

Neurobiologie von Sprache und Musik

(Paul Natterer)

Auch *Sprache* ist als Spezialfall sensorischer Integration und eines Systems zur exakten zeitlichen Steuerung, Speicherung und Wiedergabe von Tönen aufzufassen. Lesen hat wahrscheinlich mit der *pars pro toto* Erkennung beim Sehen zu tun, d.h. der Fähigkeit, aus einem Teil auf das Ganze zu schließen.

- Echte Sprache zeichnet sich dann durch diese fünf Kriterien aus: (1) Spezialisierung auf Informationsmitteilung, (2) Freie Wahl (Willkür) der Zeichen (Laute); (3) Kontextunabhängigkeit der Zeichen (Verschiebbarkeit in andere Zusammenhänge und Gebiete); (4) Produktivität (Neuheit bzw. Kreativität der Zeichenhandlungen); (5) Iteration (Wiederholung und Addition der Zeichen zu Sätzen und Texten) und Rekursion (Einbetten einer Sprachstruktur in eine weitere derselben Art). Man definiert daher oft die (menschliche) Sprache als Welthabe in objektiver Sachlichkeit und Freiheit (Distanz) im Unterschied zur tierischen Kommunikation als Umweltorientierung. Der enge Zusammenhang von multisensorischer Assoziation und Sprache legt daher nahe, dass die

„Fähigkeit, Sprache zu erwerben, als Vorbedingung die Fähigkeit hat, crossmodale [die Einzelsinne übergreifende] Assoziationen zu bilden. Bei subhumanen Formen sind die einzigen ohne weiteres bewerkstelligten sensori-sensorischen Assoziationen diejenigen zwischen einem nicht-limbischen (das heißt visuellen, taktilen oder auditorischen) und einem limbischen Reiz. Es ist nur beim Menschen der Fall, daß Assoziationen zwischen zwei nicht-limbischen Reizen ohne weiteres gebildet werden, und diese Fähigkeit liegt dem Lernen von Namen von Gegenständen zugrunde.“ Popper/Eccles (1991, 372)
- Im Zusammenhang der Sprachfähigkeit hier ein Blick auf den hier grundlegenden Sensor, das Gehör. Dieser Seitenblick hat den zusätzlichen Gewinn zu zeigen, dass sich die Grundsätze und Phasen der Informationsverarbeitung bei anderen Sensoren ähnlich darstellen wie beim Gesichtssinn. Der spezifische Reiz des Hörsinns sind längsverlaufende (longitudinale) Luftdruckschwankungen (Verdichtungen und Verdünnungen) von 16–20000 Hz und 340 m/s Ausbreitungsgeschwindigkeit, also 1 Mach oder 1224 km/h (Schallwellen).
- Menschen registrieren den Frequenzbereich von 20–16000 Hz. Der physikalische Schalldruckpegel wird auf der Dezibelskala angegeben, der physiologische Lautstärke-pegel auf der Phonskala. Töne sind Schallwellen mit einer einheitlichen Frequenz. Klänge sind Kombinationen von Grundtönen mit Obertönen. Geräusche sind Schallereignisse über den gesamten Frequenzbereich.

- Schallortung und umweltorientiertes Richtungshören erfolgt durch verbewusste Verrechnung von Zeit- und Intensitätsunterschieden zwischen beiden Ohren (Laufzeit- und Schallpegeldifferenz).
- Entfernungsbestimmung und subjektorientiertes Richtungshören wird aus der Klangfärbung entnommen. Diese entsteht durch Resonanzen und Reflektionen an Kopf und Ohrmuscheln.
- Die Reizkodierung in den Hörnervenfasern erfolgt anhand drei Reizparametern: (1) Intensität oder Entladungsrate des aktivierten Sensorpotentials. Diese entspricht etwa der Amplitude des Aktionspotentials in den weiterverarbeitenden Neuronen – (2) Zeitdauer der Aktivierung – (3) Frequenzselektivität. Letzteres bedeutet, dass das Frequenzspektrum analysiert wird, indem jeder Frequenz ein bestimmter Ort des sog. Corti-Organs in der Ohrschnecke (*Cochlea*) zugeordnet wird (Ortsselektivität). Das Frequenzspektrum wird also „verräumlicht“.
- Das primäre Hörzentrum, der primäre akustische Kortex, leistet die allgemeine Tonanalyse, das sekundäre Hörzentrum ist beim Menschen auf die Wahrnehmung von Sprachlauten spezialisiert, d.h. auf immer komplexere Schallmustererkennung, auf deren Beginn, Ende und Wechselwirkung (Interferenz). Wie aus der Rundfunktechnik bekannt, sind diese Schallmuster amplituden- oder frequenzmodulierte Töne.
- Sprachliche Informationsverarbeitung fällt im Übrigen nicht zusammen mit der allgemeinen Informationsverarbeitung des Hörzentrums. Der allgemeine Hörsinn arbeitet kontinuierlich und der Reihe nach (linear) und ist genetisch festgelegt. Sprachwahrnehmung erfolgt dagegen in getrennten Segmenten und nichtlinear, d.h. in einem Netzwerk mehrerer zusammenwirkender Ebenen und mit freier Festlegung der Reizbedeutung. Im Sprachverstehen können wir daher bis zu 50 Laut-einheiten/Sekunde unterscheiden und ordnen (normalerweise 12 Töne/Sekunde), während bei nichtsprachlichen Lauten die Unterscheidungsfähigkeit dramatisch auf 0,6 Einheiten/Sekunde absinkt. Begriffsgeleitete Information und Rekonstruktion aufgrund von Wortbedeutungen, Grammatik und Kontext macht 50% des Sprachverständnisses aus.
- Die Sprachwahrnehmung erfolgt wie bei anderen Sensoren auch durch Merkmalsextraktion: diese Merkmale sind die Laute oder Phoneme – Mustererkennung oder Gestaltwahrnehmung: Muster sind hier die Wörter oder Morpheme – begriffliche Interpretation und Rekonstruktion: begriffliche Schemata sind die Sätze. Das Lautsystem natürlicher Sprachen variiert dabei von 11 Lauten (Hawaiisch) über 40 Laute (Englisch) bis zu 60 Lauten (afrikanische Sprachen). Laute lassen sich in

sogenannten Sprachspektrogrammen durch ihre Merkmale eindeutig definieren: Vokale sind Akkorde, d.h. Lautfrequenzen mit Grundton und Obertönen, und werden Formanten genannt; Konsonanten sind schnelle Formantübergänge. Entwicklungspsychologisch ist interessant der Wechsel von uneingeschränkter Sprachunterscheidungsfähigkeit (0–1 Jahr) zur eingeschränkten spezifischen Sprachunterscheidungsfähigkeit mit abgeschwächter Unterscheidungsfähigkeit für Abweichungen von typischen Lauten der Muttersprache.

- Musik ist eine dritte besondere akustische Informationsverarbeitung. Töne sind dabei sehr ähnlich wie die Farben im Farbkreis in drei Dimensionen veränderlich: Tonhöhe – Tonart – Tonstärke (20–160 dB). Und analog zur Sprache und anderen Wahrnehmungsdimensionen erfolgt die Musikwahrnehmung durch Merkmalsextraktion: Dies betrifft gleichzeitige akustische Konturen (Frequenzspektren der Töne) und Merkmale der Abfolge von Tönen (Rhythmen) – vorbewusste angeborene Mustererkennung (Melodie und Takt) – erlernte bewusste begriffliche Interpretation: technisches und geschichtliches Vorwissen, Erwartungen und Hypothesen bzgl. Instrumenten, Stilen und Kompositionen.
- Der Aufbau des Hörsinns ist also ähnlich der fortlaufend immer komplexeren Merkmalsanalyse des visuellen Kortex oder des Sehentrums.