way" how one would express a given idea.

[...] texts with wrong collocates or with compositional expressions where collocational alternatives would exist, are perceived as unnatural; (Heid 1994, S. 228-229, 240).

Entsprechend verwundert es nicht weiter, dass Kollokationen beim Erlernen einer Fremdsprache eine äußerst wichtige Rolle spielen, zumal Kollokationen typischerweise von Sprache zu Sprache variieren, wie dies z.B. bei *einen Vortrag halten/to give a talk/faire une conférence* der Fall ist. ...

(67) Words are like people [...]

We all feel comfortable when we are surrounded by friends and acquaintances, but anxious in unfamiliar situations when we are surrounded by strangers. We have friendships of different kinds - close, intense relationships which, even if relatively infrequent, are the most important in our lives - loved ones who live abroad, for example. But we have relationships which are frequent but unimportant - the person who travels on the same train to and from work, five days a week. There is also the one-night-stand - a serendipitous one-off, creative encounter, but however sadly, not part of everyday life. The relationships between words closely resemble the relationships between people. (Hill et al. 2000, S. 88).

(68) Neben Sprichwörtern (*Wer nicht wagt, der nicht gewinnt*), Floskeln (*soviel, ich weiß,...*), Idiomen (*an den Nagel hängen*) und Funktionsverbgefügen (*zum Ausdruck kommen*) zählen insbesondere Kollokationen (*den Tisch decken*) zu diesem gerne unter dem Begriff Phraseologie zusammengefassten Phänomenbereich (s.a. Burger 1998).

Der theoretische Werdegang von Kollokationen ist dadurch gekennzeichnet, dass Kollokationen in der generativ geprägten Linguistik und infolge auch in der Computerlinguistik lange Zeit sehr stiefmütterlich behandelt wurden. Im Zentrum der Betrachtungen standen die produktiven Aspekte der Sprache.

zu Ziff. 4.6231:

Kollokationen - Statistik + Semantik?

zu:

aus: P. LUDEWIG, Korpusbasiertes Kollokationslernen. Computer-Assisted Language Learning als prototypisches Anwendungsszenario der Computerlinguistik. Sprache, Sprechen und Computer 9. Frankfurt/M 2005.

(111) ...wird bei dem vorgestellten Verfahren davon ausgegangen, dass die Abstände der Wörter, die eine diskontinuierliche Kollokation konstituieren, relativ homogen, sprich gleichbleibend sind. Entsprechend berechnet man bei diesem Verfahren den Abstandsmittelwert sowie die Varianz des Abstands, welche beschreibt, wie sehr die konkreten Abstände vom Mittelwert abweichen. Wenn der Abstand immer identisch ist, beträgt die Varianz 0. Eine geringe Varianz besagt, dass die beteiligten Wörter in der Regel in etwa dem gleichen Abstand zueinander vorkommen. Ist der Abstand ein Zufallsprodukt, nimmt die Varianz hohe Werte an. In der Regel wird an Stelle der Varianz die Standardabweichung σ als Maß für die Variabilität der Abstände genutzt. Der Mittelwert für die unter (3.46) aufgelisteten Sätze beträgt z.B. $m = 1/5(1 + 4 + 4 - 1 - 6) = 2$, die Varianz σ^2 hat den Wert

$$\sigma^2 = \frac{(1-2)^2 + (4-2)^2 + (4-2)^2 + (-1-2)^2 + (-6-2)^2}{4} \sim 34,7$$

, so dass σ hier etwa den Wert 5,89 annimmt. Dieser Wert ist relativ groß, so dass *Kontakte knüpfen* als Kollokationskandidat abgelehnt würde.

- (3.46) 1. Über Kim Philby wollte mein Vater *Kontakte knüpfen*. (Spiegel 1996) ($d_1 = 1$)
2. V-Mann Rehkopf sollte *Kontakt* mit NF-Chef Schönborn *knüpfen*. (Spiegel 1996) ($d_2 = 4$)
3. In dieser Zeit hat er diverse *Kontakte* zu einflussreichen Persönlichkeiten *geknüpft*. (konstruiert) ($d_3 = 4$)
4. Der PDS-Mann *knüpfte Kontakt* zu der Dissidentengruppe Bewegung für die Wahrheit und kam zu dem Ergebnis (Spiegel 1996) ($d_4 = -1$)
5. Er wirkte an der Rede des PLO-Chefs vor der Uno-Vollversammlung mit, *knüpfte* für ihn 1979 die ersten *Kontakte* mit der US-Regierung. (Spiegel 1996) ($d_5 = -6$)

Offensichtlich ist diese Vorgehensweise eher für positionell erstarrte Kollokationen geeignet, die kaum syntaktische Transformationen erlauben.

(134) *Wasser trinken* dürfte ein Beispiel für eine Verbindung darstellen, die lediglich aus syntaktisch-statistischer Sicht als Kollokation einzustufen wäre, wobei im Grunde zu hinterfragen wäre, ob dies wirklich gewünscht wird oder nicht. Semantisch verhält sich diese Konstruktion unauffällig, sprich klassisch kompositionell. In dieser Verbindung lässt sich das Nomen *Wasser* durch andere trinkbare Flüssigkeiten wie *Milch* ersetzen. Entsprechendes gilt für das Verb *trinken*, das z.B. durch *schlürfen* oder *zu sich nehmen* substituiert werden darf. Auch die Übersetzung ins Englische und ins Französische erfolgt gewissermaßen wortweise (*to drink water / boire de l'eau*).

Bei *Angst bekämpfen* verhält es sich gemäß der auf Seite 115 beschriebenen Studie eher entgegengesetzt. Diese Konstruktion ist in erster Linie aus semantisch-kontrastiver Sicht als Kollokation einzustufen. In besagtem Korpus war diese Wendung deutlich unterrepräsentiert und damit statistisch gesehen nicht als Kollokation zu werten. Die Verbindung verfügt aber dahingehend über eine Basis-Kollokator-Struktur, dass *bekämpfen* im Kontext von *Angst* eine Bedeutung erhält, die sich mit *etwas einzudämmen versuchen* oder *gegen etwas angehen* paraphrasieren lässt. Ein Ersetzen des Verbs durch ein anderes sinnähnliches Verb ist schwierig; eine Verbindung mit dem Verb *angreifen* ist unüblich.

(175) Der Eintrag von *Macht* enthält 10 V-N-Kollokationen, der von *Hand* sogar 24 V-N-Kombinationen, davon zwei mit *geben*, drei mit *nehmen* und sogar 6 mit *haben*.

(5.3) jdm die Hand geben,
 etw. (nicht) aus der Hand geben,
 jdn bei der Hand nehmen,
 etwas zur Hand nehmen,
 etw. in die Hand nehmen,
 jdn in der Hand haben,
 alle/beide Hände voll zu tun haben,
 jdn an der Hand haben,
 zwei linke Hände haben,
 eine ruhige/sichere Hand (bei etw.) haben,
 etwas hat Hand und Fuß,
 jdm die Hand schütteln,
 etw. liegt klar auf der Hand,
 sich etw. an beiden Händen abzählen können,
 jmd zur Hand gehen,
 etw. fällt jdm in die Hände,
 jd fällt jdm in die Hände,
 in festen Händen sein,
 jdm rutscht die Hand aus,
 die Hände über dem Kopf zusammenschlagen,
 für jdm/etw. die/seine Hand ins Feuer legen,
 sich mit Händen und Füßen gegen jdn/etw. wehren,
 seine Hände in Unschuld waschen,
 um jds Hand anhalten

(5.4) Macht ausüben,
 jdn in seiner Macht haben,
 seine Macht ausspielen,
 seine Macht ausbauen,
 alles tun, was in seiner Macht steht/liegt,
 an der Macht sein,
 an die Macht kommen,
 die Macht an sich reißen,
 die Macht ergreifen,
 jdn an die Macht bringen,

(176) Zusammenfassend bleibt für Kollokationen als fremdsprachlichen Lerngegenstand festzuhalten:

- Es genügt nicht, nur einzelne Wörter zu lernen, es müssen auch ganze Wortverbindungen memorisiert und in das lexikalische Repertoire aufgenommen werden.
- In vielen Fällen reicht die Kenntnis der Grundform einer Kollokation für eine adäquate Verwendung nicht aus. Zusätzlich müssen auch Detailkenntnisse über das grammatische Modifikationspotential erworben werden. (177)
- Die Anzahl der zu memorisierenden Kollokationen ist enorm.

- Deshalb reicht ab einem bestimmten Lernniveau die Orientierung an einem festen Kanon zu erlernender Kollokationen nicht aus. Ergänzend ist eine lebenslanges autonomes und bedürfnisorientiertes Lernen von Kollokationen unverzichtbar.
- Dazu muss der Lerner ein Bewusstsein für die sprachliche Bedeutsamkeit von Kollokationen im Allgemeinen und konkrete kollokatorische Phänomene im Besonderen entwickeln.

zu Ziff. 4.624:

Krieg der Sprichwörter und Zitate

zu:

aus: B. Brecht, Der kaukasische Kreidekreis. Frankfurt/M 1955.

(133)

AZDACK: Halt's Maul. Hab ich etwa von dir genommen?

GRUSCHE *obwohl die Köchin sie zurückhalten will*: Weil ich nichts hab.

AZDACK: Ganz richtig. Von euch Hungerleidern krieg ich nichts, da könnt ich verhungern. Ihr wollt eine Gerechtigkeit, aber wollt ihr zahlen? Wenn ihr zum Fleischer geht, wißt ihr, daß ihr zahlen müßt, aber zum Richter geht ihr wie zum Leichenschmaus.

SIMON *laut*: "Als sie das Roß beschlagen kamen, streckte der Roßkäfer die Beine hin", heißt es.

AZDACK *nimmt die Herausforderung eifrig auf*: "Besser ein Schatz aus der Jauchegrube als ein Stein aus dem Bergquell."

SIMON: "Ein schöner Tag, wollen wir nicht fischen gehn? sagte der Angler zum Wurm."

AZDACK: "Ich bin mein eigener Herr, sagte der Knecht und schnitt sich den Fuß ab."

SIMON: "Ich liebe euch wie ein Vater, sagte der Zar zu den Bauern und ließ dem Zarewitsch den Kopf abhaun."

AZDACK: "Der ärgste Feind des Narren ist er selber."

SIMON: Aber "der Furz hat keine Nase".

AZDACK: 10 Piaster Strafe für unanständige Sprache vor Gericht, damit du lernst, was Justiz ist.

zu Ziff. 4.63:

Lektüre: Ganzworterkennung / Fehlerkorrektur

zu: vprag8.202

aus: Werbeanzeige

WSEIO KNÖNEN SIE DEIESN STAZ LSEEN,

OWBHOBL DIE BCUTHSAEBN NCHIT IN DER

RITHCIEGN RIEHNEFOGLE SHETEN?

(Ausdrucks-)SYNTAX

- Erst- und Letztbuchstaben stimmen alle
- Bei "STAZ" sind Variationsmöglichkeiten nicht groß. Hat man "Satz" → "Lesen" liegt von der Wahrscheinlichkeit her nahe; ebenso "diesen".
- Wenn "Satz + lesen", dann hohe Wahrscheinlichkeit für "Buchstaben" in der Nähe.
- Laut Satzzeichen muss Frage vorliegen → erstes Wort W-Fragewort

SEMANTIK/PRAGMATIK

- Wenn Frage: Gefragter entweder mit "Du" oder "Sie" angesprochen / Erwartung
- Erwartung hoch: Frage kombiniert mit "obwohl" [=entgegen dem Wissen, "konzessiv"]. "Frage" heißt: Es geht um Wissen/Nicht-Wissen
- Wenn "Satz + lesen" - Entzifferung wird unterstützt durch die aktuelle Praxis: man schlägt sich auffallend intensiv mit "Buchstaben" herum.
- Pragmatisch ist das aktuell erlebte Problem die "richtige Reihenfolge + stehen"

Erwartbar: Ganzworterkennung erheblich erschwert, wenn

- noch stärkere Buchstaben-Permutation vorliegt
- die schriftliche Aussage von der aktuellen Lesepraxis entkoppelt wäre.

4.63 Wortnachbarschaften und ihr Ertrag

zu Ziff. 4.631:

(Ausdrucks-)SYNTAX: Muster, Nachbarschaften

zu: vsem1.551

aus: RIEGER, B, Wissenserwerb aus natürlichsprachlichen Texten. Zur Interdisziplinarität eines computerlinguistischen Forschungsprogramms, in: KÜHLWEIN, W, RAASCH, A (eds.), Angewandte Linguistik heute. forum Angewandte Linguistik Bd. 20. 1990. S.101-124.

(115f) Anhand verschiedener Corpora natürlichsprachlicher Texte und der daraus berechneten semantischen Hyperräume $\langle S, \mathcal{O}_2 \rangle$ wurde überprüft, ob das entwickelte und angewandte Analyse- und Repräsentationsverfahren tatsächlich die über Wortverwendungen in Texten konstituierten semantischen Zusammenhänge zu ermitteln vermag. Hierzu wurden zum einen die *semantischen Umgebungen* $E(z_i, r)$ derjenigen Bedeutungspunkte z berechnet und aufgelistet, die sich in den topologischen Nachbarschaften von sowohl einzelnen Bedeutungspunkten z_i als auch von durch konjunktive und adjunktive UND/ODER-Verknüpfungen entstandenen Bedeutungspunkten $z_i \wedge_j$ bzw. $z_i \vee_j$ finden. Zum anderen konnte anhand diverser, die Punktverteilungen in $\langle S, \mathcal{O}_2 \rangle$ untersuchender Clusteranalysen, welche auf der Grundlage einzig *numerisch-statistischer* Kriterien der Positionen von Punkten im Hyperraum Ähnlichkeitsklassen von Bedeutungspunkten ermitteln lassen, der Beleg erbracht werden, daß sich eben jene Bedeutungspunkte in *Cluster* versammeln, die auch intuitiv *semantisch* ähnliche Bedeutungen repräsentieren. Es ließ sich derart nachweisen, daß das vorgelegte, einzig auf der Analyse von Wortverwendungsweisen in natürlichsprachlichen Texten basierende automatische Verfahren zur Repräsentation von Wortbedeutungen und ihrer Beziehungen in *pragmatisch homogenen* Textcorpora zum Aufweis semantischer Zusammenhänge führt, die ihrerseits als Basis des in und durch Texte vermittelten (lexikalisierten) Wissens gelten können.

WIRTSCHAFT	0.000				
AUSLAND	3.785	BRITAIN	5.094	ENTWICKL	5.893
FOLGE	6.112	VERWALT	6.428	RAUM	6.903
EINSATZ	9.307	KONTAKT	9.934	HERRSCHEN	10.163
GESCHÄFT	10.931	KRANK	11.732	VERKEHR	11.984
VERANTWORTL	12.298	SPRACH	12.429	MÖGLICH	13.257
WEG	13.285	NEU	13.871	ZENTRAL	14.831
LEHR	15.131	JUNG	15.550	ALLGEMEIN	15.796
MODE	15.850	AUFTRAG	15.952	MASCHINE	16.210

Tabelle 1: Topologische Umgebung $E(z_i, r)$ von $i =$ WIRTSCHAFT mit Bedeutungspunkten aus der Hyperkugel mit Radius r im *semantischen Hyperraum* $\langle S, \mathcal{O}_2 \rangle$ berechnet anhand eines Textcorpus der Tageszeitung DIE WELT der 1964er Ausgaben (Stichprobe: 175 Artikel mit rund 7000 Worttoken und 365 Worttypen)

Die derart repräsentierten *lexikalischen* Bedeutungszusammenhänge bilden eine relationale Datenstruktur, deren sprachlich etikettierte Elemente (Bedeutungspunkte) und deren wechselseitige Abstände (Bedeutungsunterschiede) ein System von einander überdeckender *Stereotype* bilden. Die *Bedeutung* eines Elements kann daher sowohl als unscharfe Teilmenge des Vokabulars, als Vektor eines Bedeutungspunkts als auch durch die topologische Umgebung eines Bedeutungspunkts dargestellt werden. Letztere besteht aus allen Namen und Werten der nach zunehmenden Abständen

aufgelisteten Bedeutungspunkte, die sich innerhalb einer Hyperkugel des Radius r um einen Bedeutungspunkt $z_i \in \langle S, \delta_2 \rangle$ finden. Sie repräsentieren die Bedeutung des betreffenden Wortes daher indirekt als *Stereotyp*, d.h. als Zustand, Muster oder Verteilung von Bedeutungspunkt-Wert-Paaren. ...

(121) Anders als in den (propositionalen) Formaten zur (prädikativen) Bedeutungs- und Wissensrepräsentation der bisherigen KI-Forschung werden die skizzierten Prozeduren als *zeitkritische* Algorithmen unterschiedlicher (z.T. noch zu testender) Aufgabenstellungen und Operationscharakteristiken weiterzuentwickeln sein,

- welche semantische Beziehungen zwischen Konzepten nicht voraussetzen müssen, sondern diese induktiv aus den Strukturen der analysierten Corpora als Funktion des Gebrauchs von Wörtern in Texten zu berechnen gestatten;
- welche - durch die Trennung von Basisstruktur und den auf dieser Basis operierenden Prozeduren - es erlauben, semantische Beziehungen zwischen den *stereotypischen* Repräsentationen (Bedeutungspunkten im *semantischen Hyperraum*) von deren - je nach Aspekt, Perspektive, Kontext - variablen konzeptuellen Abhängigkeiten untereinander zu unterscheiden;
- welche - auf der Grundlage dieser konzeptuellen Hierarchien und der sie aktivierenden Prozeduren - *assoziativ-analoges* im Unterschied zu *deduktiv-logischem* Schließen als *semantische Inferenzen* modellieren;
- welche schließlich - durch ihre teils rekursiven, teils rückbezüglichen Strukturen - die Resultate solcher Verarbeitungsprozesse eben diesen Verarbeitungsprozessen zu unterwerfen vermögen.

zu Ziff. 4.632:

Ähnlichkeiten bei Wortketten

zu:

aus: B. GLOOR, *staat - sex - amen*. 81 Sprachbeobachtungen. Zürich 42004.

Nr. 31

Aus einer Radioansage: "Sie hören nun die h-Mess-Molle, Verzeihung, die h-Moss-Melle, ich bitte sehr um Entschuldigung, die h-Moll-Messe von Johann Sebaldrian Bach - ich häng mich auf!"

Auf tausend Wörter passiert im Durchschnitt ein Versprecher. Das ist nicht viel; etwa alle zehn Minuten. Leider. Denn wenn's im Wolken nordig wird (Hochzuckdrohne!) und die Temperaturen zwischen 16 und 22 Uhr liegen, wenn es blittert beim Gewitzer und es auf der Autobahn Frankfurt-Köln in beiden Fahrtrichtungen schneit - dann setzt sich die Sprache für einen Moment über ihre eigenen Gesetze hinweg und bringt skurril-poetische Bilder hervor, wie wir sie sonst nur aus dem Kindermund kennen: "Bitte eine Prachtel Schalinen und ne Cose Dola."

In Büchern kommen solche Wertverdrehungen noch seltener vor. Es soll ja Leute geben, die nichts Besseres zu tun wissen, als sämtliche ausgelückten Gesetzesnutzen sowie bankursen Spanken und Barkassen bis zum Tröpfelchen aufs i aus dem Verkehr zu ziehen, so dass die Texte am Ende geschnügelt und gebiegelt daherstehen.

Elende Wasserbisser! Doch zum Glück hat jemand alle rausgeschmissenen Versprecher und Verschreiber gesammelt, zwischen zwei Buchdeckel gepresst und so quasimassen mit gleicher Münze zurückgeschlagen. Die Autorin von "Reden ist Schweigen, Silber ist Gold" heisst Helen Leuninger und arbeitet als Professorin für Sprachwissenschaft an der Universität Frankfurt. Ihr Fachgebiet ist die kognitive Linguistik. Das macht aber nichts, denn selbst die "Kleine Theorie des Versprechers", die sie ihren sechshundert Lieblingsausrutschern hinterherschickt, ist erfrischend verständlich.

Für Leute wie mich ist so ein Buch natürlich das Ein und O! Denn Frau Leuninger nimmt kein Brett vor den Mund. Ihre Beispiele sind nicht aus den Haaren gesogen, sondern sie hat sich mehr als ein Ohr um die Nacht geschlagen, um ihren Lesern reinen Tisch einzuschenken: Wenn Ärzte Dünnfall, Tiefdruck, Artillerverkalkung oder Angora pectoris diagnostizieren, dann stehen uns die Berge schon mal so zu Haare, dass uns der Hut platzt. Und es besteht die Gefahr, dass dem Patienten bei der Untermeidung lebensnotwichtiger Massnahmen des Lebensknicht ausgelipst wird. Dann gibt es nur noch eine Diagnose: Tote sterben fest. Wenn dagegen am Fernsehen hintertückisch von Tot- und Mordschlag daherschwadroniert und den Zuschauern überhaupt Honig ins Gesicht geschmiert wird, dass es uns die Balken biegt, fallen wir zwar aus fast allen Socken, aber zugleich geht uns so mancher Groschen auf - und die Tatsachen bleiben im Dorf.

4.7 Phonetik/Prosodie (ausgeklammert)**Literatur**

- ANREITER, P: Akzelerierter Wortschatzzerfall durch Superstratdruck in: KLENK, U (ed./eds.): Computatio Linguae II. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache. ZDL BH.83. Stuttgart 1994 1-18. // I.2.7
- BEAUDOUIN, V; YVON, F: The Metrometer: a Tool for Analysing French Verse. Literary and Linguistic Computing 11 / 1 (1996) 23-31.
- BECKER, T: Zur Repräsentation der Vokallänge in der deutschen Hochsprache. Zeitschrift für Sprachwissenschaft 15,1 (1996) 3-21.
- DAVIS, J F: Phonetics and Phonology. Uni-Wissen Anglistik, Amerikanistik. Stuttgart 1998: Klett // I.2.7
- ELSEN, H: Rezension von: HIRSCHMANN, Irena: Vergleich der Silbenstruktur und der Akzentuierung der deutschen Standardsprache und der tschechischen Schriftsprache... Zeitschrift für Sprachwissenschaft 14,1 (1995) 101-108.
- EPPINGER, B; HERTER, E: Sprachverarbeitung. München / Wien 1993: Hanser // I.2.7
- HALLIDAY, M A K: An introduction to functional grammar. London 1994: Arnold // I.2.4
- HAMMOND, M: Is Phonology Irrelevant?. How and Why Frequency Can be Modelled in Phonology. Literary and Linguistic Computing 13 (4 / 1998) 165-175.
- HILL, E: Die Präferenztheorie in der historischen Phonologie aus junggrammatischer Perspektive. Zeitschrift für Sprachwissenschaft 28,2 (2009) 231-263.
- HUNYADI, L; KALMAN, A et. al.: Forensic Linguistics: its Contribution to Humanities Computing. Literary and Linguistic Computing 18 / 1 (2003) 49-62.
- KRÄMER, S: Sprache und Schrift oder: Ist Schrift verschriftete Sprache?. Zeitschrift für Sprachwissenschaft 15,1 (1996) 92-112.
- LESSARD, G; LEVISON, M: Computational Generation of Limericks. Literary and Linguistic Computing 20 (2005) 89-105.
- OUSAKA, Y, Z; YAMAZAKI: Automatic Analyses of the Canon in Middle Indo-Aryan by Personal Computer II. Literary and Linguistic Computing 11 / 1 (1996) 9-17.
- RAMERS, K-H; WIESE, R . Zeitschrift für Sprachwissenschaft 10,1 (ed./eds.): Prosodische Phonologie (1991)
- RITTAUD-HUTINET, C: La phonopragmatique. Sciences pour la communication 45. Bern 1995: Peter Lang // I.2.7
- ROBERTS, A: Rhythm in Prose and the Serial Correlation of Sentence Lengths: a Joyce Cary Case Study. Literary and Linguistic Computing 11 / 1 (1996) 33-39.
- SAGERER, G: Automatisches Verstehen gesprochener Sprache. Informatik 74. Mannheim // I.2.8
- SELTING, M: Phonologie der Intonation. Probleme bisheriger Modelle und Konsequenzen einer neuen

interpretativ-phonologischen Analyse. Zeitschrift für Sprachwissenschaft 11,1 (1992) 99-138.

- VOLGGER, D: Notizen zur Phonologie des Bibelhebräischen. ATS 36. St. Ottilien 1997 // I.2.7
- VÖLTZ, M: Das Rhythmusphänomen. Einige Anmerkungen zur Problematik der typologischen Klassifizierung. Zeitschrift für Sprachwissenschaft 10,2 (1991) 284-296.

zu Ziff. 4.711:

Zitate aus Heinrich Böll "Ansichten eines Clowns"

dtv 880

(52) "Er war in seinen Idealen getroffen, und meine Mutter wußte Leos Zustand nicht anders zu umschreiben als 'Le ist in einer Krise', und als er dann eine Klassenarbeit fünf schrieb, wollte sie Leo zu einem Psychologen schleppen."

(51) "Auch für mich war es ein Schock, aber kein moralischer, ich konnte mir schon vorstellen, daß es schlimm sein mußte, mit meiner Mutter verheiratet zu sein, deren trügerische Sanftmut eine I- und E-Sanftmut war. Sie sprach selten einen Satz, in dem ein A, O oder U vorgekommen wäre (52) und es war typisch für sie, daß sie Leos Namen in Le abgekürzt hatte. Ihr Lieblingssatz war: 'Wir sehen die Dinge eben verschieden' - der zweitliebste Satz war: 'Im Prinzip habe ich recht, ich bin bereit, gewisse Dinge zu ventilieren.'"

(66) "Manchmal sehne ich mich danach, rheinisch zu hören, so sehr, daß ich von irgendeinem Hotel aus eine Bonner Telefondienststelle anrufe, um diese vollkommen unmartialische Sprache zu hören, der das R fehlt, genau der Laut, auf dem die militärische Disziplin hauptsächlich beruht."

(118) "Wenn Anna mit meiner Mutter abrechnen, ihr die Bücher zeigen wollte, sagte meine Mutter immer: 'Über Geld sprechen - wie gräßlich.' Ein Ä fällt bei ihr hin und wieder, sie spricht es ganz nah an E aus."

(206) "'Sagen sie mir schnell, was sie lieben, gnädige Frau, schnell!' (207) Du wirst ihm die Wahrheit gesagt haben: 'Kinder, Beichtstühle, Kinos, gregorianischen Choral und Clowns.' - 'Und Männer nicht, gnädige Frau?' - 'Doch, einen', wirst du gesagt haben. 'Nicht die Männer als solche, sie sind so dumm.' - 'Darf ich das publizieren?' - 'Nein, nein, um Gotteswillen, nein!' Wenn sie gesagt hat einen, warum sagt sie dann nicht meinen? Wenn man einen Mann liebt, in Worten einen, kann man doch nur seinen meinen, den angetrauten. O, vergessenes, verschlucktes kleines m."

zu Ziff. 4.712:

Endreim

zu:

Hermann Hesse
1877-1962

Auf den Tod eines kleinen Kindes

Jetzt bist du schon gegangen, Kind,
Und hast vom Leben nichts erfahren,
Indes in unsern welken Jahren
Wir Alten noch gefangen sind.

aaaaa
bbb
bbb
aaaaa

Ein Atemzug, ein Augenspiel,
Der Erde Luft und Licht zu schmecken,
War dir genug und schon zu viel;
Du schliefst ein, nicht mehr zu wecken.

=
=
c
ddd
c
ddd

Vielleicht in diesem Hauch und Blick
Sind alle Spiele, alle Mienen
Des ganzen Lebens dir erschienen,
Erschrocken zogst du dich zurück.

=
ee
fff
fff

Vielleicht wenn unsre Augen, Kind,
Einmal erlöschen, wird uns scheinen,
Sie hätten von der Erde, Kind,
Nicht mehr gesehen als die deinen.

=
=
aa
gg
aa
gg

zu Ziff. 4.713:

Oper - nur Geräusch?

zu:

aus: Mark TWAIN, Bummel durch Europa. Zürich 1990.

(56) Ein andermal fuhren wir nach Mannheim und besuchten ein Spektakel - auch Oper genannt -, und zwar "Lohengrin". Das Gebumse und Gepauke, Gedröhn und Gekrache war einfach unglaublich. Der quälende und unbarmherzige Schmerz, den es verursachte, ruht in meinem Gedächtnis gleich neben der Erinnerung an die Zeit, als ich meine Zähne in Ordnung bringen ließ. Gewisse Umstände machten es notwendig, daß ich die vier Stunden bis zum Schluß dablief, und ich blieb da; aber das Andenken dieser langen, schleppenden, harten Leidenszeit ist unzerstörbar. Daß man es schweigend und stillsitzend ertragen mußte, machte es nur noch schlimmer. Ich befand mich in einem Abteil mit acht oder zehn Fremden beiderlei Geschlechts, und das legte mir Zurückhaltung auf; aber zeitweise war der Schmerz so heftig, daß ich kaum die Tränen unterdrücken konnte. In solchen Augenblicken, wenn das Heulen, Jammern und Kreischen der Sänger und das Rasen, Tosen und Krachen des gewaltigen Orchesters immer lauter, immer toller, immer wilder wurden, hätte ich weinen können, wenn ich allein gewesen wäre. Diese Fremden wären wohl nicht überrascht gewesen, jemanden weinen zu sehen, dem man allmählich die Haut abzöge, aber hier hätten sie sich darüber gewundert und zweifellos Bemerkungen darüber gemacht, während doch der vorliegende Fall gegenüber dem Geschundenwerden absolut keinen Vorzug besaß. Nach dem ersten Akt gab es eine halbstündige Pause, und ich hätte in dieser Zeit hinausgehen und mich erholen können, aber ich traute mich nicht, denn ich spürte, daß ich desertieren und draußen bleiben würde. Gegen neun Uhr gab es eine weitere halbstündige Pause, aber mittlerweile hatte ich so viel durchgemacht, daß ich keine Kraft mehr hatte und deshalb nur noch wünschte, in Frieden gelassen zu werden.

4.8 Sprachfamilien - ausdrucksformal erhoben

zu Ziff. 4.81:

Sprachfamilien via LEMPEL/ZIV

zu: vsyst4.401

aus: Studienarbeit Markus Stengel

Vgl.

<http://www-ct.informatik.uni-tuebingen.de/daten/lzxx.pdf>

5.0 Literatur zum Gesamtbereich

Literatur

- BADER, W: Simson bei Delila. Computerlinguistische Interpretation des Textes Ri 13-16. THLI 3. Tübingen 1991: Francke // I.2.7
- GRZYBEK, P; KÖHLER, R (ed./eds.): Exact Methods in the Study of Language and Text - Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday. Quantitative Linguistics 62. Berlin 2007 978-3-11-019354-1: Mouton de Gruyter // I.2.5
- KÖHLER, R u.a (ed./eds.): Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005 // I.2.8
- MITKOV, R (ed./eds.): The Oxford Handbook of Computational Linguistics. Oxford Handbooks Series. Oxford 2003: Oxford University Press // I.2.4
- PELLEN, R: L'Informatique et les Humanités. Bibliographie 1989-1993, d'après quelques périodiques spécialisés. Literary & Linguistic Computing 9,4 (1994) 303-316.
- SCHMITZ, U: Statistische Methoden in der Textlinguistik in: BRINKER, K; ANTOS, G; HEINEMANN, W; SAGER, S F (ed./eds.): Text- und Gesprächslinguistik, Linguistics of Text and Conversation. Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research. HSK 16 / 1. Berlin 2000 196-201. // I.2.4
- SCHWEIZER, H (ed./eds.): Computerunterstützte Textinterpretation. Die Josefsgeschichte beschrieben und interpretiert im Dreischritt: Syntax-Semantik-Pragmatik. THLI 7 / i-iii. Tübingen 1995: Francke // I.2.7
- SCHWEIZER, H: Metaphorische Grammatik. Wege zur Integration von Grammatik und Textinterpretation in der Exegese. ATS 15. St.Ottilien 1981.²1990 // I.2.0
- SCHWEIZER, H: Biblische Texte verstehen. Arbeitsbuch zur Hermeneutik und Methodik der Bibelinterpretation. Stuttgart 1986: Kohlhammer // I.2.0
- SCHWEIZER, H: Fantastische "Opferung Isaaks". Textanalyse in Theorie und Praxis. Lengerich 2006 // I.2.0

zu Ziff. 5.011:

Statistik bei Texten

zu: vprag8.061

aus: SCHMITZ, U, Statistische Methoden in der Textlinguistik: BRINKER, K; ANTOS, G; HEINEMANN, W; SAGER SF (eds.), Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischen Forschung. HSK 16.1. Berlin 2000: de Gruyter. S. 196-201 // I.2.4

(196) Textstatistik insgesamt (1) zählt Textelemente aus und errechnet statistische Kennwerte von Texten, (2) mißt syntaktische und lexikalische Homogenität einzelner Texte oder einer Gruppe von Texten, (3) identifiziert Brüche innerhalb von Texten (sei es aufgrund besonderer Kreativität, Themen- oder Textsortenwechsels, schlechten Stils oder der Beteiligung verschiedener Autoren), (4) vergleicht Texte hinsichtlich quantifizierbarer Eigenschaften (z.B. um Stile, Epochen, Autoren oder Textsorten zu unterscheiden), (5) beschreibt probabilistische Charakteristika von Sprachnormen sowie Abweichungen bzw. Merkmale sprachlicher Varietäten (z.B. Fachsprachen, (197) Soziolekte), Idiolekte oder einzelner Texte), (6) mißt und vergleicht lexikalische Reichhaltigkeit von Texten (z.B. durch Bestimmung der Anzahl verschiedener Wörter im Verhältnis zur Gesamtzahl der Wörter (type-token-ratio), (7) mißt Verständlichkeit von Texten, soweit diese quantifiziert werden kann (vgl. Ballstaedt/Mandl 1988; Hřebíček/Altmann (eds.) 1993, 215-252), (8) beschreibt die allmähliche Entfaltung neuer Information in Texten (vgl. Wildgen 1993), (9) untersucht die lineare Präsentation nicht-linearen Wissens in Texten und (10) sucht allgemeine Eigenschaften, Unterschiede und Gesetzmäßigkeiten in Klassen aller Art von Texten (z.B. mündlich vs. schriftlich, Nachricht vs. Kommentar, Epik vs. Dramatik, Mittelalter vs. Moderne, Dialekt vs. Hochsprache) sowie (11) in "Text" überhaupt. ...

(197) Seltener, aber meist sehr ergiebig, sind analytisch-statistische Untersuchungen zur Eigenart einzelner Texte (Orlov u.a. 1982), zu semantischen Relationen in Texten (Skorochoď'ko 1981, 120-185), zur dynamischen Entwicklung von Merkmalen im Verlauf eines Textes (z.B. Entropie und Wiederholungsrate) (Köhler/Galle 1993), zu Entwicklungslinien in der Schreibweise eines einzelnen Autors (Laffal 199/9, zu langfristigen Entwicklungen im Vokabular und damit verbundenen spezifischen Einstellungsänderungen in der Bevölkerung (z.B. Fortier/ Keen 1997).

zu Ziff. 5.012:

Rekurrenz - Ausdruck/Inhalt

zu: vsem10.5

aus: LINKE, A; NUSSBAUMER, M, Rekurrenz: BRINKER, K; ANTOS, G; HEINEMANN, W; SAGER SF (eds.), Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischen Forschung. HSK 16.1. Berlin 2000: de Gruyter. S. 305-315 // I.2.4

(307) 3.1 Rekurrenz von ausschließlich Ausdrucksseitigem

Es gibt Fälle, wo Ausdrucksseitiges rekurrent ist, ohne dass auch Inhaltsseitiges rekurriert und ohne dass Koreferenz vorliegt. Dazu gehört die Rekurrenz prosodisch-rhythmischer (Takt, Metrum) sowie lautlicher bzw. graphematischer Elemente oder Muster (Endreim, Alliteration, Stabreim). Dasselbe gilt für Rekurrenz auffälliger syntaktischer Muster (Parallelismus, Chiasmus). Solche rein ausdrucksseitige Rekurrenz trägt (auch) zur Stiftung von textuellem Zusammenhalt bei, ihr Haupteffekt ist jedoch ein rhetorischer (vgl. Abschnitt 4.). Auch in Fällen, in denen die Auswahl spezifischer grammatischer Kategorien als markiert erscheint, kann es zu einem zusätzlichen Kohärenzeffekt kommen: So kann z.B. die auffällige Häufung gleicher Wortarten oder die Häufung von Superlativen, Diminutiven etc. (ohne Koreferenz) einen Text oder eine Textpassage als in stilistischer Hinsicht kohärent erscheinen lassen.

3.2 Rekurrenz von (auch) Inhaltsseitigem

Unter dieser Rubrik unterscheiden wir verschiedene Formen von Rekurrenz, die sich alle durch Koreferenz im oben definierten weiten Verständnis auszeichnen. Diese reicht von der engen, "prototypischen" Koreferenz, bei der zwei oder mehr Aspekte sich in einem identischen Referenzpunkt treffen, also sozusagen "Protagonisten" der "Welt-Seite" rekurrenzieren, bis zur Koreferenz im weitesten Sinn, bei der hinter nicht koreferenten Protagonisten ein konzeptuelles Netz, ein gemeinsamer Einordnungsrahmen, also gewissermassen die Kulissen rekurrenzieren. Die Rekurrenz von Inhaltsseitigem (Koreferenz) kann in allen diesen Fällen mit der Rekurrenz von Ausdrucksseitigem verbunden sein, sie muss es aber nicht. Und: Von der ausdrucksseitigen Form (308) kann nicht direkt auf die Inhaltsseite geschlossen werden. So korreliert z.B. die Wiederholung identischen Wortmaterials nicht zwingend mit Koreferenz im engen Sinn.

zu Ziff. 5.013:

Quantitative Linguistik: Probleme

zu: vprag8.062

aus: R. KÖHLER, Gegenstand und Arbeitsweise der Quantitativen Linguistik in: KÖHLER, R u.a. (eds.), Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch. HSK 27. Berlin 2005

(10) Zu den wichtigen methodologischen Problemen bei der Arbeit mit Sprachdaten gehören außerdem:

(1) Inferenzprobleme

(a) Repräsentativität: Keine Stichprobe (auch keine nach den Regeln der Statistik erhobene) kann repräsentative Sprachdaten in dem Sinne liefern, dass in dem in der Statistik üblichen Sinne gültige Schlussfolgerungen auf die Population, das "Sprachganze", möglich wären. Durch Vergrößerung der Stichprobe, z.B. durch Erweiterung eines Textkorpus (11) um weitere Texte, vergrößert sich dagegen die Diversität der Daten im Hinblick auf viele Parameter (Thematik, Stilistik, Genre/Funktionalstil etc.) und damit die Inhomogenität der Daten (s. Punkt b).

(b) Die Homogenität der Daten: Nur homogene Stichproben sind für viele der meistverwendeten statistischen Verfahren geeignet. Diese Bedingung ist für Sprachdaten nur selten erfüllt, möglicher Weise z.B. im Fall von Briefen, die spontan, ohne Unterbrechung und ohne nachträgliche Überarbeitung geschrieben wurden und nicht zu lang sind, sodass über den gesamten Prozess der Textgenerierung konstante Randbedingungen angenommen werden können. Solche einzelnen, kurzen Texte sind allerdings gerade wegen ihrer Kürze nur bedingt aussagefähig (vgl. Punkt a).

(c) Die Normalverteiltetheit der Zufallsvariablen und der Abweichungen: Die wichtigsten Testverfahren, auf denen eine Schlussfolgerung von der Stichprobe auf die Grundgesamtheit ja beruht, setzen voraus, dass die beobachteten Abweichungen von den erwarteten Werten der Zufallsvariablen normalverteilt sind. Diese Voraussetzung ist in der Sprache jedoch nicht erfüllt, so dass eigentlich für jeden einzelnen Fall gesonderte Tests abgeleitet werden müssen (eine mathematisch äußerst unbequeme und in der Praxis nicht durchführbare Forderung).

(d) Die Homoskedastizität: Auch diese Bedingung, die gleichbleibende Varianz über alle Werte der betrachteten Zufallsvariablen, wird von Sprachdaten nicht generell erfüllt und muss besonders sorgfältig überprüft werden, bevor übliche Verfahren der Statistik angewendet dürfen.

(e) Gültigkeitsbedingungen für Gesetzmäßigkeiten: Von einigen Zusammenhängen und Gesetzen ist bereits bekannt, dass zu ihrer Erfüllung bestimmte Bedingungen erfüllt sein müssen. So kann in der Sprache - im Gegensatz zu anderen Phänomenbereichen - nicht vom Gesetz der großen Zahlen ausgegangen werden. Das bekannte Zipf-Mandelbrot-Gesetz gilt z.B. nur für komplette Einzeltexte - nicht aber für Textfragmente oder Textkorpora. Es ist zu vermuten, dass noch viele unbekannte Abhängigkeiten ähnlicher Art existieren, deren Kenntnis für korrekte Schlussfolgerungen unabdingbar wären.

(2) Verteilungsprobleme

(f) Die extreme Schiefe der Häufigkeitsverteilungen: Dieses zentrale und für die Sprache typische Phänomen z.B. von Lauten, Silben, Wörtern (Formen und Bedeutung) und auch von syntaktischen Konstruktionen in Texten führt dazu, dass im Bereich der seltenen Einheiten stets - wie groß die analysierte Textbasis auch sei - eine nicht vernachlässigbare Unterrepräsentation vorliegt. Ein zweites Beispiel betrifft Stichproben aus Wörterbüchern oder Textvokabularen, die zwangsläufig eine Unterrepräsentation kurzer Wörter mit sich bringen.

(g) Direkte und indirekte funktionale Abhängigkeiten zwischen den linguistischen Größen wie Länge, Polysemie, Polytextie etc. bewirken, dass sich die entsprechenden Besonderheiten von Sprachdaten auf jede linguistische Untersuchung auswirken können. Dies gilt für Signifikanztests von Verteilungsanpassungen und Regressionen ebenso wie für Verfahren des Textvergleichs u.a. Für die korrekte Anwendung statistischer Verfahren aus Sprachdaten (bzw. die Entwicklung neuer Methoden für solche Daten) ist daher außer einer guten Kenntnis der mathematischen Statistik allemeint auch immer die genaue Prüfung der jeweiligen Randbedingungen im Einzelnen erforderlich.

Verbreitete Lehrbücher über statistische Verfahren in der Korpuslinguistik suggerieren die problemlose Möglichkeit, Standardmethoden der Statistik auf die Arbeit mit Sprachdaten zu übertragen. Tatsächlich stellen manche von ihnen nichts anderes dar als eine unbedachte Adaption anderer Statistik-Einführungen.

zu Ziff. 5.02:

Ausdrucksseite: SILBE und INTONATION

zu:

Themenheft: Zeitschrift für Sprachwissenschaft 22/1 (2003)

- Silbenbegriff in der Schrift-, Laut- und Gebärdensprache
- Loanword Nativization in German
- Intonation and Syllable Structure

zu Ziff. 5.021:

Lautschrift

zu:

Schwaben Bräu / Werbung

['sabbrlodd]

[am beschda'glei]

['saugu:ds'pilz]

[de:s'me:gedse]

zu Ziff. 5.03:

Phonologie ↔ Morphologie/Syntax(alt)

zu: vgtl.312

aus: R. JACKENDOFF, The representational structures of the language faculty and their interactions, in: C. M. BROWN; P. HAGOORT (eds.), The Neurocognition of Language. Oxford 2000. S. 37-79.

(44) Prosodic units do not correspond neatly to the standard units of syntax. For instance, among prosodic units syllables and feet (a foot, as in the terminology of poetry, is a unit consisting of a stressed syllable and one or more adjacent unstressed syllables, usually to the right). These units often cut across the boundaries of morphemes (stems, prefixes, and suffixes, which are syntactic units within words). Such a situation is seen in (1):

- (1) a. Phonological: [or+ga+ni][za+tion]
 b. Morphosyntactic: [[organ]iz]ation

English articles form a phonological unit with the next word (i.e. they cliticize), whether or not they form a syntactic constituent with it, as shown in (2):

- (2) a. Phonological: [abig[house], [avery][big][house]
 b. Syntactic: [[a][[big][house]], [[a][[very]big][house]]]

And intonational phrasing cuts across syntactic phrase boundaries, as seen in the old chestnut (3):

- (3) a. Phonological: [this is the cat][that ate the rat][that ate the cheese]
 b. Syntactic: [this is [the cat [that [ate [the rat [that [ate [the cheese]]]]]]]]]

Consequently, the constituent structure of PS cannot be produced by simply erasing (or, alternatively, adjusting) syntactic boundaries. For instance, the intonational units of (3a) cannot be identified with any syntactic units such as NP (noun phrase) or VP (verb phrase), and the prosodic feet in (1a), or-ga-ni and za-tion, cannot be identified with any morphosyntactic category.

The upshot is that phonological structure is *constrained* by syntactic structure but not *derived from* it; some of its aspects are characterized by autonomous phonological principles whose structural descriptions make no reference to syntax.

zu Ziff. 6.625:

Sprichwörter und Politik/Wahlkampf

zu:

aus: MIEDER, W, "Yes We Can". Barack Obama's Proverbial Rhetoric: Bern 2009.

(3) Most of Obama's use of folk speech comes in the form of proverbial phrases without any claim to wisdom or truth. He uses these metaphorical phrases to add a certain expressiveness, emotion, color, imagery, and colloquialism to his writings and speeches. It is here where he shows himself to be part of the general population. He prides himself on listening to and thinking about "the voices of all the people" (356), and consequently he mixes their conventional and proverbial language into his utterances. As an impressive intellectual, he does well to follow in the footsteps of a Woodrow Wilson or John F. Kennedy, who also took to heart the folk speech patterns of their constituents. A few telling examples of Obama's folk speech prowess with very little context are: "I had gotten some taste of how the game had come to be played" (16), "We might've fought like cats and dogs" (35), "If we aren't willing to pay a price for our values" (68), "We have played fast and loose with constitutional principles in the fight against terrorism" (56), "It would have been typical of today's politics for each side to draw a line in the sand" (58), "Simply put, they have an ax to grind" (116), "America's schools are not holding up their end of the bargain" (159), "racial discrimination stays on the front burner" (248), "rid themselves of the instinct to throw their weight around" (306), "she asked me out of the blue if our family was rich" (351), etc.

(16) There is no doubt that Barack Obama enjoys politics, and one of his "favorite tasks of being senator is hosting town hall meetings" (101). In his lively descriptions of how politics work and what people think of political activities, he repeatedly turns to proverbial expressions to add a certain emotional expressiveness to his acute comments, i.e., "They're all [Politicians] in the pockets of special interests" (102), "Things went downhill from here" (106), "I knew in my bones" (106), "But it was too little late" (107), "people, who have the luxury of licking their wounds" (107), the "campaign plan called for a bare-bones budget" (110), "my dark-horse status protected me from some of the more dangerous pitfalls [taking bribes] of fund-raising" (113), "you're a typical, two-faced politician" (117), "my status as an underdog" (120), "loses his head with all the publicity" (124), "reporters will go out of their way to stir up the pot" (126), "maintain a straight face during debate" (128), and "they are baptized by fire" (133). The use of somatic phrases, colloquial comparisons to animals, and highly charged idioms with such words as "downhill", "pot", or "fire" add to this account of experiencing politics. And as always, Obama is upbeat about it all, stressing that the frustrations and arguments are well worth it if politicians from both parties come together and actually move things ahead: "Genuine bipartisanship, though, assumes an honest process of give-and-take, and that quality of the compromise is measured by how well it serves some agreed-upon goal, whether better schools or lower deficits" (131).