Ulrike Haß

Korpus-Hermeneutik. Zur hermeneutischen Methodik in der lexikalischen Semantik

1 Mögliche Gegenstände einer linguistischen Hermeneutik

Linguistisches Interpretieren richtet sich zuallererst auf komplexe sprachliche Einheiten, die darauf angelegt sind, von einem Gegenüber verstanden zu werden: Texte, Reden, Gesprächsbeiträge, audiovisuelle Kommunikate und mediale Ereignisse, sprachliche Äußerungen im Zusammenhang mit nonverbalen und paraverbalen Zeichen sowie Text-Bild-Ganzheiten. Dies sind in der Tat die unmittelbar gegebenen und von Individuen wie Gruppen intendierten oder intendierbaren Verstehensangebote an – wiederum – Individuen oder Gruppen von Sprechern. Eine interpretierende Linguistik stellt sich gewissermaßen in die Schuhe wirklicher Rezipienten und Interpreten, mit dem Unterschied, dass ‚natürliche‘, unmittelbare und mehr oder weniger implizite Interpretationsprozesse und -ergebnisse im linguistischen Verfahren expliziert und methodisch reflektiert werden. Gibt es darüber hinaus sinnvollerweise noch weitere Gegenstände einer linguistischen Hermeneutik? – Ja, und zwar je einen oberhalb und unterhalb der Textebene (dazu 1.1 und 1.2).

1.1 Diskurse

Diskurse sind linguistische (und literaturwissenschaftliche, philosophische, historiografische ...) Konstrukte und können als solche Gegenstand eines beabsichtigten Verstehensprozesses werden.¹ Diskurs wird dabei als intertextueller ‚Super‘-Text verstanden, der zum Zwecke der konkreten Analyse methodisch handhabbar, d. h. ‚verkleinert‘ werden muss auf eine Größe, die den eigentlichen Gegenstand des Verstehensinteresses zu repräsentieren in der Lage ist. Diese Größe ist mitunter wiederum ein ausgewählter Einzeltext (so bei S. Jäger und der ‚Duisburger Diskurswerkstatt‘) oder aber eine vernetzte Menge von Schlüsselwörtern, die als Knoten und Brennpunkte des betreffen-

¹ Zur Diskursanalyse vgl. die Arbeiten von Siegfried Jäger (z. B. Jäger 1996), Busse/Teubert (1994), Busse (2003).

den Diskurses verstanden werden (so bei der ‚Düsseldorfer Schule‘, vgl. Stötzel/Wengeler 1995, Stötzel/Eitz 2003).

1.2 Wörter und ihre Bedeutungen

Da jedes Wort respektive jeder lexikalische Ausdruck als Schlüsselwort zwecks Erschließung eines intertextuellen Zusammenhangs zum Gegenstand linguistischen Interpretierens werden kann, können auch lexikalische Ausdrücke, die aus einer Reihe von anderen Gründen interessant sind, zum Gegenstand einer linguistischen Hermeneutik werden. Das Erkenntnisinteresse kann sich auf die Kategorisierungsfunktion von im Wortschatz miteinander vernetzten Lexemen richten, wie sie im Rahmen kognitionslinguistischer Perspektiven thematisiert wird.² Manch öffentliches Interesse an Wörtern ist detaillierter und auf konkrete Lexeme bezogen und setzt so etwas wie eine ‚tief verborgene‘ Semantik auch alltäglicher Bezeichnungen voraus; deren Offenlegung wird gelegentlich von LinguistInnen erwartet (so der Anlass für Haß-Zumkehr 2001b und besonders Haß 2005b). Meist wird von Sprachteilhabern eine solche ‚tief verborgene‘ Semantik über (Re-)Motivierung gesucht, die ihrerseits regelhaft an Bedeutungsgeschichte bzw. Etymologie geknüpft wird. Aber die Linguistik kann auch synchrone Zugänge zur Wortbedeutung in allen ihren Dimensionen eröffnen.

Wortbedeutungen können sehr viel besser synchron, mit empirischen Verfahren aus ihren Verwendungskontexten ‚herausinterpretiert‘ werden. Die alltägliche Redeweise vom ‚Herausinterpretieren‘ unterstellt eine autonome Einheit von Bedeutung, die natürlich nicht gegeben ist – (eine) Bedeutung (eines Worts) ist allererst Ergebnis des jeweiligen Interpretationsprozesses. Dort, wo die Interpretation sich auf eine außerordentlich große Menge von kontextuell vollständigen *parole*-Akten zum fraglichen Ausdruck stützen kann, ist es sogar möglich, sich der *langue*-Bedeutung des Ausdrucks einschließlich seiner pragmatisch-konnotativen Funktionen empirisch zu nähern.

2 Elektronische Textkorpora als Mittel der Bedeutungsinterpretation

Die erwähnte außerordentlich große Menge interpretierbarer Verwendungen eines beliebigen lexikalischen Ausdrucks liegt in Form elektronischer Textsammlungen (Korpora) seit wenigen Jahrzehnten vor. Je größer deren Mate-

² Vgl. etwa Aitchison (1997).

rialmasse, desto notwendiger ist Software, die mittels Sortierungen und Anordnungen ein Korpus erschließt. Auch diese Software prägt die semantische Interpretation der Ausdrücke. Korpora sind Medium für die Analyse prinzipiell beliebiger einzelner Sprachphänomene. Korpora samt zugehöriger Software sind so ein Mittel, um Wortbedeutungen in einem pragmatisch-konnotativ weiteren Sinne zu interpretieren. Die technischen Möglichkeiten ändern allerdings nichts an der Tatsache, dass korpusinterpretierende Verfahren das Gleiche leisten wie das Jahrtausende alte, in den Philologien wie in den Theologien und anderen Textwissenschaften bekannte Parallelstellenverfahren, das lexikalische Ausdrücke oder auch grammatische Ausdrucksmittel anhand verzettelter Belegstellen im Hinblick auf ihre semantischen oder anderen Funktionen interpretiert. Bedingungen und Möglichkeiten der „Zettelkasten“-Hermeneutik liegen auch der technisch unterstützten Erschließung elektronischer Belege immer noch zugrunde. Doch dadurch, dass die Technik die linguistisch gesteuerte Handhabung von Tausenden und Millionen von Belegstellen für einen Ausdruck ermöglicht, bekommt das Verfahren eine andere Qualität.³

Ob Belege aus Zettelkästen oder elektronische Korpora interpretiert werden – es ist zunächst offen, ob das Erkenntnisinteresse sich auf einzelne lexikalische Ausdrücke richtet oder ob es diese Ausdrücke lediglich als Repräsentanten eines Diskurses oder einer Varietät einordnet, um den es bei der Interpretation letztlich gehen soll. Faktisch ist von einem skalenförmigen Skopus des interpretativen Interesses auszugehen, bei dem die Lexeme jedoch ständig mehr oder weniger im Zentrum stehen und der Diskurs bzw. die Varietät des Korpus den Rahmen liefern. Dass es eine kontextfreie semantische Analyse isolierter lexikalischer Ausdrücke nicht geben kann, wird in hermeneutischer Perspektive besonders deutlich. Die Frage nach dem Gegenstand linguistischer Verstehensarbeit hat am Ende zwar nur eine Antwort: Immer hat man es mit Texten zu tun. Nur gibt es an Texten Unterschiedliches zu verstehen und fürs Verstehen zu entdecken, darunter auch die Bedeutung von Wörtern.

Wenn, wie oben behauptet, eine sehr große Menge von *parole*-Akten in Korpora dergestalt interpretiert werden kann, dass der *langue* zugeschriebene konzeptuelle Strukturen bzw. Kategorisierungsfunktionen einfacher oder komplexer Lexeme erkennbar werden und zwar in einem Umfang, der über das in der Lexikografie übliche Maß hinausgeht, dann lässt die Arbeit mit großen Korpora LinguistInnen gewissermaßen zusehen bei einem Prozess, den sie sonst theoretisch voraussetzen und nur von seinem Ergebnis her se-

³ Dass Korpora keine ‚objektiven‘ Analysen ‚auswerfen‘, sondern dass es sich stets um eine plausibilisierungsbedürftige Interpretation handelt, geben auch viele Korpus- bzw. Computerlinguisten offen zu, so Čermak (2002).

hen können: Der Prozess besteht darin, dass der Gebrauch (*parole*) eines Lexems, sobald er usuell wird, das System (*langue*) des Wortschatzes verändert. Insofern mündet die Hypothese in der Frage, ob das Verfahren nicht zugleich vertiefte Einblicke in kognitive, volitive und emotive Strukturen einer, hier der deutschsprachigen, Kultur gestattet.

Das Verfahren der Bedeutungsinterpretation mittels Parallelstellen sei zunächst kurz anhand der klassischen Zettelkasten-Hermeneutik erläutert. Anschließend wird die Weiterführung dieser traditionellen Methodik im Hinblick auf konzeptuelle Strukturen und Kategorisierungen an einem Beispiel der Interpretation von Korpusdaten veranschaulicht.

3 Die klassische Zettelkasten-Hermeneutik

Gesetzt den prototypischen Fall,⁴ ein Lexikograf will sämtliche, v. a. semantische Informationen zu einem Wort erarbeiten, dessen Bedeutung er nicht oder nur vermutend kennt. Ob er sie nicht kennt, weil die zu dokumentierende und in seinen Quellen repräsentierte Varietät ihm fremd ist, oder weil es sich um ein neues oder fremdes Lexem in einer ihm ansonsten bekannten Varietät handelt – im Prinzip ist sein Verfahren immer gleich: Er liest möglichst viele und unterschiedliche Quellentexte ganz oder teilweise durch, sucht nach dem fraglichen Ausdruck und schreibt dessen Verwendungskontexte mit genauem bibliografischem Nachweis auf Zettel, die anschließend nach Stichwort alphabetisch in Kästen geordnet werden. Dieses Verfahren heißt ‚Exzerption‘ und wird wie folgt weitergeführt:

Eine Belegstelle kann aus einem Wort bestehen oder aus einem längeren Textabsatz. Heute weiß man: je mehr Kontext, desto besser, desto mehr kann man über die Bedeutung, die syntaktischen Verknüpfungen und das Wertungs- und Konnotationspotenzial eines Worts herausarbeiten. Am Ende hat man dann zwanzig oder einhundert Zettel aus unterschiedlichsten Texten etwa zum Stichwort *online* und man schreibt einen Wortartikel, in dem der Inhalt der Zettel zusammengefasst, das Typische betont und verallgemeinert wird. Was nachher zu Bedeutung und Verwendung von *online* im Wörterbuch steht, ist eine Abstraktion der vielen einzelfallgebundenen Bedeutungen und Verwendungsbedingungen. Die Frage nach der

⁴ Das nachfolgend geschilderte Verfahren scheint in der Lexikografie so selbstverständlich, dass es mir nicht gelang, eine diesbezüglich explizite und vollständige Beschreibung zu finden. Das Verfahren wird durch persönliche Einführung vermittelt und muss geübt werden. In diesem Abschnitt folgt daher die Explizierung eigener Erfahrungen mit der Lexikografie historischer, fachsprachlicher, gegenwarts- und standardsprachlicher Varietäten.

Grundlage der Abstraktionen kann durch Auflisten der Textquellen ganz genau angegeben werden. (Haß-Zumkehr 2001a: 23.)

„Lexikographische Arbeit ist zum großen Teil Exzerption von Quellen“, stellt Reichmann (1986: 48) fest und betont auch, dass „das Urteil des Lexikographen [...] nicht durch eine generelle Arbeitsanweisung zu ersetzen [ist]“. Und weiter: „Die Delegation der Exzerptionsarbeit an zeitweilig Mitarbeitende, z B. an Hilfskräfte [...] findet hier und an vielen anderen Stellen ihre Grenze“ (Reichmann 1986: 49).

Welchen Grund hat ein viel beschäftigter Wissenschaftler, dass er die geschilderte ermüdende Arbeit nicht oder nur mit Einschränkungen an Hilfskräfte übertragen möchte? – Er weiß, welche interpretativen Vorentscheidungen und Festlegungen unreflektiert schon während der Verzettelung vorgenommen werden und nachher praktisch nicht mehr revidierbar sind: Wählt der Exzerpierende eine Belegstelle, weil sie typisch ist oder gerade weil sie irgendwie besonders auffällig und ungewöhnlich ist? Wählt er einen engeren oder einen weiteren Kontext? Ein weiterer Kontext bedeutet mehr Schreib- oder Kopierarbeit und wird nur gewählt, wenn der Exzerpierer in ihm wichtige semantisch-pragmatische Informationen ausmacht, wozu linguistische Kompetenz bzw. philologisches Gespür notwendig ist. Steht die Menge der Zettel zu einem Wort in einem verlässlichen Verhältnis zur Häufigkeit seines Gebrauchs? Oder legt der Exzerpierer für schwierige und ungewöhnliche Wörter mehr Zettel an als für ihm vertraut erscheinende Wörter? Wie vertraut ist der Exzerpierer überhaupt mit der Varietät bzw. dem Diskurs, aus der bzw. dem das Wort stammt? Kann er den Stellenwert bestimmter Textsorten oder bestimmter Quellen in Relation zu anderen Textsorten/Quellen und zum fraglichen Wort einschätzen und bei der Festlegung des Kontextausschnitts berücksichtigen? All dies hat Auswirkungen auf das Bild von Usualität und Okkasionalität, das die zu einem Wort gesammelten Zettel dem interpretierenden Wortartikel-Verfasser am Ende vermitteln. Ein weiterer Komplex von Vorentscheidungen hängt mit der Identifikation des fraglichen ‚Worts‘ und vor allem mit der Identifikation der seine Polysemie bildenden Einzelbedeutungen (Lesarten) zusammen – diese Identifikationen sind linguistisch-interpretative Setzungen, die Hilfskräfte anders vornehmen als Lexikografen und Lexikologen bzw. die ‚reine‘ Exzerpierer anders vornehmen als solche, die auch die darauf aufbauenden interpretativen Schritte (das Formulieren der Bedeutungsparaphrase) vollziehen. Die Liste der interpretativen Vorentscheidungen, die je nach Vorkenntnissen bzw. Erkenntnisinteressen der Exzerpierer anders ausfallen, ließe sich noch verlängern. Das Gesagte genügt aber bereits, um qualitative Grenzen zu markieren. Auch wenn Nach-Exzerption durch einen mit den Exzerpten unzufriedenen Lexi-

kografen theoretisch möglich ist – faktisch scheitert dies am Zeitaufwand, denn es geht in der Lexikografie immer um ‚viele Wörter‘.

Zu den weiteren Schritten der Zettelkastenhermeneutik: Welche Fragen stellt der Wortartikel-Verfasser schließlich an die Zettelsammlung, die am Ende der Exzerption steht, und welche kann sie ihm überhaupt beantworten? Die ersten klärenden Sortiervorgänge dienen der Unterscheidung von Einzelbedeutungen bzw. der Feststellung von Polysemie und Homonymie und der Feststellung, ob es diachrone Veränderungen in Form und Gebrauch des Worts gibt. Die weiteren Schritte sind Feststellung der gebräuchlichen Schreibung, Genuszuordnung, Flexionsformen etc. – auf Morphologie und Syntax soll hier nicht weiter eingegangen werden, aber auch hierbei müssen mehr interpretative Verfahren eingesetzt werden als gemeinhin angenommen wird. Sprachteilhaber und auch Linguisten gehen üblicherweise von einer deutlich geringeren Variabilität bzw. von einer größeren Festigkeit des Standards aus, als es der Fall ist.

In semantisch-pragmatischer Hinsicht sind Zettelbelege interpretierbar gemäß folgenden Fragen – immer eingeschränkt bezogen auf die Menge bzw. ‚Welt‘ der exzerpierten Quellen. Diese Fragen bedienen sich einer neueren linguistischen Terminologie; implizit fokussiert die klassische Lexikografie bei der Interpretation von Zettelbelegen aber prinzipiell die gleichen Aspekte:

- Welcher semantischen Basiskategorie (Handlung, Ereignis, Zustand, Eigenschaft, Einstellung, Gegenstand, usw.) ist der Ausdruck zuzuordnen? Handelt es sich gegebenenfalls um ein Kollektivum, Kontinuum oder dergleichen? Z.B. ist dies im Gegenwartsdeutschen für die Lexeme *Wissen* und *Information* keineswegs von vornherein klar: *Wissen* wird in Belegkontexten *erworben, vertieft, gesichert, abgefragt, gespeichert und gebündelt*; *Information(en)* hingegen *geliefert, gefunden, gesammelt, ausgetauscht, verarbeitet, herausgefiltert*. Bei je sechs angenommenen Belegen ergibt sich zumindest die Hypothese, *Wissen* könnte zu den Kontinuativa gehören und *Information* zu den Individuativa.
- Was sind die üblichen Referenzobjekte des Ausdrucks? Welche Grenzen hat der gesamte Bereich der Referenz? Sind die Grenzen eher bestimmt oder eher unbestimmt? Was fällt z. B. in den Belegen (üblicherweise) unter *Abgott* oder *Star*, wie stark ist ein *Sturm* in Abgrenzung von *Orkan*, wie schwer ein *Zwischenfall* in Abgrenzung von *Störfall* und *Unfall*?
- In welchen paradigmatischen und syntagmatischen Bezügen steht der Ausdruck und welche semantischen Eigenschaften werden ihm entsprechend seiner Position in einem Wortfeld bzw. in einer Kollokationsbeziehung zugeschrieben?

- Bei Verben und Verbalsubstantiven: Welches sind die typischen Agensrealisierungen – belebte oder unbelebte? Entsprechende Fragen können sich auf weitere Argumentrollen beziehen.
- Bei Adjektiven und nominalen Zustands- und Eigenschaftsprädikatores: Welches sind die Nomina bzw. Referenzgegenstände, die mit ihnen modifiziert werden? Z. B. worauf alles kann man stolz sein bzw. Stolz empfinden?
- Bei Einstellungsprädikatores: Welche Sprecher(-Gruppen) prädisieren typischerweise was und gegebenenfalls mit welcher Wertung über welche Art von Redegegenständen?

Es ist leicht vorstellbar, dass einige der genannten Fragestellungen sich bei eher knappen Kontexten nicht oder nicht sicher beantworten lassen. Dies bedeutet entweder aufwändige Nachsuche oder aber Verzicht auf entsprechend klare Aussagen – die Qualität des Interpretationsergebnisses ist an einige Mühe gebunden. Die klassische Methodik der Zettelkastenhermeneutik stößt aber auch relativ schnell an ihre quantitativen Grenzen: Je nach Art des Wörterbuchprojekts und dessen Zeitrahmen und auch hinsichtlich der kognitiven Kapazität dürfte die Zahl der Belegstellen pro Wort, die in ein und demselben Abstraktionsschritt verarbeitet werden können, zwischen 500 und 1000 liegen.

4 Korpusbasierte Hermeneutik am Beispiel von *Wärme* und *Kälte*

Es soll nun an einem Beispiel erläutert werden, inwieweit ein methodisch geleiteter Umgang mit sehr großen digitalen Textkorpora die traditionelle Zettelkastenhermeneutik erweitert und auch übersteigt. Das gewählte Beispiel ist eine durch Wortbildung und Paradigmatik zusammenhängende Wortgruppe und entstammt dem Grundwortschatz: *warm/Wärme* und *kalt/Kälte* sind in ihrer Semantik zudem kaum ‚frag-lich‘, da ihre Konzepte auf universalen menschlichen Sinneseindrücken beruhen. Dies führt in der lexikografischen Arbeit mit einiger Sicherheit dazu, dass Kontextbelege für tendenziell überflüssig erachtet und die jeweiligen semantischen Angaben kompetenzgestützt formuliert werden. Zudem erhält man für die Beispielllexeme gleichgültig in welchem elektronisch durchsuchbaren Korpus so hohe Trefferzahlen, dass ein belegweises Interpretieren und Abstrahieren wie oben geschildert absurd erschiene.

Nachfolgenden Angaben liegt das Textkorpus des Instituts für Deutsche Sprache, Mannheim, und der zugehörigen Recherchesoftware namens *Cosmas* zugrunde.⁵ Das Textkorpus ist aus verschiedenartigen Textsorten seit der Goethe-Zeit zusammengesetzt; geprägt ist es jedoch von deutschsprachigen Zeitungen und Zeitschriften der jüngsten 20 Jahre und repräsentiert mithin eine Varietät, die zu den meist rezipierten der Gesellschaft zählt. Auf Erläuterung der technischen Verfahren, die *Cosmas* am IDS-Korpus vornimmt und v. a. der Parameter, die LinguistInnen bei der Recherche wählen können, wird hier verzichtet, da sie an anderer Stelle erläutert sind (zuletzt in Haß 2005c). Hinzuweisen ist aber auf die in *Cosmas* enthaltene sog. Kookkurrenzanalyse, die allein es erlaubt, Treffermengen von mehr als 500 oder 1000 wirklich zu nutzen. Es geht dabei um ein statistisch ansetzendes Verfahren, das die in einem definierten Kontext mit einander vorkommenden Lexeme identifiziert und nach der Stärke ihrer Affinität sortiert.

4.1 Interpretation der Wortfrequenz

Folgende gerundete Lexemfrequenzen stehen zur Diskussion:

<i>warm</i> mit sämtlichen Flexionsformen:	87.300,
<i>Wärme</i> mit sämtlichen Flexionsformen:	19.500,
<i>kalt</i> mit sämtlichen Flexionsformen:	123.300,
<i>Kälte</i> mit sämtlichen Flexionsformen:	23.400.

Da die automatische Analyse eines Textkorpus nicht den o. g. Selektionsproblemen und -unregelmäßigkeiten der klassischen Exzerption unterliegt, kann die Frequenz eines Lexems mit expliziter wie impliziter Thematisierung des jeweiligen Konzepts im Diskurs des Korpus in Verbindung gebracht werden: Augenscheinlich werden *Kälte* und *kalt* sehr viel häufiger thematisiert als *Wärme* und *warm*, was als kognitive Markiertheit ‚des Kalten‘ gedeutet werden könnte. Salopp formuliert: Normalerweise ist es warm und dieser Zustand nicht der Rede wert; wird er durch Kaltes gefährdet oder durchbrochen, beginnen Menschen auch sprachlich verstärkt zu handeln, um den Normalzustand des Warmen wiederherzustellen. Drinnen und draußen unterscheiden sich in erster Linie durch Temperaturen; Licht und Dunkel von Tag und Nacht sind eng mit warm und kalt korreliert (typisch: *die kalte Winternacht, der warme Sommersonnenschein*) – dies scheint die kognitive Basiskategorisierung zu sein, auf der nachfolgende Besonderheiten beruhen.

⁵ <http://www.ids-mannheim.de/cosmas2/>. Näheres zur Kookkurrenzanalyse: <http://www.ids-mannheim.de/kt/projekte/methoden/ka.html>.

4.2 Interpretation der Kookkurrenzpartner und typischen Syntagmen, der rechts- und linkerweiternden Komposita zu *warm/Wärme*

Zunächst (nächste Seite) ein beispielhafter und um einige Zeilen gekürzter Ausschnitt aus den Listen der Kookkurrenzen und syntagmatischen Muster (hier zu *Wärme*), mit denen es eine linguistische Interpretation zu tun hat:

Wie man hier z. T. erkennt, sind statistisch signifikante Kookkurrenzen unterschiedlichen linguistischen Kategorien zuzuordnen und letztere entsprechend anders zu interpretieren. Es finden sich:

- attributiv erweiterte Nominalphrasen (*warm* + N; Adj + *Wärme*)
- Phrasen aus einem Verb und seiner wichtigsten Argumentrealisierung
- Reihen von Synonymen bzw. Kohyponymen zu einem nicht explizit gegebenen Hyperonym
- kontrastierende Reihung antonymischer und inkompatibler Lexemkonzepte
- mehr oder weniger stark idiomatische oder lexikalisierte Syntagmen
- Partnerlexeme, die das fragliche Lexem einer Sachdomäne, einem spezifischen Verwendungskontext und damit auch einer Wahrnehmungsperspektive zuordnen.

Auf Grundlage der oben genannten Belegfrequenzen jenseits der 100.000 wurden zu *Wärme* 430, zu *warm* 519 verschiedene Kookkurrenzpartner mit z. T. mehreren syntagmatischen Mustern ermittelt.⁶ Zu den in Affinität und Frequenz dominantesten Indikatoren für Verwendungszusammenhänge gehören:

- bei *warm*: *Wasser, Jahreszeit, anziehen, Strom, Mahlzeit, Wetter, Temperaturen, Sonne, Luft, Kleidung, Licht, kalt, Winter*;
- bei *Wärme*: *Strom, Geborgenheit, Licht, Kälte, Sonne, Energie, Feuchtigkeit, Gas, Herzlichkeit, Liebe, Sehnsucht*.

Wie man sieht, verhalten sich Adjektiv und Nomen in semantisch-pragmatischer Hinsicht nahezu identisch⁷ und können hier als gemeinsames Konzept analysiert werden. Unter Berücksichtigung der gesamten, hier nicht wiederzugebenden Kookkurrenzpartnerlisten lassen sich für die nominale wie adjektivische Realisierung des Konzepts als Verwendungskontexte in der Rei-

⁶ Damit die Software Cosmas ein Lexem überhaupt als Kookkurrenzpartner wertet, muss es mit dem fraglichen Stichwort häufiger zusammen vorkommen als es nach ihrer absoluten Häufigkeit sowieso der Fall ist, denn sonst würden immer nur die Wörter der obersten Frequenzränge (dazu zählen die Funktionswörter) miteinander Kookkurrenzen bilden. D. h. man ‚darf‘ zu Recht semantische und pragmatische Gründe hinter den Kookkurrenzen annehmen.

⁷ Dies ist bei wortbildungsverwandten Lexemen nicht selbstverständlich vorauszusetzen und muss zunächst geprüft werden.

Kookkurrenzen zu *Wärme*

Zahl	LLR	Kookkurrenzen	syntagmatische Muster
36	10556	Strom Blockheizkraftwerk	einem <i>Blockheizkraftwerk</i> [...] <i>Strom</i> [und] <i>Wärme</i>
1244		Strom	<i>Strom</i> [und] <i>Wärme</i>
2	3612	Geborgenheit Liebe Sehnsucht	<i>Liebe</i> ... <i>Wärme</i> <i>Sehnsucht</i> <i>Geborgenheit</i>
27		Geborgenheit Liebe	für alle <i>Liebe</i> <i>Wärme</i> [und] <i>Geborgenheit</i> die ...
13		Geborgenheit Sehnsucht	die <i>Sehnsucht</i> nach <i>Wärme</i> und <i>Geborgenheit</i> in
302		Geborgenheit	<i>Wärme</i> [und] <i>Geborgenheit</i>
12	3585	Licht Farbe	<i>Wärme</i> [...] <i>Licht</i> [und] <i>Farbe</i>
616		Licht	<i>Licht</i> [und] <i>Wärme</i>
18	3570	wohlige sorgt	<i>sorgt</i> ... für <i>wohlige</i> <i>Wärme</i>
12		wohlige verbreitet	<i>verbreitet</i> [...] <i>wohlige</i> <i>Wärme</i>
215		wohlige	die <i>wohlige</i> [...] <i>Wärme</i>
6	3294	Kälte Schall	<i>Kälte</i> <i>Wärme</i> <i>Schall</i> und ...
361		Kälte	<i>Wärme</i> [und ...] <i>Kälte</i>
12	3227	menschliche Nähe	<i>menschliche</i> <i>Wärme</i> und <i>Nähe</i>
402		menschliche	<i>menschliche</i> [...] <i>Wärme</i>
5	2928	Sonne tanken	<i>Sonne</i> und <i>Wärme</i> [...] <i>tanken</i>
4		Sonne frühlingshafte	<i>Sonne</i> und <i>frühlingshafte</i> <i>Wärme</i>
442		Sonne	<i>Sonne</i> [und] <i>Wärme</i>
2		menschlicher Humor	<i>Humor</i> und <i>menschlicher</i> <i>Wärme</i>
3		menschlicher Ausstrahlung	<i>Ausstrahlung</i> ... und <i>menschlicher</i> <i>Wärme</i>
9		menschlicher voll	und ... voll [...] <i>menschlicher</i> <i>Wärme</i>
179		menschlicher	mit <i>menschlicher</i> [...] <i>Wärme</i> und
12	1541	erzeugt Reibung	<i>Reibung</i> [die] <i>Wärme</i> [...] <i>erzeugt</i>
2	1362	erzeugen gleichzeitig Anlagen	<i>Anlagen</i> die <i>gleichzeitig</i> [...] <i>Wärme</i> [...] <i>erzeugen</i>
1		erzeugen Blockheizkraftwerke	<i>Blockheizkraftwerke</i> <i>erzeugen</i> ... <i>Wärme</i>
93		erzeugen	<i>Wärme</i> [...] <i>zu</i> <i>erzeugen</i>
15	1268	Energie elektrische	<i>elektrische</i> <i>Energie</i> [und] <i>Wärme</i> frei
14		Energie Form	<i>Energie</i> in <i>Form</i> von <i>Wärme</i> frei ...

Legende. Spalte **Zahl**: gibt absolute Häufigkeit der jeweiligen Kookkurrenzrelation an; Spalte **LLR** (*log-likelihood-ratio*): gibt einen statistisch ermittelten Wert für die Stärke der Affinität der Kookkurrenzen an, z. B. ist erkennbar, dass die Kookkurrenz von *Wärme* und *Strom* verglichen mit der Kookkurrenz von *Wärme* und *Geborgenheit* im Verhältnis 10.556 zu 3.612 steht, dies entspricht etwa 3 zu 1; Spalte **Kookkurrenzen**: gibt zusammen mit der Spalte **Zahl** differenzierter an, welche Lexeme mit den beiden „Haupt“-Partnern sonst noch häufig und damit eng verknüpft sind; Spalte **syntagmatische Muster**: lässt syntagmatische und syntaktische Relationen zwischen den Kookkurrenzpartnern deutlich werden.

henfolge ihrer Frequenz und daraus zu schließenden kognitiven Dominanz folgende Einzelbedeutungen unterscheiden:

(1) Als Eigenschaft im Bereich „Heizung und Wasser“ im Sinne des technischen Versorgtseins: *überschüssige, nötige, wohltuende, elektrische, heimelige Wärme; Wärme erzeugen, speichern, spenden, verbreiten, produzieren, liefern, tanken, abgeben, dämmen, nutzen, umwandeln, Haushalte mit Wärme versorgen; Wärme entsteht, strahlt, strömt, durchströmt. – Warm sind typischerweise: Wasser, Temperaturen, Dusche, Füße, Bett.*

Der technische Bereich des Heizens und der Versorgung mit warmem Wasser ist durch eine große Zahl verschiedener, teilweise handwerklich-technischer Bezeichnungen und komplexer Wortbildungen (z.B. *Wärme-Kraft-Kopplung, Wärmedämm-; Wärme wird in Kilowattstunden gemessen, Blockheizkraftwerk, Holz u.v.m.*) repräsentiert, die ein offenbar herausragendes Interesse unserer Gesellschaft an warmen Wohnungen belegen und damit vielleicht auch das latente Gefühl der Gefährdung dieses Grundbedürfnisses, jedenfalls das Interesse an (neuen) Techniken der Versorgung mit Wärme.

Linkserweiterungen zu *-wärme* machen nur etwa ein Drittel der Rechtserweiterungen aus und sind ebenfalls überwiegend technisch-physikalischer Provenienz. Die *warme Stube* und *ein warmes und trockenes Plätzchen* sind keineswegs synonym mit einer ‚geheizten Wohnung‘ – es handelt sich um emotionale und auf subjektivem, aber konventionalisiertem Empfinden beruhende Prädikate, die auf die hohe Emotionalität im Hintergrund des „heiz-technischen“ Diskurses verweisen.

(2) Als Eigenschaft im Bereich ‚Wetter und Jahreszeit‘ (Frühling – typisch ist *frühlingshafte Wärme* – und Sommer – typisch ist *sommerlich warm* – werden hier in der genannten Reihenfolge thematisiert): *wohlige, angenehme, behagliche, feuchte, tropische, frühlingshafte, trockene Wärme, Bedürfnis nach Wärme, warme Haut.* Eine eher kleine Gruppe von rechtserweiternden Komposita benennt Wetterphänomene wie *Wärmeeinbruch, warme und kalte Luftmassen* oder *Luft- und Wasserströmungen* sind aus Wetterberichten und Klimaberichten vertraut. – *Warm* sind typischerweise: *Wasser, Jahreszeit, Wetter, Temperaturen, Luft, Sonne, Sommerabend, Süden, Frühlingssonne, in den warmen Süden und in wärmere Gefilde ziehen.*

Wenn man die Verwendungskontexte (1) und (2) mit der o.g. Markiertheithypothese (das Warme ist unmarkiert, das Kalte ist markiert) in Verbindung bringt, ergibt sich die Schlussfolgerung, dass uns das zivilisatorische und künstliche Wärmeklima in Gebäuden zur eigentlichen Natur geworden zu sein scheint, zum Normalfall; die natürlichen und schwankenden Außentemperaturen unserer gemäßigten Breitengrade sind dieser Normalfall gerade nicht. Thematisiert und mit negativen Emotionen verknüpft ist alles, was die

zivilisatorische Selbstverständlichkeit irgendwie einschränkt und sei es durch höhere Preise für Heizenergie.

(3) Das Adjektiv *warm* ist spezifisch assoziiert mit ‚Kleidung‘ und ‚Schlafen‘: *mollig, kuschelig, Bettwärme*. *warm* sind typischerweise: *Socken, Kleidung, Decken, Füße, Bett*; *sich warm anziehen* wird oft appellativ gebraucht, hier wörtlich, aber auch übertragen (s. u.). Keine Wanderung, Jugendfreizeit und dergl. verzichtet auf die stereotype Empfehlung, *warme Kleidung und festes Schuhwerk* mitzunehmen.

(4) Das Adjektiv *warm* wird mit Bezug auf ‚Nahrung‘ gebraucht: die *tägliche warme Mahlzeit* ist fast schon sprichwörtlich, *kalte und warme Speisen und Getränke, durchgehend kalte und warme Küche, kaltes und warmes Büfett, warme Suppe* sind allgegenwärtig, das Alltagswissen empfiehlt *warme Milch mit Honig* (mit 119 Belegen!). Dass etwas wie *warme Semmeln/Weggli* weggeht, ist ein viel gebrauchter idiomatischer Ausdruck, allerdings fast der einzige zu *warm/Wärme*.

(5) Das Nomen *Wärme* wird metaphorisiert und mit weiteren positiven Emotionsausdrücken in eine Reihe gestellt: *menschliche, viel, innere, soziale, mütterliche Wärme, sich nach Wärme sehnen. Wärme und: Geborgenheit, Nähe, Humor, Ausstrahlung, Herzlichkeit, Fürsorge, Liebe, Zuneigung, Sehnsucht, Gefühl, Gemütlichkeit, Ruhe, Witz, Lebensfreude, Güte, Innigkeit, Fröhlichkeit, Offenheit, Emotionalität*. Ähnliche Prädikationen sind in Komposita realisiert, die zum Teil sowohl physische als auch emotionale Wärme miteinander verbinden: *Gefühlswärme, Nestwärme, ein warmes Nest, Herzenswärme, Stallwärme*. Die verglichen mit *kalt/Kälte* geringe idiomatische Potenz von *warm/Wärme* steuert emotionsbezeichnende Wendungen bei: *warm ums Herz werden, warmer Geldregen, mit warmem Händedruck*. – *Zieht euch warm an!* ist ein metaphorischer warnender Appell.

(6) Als Metapher für Farbtöne und Licht sowie für Stimme und Klänge: *warm* sind Farben wie *gelb, orange, rot und braun, Stimme/Timbre, Worte*.

4.3 Interpretation der Kookkurrenzpartner und typischen Syntagmen, der rechts- und linkserweiternden Komposita zu *kalt/Kälte*

Die Tatsache, dass es zu *kalt/Kälte* (vgl. die nächste Seite) deutlich mehr idiomatische Wendungen gibt als zu *warm/Wärme*, könnte, wie schon die bloßen Frequenzen oben, als Folge von kognitivem Hintergrund und Fokus bzw. Markiertheit gedeutet werden, mit dem Zusatz, dass die Markiertheit des Kalten sprach- und kulturgeschichtlich wahrscheinlich schon älter ist, denn etliche der o.g. Idiome sind vor dem bzw. im 19. Jahrhundert belegt, einige

Kookkurrenzen von *Kälte*

Zahl	LLR	Kookkurrenzen	syntagmatische Muster
7	1023 1	klirrender eisigem	<i>klirrender Kälte</i> und <i>eisigem</i>
7		klirrender strahlendem	bei <i>strahlendem</i> Sonnenschein und <i>klirrender Kälte</i>
5		klirrender Celsius	bei <i>klirrender Kälte</i> von minus ... Grad Celsius
597		klirrender	bei <i>klirrender</i> [...] <i>Kälte</i>
1	8716	Regen Wind Sturm	<i>Sturm ... Kälte</i> Wind ... <i>Regen</i>
138		Regen Wind	<i>Regen</i> [...] <i>Wind</i> [und] <i>Kälte</i> ... die
2	8711	eisiger ausharren Freien	<i>eisiger Kälte</i> im Freien ausharren
7		eisiger ausharren	in <i>eisiger Kälte</i> [neben dem toten Piloten] ausharren
10		eisiger Freien	bei <i>eisiger Kälte</i> ... im Freien
6		eisiger strahlendem	bei <i>strahlendem</i> Sonnenschein und <i>eisiger Kälte</i>
601		eisiger	bei / in <i>eisiger</i> [...] <i>Kälte</i>
1	7478	Schnee Eis Matsch	<i>Schnee</i> Eis Matsch ... <i>Kälte</i>
1		Schnee Eis Winter-einbruch	<i>Wintereinbruch ... Kälte</i> Schnee ... Eis
750		Schnee	<i>Schnee</i> [und] <i>Kälte</i>
19	4862	klirrende herrscht	<i>herrscht</i> [...] <i>klirrende Kälte</i>
15		klirrende herrschte	<i>herrschte</i> [...] <i>klirrende Kälte</i>
426		Hitze	<i>Hitze</i> [und] <i>Kälte</i>
1	4132	Hunger Durst Seuchen	<i>Hunger</i> Durst <i>Kälte</i> ... Seuchen
12		Hunger Durst	<i>Hunger</i> Durst [und] <i>Kälte</i>
443		Hunger	<i>Hunger</i> [und] <i>Kälte</i> und
19	4049	Nässe schützen	vor <i>Kälte</i> und <i>Nässe</i> [zu] <i>schützen</i>
20		Nässe Schutz	<i>Schutz</i> vor / gegen <i>Kälte</i> [und] <i>Nässe</i>
9	3062	Wind Wetter	<i>Kälte</i> Wind und Wetter
278		Wind	<i>Wind</i> [und] <i>Kälte</i>
4	2051	klirrenden Tage	der <i>klirrenden Kälte</i> der vergangenen Tage
140		klirrenden	der <i>klirrenden</i> [...] <i>Kälte</i>
3	1913	soziale Gleichgültigkeit	<i>Gleichgültigkeit</i> und <i>soziale Kälte</i>
3		soziale herrschen	<i>soziale Kälte</i> [...] <i>herrschen</i>
5		soziale zunehmende	die <i>zunehmende soziale Kälte</i>
384		soziale	die <i>soziale</i> [...] <i>Kälte</i>

Hinweis. In LLR (Wert für Stärke der Affinität der Kookkurrenzen) ist erkennbar, dass *Kälte* mit *klirrend* etwa 5-mal stärker affin ist als *Kälte* mit *sozial*, dass *Kälte* und *sozial* aber (bereits) stärker kookkurriert als *Kälte* und *bitter*.

auch seit der Zeit nach dem Zweiten Weltkrieg.⁸ Dass sämtliche Idiome negativ wertend sind, verwundert nicht.

Über Idiomatisches hinaus lassen sich die Kookkurrenzen folgenden Einzelbedeutungen bzw. Verwendungskontexten von kalt/Kälte zuordnen; die Reihenfolge richtet sich nach Verwendungshäufigkeit und daraus zu schließender kognitiver Dominanz:

(1) Jahreszeit und Wetter als Bedrohung für Leben, Gesundheit und Wohlbefinden und in Gegenüberstellung zu warmem Wetter und wärmeren Jahreszeiten; dies wird mittels zahlreicher Ausdrucksvarianten mit dem Adjektiv *kalt*, dem Nomen *Kälte* und diversen Komposita und Ableitungen realisiert, v. a.: *kalt und nass*, *kalt und ungemütlich*, *nasskalt*, *klirrend kalt*, *kalt und grau*, *feuchtkalt*, *fußkalt*; Steigerungen: *arschkalt*, *bissig-kalt*, *bitterkalt*, *Eiseskälte*, *eiskalt*, *eisige Kälte*, *hundekalt*, *saukalt*, *Saukälte*, *arktische Kälte*, *sibirische Kälte*, *beißende Kälte*, *grimmige Kälte*, *schneidende Kälte*, *empfindliche Kälte*, *anhaltende Kälte*, *lausige Kälte*, *knackig kalt*, *klamme Kälte*, *kälteempfindlich*, *kälteresistent*. – *Kälte* wird häufig in der Agensrolle verwendet: *Sie weht*, *kriecht*, *dringt in/durch etwas*, *herrscht*, *bedroht*, *etwas/jemand ist ihr ausgesetzt*, *vor ihr muss man sich schützen*, *sie kann einem nichts anhaben*, *man trotzt ihr*, *sie wird vertrieben*, *man entflieht ihr*, *harrt in ihr aus*, *man zittert oder bibbert vor Kälte*, *man leidet unter ihr*, *etwas/jemand ist vor Kälte erstarrt/starr*; positiv ist einzig, dass *Kälte konserviert*.

(2) Das, was wir essen, wird in zwei Kategorien unterteilt: in warme und kalte Speisen. So werden *kalte Speisen und Getränke*, *kalte Platten* weniger in Gegenüberstellung als in positiv bewerteter Verbindung mit *warmen Speisen und Getränken* thematisiert. Verglichen mit dem gleichen Verwendungskontext von *warm* zeigt sich aber ein weitreichender Unterschied: Die „warme Mahlzeit“ wird als lebensnotwendig gesehen und wird neuerdings zum Ziel und Symbol einer sozialen Bewegung („XY-Tafel“); *kalte Platten*, *kaltes Büffett* usw. hingegen gehören nicht zur grundlegenden Ernährung, sondern sind luxuriöse Tafelfreuden und signalisieren einen gewissen Überfluss (hier wird *Speise(n)* und *servieren* verwendet, nicht *Mahlzeit*) oder ein Angebot an Konsumenten (*Sekt kalt stellen*). Auch in der Kochkunst spielen kalte Zutaten (*kalte Butterflocken*) und das Arbeiten mit Kälte (*kalt abschrecken*, *kalt stellen*, *kalt gepresst*) eine Rolle.

(3) Analog zu *warm/Wärme* fungiert *kalt/Kälte* als Metapher für jetzt negativ bewertete soziale und emotionale Haltungen und Zustände: *kalte Berechnung*, *kaltblütig*, *eiskalter Killer*, *eiskalte Mörderin*, *kaltblütiger Mord*, *Gefühlskälte*, *Herzenskälte*, *kaltherzig*, *eiskalte Rache*, *kalt lassen*, *kaltschnäu-*

⁸ Nicht ganz eindeutig hierzu Röhrich (1992).

zig, kalt und abweisend, kalt und unfreundlich, kalt und herzlos, soziale Kälte (hierfür allein ca. 650 Belege), menschliche Kälte, familiäre Kälte, emotionale Kälte, Kälte und Egoismus, Einsamkeit und Kälte. Mit Übergang in die Politik: kalte Aussperrung, kalter Staatsstreich, Kalter Krieg. Wetterbedingte Temperaturempfindungen (Verwendungskontext 1) hängen offenbar auch metonymisch mit der metaphorischen Verwendung zusammen, denn bei typischen Syntagmen wie *Einsamkeit und Kälte*, *Kälte und Unterernährung*, *Hunger und Kälte*, *Kälte und Schmutz* wird zwar auf einen Temperaturzustand referiert, aber zugleich negativ bewertet, dass Menschen diesem Zustand ausgesetzt sind und nicht davor bewahrt werden (können).

(4) *Kalt/Kälte* in physikalischen (auch fachlich meteorologischen) und technischen Zusammenhängen. Dieser Verwendungskontext wird wie auch bei *Wärme* überwiegend mittels Komposition realisiert (*Verdunstungskälte*, *kalte Strömung*, *Kältetechnik* u. v. m.).

(5) In der medizinischen Anwendung: *kalte Umschläge*, *Wickel*, *Kompressen*.

(6) Mit Bezug auf Farben und Licht wird Kaltes sehr viel seltener thematisiert als Warmes: *kaltes blaues Licht*, *kaltes Neonlicht*. Fokus und Hintergrund sind hier vertauscht, mit anderen Worten: In der zivilisatorisch-künstlichen Umwelt bilden warmes Licht und Farben, wie sie in der Natur vorkommen, die Ausnahme und müssen eigens hergestellt werden.

4.4 Wortfeldbeziehungen zu *kühl* und *heiß*

Da als Beispiel für die Korpusinterpretation Lexeme ausgewählt wurden, die in semantisch-paradigmatischer Beziehung zu einander stehen und für die sich Wortfeldbeziehungen im Sinne der *valeur* de Saussures auch in der bisherigen Analyse bestätigt haben, müssen Analysen möglicher weiterer Elemente des Wortfeldes bzw. der Wortfelder mindestens vergleichend hinzugezogen werden. Unter den Kookkurrenzpartnern von *warm* war auch *kühl* und unter denen von *kalt* auch *heiß*. Weitere vergleichend hinzuzuziehende Wortfeldelemente können also ebenfalls korpusgestützt ermittelt werden, nicht immer decken sich die Korpusbefunde mit dem, was Linguisten und/oder Sprachteilhaber aus der eigenen Kompetenz heraus an partiellen Synonymen oder Kohyponymen nennen können.

4.4.1 *kühl*

Das Lexem *kühl* (mit *Kühle*) ist mit ca. 33.500 Belegen das relativ ‚seltenste‘ Konzept im Feld. Analog zu *kalt* und *warm* lassen sich 4 Verwendungskontexte feststellen und wie folgt interpretieren:

(1) Wetter (*wechselhaft und kühl, kühle Luft, Temperaturen, Meeresluft, kühler Wind, ein kühles Lüftchen*). Da es auffälligerweise die Monate April und Mai sind, die in den syntagmatischen Mustern vorkommen, und *kühl* in Wetterberichten regelmäßig in Sommermonaten, in Wintermonaten hingegen *kalt* verwendet wird, kann die These formuliert werden, dass *kühl* eine relationale Bedeutung in Bezug auf eine relativ wärmere Umgebung oder Wettererwartung hat. Eine ergänzende, aber nur exemplarische Recherche in meteorologischer Fachliteratur ergab, dass dort *Warmluft* und *Kaltluft*, *Warmfront* und *Kaltfront* nicht absolut in Grad Celsius bestimmt werden, sondern relational in Abhängigkeit von benachbarten Luftschichten.⁹ *Kühl* ist allerdings kein Ausdruck der meteorologischen Fachsprache. Neben negativen Bewertungen finden sich auch einige positive wie *schön kühl* und das Idiom *der Sprung ins kühle Nass*. Die negative Bewertung des Kühlens in Bezug auf Umgebungstemperatur ist jedoch häufiger.

(2) Lebensmittel und Getränke, v.a. *ein kühles Bier, kühle Drinks, erfrischend kühl*; häufig kontrastiert mit *heißem Wetter*; auch das lexikalisierte *Kühlschrank* und weitere dazugehörige Komposita gehören in diesen Kontext. *Kühl*, nicht *kalt* ist der empfohlene Zustand für die Lagerung v.a. von Nahrungsmitteln *kühl und trocken, dunkel, luftig*.

(3) Zur Charakterisierung der mentalen Haltung bzw. Eigenschaft eines Menschen mit *kühlem Kopf, kühler Rechner, Stratege, Analytiker, kühl kalkulieren, kühl und distanziert, kühl und berechnend, kühl und emotionslos, kühl und rational, kühl reagieren*, gelegentlich dem *heißen Herzen* oder der *Leidenschaft* kontrastierend gegenübergestellt. Anders als die kontextanalogen prototypischen Kookkurrenzen *soziale Kälte* und *Wärme und Geborgenheit* wird *kühl* hier sowohl negativ als auch positiv wertend verwendet: Es scheint, als schwanke die Sprachgemeinschaft bzw. der Diskurs im/hinter dem Korpus, ob sie konsequent rationales Verhalten gut oder doch eher unsozial und schlecht finden soll.

(4) Zur Kennzeichnung von Farben, z.B. *kühles Blau*. Verglichen mit „kalten“ Farben sind „kühle“ positiv bewertet.

Insgesamt ist *kühl* zwar nicht ausschließlich, aber doch eher positiv wertend als *kalt*, das ein eher negatives Prädikat darstellt.

4.4.2 heiß/Hitze

Das Lexem *heiß* ist mit über 100.000 Belegen dreimal so häufig wie *kühl*; nur *kalt* ist noch etwas häufiger belegt (s. o.), dazu kommen über 25.000 Be-

⁹ Vgl. z.B. <http://www.top-wetter.de/lexikon/lexikon.htm>; <http://www.dwd.de/de/-SundL/Freizeit/Hobbymeteorologen/Wetterlexikon>.

lege für *Hitze*. Ähnlich wie bei *kalt* sind die dominanten Verwendungskontexte von *heiß* idiomatischer Natur:

heißes Eisen (mit über 30.000 Belegen das dominanteste Syntagma überhaupt), *heiße Spur*, *heiße Ware*, *heiße Fracht*, *heißes Pflaster*, *in der Hitze des Gefechts*, *die Suppe/etwas wird nicht so heiß gegessen wie sie/es gekocht wird*, *mit heißer Nadel gestrickt*, *heiß laufende Telefone/ Drähte*, *die Hölle heiß machen*, *heißer Herbst*, *heiß hergehen*, *heiß gehandelt*, *fallen lassen wie eine heiße Kartoffel*, *um den heißen Brei herumreden*, *heiß und fettig*,¹⁰ *heiße Sohle aufs Parkett legen*. – Zitathaft: *Manche mögen's heiß, in der Hitze der Nacht*.

Die wichtigste Funktion von *heiß* ist kontextunabhängig die eines Intensivierers: *heiße Phase*, *heiß begehrt*, *ersehnt*, *heiß umkämpft*, *heiß hergehen*, *heiß diskutiert*, *umstritten*; *siedend heiß*, *glühend heiß*, *kochend heiß*, *heiß geliebt*, *heißer Kandidat*.

Analog zu *warm*, *kalt* und *kühl* wird *heiß* auch bei Temperaturen von Luft (*heiß und stickig*, *brütend heiß*, *unerträglich heiß*), Wasser (*heißes Wasser*) und Klima (*heiß und trocken*, *heiß und schwül*) verwendet; *brütende*, *sengende*, *extreme*, *glühende*, *sommerliche*, *drückende Hitze* bezieht sich immer auf mindestens 30 Grad Celsius. *Hitze* ist das einzige Feldelement, das gelegentlich mittels absoluter Temperaturangaben spezifiziert wird. Wie *kalt* ist auch *heiß/Hitze* überwiegend negativ wertend.

Ebenso analog ist der Kontext des Essens und Trinkens: *heiße Tasse Tee*, *heiße Würstchen*, *heiße Suppe*, *in heißem Öl anbraten*, *bei mittlerer, schwacher, mäßiger Hitze braten*, *backen*, *köcheln*, *Hitze reduzieren*, *heiß servieren*, *heiße Milch mit Honig* (mit 134 Belegen annähernd so häufig wie die *warme Milch mit Honig*, s. o.). Das Nomen *Hitze* ist im Kontext des Kochens deutlich häufiger vertreten als alle anderen Feldelemente und wird hier auch durch Attribute noch modifiziert – es kommt beim Kochen erfahrungsgemäß sehr darauf an, ob Lebensmittel 90, 150, 200 oder 300 Grad ausgesetzt werden.

¹⁰ Seit 1994 im und durch den Kulturteil der *taz* auf Musik bezogen verwendet und dann auch in anderen Zeitungen auf politische Probleme übertragen, so von der Frankfurter Rundschau (2.7.1997, S. 2, Ressort *Nachrichten*): „Luxemburg übernimmt mit dem EU-Vorsitz die Aufgabe, Erweiterung und Währungsunion voranzutreiben“: „Weil kein Land mehr zahlen und jedenfalls keines weniger bekommen will als bisher, sind zähe Verteilungskämpfe unvermeidlich. Dazu hat Juncker zwei weitere Menüs von der Sorte ‚heiß und fettig‘ zu balancieren: Einerseits gilt es, das Projekt Währungsunion einigermaßen unbeschädigt durch die entscheidende Phase der Qualifikation zu steuern. Zum andern hat der Amsterdamer Gipfel Order gegeben, der EU beim Kampf gegen die Arbeitslosigkeit endlich ein erkennbares Profil zu verschaffen.“

Die Wortfolge *heiße Herzen* in Kontrast zum *kühlen Kopf* ist nur selten belegt; die assoziativen *Heißsporn* und *Hitzkopf* wurden hier nicht untersucht.

Das Verwendungsspektrum von *heiß/Hitze* weicht erkennbar von denen der anderen Feldelemente ab, insofern *heiß* für eine Intensivierungsfunktion ‚reserviert‘ ist und *Hitze* den Status eines Fachausdrucks der Sprache des Kochens besitzt. In Bezug auf den allen Feldelementen analogen Kontext Wetter/Jahreszeit/Klima und bei den meisten der idiomatischen Wendungen besetzen *heiß/Hitze* die Position des abwertenden Prädikators. Grosso modo, aber mit den dargelegten Einschränkungen und Besonderheiten könnte man sagen, dass *kühl* das Positive am Kalten und *heiß* das Negative am Warmen kennzeichnen.

5 Fazit

Der Umgang mit elektronischen Textkorpora ist ein hermeneutisches Verfahren und unterscheidet sich darin nicht von der klassischen Methode des Parallelstellenvergleichs. In der Korpuslinguistik sind immer wieder Versuche unternommen worden, die semantische Analyse von Lexemen in bzw. aus Textkorpora zu algorithmisieren, aber ebenso oft sind diese Versuche gescheitert, diesbezügliche Absichten revidiert worden (vgl. Renouf/Bauer 2000; Čermak 2002).

Die Interpretation aller Feldelemente von *kalt* bis *heiß* insgesamt zielte hier über dasjenige hinaus, was lexikografisch ‚verwertbar‘ wäre. Die Interpretation ging nicht nur feiner und differenzierter vor, sondern ‚verstieg‘ sich auch immer wieder zu Hypothesen zum Zusammenhang zwischen Wortverwendungshäufigkeit und kognitiven, emotiven und evaluativen semantischen Eigenschaften. So habe ich oben die Frequenz einer Kookkurrenzbeziehung als kognitive Dominanz und letztere als Prototypikalität gedeutet. An anderer Stelle habe ich kookkurrierende nominale Reihungen wie *Wärme und Geborgenheit* als Prädikation im Sinne einer kognitiven Kategorisierung als „diese Wärme hat mit Geborgenheit zu tun, ist so etwas Ähnliches wie Geborgenheit“ interpretiert. Diese und ähnliche Schlussfolgerungen sind der traditionellen Lexikografie insofern fremd, als dort allenfalls von hunderten von *parole*-Akten auf eine Regel der z.T. varietätenspezifischen *langue*, nicht jedoch auf kognitive oder emotive ‚Regularitäten‘ geschlossen wird. Genau dies tue ich aber, wenn ich hunderttausende von *parole*-Akten mit statistischer Hilfe analysiere. Ist dies sprachtheoretisch zu rechtfertigen?

Die Antwort lautet: ja. Linz (2002) hat „Performanz als Ort der Bedeutungserzeugung“ plausibel gemacht und begründet, warum sprachliche Erfahrungen Bedingung kognitiver Strukturierungen sind und nicht umgekehrt (Linz 2002: 227f.). In sehr großen Textkorpora, die aus viel rezipierten, thematisch breit gefächerten und datierbaren Texten einer zentralen Varietät bestehen, sind sprachliche Erfahrungen enthalten, die sich als diverse semantische Dimensionen (Kognition, Emotion, Volition bzw. Einstellung, nach Hermanns 2002) von Lexemen ‚sedimentieren‘. Die Kookkurrenzanalyse macht es möglich, diese ‚Sedimente‘ empirisch nachzuweisen. Die kollektiven Erfahrungen lassen sich in Form syntagmatischer Lexembeziehungen untersuchen und als relativ zeitgebundene kognitive oder emotive Strukturierungen interpretieren. So weit ging die Lexikografie meines Wissens bisher an keiner Stelle. Das hier vorgestellte Verfahren hat denn auch ein gänzlich anderes Ziel als Lexikografie: Es will die kollektiv verbreiteten, also usuellen Bedeutungsdimensionen einzelner Lexeme und Lexemfelder quasi mit einer neuen ‚Lupe‘ erfassen und damit in einer bisher nicht erreichbaren Detailliertheit explizit machen. Dieses Explizieren ist Hermeneutik im eigentlichen und engeren Sinne des Wortes, es ist womöglich eine neue Konstruktion, die Verfahren der neurowissenschaftlichen Analyse der Aktivierung sprachbeteiligter Gehirnareale und einigen Methoden der experimentellen Psychologie wie dem *priming* und *gating* hinsichtlich der Validität ihrer Erkenntnisse zum mentalen Lexikon nicht nachsteht.¹¹ Das korpusermeneutische Verfahren erlaubt es nicht nur zu erkennen, *dass* Lexemkonzepte miteinander verknüpft sind, sondern auch in welcher Art und Weise Lexemvernetzung funktioniert.

Eingeschränkt wird die hier vorgestellte Methode lediglich durch zwei hervorzuhebende Punkte. Erstens: Kein Korpus kann sämtliche sprachlichen Erfahrungen einer Sprachgemeinschaft enthalten – die Grenzen (und Möglichkeiten) des Korpus sind die Grenzen (und Möglichkeiten) der Interpretation. Zweitens: Die Analyse muss sich auf typische Lexemverbindungen (Kookkurrenzen, Kollokationen, Ausdrucksgestalten oder Sinnschemata¹²) konzentrieren, nicht auf einzelne Lexeme mit einer unterstellten vorgängigen ‚wörtlichen‘ Bedeutung, die nachgängig in Kontexten konkretisiert oder modifiziert wird. Der eigentliche Gegenstand der Analyse linguistischer Korpora im Sinne der eingangs behandelten Frage sind, sofern man es will, weder

¹¹ Beim *priming* und *gating* wird im Millisekundenbereich gemessen, wie schnell ein Wort in Abhängigkeit von vorher gesehenen / gehörten Wörtern verstanden wird. Näheres in Aitchison (1997: 30–33).

¹² Hiermit soll an Feilke (1996) und seine Theorie idiomatischer Prägung angeknüpft werden, die für eine theoretische Fundierung der hermeneutischen Methode, die in diesem Beitrag vorzuführen war, ausführlich herangezogen werden müsste.

Wörter noch Texte, sondern das, was Diskurse als intertextuelle Zusammenhänge aus Wörtern machen.

Literatur

- Aitchison, Jean (1997): Wörter im Kopf. Eine Einführung in das mentale Lexikon. – Tübingen: Niemeyer.
- Busse, Dietrich (2003): Diskursanalyse in der Sprachgermanistik. Versuch einer Zwischenbilanz und Ortsbestimmung. – In: Ulrike Haß, Christoph König (Hg.): Literaturwissenschaft und Linguistik von 1960 bis heute. Göttingen: Wallstein, 175–188.
- Busse, Dietrich / Wolfgang Teubert (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. – In: Dietrich Busse, Fritz Hermanns, Wolfgang Teubert (Hg.): Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik. Opladen: Westdeutscher Verlag, 10–28.
- Čermak, František (2002): Today's corpus linguistics. Some open questions. – In: International Journal of Corpus Linguistics 7, 265–282.
- Feilke, Helmuth (1996): Sprache als soziale Gestalt. Ausdruck, Prägung und die Ordnung der sprachlichen Typik. – Frankfurt/M.: Suhrkamp.
- Haß, Ulrike (2005a): Wortbedeutungen – Verstehen – Lernen. – In: Geisteswissenschaften. Essener Unikate. Berichte aus Forschung und Lehre 26/2005, 76–87.
- Haß, Ulrike (2005b): Das Konzept Wasser in der Sprache der alltäglichen und öffentlichen Diskussion. – In: José L. Lozán u. a. (Hg.): Warnsignal Klima. Genug Wasser für alle? Wissenschaftliche Fakten. Hamburg: Wissenschaftliche Auswertungen, 32–35.
- Haß, Ulrike (Hg.) (2005c): Grundfragen der elektronischen Lexikographie. – Berlin, New York: de Gruyter.
- Haß-Zumkehr, Ulrike (2001a): Deutsche Wörterbücher. Brennpunkt von Sprach- und Kulturgeschichte. – Berlin, New York: de Gruyter (de Gruyter Studienbuch).
- Haß-Zumkehr, Ulrike (2001b): Die Grenzen des Stolzes im Deutschen. – In: Sprachreport 2/2001, 2–4.
- Haß-Zumkehr, Ulrike (2002): Das Wort in der Korpuslinguistik. Chancen und Probleme empirischer Lexikologie. – In: Vilmos Ágel u. a. (Hg.): Das Wort. Seine strukturelle und kulturelle Dimension. Festschrift für Oskar Reichmann zum 65. Geburtstag. Tübingen: Niemeyer, 45–70.
- Hermanns, Fritz (2002): Dimensionen der Bedeutung I. Ein Überblick. – In: D. Alan Cruse u. a. (Hg.): Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen. Bd. 1. Berlin, New York: de Gruyter, 343–350.
- Jäger, Siegfried (1996): Diskursanalytische Methoden. – In: Helmwart Hierdeis, Thomas Hug (Hg.): Taschenbuch der Pädagogik. 4. Aufl. Bd. 2. Baltmannsweiler: Schneider-Verlag Hohengehren, 544–554.
- Linz, Erika (2002): Indiskrete Semantik. Kognitive Linguistik und neurowissenschaftliche Theoriebildung. – München: Fink.

- Reichmann, Oskar (1986): Lexikographische Einleitung. – In: Oskar Reichmann, Ulrich Goebel, Robert A. Anderson (Hg.): Frühneuhochdeutsches Wörterbuch. Bd. 1. Berlin/New York: de Gruyter, 10–285.
- Renouf, Antoinette / Laurie Bauer (2000): Contextual Clues to Word-Meaning. – In: *International Journal of Corpus Linguistics* 5, 231–258.
- Röhrich, Lutz (1992): Das große Lexikon der sprichwörtlichen Redensarten. 3 Bände. – Darmstadt: Herder (Lizenzausgabe der Aufl. Freiburg 1992).
- Stötzel, Georg / Thorsten Eitz (Hg.) (2003): Zeitgeschichtliches Wörterbuch der deutschen Gegenwartssprache. Schlüsselwörter und Orientierungsvokabeln. – 2. Ausgabe. Hildesheim: Olms.
- Stötzel, Georg / Martin Wengeler (1995): Kontroverse Begriffe. Geschichte des öffentlichen Sprachgebrauchs in der Bundesrepublik Deutschland. – Berlin, New York: de Gruyter.

Korpus-Hermeneutik

Haß, Ulrike

Dieser Text wird über DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt.

Die hier veröffentlichte Version der E-Publikation kann von einer eventuell ebenfalls veröffentlichten Verlagsversion abweichen.

DOI: <https://doi.org/10.17185/duepublico/48648>

URN: <urn:nbn:de:hbz:464-20190426-111738-0>

Link: <https://duepublico.uni-duisburg-essen.de:443/servlets/DocumentServlet?id=48648>

Rechtliche Vermerke:

Ulrike Haß dankt dem De Gruyter Verlag für die freundliche Genehmigung, diesen Text in der Verlagsversion online veröffentlichen zu dürfen. Zudem dankt Frau Haß den Herausgebern für die Aufnahme des Beitrags in den Sammelband.

Quelle: In: Hermanns, Fritz / Holly, Werner (Hg.): Linguistische Hermeneutik. Theorie und Praxis des Verstehens und Interpretierens. Tübingen: Niemeyer, 2007, S. 241-261. - ISBN 978-3-484-31272-2