

WIA BRING' MA DA KÜNSTLICHEN INTELLIGENZ STEIRISCH BEI?

Dialekt trifft auf Daten. Wie Forscherin Barbara Schuppler von der TU Graz der künstlichen Intelligenz heimische Dialekte beibringt – und dabei neue Wege in Medizin, Inklusion und Dialektpflege aufzeigt.

Text: Ulrich Dunst | Fotos: Jürgen Fuchs

Diese Entwicklung der TU Graz „gschlaunt“ ziemlich gut. Es gibt Dialektwörter, die man kaum übersetzen kann. „Gschlaun“ zum Beispiel. Südsteirisch. Bedeutet so viel wie: Etwas gelingt mit Leichtigkeit und dazu noch schnell. Für Barbara Schuppler ist das nicht nur ein Lieblingswort, sondern auch ein Beispiel für sprachliche Schatzkisten, die Maschinen bislang nicht ausreichend öffnen können.

Die Forscherin am Institut für Signalverarbeitung und Sprachkommunikation der TU Graz arbeitet mit daran, das zu ändern. In einem mehrjährigen Projekt verbesserte sie gemeinsam mit einem interdisziplinären Team (gemeinsam mit der Linguistik der Uni Graz) bestehende Spracherkennungssysteme, damit diese auch mit Dialekt zurechtkommen. Kurz: Sie bringen der KI Steirisch bei. Klingt nach einer kuriosen Spielerei – ist aber ein technologischer Meilenstein mit enormem Einsatzpotenzial.

Der Impuls kam aus der Praxis. Barbara Schuppler und ihr Team haben untersucht, wie Menschen private Gespräche gestalten, bei denen keine fremden Personen zuhören. Wie sie sich ins Wort fallen, wie sie gemeinsam Bedeutungen formen, wie sie Sätze abbrechen. Doch beim Transkribieren der spontanen Gespräche stieß sie auf ein Problem: Die gängigen Spracherkennungsprogramme kapitulierten reihenweise. „Die Systeme sind auf standardisierte, mediennahe Sprache trainiert – Dialekte kommen da nicht wirklich vor“, sagt sie. Also ging man einen Schritt weiter.

Was Dialekte schwierig macht, ist nicht ihre Komplexität, sondern ihre Unsichtbarkeit in den digitalen Trainingsdaten. „Es ist nicht so, dass Dialekt für Spracherkennungssysteme weniger verständlich als Englisch wäre“, erklärt Schuppler. Aber Dialekte kommen in Fernsehnachrichten, Filmen, Podcasts kaum vor – also wurden die Spracherkennungssysteme in ihrer Trainingszeit fast nur mit Wörtern der Hochsprache angeleitet. Zum Wortschatz hinzu kommt noch die Aussprache: „Hom“, „hobn“, „ham“ statt „haben“ – wer das nicht kennt, versteht nur Bahnhof.

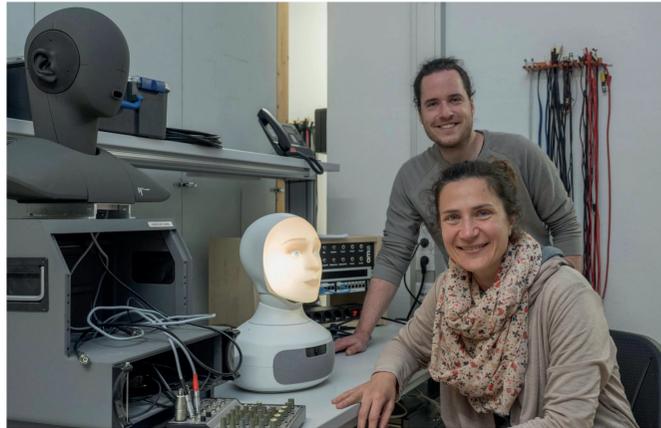
Noch schwieriger wird es, wenn schnell gesprochen wird. „Wir haben



Gespräche analysiert, in denen Sätze im Schnitt eine Sekunde lang waren. Dann übernimmt schon der Nächste.“ Für die KI, die idealerweise 30 Sekunden Kontext braucht, um die Wahrscheinlichkeit gut berechnen zu können, welches Wort als nächstes kommt, ist das ein Desaster.

Also sammelte das Team der TU und Uni Graz eigene regionale Sprachdaten und erweiterte damit bestehende riesige Sprachdatenbanken. 38 Menschen aus dem österreichischen Raum – Steiermark, Wien, Ober- und Niederösterreich, Salzburg, Ostkärnten – wurden eingeladen, in Paaren zu sprechen. Mutter und Tochter, Arbeitskollegen, Menschen, die sich gut kennen, weil sie dann eher so sprechen, wie ihnen „da Schnabel gwaxn is“. Ohne Vorgaben, einfach plaudern. Hinterher baten die Forschenden ihre Probanden, ein paar Sätze vorzulesen, um Referenzen zu haben – der Kontrast zwischen gesprochener und gelesener Sprache war frappierend. „Je schneller und umgangssprachlicher jemand sprach, desto schlechter wurde die Erkennungsrate“, erzählt Schuppler.

Der Clou der Grazer: Statt bestehenden Systemen bloß neue Wörter „anzutrainieren“, entwickelten Schuppler und ihr Team ein wissenschaftliches Lexikon: Darin sind Aussprachevarianten für jedes Wort vermerkt – etwa „haben“ in all seinen dialektalen Formen. „So kann der Spracherkennner auch Wörter erkennen, die im Trainingsmaterial nie vorkamen – weil



skomma5inne

Im Angesicht der KI. Barbara Schuppler und Michael Paier mit dem „Social Robot“ namens „Furhat“, der an der TU Graz darauf trainiert wird, je nach Gesagtem die richtige Mimik zu zeigen. Unten: 5komma5inne beim TU Graz-Besuch.



„DIE SYSTEME SIND AUF STANDARDSPRACHE TRAINIERT. DIALEKTE KOMMEN DA KAUM VOR.“

er weiß, dass ‚hohn‘, ‚hom‘ und ‚haben‘ dasselbe bedeuten können.“ Dieser Ansatz ist nicht nur für Dialekterkennung spannend – sondern auch für viele andere Sprachtechnologien, die jenseits des Standarddeutschen funktionieren sollen. 25 Forschungsinstitute weltweit arbeiten schon mit den in Graz erstellten Sprachmodellen. „Das sind alles nicht kommerzielle Anwendungen“, wird betont.

„DANKE, DASS DU MEINE WELT VERSTÄNDLICH MACHST“

Aber was sind jetzt die konkreten Anwendungen? Versteht uns endlich unser Auto? „Das Problem bei Spracherkennung im Auto ist die Geräuschkulisse. Aber das ist eine starke kommerzielle Anwendung, wo sich große Firmen drauf stürzen“, erklärt Barbara Schuppler. „Als kleines Team wenden wir uns daher lieber der Medizin zu.“ Denn gerade Spracherkennung in

assistierenden Technologien (zum Beispiel zur besseren Verständigung für Menschen nach Schlaganfällen) können mit Dialekt bis jetzt nichts anfangen. „Wir arbeiten derzeit mit einem Kind mit Dysarthrie – das ist eine Sprachmotorikstörung.“ Auf einem Tablet erscheint live ein Untertitel, während das Kind spricht. So können andere Kinder leichter mitlesen – und sich langsam an die Sprache gewöhnen. Und das funktioniert: „Danke, dass du meine Welt verständlich machst“, hat das betroffene Kind auf ein Plakat gemalt, das in Schupplers Büro hängt. Dank und Ansporn. Auch bei neurologischen Erkrankungen gibt es viel Potenzial. In Kooperation mit der Universität Salzburg untersucht das Team etwa, wie Sprache zur Früherkennung von Alzheimer oder Depressionen genutzt werden kann. Die Vision: Angehörige könnten

zu Hause kurze Sprachaufnahmen machen – und durch die Sprachanalyse (deren Ergebnisse durch die gewohnte Umgebung des Patienten nicht verfälscht sind) frühzeitig Veränderungen erkennen.

Dabei geht es nicht nur um Technik, sondern auch um Kultur. Wenn Dialekte im digitalen Raum nicht vorkommen, könnten sie verloren gehen. „Wir haben in Österreich eine enorme sprachliche Vielfalt – aber kaum Bewusstsein für ihren Wert“, sagt Schuppler. Anders in der Schweiz: Dort gibt es Plattformen, auf denen Menschen ihre Dialektwörter einsprechen, vergleichen, dokumentieren. Ein kollektives Sprachgedächtnis, eine Dialekt-Landkarte, die Generationen überdauert. Wie würde ein Südtiroler sagen: Des gschlaunt.