

Henning Lobin

EIN CERN DER DEUTSCHEN SPRACHE

ÜBERLEGUNGEN ZU EINEM DOKUMENTATIONSZENTRUM DER DEUTSCHEN SPRACHE

Sprache ist überall. Sie findet sich auf unseren Smartphone-Bildschirmen, in Zeitungen, Zeitschriften und Druckerzeugnissen aller Art, in unseren privaten Notizen und in hunderttausendfach verkauften Romanen. Unsere Umwelt ist von Sprache durchdrungen – Beschriftungen und Beschilderungen reichern Orte mit Namen, Informationen und Verhaltenshinweisen an. Mündlich gebrauchen wir Sprache von „Guten Morgen“ bis „Gute Nacht“, zehntausendfach entweichen die Wörter jeden Tag aufs Neue unseren Mündern, um in noch größerer Zahl in anderer Menschen Ohren zu dringen. Technisch vermittelt trägt gesprochene Sprache auch über große Distanzen, ist heute im Internet leicht speicher- und auffindbar. Digitale Geräte unterstützen uns in unserer sprachlichen Kommunikation und werden zu neuen sprachlichen Akteuren, wenn uns beim Schreiben auf dem Handy Wortvorschläge unterbreitet werden oder wir Sprachassistenten Befehle erteilen. Lesen, schreiben, sprechen und zuhören werden technisiert und bringen neuartige sprachliche Phänomene und Verhaltensweisen hervor.

NEUARTIGE SPRACHLICHE PHÄNOMENE UND VERHALTENSWEISEN DURCH TECHNISIERUNG

Das Hochdeutsche hat sich seit dem Ende des Mittelalters in einem lang andauernden, ungesteuerten Ausgleichsprozess aus den mittel- und oberdeutschen Dialekten entwickelt. Die neuhochdeutsche Schriftsprache, spätestens ab etwa 1800 im Großen und Ganzen in der Gestalt,

wie wir sie heute kennen, hat sich ohne eine staatliche Einheit gewissermaßen von selbst ausgeformt. Bedeutende Autoren, allen voran zunächst Martin Luther, später nach vielen anderen Johann Wolfgang Goethe, hatten großen Anteil daran, aber niemand konnte Sprachgesetze oder Regeln für den gesamten deutschen Sprachraum erlassen und mehr als nur eine kulturelle Wirkung entfalten.

BEDINGT DURCH DIE TRADITION DER DEUTSCHEN SPRACHE GIBT ES KEINE AKADEMIE

Die deutsche Sprache hat keine Akademie, die sie wie die berühmte *Académie française* zu vereinheitlichen und zu pflegen beauftragt ist. Es gibt auch keine Standardgrammatik und kein staatlich anerkanntes Wörterbuch. Der Duden wird von einem privaten Verlag herausgegeben und hat schon lange keinen amtlichen Auftrag mehr. All das kommt nicht von ungefähr, sondern ist durch die Tradition der deutschen Sprache bedingt. Lediglich im Bereich der Rechtschreibung werden heute gemeinsam von Abgesandten der sieben offiziell deutschsprachigen Länder (Deutschland, Österreich, Schweiz, Belgien, Luxemburg, Liechtenstein, Italien/Südtirol) Regeln definiert, die danach durch die jeweils zuständigen staatlichen Stellen für den öffentlichen Schriftverkehr als verbindlich erklärt werden. Seine Geschäftsstelle hat dieser Rechtschreibrat am Leibniz-Institut für Deutsche Sprache in Mannheim (IDS), der zentralen Forschungseinrichtung für die deutsche Sprache.

Wie das Wasser in einem Feuchtbiotop Tiere und Pflanzen umgibt, so umgibt die Sprache seit jeher die Mitglieder der deutschen Sprachgemeinschaft und entwickelt sich mit ihnen und in ihnen. Sie erscheint uns dabei oft so selbstverständlich gegeben, dass wir sie gar nicht wahrnehmen. Wie das biologische Ökosystem ist auch dieses kulturelle Ökosystem dazu in der Lage, sich an Umweltveränderungen anzupassen, und deshalb ständig im Wandel. Das kulturelle Fluidum der Sprache verändert kontinuierlich seine Zusammensetzung und sucht sich neue Strömungswege durch die Ebenen der Kommunikation. Es ist von großer Bedeutung, die Eigenschaften dieses Fluidums zu erforschen, nicht zuletzt, um Veränderungen zu erkennen, die eine Störung des Stoffwechsels anzeigen. Dies ist etwa der Fall, wenn Teile des kulturellen Biotops nicht mehr miteinander verbunden sind und sich die Mitglieder der Sprachgemeinschaft nicht mehr miteinander verständigen können.

Ein neues Bild der Sprache

Das Bild von Sprache, das damit gezeichnet wird, ist eines der Variation und des Wandels. Seitdem mithilfe des Computers große digitale Text- und Sprachkorpora gesammelt und ausgewertet werden können, gibt es in der Sprachwissenschaft die Möglichkeit, derartige Wandlungs- und Differenzierungsprozesse systematisch zu erforschen. Untersuchungen werden auf der Basis umfangreicher Bestände realer Sprachdaten vorgenommen,

Der Autor ist der Direktor des Leibniz-Instituts für Deutsche Sprache, Mannheim.

ohne dabei, wie es früher vorkam, die Vielfalt sprachlicher Phänomene durch eine Auswahl passender Beispiele glätten zu müssen.

GROBE DIGITALE TEXT- UND SPRACHKORPORA MACHEN AUS DER SPRACH- WISSENSCHAFT EINE DATENWISSENSCHAFT

Die Sprachwissenschaft ist in fast allen ihren Bereichen zu einer Datenwissenschaft geworden, in der mit statistischen Mitteln die Feinheiten der sprachlichen Variation in einer Weise betrachtet werden, wie es bis vor wenigen Jahren noch undenkbar erschien. Dies betrifft Variation in unterschiedlichen Dimensionen: geografisch (Dialekte gegenüber der Standardsprache), gesellschaftlich (beispielsweise Jugendsprache, „Kiezdeutsch“ oder Fachsprachen), medial (etwa die Sprache in WhatsApp-Nachrichten) und bezüglich weiterer Dimensionen, in denen ein „Fingerabdruck“ des Sprachgebrauchs in den Datensammlungen zu ermitteln ist. Sprache wird dabei erstmals als das sichtbar, was sie trotz aller Steuerungs- und Bildungsanstrengungen immer gewesen ist: die evolutionär geformte kulturelle Hervorbringung einer Gemeinschaft von Menschen, und zwar einer Gemeinschaft in allen ihren Teilen.

Ins Goldene Zeitalter mit dem Internet?

Man könnte nun meinen, dass mit dem Aufkommen des Internets ein goldenes Zeitalter für die Dokumentation und Erforschung der deutschen Sprache angebrochen sei.

Einerseits stimmt dies auch, weil tatsächlich Texte in riesiger Menge zur Verfügung stehen. Auf der anderen Seite eignen sich diese Texte nur sehr begrenzt dafür, die Vielfalt der Sprache zu erfassen. Beide Punkte verdienen nähere Betrachtung. Zunächst zur Menge: Die Menge an Texten, die im Internet prinzipiell zum Download zur Verfügung steht, ist unüberschaubar. Dies betrifft auch ältere Buch- und Zeitschriftenpublikationen, die inzwischen in großer Zahl digitalisiert vorliegen. Gesprochene Sprache wird in Millionen von Videos auf Youtube und anderen Plattformen dokumentiert – die aktuelle Sprachverwendung, aber auch die historische in Fernseh- und Filmbeiträgen bis zurück zu den Anfängen dieser Medien in erstaunlich umfangreichem Maße.

ZUR ERFORSCHUNG DES TATSÄCHLICHEN SPRACHGEBRAUCHS FEHLEN NOCH TEXTE AUS SCHULE, VERWALTUNG, BERUFLICHEM UND PRIVATEM ALLTAG

Im schriftsprachlichen Bereich handelt es sich allerdings größtenteils um professionell geschriebene, redigierte und lektorierte Texte, die von vornherein für eine Veröffentlichung bestimmt waren. Was fehlt, sind Texte aus Schule, Verwaltung und Wirtschaft, aus dem beruflichen Alltag oder dem privaten Bereich, die alle keiner systematischen

Qualitätskontrolle unterliegen und sich nur an einen kleinen Adressatenkreis richten. Gerade deshalb sind sie so wichtig, wenn der tatsächliche Sprachgebrauch erfasst werden soll. An diese Texte ist keineswegs durch eine einfache Internet-Recherche heranzukommen. Auch spontan, unkontrolliert und nichtöffentlich geäußerte gesprochene Sprache finden wir im Fernsehen oder auf Youtube nicht in nennenswerter Menge. Und die Kommunikation in den sozialen Medien bildet ebenfalls einen der vielen weißen Flecken auf der Landkarte der Sprachdokumentation, weil hier Persönlichkeitsrechte und die Bestimmungen der privaten Betreiber Download, Speicherung und Bereitstellung unterbinden. Die Übertragung all dieser Sprachbelege auf Server mit Langzeitarchivierung ist nämlich für die Dokumentation einer Sprache unerlässlich: Was heute gefunden oder erhoben wird, kann morgen schon wieder verschwunden sein, und deshalb ist die Übernahme der Sprachdaten auf spezielle Forschungsserver so wichtig.

Das Leibniz-Institut für Deutsche Sprache in Mannheim betreibt schon heute solche Server und stellt damit seine Sprachdatensammlungen der Wissenschaft und der Öffentlichkeit zur Verfügung. Doch handelt es sich dabei überwiegend um lizenzierte Zeitungs- und Zeitschriftentexte und nur zu einem geringen Teil um andere Textsorten. Wikipedia-Diskussionen finden sich zwar auf den IDS-Servern und auch einige Blogs mitsamt den jeweiligen Diskussionsbeiträgen, nicht aber größere Bestände von WhatsApp- oder Facebook-Chats, des Mailver-

kehr aus Unternehmen und Behörden oder rein privater Nachrichten. Auch reale Gesprächssituationen werden heute zwar erfasst, aber wegen des enormen Aufwands bei Erhebung und Aufbereitung nur in einem geringen Umfang.

Dokumentation der Vielfalt sprachlicher Kommunikation

Die hier geschilderte Situation muss den Anlass dafür bilden, die deutsche Sprache der Gegenwart endlich auch in den vielen Bereichen systematisch zu dokumentieren, die bislang keine Berücksichtigung gefunden haben. Das kann nur durch ein Dokumentationszentrum der deutschen Sprache geschehen, das in persönlichem Kontakt mit den Mitgliedern der Sprachgemeinschaft um „Sprachspenden“ aus dem Bereich nichtöffentlicher Kommunikation wirbt und diese systematisch erhebt. Die erhobenen Sprachdaten werden der Wissenschaft zur Verfügung gestellt – im Zugriff abgestuft nach den von den Spendenden erteilten Berechtigungen. Nur ein solches Zentrum kann das notwendige Vertrauen bezüglich des rechtmäßigen Umgangs mit diesen sensiblen, oftmals privaten Daten aufbauen.

EINE INTERAKTIVE AUSSTELLUNG ZUR DEUTSCHEN SPRACHE ALS GEMEINSAMES KULTURELLES GUT EINER PLURALISTISCHEN GESELLSCHAFT

Aber wie werden die Menschen dazu gebracht, ein solches Dokumentationszentrum überhaupt auf-

zusuchen? Wie werden sie zu Sprachspenden motiviert und davon überzeugt, dass vertraulich und rechtlich einwandfrei mit diesen umgegangen wird? Eine attraktive Ausstellung zur deutschen Sprache ist es, die wir den Menschen bieten wollen und die sie veranlasst, zum Dokumentationszentrum zu kommen. In dieser Ausstellung soll es um Themen gehen, die in der Öffentlichkeit erfahrungsgemäß besonders großes Interesse hervorrufen, etwa Sprachwandel, der deutsche Wortschatz, das Deutsche in Europa und der Welt, Sprachen in Deutschland, Mehrsprachigkeit, politischer Sprachgebrauch und Propaganda, digitalisierte Kommunikation oder die Entstehung der deutschen Standardsprache und ihr Verhältnis zu Dialekten und anderen Varietäten. All das wird auf der Grundlage realer Sprachdaten dynamisch visualisiert und eröffnet vielfältige Möglichkeiten für einen interaktiven Umgang mit den Exponaten.

In einer solchen Ausstellung, die sich verbindet mit der Erhebung von Sprachdaten, werden die Besucherinnen und Besucher zu aktiv Beteiligten, denn sie selbst bringen den Ausstellungsgegenstand in das Dokumentationszentrum mit. Dadurch wird ein solches Zentrum zu einem einzigartigen Ort, an dem sich die deutsche Sprachgemeinschaft manifestiert. Darüber hinaus soll es den Besucherinnen und Besuchern ermöglicht werden, nicht nur durch ihre Sprachspenden am wissenschaftlichen Erkenntnisprozess mitzuwirken, sondern auch im Rahmen von *Citizen Science*-Projekten selbst sprachforschend tätig zu werden, beispielsweise im Be-

reich der Dokumentation von Dialekten. Anders als in einem „normalen“ Museum ist es dazu erforderlich, neben sprachbezogenen Informationen auch das Verständnis grundlegender sprachwissenschaftlicher Untersuchungsmethoden, etwa die Auswertung von Texten, die Analyse von Gesprächen oder die Durchführung linguistischer Befragungen, in der Ausstellung aufzubauen.

Die deutsche Sprache muss in diesem Dokumentationszentrum als das gemeinsame kulturelle Gut einer pluralistischen Gesellschaft dargestellt werden. Die gegenwärtig und auf kaum absehbare Zeit stark durch Migration geprägten deutschsprachigen Länder benötigen ein Bild von Sprache, das nicht exkludierend wirkt, sondern Wandel und Vielfalt der Sprache als etwas ansieht, das seit jeher in die deutsche Sprache eingeschrieben ist. Das Konzept „Sprache“ wird in der öffentlichen Diskussion oftmals als eine Metapher für die innergesellschaftliche Kommunikation überhaupt genutzt. Das Bild der Sprache, das sich in der Öffentlichkeit verbreitet, ist somit von entscheidender Bedeutung, wenn es gilt, eine geeignete Auffassung dieser gesellschaftlichen Kommunikation zu entwickeln, eine Auffassung, die inkludierend wirkt, gerade weil sie sprachwissenschaftlich fundiert ist.

Das CERN der deutschen Sprache in Mannheim

Ein solches Dokumentationszentrum der deutschen Sprache, das eine Ausstellung, eine Infrastruktur

zur Datenerhebung, Forschungs- und Veranstaltungsmöglichkeiten umfasst, wird gegenwärtig in Mannheim geplant. Wir stellen uns dabei nichts weniger vor als ein CERN der deutschen Sprache, das zu einem Zentrum der datenorientierten Sprachforschung werden soll. Das physikalische Großforschungszentrum CERN in Genf kann als der Inbegriff einer wissenschaftlichen Institution angesehen werden, in der Daten erzeugt, erhoben und erforscht werden. Mittels riesiger Teilchen-Beschleuniger wird die Materie dazu gebracht, ihre Eigenschaften preiszugeben. Auf einem ausgedehnten Campus genau an der Grenze der Schweiz zu Frankreich werden die Datenbestände von internationalen Forschungsteams analysiert und zu Theorien über unsere materielle Welt verdichtet.

EIN MITTELEUROPÄISCHES ZENTRUM DER DATEN- ORIENTIERTEN ERFOR- SCHUNG DER DEUTSCHEN SPRACHE IN MANNHEIM

In einem CERN der deutschen Sprache werden die Menschen durch eine Ausstellung dazu gebracht, ihre Daten zu spenden, damit sprachwissenschaftliche Forschungsgruppen diese analysieren und daraus Theorien über unsere kulturelle Welt ableiten können. Ein solches Dokumentationszentrum der deutschen Sprache muss ein Ort sein für alle, die sich für die deutsche Sprache interessieren oder sie sprechen, wo auch immer sie leben. Dieses „kulturelle CERN“ muss

als ein mitteleuropäisches Zentrum der deutschen Sprache verstanden werden, in dem das Deutsch aller sieben deutschsprachigen Länder und aller weiteren Länder, in denen es als Regional- oder Minderheitensprache anerkannt ist, berücksichtigt wird. Deutschland als das mit Abstand größte Land des deutschsprachigen Raums sollte sich dem in besonderer Weise verpflichtet fühlen.

Das Leibniz-Institut für Deutsche Sprache, Teil der von Bund und allen 16 Bundesländern finanzierten Leibniz-Gemeinschaft, hat vom Land Baden-Württemberg Mittel erhalten, mit denen gegenwärtig ein Masterplan für die inhaltliche und organisatorische Umsetzung erstellt wird. In der Stadt Mannheim gibt es seit geraumer Zeit Überlegungen, für dieses Vorhaben ein repräsentatives Grundstück am Rande der Innenstadt direkt am Neckar bereitzustellen. Mit einer namhaften Stiftung aus der Region wurden vielversprechende Vorgespräche geführt hinsichtlich der Errichtung eines Gebäudes, in dem die Ausstellung und die anderen Funktionsbereiche realisiert werden könnten. Über die Leibniz-Gemeinschaft werden Gelder beantragt, mit denen die wissenschaftliche Seite des Vorhabens gemeinsam von Bund und Ländern finanziert werden sollen. Weitere Finanzzusagen werden notwendig sein, um den an die Öffentlichkeit gerichteten Teil – Wechselausstellungen, Veranstaltungen, Vermittlung, Betrieb – zu verwirklichen. Trotzdem ist der Zeitpunkt gekommen, mit der Realisierung dieses Vorhabens nun tatsächlich zu beginnen.

In Mannheim besteht aber nicht nur wegen des Interesses und der Möglichkeiten relevanter Akteure – Leibniz-Institut, Stadt, Land und Stiftung – die einmalige Situation, ein solches, weltweit einzigartiges Zentrum für eine bedeutende Kultursprache zu realisieren. Mannheim, seit jeher eine internationale Stadt, nicht als eine Hauptstadt mit nationalem Anspruch versehen, unweit der Wiege der neuhochdeutschen Schriftsprache an der Grenze des mitteldeutschen zum oberdeutschen Sprachraum gelegen, mit hervorragenden Verkehrsverbindungen in alle deutschsprachigen Regionen, ist ein idealer Ort, um jetzt dort ein solches Dokumentationszentrum zu verwirklichen.

Ein Dokumentationszentrum der deutschen Sprache kann vielleicht einmal ein Ort werden, an dem die Sprache zu sich selbst kommt, weil sich die Sprachgemeinschaft dort für ihre Sprache zusammenfindet. Dabei muss es uns für den deutschsprachigen Raum um die Erfassung der Vielfalt und des Wandels unserer Sprache gehen, aus dem schon immer ihre Leistungsfähigkeit erwuchs. Und damit kann ein solches Dokumentationszentrum als das sprachliche Abbild der Pluralität unserer freiheitlichen Gesellschaft verstanden werden, die sich immer wieder auf eine Vorstellung dessen, was das Gemeinsame ausmacht, zu einigen hat. ■