

Ich bin ein Berliner - I'm a doughnut - Je suis un beignet

Kurt Eberle
Lingenio GmbH
Hebelstr. 14
D-69115 Heidelberg
Germany

5.11.2004

Übersetzungsprobleme und ihre Lösung in einem lexikalistischen Transfer-Modell auf der Basis flacher semantischer Repräsentationen

Mehrdeutigkeit ist ein zentrales Problem für die maschinelle Übersetzung wie das Titel-Beispiel zeigt. Dass es dort viel wahrscheinlicher ist, dass *Berliner* in der Bedeutung *Person* und nicht als *Gebäck* zu verstehen ist (und somit als *Berliner* oder *Berlinois* zu übersetzen ist), resultiert dabei aus dem Wissen, dass in prädikativen Konstruktionen von Eigenschaften die Rede ist, die für das Subjekt Sinn machen. Gute maschinelle Übersetzung setzt also Berücksichtigung der Bedeutung der Satzteile und Kenntnis der inhaltlichen Bezüge zwischen diesen Teilen im Quelltext voraus. Allerdings ist eine sorgfältige semantische Interpretation nur für solche Mehrdeutigkeiten notwendig, für die die Zielsprache keine äquivalenten mehrdeutigen Wörter und Strukturen vorsieht.

In dem Vortrag wird ein Transfer-Modell vorgestellt, das versucht, dieser Erkenntnis ökonomisch Rechnung zu tragen, indem es die Übersetzungsbeziehung für die Ebene flacher semantischer Repräsentation der Satzteile definiert und tiefere semantische Interpretation dieser Teile und ihrer Bezüge untereinander nur dort vornimmt, wo es ein aktuelles Übersetzungsproblem gibt. Ein alltäglicher Fall ist etwa die Übersetzung eines Pronomens. Wenn in dem Text '*Enfin il trouva la balle. Elle était dégonflée.*' das Pronomen richtig übersetzt werden soll, muss man wissen, worauf es sich bezieht. Hier bezieht es sich auf *balle*, was *Kugel* bedeuten kann oder *Ball*. Abhängig von dieser Entscheidung ist dann *elle* mit *sie* oder *er* zu übersetzen. Die prädikative Kennzeichnung aus dem zweiten Satz (*ohne Luft*) gibt genügend Information, um auf *Ball* und *er* zu schließen. Im Vortrag wird gezeigt, wie das beschriebene Modell solche Bezüge erkennen und verwalten kann.

In dem Transfer-Modell wird die Übersetzungsaufgabe in der Regel auf der Ebene des Lexikons gelöst. Dort werden tiefe semantische Repräsentationen durch kontextuelle Bedingungen der Form *wenn A gegeben ist, dann hat das Wort die Bedeutung B* ausgelöst und nahezu alle wichtigen Wort-abhängigen Umstrukturierungen werden dort beschrieben. Das ist von großem Vorteil unter kommerziellem Aspekt, aber auch unter theoretischem. Beides wird versucht, in dem Vortrag nachzuweisen, letzteres durch einen Vergleich mit alternativen, aus der Literatur bekannten Transfer-Ansätzen.