



Zur Statistik des deutschen Wortschatzes

Reinhard Güll



Reinhard Güll ist Büroleiter der Abteilung „Informationsdienste, Öffentlichkeitswesen, sozial- und regionalwissenschaftliche Analysen“ im Statistischen Landesamt Baden-Württemberg.

Die Darstellung und Analyse statistischer Ergebnisse bedürfen einer sprachlich-textlichen Kommentierung. Auch wenn es zutreffen mag, dass „ein Bild mehr sagt als 1 000 Worte“, kommt man in dem Bestreben, „Zahlen sprechen zu lassen“ nicht ohne Worte aus. Der Wortschatz einer Sprache ist keine statische Größe, sondern er unterliegt einem permanenten Wandel. Daher fällt es schwer, den Umfang des Wortschatzes einer Sprache auch statistisch zu ermitteln, zumal eine Statistik des deutschen Wortschatzes nicht zu den Aufgaben der amtlichen Statistik in Deutschland gehört. Derzeit ist gerade der deutsche Wortschatz im Vergleich zu anderen Sprachen in einer starken Umbruchsituation – dies als Folge etwa der Internationalisierung von Kommunikationsinstrumenten oder auch eines wachsenden Bevölkerungsanteils mit Migrationshintergrund.

Was ist ein Wortschatz?

Sprachwissenschaftler weisen dem Begriff Wortschatz zwei unterschiedliche Bedeutungen zu. Zum einem ist darunter die Gesamtheit aller Wörter einer Sprache zu einem bestimmten Zeitpunkt zu verstehen, zum anderen wird damit die Gesamtheit aller Wörter bezeichnet, die eine einzelne Person kennt oder verwendet. In feinerer Untergliederung beim Wortschatz einer Person wird differenziert zwischen dem rezeptiven und dem produktiven Wortschatz. Zum rezeptiven oder passiven Wortschatz einer Person gehören all die Wörter einer Sprache, die sie vom Sinn her richtig versteht. Der aktive oder produktive Wortschatz eines Menschen wird vom ihm auch beim Sprechen oder Schreiben genutzt, indem er damit sinnvolle und verständliche sprachliche Äußerungen formuliert.

Langfristig gesehen zeigen sich sprachgeschichtliche Entwicklungen in der Veränderung des Wortschatzes. Im Laufe eines Jahrhunderts entstehen bei einer lebenden Sprache viele neue Wörter, andere sterben ab und viele werden verändert. Dies ist die Hauptschwierigkeit, den Wortschatz einer Sprache zu einem bestimmten Zeitpunkt buchhalterisch korrekt zu quantifizieren.

Historische Vorläufer

Als in der zweiten Hälfte des 19. Jahrhunderts die Germanistik an deutschen Universitäten fest etabliert war und eine erste Blütezeit erlebte, haben sich auch die Statistiker der Vorgängerinstitutionen der heutigen statistischen Ämter mit der Materie des deutschen Wortschatzes beschäftigt. In der Zeitschrift des Königlich Preußischen Statistischen Bureaus des Jahres 1896 erschien ein umfangreicher Beitrag zur Statistik des deutschen Wortschatzes.¹ In diesem historischen Beitrag geht es weniger um die statistische Ermittlung des Wortschatzumfangs, vielmehr wurde hier eine philologische Methode zur Quantifizierung von Worthäufigkeiten praxisorientiert vorgestellt und zur Verwendung empfohlen. Die Autoren sahen den künftigen Nutzen ihrer Arbeit vor allem bei den Stenografen und Druckereien. Die Ergebnisse boten Möglichkeiten zur Verbesserung der bestehenden Stenografiesysteme und zur effizienteren Ausstattung der Setzkästen in Druckereien.

Bereits in der Ausgabe 1884 der württembergischen Jahrbücher für Statistik und Landeskunde des Königlich Statistisch-Topographischen Bureaus erschien ein Artikel von *Hermann Fischer* „Über den schwäbischen Dialekt und die schwäbische Dialektdichtung“²

In diesem Artikel wird auf eine sehr komplizierte Materie eingegangen, denn deutsche Dialekte werden im Regelfall nur in der mündlichen Kommunikation und nur selten in gedruckter Form verwendet. Inhaltlich wird schwerpunktmäßig die geografische Ausbreitung der schwäbischen Dialektformen analysiert. Da Dialekte zumeist nur mündlich verbreitet und überliefert werden, ist es für diese Sonderformen der deutschen Sprache sehr viel schwieriger als bei der Hochsprache, Quantitäten des Dialektwortschatzes zu ermitteln. Sowohl die badischen als auch die schwäbischen Dialekte kommen nicht in einer einheitlichen Sprachform vor, sie bestehen vielmehr aus einer Vielzahl von geografisch eng begrenzten und genutzten Sprachvarianten. So gibt es im Sprachraum der beiden badischen Regierungsbezirke Karlsruhe und

1 Siehe Statistisches Bundesamt: *Wirtschaft und Statistik* 8/2007, Seiten 797 ff.

2 Siehe Fischer, Hermann: *Über den schwäbischen Dialekt und die schwäbische Dialektdichtung*. In: *Württembergische Jahrbücher für Statistik und Landeskunde* des Königlich Statistisch-Topographischen Bureaus Jahrgang 1884. Stuttgart 1885, Seiten 56 ff.

WÜRTEMBERGISCHE JAHRBÜCHER

FÜR

STATISTIK UND LANDESKUNDE.

HERAUSGEGEBEN

VON DEM

K. STATISTISCH-TOPOGRAPHISCHEN BUREAU.

JAHRGANG 1884.



STUTTGART.

W. KOHLHAMMER.

1885.

Freiburg folgende Dialektgruppen: Südfränkisch, Kurpfälzisch, Hochalemannisch, Niederalemannisch, Schwäbisch und Ostfränkisch. Das Niederalemannische wird oftmals noch unterteilt in das Oberrheinalemannische und das Seealemannische. Im mittleren und südöstlichen Teil von Baden-Württemberg, hauptsächlich in den Regierungsbezirken Stuttgart und Tübingen, wird dagegen Schwäbisch mit seinen vielen Unterformen gesprochen. Das Schwäbische gehört zu den alemannischen Dialekten und somit zu der oberdeutschen

Dialektgruppe. Alleine schon diese kleine Aufzählung zeigt, dass es nahezu unmöglich ist, Näherungswerte des Wortschatzumfangs für die in Baden-Württemberg gesprochenen Dialekte zu ermitteln.

Der Wortschatzumfang

Nach führenden Sprachwissenschaftlern der deutschen Sprache umfasst der Wortschatz der deutschen Standardsprache etwa 75 000 Wör-

ter. Je nach Quelle und Zählweise wird die derzeitige Gesamtgröße des deutschen Wortschatzes auf ca. 350 000 bis 500 000 Wörter geschätzt. Das für die deutsche Sprache wohl aufwändigste und anspruchvollste Nachschlagewerk, das Deutsche Wörterbuch, das von den Brüdern *Jacob* und *Wilhelm Grimm* begonnen und erst in der zweiten Hälfte des letzten Jahrhunderts von der nachfolgenden Germanistengeneration fertig gestellt wurde, enthält annähernd 350 000 Stichwörter der deutschen Sprache. In der aktuellen Ausgabe des Duden, dem deutschen Universalwörterbuch, wird der Wortschatz der deutschen Alltagssprache auf 500 000 Wörter geschätzt und der zentrale Wortschatz auf rund 70 000 Wörter. Die Ausgabe 2008 des anderen namhaften deutschen Wörterbuchs von *Wahrig* enthält rund 260 000 Stichwörter. Diese Richtgrößen zeigen, auf welchen Umfang der deutsche Wortschatz mindestens geschätzt werden muss. In all diesen Wörterbüchern gibt es nur geringe Anteile der vielen Fachwortschätze. So gibt es realistische Schätzungen, dass alleine

die deutsche Fachsprache der Chemie einen Umfang von 20 Mill. Begriffen beinhaltet. Dies zeigt, dass der Gesamtwortschatz der deutschen Sprache mit der Größe $500\,000 + n$ zu bezeichnen ist, wobei n eine unbestimmte Größe bleibt. Was den individuellen Wortschatz einer Person anbelangt, kann aufgrund seriöser sprachwissenschaftlicher Untersuchungen davon ausgegangen werden, dass mit zunehmendem Bildungsstand auch der Wortschatz eines Menschen zunimmt. Schätzungen gehen hier von einer Breite von 3 000 bis hin zu 200 000 Wörtern je nach Bildungsstand einer Person aus. Bei *Goethe*, dem wohl berühmtesten deutschen Dichter, konnte auf Grund einer Auszählung in seinem Gesamtwerk ein Wert von gut 90 000 aktiv gebrauchten Wörtern ermittelt werden. ■

Weitere Auskünfte erteilt
Reinhard Güll, Telefon 0711/641-20 08,
Reinhard.Guell@stala.bwl.de

Buchbesprechung

Rebecca Bihler



Peter Huber
Visionär in Zeiten des
Umbruchs
Johannes Mährlen:
Ökonom, Sozialreformer
und Freund Möhrkes
395 Seiten, 38 Abbildungen
Hohenheim Verlag, Stuttgart
2008, 22,00 Euro
978-3-89850-161-3

Visionär in Zeiten des Umbruchs – Johannes Mährlen

Johannes Mährlen, ein Freund Möhrkes, wurde 1803 in Ulm geboren und gilt als eine der treibenden Kräfte der Industrialisierung in der biografischen Abhandlung wird er unter anderem als Grundsteinleger des heutigen Statistischen Landesamtes bezeichnet. 1855 wird Mährlen erster Hauptgeschäftsführer der neugegründeten „Handels- und Gewerbekammer Stuttgart“ und verfasst dort als angesehener Statistiker der Königlichen Centralstelle für Gewerbe ab 1858 den jährlichen Bericht. Im Zuge dessen legt der Professor für Nationalökonomie und Gewerbestatistik laut Huber die Grundlagen für Gewerbestatistik von Württemberg und auch für das spätere Statistische Landesamt.

Als Mährlen 1860 in das Statistisch Topographische Institut eintritt, erscheinen die Jahresberichte der Handels- und Gewerbekammern in Württemberg unter seinem Einfluss in veränderter Form. Mährlen, als einziger Volkswirt unter den Kammerleuten, redigiert die Jahresberichte, fertigt den Extrakt aus den Berichten der einzelnen Kammern und erstellt eine eigene statistische Arbeit, den „Statistischen Anhang“, in der Spezialfragen behandelt werden. Dort findet sich 1861 der erste Energiereport für Württemberg in Form einer Auflistung aller Flüsse Württembergs. Zudem gibt Mährlen

einen guten Einblick in die finanzielle Situation der Staatshaushalte der deutschen Länder dieser Zeit. Im Jahr 1863 beteiligt er sich außerdem an einer umfangreichen Beschreibung des Königreichs Württemberg. Damit sind Mährlens Arbeiten einzigartige Dokumente des technischen, wirtschaftlichen und gesellschaftlichen Umbruchs in der Mitte des 19. Jahrhunderts. Drei Jahrzehnte nach dem Tod Mährlens im Jahr 1871 wurde die wirtschaftsstatistischen Forschungen, der Kern seiner reform der Handelskammer-Berichterstattung, eingestellt.

Die Historie des Statistischen Landesamtes beginnt jedoch schon vor Mährlens Zeit in der Zentralstelle für Statistik. Bereits 1820 wird ein Königliches Dekret zur Einrichtung eines statistischen Bureaus in Württemberg erlassen, welches als unentbehrliche Informationsquelle für Gesetzgebung und Verwaltung dienen soll. Dieses Bureau wird 1856 in Zentralstelle für Landesstatistik umbenannt und soll gewerbestatistische Aufnahmen und Viehzählungen durchführen. Unter Mährlens Einfluss wird 1860 das Aufgabenspektrum um die Bearbeitung der internationalen Finanzstatistik erweitert. Durch eine Bekanntmachung des Innenministeriums Baden-Württemberg entstand dann schließlich 1953 aus vier Statistischen Landesämtern mit Sitz in Stuttgart, Tübingen, Karlsruhe und Freiburg ein Statistisches Landesamt für das gesamte Bundesland mit Sitz in Stuttgart. ■