

WIR VERSTEHEN UNS

Wie Spracherkennungssoftware die Flugsicherung verbessern kann

von Prof. Dr. Hartmut Helmke

Spätestens seit Alexa, OK Google oder Siri ist die Spracherkennung im Alltag angekommen. Sie erlaubt freie Hände, wenn wir eine Adresse ins Navi einsprechen, aber nicht nur das: Software kann die Flugsicherung in der Luft und am Boden unterstützen. Sie spart dem Lotsenpersonal Zeit und den Flugzeugen Kerosin. Das DLR-Institut für Flugführung forscht schon seit vielen Jahren daran, Spracherkennungssoftware weiterzuentwickeln und in neuen Einsatzgebieten zu testen. Aber wie funktioniert eine solche Software eigentlich und welche weiteren Vorteile bringt sie?

Lotsenpersonal kommuniziert mit Pilotinnen und Piloten in Fachsprache. Bei Sätzen wie „Speedbird two zero zero zero, reduce one eight zero knots until four miles final, contact tower on frequency one one eight decimal seven zero zero, bye bye!“ erkennt eine herkömmliche Sprachsoftware wie Siri gerade einmal die Hälfte der Worte – diese komplexe Anweisung wird sehr schnell gesprochen. Hinzu kommen die unterschiedlichen englischen Akzente. Für den realen Betrieb reicht eine solche Erkennungsrate bei Weitem nicht aus. Zusammen mit verschiedenen Flugsicherungen und Forschungseinrichtungen, wie der Universität des Saarlandes, dem Schweizer Forschungsinstitut Idiap und der Universität von Brüssel, hat das DLR-Institut für Flugführung Systeme für das Spracherkennen und -verstehen für Flugsicherungsanwendungen entwickelt. Hierzu werteten die Forschenden mehr als 50 Stunden Sprachdaten aus und verschriftlichte diese. Zum Vergleich: Google kann auf etwa 200.000 Stunden verschriftlichte Sprachdaten zugreifen. Mit den Daten trainierten die Forschenden ein neuronales Netz. Für den Vorfeldbereich des Flughafens Frankfurt erreichte das System danach Worterkennungsraten von über 97 Prozent.

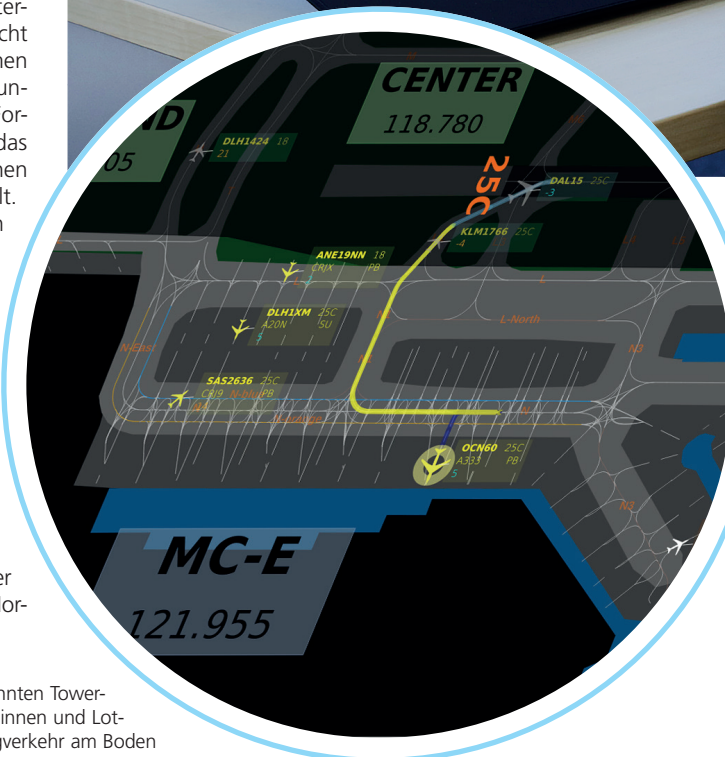
Spracherkennung versus Sprachverstehen

Neben der Erkennungsrate entscheiden auch die nicht erkannten Worte über die Qualität einer Anwendung. Ein Fehler bei „good morning“ ist deutlich weniger dramatisch als eine Verwechslung in „heading two six zero“. Wird hier „three“ statt „two“ erkannt, fliegt das Flugzeug Richtung Norden anstatt nach Westen.

Mit dem sogenannten Tower-Pad regeln die Lotsinnen und Lotsen den Flugverkehr am Boden



Freiport AG



TowerPad™/ATRICS Advanced Traffic Solutions GmbH

Bei diesem Test wurde die Lotsin von einer Spracherkennungssoftware dabei unterstützt, den Verkehr im Vorfeldbereich des Frankfurter Flughafens zu regeln.

Eine Sprachäußerung, kommt sie nun aus dem Cockpit oder aus dem Tower, setzt sich aus verschiedenen Instruktionen zusammen – einem Rufzeichen, einem Kommando sowie möglichen Bedingungen. Jede Lotsenanweisung muss vom Piloten oder der Pilotin zurückgelesen werden, um sie zu bestätigen. Dabei kann sich die Reihenfolge der Worte ändern oder es werden andere Ausdrücke genutzt. Damit auch eine Software die Äußerungen „verstehen“ kann, einigten sich 22 Partnerorganisationen aus 15 europäischen Ländern unter Federführung des DLR 2018 auf Regeln zur semantischen Interpretation von Sprechfunkäußerungen, also eine Ontologie. Zusammen mit Flugsicherungen aus ganz Europa hat das DLR diese Ontologie in verschiedenen Projekten zum Sprachverstehen eingesetzt und ihre Tauglichkeit bewiesen. Derzeit wird sie unter Leitung des DLR weiterentwickelt.

WIE WIRD IN DER FLUGSICHERUNG KOMMUNIZIERT?

speedbird two zero zero zero
reduce one eight zero knots until four miles final
contact tower on frequency
one one eight decimal seven zero zero bye bye



one eighty to four
tower one eighteen seven
speed bird two thousand

Flugzeugkennung	BAW 2000
Anweisung 1	Geschwindigkeit bis spätestens vier nautischen Meilen vor der Landung auf 180 Knoten reduzieren
Anweisung 2	Tower kontaktieren
Anweisung 3	Zur Frequenz 118.700 des Towers wechseln
Verabschiedung	

Oben sieht man eine typische Lotsenäußerung aus der Anflugkontrolle. Sie beginnt mit dem Rufzeichen, also dem Namen des Fluges. Es folgt eine Anweisung, die die Geschwindigkeit auf 180 Knoten zu reduzieren, was spätestens bis vier nautischen Meilen vor der Landung erreicht sein sollte. Anschließend soll der Pilot oder die Pilotin zur Frequenz 118.700 des Towers wechseln. Der Pilot oder die Pilotin wiederholt die Geschwindigkeit sowie die Tower-Frequenz und zum Schluss das Rufzeichen.

Die Vorfeldlotsinnen und -lotsen des Frankfurter Flughafens werden ab 2024 sämtliche gesprochenen Rollführungskommandos zusätzlich per Maus oder Tastatur in ein Kontrollsystem eingeben. Dieser Prozess wird die Sicherheit erhöhen. Er steigert allerdings die Arbeitsbelastung erheblich – schlimmstenfalls könnten sogar weniger Flugzeuge starten und landen. Eine Spracherkennungssoftware kann die gesprochenen Befehle verschriftlichen und automatisch interpretieren. Das Lotsenpersonal muss lediglich verbleibende Fehler korrigieren. Bei einer Erkennungsrate von 90 Prozent müssten sie nur noch jede zehnte Anweisung eingeben. Das erforschten die DLR-Wissenschaftlerinnen und -Wissenschaftler im Project STARFiSH (Safety and Artificial Intelligence Speech Recognition). Idiap entwickelte den Spracherkenner, das DLR lieferte das Modul zum Sprachverstehen, die Freiburger Firma ATRiCS

Advanced Traffic Solutions GmbH entwickelte den Simulator sowie das TowerPad und die Frankfurter Flughafen AG (Fraport) stellte die Versuchslotsinnen und -lotsen. Getestet wurde das System zunächst im Simulator der Fraport AG. Die Versuche im Sommer 2022 zeigten, dass eine gute Spracherkennungssoftware das Lotsenpersonal deutlich entlastet, weil sie über 50 Prozent weniger Eingaben tätigen mussten.

Von der Simulation in den Kontrollraum

Wurde im STARFiSH-Projekt die Software noch mit Sprachdaten aus dem Simulationsbetrieb getestet, ist das vom DLR geleitete europäische Projekt HAAWAI schon einen Schritt weiter. Hier wurde eine Spracherkennungssoftware eingesetzt, um Lotsenanweisungen aus dem operationellen Umfeld der isländischen und der britischen Flugsicherung zu erkennen und zu verstehen. HAAWAI steht für Highly Automated Air Traffic Controller Workstation with Artificial Intelligence Integration. Beteiligt sind Idiap, die Universität Brüssel und die Flugsicherungen aus Island, Großbritannien, Österreich und Kroatien. Die Software erkennt nicht nur die Äußerungen des Lotsenpersonals im operationellen Umfeld, sondern auch die der Pilotinnen und Piloten. Dies ist nicht nur wegen dem hohen Geräuschpegel im Cockpit herausfordernd, sondern auch aufgrund der verschiedenen Sprachakzente. Hinzu kommt, dass der Sprachkanal vom Flugzeug zum Boden meist stark verrauscht ist. Das Projekt begann 2020 mit einem Spracherkennungssystem, das zwar schon mit 3.000 Stunden Alltags-Englisch und auch mit Flugsprechfunk trainiert war, aber weder mit einer Sprachäußerung der britischen Flugsicherung noch der isländischen Flugsicherung konfrontiert worden war. Entsprechend schlecht waren die Erkennungsraten auf Wortebene zu



Für den Versuch wurden die Pilotinnen und Piloten mehrerer Flugzeuge simuliert. Hier steuert eine Person zehn Flugzeuge gleichzeitig als „Pilotin“.

Beginn mit Wortfehlerraten von 30 bis 40 Prozent. Britische und isländische Fluglotsinnen und -lotsen verschriftlichten daraufhin ihren Sprechfunk. Mit den Daten trainierten die Forschenden die neuronalen Erkennungsnetze. Danach lag die Worterkennungsrate für die Lotsinnen und Lotsen bei über 95 Prozent und für die Pilotinnen und Piloten bei über 90 Prozent. Die isländische Flugsicherung nutzte in HAAWAI Sprachverstehen, um das Lotsenpersonal dabei zu unterstützen, Fehler der Pilotinnen und Piloten beim Zurücklesen von Lotsenkommandos zu erkennen. Die britische Flugsicherung nutzte Sprachverstehen zur Vorhersage der Arbeitsbelastung der Fluglotsinnen und Fluglotsen.



Die Tests der Software zeigten, dass die Lotsinnen und Lotsen über 50 Prozent weniger Befehle selbst eingeben müssen.

Treibstoff sparen mit Spracherkennung

Dass der Einsatz einer solchen Software auch Treibstoff sparen kann, zeigten die Forscherinnen und Forscher des DLR-Instituts für Flugführung schon im Jahr 2015. Wenn die Lotsinnen und Lotsen die Befehle nicht mehr manuell eingeben müssen, sondern einfach einsprechen, haben sie mehr Ressourcen für ihre eigentliche Aufgabe, sprich für die sichere und effiziente Führung des Flugverkehrs. Wird ein Kommando im Landeanflug auch nur geringfügig zu spät gegeben, kann dies dazu führen, dass das Flugzeug einige Kilometer zu

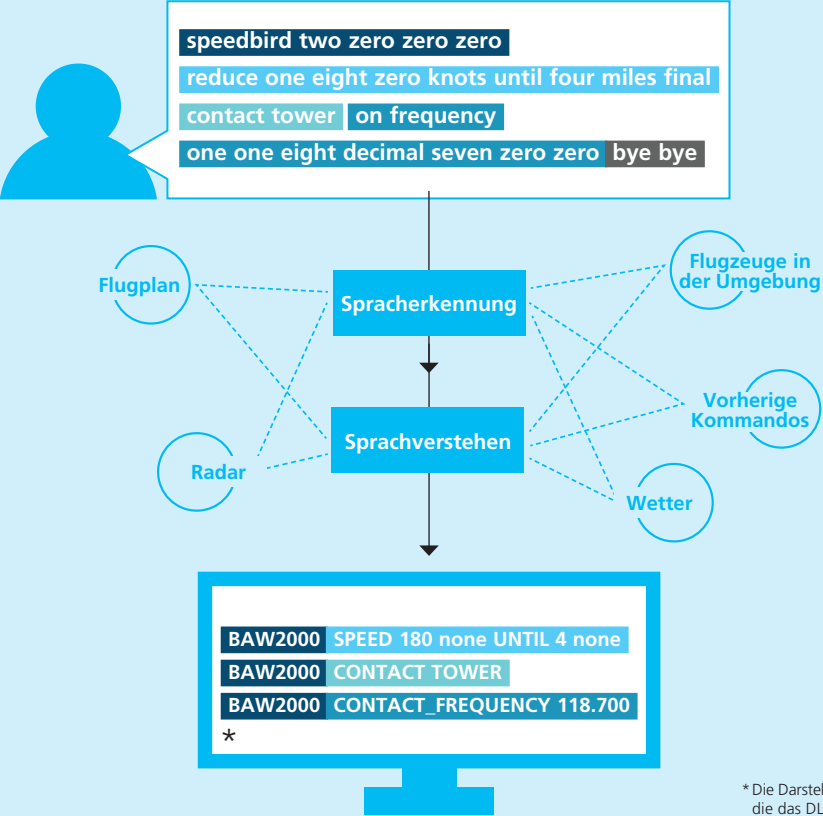
weit in die falsche Richtung fliegt. Dies bedeutet wiederum mehr Arbeit für das Lotsenpersonal. Tests mit Fluglotsinnen und Fluglotsen aus Deutschland, Österreich, Kroatien, Tschechien, Dänemark, Schweden und Irland zeigten zwischen 2015 und 2017, dass sich die Flugzeit mit einer Spracherkennungssoftware um circa 77 Sekunden verkürzen lässt. Dieses entspricht 60 Liter Kerosin pro Flug.

Erster Live-Einsatz und Ausblick

Im Sommer 2022 integrierte die isländische Flugsicherung das HAAWAI-System im Kontrollraum des Flughafens Reykjavik, um die gesprochenen Kommandos anzeigen zu lassen. Ziel dieser Demonstration war es, Rücklesefehler automatisch zu erkennen, was das System erfolgreich leistete.

Derzeit planen das DLR mit Flugzeugherstellern die Software in Cockpits zu integrieren, sodass die Pilotinnen und Piloten beim Rollen auf dem Flughafen keine Wege mehr manuell in das System eingeben müssen. Die Vision zum Ein-Personen-Cockpit rückt damit ein Stück näher. Der Flughafen Frankfurt plant das System nicht nur im Simulator einzusetzen, sondern spätestens ab 2026 auch die Lotsinnen und Lotsen direkt im Kontrollraum bei der Kommandoeingabe zu entlasten.

Prof. Dr. Hartmut Helmke arbeitet im DLR-Institut für Flugführung in der Abteilung Lotsenassistenz und ist Spezialist auf dem Gebiet des Sprachverstehens.



* Die Darstellung der Anweisungen basiert auf der Ontologie, die das DLR gemeinsam mit Flugsicherungen aus ganz Europa entwickelt hat.

WIE ARBEITET EINE SPRACHERKENNUNGS SOFTWARE?

Ein gesprochenes Sprachsignal wird von dem Spracherkennungsalgorithmus der Software in eine Folge von Worten transformiert. Der Algorithmus für das Sprachverstehen wandelt die einzelnen Worte in semantische Einheiten um. Das können Rufzeichen, Kommandotypen oder Kommandowerte sein. Um die Spracherkennung und das Sprachverstehen zu verbessern, nutzt das System die Kommandovorhersage. Diese ermittelt zum einen die Rufzeichen der Flugzeuge, die sich in der Nähe befinden und somit in nächster Zeit angesprochen werden könnten, zum anderen für jedes Rufzeichen mögliche Kommandotypen mit Werten. Kommandotypen sind zum Beispiel Richtungs-, Geschwindigkeits- oder Flughöhenänderungen. Dazu greift das Kommandovorhersage-System auf Radar-, Flugplan- sowie auf Wetterdaten zurück. Die Ergebnisse der Spracherkennung werden letztendlich automatisch auf dem Arbeitsmonitor des Lotsenpersonals angezeigt.

VOO QQQ/Unsplash



Den Verkehr auf einem Flughafen zu regeln, ist keine leichte Aufgabe.



Dieses Projekt wird durch das SESAR Joint Undertaking innerhalb des EU-Forschungs- und Innovationsprogramms Horizon 2020 unter der Nummer 884287 gefördert.