



ZDL

Zentrum für digitale Lexikographie
der deutschen Sprache

Paradigmenwechsel in der Lexikographie

Herausforderung und Chance

Vorträge zum Auftakt des Zentrums für
digitale Lexikographie der deutschen Sprache (ZDL)

Herausgeber

Die Präsidenten der Berlin-Brandenburgischen
Akademie der Wissenschaften und der Akademie
der Wissenschaften zu Göttingen

Herausgeber: Die Präsidenten der Berlin-Brandenburgischen Akademie der
Wissenschaften und der Akademie der Wissenschaften zu Göttingen

Redaktion: Julia Naji, Bernhard Roll

Grafik: angenehme gestaltung/Thorsten Probst

Druck: Druckerei Conrad, Berlin

© Berlin-Brandenburgische Akademie der Wissenschaften, 2019

Jägerstraße 22–23, 10117 Berlin, www.bbaw.de

Nachdruck, auch auszugsweise, nur mit ausdrücklicher Genehmigung der Herausgeber.

ISBN: 978-3-939818-87-8

INHALTSVERZEICHNIS

Vorwort	6
Begrüßung	
Martin Grötschel, Präsident der BBAW	8
Grußworte und Einführung	
Andreas Gardt, Präsident der AdWG	13
Thomas Rachel, Staatssekretär BMBF	16
Steffen Krach, Staatssekretär für Wissenschaft und Forschung, Berlin	20
Gegenwart und Zukunft der deutschen Lexikographie	
Wolfgang Klein, Akademiemitglied der BBAW, Gesamtkoordinator des ZDL	23
Den Wortgebrauch verlässlich und umfassend beschreiben: die Ziele der ZDL-Plattform	
Alexander Geyken, Leiter der Berliner Arbeitsstelle des ZDL	35
„Wortgeschichte digital“: Aufgaben und Umsetzungsstrategien	
Volker Harm, Leiter der Göttinger Arbeitsstelle des ZDL	47
Neologismen im Kontext der lexikalischen Forschung am Leibniz-Institut für Deutsche Sprache	
Henning Lobin, Leibniz-Institut für Deutsche Sprache	55
Variatio delectat? Zum Verhältnis von Lexikographie und sprachlicher Vielfalt	
Alexandra Lenz, Österreichische Akademie der Wissenschaften, Universität Wien	64
Sprachpolitik ist Kulturpolitik: Gegenwartslexikographie als öffentliche Dienstleistung in Skandinavien	
Lars Trap-Jensen, Ordnet.dk	89
Digital Lexicography and Oxford Dictionaries: The International Context	
Sarah Ogilvie, OUP Dictionaries	96
Abendvortrag mit anschließendem Gespräch Die deutsche Sprache, ihr Wortschatz und internationale Spracharbeit	
Klaus-Dieter Lehmann, Goethe-Institut, Akademiemitglied BBAW	101

Lars Trap-Jensen ist leitender Redakteur bei ordnet.dk, verfügt über mehr als 25 Jahre Erfahrung in der praktischen Lexikographie und betreut die digitale Ausgabe des Dänischen Wörterbuchs und des Wörterbuchs der dänischen Sprache sowie die Belegstellen-Ressource KorpusDK. Er war und ist Mitglied in verschiedenen lexikographischen Verbänden im In- und Ausland.



Foto: BBAW, Judith Affolter

SPRACHPOLITIK IST KULTURPOLITIK: GEGENWARTSLEXIKOGRAPHIE ALS ÖFFENTLICHE DIENSTLEISTUNG IN SKANDINAVIEN

Die lexikographischen Traditionen in Deutschland und in den skandinavischen Ländern haben mehr gemeinsam, als viele im Allgemeinen annehmen. Die germanischen Sprachen bilden nicht nur einen gemeinsamen Sprachraum, sondern haben auch kulturell und wissenschaftsgeschichtlich sehr viele Ähnlichkeiten. Auf dem Gebiet der Lexikographie lässt sich eine relativ parallele Entwicklung der nationalen Wörterbuchwerke in den germanischsprachigen Ländern verfolgen.

So findet man unter den historischen Wörterbüchern in England das *Oxford English Dictionary*, in den Niederlanden das *Woordenboek van der nederlandsche Taal* und in Deutschland Jacob und Wilhelm Grimms *Deutsches Wörterbuch*. Diese Werke sind allgemein bekannt, aber welche Parallelen gibt es in den skandinavischen Ländern?

In Schweden wurde die Vorbereitung eines nationalen Wörterbuchs schon bei der Gründung der schwedischen Akademie im Jahr 1786 in ihrer Satzung festgelegt. So heißt es in Ziffer 22, dass „die wichtigste Aufgabe der Akademie“ darin bestehe,

„sich für die Reinheit, Stärke und Hoheit der schwedischen Sprache einzusetzen“. Heute würde man wohl eher sagen: für die Klarheit, die Ausdruckskraft und das Ansehen. Wie dem auch sei, die schwedische Akademie arbeitet seitdem an dem Wörterbuch, die ersten 100 Jahre zwar nur mit relativ mäßigem Engagement, der erste Band erschien 1898, aber danach sind die Bände in regelmäßigem Rhythmus herausgebracht worden. Bisher sind 36 Bände und zwei Lieferungen des Bands 37 bis zum Wort „vret“ („Grundstück, Bodenparzelle“) veröffentlicht, und die Vervollständigung des Werkes wird voraussichtlich 2024 erfolgen.

In Dänemark und Norwegen begann die Entwicklung etwas später und mit weniger ehrgeizigen Projekten, die sich jedoch eindeutig von der Entwicklung in den eben genannten Ländern inspirieren ließen: In Dänemark wurde das *Ordbog over det danske Sprog* (Wörterbuch der dänischen Sprache) in der Zeit von 1918 bis 1956 herausgegeben und erschien zunächst in 28 Bänden, zu denen sich später fünf Ergänzungsbände hinzufügten. Insgesamt umfasste es etwa 220.000 Stichwörter. In Norwegen wurden in den 1930er Jahren gleich zwei nationale Wörterbuchprojekte begonnen, da es in Norwegen zwei Hauptvarianten der Schriftsprache gibt. So wurde in den Jahren 1937–1957 das norwegische *Riksmål-Wörterbuch* veröffentlicht, zu dem später noch zwei Ergänzungsbände erschienen, während die Arbeit an dem Wörterbuch für die Variante Nynorsk, *Norsk Ordbok* (Norwegisches Wörterbuch) 1930 begann und erst 2016 abgeschlossen wurde. Insgesamt umfasst letztgenanntes Werk 300.000 Stichwörter in zwölf Bänden.

Auch im Rahmen der sogenannten digitalen Revolution während der vergangenen zwei Jahrzehnte weisen die nationalen Wörterbuchprojekte gemeinsame Merkmale auf. Während dieser ‚zweiten Welle‘ wurden die ursprünglich gedruckten Wörterbücher in digitale Formate umgesetzt, retrodigitalisiert und es wurden in mehreren Ländern zum Teil auch neue Projekte ins Leben gerufen. Dahinter steht die Erkenntnis, dass die großen alten Wörterbuchwerke aus synchroner Perspektive veraltet waren. Dies führte zu neuen Projekten in den Niederlanden, in Großbritannien, Schweden, Dänemark und teilweise in Norwegen und Deutschland. Beispiele sind das *Algemeen Nederlands Woordenboek* in den Niederlanden, OxfordDictionaries.com (eine Sammlung von Wörterbüchern, die inhaltlich u. a. auf dem *Oxford Dictionary of English* und dem *New Oxford American Dictionary* basiert) in Großbritannien, und das DWDS in Deutschland – all dies sind Projekte, die im Rahmen des ZDL natürlich bekannt sind. Weniger bekannt sind vielleicht einige skandinavische Projekte, von denen ich deswegen die prominentesten kurz vorstellen möchte.

Skandinavische Gegenwartslexikographie

In Schweden wurde ein modernes schwedisches Wörterbuch in Verbindung mit der schwedischen Nationalenzyklopädie erstellt. Es wurde in drei Bänden von 1995 bis 1996 veröffentlicht und ist später unter dem Titel *Svensk ordbok* („Schwedisches Wörterbuch“) erschienen, von Redakteuren der lexikographischen Abteilung der Universität Göteborg erarbeitet und von der Schwedischen Akademie finanziert.

Abbildung 1 zeigt die digitalen Ausgaben der drei Wörterbücher der Akademie. Links die SAOL (Abkürzung von *Svenska Akademiens ordlista*, „Wortliste der schwedischen Akademie“), ein orthographisches Wörterbuch mit etwa 120.000 Stichwörtern und deren offiziellen Schreibweisen. In der Mitte das eben erwähnte *Svensk ordbok* in digitaler Form und rechts das historische Wörterbuch der schwedischen Akademie. Diese drei Wörterbücher werden den Benutzern kostenlos zur Verfügung gestellt. Das *Svensk ordbok* wird regelmäßig, aber in begrenztem Umfang, aktualisiert. Der Betrieb der Website und die Entwicklung des Wörterbuchs sind bis 2060 durch eine Vereinbarung zwischen der Akademie und der Universität Göteborg gewährleistet.



Abb. 1: Portalseite der drei schwedischen Akademiewörterbücher, svenska.se

In Norwegen gibt es zwei erwähnenswerte moderne Wörterbuchprojekte. Zunächst das *Wörterbuch der norwegischen Akademie NAOB*, das 2018 erschienen ist. Dies ist eine digitale und völlig modernisierte Version des eingangs erwähnten historischen *Riksmål-Wörterbuchs*, das ursprünglich seit den 1930er Jahren erarbeitet und in einem etwas unregelmäßigen Rhythmus in Einzelteilen herausgegeben wurde (siehe Abbildung 2).



Abb. 2: Startseite des Wörterbuchs der norwegischen Akademie, *naob.no*

Das NAOB-Wörterbuch ist also in erster Linie ein Digitalisierungsprojekt, bei dem das Originalwerk mit seinen Ergänzungsbänden retrodigitalisiert und zu einem digitalen Wörterbuch zusammengesetzt wurde. Darüber hinaus wurde die Artikeldatei mittels einer sowohl strukturellen wie inhaltlichen Überarbeitung und Ergänzung modernisiert. Das Ergebnis ist ein Wörterbuch mit etwa 225.000 Stichwörtern in aktualisierter, Bokmål-naher Orthographie.

Neben dieser Arbeit gibt es auch ein kleineres Wörterbuch, das die beiden offiziellen Standardsprachen Bokmål und Nynorsk abdeckt und ursprünglich als Gemeinschaftsprojekt des Norwegischen Sprachrats und der Universität Oslo entstand. Damals, im Jahre 1986, handelte es sich um zwei eigenständige gedruckte Wörterbücher, die jedoch mit identischer Struktur und gleichem Inhalt für die beiden Hauptvarianten des Norwegischen parallel erarbeitet wurden. Zwischenzeitlich haben sie sich aber etwas unterschiedlich entwickelt, was insbesondere aus dem Stichwortbestand hervorgeht: So gibt es für Bokmål 70.000 Stichwörter, aber 90.000 für Nynorsk. Der Unterschied macht sich vor allem im Moment der digitalen Nutzung bemerkbar, bei der man die Artikel aus beiden Werken parallel betrachten kann. Eine Überprüfung der beiden Wörterbücher wurde im vergangenen Jahr eingeleitet, um den Wortbestand beider Wörterbücher zu vereinheitlichen und auf 100.000 zu erhöhen. Das Projekt ist auf fünfzehn Jahre angelegt und wird an der Universität Bergen durchgeführt, die im Zusammenhang mit einer umfassenden Neuorganisation der lexikographischen Arbeit in Norwegen die redaktionelle Verantwortung dafür übernommen hat (siehe Abbildung 3).

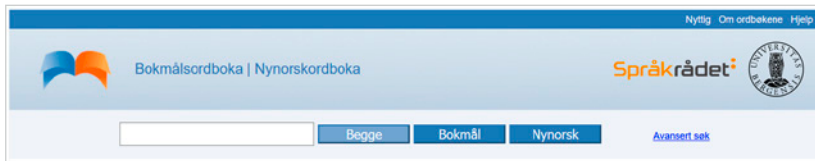


Abb. 3: Startseite des *Bokmåls-/Nynorskboka*, <https://ordbok.uib.no>

In Dänemark deckt die Website *ordnet.dk*, und hier insbesondere *Den Danske Ordbog*, DDO (Das Dänische Wörterbuch), den modernen Wortschatz ab. DDO wurde ursprünglich als gedrucktes Wörterbuch in sechs Bänden 2002–2005 veröffentlicht, wurde jedoch bereits bei Aufnahme der redaktionellen Arbeiten im Jahre 1991 für die digitale Weiterverwendung konzipiert und als komplex strukturiertes SGML-/XML-Dokument erstellt. Es war außerdem das erste korpusbasierte Wörterbuch des Dänischen.

Seit 2009 ist eine digitale Version im Web unter *ordnet.dk* verfügbar, wo Nutzer ebenfalls auf eine retrodigitalisierte Version des historischen *Ordbog over det danske Sprog* sowie auf Korpusmaterial zugreifen können. Da es ein ständig aktualisiertes Wörterbuch mit der Möglichkeit, auf Korpora zuzugreifen, kombiniert, ist das *ordnet.dk*-Projekt dem heutigen DWDS-Projekt sehr ähnlich (siehe Abbildung 4).

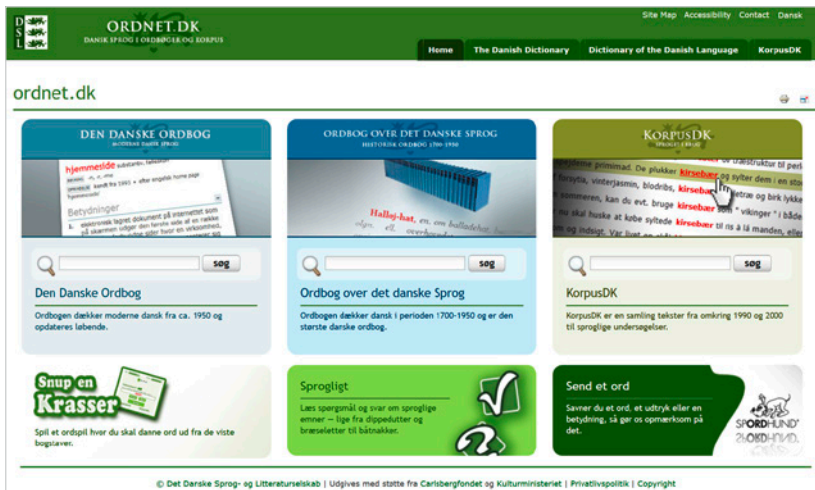


Abb. 4: Startseite des *ordnet.dk*

Die großen historischen Nationalwörterbücher dürften kaum jemals wirtschaftlich profitabel gewesen sein. So wird berichtet, dass das *Oxford English Dictionary*, obwohl es von Hunderttausenden, wenn nicht Millionen genutzt wird, noch nie einen Gewinn für den Verlag Oxford University Press erzielt hat (so Judy Pearsall in einem persönlichen Gespräch). Es versteht sich fast von selbst: Je kleiner ein Markt ist, desto schwieriger ist es, Gewinne zu erzielen. In Skandinavien ist es auch immer so gewesen, dass Wörterbücher mit über etwa 60.000 Stichwörtern nicht rentabel waren; sämtliche der hier vorgestellten Werke wurden deshalb aus öffentlichen Mitteln finanziert, entweder direkt aus der Staatskasse oder aus Stiftungen mit kulturellem oder wissenschaftlichem Auftrag. In diesem Sinne ist die Dokumentation und Beschreibung des Wortschatzes nach wie vor eine kulturpolitische Aufgabe und eine öffentliche Verantwortung.

Ich möchte in diesem Zusammenhang auf ein letztes Projekt hinweisen, das eindeutig zeigt, dass Wörterbucharbeit Kulturpolitik ist. Die färöische Website *sprotin.fo* (vgl. Abbildung 5), die 2018 online ging, wurde öffentlich finanziert, jedoch in Zusammenarbeit mit einem Verlag, der für den Betrieb der Website verantwortlich ist. Hier haben färöische Benutzer Zugriff auf eine ganze Reihe ein- und zweisprachiger Wörterbücher, darunter sehr gute Wörterbücher, denn die Färinger waren sich immer der Rolle der Sprache als nationales Symbol bewusst. So ist z. B. das Deutsch-Färöische Wörterbuch ein außerordentlich umfangreiches Werk mit fast 150.000 Stichwörtern. Es ist beeindruckend und bewundernswert, dass so etwas in einem kleinen Land mit 50.000 Einwohnern, die gleichzeitig auch fast alle färöischen Sprecher ausmachen, überhaupt möglich ist.

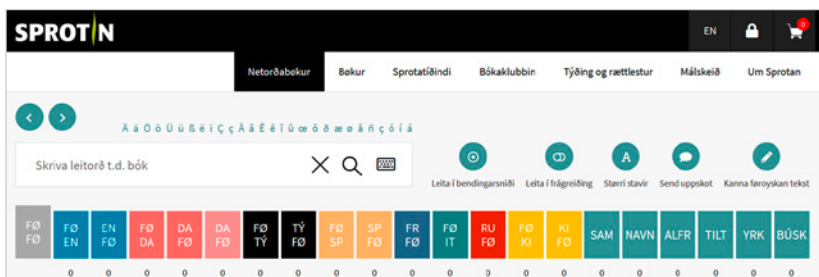


Abb. 5: Startseite des färöischen Wörterbuchportals *sprotin.fo*

Aber zurück zu den großen Nationalwörterbüchern und ihrer Rolle als Kulturträger: Eine der gewaltigsten Veränderungen durch die Digitalisierung, die nicht nur die erwähnten germanischsprachigen Wörterbücher betrifft, sondern

ein allgemeiner Trend in der Vermittlung kultureller Inhalte zu sein scheint, ist die, dass es kommerziellen Verlegern kaum noch möglich ist, mit der Produktion von Wörterbüchern Gewinne zu erzielen. Überall in Europa ist zu sehen, wie traditionelle Wörterbuchverlage ihre Marktposition nicht oder nur unter großen Schwierigkeiten behaupten können; viele müssen ihre Wörterbuchsparte entweder ganz aufgeben oder erheblich reduzieren. Auffällige europäische Beispiele sind *Chambers Harrap* in Großbritannien und *Duden* sowie *Wahrig* in Deutschland; und in Dänemark hatte der Verlag *Politiken* zunächst die Wörterbuchsparte stillgelegt und sie 2018 verkauft. Dies hat allmählich zu der Erkenntnis geführt, dass die öffentliche Hand eine größere Verantwortung übernehmen muss, wenn weiterhin aktuelle und qualitativ hochwertige Wörterbücher verfügbar sein sollen. Für die Wörterbucharbeit der Zukunft sollte dabei auch berücksichtigt werden, dass ein großer Bedarf an strukturierten und formalisierten Sprachdaten für sprachtechnologische Zwecke besteht. Denn Spracherkennung, Sprachsynthese, maschinelle Übersetzung usw. sind nur begrenzt möglich ohne geeignete lexikalisch-semantische Daten. Bei der Erarbeitung der Datengrundlage der Wörterbücher der Zukunft sollte auch dem sprachtechnologischen Aspekt Rechnung getragen werden.

Ich möchte deshalb abschließend noch einmal betonen, wie wichtig es gerade jetzt ist, dass Lexikographen, Computerlinguisten und Informatiker zusammenarbeiten, um lexikographische Daten in Standardformaten zusammenzustellen und diese untereinander und mit Kollegen außerhalb zu teilen. Auch für diese Aufgaben dürfte das ZDL hervorragend geeignet sein, da es schon im Grundsatz als eine Zusammenarbeit zwischen verschiedenen Arbeitsstellen und Mitarbeitern aus unterschiedlichen Disziplinen konzipiert ist.

Ich möchte zuletzt noch auf die EU-Infrastruktur ELEXIS aufmerksam machen. ELEXIS zielt darauf ab, solche Bemühungen zu harmonisieren, Werkzeuge zu entwickeln, die von allen genutzt werden können und dadurch die Kosten und den Zeitaufwand für die Aktualisierung vorhandener oder die Entwicklung neuer Ressourcen zu reduzieren. In der ELEXIS-Initiative werden gemeinsame Standards eingeführt, um inkompatible Daten kompatibel zu machen, mit der Perspektive, vorhandene Ressourcen miteinander zu verknüpfen.

Ich wünsche allen Mitarbeitern und anderen Beteiligten viel Erfolg bei der bevorstehenden Arbeit. Wir vom Projekt *ordnet.dk* werden die Arbeit des ZDL mit großem Interesse verfolgen und ich hoffe, dass unsere beiden Wörterbuchredaktionen auch künftig Erfahrungen, Herausforderungen und Ergebnisse unserer Arbeit miteinander teilen und besprechen können.