

Stefan Engelberg

GEGENWART UND ZUKUNFT DER ABTEILUNG LEXIK AM IDS: PLÄDOYER FÜR EINE LEXIKOGRAPHIE DER SPRACHDYNAMIK

Lexikographie, Linguistik und das Internet

Gute lexikographische und metalexikographische Praxis ist vor allem auch sprachwissenschaftliche Praxis. Sie involviert das Analysieren sprachlicher Daten, das Formulieren sprachdeskriptiver und sprachtheoretischer Erkenntnisse, das Ergründen kognitiver Prozesse der Sprachverarbeitung und das Erforschen des Einflusses von Wörterbüchern auf das sprachliche Wissen ihrer Benutzer.¹ Entwicklungen in der Linguistik sollten daher auch Entwicklungen in der Lexikographie nach sich ziehen. Ausgehend von dieser Überzeugung sollen auf den folgenden Seiten ein Blick in die Zukunft der Lexikographie geworfen und mögliche Konsequenzen für die Arbeit der Abteilung Lexik am Institut für Deutsche Sprache skizziert werden.

Die Abteilung Lexik hat als Abteilung eines sprachwissenschaftlichen Forschungsinstituts schon immer lexikologische und lexikontheoretische Forschung in vielfältiger Weise mit lexikographischer Theorie und Praxis verknüpft. So ist die lexikographische Vergangenheit und Gegenwart der Abteilung durch eine Reihe von Wörterbüchern gekennzeichnet, die auf intensiver sprachwissenschaftlicher Forschung basieren. Zu diesen lexikographischen Werken, die zu einem großen Teil in dem Wörterbuchportal OWID (www.owid.de; s. auch Müller-Spitzer, dieser Band) miteinander vernetzt dargeboten werden, gehören (i) Wörterbücher im Bereich des Sprachwandels und Sprachkontakts wie die Neubearbeitung des Deutschen Fremdwörterbuchs (s. auch Schmidt, dieser Band), das Lehnwortportal (lwp.ids-mannheim.de), die Neologismenwörterbücher und die am IDS erstellten Teile des Frühneuhochdeutschen Wörterbuchs, (ii) Hand- und Wörterbücher zu syntagmatischen und paradigmatischen Aspekten von Wörtern wie das Handbuch deutscher Kommunikationsverben, die Wortverbindungen online (wvonline.ids-mannheim.de) und das in Arbeit befindliche Handbuch zu Argumentstrukturmustern, – in diesem Zusammenhang sind natürlich auch die außerhalb der Abteilung Lexik erstellten Valenzwörterbücher des IDS zu nennen –; (iii) Wörterbücher, die den öffentlichen Sprachgebrauch erfassen, wie die Reihe von Diskurswörterbüchern (s. Kämper 2008), das Wörterbuch zu Schlüsselwörtern der Wendezeit und natürlich das genuin fürs Internet konzipierte, korpusbasierte Wörterbuch *lexiko* (s. Klosa 2011).

¹ Vgl. Rundell (2012) für einen ähnlichen Standpunkt.

Natürlich hat die Lexikographie neben sprachwissenschaftlichen Anteilen auch Aspekte, die nicht mit genuin sprachwissenschaftlichen Methoden erfasst werden können, etwa benutzungsfunktionale, didaktische, informationswissenschaftliche, kulturhistorische oder wörterbuchstrukturelle. Die folgende Skizze einer zukünftigen Entwicklung der Lexikographie wird aber meiner Überzeugung Ausdruck verleihen, dass die Lexikographie in den gegenwärtigen medialen und linguistischen Umbruchzeiten ihr innovatives Potenzial vor allem dann entfalten wird, wenn wir sie eng mit sprachwissenschaftlicher Theorie und Praxis verknüpfen. Darüber hinaus werden wohl vor allem informationstechnologische Aspekte die Lexikographie von morgen bestimmen.

Lexikographie im Umbruch

Die gegenwärtige Lexikographie ist in einschneidende Veränderungen involviert, denen sie teils ausgesetzt ist und die sie teils mitgestaltet. Da ist erstens der empirische Turn in der Linguistik, der durch quantitative Zugänge zu Korpora neue Arten von Daten verfügbar macht. Zweitens haben wir den Weg beschritten von printbasierten Publikationsformen hin zu internetbasierten Wörterbüchern mit ihren neuen Formen der Strukturierung, Suche und Präsentation lexikographischer Daten. Drittens schließlich erfüllen bestimmte Programme der automatischen Sprachverarbeitung und des Information Retrievals heutzutage Funktionen, die traditionell von Wörterbüchern geleistet werden. Es ist also eine spannende Zeit für die Lexikographie, eine lebensgefährliche geradezu.

Um die Konsequenzen dieser Veränderungen für die Lexikographie einschätzen zu können, ist es nützlich, zwei grundlegende Typen der Benutzung von Wörterbüchern zu unterscheiden. Zum einen sollen Wörterbücher uns helfen, bestimmte sprachliche Probleme in konkreten Kommunikationssituationen zu lösen. Sie sollen uns die Bedeutung eines fremdsprachigen Wortes erklären, einen Ausdruck bei der Textproduktion finden helfen oder die Richtigkeit einer Schreibung bestätigen. Wörterbücher dienen hier der Überwindung von Problemen bei der Sprachverwendung. Zum anderen werden Wörterbücher dazu konsultiert, unabhängig von kommunikativen Problemen Wissen über Sprache zu erlangen. Sie beantworten Fragen nach der Wortgeschichte, nach Sprachkontaktphänomenen oder nach semantischen Zusammenhängen im Wortschatz. In diesem Sinne will ich hier zwischen sprachverwendungs- und sprachwissensorientierter Lexikographie unterscheiden.²

² Vgl. auch Tarp (2008) hinsichtlich einer ähnlichen Unterscheidung von kommunikativen und kognitiven potenziellen Wörterbuchbenutzungssituationen.

Wagt man nun einen Blick in die Zukunft der sprachwissensorientierten und der sprachverwendungsorientierten Lexikographie, so kann man etwas überspitzt formulieren: Sprachwissensorientierte Wörterbücher werden größer werden, vernetzter, bunter und diskursiver, sprachverwendungsorientierte werden situationsspezifischer werden, eher im Hintergrund wirken und weitgehend aus der bewussten Wahrnehmung verschwinden. Oder noch knapper: Sprachwissensorientierte Wörterbücher werden immer auffälliger werden und sprachverwendungsorientierte immer unauffälliger.

Wo man früher beim Verfassen eines fremdsprachigen Textes ein gutes zweisprachiges oder vielleicht auch ein Valenz-, Kollokations- oder Verbwörterbuch herangezogen hätte, um den syntaktischen Anschluss eines Verbs zu ermitteln – oder auch auf die Benutzung eines Wörterbuchs ganz verzichtet hätte –, checkt man heute gebräuchliche präpositionale Anschlüsse über eine schnelle Google-Abfrage. Die Rechtschreibung überprüft uns beim Briefeschreiben die Word-Rechtschreibhilfe in freundlicher Anlehnung an das früher konsultierte Duden-Rechtschreibwörterbuch. Und ein Verständnisproblem in einem türkischen Text löst der Google-Übersetzer in zumindest hinreichender Qualität schneller als der Griff nach einem zweisprachigen Rezeptionswörterbuch. *Ve eğer bana inanmıyorsanız, bir de kendiniz deneyin. Mucize beklemeyin ama.* [Google-Übersetzer Türkisch-Deutsch: *Und wenn Sie mir nicht glauben, versuchen Sie es für sich selbst ein. Aber keine Wunder erwarten.* Sic!] Es stellt sich hier die Frage, ob die Lexikographie nicht eher die Verbesserung der Qualität solcher Werkzeuge unterstützen sollte, anstatt sich an althergebrachte lexikographische Kulturgüter zu klammern. Letztlich liegen ja vielen, wenn auch nicht allen solcher Anwendungen lexikographische Ressourcen zugrunde.

Und wie die Entwicklung weitergeht? Vielleicht so: Auf der Basis kookkurrenzgetriebener Analysen von Kollokationen und semantischen Relationen (s. etwa Belica 2011) werden uns adäquate Formulierungsvorschläge simultan beim Schreiben fremdsprachiger Texte angeboten; gestützt durch neue Methoden des Visual Computing wird über die Fotofunktion von Smartphones eine visuelle Analyse des Kommunikationsumfelds durchgeführt, die unmittelbare Wort-zu-Bild-Zuordnungen erzeugt; die Spracherkennungsfunktion von Google Glasses wird es erlauben, die muttersprachlichen Äquivalente gehörter Wörter direkt in das Brillendisplay einzublenden ... Nun gut: Übertreiben wir es nicht. Vielleicht wird es so kommen, vielleicht auch anders. Aber wie auch immer die Zukunft der sprachverwendungsorientierten Lexikographie aussieht, es deutet sich doch eines an: Letztlich werden wir keine Wörterbücher mehr konsultieren, sondern die Wörterbücher werden uns konsultieren und uns unauffällig und situationsgerecht ihre Dienste anbieten.

Um das Kind nicht mit dem Bade auszuschütten: Auch die sprachverwendungsorientierte Lexikographie wird in bestimmten Bereichen bis auf Weiteres noch (elektronische) Wörterbücher produzieren, etwa für spezifische Ausbildungs- und Prüfungssituationen, also etwa im Bereich des frühen Fremdspracherwerbs oder des muttersprachlichen Unterrichts. Einschneidende Weiterentwicklungen des Konsultationswörterbuchs, also des Wörterbuchs als eines vom Menschen zu benutzenden elektronischen Nachschlagewerks, werden sich aber vor allem im Bereich der sprachwissensorientierten Wörterbücher auf tun. Dazu will ich im Folgenden einen Blick auf die Rolle der Linguistik für die Weiterentwicklung der Lexikographie werfen.

Sprachdokumentation und Lexikographie

Die großen einsprachigen Wörterbücher einer Sprache erfüllen immer auch sprachwissensorientierte Funktionen in dem Maße, in dem sie nicht nur für die Konsultation bei Problemen in konkreten Sprachverwendungssituationen konzipiert sind, sondern auch als Dokumentationen des Wortschatzes einer Sprache dienen. Im Zeitalter von Korpuslinguistik und Internetinformationssystemen kann man einen gegenwärtigen Prototypen eines solchen Wörterbuchs vielleicht wie folgt charakterisieren: (i) Das Wörterbuch ist synchron ausgerichtet. (ii) Es werden in ihm Wörter als zentrale Einheiten auf der Basis großer elektronischer Textkorpora beschrieben. (iii) Diese Korpora fokussieren dabei auf Texte, von denen angenommen wird, dass sie den allgemeinen Sprachgebrauch gut widerspiegeln, d.h. sie fokussieren vor allem auf Presstexte. (iv) Die in den Korpora versammelten Massendaten erlauben es dabei, auf der Basis von Okkurrenzen und Kookkurrenzen bei der Auswahl von Lemmata, Kollokationen oder Gebrauchskontexten Sprachübliches von Unüblichem zu trennen und Ersteres im Wörterbuch zu repräsentieren. Die Umsetzung einer durchdachten Konzeption eines solchen Wörterbuchs stellt etwa das in der Abteilung entwickelte Wörterbuch *lexiko* dar (vgl. Haß 2005; Klosa 2011).

Eine Reihe von neueren Entwicklungen in der Linguistik und auch viele Arbeiten, die in der Abteilung Lexik in den letzten Jahren entstanden sind, deuten aber darauf hin, dass für ein zum Zwecke der Sprachdokumentation geschaffenes lexikalisches Informationssystem – und vielleicht auch für viele sprachverwendungsorientierte Wörterbücher – die oben angeführten vier Charakteristika neu überdacht werden sollten. Ich will dies hier kurz unter den Stichwörtern Musterbildung, Variation, Verteilungen, Exemplarbasiertheit, Dynamik und Vernetzung erläutern und skizzieren, wie man sich im Lichte der gegenwärtigen Linguistik die nächste Generation korpusbasierter Internetwörterbücher vorstellen kann.

Musterbildung: Dass Wörter als semantisch funkelnde, aus ihrem Satzzusammenhang losgelöste Solitäre beschrieben werden können, ist eine Auffassung, die sicherlich nicht mehr viele Anhänger finden wird. Die Lexikographie hat zu diesem Einstellungswandel viel beigetragen, etwa dort, wo sie aus dem britischen Kontextualismus gespeist wurde, wo sie die Erstellung elektronischer Textkorpora und Korpusanalysemethoden vorangetrieben hat oder wo sie valenztheoretische Konzepte verbreitet hat. An der Zentrierung der Lexikographie auf das Wort als Beschreibungseinheit hat das allerdings nicht sehr viel geändert.

Mit dem Aufkommen konstruktionsgrammatischer Sprachtheorien wird nun mancherorts versucht, einen weiteren Schritt zur „Entlexikalisierung“ sprachlichen Wissens zu vollziehen. Die Trennung von Lexikon und Grammatik ist dort weitgehend aufgehoben; sprachliches Wissen konstituiert sich in Form mehr oder weniger abstrakter Konstruktionen, von denen Wörter nur ein Typ unter vielen sind.

Man mag konstruktionsgrammatischen Konzepten durchaus kritisch gegenüberstehen; an der kognitiven Relevanz sprachlicher Muster – seien es nun Kollokationen, Wortbildungsmuster, Argumentstrukturmuster (s. Engelberg, erscheint) oder abstrakte Wortverbindungsmuster (s. Steyer 2011) – können angesichts auffälliger Verteilungsdaten in Korpora eigentlich kaum Zweifel bestehen. Solche Einheiten werden in lexikalischen Informationssystemen aber entweder kaum erfasst oder sie werden nicht als eigenständige Einheiten behandelt, sondern den darin auftretenden Wörtern zugeordnet. Mit der Einführung solcher Phänomene als Beschreibungseinheiten würde sich natürlich auch die nicht triviale Aufgabe stellen, einfach zu nutzende Strukturen des lexikographischen Zugriffs auf solche Daten zu entwickeln.

Variation: Große elektronische Textkorpora legen ihren Schwerpunkt meist auf Zeitungstexte. Das hat unter anderem damit zu tun, dass die für viele Arten sprachwissenschaftlicher Forschung erforderliche Korpusgröße auf andere Weise bisher nicht erreicht werden konnte. Es dürfte andererseits natürlich kaum bestritten werden, dass gerade lexikalische Phänomene sehr stark hinsichtlich Textsorte und Medialität variieren. Das betrifft den lexikalischen Bestand als solches, aber natürlich auch Bedeutungsvarianten von Wörtern. In Zeitungskorpora finden sich zwar viele Lesarten des Verbs *abziehen*, im Sinne des Lösens von Heftgaze bei Verlageinbänden wird man es aber wohl nur in Fachtexten zum Buchbinderwesen finden, und Lesarten wie in *Leider kannst so'n Scheiß ja nich wirklich abziehen* kommt man eher in Chatkorpora auf die Spur. Noch einschneidender ist die Abhängigkeit lexikalischer Phänomene von Textsorte und Medialität allerdings, wenn man Verteilungsdaten für diese Phänomene betrachtet – und zwar weit über die

Verteilung von Lexemen und ihren Bedeutungen hinaus. So sind etwa – um nur ein Beispiel aus den Forschungsarbeiten der Abteilung zu nennen – bei bestimmten Verben die Präferenzen für bestimmte Valenzmuster von Textsorte und Medialität völlig unabhängig, während andere diesbezüglich extreme Verteilungsunterschiede zeigen (vgl. Engelberg et al. 2012).

Verteilungen: In vielen Bereichen der Sprache sind Phänomene in etwa gemäß Zipfscher Regularitäten verteilt. Zipfsche Kurven veranschaulichen, dass in einer Menge von sprachlichen Typen einige wenige Typen relativ zu einer Bezugsgröße sehr häufig auftreten, während sehr viele Typen in dieser Menge extrem selten realisiert werden. Hierzu drei Beispiele: (i) Im Deutschen Referenzkorpus des IDS finden sich ca. 3.000 Lexeme, die als Wortbildungen das Wort *Mond* enthalten. Davon treten acht häufiger als zweitausendmal auf, und zwar (ohne Eigennamen): *Vollmond* (8858), *Mondaufgang* (5227), *Monduntergang* (5121), *Mondlandung* (4299), *Mondfinsternis* (3243), *Neumond* (2516), *Mondschein* (2342) und *Mondlicht* (2252). Etwa tausend „Mond“-Wörter sind dagegen nur einmal im Korpus vertreten, zum Beispiel: *Apfelsinenmond*, *Bilderbuchmondsichel*, *Buttergebäckmond*, *Eichendorfmund*, *Frühlingsneulichtmond*, *Halbmondantilope*, *halbmondbekrönt*, *Heumondkäse*, *Honigmelonenmond*, *Kieselsteinmondlandschaftslust*, *Mondapproximation*, *mondautoartig*, *Mondbärenrettungsprogramm*, *mondscheibleich*, *Mondfinsterniszielgruppe*, *Mondlandungsfürrichtighalter*, *Mondlandungslügenanhänger*, *mondscheinbleich*, *mondscheindurchflutet*, *mondscheinschwanger*, *mondscheinsonatenhaft*, *Mondscheintraumtüllkleid*, *Mondtrottel*, *mondweich*, *Musikdosenmond*, *Neumondgöttin*, *Olivenbrotmondphasenmassage*, *Pergamentmond*, *seelenmondfinster und Vanillemond*. – *Mondsteinerweichend* (1 Beleg), die deutsche Wortbildung, nicht wahr? Im Grunde alles Material für *Mondlyrik* (1 Beleg). (ii) Schaut man sich Argumentstrukturmuster des Typs „etwas [begeistert/ärgert/wundert/...] jemanden an etwas“ in einer Stichprobe von ca. 1.000 Belegen zu insgesamt 39 Verben an, so ist zu beobachten, dass *gefallen* etwa 400 dieser Belege ausmacht, *faszinieren* ca. 200 und *stören* etwa 120, während 25 Verben jeweils weniger als dreimal in dieser Konstruktion vorkommen (s. Engelberg, i. Vorb.). (iii) Geräuschverben treten vielfach auch als Bewegungsverben auf, so zum Beispiel *knattern* (vgl. Engelberg 2009). Schaut man sich in einer Zufallsstichprobe von 100 Belegen die inhaltliche Füllung der Argumentstellen an, so kann man feststellen, dass in fast 90% der Belege ein Fahrzeug (Auto, Motorrad, Flugzeug etc.) relativ zu einem Gebiet, einem Gebäude oder einer Straße knattert, z.B.: *Über die Wiesen hinter dem Feuerwehrhaus knattert inzwischen ein Hubschrauber*. Dies repräsentiert das typische propositionale Muster für *knattern* als Bewegungsverb. Zwei weitere Varianten treten ebenfalls musterhaft, aber nur in geringer Anzahl auf, und zwar Schüsse, die über Personen knattern, und Filme, die durch Projektoren knattern. Dazu kommen

einzelne nicht musterhafte Belege wie: *Bleiern lastet dann der Himmel auf der Insel. Die Krähen knattern durch die klebrige Luft.*

Man darf vermuten, dass solche Verteilungen Grundeigenschaften sprachlichen Wissens widerspiegeln, wobei das Häufige dem kognitiv einfach zu Verarbeitenden, Unauffälligen entspricht, und das Seltene dem kognitiv Aufwändigen, dem Auffälligen und Kreativen (s. Engelberg, erscheint). Wie auch immer wir solche Verteilungen interpretieren, es ist offensichtlich, dass sie charakteristisch für lexikalische Phänomene sind. Ein Wörterbuch, das dies ignoriert und sich auf die Darstellung des Gebrauchshäufigen beschränkt, ist in dem Maße, in dem es seine Aufgabe in der Erfassung der grundlegenden Strukturen des lexikalischen Systems sieht, unvollständig oder – genauer gesagt – falsch. Nun mag man einwenden, dass das Gebrauchsseltene im Normalfall allenfalls die Manifestation der Regelmäßigkeit des Sprachsystems ist und insofern im Wörterbuch nichts zu suchen hat. Auch das möchte ich im Folgenden bestreiten.

Exemplarbasiertheit: Exemplarbasierte Sprachtheorien gehen davon aus, dass einzelne im Gebrauch auftretende sprachliche Ausrücke mit ihren spezifischen Eigenschaften und in ihrem spezifischen Gebrauchskontext memoriert werden und Einfluss nehmen auf die Konstituierung und Produktivität von Kategorien und Konstruktionen. Sie werden damit unmittelbar als Faktoren der Dynamik des sprachlichen Systems aktiv, wodurch eine Unterscheidung zwischen Sprachgebrauch und Sprachsystem weitgehend aufgehoben wird. Jede sprachliche Ausdrucksinstanz ist nicht nur Output einer Regel (oder Abdruck eines Musters), sondern auch unmittelbar Input für die Konsolidierung oder Veränderung dieser Regel. Auch gestaltet sich unter einer exemplarbasierten Sprachauffassung die Dichotomie zwischen dem Idiosynkratischen und dem Regelmäßigen nicht mehr als eine zwischen dem Memorierten und dem im Sprachgebrauch ad hoc Prozessierten: Unter der plausiblen kognitiven Annahme einer hohen Redundanz in der Speicherung sprachlichen Wissens kann davon ausgegangen werden, dass vielfach auch regelhafte Aspekte sprachlicher Phänomene memoriert werden.

Betrachtet man noch einmal die wohl nirgendwo in Wörterbüchern verzeichneten Hapax Legomena aus der Liste unserer „*Mond*“-Wörter, so würden semantisch und formal strukturierte Anführungen solcher Bildungen unter der hier vertretenen Auffassung eine wichtige Rolle in einer Dokumentation des lexikalischen Systems spielen. Denn in ihrer Summe gilt für diese Bildungen: (i) Sie drücken die sprachliche Dynamik im Wortbildungspotenzial von *Mond* aus und sind an der Bestimmung seiner morphologischen Produktivität beteiligt. (ii) Sie manifestieren die sich projektiv auswirkende Relevanz bestimmter semantisch-enzyklopädischer Eigenschaften von *Mond* – etwa in

ihrer produktiven Fokussierung auf assoziative formbezogene (*Kugelmond*, *vollmondförmig*, *Halbmondfrisur*, *mondsichelgleich*, ...) oder visuelle Merkmale (*mondbleich*, *mondlichtfarben*, *mondsilbrig*, *vollmondbleich*, ...) und (iii) sie sind hinsichtlich mancher Formen in ihrer spezifischen Abweichung vom Bestehenden Ausdruck der Mechanismen sprachlicher Kreativität.

Dynamik: Trotz des Wissens um die Dynamik des Lexikons präsentieren sich elektronische Wörterbücher nach wie vor als statisch konzipierte Repositorien lexikalischer Einheiten. Dort, wo Sprachwandelaspekte berücksichtigt werden, etwa im Bereich von Neologismen oder Obsoletheitsmarkierungen, dokumentieren sie eher bereits vollzogenen Wandel, als dass sie das dynamische Potenzial des Wortschatzes erschließen. Solche Wörterbücher sind also partiell historischen, nicht aber dynamischen Charakters. Dabei bieten große Textkorpora und quantitative Analyseverfahren in zunehmendem Maße auch die Möglichkeit, das aktuelle Sprachwandelpotenzial des Wortschatzes darzustellen – in gewissem Maße vielleicht sogar in projektiver Form.

Vernetzung: Die Konzeption des Onlinewörterbuchs als Hypertext hat das nachschlageintensive Verweissystem gedruckter Wörterbücher ersetzt. Damit können Vernetzungen im Wortschatz weit besser dargestellt werden als in Printwörterbüchern, und die elektronische Lexikographie macht auch intensiv Gebrauch davon: Insbesondere semantische Relationen und Wortbildungszusammenhänge sind davon betroffen. Dazu kommen wörterbuchübergreifende Verlinkungen in Wörterbuchportalen und Links in zugeordnete elektronische Textkorpora. So entsteht schnell ein lexikographisches Netz in einem Umfang, der durch die in erheblichem Maße erforderliche IT-gestützte lexikographische Vernetzungsverwaltung widerspiegelt wird (vgl. Storjohann/Meyer 2012). Das Netz ist allerdings üblicherweise nur von innen erschließbar, das heißt ausgehend von lexikalischen Items können Pfade zu anderen lexikalischen Einheiten verfolgt werden. Nach wie vor bieten elektronische Wörterbücher, die auf diesem Wege längst zu hochgradig vernetzten Repositorien von Massendaten geworden sind, aber kaum Zugänge, die das Netz von außen, als Ganzes erschließen und damit Generalisierungen über den erfassten Datenbestand darbieten. Strukturierte Darstellungen von Teilwortschatzen, semantisch-ontologischen Verteilungen, Entwicklungen von Wortbildungsmustern oder des Ausbaus von Systemen semantischer Relationen sind als integrale Bestandteile lexikalischer Informationssysteme die Ausnahme. Dabei könnte gerade dadurch die durch die Wortzentriertheit und alphabetische Anordnung atomisierte Struktur des Wortschatzes in einer Weise aufgebrochen werden wie es der tatsächlichen Vernetztheit des Lexikons entspricht.

Das Potenzial sprachdokumentarischer Internetlexikographie

Wie nun darf man sich ein linguistisch adäquates sprachdokumentarisches Internetwörterbuch in Zukunft vorstellen? Der Weg sollte hinführen zu einem Wörterbuch der Sprachdynamik, einem Wörterbuch also, das Sprache in ihren lexikalischen Aspekten und im Übergang zur Grammatik als ein dynamisches System darstellt. Dazu bedarf es auf linguistischer Seite i) einer stärkeren Orientierung an quantitativen Verteilungen, ii) einer Darstellung des Spannungsbogens zwischen dem Üblichen und dem Kreativen, iii) einer Berücksichtigung textsorten- und medialitätsabhängiger Variation, iv) einer stärkeren Einbeziehung von musterhaften Phänomenen in den Übergangsbereichen zwischen Wortbildung, Wortschatz und Grammatik und (v) damit verbunden auch einer zunehmenden Loslösung vom Wort als einziger oder zentraler Beschreibungseinheit. Auf lexikographisch-struktureller Seite sind wiederum Verfahren der geeigneten Repräsentation sprachlicher Dynamik zu entwickeln. Dazu gehören neuartige Visualisierungstechniken, Adressierungsformen für nicht wortbasierte Beschreibungseinheiten, die Umsetzung eines sich dem Gegenwartsmoment annähernden lexikalischen Sprachwandel-Monitorings und Verfahren zur repositorischen Erfassung des Spektrums zwischen dem Usuellen und dem Kreativen.

Ob man das skizzierte Dokumentations- und Informationssystem zu lexikalisch basierter Sprachmusterdynamik dann noch als Wörterbuch verstehen und bezeichnen möchte, sei an dieser Stelle dahingestellt. Der Grundstein für Entwicklungen dieser Art ist jedenfalls bereits gelegt. Die Abteilung hat partiell standardisierte Arbeitsabläufe für die Erstellung, Vernetzung und Präsentation von Internetwörterbüchern etabliert (Klosa 2011; Engelberg; Klosa/Müller-Spitzer 2009). Lexikographische Präsentationsformen und Zugriffsstrukturen werden kontinuierlich weiterentwickelt (Meyer/Müller-Spitzer 2013) und auf ihre Adäquatheit mithilfe neuer Benutzungsforschungsmethoden überprüft (Müller-Spitzer/Koplenig/Töpel 2012; Müller-Spitzer, erscheint). Zur Darstellung sprachlicher Dynamik werden zurzeit neue Arten interaktiver und animierter Visualisierungen erprobt, die über entsprechende Grafik-Bibliotheken (z.B. www.d3js.org) erzeugt werden können.

Dynamische Aspekte des Wortschatzes sind bereits Gegenstand verschiedener Projekte der Abteilung. Seit längerem wird etwa neben der Erfassung des Fremdwortschatzes im Deutschen Fremdwörterbuch auch die Erweiterung des Wortschatzes durch Neologismen beobachtet und lexikographisch dokumentiert, seit einiger Zeit auch unter Einschluss der Darstellung dynamischer Zeitgebrauchsverläufe. Neben einem noch stärker am Gegenwartsmoment orientierten Monitoring von Veränderungen im deutschen Wortschatz werden in Zukunft auch produktive Prozesse im Wortbildungsbereich dokumentiert werden. Mit dem Ausbau schrift- und gesprochensprachlicher Korpora

und der damit einhergehenden Auffächerung der textuellen Grundlage der lexikographischen Arbeit werden sich auch dem Wortschatz-Monitoring neue Möglichkeiten eröffnen. Insbesondere die Einbeziehung internetbasierter Kommunikation und gesprochener Sprache wird die Beobachtung und Erklärung aktueller lexikalischer Wandelprozesse unterstützen.

Die Berücksichtigung von Okkurrenzen und Kookkurrenzen, insbesondere auf der Basis der im Programmbereich Korpuslinguistik entwickelten Analysemethoden, gehört schon seit längerem zum Standard unserer korpuslexikographischen Arbeit. Seit jüngerer Zeit befasst sich die Abteilung auch mit anderen Verteilungen im Wortschatz, die es in Zukunft in zunehmendem Maße möglich machen werden, an Produktivitätsmaßen und Zipfschen Verteilungsauffälligkeiten orientierte Entwicklungstendenzen im Bereich von Wortschatz und Sprachmustern in dynamische lexikographische Repräsentationen zu überführen. Die in dem Zusammenhang auftretenden Fragen der Musterbildung im Bereich von Argumentstrukturen (Engelberg et al. 2011) und Wortverbindungen (Steyer 2013) sind bereits zentrale Forschungsgegenstände der Abteilung, und entsprechende sprachdokumentarische Präsentationsformen sind in der Entwicklung begriffen (s. etwa wvonline.ids-mannheim.de). Es gibt viel zu tun:



Auf die nächsten 50 Jahre!

Literatur

- **Belica, Cyril** (2011): Semantische Nähe als Ähnlichkeit von Kookkurrenzprofilen. In: Abel, Andrea/Zanin, Renata (Hg.): Korpora in Lehre und Forschung. Bozen, S. 155-178.
- **Engelberg, Stefan** (2009): *Blätter knistern über den Beton*. Zwischenbericht aus einer korpuslinguistischen Studie zur Bewegungsinterpretation bei Geräuschverben. In: Winkler, Edeltraud (Hg.): Konstruktionelle Varianz bei Verben. OPAL, 4/2009. Mannheim, S. 75-97. Internet: <http://www.ids-mannheim.de/pub/laufend/opal/privat/pdf/opal09-4_engelberg.pdf>.
- **Engelberg, Stefan** (im Erscheinen): The argument structure of psych-verbs: A quantitative corpus study on cognitive entrenchment. In: Boas, Hans/Ziem, Alexander (Hg.): Constructional approaches to argument structure in German. Boston/Berlin.
- **Engelberg, Stefan** (in Vorb.): Gespaltene Stimulus-Argumente bei Psych-Verben. Quantitative Verteilungsdaten als Indikator für die Dynamik sprachlichen Wissens über Argumentstrukturen. In: Engelberg, Stefan et al. (Hg.): Argumentstruktur – Valenz – Konstruktionen. Tübingen.
- **Engelberg, Stefan/Klosa, Annette/Müller-Spitzer, Carolin** (2009): Challenges to Internet lexicography: The Internet dictionary portal at the Institute for German Language. In: Kernerman Dictionary News, (July), S. 10-19.
- **Engelberg, Stefan et al.** (2011): Argumentstrukturmuster als Konstruktionen? Identität – Verwandtschaft – Idiosynkrasien. In: Engelberg, Stefan/Holler, Anke/Proost, Kristel (Hg.): Sprachliches Wissen zwischen Lexikon und Grammatik. Berlin/New York, S. 71-112.

- **Engelberg, Stefan et al.** (2012): Argument structure and text genre: cross-corpus evaluation of the distributional characteristics of argument structure realizations. In: *Lexicographica* 28, S. 13-48.
- **Haß, Ulrike** (Hg.) (2005): Grundfragen der elektronischen Lexikographie. *elexiko – das Online-Informationssystem zum deutschen Wortschatz*. Berlin/New York.
- **Kämper, Heidrun** (2008): Diskurswörterbuch – Zur Konzeption eines neuen Wörterbuchs. In: Bernal, Elisenda/DeCesaris, Janet (Hg.): *Proceedings of the XIII Euralex International Congress*. Barcelona: Institut Universitari de Linguística Aplicada / Universitat Pompeu Fabra, S. 689-695.
- **Klosa, Annette** (Hg.) (2011): *elexiko*. Erfahrungsberichte aus der lexikographischen Praxis eines Internetwörterbuchs. Tübingen.
- **Meyer, Peter/Müller-Spitzer, Carolin** (2013): Überlegungen zur Visualisierung von Wortbildung in elektronischen Wörterbüchern. In: Klosa, Annette (Hg.): *Wortbildung im elektronischen Wörterbuch*. Tübingen, S. 255-279.
- **Müller-Spitzer, Carolin** (Hg.) (im Erscheinen): *Using Online Dictionaries*. Berlin/Boston.
- **Müller-Spitzer, Carolin/Koplenig, Alexander/Töpel, Antje** (2012): Online dictionary use: Key findings from an empirical research project. In: Granger, Sylviane/Paquot, Magali (Hg.): *Electronic Lexicography*. Oxford, S. 425-457.
- **Rundell, Michael** (2012): 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. In: Vatvedt Fjeld, Ruth/Torjusen, Julie Matilde (Hg.): *Proceedings of the 15th EURALEX International Congress*. Oslo, S. 47-92.
- **Steyer, Kathrin** (2011): Von der sprachlichen Oberfläche zum Muster. Zur qualitativen Interpretation syntagmatischer Profile. In: Elminger, Daniel/Kamber, Alain (Hg.): *La linguistique de corpus – de l'analyse quantitative à l'interprétation qualitative*. Université de Neuchâtel, S. 219-239.
- **Steyer, Kathrin** (2013): Usuelle Wortverbindungen. Zentrale Muster des Sprachgebrauchs aus korpusanalytischer Sicht. Tübingen.
- **Storjohann, Petra/Meyer, Peter** (2012): Konsistent vernetzte Wörterbucheinträge. In: *Zeitschrift für Angewandte Linguistik* 56, 1, S. 121-134.
- **Tarp, Sven** (2008): *Lexicography in the Borderland between Knowledge and Non-Knowledge*. Tübingen.