

WAS VERBIRGT SICH HINTER LEDA?

Wer bei Leda nur an die Dame denkt, der sich Zeus in Schwanengestalt genähert hat, kann jetzt dazulernen. LEDA ist auch das Akronym für die LEXikographische DATenbank, an der am Institut für deutsche Sprache seit nunmehr knapp zwei Jahren eifrig gearbeitet wird und die die Arbeit der Wörterbuchmacher grundlegend revolutionieren soll.

Große Wörterbuchvorhaben brauchen bekanntlich viel Zeit und Geld. Zwischen dem ersten und dem letzten Band des Grimmschen Wörterbuchs lagen hundert Jahre. Generationen von Germanisten haben ihren Schweiß vergossen, während sich der Wortschatz rasant vergrößerte, ohne daß mit seiner Beschreibung auch nur annähernd Schritt gehalten werden konnte.

Zeit und Geld sind knapp in unseren Tagen. Wörterbuchmacher müssen schon deshalb im großen Umfang auf vorhandene Wörterbücher zurückgreifen. Fehler werden übertragen, neue Wörter und Bedeutungen fehlen, Unterschiede zwischen sinnverwandten Wörtern (etwa *Kummer*, *Gram*) werden nicht erfaßt, und sei es nur, weil der Buchstabe G einen anderen Bearbeiter hatte als der Buchstabe K. Lexikologen wie der Heidelberger Germanist Herbert E. Wiegand und der Erlanger Romanist Franz Josef Hausmann haben die Unzulänglichkeiten der deutschen Wörterbücher zu Recht bemängelt. An Entwürfen für bessere Wörterbücher fehlt es nicht.

Grundlage jeder Wortbeschreibung müssen Originalbelege aus Texten sein. Natürlich können treffende Definitionen aus vorhandenen Wörterbüchern durchaus die eigene Arbeit erleichtern. Ganz wichtig ist jedoch, daß in jedem Bearbeitungszustand jedes Mitglied eines Wörterbuchteams auf alles Vorhandene zugreifen kann. Nur so läßt sich erreichen, daß vergleichbare Wörter ähnlich beschrieben und die assoziativen und hierarchischen Beziehungen zwischen den Wörtern dargestellt werden.

Diesen Forderungen entspricht LEDA durch ihren Aufbau in den drei Komponenten *Textdatei*, *Worddatei* und *Arbeitsdatei*. Die *Textdatei* enthält in maschinenlesbarer Form Texte aller Art: Romane, Zeitungen, Sachbücher, gesprochene Sprache usw. (zur Zeit knapp 10 Mio. Wörter; der Ausbau auf 50 Mio. Wörter ist in Arbeit). Sucht der

Lexikograph Belege zu einem Wort oder einer Wortkombination, erhält er in Sekundenschnelle an seinem Bildschirm alle Sätze, in denen sie vorkommen, dazu statistische Angaben über Häufigkeit und Verteilung auf Textsorten. Was ihn davon interessiert, kann er sich über seinen Arbeitsplatzdrucker am Schreibtisch ausdrucken lassen. Geeignete Beispielsätze kann er samt Kommentar in die Arbeitsdatei kopieren. Textdatei und ein interaktives Zugriffsprogramm sind bereits im Einsatz.

In einer *Worddatei* sollten die wichtigen der bereits vorhandenen Wörterbücher abgespeichert sein, auf die der Lexikograph bei der Abfassung seiner Wortartikel über Bildschirm zugreifen könnte. Da die Verlage aus urheberrechtlichen Gründen die Abspeicherung ihrer Wörterbücher (noch) nicht zulassen, bleibt diese Komponente von LEDA vorerst zurückgestellt.

Mit der Programmierung der *Arbeitsdatei* ist inzwischen begonnen worden. Zu dem Aufbau dieser Datenbankkomponente gehört auch die Entwicklung von Programmen, mit denen der Bearbeiter jederzeit die Konsistenz überprüfen kann, indem er kontrolliert, daß jedes Wort, das er zur Beschreibung eines anderen Wortes verwendet, auch selbst wieder erklärt wird. Auch Ver-

weise auf sinnverwandte und übergeordnete Begriffe lassen sich beispielsweise mit diesem Instrument automatisch erzeugen. Nach Arbeitsabschluß liefern dann Druckaufbereitungsprogramme die fertigen Setzbänder für den Druck. Der Einsatz einer ersten Version der Arbeitsdatei ist bereits für 1986 vorgesehen.

LEDA macht den Lexikographen nicht überflüssig; im Gegenteil, sie gestattet ihm, sich auf den schöpferischen Teil seiner Arbeit zu konzentrieren, indem sie ihn von mechanischen Such- und Prüfvorgängen entlastet. Mit LEDA rückt eine umfassende und zugleich aktuelle Bestandsaufnahme unseres deutschen Wortschatzes einschließlich der unzähligen Fach- und Sondersprachen, das schon seit Jahren geforderte »Große Interdisziplinäre Wörterbuch des Deutschen«, endlich in den Bereich des Machbaren.

Zwei Wörterbuchprojekte des Instituts für deutsche Sprache sind Pilotanwender von LEDA. Schon haben andere Institute ihr Interesse daran bekundet, u. a. für ein Wörterbuch der Byzantinistik und eines zu den Schriften von Karl Marx. LEDA: nicht mehr nur ein Name aus der griechischen Mythologie.

Wolfgang Teubert