

Kaffee mit Hund

Milena Rabovsky untersucht
die Sprachverarbeitung
im Gehirn mithilfe neuronaler
Netzwerke und setzt dabei
vor allem auf Lerneffekte durch
Vorhersagefehler



Aufgepasst: „The dog was bitten by the man.“ – Alles klar? Nein? Dann ist gut. Wenn wir etwas lesen oder hören, weiß unser Gehirn oft schon, was als nächstes kommt. Dank dieses Wissens müssen wir Sätze nicht immer umständlich analysieren, sondern erfassen ihren Sinn meist schon „auf halbem Weg“. Wenn es doch einmal anders kommt, als wir erwarten, merkt das Hirn dies in der Regel zuverlässig – und messbar. Die Neurowissenschaftlerin Milena Rabovsky erforscht genau diese Abweichungen vom Erwartbaren, denn sie helfen dabei zu verstehen, wie wir Sprache lernen und verarbeiten.

Dabei interessiert sie sich gerade nicht so sehr für die Strukturen und Regeln, mit denen Sprache arbeitet. „Sprache wird zwar durch Syntax strukturiert. Ihr eigentlicher Zweck ist aber, Bedeutung zu vermitteln“, so die Wissenschaftlerin. In ihrer Forschung setzt sie sogenannte neuronale Netzwerkmodelle ein, die auch als Deep-Learning-Modelle bezeichnet werden. Dies sind Computermodelle, die unter anderem nachbilden sollen, wie das menschliche Gehirn Sprache verarbeitet. Mit ihnen versucht sie, Hirnsignale, die bei menschlicher Sprachverarbeitung gemessen werden, zu reproduzieren und zu erklären. „Wenn das gelingt, ist das ein Hinweis darauf, dass es gute Modelle sein könnten.“

Das Besondere ist: Das Verständnis von Sprache, das neuronalen Netzwerkmodellen zugrunde liegt, unterscheidet sich wesentlich von klassischen linguistischen Theorien der Sprachverarbeitung. In diesen wird angenommen, dass die Funktionalität von Sprache auf syntaktischen Regeln basiert. Und häufig auch, dass diese nicht gelernt werden können, sondern basale Aspekte dieser syntaktischen Regeln angeboren sind. Dahinter steckt die Idee einer Universalgrammatik, wie sie etwa der Linguist Noam Chomsky beschrieben hat. Für die Neurowissenschaftlerin Milena Rabovsky ist die Rolle syntaktischer Regeln nicht mehr so primär. Für die Aufgabe von Sprache, Bedeutung zu vermitteln, seien die statistischen Regelmäßigkeiten oft wesentlicher und vor allem wirkmächtiger als syntaktische Regeln. Deshalb verstünden viele Menschen den Satz „The dog was bitten by the man.“ zunächst falsch. Sie verstehen, was sie erwarten, dass nämlich der Hund den Mann beißt, und nicht, wie die Syntax vorgibt, dass der Hund vom Mann gebissen wird.

Sprache im Gebrauch lernen

„Neuronale Netzwerkmodelle lernen Sprache aus den statistischen Regelmäßigkeiten der Umwelt“, erklärt Milena Rabovsky. Und zwar je nach Herangehensweise auf verschiedene Art und Weise. „Mein Modell

etwa lernt, sprachlichen Input auf Bedeutung zu mappen, indem es – basierend auf sequentiell präsentierten Wörtern – schätzt, welches Ereignis der jeweilige Satz beschreibt, d.h. wer was mit wem (oder was) macht, und dazu Feedback bekommt, um seine Schätzungen zu verbessern.“ Andere Modelle versuchen, das jeweils nächste Wort im Satz vorherzusagen. „Die Modelle lernen Sprache einfach, indem sie ihre Aufgabe ausführen.“ Dafür seien syntaktische Regeln zwar auch wichtig, immerhin enthalten die Texte, mit denen die Modelle Sprachverstehen lernen und trainieren, normale und damit regelgerechte Sätze. Aber die Regeln werden nicht vorab in den Modellen verankert. Vielmehr lernen sie diese mit dem Sprachgebrauch. „Damit unterscheidet sich die Art, wie ich Sprachverarbeitung sehe, stark davon, wie viele Linguisten sie verstehen“, so die Forscherin. „Und ich denke, durch die beeindruckende Leistungsfähigkeit aktueller Deep-Learning-Modelle der Sprachverarbeitung (z.B. chatGPT) wird es für die linguistische Idee, Sprache könne ohne ein gewisses Vorwissen nicht gelernt werden, zumindest eng.“

Milena Rabovskys Modelle lernen Sprache letztlich so, wie auch Menschen sie lernen: Sie hören zu, immer wieder. Genauer gesagt werden sie mit sogenannten Korpora „gefüttert“, Sammlungen von typischen gedruckten Texten, die oft aus Zeitungsausgaben bestehen. Irgendwann erkennen – oder besser: beherrschen – sie Regelmäßigkeiten. Und dann wird es für die Wissenschaftlerin erst richtig spannend. Denn im Fokus der Forschung von Milena Rabovsky stehen eigentlich jene sprachlichen Phänomene, die vom Normalen abweichen. Wie der Satz vom bissigen Mann und dem armen Hund. Die Neurowissenschaftlerin erforscht nämlich die N400. Das ist kein Überschallflugzeug und auch keine Küchenmaschine, sondern eine Welle. Die N400-Welle lässt sich beim Menschen mithilfe der Elektroenzephalografie (EEG) messen und ist eine Komponente des sogenannten ereigniskorrelierten Potenzials. Die Welle gibt Auskunft über unsere Sprachverarbeitung, genauer: Sie zeigt an, wenn unser Gehirn dabei Schwierigkeiten hat. Die N400-Welle wird nämlich immer dann ausgelöst, wenn Worte auftauchen, die nicht in den Kontext passen. Ein klassisches Beispiel: „Ich trinke meinen Kaffee mit Sahne und Hund.“ Wieder dieser Hund. Ihren Namen hat die Welle von der Verzögerung, mit

Was Deep-Learning-Modelle der Sprachverarbeitung können, beweist der Ende 2022 an den Start gegangene chatGPT der Firma Open AI, der für alle frei verfügbar und in der Lage ist, auf der Grundlage von Eingaben Texte aller Art zu generieren – von romantischen Gedichten bis zu wissenschaftlichen Aufsätzen.

<https://openai.com/blog/chatgpt>

der unser Hirn die Abweichung registriert: 400 Millisekunden nachdem wir „Hund“ lesen oder hören, wird sie ausgelöst. Dabei fällt das Signal, also die Welle, größer aus, je stärker das Wort von der Erwartung abweicht. Trinken wir unseren Kaffee mit Honig ist das weniger auffällig als mit Hund.

Mensch und Modell im Experiment

Mit der N400 beschäftigt sich Milena Rabovsky schon lange, denn die studierte Psychologin ist in der EEG-Forschung „groß“ geworden. Dort bildet die Welle eines der wichtigsten Instrumente zur Untersuchung der Sprachverarbeitung. Zeigt sie doch zuverlässig an, wie unser Hirn mit sprachlichen Phänomenen umgeht, die es (noch) nicht kennt. „Wir konnten in zahlreichen Studien N400-Effekte nachweisen“, sagt sie. „Doch es war lange nicht klar, was genau dahintersteckt, was das Hirn eigentlich macht, wenn die Welle größer oder kleiner wird. Das war irgendwie frustrierend.“ Auf der Suche nach neuen Ansätzen wandte sich die Kognitionswissenschaftlerin der Modellierung zu, ging dafür nach Kanada an die University of Western Ontario, später nach Stanford. Seit 2019 ist sie an der Universität Potsdam. Hier hat sie dank der Förderung der Deutschen Forschungsgemeinschaft im Emmy-Noether-Programm eine Gruppe aufgebaut, mit der sie experimentelle Forschung und Modellierung zusammenbringt – aktuell vor allem mit Blick auf die N400. „Es besteht die Theorie,

dass die N400 einen Vorhersagefehler und den damit verbundenen Lerneffekt anzeigt“, erklärt sie. Wenn wir Sprache hören, versuche das Hirn ständig, das nächste Wort und dessen Bedeutung vorherzusagen. Bei Abweichungen kommt es zur Fehlermeldung, die N400-Welle „schlägt“ aus – und das Hirn muss seine Vorhersage korrigieren. Es lernt.

„Diese Annahmen haben wir in unsere Modelle implementiert und arbeiten nun daran, sie zu belegen“, so die Forscherin. Nachdem die Modelle einige Zeit mit Korpora trainiert haben, müssen sie sich mit Menschen messen. Dafür führen Milena Rabovsky und ihr Team im Prinzip parallel zwei Experimente durch: Im EEG-Labor werden Testpersonen mit Sätzen konfrontiert, die immer wieder unterschiedlich starke Abweichungen enthalten. Mal gibt es Milch mit Honig, mal mit Hund. Oder ähnliches. Die aufgezeichneten N400-Wellen vergleichen die Forschenden dann mit den Aktivierungsmustern der Modelle, die mit demselben sprachlichen Input „gefüttert“ wurden. „Wenn wir bei den Modellen dieselben N400-Effekte nachweisen können wie bei unseren Testpersonen, ist dies ein Hinweis darauf, dass es gute Modelle menschlicher Sprachverarbeitung sein könnten“, so Milena Rabovsky. Experimentell belegen wollen sie auch den Lerneffekt: Auf einen typischen N400-Test folgt ein zweiter „Lerntest“, bei dem sich zeigen soll, ob die Größe der N400 im ersten Teil spätere Lerneffekte vorhersagt. „Ist die N400 tatsächlich ein Lernsignal, müssten Probanden Worte, bei denen sie eine größere N400-Welle zeigten, beim zweiten Versuch schneller erkennen.“ Neu Gelerntes wird zum alten Hund.

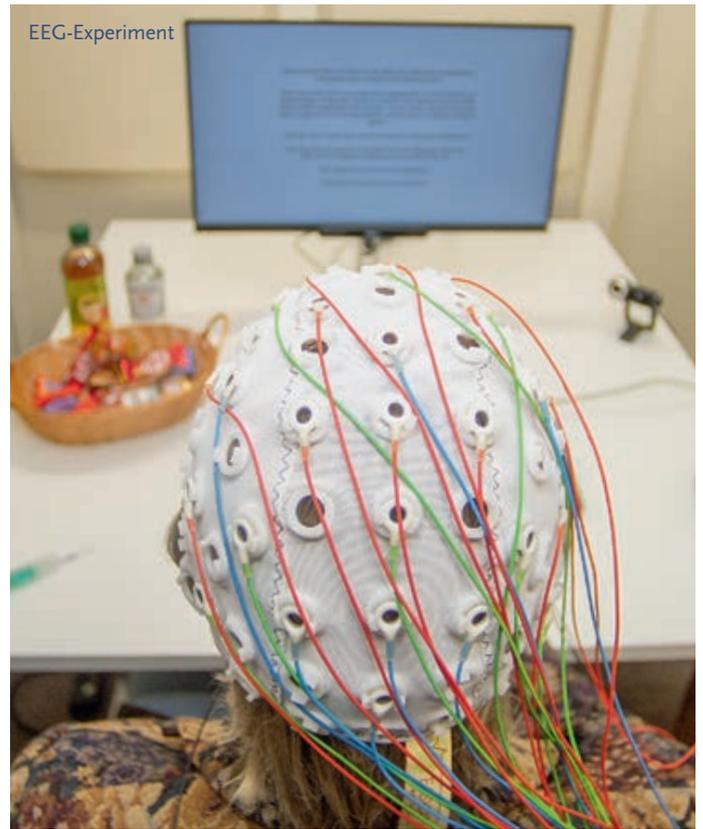
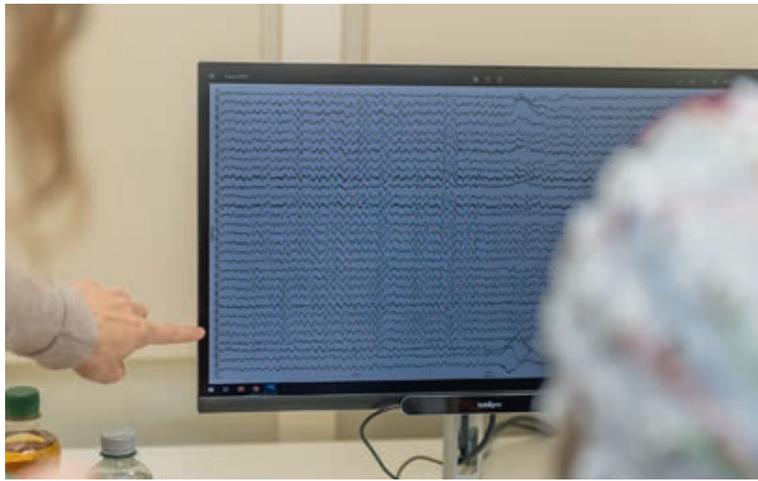
Im Idealfall lässt sich diese Erklärung dann auch am Modell verifizieren. Bislang hat es die Härtestests alle bestanden. Milena Rabovsky ist zufrieden und wendet sich bereits neuen Zielen zu. So plant sie gemeinsam mit dem Max-Planck-Institut für Kognitions- und Neurowissenschaften eine MEG-Studie. Dank der dabei eingesetzten Magnetoenzephalografie (MEG) können die Forschenden nicht mehr nur – wie mit dem EEG – erfassen, wann etwas im Gehirn geschieht, sondern sogar wo.

Prof. Milena Rabovsky

DIE FORSCHERIN

Prof. Dr. Milena Rabovsky studierte Psychologie an der Humboldt-Universität zu Berlin. Nach der Promotion an der Berlin School of Mind and Brain und Stationen in London (Ontario), Stanford und an der Freien Universität Berlin ist sie seit 2019 Professorin für Kognitive Neurowissenschaften an der Universität Potsdam.

✉ milena.rabovsky@uni-potsdam.de



DAS PROJEKT

Kombination von Elektrophysiologie und neuronalen Netzen bei der Untersuchung des Sprachverstehens (Emmy Noether-Nachwuchsgruppe)

Förderung: Deutsche Forschungsgemeinschaft (DFG)
Laufzeit: 2019–2024

Beteiligt: Prof. Dr. Milena Rabovsky (Leitung), Professor Dr. James McClelland (Stanford University)

Vom Fehler zum Lernen

Außerdem soll künftig eine zweite Welle ins Modell einfließen: Die P600. „Während die N400 so etwas wie den automatischen Prozess der Fehlererkennung widerspiegelt, zeigt die P600 möglicherweise die Reanalyse an“, erklärt Milena Rabovsky. „Also einen relativ kontrollierten kognitiven Prozess, in dem Menschen den Fehler korrigieren und Missverständenes richtig erfassen.“ Wer etwas falsch versteht, anschließend korrigiert und dann richtig erfasst, sollte sich an den beiden Wellen zeigen lassen. Da dieser Vorgang durchaus kognitiv herausfordernd ist, müssen die Experimente auch verschiedene Faktoren berücksichtigen. Ob die Versuchspersonen aufmerksam sind, etwa. Dafür sollen künftig nicht nur EEG-, sondern auch Augenbewegungsmessungen eingesetzt werden. Und natürlich das „gewachsene“ Modell.

Von dem ist Milena Rabovsky jedenfalls überzeugt. Auch wenn es natürlich momentan ein stark vereinfachtes Modell sei, würde die Vielzahl der erklärten neuronalen Effekte nahelegen, dass die dem Modell zugrundeliegenden Prinzipien wesentliche Aspekte der menschlichen Sprachverarbeitung erfassen. Und auch wenn sie betont, dass sie Grundlagenforschung betreibt, deren Aufgabe es ist, das menschliche Sprachverstehen zu untersuchen, sieht sie bereits mögliche Anwendungen: „Wenn man ein wirklich gutes Modell hat, kann man mit dessen Hilfe etwa Aphasien besser verstehen oder Mechanismen von Krankheiten erklären, die sich auf unser Sprachvermögen auswirken.“

Bis dahin sei der Weg zwar noch weit, aber gerade die enge Zusammenarbeit mit den Potsdamer Kolleginnen und Kollegen am Institut für Linguistik in dieser Hinsicht inspirierend. Denn Berührungspunkte gibt es zwischen Linguistik und kognitiven Neurowissenschaften keine. Im Gegenteil: So arbeitet Milena Rabovsky etwa im linguistischen Sonderforschungsbereich „Die Grenzen der Variabilität der Sprache“ mit einigen sehr produktiv zusammen: „Wir haben ähnliche Interessen und es passt gerade experimentell sehr gut. Außerdem sind wir alle tolerant. Und vielleicht denkt jeder im Geheimen, dass er die anderen irgendwann doch noch überzeugt.“

MATTHIAS ZIMMERMANN