häufigsten verwendeten Wörtern, die aus einer etwas mehr als eine Million Wörter umfassenden Tagesstichprobensammlung gewonnen wurden, konnte der Autor lediglich 25 Anglizismen ausmachen. Auch in einer Stichprobe von 100 zufällig ausgewählten Wörtern aus der *Wortwarte* fand er nur 10 rein englische Lehnwörter. Der Verfasser macht außerdem darauf aufmerksam, dass viele englische Entlehnungen wie z. B. *Kid* das deutsche Wort (in dem Fall *Kind*) nicht verdrängen, sondern dessen Bedeutung um neue Aspekte ergänzen. Er bestreitet also keineswegs die sinnvolle Eindeutschung von Anglizismen, übt jedoch Kritik an deren inhaltsleerem Gebrauch.

Zusammenfassend ist festzuhalten, dass es Lothar Lemnitzer gelungen ist, ein sehr lesenswertes Buch über neue deutsche Wörter und lexikalische Wandlungsprozesse zu schreiben. Die vom Autor ausgewählten Wörter werden in Einträgen präsentiert, die den Leser nicht nur mit der Bedeutung der Neubildungen, sondern auch mit deren Ursprungsgeschichte sowie dem kontextuellen Gebrauch vertraut machen. Dank des gut fundierten Wissens wagt der Autor viele interessante Wortvergleiche und stellt die Problematik der Neuschöpfungen in ihrer ganzen Vielfalt dar. Man erlebt die Sprache als eine lebendige Struktur, die sehr anpassungsfähig auf die veränderte Umwelt und die Bedürfnisse der Sprecher reagieren kann. Das Thema selbst und vor allem die klare, leserfreundliche Gedankenführung des Autors machen das Buch attraktiv nicht nur für Sprachwissenschaftler und Studierende der sprachorientierten Fächer, sondern auch für sprachinteressierte Laien. Man kann nun die Hoffnung hegen, dass die zukünftigen Wortsammlungen in der *Wortwarte* den Autor dazu veranlassen, sich dem Thema abermals in einem neuen Buch zu widmen.

Lenz, Friedrich; Schierholz, Stefan J. (Hrsg.):
Corpuslinguistik in Lexik und Grammatik. 2. Auflage. Tübingen: Stauffenburg, 2008. – ISBN 978-3-86057-785-1. 207 Seiten, € 33,00

(Werner Heidermann, Florianópolis / Brasilien)

Google steht drauf, und Corpuslinguistik ist drin. Das macht vielleicht die Relevanz dieser noch immer als recht neu geltenden Teildisziplin deutlich. Eine Suchmaschine macht ja nichts anderes als ein möglichst komplettes Sprachaufkommen als Corpus zu begreifen und es nach vorgegebenen (und vermutlich sehr komplexen) Regeln zu durchsuchen. Stefan J. Schierholz macht auf diesen Zusammenhang bereits in seiner Einführung aufmerksam, nicht ohne zu Recht darauf hinzuweisen, dass die computerlinguistisch ermittelten Suchergebnisse einer professionellen (also linguistischen) Bewertung bedürfen. Das »gesamte Internet als Corpus« (10) zu verstehen, wird in Sonderfällen für legitim erklärt; Schierholz nennt hier die Neologismenforschung als Beispiel.

Bereits an dieser Stelle lässt sich einmal mehr auf das rasante Tempo gegenwärtiger technologischer Entwicklungen hinweisen und damit zusammenhängend auf den unerhörten Aktualisierungsbedarf auch innerhalb der linguistischen Forschung. Der Titel *Corpuslinguistik in Lexik und Grammatik* ist der gerade herausgekommene Nachdruck des ursprünglich 2005 erschienenen Bandes, der wiederum auf die GAL-Tagung 2003 in Tübingen zurückgeht. Das »Googeln« bedurfte 2003 noch der Anführungszeichen – ganz anders heute, wo zahllose Internetnutzer eine der gängigen Suchmaschinen als Startseite definiert haben. Das Googeln selbst wird den Bedürfnissen der Suchenden immer nähergerückt.

Erinnert sich überhaupt noch jemand, wie man früher, prä-*Google* sozusagen, an Informationen gelangt ist?
Wenn das Internet als Corpus definiert und für sprachwissenschaftliche Fragestellungen genutzt werden soll, so Stefan Schierholz, muss wegen des unvermeidlichen »Rauschfaktors« (10) ein Experte mit der Auswertung beauftragt werden. So richtig das ist, so leicht ist vorstellbar, dass die nächsten Generationen von Suchmaschinen diesen Experten nicht erst mit den Ergebnissen arbeiten lassen, sondern so konzipiert werden, dass Corpuslinguisten bereits zu dem Zeitpunkt einbezogen werden, an dem die Suchprozeduren selbst definiert werden. Nicht mehr das »gesamte Internet als Corpus« zu benutzen, sondern das gesamte Internet als Corpus zu konzipieren – hierin könnte ein nächster (qualitativer) Schritt liegen. Ein sehr einfaches Beispiel ist die Frequenz. Ein Frequenzwörterbuch (etwa der deutschen Sprache) lässt sich wahrscheinlich in wenigen Jahren tagesaktuell (!) abrufen. Wichtig wäre es wohl, gar nicht erst die unerhörten Daten- und also Sprachmengen zu generieren, sondern von vornherein sehr gezielt vorzugehen und derart auf Datenqualität zu achten, dass der Rauschfaktor entsprechend reduziert würde.
»Einige grundlegende Überlegungen zur Corpuslinguistik« nennt Schierholz seine Einführung, die sehr hilfreich und instruktiv, kurz und klar ist. Sehr plausibel und überzeugend sind die Grundregeln der Belegauswertung (12) sowie das kurze Votum für die Einbeziehung corpuslinguistischer Verfahren in die linguistische Analyse: »Viele Resultate aus der aktuellen linguistischen Forschung, insbesondere der Grammatikforschung, zeigen, dass eine verstärkte Berücksichtigung der tatsächlichen Sprachdaten dazu beitragen könnte, das freie Theoretisieren, das nur auf Fallbeispiele gestützt ist, abzulösen.« (13) Bemerkenswert ist hier der berechtigte Akzent auf der Grammatikforschung; es geht der Corpuslinguistik ja nicht ums bloße Wörterabzählen. Und das »freie Theoretisieren«, stellt man sich vor, hätte nicht jeder Autor so höflich hingekriegt. Markus Hundt nennt später in etwa dasselbe, ebenfalls noch sehr diskret, »Phantasiekonstruktionen« (27).
Der Sammelband umfasst neben der Einführung sieben Arbeiten, die für eine größere Zahl von Vorträgen stehen, die auf der 34. Jahrestagung der *Gesellschaft für Angewandte Linguistik* gehalten wurden, und zwar allesamt in der Sektion »Lexik und Grammatik«. In mehreren Arbeiten erscheinen denn auch Lexik *und* Grammatik im Titel, etwa in »Corpusbasierte Gewinnung von Daten zur Interaktion von Lexik und Grammatik: Kollokation – Distribution – Valenz« von Ulrich Heid oder in Daniela Wawras »Lexik und Syntax in englischen Jobinterviews: Eine qualitative und quantitative korpuslinguistische Studie«. In den meisten Fällen hingegen steht entweder die Lexik *oder* die Grammatik im Vordergrund. Und nur in einem Beitragstitel ist weder von Lexik noch von Grammatik überhaupt die Rede, in Götz Kaufmanns Aufsatz nämlich: »Der eigensinnige Informant. Ärgernis bei der Datenerhebung oder Chance zum analytischen Mehrwert«. Auf diese Arbeit wird hier als erstes eingegangen, auch weil das Sprachmaterial, mit dem Kaufmann gearbeitet hat, nicht eigentlich ein Corpus ist, sondern eine Belegsammlung, wie bereits in der Schierholz-Einführung (9) und auch noch einmal von Markus Hundt (26) klargestellt wird.
In seiner engagierten und sehr anregenden Analyse beschäftigt Kaufmann sich mit der Verbabfolge im Plattdeutschen von Mennoniten in Brasilien, Paraguay, Bolivien, Mexiko und den USA. Auf die linguistischen Einzelheiten der sehr um-

fassenden Erhebung (mit 300 Informanten und 15.000 untersuchten Sätzen) kann hier aufgrund der Dichte von Daten und der Breite der Darstellung nicht einmal resümierend eingegangen werden. Es muss ein Hinweis auf die teilweise sehr kreativen Erklärungen ausreichen, die Kaufmann für verschiedene sprachliche Realisierungen bereithält, zum Beispiel »die unterschiedliche akustische Deutlichkeit der Präpositionen« (79) als mögliche Begründung für signifikant unterschiedlich frequentes Auftreten von Präpositionalobjekten in den verschiedenen Populationen der Studie. Aus der Studie von gut 30 Seiten soll nur derjenige Punkt angesprochen werden, der mit den methodischen Grundlagen der Corpuslinguistik zu tun hat. Kaufmann nämlich hat eine Belegsammlung geschaffen, indem er Informanten Sätze vom Englischen, Spanischen und Portugiesischen ins Plattdeutsche hat übersetzen lassen. In den sehr sorgfältig dargelegten methodischen Überlegungen geht der Autor selbst mehrmals (62, 65, 67, 81, 90) auf das Problematische dieser Art der Datenermittlung ein. Und doch bleibt ein Zweifel, denn, und das gilt selbstverständlich für die gesamte Branche: Lässt man Informanten übersetzen, dann erfährt man, wie diese übersetzen! Man erfährt nicht (unbedingt), wie sie reden! Und wenn man Informanten »mit einer eher geringen formalen Schulausbildung« (67) Sätze übertragen lässt wie diesen »I have found the book that I have given to the children« (77), dann mag man froh sein, wenn es der Informant in seiner Reaktion beim Eigensinn belässt. Auch die Akzeptabilitätstests lösen das Dilemma nicht; Tests sind ja auch nur Simulationen. So bleibt die Hoffnung, dass irgendwann ähnliche Untersuchungen »im direkten Vergleich mit Sprachmaterial geleistet werden, das im freien Gespräch erhoben wurde« (73).

Dass es der Computerlinguistik nicht um die blindwütige Produktion von Zahlenkolonnen geht, belegt der Beitrag von Markus Hundt mit dem Titel »Grammatikalität – Akzeptabilität – Sprachnorm. Zum Verhältnis von Korpuslinguistik und Grammatikalitätsurteilen«. Hundt geht hier vom System-Norm-Rede-Modell von Hans-Werner Eroms aus und reflektiert sehr gründlich das Wesen der Grammatikalität, die sich »am besten an den Grenzen der Grammatikalität erkunden« (21) lasse. Die Beschreibungen der Kriterien für die Grammatikalitätsurteile (»Varietätenbezug«, »Individualität«, »Gradierbarkeit«, »Wandelbarkeit«, »Hohe Streuung«, 17–21) sind lesenswert. In der Diskussion von »Akzeptabilität und Sprachnorm« stellt sich Hundt das Problem der Legitimität von Setzungen, die sich mit der feststellbaren Frequenz erhöht. Wenn es jedoch hier konkret wird und der Autor selbst nach einer Frequenz fragt, die der »Sprachnormhaftigkeit« (23) Genüge tut, so wird es eng. Und überhaupt: Bei der sehr informativen Bewertung und Kommentierung grammatischer Phänomene (Reflexivpassiv, Dativpassiv, Wortartenübergänge, Entstehung neuer Modalpartikeln, diskontinuierliche Pronominaladverbien und *weil*-V2-Stellung) hätte er doch sehr gern ein paar Zahlenwerte angeben können statt der Informationen »niederfrequent« (31, 37), »relativ frequent« (34) und »praktisch nicht belegt« (36, 37). Trotz der beinahe zwangsläufigen Überschneidung mit der Einführung ist auch das Kapitel »Korpusfragen« ausgesprochen informativ. Die Funktion des Corpus als Korrektiv (27), als Gegenstück also zu den bekannten syntaktisch möglichen, inhaltlich absurden Äußerungen, ist so plausibel beschrieben wie die »Korpus-Lücke« (27) terminologisch begründet ist.

Das relativ Neue der Corpuslinguistik zeigt sich u. a. in der Tatsache, dass viele der Autoren einleitend ihr Corpus- bzw. Korpus-Verständnis darlegen. So auch Carolin Müller-Spitzer in ihrer Arbeit »Erstellung lexikografischer Daten aus Korpora. Eine neue Art elektronischer Wörterbücher«. Es geht um die Arbeit mit dem Wortschatz-Lexikon der Universität Leipzig; es geht einmal mehr um den technologischen Wandel: War nämlich 1997 ein elektronisches Wörterbuch bloß die elektronische Aufbereitung eines konventionellen Buches, so war nur »sieben Jahre später« (43) klar, dass die Informatik sowohl die Wörterbuch-Produktion als auch die Wörterbuch-Nutzung wesentlich (!) verändert hat. Auch in Müller-Spitzers Reflektion geht es berechtigterweise um die Legitimität jedweder Setzung, um den »Anspruch an die Zuverlässigkeit der lexikografischen Daten« (46). Viel Mühe wird verwandt auf die Definition dessen, was überhaupt ein »elektronisches Wörterbuch« ist, für das Müller-Spitzer die alternative Benennung »Wortschatzinformationssystem« (51) vorschlägt. Gewöhnungsbedürftig! In terminologischer Hinsicht geht es ohnehin zuweilen etwas verspielt zu. So steht *elexiko* nicht etwa einfach für ein elektronisches Lexikon, sondern für »elektronisches, lexikologisch-lexikografisches, korpusbasiertes Wortschatzinformationssystem« (55). Solche Begriffsmonster werden die deutsche Corpuslinguistik nicht gerade zu einem Exportschlager machen; die »Lemmazeichengestaltangaben« (56) könnte man ebenfalls entspannter zum Ausdruck bringen. Erfreulich ist, wie die definitorischen Überlegungen im zweiten Teil der Arbeit anhand der *elexiko*-Präsentation exemplifiziert werden.

Zentral (und damit richtig) platziert ist der Beitrag von Ulrich Heid, in dem Aspekte von Lexik und Grammatik zusammengeführt werden. Die Arbeit bezieht mehrere Bereiche der Corpuslinguistik ein. Es geht um die Frage, wie morphosyntaktische Gegebenheiten lexikalischer Kollokationen effektiv ermittelt und sinnvoll (z. B. in einem Lernerwörterbuch) dargestellt werden sollen. Die wegen ihrer Beispiele sehr gut nachvollziehbare Argumentation plädiert für eine noch immer als neu geltende Sichtweise: »lexikalisches Material im Kontext« (102), im grammatischen Kontext. Das Beispiel »baby: be teething« (105) macht klar, wieso diese Form der Darstellung relevanter ist als der herkömmliche Wörterbuchinfinitiv. Interessant der Hinweis auf die dreiwertigen Kollokationen, »signifikante Tripel« (110) genannt, und die (un-)vorstellbaren mathematisch-statistischen Operationen bei der Auswertung dieser Konstruktionen. Sehr bemerkenswert ist, wie dieser Aufsatz sowohl dem lexikografischen Laien (mit dem Kapitel über die automatische Exzerption) wie dem computerlinguistischen Spezialisten (im Kapitel über rekursives Chunking) lohnende Lektüre sein mag.

Der Band beinhaltet weitere drei Beiträge, die sich sehr speziellen Themen widmen. In Torsten Drevers Aufsatz »Markierte Satzstrukturen im Englischen: die Präpositionalphrasen-Initialstellung« ist von eigentlicher Corpuslinguistik nicht die Rede. Dmitrij Dobrovol'skij überschreibt seine Arbeit mit »Paralleles Textcorpus bei der Untersuchung lexikalischer Semantik«, arbeitet mit dem Dostoevskij-Corpus der Österreichischen Akademie der Wissenschaften und ist besonders dort interessant, wo es um Prinzipielles zum Thema Parallelcorpus (154–159) geht. Daniela Wawras Artikel trägt den Titel »Lexik und Syntax in englischen Jobinterviews: Eine qualitative und quantitative korpuslinguistische Studie« und diskutiert, was in allen anderen Vorträgen allenfalls am Rande zur Sprache kommt, nämlich die Aspekte

Authentizität und Repräsentativität des Datenmaterials. Dabei sollte diese Diskussion doch wohl am Anfang jeder computerlinguistischen Erhebung stehen.
Am Ende eine ganz laienhaft computerlinguistisch ermittelte Information zum Fachsprachgebrauch von Linguisten. Was halten Sie für frequenter: »gesprochensprachlich« (Hundt, 29) oder »sprechsprachlich« (Kaufmann, 75)?
Antwort: Ende Oktober 2008 ergibt »gesprochensprachlich« 247, »sprechsprachlich« 3.790 Googles. Die Frage ist, und damit sind wir am Anfang der Computerlinguistik, welche Konsequenzen Frequenzen haben. Es wird ohnehin der Tag kommen, an dem wir aufgrund von ins Uferlose gehenden quantitativen Erhebungen hin und wieder ›lexikalischen (und auch grammatischen) Minderheitenschutz‹ fordern und darauf bestehen werden, so zu reden und zu schreiben, wie es computerlinguistisch für gar nicht möglich gehalten werden wird.

Li, Yuan:
Integrative Landeskunde. Ein didaktisches Konzept für Deutsch als Fremdsprache in China am Beispiel des Einsatzes von Werbung. München: iudicium, 2007. – ISBN 978-3-89129-587-8. 288 Seiten, € 30,00

(Conny Bast, Straßberg)

Die meisten Fremdsprachenlehrer kennen das Problem: Unterrichtet man Landeskunde innerhalb des Sprachunterrichts anhand von authentischen Texten, sind die Themen (besonders im Anfängerunterricht) insgesamt beschränkt. Unterrichtet man faktische Landeskunde als eigenständiges Fach, erhalten die Studenten zwar viel theoretisches Wissen zur Zielkultur, es fehlt jedoch häufig der Bezug zur Praxis und zum späteren Beruf. Aus diesem Grund stellt Yuan Li in der hier vorliegenden Dissertation das von ihr entwickelte Konzept der integrativen Landeskunde vor, welches sie am Beispiel des Einsatzes von Werbung in China erklärt. Das Buch besteht aus 6 Kapiteln nebst einer Einleitung, einem Anhang, der die besprochene Werbung zeigt, und einem Literaturverzeichnis, welches in Literatur auf Deutsch und Literatur auf Chinesisch unterteilt ist.
Kapitel 1 beginnt mit der Darstellung der Didaktik der Landeskunde in Deutschland und gibt zunächst einen Überblick über die Probleme der Benennung und der Definition des Begriffs Landeskunde (19). Li beschreibt anschaulich, dass sich die Landeskunde in drei Blöcken entwickelt hat. Sie galt zunächst als Kulturkunde zur Geistesbildung und sollte den Beweis der Überlegenheit des deutschen Wesens liefern (26), ein Ansatz, der im Dritten Reich von den Nationalsozialisten zu Propagandazwecken missbraucht wurde. Ein weiterer Ansatz ist die Landeskunde als Realienkunde. Sie sollte Informationen über andere Nationen liefern, um auf die Globalisierung bzw. den Konkurrenzkampf vorzubereiten. Seit den 60er Jahren gibt es den faktenorientierten Ansatz, bei dem enzyklopädisches Wissen vermittelt wird, die kommunikative Landeskunde im Dienst des sprachlichen Handelns und den interkulturellen Ansatz zum Verstehen des Fremden und des Eigenen. Jeder Ansatz wird beschrieben und kritisch beleuchtet.
In Kapitel 2 stellt Li dar, wie sich das Herangehen an Landeskunde in China von dem in Deutschland unterscheidet. Bis zum Ende der 80er Jahre kam es in erster Linie auf die Beherrschung von Wortschatz und Grammatik und nicht auf das Wissen über das jeweilige Land an. Bis heute ist der faktenorientierte Ansatz vorherrschend und starken politischen Erwartungen unterworfen (70). Be-