

Einleitung

Rita Nübel (IAI Saarbrücken)
Uta Seewald-Heeg (Hochschule Anhalt)

Maschinelle Übersetzungswerkzeuge haben in Übersetzungsunternehmen, vor allem im Bereich der Lokalisierung, inzwischen einen festen Platz erhalten. Auch freiberuflich arbeitende Übersetzer setzen sich zunehmend mit neuen Technologien auseinander, die eine Effizienzsteigerung ihrer Übersetzungsleistung versprechen. In diesem Zusammenhang rücken vor allem satzspeicherbasierte Systeme, sogenannte Translation-Memory-Systeme, in das Zentrum des Interesses; denn wie die Ergebnisse einer 1998 vom Arbeitskreis „Maschinelle Übersetzung“ der GLDV durchgeführten Evaluierung der linguistischen Performanz vollautomatischer maschineller Übersetzungssysteme nahelegen [Nübel/Seewald-Heeg98], muss beim Einsatz maschineller Übersetzungssysteme für eine qualitativ hochwertige Übersetzung von Dokumenten mit einer relativ aufwendigen Postedition gerechnet werden, die den Nutzen der maschinellen Übersetzung (MÜ) vielfach relativiert, zum Teil sogar in das Gegenteil verkehrt. Die Frage nach den Leistungsmerkmalen von Translation Memories (TM) ist daher sowohl bei zahlreichen industriellen Nutzern als auch bei potentiellen Anwendern von zentralem Interesse, zumal die Produktpalette bei diesem Systemtyp in den vergangenen Jahren ähnlich wie im Bereich der MÜ-Produkte deutlich angewachsen ist.

Für den Arbeitskreis „Maschinelle Übersetzung“ war das Anlass, sich 1999 mit TM-Systemen auseinanderzusetzen und Möglichkeiten zu erörtern, die Leistungsfähigkeit von TMs für den professionellen Einsatz zu bewerten.

Anders als bei vollautomatischen Übersetzungssystemen basiert die TM-Technologie auf der Wiederverwendung bereits übersetzter Textsegmente, die im Verlauf vorausgehender Übersetzungsprozesse als Satzpaare gespeichert werden. Da TMs ohne Daten geliefert werden, sind sie ausschließlich für professionelle Nutzer geeignet, bei deren Übersetzungstätigkeit Texte mit einem gewissen Wiederholungsanteil anfallen. Eine Evaluierung der Leistungsfähigkeit von TMs muss daher die Trefferrate berücksichtigen, mit der Textsegmente im bereits vorhandenen Material (Referenzmaterials) aufgefunden werden. Um dies zu bewerkstelligen, sind geeignete Testdaten und Bewertungsmethoden erforderlich. Linguistisch orientierte Qualitätskriterien, die für die Bewertung der Übersetzungsqualität von MÜ-Systemen verwendet werden, wie z.B. Grammatikalität, Lesbarkeit,

Stil oder Verständlichkeit, sind hier nicht von Bedeutung, da TMs selbst keine Übersetzung erzeugen. Für die Bewertung von TMs ist vielmehr zentral, inwieweit sie aus einer möglichen Vielfalt ähnlicher, bereits archivierter Übersetzungen die besten bzw. „ähnlichsten“ Kandidaten auswählen, die als Basis für die jeweils aktuelle Übersetzung eines Satzes dienen können. In diesem Zusammenhang ist zunächst der Begriff der Ähnlichkeit zu klären, den auch Reinke in seinem Artikel zur Evaluierung der linguistischen Leistungsfähigkeit von Translation-Memory-Systemen problematisiert.

In einem ersten Treffen des Arbeitskreises „Maschinelle Übersetzung“ tauschten Teilnehmer aus der Übersetzerbranche sowie aus Forschung und Entwicklung ihre in unterschiedlichen Zusammenhängen gesammelten Erfahrungen mit TM-Systemen aus. Aufgrund des hier bekundeten Interesses, die Performanz von TMs vor allem auch hinsichtlich ihrer Retrieval-Leistung im Rahmen eines unabhängigen Gremiums wie des GLDV-Arbeitskreises zu untersuchen, fand auf der GLDV-Jahrestagung 1999 in Frankfurt ein Workshop statt, auf dem TM-Systeme aus Nutzer- und Entwicklerperspektive zunächst vorgestellt und Ideen für eine Evaluierung diskutiert wurden.¹

Aufgrund der unterschiedlichen Provenienz der Workshop-Teilnehmer ist es gelungen, mit dem vorliegenden Band verschiedene Facetten der Beschäftigung mit TM-Technologie und ihre gegenseitigen Berührungspunkte zu präsentieren:

Im ersten Beitrag werden von H. Elsen die typischen Aufgaben eines im Bereich der Lokalisierung tätigen Übersetzungsunternehmens skizziert und die technologiespezifischen und organisatorischen Anforderungen aus Management-sicht beschrieben. Der Beitrag verdeutlicht, dass ohne den professionellen Einsatz von Sprachtechnologiewerkzeugen wie TMs heute kaum ein Übersetzungsdienstleister in der Lage ist, wirtschaftlich erfolgreich zu arbeiten. Da durch den Einsatz von TMs die Produktivitätssteigerung unterstützt werden soll, haben potentielle Nutzer sowohl ein großes Interesse an der Transparenz von Funktionalitäten solcher Werkzeuge als auch an verlässlichen Aussagen zur Systemperformanz. Diese beiden Fragen bestimmten schwerpunktmäßig auch die Beiträge des Workshops und der anschließenden Diskussion (siehe „Ausblick“) und werden in den nachfolgend kurz vorgestellten Beiträgen dieses Bandes aus unterschiedlichen Perspektiven diskutiert.

U. Seewald-Heeg und R. Nübel betrachten in ihrem gemeinsamen Beitrag zwei Systeme, die die Funktionalität der vollautomatischen Übersetzung mit jener der Satzarchivierung (im Sinne einer einfachen TM-Funktionalität) verbinden.² Der

Beitrag geht zunächst der Frage der Leistungsfähigkeit der Satzarchive beim Abgleich von Sätzen mit dem archivierten Referenzmaterial nach und mündet in die Frage, inwieweit sich die Satzarchive der untersuchten Systeme als Erweiterung des Wörterbuchs für komplexe lexikalische Strukturen eignen.

Die beiden nachfolgenden Beiträge stellen die zentralen Funktionsweisen zweier auf dem Markt verbreiteter TM-Systeme vor: Der Beitrag von H. Bohn illustriert die wichtigsten Funktionen der Translator's Workbench von Trados, und J. Klein beschreibt die Funktionsweise von Transit, dem TM-Werkzeug der Firma STAR.

Erfahrungen mit verschiedenen TM-Systemen (TranslationManager von IBM und Translator's Workbench von Trados) aus Sicht eines professionellen Anwenders dokumentiert A. Pesch in ihrem Beitrag.

U. Reinke diskutiert in seinem ersten Beitrag Ideen für eine Optimierung von TMs durch eine intelligente Verwendung bereits vorhandener Wissensquellen wie paralleler Korpora und Terminologie. K. Lee et al. stellen in ihrem Beitrag das Forschungssystem MIRAC vor, mit dem multilinguale Dokumente (semi-)automatisch aligniert und auf ihre Konsistenz auf Termebene sowie auf einer semantisch abstrakteren Ebene überprüft werden.

Die Evaluation von TM-Systemen ist Gegenstand des zweiten Beitrags von U. Reinke, der im Zusammenhang mit der Frage nach der Bewertung der linguistischen Leistungsfähigkeit von TMs zum einen die Verwendbarkeit von Methoden, die für die Evaluation von *Information Retrieval* Systemen angewendet werden, erörtert, zum anderen darstellt, wie bestimmte Eingabemodifikationen die Trefferraten der TMs beeinflussen und schließlich beispielhaft illustriert, wie die Berücksichtigung linguistischer Eigenschaften beim Abgleich von Sätzen mit dem Referenzmaterial die Retrievalleistung verbessern kann.

Der „Ausblick“ am Ende fasst die Abschlussdiskussion zusammen, in der Ideen und Anforderungen an eine TM-Evaluation erörtert und in ein Evaluierungsszenario eingearbeitet wurden, das die folgenden für die Evaluierung erforderlichen Schritte präzisiert.

Literatur

[Nübel/Seewald-Heeg98] Nübel, R. und Seewald-Heeg, U. (Hrsg.) (1998): *Evaluation of the Linguistic Performance of Machine Translation Systems*. Proceedings of the KONVENS-98 Workshop in Bonn. St. Augustin: Gardez! Verlag.

ANMERKUNGEN

- ¹ Präsentationen einzelner Beiträge können über die Programmübersicht des Workshops unter folgendem URL eingesehen werden:
<http://www.heeg.de/~uta/AK-Programm-Jahrestagung.htm>.
- ² Personal Translator 2000 Office plus von Linguattec und IBM sowie Langenscheidts T1, eine Entwicklung aus dem Haus Lernout & Hauspie.