

Der Computer versteht nichts, er rechnet nur

Schimpft der Kollege wieder mit seinem PC, ist man versucht zu glauben, die beiden verstünden sich. Spracherkennungsprogramme funktionieren jedoch nur dann gut, wenn ihr Anwender sie trainiert hat. Sprecherunabhängige Systeme sind noch fehleranfälliger, besonders bei Menschen mit fremdem Akzent.

Stefan Schaden bekommt oft Besuch von Studierenden. Gleich ob sie aus Italien, Spanien, England, Frankreich oder Deutschland kommen, der Doktorand am Institut für Kommunikationsakustik der Ruhr-Universität Bochum setzt jeden seiner Gäste vor ein Aufzeichnungsgarät und lässt ihn lesen: Städtenamen und kurze Sätze.

„Ambüroh“, haucht die Französin. „Munshengladback“, umlaftet der Engländer. Und der in der Bestellung von Café Macchiato geschulte Germane spricht weltgewandt: „Ottschiobello“.

Der Sprach- und Kommunikationswissenschaftler arbeitet im Auftrag der Deutschen Forschungsgemeinschaft an der Entwicklung eines Systems zur automatischen Spracherkennung. Etwa für

Ticketautomaten, Navigationssysteme, Fahrplankünfte. „Natürlichsprachige Dialogsysteme“ heißt das im Fachjargon.

„Kein glücklicher Begriff“, meint der 36-Jährige. Denn eigentlich gehe es nur um das Erkennen und die künstliche Nachbildung von Lauten. Von „verstehen“ und „sprechen“ könne bei einem Computer nicht die Rede sein. „Ein gesprochenes Satz ist für ein Erkennungssystem nur eine bestimmte Folge von Lauten“, so Schaden. Das darin repräsentierte Wissen - der Inhalt - sei für den Rechner irrelevant. Er vergleiche lediglich die Lautfolgen mit jenen, die in der Datenbank abgespeichert sind.

Findet er eine Ähnlichkeit mit einem der durch viele Sprechproben ermittelten Durchschnittsmuster, gibt der Computer jene Lautfolge zur „Antwort“, die der Programmierer mit dem erkannten Klangmuster verknüpft hat: „Der Zug nach Kiel fährt um sechsuhrsieben ab Gleis drei.“ Schaden: „Die Kunst ist, die Toleranz des Programms richtig einzustellen.“ Sonst „verstehe“ der Computer entweder kaum etwas oder „deute“ beispielsweise ein undeutliches „bestimmt nicht“ flott in einen „Imbiss“ oder gar zur „Stiftung“ um.

„Sprecherunabhängige Sys-



Der Bochumer Sprachwissenschaftler Stefan Schaden trainiert seinen Computer. Auch wenn die Französin „Ambüroh“ haucht, soll er Hamburg verstehen und Antwort geben. **WAZ-Bild: hac**

teme allein auf individuelle Sprechweisen und regionale Akzente abzustimmen, ist schon eine komplexe Aufgabe“, weiß Schaden.

Noch aufwändiger ist sein Projekt mit den diversen Fremdsprachen. „Dabei versuchen wir bei den fremdsprachlichen Probanden Regularitäten in den Abweichungen gegenüber der Aussprache eines Standardsprechers ausfindig zu machen und die Ergebnisse

technisch zu operationalisieren.“

Das werden die Sprechproben in die computerlesbare

Lautschrift „SAMPÄ“ übersetzt. Aus der Niederschrift leitet Schaden Regeln ab. Diese integriert er in das vorhandene Erkennungsprogramm,

das so nach und nach in die Lage versetzt wird, fehlerhaft ausgesprochene Lautkombinationen anhand ihrer charakteristischen Abweichun-

gen zu erkennen und auf andere Worte zu übertragen.

Mit zwei studentischen Hilfskräften muss der Doktorand das Projekt stemmen. Wissend, dass Geld und Zeit zu kaum mehr reichen werden, als ein Verfahren zu testen, das sich hoffentlich später als Standard durchsetzen wird. Bis dahin ist es noch ein weiter Weg. Auf dem der Hobbymusiker immer wieder erklären muss, dass Spracherkennungssysteme nicht besser arbeiten, wenn man gaanz laangsaam spricht. Noch, dass es hilft sie anzuschreiben. Denn dann verstehen sie nur Bahnhof - statt Bangkok.

Mirko Hackmann



wissenschaft@waz.de