

Der Balanceakt des räumlichen Gehörs

Robert Baumgartner

Das räumliche Gehör ermöglicht es, Schallquellen von Interesse zu lokalisieren und selektiv auf sie zu achten. Es erleichtert somit die Kommunikation in komplexen Umgebungen und stellt gleichzeitig einen wichtigen Warnmechanismus dar. Das Gehör schließt aus mehreren akustischen Hinweisen auf die räumliche Lage einer Quelle und nimmt diese als externalisiertes Hörobjekt wahr. Vorwissen fließt ein, um die Wahrnehmung robuster zu gestalten, wobei deren Reliabilität berücksichtigt werden muss, um dennoch anpassungsfähig zu bleiben. Auch Lernprozesse müssen so reguliert sein, dass Robustheit und Anpassungsfähigkeit gleichermaßen gewährleistet sind. Es gibt also eine Reihe an teils gegenläufigen Mechanismen, die es im Zusammenhang mit räumlichem Hören zu verstehen gibt. Inwiefern unnatürliche akustische Gegebenheiten das Zusammenspiel beeinträchtigen können, wird an Beispielstudien zur Externalisierungswahrnehmung und zu räumlich selektiver Aufmerksamkeit aufgezeigt.

Einleitung

Wir sind häufig komplexen akustischen Umgebungen mit mehreren Schallquellen ausgesetzt. Räumliches Hören ermöglicht uns dabei einzelne Schallquellen zu lokalisieren und unsere Aufmerksamkeit gezielt auf sie zu richten. Die Aufmerksamkeit kann bewusst von uns gelenkt werden, um beispielsweise einem Gespräch zu folgen, sie kann aber auch durch bestimmte Reize gelenkt sein, um uns beispielsweise vor Gefahrenquellen zu schützen. Man unterscheidet demnach zwischen endogener oder auch absteigender und exogener oder aufsteigender Aufmerksamkeitslenkung. Ein stimmiges Zusammenspiel dieser beiden gegenläufigen Prozesse ist wichtig, um lokale Fokussierung und globale Überwachung gleichzeitig zu realisieren. Darüber hinaus ist es wichtig, Robustheit und Anpassungsfähigkeit von räumlichen Repräsentationen gleichermaßen zu bewerkstelligen, sowohl hinsichtlich akuter als auch langfristiger Veränderungen von akustischen Gegebenheiten. Wie die Balance auf unterschiedlichen Ebenen stattfinden kann, wird in weiterer Folge genauer beleuchtet.

Räumliche Inferenz

Sowohl Lokalisierung als auch räumliche Aufmerksamkeitslenkung basieren auf der Fähigkeit, räumli-

The balancing act of spatial hearing

Spatial hearing makes it possible to locate sound sources of interest and selectively attend to them. It thus facilitates communication in complex environments while providing an important warning mechanism. Hearing infers the spatial location of a source from multiple acoustic cues and results in the perception of an externalized auditory object. In order to make perception more robust, prior knowledge is incorporated, and prior reliability must be taken into account in order to remain adaptive. Learning processes must also be regulated in such a way that robustness and adaptability are equally guaranteed. Thus, there are a number of partly counteracting mechanisms to be understood in the context of spatial hearing. The extent to which unnatural acoustic conditions can interfere with this interaction will be illustrated by example studies on externalizing perception and spatial selective attention.

che Information aus dem Gehörten zu extrahieren. Im Sinne von perzeptiver Inferenz versucht das Gehirn basierend auf sensorischer Information und Vorerfahrung die unbekanntesten Zustände in der Umgebung zu ermitteln. Die angenommenen Zustände bilden ein internes Modell der Umgebung [1]. Dieses Modell kann wiederum verwendet werden, um Vorhersagen zu generieren und damit eine möglichst effiziente und kontrollierte Interaktion mit einer sich ständig veränderbaren Umgebung zu ermöglichen [2]. Illusionen ergeben sich als logische Konsequenz der Modellannahmen. Rein reaktive sensorische Verarbeitung wäre langsam und ineffizient. Vorhersagegetriebene Wahrnehmung ist hingegen schnell und bei guter Vorhersage insofern effizient, dass immer nur der Vorhersagefehler, nicht aber die gesamte sensorische Information ausgewertet werden muss, um sich Veränderungen anzupassen [3].

Robustheit und Anpassungsfähigkeit

Durch die gleichzeitige Berücksichtigung von Vorwissen und akuter sensorischer Evidenz lässt sich Robustheit erhöhen und Unsicherheit reduzieren. Statistisch betrachtet lässt sich die optimale Kombination dieser beiden Komponenten durch den Satz von Bayes formalisieren (siehe Abbildung 1a).

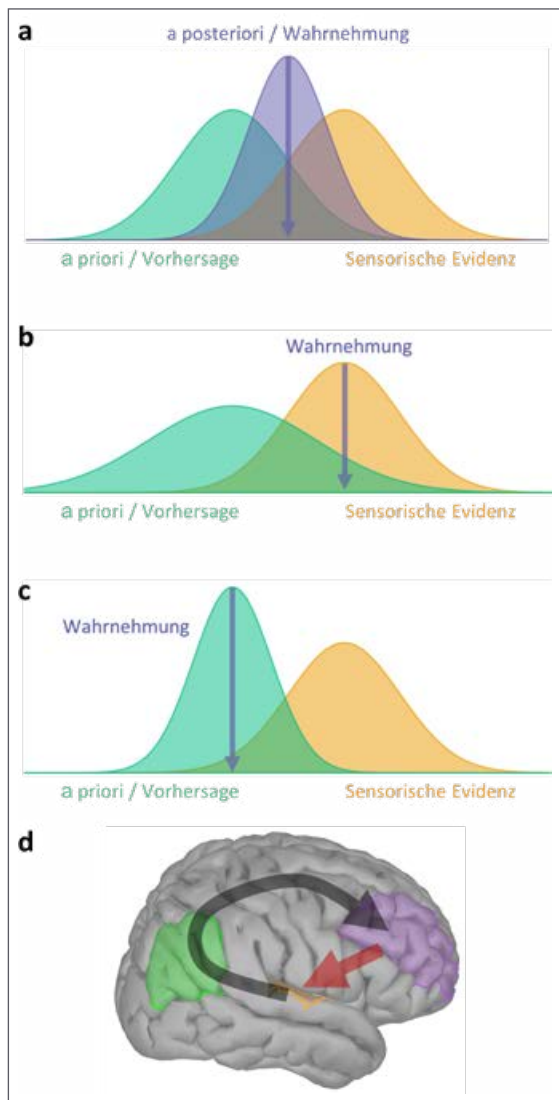


Abb. 1: Probabilistische Ansätze zur Entscheidungsfindung und das zugehörige kortikale Netzwerk für räumliches Hören.

a) Statistisch optimale Wahrnehmungsentscheidung durch Bayessche Auswertung von a-posteriori-Wahrscheinlichkeiten mittels Multiplikation von Wahrscheinlichkeitsverteilungen der sensorischen Evidenz und a-priori-Information.

b+c) Statistische Selektion zwischen sensorischer Evidenz und a-priori-Information, abhängig von deren relativen Reliabilitäten.

d) Kortikales Netzwerk bestehend aus primärem auditiven Kortex und Planum Temporale im oberen Temporallappen (gelb), sensorischen Assoziationszentren im unteren Parietallappen (grün) und Zentren zur Handlungsplanung im vorderen Frontallappen (violett). Diese Areale sind über einen dorsalen Verarbeitungspfad verbunden (schwarzer Pfeil) und es gibt mehrere direkte und indirekte Pfade, über die zielgerichtete Anforderungen zurück an die sensorischen Areale übermittelt werden können (roter Pfeil).

Die Berechnung dieser Lösung erfordert zu jedem Auswertungszeitpunkt eine Multiplikation zweier Wahrscheinlichkeitsverteilungen, eine für die sensorische Evidenz und eine für das gesammelte Vorwissen (a-priori-Wahrscheinlichkeit). Da diese Berechnung zu aufwändig erscheint, um tatsächlich neuronal implementiert zu sein, ist man auf der Suche nach heuristischen Algorithmen, die diese Entscheidungsfindung effizienter erklären können [4]. Eine alternative, effizientere Entscheidungsstrategie wäre eine dynamische Selektion, die je nach relativer Reliabilität entweder die sensorische Evidenz (Abb. 1b) oder das Vorwissen (Abb. 1c) auswählt. Es existieren bereits Studien, z. B. [5], die dieses dynamische Verfahren unterstützen und aufzeigen, dass dadurch ein nahezu optimales Verhalten erzielt werden kann. Diese Studien beschäftigen sich aber fast ausschließlich mit dem Sehen, über das Hören weiß man vergleichsweise wenig.

Der Robustheit steht die Anpassungsfähigkeit gegenüber, um rasch auf Veränderungen in der Umgebung reagieren zu können. Auch unter diesem Gesichtspunkt hätte ein dynamisches Selektionsprinzip einen Vorteil in der Verarbeitungslatenz, denn starke sensorische Evidenz beziehungsweise ein großer Fehler in der Vorhersage dieser Evidenz kann hier den Einfluss von Vorwissen sofort eliminieren. Neben dieser akuten Wahrnehmungsentscheidung stellt sich auch die Frage, wie das Gehirn mit Mehrdeutigkeiten aus der Vergangenheit umgeht. Im Sinne des Bayesschen Ansatzes können hierfür alle potentiellen Möglichkeiten berücksichtigt werden oder aber nur eine oder wenige dominierende perzeptive Hypothesen verfolgt werden [6]. Es existiert also eine große mögliche Bandbreite hinsichtlich der Komplexität des berücksichtigten Vorwissens bzw. der internen Vorhersage, welche der sensorischen Evidenz gegenübergestellt wird.

Auf neuronaler Ebene werden diese Prozesse vermutlich durch ein Netzwerk mehrerer Gehirnareale realisiert (Abb. 1d) [7]. Die sensorische Evidenz wird zunächst vom auditiven Kortex und dem angrenzenden Planum Temporale im oberen Teil des Temporallappens verarbeitet (gelb). Entlang einem dorsalen Pfad folgen dann sensorimotorische Areale im Parietallappen, die als sensorische Assoziationszentren vermutlich das interne Modell und die daraus resultierenden Vorhersagen generieren (grün). Die motorischen Areale im oberen und vorderen Frontallappen bilden das Ende dieses Pfades und dienen vermutlich der Entscheidungsfindung (violett). Es gibt mehrere Möglichkeiten, wie ausgehend vom präfrontalen Kortex aufgabenspezifische Anforderungen an die sensorischen Areale rückgemeldet werden könnten (roter Pfeil). Es sind sowohl Verbindungen zum Planum Temporale und auditiven

Kortex als auch zu subkortikalen und limbischen Arealen bekannt. Somit ergibt sich ein komplexes rückgekoppeltes Gesamtsystem, das unter Einbezug der emotionalen Funktionseinheiten im limbischen System zielgerichtet in die Verarbeitung der sensorischen Evidenz einwirken kann.

Sensorische Evidenz

Welche Arten von sensorischer Evidenz stehen aber nun dem räumlichen Gehör zur Verfügung, um dieses Inferenzproblem zu lösen? Hinsichtlich der akustischen sensorischen Information kann zwischen zwei Arten von primären Richtungsmerkmalen unterschieden werden. Dabei handelt es sich zum einen um interaurale Unterschiede im Schalldruckpegel (Interaural Level Difference, ILD) und Ankunftszeitpunkt (Interaural Time Difference, ITD), die durch die räumliche Trennung der beiden Ohren verursacht werden. Zum anderen geht es um spektrale Richtungsprofile, die durch jegliche Reflexionen der eintreffenden Schallwelle am Körper erzeugt werden, bevor sie das Trommelfell erreichen. Eine sehr markante Richtcharakteristik bewirkt hierbei die menschliche Pinna im hohen Frequenzbereich ab ca. 5 kHz.

Die primären Richtungsmerkmale werden relativ zur Orientierung des Kopfes registriert. Sie befinden sich also zunächst in einem egozentrischen Bezugsrahmen mit dem Kopfmittelpunkt als Ursprungskordinate (siehe Abbildung 2). Der binaurale Übertragungspfad für eine Quellposition wird im Englischen folglich als Head-Related Transfer Function (HRTF) bezeichnet. Interaurale Unterschiede geben Aufschluss über den Lateralwinkel der Quelle, welcher Richtungen zwischen links und rechts erfasst, nicht aber über die Distanz und den Polarwinkel, welcher sich über die Richtungen oben und unten sowie vorne und hinten erstreckt. Es ergibt sich demnach eine kegelförmige Oberfläche um die interaurale Achse zentriert – im Englischen als Cone of Confusion geläufig. Um diese räumliche Unsicherheit aufzulösen, können die spektralen Profile ausgewertet werden. Sie lassen jedoch nur auf die Distanz der Quelle schließen, sofern sie sich im akustischen und persönlichen Nahfeld befindet, also nicht weiter entfernt ist als etwa 1 m [8].

Die Verarbeitungsebenen dieser Merkmale sind recht unterschiedlich. Bereits im Hirnstamm werden interaurale Unterschiede frequenzspezifisch ausgewertet und spektrale Profile für die spätere Auswertung vorverarbeitet. Auf kortikaler Ebene existieren sowohl getrennte als auch kombinierte Repräsentationen der einzelnen Merkmale [9]. Entsprechend der Komplexität der Merkmalsextraktion ergeben sich auch unterschiedliche Verarbeitungszeiten. Veränderungen von interauralen Unterschieden wer-

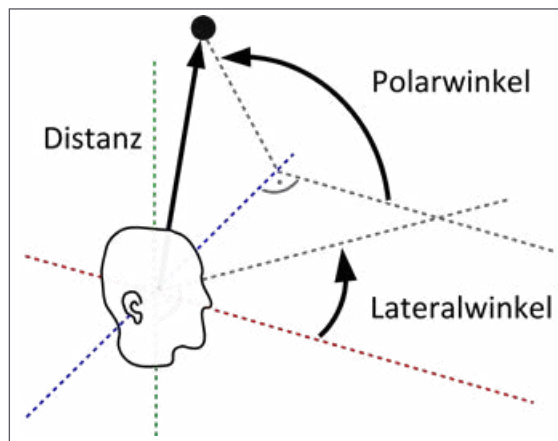


Abb. 2: Egozentrisches sphärisches Koordinatensystem mit Polen auf der interauralen Achse (blau). Der Lateralwinkel erstreckt sich von rechts (-90°) nach links ($+90^\circ$), der Polarwinkel von unten (-90°) über vorne (0°) und oben (90°) nach hinten (180°). Die Distanz wird ausgehend vom Zentrum des Kopfes gemessen.

den beispielsweise bereits nach 100 ms registriert, wohingegen die räumliche Zuordnung basierend auf spektralen Profilen in etwa die doppelte Verarbeitungszeit aufweist [10]. Die Auswertung spektraler Profile beruht folglich auf Assoziationen, die in höheren Ebenen verarbeitet werden. Neue Assoziationen dieser Art sind erlernbar, ohne mit bereits existierenden zu interferieren [11], ähnlich wie das Erlernen neuer Synonyme in der Sprache.

Überwachtes Lernen und statistisches Lernen sind zwei unterschiedliche Mechanismen, die neuronale Plastizität ermöglichen. Statistisches Lernen bezieht sich auf die implizite Fähigkeit – mit oder ohne Aufmerksamkeit – Regelmäßigkeiten zu extrahieren und zu repräsentieren, wie z. B. häufig auftretende Klangmuster oder häufige akustische Übergänge. Überwachtes Lernen bezieht sich hingegen auf die Fähigkeit, Hörereignisse auf der Grundlage expliziter Rückmeldungen aufmerksam zu kodieren. Klar ist, dass die neuronale Verarbeitung in frühen Entwicklungsstadien deutlich plastischer ist als bei Erwachsenen. Wie die beiden Lernmechanismen beteiligt sind und reguliert werden, um eine dem Entwicklungsstand angepasste Balance zwischen Anpassungsfähigkeit und Robustheit zu gewährleisten, ist ein weiteres spannendes Thema [12].

Das Externalisierungsproblem

Unter künstlichen Hörbedingungen kann es leicht zu Reizkombinationen kommen, die sich nicht mit gelernten Mustern in Einklang bringen lassen. Das auditorische System ist in diesem Fall nicht in der Lage, eine Abbildung in einen allozentrischen, also welt-zentrierten, Bezugsrahmen vorzunehmen, und es resultiert keine externalisierte Distanzwahrneh-

mung. Während eine Lateralisierung entlang der interauralen Achse auch mit höchst artifiziellen Schallreizen (wie beispielsweise Sinustönen) funktioniert, erfordert eine scharfe Lokalisation im dreidimensionalen Raum zumeist akustische Gegebenheiten, die mit Erwartungen basierend auf gewohnten Hörumgebungen in Einklang zu bringen sind.

Um quantifizieren zu können, wie stark die Externalisierbarkeit von einzelnen Merkmalen abhängt und wie diese kombiniert werden, wurden Modelluntersuchungen vorgenommen [13, 14]. Dabei hat sich gezeigt, dass um eine hohe Externalisierbarkeit zu erreichen, Assoziationen basierend auf spektralen Profilen nicht nur monaural sondern auch interaural erweckt und Erwartungen bezüglich des interauralen Nachhallverhaltens erfüllt werden müssen. Die relative Gewichtung zwischen diesen spektralen Assoziationen und der Nachhallerwartung verschiebt sich mit zunehmender Intensität des Nachhalls. Diese Verschiebung lässt sich damit begründen, dass Nachhall Unsicherheit in der Auswertung der spektralen Profile hervorruft und das auditorische System ganz im Sinne der Bayesschen Inferenz dieser Information somit weniger Wichtigkeit zuordnet. Alternative Modellierungsansätze, die sich situationsbedingt immer nur auf ein Merkmal stützen, konnten die Versuchsergebnisse weniger gut erklären als dieser graduelle Gewichtungsansatz.

Aktive Inferenz und andere Informationsquellen

Trotz der Vielzahl an zur Verfügung stehender Information kann es zu Unsicherheiten kommen. Das Gehirn versucht Mehrdeutigkeiten daher auch durch gezielte Aktionen aufzulösen, man spricht von aktiver Inferenz [15]. So ermöglicht beispielsweise eine Rotation des Kopfes zu unterscheiden, ob sich eine Quelle vor oder hinter uns befindet. Um sensorimotorische Information von Kopfbewegungen und komplexere Assoziationen miteinzubeziehen, müssen unterschiedliche Bezugsrahmen kombiniert werden. Auf Wahrnehmungsebene ist letztlich das Ziel, eine allozentrische Repräsentation zu erreichen, die stabil ist gegenüber eigenen Bewegungen.

Über die primären Merkmale hinaus werden viele allgemeine Assoziationen und relative Vergleiche mit einbezogen, die in einem allozentrischen Bezugsrahmen registriert werden. So stellt zum Beispiel die Schallintensität ein wichtiges Distanzmerkmal dar, obwohl sie absolut gesehen keine Auskunft über die Distanz gibt, sondern nur wenn ihre Änderung über die Zeit oder bezüglich einer internen Referenz für bekannte Stimuli ausgewertet wird. Ein weiteres Beispiel ist die Frequenz-Elevations-Assoziation, bei der höhere Töne mit höherer Elevation in Verbindung gebracht werden [16]. Dieser Zusammenhang

hat sich sogar in unserer Sprache manifestiert. Darüber hinaus können auch kausale Zusammenhänge die räumliche Wahrnehmung beeinflussen. Eine flüsternde Stimme wirkt bei gleicher Schallintensität zum Beispiel näher als eine schreiende Stimme [17], weil das Gegenteil schlichtweg unplausibel wäre. Selbst unter Ausschluss von anderen sensorischen Modalitäten wie dem Sehsinn wird also eine Vielzahl unterschiedlicher Informationsquellen verwendet, um die Position einer Schallquelle zu bestimmen.

Räumliche Aufmerksamkeitslenkung

Um lokale Fokussierung und globale Überwachung gleichzeitig zu realisieren, geht man von einer stetigen Konkurrenz zwischen Wahrnehmungsobjekten aus, die sowohl durch den Stimulus selbst als auch durch bewusst gesteuerte Aufmerksamkeitslenkung beeinflusst werden kann [18]. Zu einem Zeitpunkt dominiert immer ein Objekt und rückt sozusagen in den Vordergrund, während der Rest als Hintergrund verschwimmt. Eine Veränderung dieses Zustands, also die Aufmerksamkeit auf ein anderes Objekt zu lenken, ist mit Aufwand verbunden.

Gewisse Reize schaffen es dennoch, diese Veränderung hervorzurufen. Das Gehör verarbeitet zum Beispiel Reize, die auf eine Annäherung hinweisen könnten, mit besonderer Priorität. Dieser Effekt wird in der Fachliteratur als Looming Bias bezeichnet. Auf Verhaltensebene zeigt sich diese Priorisierung oder erhöhte Salienz durch schnellere und akkuratere Antworten in Diskriminationsaufgaben, Überschätzungen von Reizänderungen und sogar schnelleren Lernfortschritten bei klassischer Konditionierung. Diese Tendenzen treten bei unterschiedlichen Spezies auf und wurden auch bereits bei Neugeborenen mittels audiovisuellen Verhaltensexperimenten nachgewiesen. Auf neuronaler Ebene geht die Priorisierung einher mit erhöhter Aktivität ab ca. 100 ms nach dem Beginn der Bewegung [19]. Welches Netzwerk an Gehirnarealen dafür verantwortlich gemacht werden kann, ist noch nicht endgültig geklärt, aber bisherige Hinweise sind im Einklang mit dem zuvor beschriebenen Netzwerk aus temporalen, parietalen und frontalen Arealen [20] (vgl. Abbildung 1d) und dem Einfluss von limbischen Zentren wie der Amygdala [21]. Während die exogene Aufmerksamkeitslenkung als Schutzmechanismus erklärt werden kann, ist die endogene Steuerung von räumlich selektiver Aufmerksamkeit wesentlich, um in vielen sozialen Alltagssituationen kommunizieren zu können. Diese Situation stellt zugleich ein fundamentales Problem vieler Nutzer von Hörhilfen dar. Es wurde deshalb untersucht, inwiefern inkonsistente Richtungsmerkmale Aufmerksamkeitsmechanismen in einer Hörsituation mit zwei konkurrierenden, örtlich getrennten (ca.

$\pm 30^\circ$ Azimut) Sprechern beeinträchtigen [22]. Die zu beachtende Richtung wurde über einen räumlichen Hinweisreiz vorgegeben. Sowohl der Hinweisreiz als auch die Ziel- und Distraktorreize wurden räumlich entweder durch die natürlich konsistente Information von ITD und ILD in Form von individuell gemessenen HRTFs kodiert oder es wurde diese ITD oder ILD unnatürlich isoliert verwendet. Es zeigten sich deutliche Unterschiede zwischen natürlichen und unnatürlichen Repräsentationen in der neuronalen Verarbeitung. Nur unter der natürlichen Bedingung rief der Hinweisreiz in Vorbereitung zum Zielreiz eine signifikante Lateralisierung von parietaler alpha-Band-Aktivität (8–13 Hz) hervor, die als Indikator für räumlich selektive Aufmerksamkeitsprozesse allgemein bekannt ist. Ebenfalls nur unter der natürlichen Bedingung zeigte sich eine räumliche Selektivität in der Verarbeitung des Zielreizes mit einer Latenz von nur etwa 50 ms. Natürlich konsistente Richtungsmerkmale ermöglichen also eine sehr frühe Selektion von aufgabenrelevanter Information und es erscheint entsprechend wichtig zu sein, diese Konsistenz auch bei unterstütztem Hören zu erhalten.

Ausblick

Es gilt, noch viele grundlegende Mechanismen des räumlichen Gehörs zu verstehen. Da die Integration unterschiedlicher räumlicher Merkmale bisher zu meist unter einfachen, statischen Bedingungen untersucht wurde, sind noch viele Fragen offen, wenn es darum geht, wie sich diese Integration an Veränderungen anpasst und unter komplexeren realitätsnahen Bedingungen agiert [7]. Auch sind Interaktionen zwischen räumlichen und inhaltlichen auditiven Verarbeitungsprozessen noch weitestgehend ungeklärt [1]. In aktuellen Kollaborationsprojekten untersuchen wir am Institut für Schallforschung einige dieser Aspekte, indem wir psychoakustische und neurophysiologische Messmethoden kombinieren und versuchen, die Algorithmen des Gehirns durch funktionale Computermodelle abzubilden. Zum einen möchten wir grundlegende kognitive Mechanismen für statistisches und überwachtetes Lernen von auditiver räumlicher Information untersuchen, die zugehörigen kognitiven Netzwerke identifizieren und offenlegen, wie sich diese Mechanismen und ihre neuronale Repräsentation über die gesamte menschliche Lebensdauer verändern (siehe <http://oeaw.ac.at/isf/born-2hear>). Einen Fokus legen wir dabei auf den Looming Bias, da dieser über alle Altersstufen hinweg messbar ist und Überlebensrelevanz besitzt. Zum anderen untersuchen wir, wie akute Vorerfahrung die sensorische Verarbeitung beeinflusst, welche Entscheidungsprozesse damit einhergehen und wie diese evolutionär geformt wurden (siehe <http://oeaw.ac.at/isf/>

dynamates). Zu diesem Zweck entwickeln wir probabilistische Modelle, die das Lokalisationsverhalten von Menschen und nicht-menschlichen Primaten in dynamischen Umgebungen abbilden und somit auf algorithmischer Ebene vergleichbar machen können. Mittels Elektroenzephalographie können wir zudem die neuronale Implementierung dieser Algorithmen im Menschen untersuchen.

Fazit

Räumliches Hören ist vielschichtig und wichtig in vielen Lebenssituationen. Wesentliche Aufgaben stellen die Lokalisierung von Schallquellen, die räumlich selektive Aufmerksamkeit und die Schutzfunktion basierend auf exogener Aufmerksamkeitslenkung bei Erkennung potentieller Gefahrenquellen dar. In all diesen Situationen versucht das Gehör, basierend auf den verfügbaren akustischen Merkmalen, dem akkumulierten Vorwissen und etwaigen Grundannahmen, die relevanten Zustände der Umgebung zu ermitteln. Dabei muss die Reliabilität der unterschiedlichen Informationsquellen abgewogen werden, um eine gute Balance zwischen Robustheit und Anpassungsfähigkeit der Wahrnehmung zu garantieren.

Basierend auf den Zustandsannahmen wird aus zunächst egozentrisch kodierten Merkmalen eine externalisierte räumliche Wahrnehmung mit allozentrischem Bezugsrahmen erzeugt. Räumlich verarmte Hörbedingungen, wie sie nicht nur unter Laborbedingungen, sondern auch durch andere technische Einflüsse wie beispielsweise von Hörhilfen auftreten können, führen oft zu einem Externalisierungsverlust und eingeschränkter räumlicher Selektivität in der sensorischen Verarbeitung. Wie stark die präsentierten Merkmale von der Realität abweichen dürfen, um keine solchen Einbußen zu verursachen, ist abhängig von vielen akustischen aber auch nicht-akustischen Faktoren, denn maßgebend ist, welches Ziel das Gehör gerade verfolgt.

Danksagung

Ich möchte diese Gelegenheit nutzen, um mich bei mehreren Personen zu bedanken, und zwar Piotr Majdak und Barbara Shinn-Cunningham für Ihre Unterstützung und das hervorragende wissenschaftliche Mentorat, Ulrich Pomper, Michelle Spierings und Brigitta Tóth für die fruchtbare Kollaboration, und Diane Baier, Roberto Barumerli, Yuqi Deng, Tobias Greif, Sophie Hanke, Karolina Ignatiadis und David Meijer für die engagierte Projektmitarbeit. Zudem danke ich Peter Balazs stellvertretend für die angenehmen und anregenden Arbeitsbedingungen am Institut für Schallforschung, der DEGA für die Ehrung mit dem Lothar-Cremer-Preis vor zwei Jahren und dem Österreichischen Wissenschaftsfonds

(FWF) für die mehrmalige und andauernde finanzielle Unterstützung (J3803-N30, I4294-B, ZK66).

Literatur

- [1] Majdak, P.; Baumgartner, R.; Jenny, C.: Formation of Three-Dimensional Auditory Space. In: Blauert, J. and Braasch, J. (eds.) *The Technology of Binaural Understanding*. pp. 115–149. Springer International Publishing, Cham, 2020.
- [2] Friston, K.: A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, pp. 815–836, 2005. <https://doi.org/10.1098/rstb.2005.1622>
- [3] Rao, R. P. N.; Ballard, D. H.: Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, pp. 79–87, 1999. <https://doi.org/10.1038/4580>
- [4] Gardner, J. L.: Optimality and heuristics in perceptual neuroscience. *Nat. Neurosci.* 22, pp. 514–523, 2019. <https://doi.org/10.1038/s41593-019-0340-4>
- [5] Laquitaine, S.; Gardner, J. L.: A Switching Observer for Human Perceptual Estimation. *Neuron*. 97, pp. 462–474.e6, 2018. <https://doi.org/10.1016/j.neuron.2017.12.011>
- [6] Skerritt-Davis, B.; Elhilali, M.: Detecting change in stochastic sound sequences. *PLOS Comput. Biol.* 14, e1006162, 2018. <https://doi.org/10.1371/journal.pcbi.1006162>
- [7] van der Heijden, K.; Rauschecker, J. P.; de Gelder, B.; Formisano, E.: Cortical mechanisms of spatial hearing. *Nat. Rev. Neurosci.* 20, pp. 609–623, 2019. <https://doi.org/10.1038/s41583-019-0206-5>
- [8] Brungart, D. S.; Durlach, N. I.; Rabinowitz, W. M.: Auditory localization of nearby sources. II. Localization of a broadband source. *J. Acoust. Soc. Am.* 106, pp. 1956–1968, 1999. <https://doi.org/10.1121/1.427943>
- [9] Edmonds, B. A.; Krumbholz, K.: Are Interaural Time and Level Differences Represented by Independent or Integrated Codes in the Human Auditory Cortex? *J. Assoc. Res. Otolaryngol.* 15, pp. 103–114, 2014. <https://doi.org/10.1007/s10162-013-0421-0>
- [10] Fujiki, N.; Riederer, K. A. J.; Jousmäki, V.; Mäkelä, J. P.; Hari, R.: Human cortical representation of virtual auditory space: differences between sound azimuth and elevation. *Eur. J. Neurosci.* 16, pp. 2207–2213, 2002. <https://doi.org/10.1046/j.1460-9568.2002.02276.x>
- [11] Trapeau, R.; Aubrais, V.; Schönwiesner, M.: Fast and persistent adaptation to new spectral cues for sound localization suggests a many-to-one mapping mechanism. *J. Acoust. Soc. Am.* 140, pp. 879–890, 2016. <https://doi.org/10.1121/1.4960568>
- [12] Seitz, A. R.; Dinse, H. R.: A common framework for perceptual learning. *Curr. Opin. Neurobiol.* 17, pp. 148–153, 2007. <https://doi.org/10.1016/j.conb.2007.02.004>
- [13] Li, S.; Baumgartner, R.; Peissig, J.: Modeling perceived externalization of a static, lateral sound image. *Acta Acust.* 4, p. 21, 2020. <https://doi.org/10.1051/aacus/2020020>
- [14] Baumgartner, R.; Majdak, P.: Decision making in auditory externalization perception. *bioRxiv*. 2020.04.30.068817, 2020. <https://doi.org/10.1101/2020.04.30.068817>
- [15] Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; O’Doherty, J.; Pezzulo, G.: Active inference and learning. *Neurosci. Biobehav. Rev.* 68, pp. 862–879, 2016. <https://doi.org/10.1016/j.neubiorev.2016.06.022>
- [16] Parise, C. V.; Knorre, K.; Ernst, M. O.: Natural auditory scene statistics shapes human spatial hearing. *Proc. Natl. Acad. Sci.* 111, pp. 6104–6108, 2014. <https://doi.org/10.1073/pnas.1322705111>
- [17] Brungart, D. S., Scott, K. R.: The effects of production and presentation level on the auditory distance perception of speech. *J. Acoust. Soc. Am.* 110, pp. 425–440, 2001. <https://doi.org/10.1121/1.1379730>
- [18] Shinn-Cunningham, B. G.: Object-based auditory and visual attention. *Trends Cogn. Sci.* 12, pp. 182–186, 2008. <https://doi.org/10.1016/j.tics.2008.02.003>
- [19] Baumgartner, R.; Reed, D. K.; Tóth, B.; Best, V.; Majdak, P.; Colburn, H. S.; Shinn-Cunningham, B.: Asymmetries in behavioral and neural responses to spectral cues demonstrate the generality of auditory looming bias. *Proc. Natl. Acad. Sci.* 114, pp. 9743–9748, 2017. <https://doi.org/10.1073/pnas.1703247114>
- [20] Ignatiadis, K.; Baier, D.; Tóth, B.; Baumgartner, R.: Neural Mechanisms Underlying the Auditory Looming Bias. *Audit. Percept. Cogn.* 2021. <https://doi.org/10.1080/25742442.2021.1977582>
- [21] Bach, D. R.; Schächinger, H.; Neuhoff, J. G.; Esposito, F.; Salle, F. D.; Lehmann, C.; Herdener, M.; Scheffler, K.; Seifritz, E.: Rising Sound Intensity: An Intrinsic Warning Cue Activating the Amygdala. *Cereb. Cortex.* 18, pp. 145–150, 2008. <https://doi.org/10.1093/cercor/bhm040>
- [22] Deng, Y.; Choi, I.; Shinn-Cunningham, B.; Baumgartner, R.: Impoverished auditory cues limit engagement of brain networks controlling spatial selective attention. *NeuroImage*. 202, p. 116151, 2019. <https://doi.org/10.1016/j.neuroimage.2019.116151> ■



DI Dr. Robert Baumgartner
Österreichische Akademie der Wissenschaften,
Institut für Schallforschung, Wien