

Könnten Sie einen Termin um sechzehn Uhr reinschieben?" fragt Özlem Senay in ihr Mikrofon und blickt erwartungsvoll auf den Computermonitor. Gelbe Kästchen mit kryptischen Inschriften wie „Prosodie-Analyse“ beginnen zu blinken. Dann öffnet sich auf dem Schirm ein Fenster mit dem Schriftzug „ahm k'onnen sie einen Termin um sechzehn Uhr reinschieben“. Gleichzeitig plärrt eine Roboterstimme die Frage – nunmehr stark elektronisch verfremdet – in den noch leeren Konferenzsaal im Haus 4 der Daimler-Benz AG in Stuttgart. Einige Rechensekunden später ist das Werk vollbracht. Das „System“ intoniert, jetzt mit weniger Akzent: „Can you squeeze in an appointment at four o'clock?“

„Wunderbar, schön“, freut sich der Kameramann des Süddeutschen Rundfunks. Zweimal schon hatten das Fernsehteam und Özlem Senay, frischgebackene Informatikerin vom Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI), vergeblich versucht, den Computer zur Übersetzung des Terminwunsches zu überreden. Beim ersten Mal reklamierte der Rechner: „Bitte

sprechen Sie etwas lauter!“ Beim zweiten Mal hatte er statt „reinschieben“ merkwürdigerweise „und ihnen“ verstanden, was er mit dem Kommentar quittierte, er könne bei der Analyse „keinen Satz finden“. Dabei hatte Özlem Senay ihre Frage stets klar und akkurat wie eine Nachrichtensprecherin vorgetragen.

„Verbomobil“ nennt sich das ehrgeizige Projekt, entstanden in enger Zusammenarbeit von sieben großen Industrieunternehmen, neunzehn deutschen Universitäten, zwei amerikanischen Forschungseinrichtungen und dem DFKI. Über hundert Forscher, darunter Computerlinguisten, Sprachwissenschaftler, Informatiker und Spezialisten der „Künstlichen Intelligenz“ (KI), haben, auf 21 Standorte verteilt, an der Software des Systems gestrickt. Um die Programmierung der 25 „Sprachverarbeitungsmodulare“ perfekt aufeinander abzustimmen, pflegten sie regen Datenaustausch über das Internet.

Ziel des kooperativen Kraftaktes ist ein kommerzielles Produkt: ein wie ein Laptop tragbarer Dolmetschelfer, der vom Deutschen und vom Japanischen ins Englische übersetzt und sich mit gesprochenen Kommandos steuern läßt. „Wir denken auch an eine Erweiterung von Funktelefonen“, sagt Wolfgang Wahlster, DFKI-Direktor und wissenschaftlicher Leiter des Projekts: „Benutzer könnten einen Rechner, auf dem Verbomobil läuft, per Funk anrufen und so bei internationalen Gesprächen Übersetzungshilfe erhalten, wenn sie mit ihrem Englisch am Ende sind.“ Zudem sei ein Dialog zwischen einem Amerikaner, einem Deutschen und einem Japaner denkbar, bei dem jeder in seiner Landessprache in sein Handy spreche.

Wahlster ist trotz der enttäuschend schlechten Performance von Verbomobil zuversichtlich: „Derzeit befindet sich das System in einem Forschungsstadium. Wir brauchen sicherlich noch etwa fünf Jahre, um es zur Prototypenstufe zu entwickeln.“ Solche Sätze hat das Publikum freilich schon oft aus dem Munde der KI-Forscher vernommen.

Gemeinsam mit einigen Teamkollegen präsentierte Wahlster, die hörende Übersetzungsmaschine am vergangenen Freitag bei Projektpartner Daimler-Benz der Presse. So recht mobil ist das gegenwärtige „Experimental-System“ noch nicht. Auf dem Tisch steht eine ausladende „Space“-Workstation – ein handelsüblicher Arbeitsplatzrechner mit zwei schnellen Spezialprozessoren, einem 224 Millionen Byte großem Speicher (224 MB RAM) und einer ein Gigabyte (1000 MB) fassenden Festplatte.

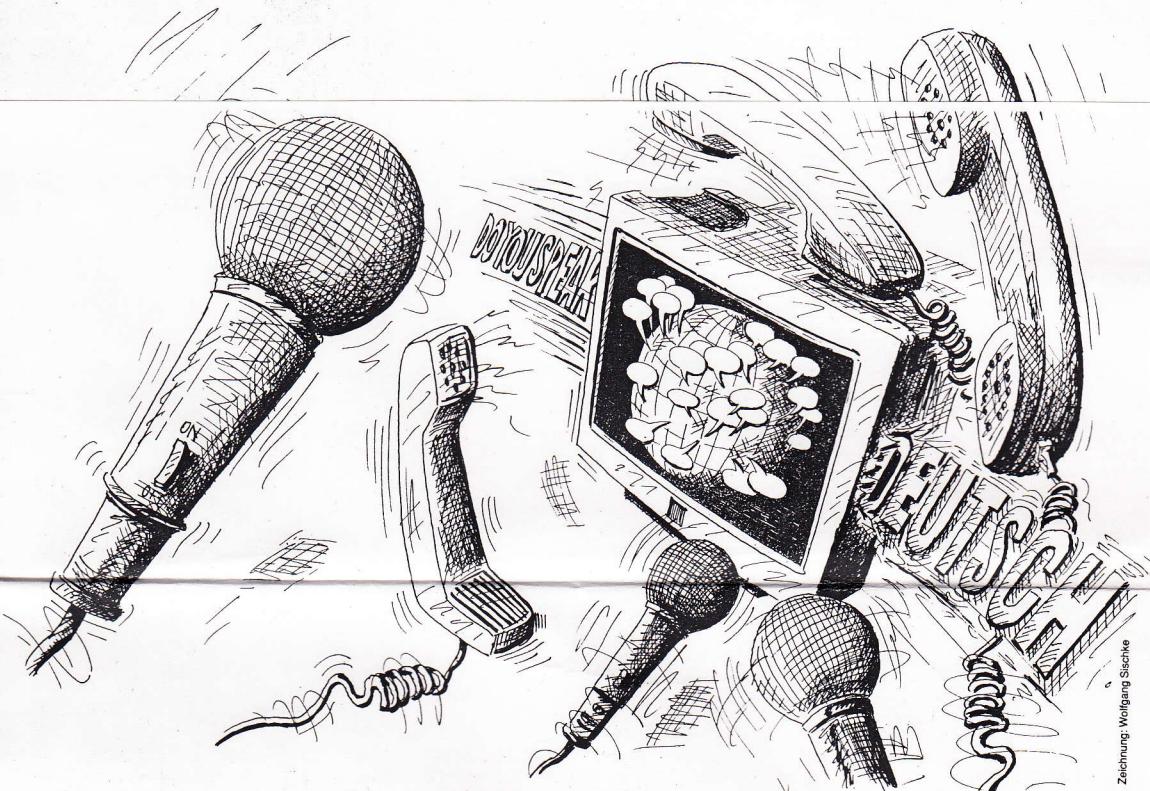
Verbal ist Verbomobil auch noch nicht auf der Höhe. Bei der offiziellen Vorführung am Nachmittag klappt so gut wie nichts. Projektkoordinator Reinhard Karger vom DFKI ist hörbar nervös. Immer wieder spricht er – unterhalb von klackernden Kameraverschlüssen – den Satz: „Ich habe einen Termin beim Zahnarzt um vierzehn Uhr dreißig“ ins Mikrofon. Doch gerade jetzt, wo es darauf ankommt, stellt Verbomobil sich störisch und versteht nur Bahnhof. Sogar im „Buchstabier-Modus“ – Karger spricht Letter für Letter einzeln aus – ist der Apparat schwer von Begriff. „Da ist wohl ein Kabel draußen oder ein Wackelkontakt“, entschuldigt er das Versagen. „Außerdem spreche ich jetzt übertrieben laut, das System ist für Bürolautstärke ausgelegt.“

Dennoch sind die bislang erreichten Fortschritte beachtlich. Denn das Erkennen und Über-

# Kannitverstan aus dem Computer

Europas größtes Projekt der „Künstlichen Intelligenz“ lud zur Präsentation.

Das Verbomobil soll Sprache verstehen und übersetzen können / Von Claus-Peter Sesin



setzen von Sprache ist für Computer seit je eines der schwierigsten Probleme. Vor zehn Jahren noch verstanden die Spracherkennung meist nur einzelne, jeweils mit deutlichen Sprechpausen getrennte Wörter – und selbst dies nur, wenn sie in mühseligem Training an Aussprache und Dialekt des Benutzers gewöhnt worden waren. Deutlich schwieriger schon ist die Aufgabe sogenannter Verbundwörterkennung, die aus einem kontinuierlichen Strom von Lauten die einzelnen Wörter herausfiltern. Sie erfordern eine peinlich präzise und grammatikalisch korrekte Aussprache.

Verbomobil nun versucht sich in der vertracktesten aller Varianten – dem Verstehen sogenannter

**Wenn der Saarländer „Kirsche“ sagt, meint er oft die Kirche. Rechner verwirrt so was**

Spontansprache, die so rauskommt, wie dem „User“ der Schnabel gewachsen ist: Selbst während des Sprechens korrigierte Wörter und umgestellte „ungrammatische“ Sätze sollen es – so die Wunschvorgabe – nicht aus dem Konzept bringen, ebensowenig die vielen „Ähs“ und „Hms“, Schmatzer und Schnalzer, Geräusper und Gehüstel. Zudem arbeitet es bereits heute „sprecherunabhängig“, also ohne vorheriges Training. Diese hochkomplizierte Aufgabe kann es nur dank extremer Beschränkungen bewältigen: Der Wortschatz umfaßt derzeit lediglich 1200 Wörter, und

inhaltlich müssen alle Dialoge um das Thema „Terminabsprache im Büro“ kreisen.

Zunächst werden die vom Mikrofon kommenden Sprachsignale in für den Computer verständlichen Digital-Code umgewandelt. Nun stellt der Rechner Hypothesen auf, welche Wörter gesprochen sein könnten, und analysiert mögliche Sätze nach grammatischen und semantischen Regeln. Macht der gefundene Satz im Kontext des Dialogs Sinn, wird er ins Englische übersetzt und von einem Sprachsynthesizer über Lautsprecher ausgegeben – Frauenstimmen hoch, Männerstimmen tief. Anderenfalls und bei Störungen gibt es eine gesprochene Fehlermeldung.

Damit Verbomobil jeden beliebigen Sprecher möglichst auf Anhieb versteht, wurde es zuvor mit Umengen unterschiedlichster Sprechproben trainiert. „Denn ein Hamburger spricht das Wort ‚Stein‘ anders als ein Münchner“, weiß Wahlster. Außerdem benötigt das System einprogrammiertes Hintergrundwissen. Nur damit kann es aus dem Kontext heraus schließen, ob etwa ein Saarländer, der das Wort „Kirche“ ähnlich wie „Kirsche“ ausspricht, das Gotteshaus oder das Steinobst meint.

Auch beim Übersetzen geht es ohne solches Wissen nicht. Sonst könnte der Computer den Satz „Unser Chef wohnt in einem Schloß“ fälschlicherweise mit „Our boss lives in a lock“ übersetzen – statt korrekt mit „Our boss lives in a castle“. Jedem Menschen ist sofort klar, daß ein Vorhängeschloß schon aus Platznot ausscheidet. Dem Rechner muß dies mühsam mit Wenn-dann-Regeln eingetrichtert werden.

Zum Leidwesen der Informatiker und KI-Forscher haben sich Computer beim „Lernen“ sol-

chen Wissens bislang enttäuschend dumm angestellt. Ob die „Wissensverarbeitung“ in Zukunft besser läuft und Computer eines Tages sogar den Sinn des Gesprochenen verstehen können? „Prinzipiell gibt es kein Naturgesetz, das uns sagt, so etwas gehe nicht“, meint Wahlster. „Aber ich bin überzeugt, daß es schon eine Grenze gibt. Sprache setzt auch einen kulturellen Erfahrungshorizont voraus. Die Feinheiten der Sprache in Literatur und Poesie werden sich einem Computer nie erschließen. Aber damit beschäftigen wir uns ehrlich gesagt auch nicht. Wir sind ja im Bereich der Fach- und Techniksprache tätig – und froh, wenn wir unseren Job dort gut machen. Die anderen Dinge wollen wir auch in Zukunft dem Menschen überlassen.“

Finanziell ist das Verbomobil-Projekt vorerst abgesichert. Für die erste, 1996 auslaufende Testphase hat das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) insgesamt 66,4 Millionen Mark an Fördermitteln bewilligt. Das Projekt, meint Wahlster, sei nicht nur wissenschaftlich, sondern auch politisch und marktstrategisch von großer Bedeutung – und er findet auf Anhieb das klassische Argument für Forschungsförderung: „Deutsch ist heute in der Europäischen Union die am meisten gesprochene Sprache. Wir können in dieser Technologie nicht darauf warten, daß wir von den Amerikanern oder Japanern, die ja in der Sprachverarbeitung traditionell sehr stark waren, Produkte geliefert bekommen. Wo es um unsere Muttersprache geht, sollten wir auf eigenen Füßen stehen. Die Kompetenz dafür liegt nun mal in Deutschland und in deutschsprachigen Ländern, und das sollten wir uns nicht abnehmen lassen.“