

**COMPUTERLINGUISTIK ALS EIN ERGEBNISS
DER GEGENWÄRTIGEN ÄNDERUNGEN
IN DER LINGUISTIK DER GANZEN WELT**

In unserem Artikel zeigen wir die Sprachforschungen nicht nur bestimmter Sprache (Deutsch, Ukrainisch, Englisch usw.), sondern digitale Sprache, die im globalen Weltnetz existiert.

Heutige Sprache ist eng mit der Technik verbunden und das gibt neue Möglichkeiten für Sprach-, Text- und zunehmend integrierte Informationstechnologie. Das gilt nicht nur im praktischen Leben, wo Sprache neben ihren herkömmlichen Verwendungsweisen immer mehr auch im Verbund mit anderen, vor allem technisierten Informations- und Kommunikationsmitteln verwendet wird (z.B. in Navigations-Systemen für den Straßenverkehr, in Online-Zeitungen und multimedialen Lernumgebungen), sondern auch für die wissenschaftliche Untersuchung von Sprache und ihren Verwendungen. Alle diese fachspezifischen Anwendungen der Sprache erweitern und verändern unsere Wissen von Sprache und fordern von der Erforschungen der Linguistik zusammen mit der Zeit zu ändern. Der Umfang dieses Artikels ist zu eng, um in vollem Masse das Problem dieser Änderungen zu zeigen, aber es ist offensichtlich, dass Linguistik ist heutzutage mehr international, als niemals in der Geschichte der Menschheit.

Wissen über Phonetik und Phonologie eignet sich vorzüglich für multimediale Darstellung und unterstützt Sprachsynthese. Bei morphologischer Analyse helfen maschinelle Systeme. Auf die fundamentale Erweiterung lexikographischer Arbeitsmöglichkeiten durch Hypertextsysteme wurde oben schon hingewiesen. Für den Alltagsgebrauch stehen im Internet über 800 verschiedenartige Lexika von unterschiedlichster Qualität und Zweckbestimmung zur Verfügung, und zwar für circa 170 Sprachen. Oft handelt es sich um nicht mehr als praktikable Online-Versionen gedruckter Wörterbücher oder um ungedruckte einfache (ein- oder mehrsprachige) Wörtersammlungen ggf. mit einfachen Erläuterungen. Eine interessante Anwendung ist die Zusammenführung mehrerer gedruckter Wörterbücher im elektronischen Medium [2]. Ein besonders geeignetes Anwendungsfeld sind Wörterbücher, die sämtliche Arten lexikographischen Wissens in einem einzigen umfassenden und permanent erweiterbaren lexikalischen Informationssystem ausgesprochen benutzerfreundlich zusammenführen und dabei vielfältigen 'linguistischen Mehrwert' erzielen können [8].

In dem Maße, wie umfassende Sprachdaten zur Verfügung stehen, könnten sich auch linguistische Forschungsinteressen und damit Schwerpunkte der Theoriebildung verlagern, beispielsweise weg von intuitiv begründeten Kompetenz- hin zu beobachtungsgestützten Performanzmodellen in der Grammatiktheorie.

Erst in jüngster Zeit beginnt man, die weit darüber hinausgehenden Möglichkeiten hypermedialer Wissensrepräsentation intensiver auszunutzen [3]. Im Aufbau befinden sich zum Beispiel ein deutsches Wortschatz-Lexikon in Leipzig [4] und ein lexikalisch-lexikologisches, korpusbasiertes Such- und Informationssystem am Institut für deutsche Sprache in Mannheim [5]. Alles führt selbstständig dazu, (1) dass Grammatik wird leichter erlernen und sogar kann man wenig sie wissen (das macht Grammatikkontrollprogramm), (2) dass Phonetikkenntnisse brauchen wir nicht so oft (Wir sprechen nicht, wir schreiben), (3) dass gedruckte Wörterbücher gebraucht man jedes Jahr weniger und weniger. Ausserdem sagen wir kein Wort über die Beschleunigung des Entlehnungseindringens in verschiedene Sprachen.

Spracheinsatz konzentrierte sich zunächst in zwei Bereichen, die ohne Computer gar nicht entstanden wären. Das sind «Linguistische Datenverarbeitung» und «Computerlinguistik». «Linguistische Datenverarbeitung» versammelte sämtliche Techniken zur maschinellen Be- und Verarbeitung sprachlicher Daten zunächst eher im Sinne einer Hilfswissenschaft, bevor sich unter dem Titel «Computerlinguistik» allmählich auch computerabhängige Theoriebildungen entwickelten letzten Endes mit dem Ziel, menschliche Sprachfähigkeit oder Teile davon auf elektronischen Rechnern nachbilden zu können. Im Internet stehen eine Reihe interaktiver Online-Demonstrationen computerlinguistischer Systeme zur allgemeinen Erprobung bereit. Das theoretisch aufwendigste wie praktisch anspruchsvollste Fernziel computerlinguistischer Anwendungen jedoch, die maschinelle Übersetzung, konnte angesichts der Komplexität natürlicher Sprachsysteme und der unendlichen Vielfalt ihrer Verwendungen bisher allenfalls für vergleichsweise einfache Texte und auch dort nur in Ansätzen befriedigend gelöst werden. Ein Gang durch verschiedene sprachwissenschaftliche Teildisziplinen mag einen Eindruck von der Vielfalt produktiver Verwendungen computergestützter Arbeitstechniken in der Linguistik geben.

Digitale Information leisten hier ungleich bessere Dienste als herkömmliche. Seit etwa 1990 liegen geschriebene Texte aller Art oft ohnehin schon als digitale Quellen vor und das forderte die Änderungen in Texttypologie. Lineare Texte, Textdesignstrategien und nichtsprachlichen Zeichen füllen eine Fläche, auf der sich zunächst nichts bewegt. So entsteht der Begriff «Hypertext». Die Organisationsmethode der grossen interaktiven Informationskomplexe, die aus verschiedenartigen Umgebungen (Text, Textdesign, Seitenlayout, Grafik, Foto, Ton, Animation) besteht, wird in den Multimediasystemen Hypertext genannt. Vom linguistischen Aussichts, der Begriff «Hypertext» bedeutet die Organisationsmethode solcher Textinformation, die als ein Netz, dient, dessen Knoten korrelativ zu bestimmten Textfragmenten sind; und Bogen verbinden diese Knoten und symbolisieren vielfältige unlineare Tiefverbindungen [1: 68]. Wandernd durch solchen Verbindungen, gibt es Möglichkeit, dass Material in jeder beliebige Aufeinanderfolge zu beobachten, neue Textstrukturen, gleichzeitig verschiedene Informationsfragmente gegreifend und vergleichend, zu bilden.

Schon ein erster Blick in Textcorpora lässt ahnen, in welcher Weise neue Techniken es erlauben, die Menge zuverlässiger Sprachdaten erstens zu vervielfachen, zweitens anspruchsvoller aufzubereiten und drittens universeller zugänglich zu machen. Dadurch gewinnt Corpuslinguistik [5] einen neuen Stellenwert innerhalb der Sprachwissenschaft. Introspektion als empirische Quelle wird (und darf) nicht verschwinden, wohl aber gegenüber der Beweiskraft umfangreicher Textbelege an Bedeutung verlieren.

Hypertextuelle und multimediale Wissensspeicher schließlich erweitern linguistische Arbeits-, Darstellungs- und in der Folge vielleicht auch Denkweisen von der vorwiegend zweidimensional (linear) schriftgebundenen hin zu dreidimensionalen (netzwerkartigen) semiotischen Universalformen. Allen sprachwissenschaftlichen Teildisziplinen werden sich hier neue Dimensionen eröffnen [8].

Alle neuen Arbeitsmittel und -techniken ermöglichen erheblich umfangreichere, solider durchgearbeitete und leichter kontrollierbare empirische Grundlagen für sprachwissenschaftliche Forschung, als sie mit herkömmlichen Mitteln erreicht werden konnten. Dadurch sind sie geeignet, sprachwissenschaftliche Interessenschwerpunkte zu verlagern und Fragestellungen zu erweitern.

Literatur

1. Штерн І,Б. Вибрані топіки та лексикон сучасної лінгвістики: Енциклопідичний словник для фахівців з теоретичних гуманітарних дисциплін та гуманітарної інформатики / Міжнародний фонд «Відродження»ю - К. АрТЕК, 1998.

2.<http://www.uni-trier.de/uni/fb2/germanistik/aedph/I.exer-Projekt.htm>

3.<http://aspra9.informatik.uni-leipzig.de/wort/inhalt.htm>

4.Fraas/Haß-Zumkehr 1998; <http://w^vw.ids-mannheim.de/lekses/>

5.Schlobinski Peter. Empirische Sprachwissenschaft. Opladen: Westdeutscher Verlag

6.Leitner Gerhard. New Directions in English Language Corpora . Methodology, Results, Software Developments. (= Topics in English Linguistics 9). Berlin, New York: Mouton de Gruyter. - 1992.

7.Fraas Claudia, Haß-Zumkehr Ulrike. Vom Wörterbuch zum lexikalischen Informationssystem. LEXX1S - ein neues Projekt des Instituts für deutsche Sprache. In: Deutsche Sprache 26, S. 289-303. - 1998.

8.Storrer Angelika, Harriehausen Bettina. Hypermedia für Lexikon und Grammatik. Tübingen: Narr. - 1998.