

# SOUNDCHECK

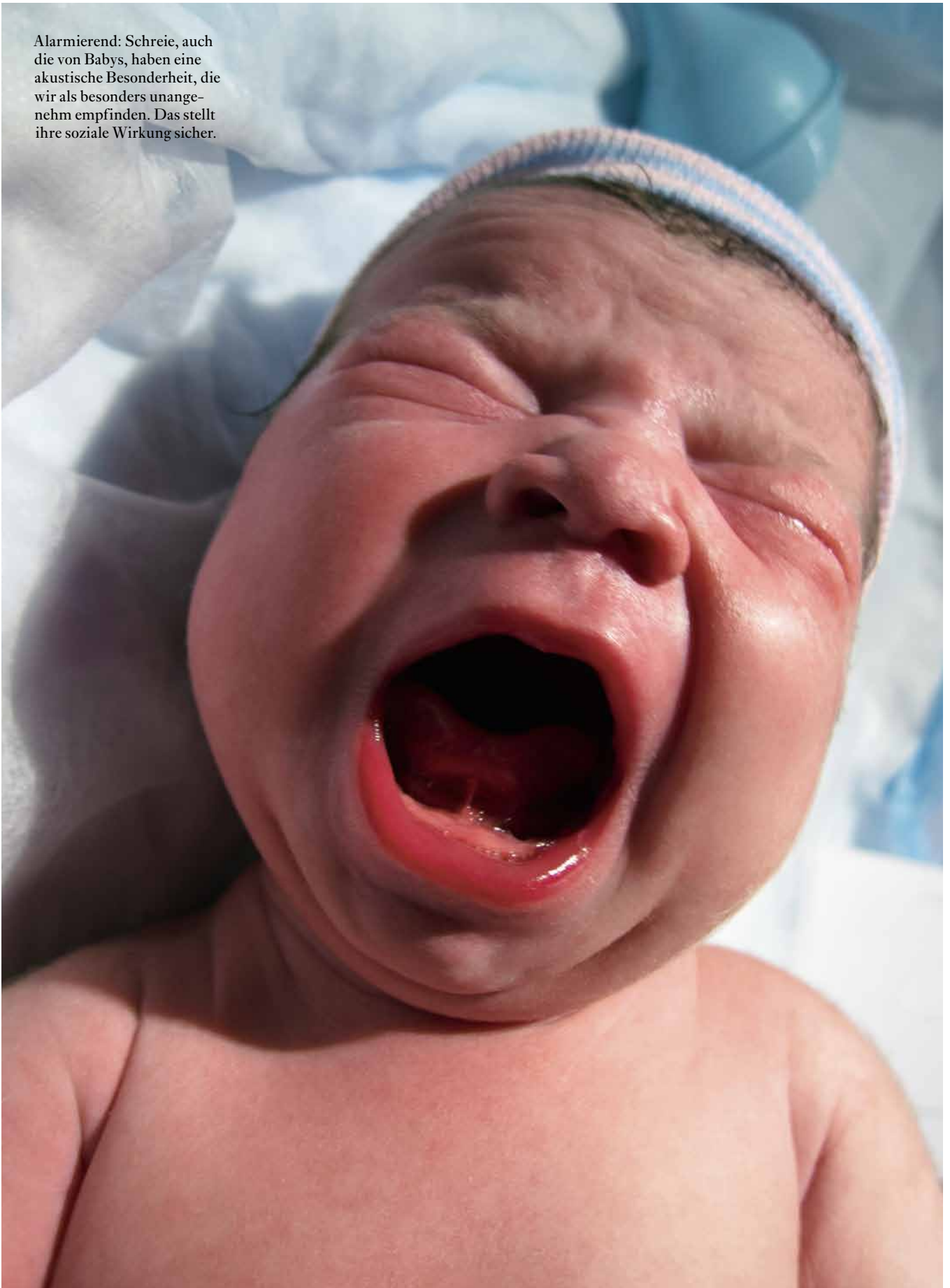
TEXT: TOBIAS HERRMANN

32

Operngesang. Vogelgezwitscher. Lautes Schreien. Eine nicht gestimmte Geige. Manche Töne finden wir instinktiv angenehm, andere unangenehm. Doch wie entscheiden wir, ob etwas gut oder schlecht klingt? Und wie wird Schall im Gehirn überhaupt verarbeitet? Um diese Fragen zu beantworten, versucht das Team um David Poeppel am Max-Planck-Institut für empirische Ästhetik in Frankfurt, Sprache und Musik in ihre elementarsten Bestandteile zu zerlegen. Und am Max-Planck-Institut für Kognitions- und Neurowissenschaften in Leipzig ergründen Forschende das Geheimnis von Superhits.

Alarmierend: Schreie, auch die von Babys, haben eine akustische Besonderheit, die wir als besonders unangenehm empfinden. Das stellt ihre soziale Wirkung sicher.

FOTO: PLAINPICTURE/BLEND IMAGES/ROBERTO WESTBROOK





Einzigartige Messbühne: Das ArtLab am Frankfurter Max-Planck-Institut ist zugleich Konzertsaal und Labor. Es ermöglicht, während der Darbietung verschiedene physiologische Daten von Künstlern – hier das Vokalensemble Cut Circle – und Zuhörenden aufzuzeichnen.

34

Auf die Frage: „Was machen Neurowissenschaften?“ antwortete David Poeppel in einem Interview einmal: „Ein Ding auseinandernehmen und in seine Einzelteile aufbrechen.“ Diese Aussage unterstreicht sowohl die übliche Herangehensweise des Forschers als auch die generelle Ausrichtung des Max-Planck-Instituts für empirische Ästhetik in Frankfurt, an dem Poeppel seit 2014 Direktor ist. An jenem Apriltag im Jahr 2020 erreicht man ihn dort allerdings nicht, stattdessen muss man eine Telefonnummer wählen, der +1 vorangestellt ist – die internationale Vorwahl der USA. Seit 2009 hat der Wissenschaftler im Nebenamt eine Professur für Psychologie und Neurowissenschaften an der New York University inne. Als die Corona-Pandemie ausbrach, verließ Poeppel mit seiner Familie den Hotspot New York City gen Connecticut, wo er bis auf Weiteres von zu Hause arbeitet – was er als „blessing in disguise“ bezeichnet, Fluch und Segen zugleich also, denn wenigstens könne er nun einige Vorhaben angehen, für die er im Tagesgeschäft nie Zeit hatte.

Zeit spielt auch in seiner Forschung eine relevante Rolle, eines seiner Forschungsgebiete umfasst die zeitliche Verarbeitung von Sprache und Musik. Poeppel formuliert es platt: „Eine Schallwelle kommt ans Ohr, wird in ein elektrisches Signal umgewandelt und in Schaltstellen im Gehirn auseinandergefummelt.“

Am Schluss erhält man dann kleine Elementarteilchen, die – wenn man sie korrekt verarbeitet – die entsprechende Information liefern.“ Ihn interessiert also, wie akustische Signale im menschlichen Gehirn verarbeitet werden. Anhand dieser Überlegungen erhofft sich Poeppel dann weiterführende Erkenntnisse für Theorien in den Sprachwissenschaften oder zu ästhetischen Aspekten von Sprache und Musik.

## Viele Sprachen, ein Tempo

Man kommt nicht umhin, ein Gespräch mit David Poeppel inspirierend zu nennen. Viele Fragen, die er formuliert, hat man sich vielleicht selbst schon einmal gestellt, andere treffen einen ziemlich unvorbereitet. Manche klingen sehr komplex, andere beinahe schon verblüffend simpel. So erscheint die Frage: „Was ist eine Silbe?“ zunächst banal, schließlich muss man ja kein Linguist sein, um zu wissen, dass jedes Wort aus einer oder mehreren Silben besteht. Doch in wissenschaftlich-technischer Hinsicht ist die Sache weit weniger klar. Seit etwa 70 Jahren, so Poeppel, diskutieren Sprachwissenschaftler nämlich, ob Silben als Elementarteilchen der Sprache gelten sollten oder doch mehr als eine Art Nebenerscheinung kleinerer akustischer Elemente wie der Phoneme, also einzelner Laute.

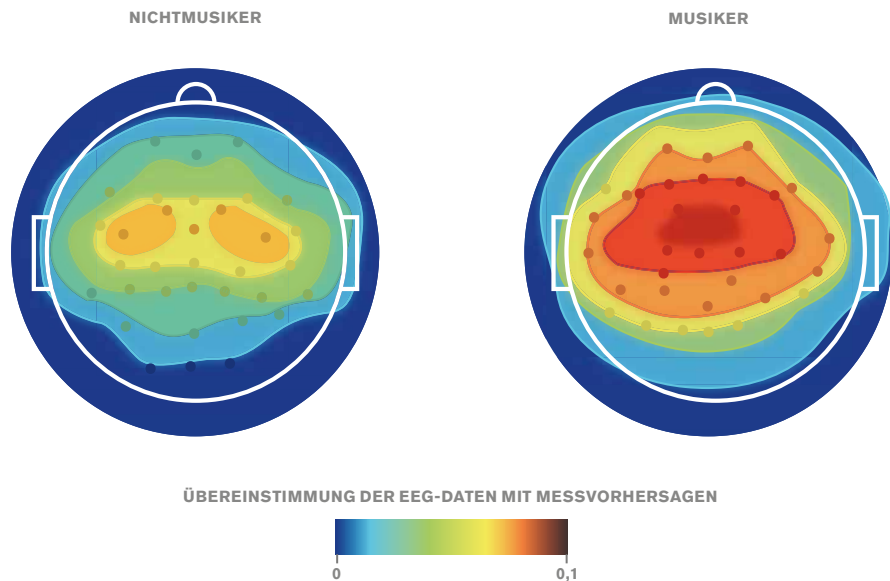
Unbestritten spielen Silben für die Sprachwahrnehmung und Sprachproduktion eine fundamentale Rolle. In einem längeren Projekt analysierte Poeppel mit Kolleginnen und Kollegen für zahlreiche Sprachen das jeweilige Sprachtempo in Relation zur Silbenanzahl. Dabei zeigte sich, dass die durchschnittliche Sprachrate der Geschwindigkeit entspricht, mit der eine Silbe auf die nächste folgt. Das Erstaunliche daran: Dieses Tempo war über unzählige Sprachen hinweg fast identisch. „Das Gefühl, wonach manche Sprachen deutlich schneller gesprochen werden als andere, ist also falsch“, sagt Poeppel.

vorne oder hinten betont ist. Erst diese sogenannte prosodische Kontur macht Sprache interessant und lebendig. Wie wichtig diese Faktoren sind, zeigt sich besonders bei Sarkasmus und Ironie. Je nach Betonung ist dem Gegenüber sofort klar, ob die Aussage „Gut gemacht!“ lobend oder ironisch gemeint ist.

Um solche Nuancen wahrnehmen zu können, sind längere Zeitfenster notwendig, statt einer zeitlichen Auflösung muss das Gehirn zusätzlich noch eine spektrale Auflösung generieren; der Frequenzbereich liegt hier bei einigen Hertz, die entsprechenden Zeitab-

GRAFIK: GGO NACH ACOUSTICAL SCIENCE AND TECHNOLOGY, 41, 361-364

Musikalische Köpfe: Forschende am Max Planck-NYU Center for Language, Music and Emotion untersuchten mittels Elektroenzephalografie (EEG), wie das Gehirn einer Melodie folgt. Auf dieser Basis erstellten sie Vorhersagen über die messbaren Reaktionen. Die Prognose traf bei Musikern genauer zu, und die Korrelation mit den Melodien war besser als bei Nichtmusikern.



In einer Sekunde kann ein Mensch problemlos vier bis fünf mehrsilbige Worte sprechen. Um jedes Wort eindeutig zu verstehen, muss der Zuhörer jeden einzelnen Laut erfassen. Zudem ist die Abfolge der Laute entscheidend. Schon minimale Ungenauigkeiten sorgen für Chaos wie im beliebten Kinderspiel Stille Post. Da wird aus „Wald“ schnell mal „Ball“, aus „Schule“ wird „Schuhe“, „stinken“, „blinken“ und „pinkeln“ lassen sich geflüstert so gut wie nicht unterscheiden.

schnitte sind zwischen 200 und 300 Millisekunden lang. „Letztlich müssen im Gehirn also zwei parallele Prozesse ablaufen“, erklärt Poeppel. „Die richtige Reihenfolge der Laute erfahre ich dabei durch die kurzen, die Betonung und die Sprachmelodie durch die langen Zeitfenster.“ Will man verstehen, wie diese unterschiedlich langen *samples* nun genau analysiert und in konkrete Informationen umgewandelt werden, muss man tief in die Neurobiologie eintauchen. Dort spielen sogenannte neuronale Oszillationen eine wesentliche Rolle.

Damit solche feinen Unterschiede in normalen Gesprächen wahrgenommen werden, ist im Gehirn eine zeitliche Auflösung zwischen 20 und 80 Millisekunden notwendig. „Das Gehirn muss so beschaffen sein, dass es ganz kleine Zeitfensterchen baut, sogenannte *samples*“, sagt Poeppel. Dann habe man jedoch nur eine stakkatoartige Abfolge von Lauten. Sprache hängt aber auch ganz wesentlich von einer bestimmten Betonung, von Pausen und einer Sprachmelodie ab. Für die Bedeutung des Satzes „München wird modern“ ist es zum Beispiel relevant, ob „modern“

Darunter verstehen Neurobiologen die synchrone Aktivierung bestimmter Zellverbände. Sobald ein Satz, eine Melodie oder ein Geräusch als Schallwelle am Ohr ankommt und in ein elektrisches Signal umgewandelt wird, werden im Gehirn gewisse Nervenzellen gleichgeschaltet und in definierten Zyklen an- oder abgeschaltet. Um die zuvor erwähnten kurzen Zeitfenster unter 100 Millisekunden zu verarbeiten, schwingen die dafür zuständigen Zellen mit einer Frequenz zwischen 25 und 35 Hertz, man spricht von



Gammawellen. Für die längeren Abschnitte wird eine andere Art von Zellen mit der Frequenz zwischen drei und acht Hertz aktiviert, diese Schwingungen werden als Thetawellen bezeichnet.

Neuronale Oszillationen spielen nicht nur für die Wahrnehmung von Sprache eine Rolle. Sie unterstützen im Gehirn auch die Verarbeitung von Musik, wie David Poeppel gemeinsam mit Keith B. Doelling von der New York University herausgefunden hat. Die beiden verglichen in ihrer Untersuchung aktive Musiker mit mindestens sechs Jahren musikalischer Ausbildung mit Nichtmusikern. Die Testpersonen hörten mehrmals hintereinander 13 Sekunden lange Ausschnitte aus verschiedenen klassischen Klavierstücken von Johann Sebastian Bach, Ludwig van Beethoven oder Johannes Brahms, wobei sich die Stücke hinsichtlich des Tempos unterschieden – von einer Note in zwei Sekunden bis zu acht Noten pro Sekunde.

## Menschen verarbeiten Musik ähnlich wie Sprache

Bei den Musikstücken mit einem schnelleren Rhythmus als eine Note pro Sekunde ließen sich sowohl bei Musikern als auch bei Nichtmusikern kortikale Oszillationen messen, die sich mit der Geschwindigkeit der Töne im gehörten Stück synchronisierten. „Die Befunde zeigen, dass das Vorhandensein dieser Schwingungen unsere Wahrnehmung von Musik und von Tonhöhenänderungen verbessert“, erklärt Keith Doelling.

Dabei war auch zu beobachten, dass sich die Gehirne der Musiker deutlicher mit dem Rhythmus der Musik synchronisierten als die der Teilnehmenden ohne musikalische Ausbildung. Zusätzlich fanden sich nur bei Musikern Oszillationen, die auch mit ungewöhnlich langsamen Stücken in Einklang kamen. Dieser Unterschied weist darauf hin, dass Menschen ohne musikalische Ausbildung kontinuierliche Melodien möglicherweise schlechter erkennen können und Musik eher als aneinandergereihte Töne wahrnehmen. Für die Forschung zeigen die Ergebnisse zudem, dass niederfrequente Oszillationen dem Gehirn dazu dienen, sich Sprache oder Musik zu erschließen.

Pauline Larrouy-Maestri, Wissenschaftlerin in Poeppels Abteilung, interessiert sich ebenfalls für die Zusammenhänge zwischen Sprache und Musik. Angesichts ihres breit gefächerten Hintergrunds scheint sie für diese Art der Forschung geradezu prädestiniert. Sie studierte Psychologie und Musik, spielt Klavier und arbeitete als Sprachtherapeutin. Ein typisches Experiment wäre etwa: Larrouy-Maestri bittet Freiwillige, bestimmte Musik zu hören und die Darbietung zu bewerten. Dazu greift sie entweder

auf natürliches oder synthetisches Musikmaterial zurück, von denen beide Vor- und Nachteile haben. „Synthetisches Material lässt sich sehr einfach manipulieren und kontrollieren, ist aber weniger natürlich, weshalb sich nur eingeschränkt sagen lässt, wie Menschen Musik tatsächlich wahrnehmen“, sagt Larrouy-Maestri.

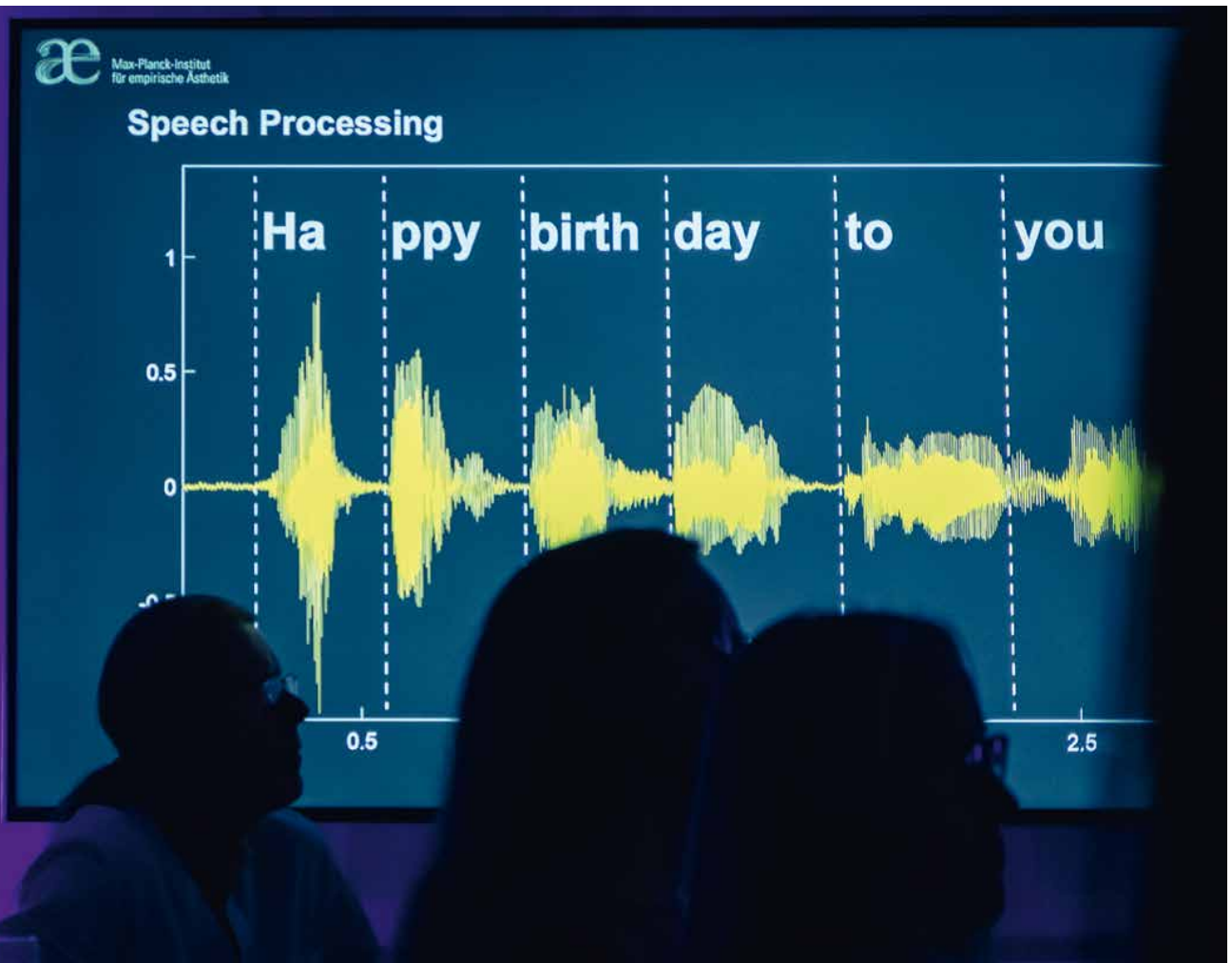
In einem Experiment, das sie mit ihrem Kollegen Xiangbin Teng durchführte, hörten Freiwillige Musikstücke von Bach, die jedoch an bestimmten Stellen verändert waren. Währenddessen beobachteten die Forschenden, wie die Gehirne der Probanden auf die manipulierten Stellen reagierten. Es zeigte sich: Die Testpersonen waren in der Lage, harmonische Strukturen wahrzunehmen, und konnten daher genau sagen, welche Stellen manipuliert waren. Mit dem Experiment zeigten Pauline Larrouy-Maestri und Xiangbin Teng, dass Menschen musikalische Werke ähnlich verarbeiten wie Sprache. Während fließend Gesprochenes in linguistische Abschnitte unterteilt wird, wie etwa Sätze, Wörter oder Silben, wird ein zusammenhängendes Musikstück in musikalische Einzelteile wie Melodien, Akkorde oder Noten zerlegt. Je musikalischer die Probandinnen und Probanden waren, desto besser konnte ihr Gehirn die musikalischen Teilabschnitte voneinander unterscheiden.

Ein weiterer Schwerpunkt von Larrouy-Maestris Forschung ist die Frage, ob musikalische Laien in der Lage sind, schiefe Töne in Liedern festzustellen – und welche kognitiven Prozesse dabei beteiligt sind. Eine Erkenntnis: Man muss kein Profimusiker sein, um falsche Töne in einem Musikstück zu entdecken. Ebenso wenig ist es notwendig, ein Expertengehör zu haben, um herauszuhören, ob ein Sänger oder eine Sängerin in der korrekten Tonhöhe singt. Praktisch jeder kann Gehörtes in richtig oder falsch einordnen – unabhängig von der jeweiligen Musik.

**„Man muss  
kein Profimusiker  
sein, um falsche  
Töne in einem  
Musikstück  
zu entdecken.“**

PAULINE LARROUY-MAESTRI





**Kleinteilige Verarbeitung:** Damit wir verstehen, was jemand sagt, muss unser Gehirn jedes Detail wahrnehmen. In Zeitfenstern, die nur 20 bis 80 Millisekunden lang sind, wird das Gehörte analysiert. Das entspricht in etwa der Länge einzelner Silben.

Wie David Poeppel beginnt auch Pauline Larrouy-Maestri ihre Forschung meist mit Alltagsbeobachtungen. Man macht das Radio an, wechselt vielleicht ein paarmal den Sender, bis man ein Lied findet, das man gern hören mag. „Egal welche Art von Geräusch wir wahrnehmen, wir können sofort sagen, ob wir es mögen oder nicht. Erstaunlicherweise besitzt jeder diese Fähigkeit. Und ich habe mich gefragt: Wie kann das sein?“ Um darauf eine Antwort zu finden, greift Larrouy-Maestri bei ihren Experimenten auf natürliche Musik zurück. „Wir laden Freiwillige – professionelle Sänger oder Laienmusiker – in unser Labor ein und bitten sie, uns etwas vorzusingen.“ Andere Freiwillige sollen dann die Performance beurteilen.

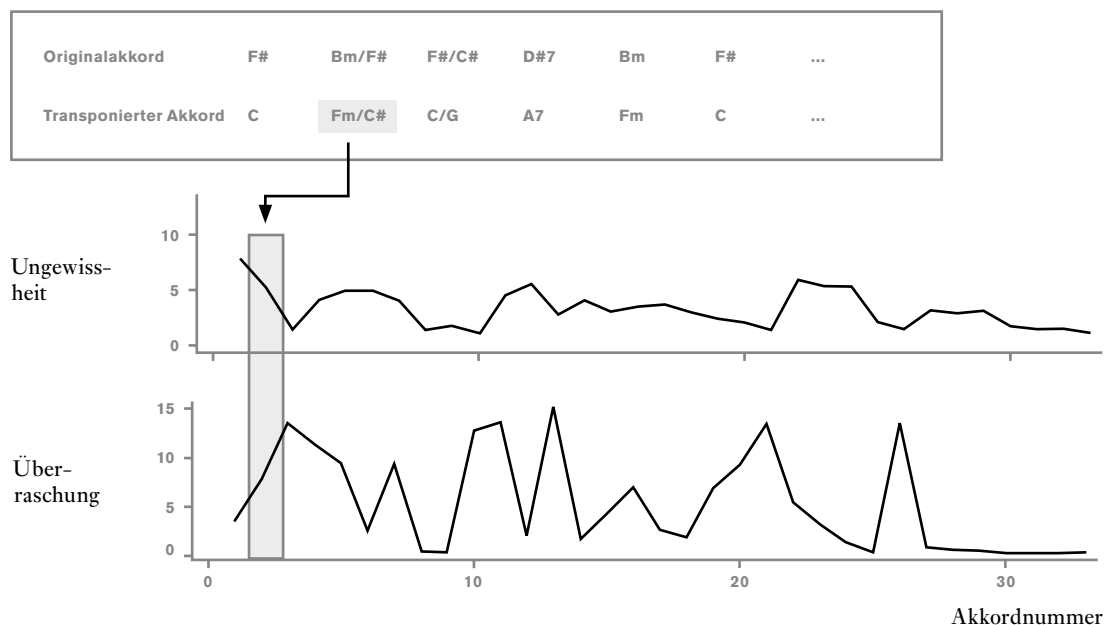
In einer weiteren Studie versuchte Larrouy-Maestri herauszufinden, welche Faktoren die Wahrscheinlichkeit bestimmen, ob eine Melodie als angenehm oder eher unangenehm empfunden wird. Bei Gesang zählt dazu, ob die Melodie korrekt wiedergegeben wird, zum Beispiel wie genau die einzelnen Tonhöhen ge-

troffen werden. Die Genauigkeit ist jedoch nicht das einzige Kriterium. Die Geschwindigkeit der Musik scheint ebenfalls eine Rolle zu spielen. Weder sehr schnelle noch sehr langsame Musik wird von der Mehrheit als angenehm empfunden. Das könnte damit zusammenhängen, wie das menschliche Gehirn Musik generell verarbeitet. Diese zugrunde liegenden Mechanismen versucht Larrouy-Maestri zu entschlüsseln, um so herauszufinden, was Menschen dazu bringt, etwas gut oder schlecht zu finden. „Ich versuche nicht, den geheimen Code zu knacken und den perfekten Song zu kreieren, den jeder mag. Mich interessiert vielmehr, wie Zuhörer zu dem Schluss kommen, dass ihnen dieses oder jenes Stück besonders gut gefällt“, sagt die Wissenschaftlerin.

Was ein erfolgreicher Hit im menschlichen Gehirn auslöst – diesem Geheimnis ist wiederum ein Forschungsteam am Max-Planck-Institut für Kognitions- und Neurowissenschaften in Leipzig auf der Spur. Vincent Cheung, Doktorand am Institut, ist



GRAFIK: GCO NACH VINCENT CHEUNG



Wissenschaftlicher Soundcheck: Mit Hilfe eines lernenden Algorithmus analysierten Leipziger Forscher die Akkordfolgen von 745 Superhits aus den USA. Für bessere Vergleichbarkeit wurden die Originalakkorde transponiert. „Ungewissheit“ bedeutet, dass sich der nächste Akkord nicht gut vorhersagen lässt. „Überschung“ zeigt an, wie sehr der eben gehörte Klang von den Erwartungen abweicht.

38 nicht nur großer Musikfan, sondern spielt auch selbst Geige. Er wollte wissen, warum bestimmte Stücke ihn und andere Fans so besonders berühren. Gemeinsam mit Stefan Koelsch suchte er nach dem Erfolgsrezept von Songs wie „Yesterday“ von den Beatles, „Nutbush City Limits“ von Tina Turner oder „The Look“ von Roxette – und wurde fündig. Mit Hilfe eines maschinellen Lernmodells analysierten die beiden 745 Superhits der US-Billboard-Charts aus den Jahren 1958 bis 1991. Dafür entfernten sie Elemente wie Text und Melodie aus den Stücken, sodass nur die Akkordfolgen übrig blieben. Wer Gitarre spielt, kennt solche Drei- und Mehrklänge aus der Begleitung von Liedern. Vom Hören sind sie wohl den meisten Menschen in der westlichen Welt vertraut, denn bestimmte Akkordfolgen finden sich seit Langem auf ähnliche Art und Weise in der westlichen Musik – vom einfachen Volkslied bis hin zur modernen Popmusik.

Die Wissenschaftler berechneten nun, wie vorhersehbar oder überraschend die Akkordfolgen in den jeweiligen Hits sind, und untersuchten die Reaktionen von Testpersonen auf die Klangfolgen. Das Ergebnis: Am größten ist der Hörerfolg, wenn es immer wieder Überraschungen gibt, aber zugleich nicht zu viel Ungewissheit entsteht. Waren die Zuhörer relativ sicher, welche Akkorde als Nächstes erklingen würden, fanden sie es angenehm, wenn sie stattdessen überrascht wurden – wenn ihre Erwartungen also verletzt wurden. Waren sie dagegen unsicher, was als Näch-

tes kommt, war es ihnen lieber, wenn die nachfolgenden Akkorde sie nicht überraschten. Untersuchungen der Probanden mithilfe funktioneller Magnetresonanztomografie untermauerten das Ergebnis: Der Nucleus accumbens, die Gehirnstruktur, die für die Erwartung von Glücksgefühlen zuständig ist, reagierte bei den Testpersonen nur dann, wenn sie besonders gespannt waren, wie die Musik weitergeht.

Aber natürlich ist nicht alles, was uns zu Ohren kommt, angenehm – unser Hörsystem dient zum Beispiel auch dazu, uns vor Gefahren zu warnen. Auch damit beschäftigen sich die Forschenden am Max-Planck-Institut für empirische Ästhetik. David Poeppel hat viel Aufmerksamkeit mit einer Studie erregt, in der er untersuchte, warum Schreie uns bis ins Mark erschüttern. Wieder so eine Frage, die einfach klingt, bis man sich auf die Suche nach einer Antwort macht. „Jeder kennt Schreie, und jeder hat eine ungefähre Vorstellung davon, was Schreie ausmacht – sie sind

## AUF DEN PUNKT GEBRACHT

Das Gehirn analysiert gesprochene Sprache in verschiedener zeitlicher Auflösung, um sowohl einzelne Laute als auch größere Zusammenhänge verstehen zu können.

Musik wird in ähnlicher Weise vom Gehirn verarbeitet: Dazu wird ein Musikstück in einzelne Bestandteile wie Melodien, Akkorde und Noten zerlegt.

Profimusiker können musikalische Strukturen teilweise besser erkennen als Menschen ohne musikalische Ausbildung. Aber auch Laien erkennen falsche Töne oder manipulierte Harmonien.

Beliebte Songs zeichnen sich durch eine Mischung vorhersehbarer und überraschender Akkordfolgen aus.

laut, hoch und schrill“, beschreibt Poeppel den Ausgangspunkt seiner Analyse. In mehreren Studien, die er gemeinsam mit seinem New Yorker Kollegen Adeen Flinker sowie Luc Arnal, Andreas Kleinschmidt und Anne-Lise Giraud von der Universität Genf durchführte, fand das Team eine akustische Besonderheit, die nur Schreie aufweisen.

„Schreie haben ein Merkmal, das als ‚Rauigkeit‘ bezeichnet wird“, erklärt Poeppel. „Rauigkeit entsteht, wenn Geräusche durch Änderung der Amplitude eine bestimmte zeitliche Struktur erhalten. Wenn diese Änderungen sehr schnell erfolgen, ist das Gehör nicht mehr in der Lage, diese zeitlichen Veränderungen aufzulösen – man empfindet ein solches Geräusch dann als rau und damit als unangenehm.“ Normale Sprache hat eine Modulationsfrequenz von etwa vier bis fünf Hertz, für Rauigkeit liegt die Frequenz dagegen zwischen 30 und 150 Hertz – die Veränderungen sind also wesentlich schneller.

In einer Studie erstellte das Forschungsteam eine Geräuschdatenbank, die viele verschiedene Arten von menschlichen Lauten, etwa Schreie und Sätze, und

künstliche Töne, zum Beispiel den Alarm eines Weckers, umfasste. Dabei fanden sie heraus, dass sowohl Schreie als auch künstliche Töne wie die des Alarms eines Weckers und dissonante Intervalle, etwa eine unreine Quinte, in den Frequenzbereich der Rauigkeit fallen – ein Ergebnis, das zeigt, dass die Hersteller von Weckern mit ihrem Alarmton sehr gut die Modulation eines menschlichen Schreies nachempfunden haben.

Der Schall, den wir wahrnehmen, kann also von verschiedenster Bedeutung für uns sein. Es ist eine enorme Leistung des Gehirns, nicht nur verschiedene Schallquellen zu unterscheiden, sondern auch zu filtern, was gerade wichtig für uns ist, und das Gehörte richtig zu entschlüsseln. Auch wenn die Forschung in den vergangenen Jahren zahlreiche neue Erkenntnisse gewonnen hat, sind auch grundlegende Rätsel noch nicht gelöst, wie David Poeppel betont – etwa das der Interaktion von Schall und Gedächtnis. Es bedarf also noch vieler kluger Köpfe und Visionäre, die die richtigen Fragen stellen.

 [www.mpg.de/podcasts/schall](http://www.mpg.de/podcasts/schall)



ANZEIGE



//

*Es gibt viele Magazine für die deutsche Wissenschaft: Jede Hochschule, jede außeruniversitäre Einrichtung, jede Institution der Forschungsförderung, jede Stiftung präsentiert in ihnen das je eigene Profil. Hinzu kommen die Wissenschaftsseiten der großen Zeitungen und sehr gute Blogs, die die Forschung kritisch begleiten. Die DUZ aber kann und macht alles: Sie ist Fundus und Dach, Seismograph und Reflektor. Sie zeigt Forschung, Lehre, Transfer, Administration. Und die Geschichten und Gesichter, die dahinterstehen.*

//



Prof. Dr. h.c. Jutta Allmendinger, Ph.D.  
Präsidentin des Wissenschaftszentrums Berlin für Sozialforschung (WZB)

