

"JAZYKO-ZNANIE" IN MOSKAU

Reinhard Köhler, Universität Trier

In Moskau gibt es mehrere computerlinguistisch einschlägige Gruppen, von denen z.B. die Gruppe "Smysl \leftrightarrow Tekst" (Melcuk, Apresjan) seit Jahren international hohes Ansehen genießt, während andere bei uns noch wenig bekannt sind, so z.B. das Institut für Russische Sprache (über das noch zu berichten sein wird). Zunehmende Bedeutung kommt auch dem *Jazyko-Znanie*¹ genannten Zentrum zu, dem der vorliegende Bericht gewidmet ist.

Seit 1980 arbeitet die Forschungsgruppe *Theoretische und angewandte Lexikologie*, ein Zweig des *Labors für angewandte Linguistik* der Philologischen Fakultät der Moskauer Universität, an theoretischen linguistischen Problemen und an der elektronischen Verarbeitung natürlicher Sprache.

Seit 1990 wird diese Arbeit in größerem Rahmen organisatorisch als *Internationales Forschungszentrum Jazyko-Znanie* unter der Leitung von A. A. Polikarpov weitergeführt. Diese Organisation wiederum ist einer der Träger von *Gumanitarnoe Znanie* (etwa: Menschliche Kognition).

Zwischen dieser Gruppe und dem Fach LDV in Trier bestehen inzwischen intensive Kontakte; so wurde im vergangenen September ein Vertrag über Kooperation und Austausch (vor allem im Bereich quantitative und synergetische Linguistik) zwischen der Philologischen Fakultät der Moskauer und dem Fachbereich II der Universität Trier geschlossen. Der Autor hatte

1991 und 1992 Gelegenheit² zu Besuchen und Gesprächen mit den dortigen Kollegen.

Jazyko-Znanie ist zur Zeit mit etwa zwölf Mitarbeitern (Linguisten, Programmierern und Wissenstechnikern) ausgestattet. Darüber hinaus koordiniert und integriert die Einrichtung fallweise die Arbeit externer Wissenschaftler (Linguisten, Mathematiker, Psychologen, Musikologen und KI-Forscher) und Institute, darunter die Fakultät für Mathematische Informatik und Kybernetik der Moskauer Universität, *das Institut für Russische Sprache* der (vormals sowjetischen) Akademie der Wissenschaften, die Institute für Linguistik der Ukrainischen Akademie der Wissenschaften, und der Universität Tartu und das Staatliche Konservatorium Tiflis.

Jazyko-Znanie gliedert sich in mehrere problemorientierte Arbeitsgruppen und unterhält zwei Labors - eins für Computerlinguistik und eines für 'systemische Linguistik'. Publikationsorgan ist die Reihe "Kvantitativnaja lingvistika i avtomaticheskij analiz tekstov" (Quantitative Linguistik und automatische Textanalyse), Universitätsverlag Tartu, herausgegeben von J. A. Tuldava.

Hauptaktivitäten der Arbeitsgruppen sind zur Zeit:

=> Modellbildung für den Bereich Evolution der natürlichen Sprache(n) und für Kommunikationsprozesse auf der Grundlage kognitiver Mechanismen und

¹ nicht übersetzbares Wortspiel, ungefähr 'Sprach- Wissen[schaft]'

² Mit Unterstützung durch die Deutsche Forschungsgemeinschaft

Randbedingungen für die Lösung von Polysemieproblemen und die semantische Interpretation auf der lexikalischen, morphemischen und syntaktischen Ebene.

- => Theoretische Analyse grundsätzlicher Fragen zur Organisation des Vokabulars und Besonderheiten der Repräsentation in verschiedenen Arten von Wörterbüchern.
- => Entwicklung von Algorithmen und von Software für die automatische Analyse (morphologisch, syntaktisch und lexikalisch-semantisch) von Texten in russischer und englischer Sprache.
- => Vergleichende Studien von Russisch und Englisch auf allen linguistischen Analyseebenen;
- => Modellierung des bilingualen Sprecher/Hörers; Untersuchungen zu Regularitäten von Interferenzen bei der Fremdsprachenverwendung, Entwicklung einer Fehlertypologie.
- => Entwicklung eines Moduls für idiomatische Ausdrücke für Übersetzungssysteme.
- => Erstellung von Softwarewerkzeugen für die maschinengestützte Lexikographie.
- => Aufbau von Textkorpora und lexikalischen Datenbanken.

Zu den konkreten Projekten gehören u. a. maschinenoperable Wörterbücher der Antonyme und der Synonyme des Russischen und weitere Spezialwörterbücher, eines Englisch-Russischen und Russisch-Englischen Übersetzungswörterbuchs (je ca. 100000 Einträge) mit Zugriffs- und Pflegesoftware, der "Thesaurus der russischen Sprache", lexikalische Datenbanken für weitere Sprachen (u. a. Chinesisch und Ukrainisch) und Software für verschiedene Zwecke, z.B. eine Volltext-Datenbank, ein elektronischer Redaktionsassistent zur Überprüfung von Orthographie, Vokabular, Grammatik, Stil und Interpunktion, ein Statistik-Paket für die quantitative Textanalyse und Module für Lehr- und Lernsysteme.

Die Forschungseinrichtung beteiligt sich laufend an internationalen Projekten (z.B. dem Bochum-Trierer Projekt zur synergetischen Linguistik) und Konferenzen (so das erste internationale Kolloquium über synergetische Linguistik, das von Jazyko-Znanie im September 92 ausgerichtet wurde). Die zweite internationale Konferenz zur quantitativen Linguistik, QUALICO-94, wird von dieser Organisation, zusammen mit der Philologischen Fakultät der Moskauer Universität, organisiert werden.