

## **Von Zuständen und Merkmalen des Zuhörens**

Jonas Obleser

Hören, zuhören, verstehen: Was ist daran anders, gar komplizierter als beim Sehen? Der Hörsinn und unsere Fähigkeit zuzuhören stellen einen besonderen Testfall dar, wenn man Hypothesen über Gehirnfunktionen anstellen und überprüfen will: Wie leistet es das Gehirn zunächst, die eindimensionalen Luftdruckschwankungen zu wandeln in ortbare auditive Objekte und sodann in interpretierbare Codes, sei es Geräusch, Sprache, Musik? In der hier vorgestellten psychologischen und neurowissenschaftlichen Forschung steht die einzelne Zuhörende im Mittelpunkt, ihre biologisch-psychologischen Zustände, in denen sie sich befindet («states») sowie die überdauernden Merkmale, die sie auszeichnen und sie von anderen Zuhörern unterscheiden («traits»). Unser Zuhören ist individuell, denn es dient immer auch bestimmten Verhaltenszwecken, und diese Zwecke unterscheiden sich und formen unsere Strategien des Zuhörens. Nicht zuletzt werden wir sehen, dass im Zuhören, Weghören, und Zusammenführen unserer Hör- und Sehsinneseindrücke der psychologische Prozess der Aufmerksamkeit eine Schlüsselrolle einnimmt.

### **I. Einleitung**

Menschen hören. Und Menschen hören zu. Was im Deutschen dieses fast beiläufige Wort «zu» markiert, ist eine vollkommen andere, weitere Maschinerie an psychologischen – und für mich immer auch: neurobiologischen – Prozessen, die sich hier anschließt. Das Englische unterscheidet deutlicher «hearing» von «listening». Schnell wird uns klar, dass es die Hörerin selbst sein muss, die das Zuhören erst bewerkstelligt, während sie sich gegen das Hören gar nicht wehren kann.

Hören, dieses «Feeling at a distance», benutzt jenes kuriose Sensorium, das unser Ohr darstellt – insbesondere dieses durch kleinste Knöchelchen stimulierte, sehr empfindliche und somit durch Lärm, toxische Substanzen wie auch vom zunehmenden Lebensalter leicht zu schädigende Innenohr. Aber es ist ein Sinn, der es dem Augentier Mensch erlaubt, auf die Ferne sowie rück- und seitwärts Gefahren ebenso wie Interessantes überhaupt erst zu bemerken.

Wer nun hingegen zuhören, hinhören will, der hat auch nur jenen Hörsinn zur Verfügung. Aber er muss einen «Prozess», ein «Modul» (wie man es in der uns vertrauten Sprache des Informationszeitalters nennen könnte), eine ganze Familie von Prozessen zum Einsatz bringen, die ihm erlauben, dieses gehörte Signal zu filtern und zu gewichten. Es muss physikalische Eigenschaften des uns interessierenden Höreindrucks geben, die uns erlauben, ihn zu unterscheiden von anderen Eindrücken und so gewisse Anteile dieses Gehörten zum «Signal» zu deklarieren und andere zum «Noise» oder Hintergrund. Solch ein «Filter» (wie ich ihn kurzerhand nennen möchte), der dies bewerkstelligt, und zwar fast nach Belieben, wie es unseren Absichten und Zielen entspricht – dieser Filter hat natürlich in der Psychologie seit nunmehr 130 Jahren (James 1890/1927) einen Namen, auch wenn dieser Begriff allein uns erstaunlich wenig weiterhelfen wird: Es ist ebendies die Aufmerksamkeit.

Ist Zuhören also nichts anderes als aufmerksames Hören? Die kurze Antwort ist: ja. Es wird aber an dieser Stelle für uns nicht langweilig, denn die seit eben jenen 130 Jahren mindestens anhängige Frage ist dann also: Was ist Aufmerksamkeit? Wie gelingt es mir, zuzuhören? Was ist mithin wichtiger, das «In-den-Fokus-Nehmen» – wie hilflos unsere visuell geprägte Sprache hier ist – eines auditiven Objekts oder doch eher das erfolgreiche Suprimieren, Unterdrücken, Herausfiltern jener uns eben nun nicht interessierenden, uns vor allem auch potentiell ablenkenden Höreindrücke? Sodann: Welche Macht haben erstens die akustischen Signale selbst, unsere Aufmerksamkeit zu ergattern, und warum gelingt dies manchen Geräuschen oder Geräuschquellen zuverlässiger als anderen? («Salienz» ist der hier oft gebrauchte Fachterminus, und «bottom-up», also von unten oder außen her ist ihre Wirkrichtung auf unseren Zuhörapparat.) Und wie stark oder wie komplett gelingt es, zweitens, einer ZuhörerIn, diesen Filter zu «implementieren», also ein-, aber auch durchzusetzen – mitunter also trotz der hohen Salienz eines zu ignorierenden Signals (z. B. einer Unterhaltung am Nachbartisch)? («Top-down», also von oben nach unten/außen, muss die ZuhörerIn ihren Filter hier einsetzen).

Die meisten der hier behandelten Fragen zum Verhältnis von Hören und Sehen, von Zuhören und Betrachten sind allesamt psychologische Fragen, also zunächst Fragen des menschlichen Erlebens und Verhaltens. Sie lassen sich auch als solche mit Methoden der experimentellen Psychologie adressieren, wenn nicht gar erschöpfend beantworten: mit aufwändigen Versuchsaufbauten, mit sogenannten Stör- und Hinweisreizen, mit gemessenen Reaktionszeiten und Fehlerraten, mit mathematischen Modellen unserer Wahrnehmungs- und Entscheidungsprozesse. Und wenn die Aufmerksamkeit rein auditiv aufgefasst und vermessen wird, dann involviert dieses Reizmaterial selbstredend akustische Reize – Musik, Sprache, Töne, Rauschen. Die berühmte «Cocktail Party» – längst nach der

Mitte des zwanzigsten Jahrhunderts wieder aus der Mode gekommen – stand in den goldenen Jahrzehnten der Aufmerksamkeitsforschung nach dem zweiten Weltkrieg hier Pate: Wie gelingt es uns, das Gemurmel und Gekirre um uns so gut zu ignorieren und stattdessen fokussiert unserem Gegenüber im Gespräch zu folgen? Aber mindestens so nützlich und noch eindrücklicher, wie gelingt es uns, eigentlich nun doch unauffällig dem neuesten Tratsch im Gespräch hinter uns zuzuhören statt dem langweiligen Monolog der Geschäftspartnerin vor einem? (Cherry, 1953) Es war also das Hörsystem als Modell, in dem die damals und teilweise bis heute zentralen Modellvorstellungen der Aufmerksamkeit (z. B. das Flaschenhals-Modell nach Broadbent; für eine Einführung siehe Serences & Kastner, 2014) entstanden.

Interessanterweise hat sich sodann auch in den folgenden Jahrzehnten etwas vollzogen, was man einen «visual turn» in der Aufmerksamkeitsforschung nennen könnte (wenn nicht ohnehin der Sehsinn im Grunde seit jeher die Psychologie und die Hirnforschung dominiert hätte). Und so wissen wir heute sehr viel besser, wie der Sehsinn und die Transduktion von visueller Information in neurale Signale funktioniert und wie das sehende Gehirn in den verschiedenen Spezies und auch insbesondere im Menschen sich organisiert. Wie unsere Aufmerksamkeit Objekte ins «Scheinwerferlicht» (so der Terminus technicus des attentional spotlight) nimmt und welche Rolle dabei Merkmale dieses Objekts oder sein Ort im Raum spielen und welche Rolle ablenkende Reize, dies ist erstaunlich gut verstanden. Im Auditiven hingegen ist schon die Formation eines auditiven Objekts – anhand welcher physikalischen Merkmale werden Schallreize als zusammengehörig gruppiert? – nicht unstrittig. Und so erschweren also die Physik des akustischen Reizes und die Beschaffenheit des an ihn angepassten sensorischen Epithels schon von der ersten Wandlung von Schall in elektrische Signale an unser wissenschaftliches Verständnis.

Zudem wurde diese Erforschung der verschiedenen Sinnesmodalitäten lange rigoros getrennt betrieben, in sehr reduktionistischen Forschungsparadigmen. Diese verloren – vielleicht notwendigerweise – zeitweise aus den Augen (und Ohren), dass Wahrnehmung und Aufmerksamkeit aber in einer uns natürlicherweise multi-sensorisch stimulierenden Umwelt erfolgen. Und erst in den letzten circa zwanzig Jahren, auch dank erheblicher Fortschritte in den mathematisch-statistischen Methoden zur Auswertung, findet auch in den psychologischen und neurowissenschaftlichen Studien wieder zusammen, was für das Individuum quasi nur gemeinsam erlebbar wird: Hörendes Zusehen oder sehendes Zuhören.

## II. Hörendes Zusehen, Sehendes Zuhören

Die multi-sensorische Verarbeitung von Umweltreizen ist also zunächst die einzig für den Menschen denkbare Verarbeitung, auch wenn das reduktionistische Forschungsparadigma dies aus guten Gründen oft hintan stellt. Spätestens aber sobald ein rudimentäres Verständnis darüber erreicht ist, wie zum Beispiel ein Areal der Hörrinde auditive Reize enkodiert, fragen wir uns: Was ändert sich eigentlich für Neuronen dieses Hörrindenareals, wenn nun auch begleitend dazu visuelle Reize eintreffen? Integrieren sie die zusätzliche, zeitgleiche oder zumindest zeitlich eng koordinierte Präsenz visueller Reize? Wie «früh» im Sinne der Verarbeitungshierarchie (Innenohr, Stammhirn, Zwischenhirn, primäre Hörrinde, usf.) erfolgt eine folgenreiche, d. h. den neuronalen Code wirklich verändernde Wirkung solcher multi-sensorischer Signale?

Die hier mithin interessanteste Frage ist derzeit, wie es uns (oder reduktionistischer: unserem Gehirn) überhaupt gelingt, überhaupt zu schließen, dass emittierter Schall und reflektiertes Licht von ein und demselben Objekt ausgehen, z. B. der zweiten Geigerin dort vorn, deren Körperbewegungen so gut zum Klang ihres Instruments passen. Wie leicht zu stören dieser Prozess der sogenannten kausalen Inferenz ist («Gibt es eine gemeinsame Ursache oder Quelle meiner auditiven und visuellen Sinneseindrücke?»), wird uns bewusst, wenn wir Bruce Springsteen und die E-Street-Band im Olympiastadion Berlin in hunderten Metern Entfernung sehen und der Schalldruck uns erst einige Hunderte Millisekunden später erreicht, oder wenn eine Audio-Spur sich schnell als offensichtlich verschoben zum Video herausstellt auf einer Plattform wie Youtube. Beides sind aber auch von unserem Gehirn sehr schnell wieder «reparierte» Missstände, nach wenigen Minuten bei gleichbleibenden Bedingungen merken wir oft nichts mehr von solchen Verschiebungen und nehmen wieder eine gemeinsame Quelle oder Ursache dieser audiovisuellen Stimulation wahr. Dieses Gemeinsame wird also (re-)konstruiert, und so falsch liegt also ein reduktionistischer Blick nicht, der zunächst das Trennende, das der auditiven wie der visuellen neuronalen Verarbeitung zunächst Eigene beschreibt.

Ob es bei dieser Konstruktion von Wahrnehmung ein Primat des Visuellen gibt? Einige Befunde legen dies nahe, aber zunächst haben die schiere Physik von Sehen und Hören hier ihren Anteil: Licht ist schneller als Schall. Das Abbild des Mannes dort vorn auf der Bühne mit seiner Telecaster-Gitarre erreicht meine Netzhaut und auch meine Sehrinde zwangsweise schneller als der Schall seines Gitarrenverstärkers mein Innenohr und meine Hörrinde. Als «natürlich» empfinden wir entsprechend auch ein dem Gehörten um Millisekunden vorausseilendes Gesehenes. Dieser zeitliche Versatz ist es auch, der mir als

Wahrnehmender erlaubt, Vorhersagen über das zu erwartende auditive Signal aufzustellen anhand des visuellen. Dieses Predictive Coding, also vorhersagende Codieren von sensorischer Information durch eine abgleichende Integration eines internen «generativen Modells» («Was erwarte ich eigentlich an auftretenden Sinnesreizen?») einerseits und des lediglich davon Abweichenden, das wirklich Neuheitswert hat und etwas technisch als «Vorhersagefehler» bezeichnet wird, andererseits – diese Vorstellung vom Ablauf auch multisensorischer Wahrnehmungsprozesse beherrscht nicht nur den neurowissenschaftlichen Diskurs unserer Zeit (Friston, 2010), sondern ist seit Hermann von Helmholtz de rigueur.

Beispielhaft betrachtet seien die Lippen-, Zungen-, und Kieferbewegungen einer Sprecherin, auf die ich unwillkürlich schaue in der natürlichen Kommunikation, und die Aufschluss darüber bieten, welche akustischen Merkmale sogleich an mein Ohr dringen werden und welche bereits jetzt ob dieser visuellen Information gänzlich unwahrscheinlich geworden sind. Der berühmt gewordene, wenn auch in der Praxis nicht ganz so reliabel zu reproduzierende McGurk-Effekt (McGurk & MacDonald, 1976) sei illustrierend hier genannt: Das Sehen eines Mundes, der eine Silbe produziert («ga»), während eine andere Silbe erklingt («ba»), lässt uns eine dritte Silbe wahrnehmen («da»). Das Fehlen des für «ba» charakteristischen Lippenschlusses schließt die eigentlich erklingende Silbe als Kandidat bereits aus, und das stattdessen gehörte «da» stellt phonetisch betrachtet eine Art Kompromiss dar zwischen der widerstreitenden visuellen und auditiven Evidenz. Eine kontemporäre und etwas elaboriertere Variante dieses Effekts aus dem Jahre 2020, das Spielerische der Sozialen Medien sowie die Macht der Schriftsprache auf unsere Wahrnehmung auskostend, findet sich nach kurzer Internet-Recherche der Wörter «green needle» und «brainstorm»: Eine junge Frau schaut in die Kamera und deutet uns mit der Hand an, jeweils eine der beiden über ihr gezeigten Phrasen zu lesen, eben links oben «green needle» und rechts oben «brainstorm» – erstaunlicherweise wird dazu in einer Endlosschleife stets dasselbe, leicht an Science-Fiction-Fernsehmusik der 1960er-Jahre gemahnende Geräusch abgespielt. Und tatsächlich «hört» unser Gehirn mühelos aus jenem stets identischen, wenn auch seltsamen Klang das jeweils gelesene Wort recht eindeutig heraus. Was ist geschehen? Zauberei, ein alchemistischer Trick? Mitnichten; hier wurde eine «akustische Chimäre» erzeugt (Smith et al., 2002), die tatsächlich akustische Merkmale zweier Aufzeichnungen derart vereint, dass diese Kombination an und für sich (also rein akustisch) für uns nicht mehr zu verstehen ist. Wiederum kann dann aber der Sehsinn (hier, das Gelesene) eine Art von Wahrnehmungs-Apriori bereitstellen – anhand dieses Leitsignals

können wir sodann die hierzu passende auditive Teilinformation durchaus extrahieren und nutzen, sprich, wir «verstehen».

In besonderem Maße produktiv für unsere Vorstellungen davon, was multisensorische Wahrnehmung wirklich leisten kann, wird es, wenn wir auf die hier immer wieder behandelten Fragen von Aufmerksamkeit und ihrem Gelingen zurückkommen: Wenn ich aufmerksam hören, also zuhören kann, und aufmerksam schauen, also betrachten – wie tue ich dies dann beides auf einmal? Und, kann ich das überhaupt? Wenn visuelle Aufmerksamkeit rhythmisch ist, also die Abbilder unserer Retina unsere Bewusstheit nur in getakteten Zeitfenstern priorisierter Verarbeitung erreichen (vgl. z.B. Fiebelkorn & Kastner, 2019; VanRullen & Koch, 2003) – was bedeutet dies für das zeitgleiche Zuhören? Solange es um jenen selben zweiten Geiger geht, dem man lauscht und dessen Körperbewegungen man folgen will, können wir sicherlich von einem sogenannten supramodalen, also über die Sinnesmodalitäten hinweg einheitlich agierenden Aufmerksamkeitspuls ausgehen. Was aber, wenn unser Hörsinn einerseits und unser Sehsinn andererseits auf unterschiedliche Objekte oder Quellen achten sollen, achten wollen? Was für uns zunächst im Alltag wenig Nutzen zu erbringen scheint (warum den Bassisten genau beobachten und dabei der Cellistin lauschen?), gelingt Probanden im Bachelor-Alter im psychologischen Labor problemlos.

Und auch in unseren alltäglichen Lebenswelten lassen sich zahlreiche, in ihrer Natur auditiv-visuelle Kommunikations- oder Stimulationssituationen finden. Das Kind lauscht gebannt einer Geschichte von Artus und den Rittern der Tafelrunde, während es mit Auge und Hand ein Elfenreich auf dem Papier erschafft. Die Wissenschaftlerin am Klinikum telefoniert angeregt über Fachliches mit einer Kollegin und betrachtet dabei durchaus zufrieden das Kommen und Gehen der Lieferanten, Paketboten, Krankenwägen vor ihrem Büfenster. In derartigen Situationen finden wir eine Flüssigkeit und Leichtigkeit des Aufmerksamkeitswechsels zwischen den Modalitäten vor, der den Menschen weit weniger eindeutig als «Augentier» erscheinen lässt.

### **III. Gibt es eine neue Auralität?**

Ohne Zweifel hängen Hören und Zuhören Attribute des Gemäßigten, des Ruhigen, des Kontemplativen an. Sowohl das oben beschriebene malende Kind, das einer Geschichte lauscht, wie auch die versonnen aus dem Fenster schauende Telefonierende erscheinen uns eher versunken denn agitiert. Aber dies ist natürlich ein Klischee. Der Mensch weiß auch das

gesprochene Wort zu konsumieren, zu extrahieren, zu verwerten und auszuschlachten. Meine Studierenden im Pandemiejahr 2020 lauschen meinen Statistik-Lehrvideos gerne in 1,5-facher Geschwindigkeit, ein gegenüber dem echten Vortrag im Hörsaal sicher unschätzbare Vorteil. Umgekehrt ist es vielen Menschen zur liebsamen Gewohnheit geworden, ihre Mitteilungen an andere als kurze Sprachnachrichten zu versenden – nach einer Phase der Mobilkommunikation, in der man meinte, das gesprochene Wort sterbe aus zugunsten von Textnachrichten. Dieses Sprechen geht in der Tat schneller als die händische Übersetzung unserer Gedanken in die Kulturtechnik der Schriftsprache. Dies ist aber ein Pyrrhussieg für den Schall und die gesprochene Sprache, denn es ist nun an der Rezipientin, sich ebenjene Sprachnachrichten in Gänze, quasi wieder linearisiert, anhören zu müssen, statt sie in halsbrecherischer Geschwindigkeit mit den Augen «scannen», also: lesen zu können, wie bei Emails oder Textnachrichten möglich.

Und die vielleicht spannendste, sicher aber neueste Volte dieses medialen Hin und Hers ereignete sich just ebenso im Pandemie-Jahr 2020, als nebst dem erstaunlich populären Medium des sogenannten Podcasts, also des Austauschs gesprochener Worte als Information und Unterhaltung (dem guten alten Radio nicht unähnlich), eine neue (und im Herbst 2021 sei nachgetragen: allem Anschein nach sehr kurzlebige) Social-Media-App namens Clubhouse aufschlug: Nicht geschrieben oder gezeigt wurden dort die eigenen Ausdrücke, sondern es wurde nur – gesprochen. Ein in Gänze vom Bild befreites soziales Medium, eine Art panel, call-in show oder gar talk radio für alle (oder zumindest für jene mit Einladung; ein Clubhaus eben). Die abschließende Einordnung dieser vermeintlich neuen oder zumindest erfrischten auralen Phänomene, und ob sie tatsächlich die Rückkehr oder Re-Affirmation des Auditiven in einer zunächst immer visuelleren Welt bedeuten, sei gerne den Medientheoretikerinnen überlassen an dieser Stelle – aber als Hörforscher darf ich es zumindest der Bemerkung wert finden.

## **VI. Wie höre ich zu? Zuhören als Zustand und Merkmal**

Allgemeinpsychologisch ist es uns daran gelegen, die Gesetzmäßigkeiten zu beschreiben, die das Verhältnis von auditiver und visueller Wahrnehmung und Aufmerksamkeit bestimmen. Differentialpsychologisch aber würden wir fragen: Erstens, warum gelingt es der einen morgens besser, zuzuhören, und dem anderen eher im Auto auf dem Weg nach Hause am Abend? («Zustände» oder sogenannte «states»). Dem gehen wir in verschiedenen Laboruntersuchungen nach, in dem wir die momentane Hirnaktivität messen

und beschreiben und versuchen, daraus beispielsweise die Encodierungstiefe des sodann Gehörten mathematisch vorherzusagen (Alavash et al., 2019; Waschke et al., 2019).

In Alavash, Tune und Obleser (2019) maßen wir beispielsweise die momentane Vernetztheit verschiedener Hirnareale, z. B. klassisch auditiver Areale im Schläfenlappen des Gehirns, mit jenen, denen wir eher kontrollierende oder steuernde Funktionen zuschreiben im Stirn- und im Hinterhauptlappen. Unsere Probanden unterzogen sich einer anstrengenden Zuhöraufgabe mit zwei konkurrierenden Stimmen, während wir die Vernetztheit des gesamten Gehirns als mehr oder weniger synchronisierte Fluktuationen der Sauerstoffsättigung all dieser Gehirnareale abbildeten (dies gelingt mit der sogenannten funktionellen Kernspintomographie, fMRT). Die Fähigkeit eines Zuhörers, gleich einer plötzlichen Planänderung in einem großen Logistiknetz wie einer Luftfahrtgesellschaft, diese neuronalen Kommunikationswege aus einem Zustand der relativen Ruhe heraus anzupassen und zu re-konfigurieren, wenn diese schwierige Zuhörsituation eintrat, erwies sich als ein interessantes Maß: Es sagte vorher, wie erfolgreich (wie korrekt, und wie schnell) der einzelne Proband eine solche «Cocktail-Party»-artige Höraufgabe bewältigte.

Zweitens aber auch, warum gelingt es ganz generell (bzw. im Durchschnitt aller Lebenslagen) Ihnen vielleicht besser als Ihrem Partner, in schwierigen Hörsituationen zurechtzukommen – im Restaurant, am Autotelefon, auf dem Bahnsteig. Die wäre ein Hinweis auf überdauernde «Merkmale» oder «traits», die dann zwei Menschen unterscheiden helfen. Auch hier interessiert uns, wie wir die Vielfalt des Zuhörens aufklären können. Natürlich sind hier alle Merkmale potentiell relevant, die mich von Ihnen, den einen von der anderen unterscheiden: Zuvorderst unser einfaches Hörvermögen im audiometrischen Sinne und damit eng zusammenhängend unser biologisches Alter.

Aber eine solche Unterscheidung unserer Zuhörleistungen kann erstaunlicherweise auch durch Merkmale unserer Persönlichkeit geschehen. So neigen emotional labilere Menschen dazu, ihre eigene Störanfälligkeit durch Geräusche zu überschätzen; objektiv «performen» sie dann aber eher überdurchschnittlich in solchen Situationen, wie wir derzeit in einer groß angelegten Studie mit über 1000 Proband\*inn\*en zeigen [Wöstmann et al., in Begutachtung]. Die Implikationen solcher Persönlichkeitsauswirkungen auf unser Hörerleben sind nicht zu unterschätzen, wenn wir Menschen individualisierte Hör- und Wahrnehmungshilfen im Alter anbieten wollen und auch, wenn wir der relativen Relevanz von Hören und Sehen für die Einzelne in Zukunft stärker Rechnung tragen wollen.

## V. Nachtrag: Das Gehirn als Metapher oder als letzter Grund?

Je nach eigenem ideentheoretischem Standpunkt (vgl. Miller, 2010) reizt und lockt uns neben, unter oder über dieser sehr fruchtbaren psychologischen Beschreibungsebene stets auch eine biologisch-neurowissenschaftliche – eine Betrachtungsweise also, die ansetzt, das Zuhören und mithin die auditive Aufmerksamkeit zuvorderst als Leistung des Gehirns zu beschreiben und zu erklären. Dies bringt seine eigenen Probleme mit sich. Erstens, ohne die psychologischen und informationstheoretischen Begriffe (Aufmerksamkeit, Filter, usw.) geht es schlechterdings kaum (Obleser & Kayser, 2019). Zweitens, die Pathologien des Hörens und Zuhörens sind in Gefahr, besonders schnell auch «biologisiert» zu werden – als wäre mit einem auch noch so glasklaren biologischen Abbild bereits das psychologische und oft zwischen-menschliche Phänomen verstanden. Dies betrifft z. B. den Tinnitus oder die sogenannten verbalen Halluzinationen, aber auch die vieldiagnostizierte «zentrale Hörstörung», wie sie insbesondere Kindern und Jugendlichen anhängt und wie sie unspezifischer kaum benannt sein könnte (mit «zentral» ist hier lediglich bezeichnet, dass es keine Störung des Innenohres ist). Ein erster Schritt für die einzelne Forscherin liegt sicherlich darin, sich der eigenen Metaphern und Begrifflichkeiten gewahr zu werden, wenn man ansetzt, menschliches Hören und Zuhören in diesen unterschiedlichen Domänen – z. B. geht es mir gerade um den audiologisch-psychischen «Algorithmus» der audiovisuellen Integration oder um ihre neurobiologisch umgesetzte «Implementation»? – beschreiben und erklären zu wollen (vgl. Marr, 1982).

Zusammengefasst haben sowohl die genuin psychologische Beschreibung und Erklärung des Zuhörens wie auch die neurobiologische Untersuchung desselben in den vergangenen Jahrzehnten uns ein Instrumentarium erbracht, mit dem wir auch zukünftig das relative und veränderliche Gewicht werden vermessen können, das unser Hörsinn im Orchester unserer Sinneswahrnehmung und unserer Kognition einnimmt – gerade unter sich verändernden Bedingungen von Kommunikation und medialer Vermittlung.

### Literatur

Mohsen Alavash, Sarah Tune & Jonas Obleser: «Modular reconfiguration of an auditory control brain network supports adaptive listening behavior», in: *Proceedings of the National Academy of Sciences of the United States of America* 116, 2 (2019, Jan 8) DOI: 10.1073/pnas.1815321116

E. Colin Cherry: «Some experiments on the recognition of speech, with one and two ears», in: *Journal of the acoustical Society of America* 25, 5 (1953), S. 975-979.

Ian C. Fiebelkorn & Sabine Kastner: «A Rhythmic Theory of Attention», in: *Trends in Cognitive Sciences* 23, 2 (2019 Feb.), S. 87-101. DOI: 10.1016/j.tics.2018.11.009

Karl Friston, «The free-energy principle: a unified brain theory?» in: *Nature Reviews Neuroscience* 11, 2 (2010 Feb.), S. 127-138. DOI: 10.1038/nrn2787

William James, «Attention», in: ders.: *Principles of Psychology*, Bd. 1, New York 1890/1927, S. 402-458.

David Marr: *Vision: a computational investigation into the human representation and processing of visual information*, New York 1982.

Harry McGurk & John MacDonald: «Hearing Lips and Seeing Voices», in: *Nature* 264, (1976), S. 746-748. DOI: 10.1038/264746a0

Gregory A. Miller: «Mistreating Psychology in the Decades of the Brain», in: *Perspectives of Psychological Science* 5, 6 (2010, Nov.), S. 716-743. DOI: 10.1177/1745691610388774

Jonas Obleser & Christoph Kayser: «Neural Entrainment and Attentional Selection in the Listening Brain», in: *Trends in Cognitive Sciences* 23, 11 (2019, Nov.), S. 913-926. DOI: 10.1016/j.tics.2019.08.004

John T. Serences & Sabine Kastner: «A Multi-level of Selective Attention», in: Anna C. Nobre & Sabine Kastner (Hg.): *The Oxford Handbook of Attention*, New York 2014. DOI: 10.1093/oxfordhb/978019967511.013.022

Zachary M. Smith, Bertrand Delgutte & Andrew J. Oxenham: «Chimaeric sounds reveal dichotomies in auditory perception», in: *Nature* 416 (2002, March 7), S. 87-90. DOI: 10.1038/416087a

Rufin van Rullen & Christof Koch: «Is perception discrete or continuous?» in: *Trends in Cognitive Sciences* 7, 5 (2003, May), S. 207-213. DOI: 10.1016/s1364-6613(03)00095-0

<https://twitter.com/ThePrisonLawyer/status/1347626960529326080?s=20>; abgerufen am 24.03.2021

Leonhard Waschke, Sarah Tune & Jonas Obleser: «Local cortical desynchronization and pupil-linked arousal differently shape brain states for optimal sensory performance», in: *eLife* 8 (2019, Dec 10), DOI: 10.7554/eLife.51501