

Diskussion von Mechanismen der Wahrnehmung geflüsterter Sprache und des Brummtons

Eckard Blumschein

Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik, Otto-von-Guericke-Universität Magdeburg
blumschein@et.uni-magdeburg.de

Es ist unstrittig, dass unsere Ohren bei der Wahrnehmung von Informationen viel mehr leisten als eine simple Frequenzanalyse [1, 2]. Man hat sich daran gewöhnt, den in traditionellen Theorien, Modellen und Tools unberücksichtigten Rest unendifferenziert einer vagen Vorstellung von nicht-spektraler alias zeitlicher Kodierung zuzuschreiben. Diese Studie betrachte zwei Fälle, bei denen der „Rest“ in den Vordergrund tritt, weil die sonst dominierende tonale Kodierung bei geflüsterter Sprache nur eine geringe bzw. beim Brummtton vermutlich gar keine Rolle spielt. Die Erörterung greift auf www.auditorymodels.org und Messages (Mxxx) in <http://iesk.et.uni-magdeburg.de/~blumsche.AuditoryFunction.html> zurück, vor allem zur Funktion der äußeren Haarzellen. Trotz Beschränkung auf Fragen von aktuellem Interesse zeigen die beiden Beispiele, worin sich fünf Mechanismen vorteilhaft voneinander unterscheidenden: der zeitlich-spektrale, der rein spektrale, der rein zeitliche, ein direkter und ein unterschwelliger. Man erkennt, wo ihre Grenzen liegen und wie sie sich gegenseitig ergänzen.

1 Signalvorverarbeitung im Innenohr (ohne Sacculus)

Bei nicht zu großem Schallpegel arbeiten die äußeren Haarzellen als Vorverstärker und bilden die Eigenfrequenzen (Formanten) im Sprechorgan als Muster im Hörnerv ab. Auf der Stelle „rollende“ Wellen beobachtet man als Schwebung. Die Hörnervverregung unterscheidet sich erheblich vom entsprechenden Spektrogramm. Formanten sind keine Obertöne. Ihre Beschreibung als diskretes Spektrum der Schwingungen der Stimmbänder mit modulierter Hüllkurve ist ungenau. Tatsächlich kann das Innenohr nur Frequenzgruppen auflösen. Innerhalb einer kritischen Bandbreite liegen mehrere Spektrallinien. Die scheinbar mäßige Auflösung reicht aus, um alle drei Formanten zu unterscheiden. Beispielsweise ist der Abstand zwischen $F_2 = 1350$ Hz und $F_3 = 1690$ Hz beim /R/ wie in Rose größer als die kritische Bandbreite von rund 210 Hz. Obwohl die *spektralen* Positionen tonotopisch bis ins Gehirn erhalten bleiben, vermutet der Autor, dass das endgültige Tonhöhenmuster im tonalen Frequenzbereich erst infolge *zeitlicher* Koinzidenz zwischen unterschiedlich verzögerten Signalen entsteht. Anders könnte er sich die gleichartige Wahrnehmung von Tönen nicht erklären, die sich um ganze Oktaven unterscheiden. Anders verstünde er weder die Robustheit des Gehörs gegenüber Rauschen noch seine Fähigkeit winzige Frequenzdifferenzen wahrzunehmen. Møller [3] schreibt: Cochlear spectral filtering may be important for temporal coding. Oberhalb etwa 4500 Hz ist dieser an die Phasenkopplung gebundene zeitliche Aspekt des spektralen Codes verschwunden.

Es verbleibt dann ein zweites Merkmal, das nicht zeitlich auflösbare *rein spektrale* Profil. Über die spektrale „Farbe“ des Rauschens sind die Zischlaute verhältnismäßig unscharf bestimmt.

Als dritte Information gibt es einen andersartigen Kode, der nicht von jenen Fasern des Hörnervs (tonotopisch) übertragen wird, die entsprechend der jeweiligen charakteristischen Frequenz (CF) der äußeren Haarzellen selektiv erregt werden sondern ziemlich „global“ von der Gesamtheit aller äußeren Haarzellen, vor allem von denjenigen, die sich nahe der Basis befinden und mit der geringsten Verzögerung fast gleichzeitig auf Schallsignale reagieren. Es ist üblich, dieses dritte Muster „zeitlich“ zu nennen. Verwechslungen mit dem zeitlichen Aspekt des spektralen Musters vermeiden die Bezeichnungen „*rein zeitlich*“ oder „nicht spektral“. Eine absolute Trennung ist jedoch nicht gerechtfertigt. Beispielsweise ist der Einsatz eines Vokals zunächst rein zeitlich beschrieben, denn die spektrale Analyse braucht Zeit. Der zeitliche Kode ist zwar grundsätzlich weder an die untere noch an die obere Grenze der hörbaren Frequenz gebunden, funktioniert aber für Intervalle unter 200 μ s Dauer nur bei nichtstationären

Signalen und sonst für Vergleiche zwischen beiden Ohren. Steigt die Dauer über 2 ms, dann stört die überschrittene Refraktärzeit.

Übersteigt viertens der Pegel einen durch Sättigung der äußeren Haarzellen gekennzeichneten Wert zwischen etwa 60 und 90 dB je nach Bandbreite der Erregung, dann werden die inneren Haarzellen *direkt* stimuliert. Die Basilarmembran schwingt nun ohne die zusätzliche Latenz der äußeren Haarzellen, sowie mit am gleichen Ort deutlich niedrigerer Frequenz, vgl. M67. Der von konstanter Frequenz maximal angeregte Ort verschiebt sich also zur Basis. Damit ändert sich die spektral kodierte Klangfarbe. Bei rein tonotopischer - also nicht zeitlicher - Wahrnehmung muss die empfundene Tonhöhe drastisch zunehmen. Dies gilt speziell für Frequenzen über 4,5 kHz. Anregungen mit Frequenzen über etwa 8 kHz beanspruchen bei hohem Pegel die äußeren Haarzellen ohne dass die inneren Haarzellen dies weiterleiten.

2 Besonderheiten geflüsterter Sprache

Tiere und sehr kleine Kinder beherrschen das Flüstern noch nicht. Die besondere Schwierigkeit liegt dabei weniger im stimmlosen, leise zischenden Sprechen als im Verstehen geflüsterter Worte. Man erkennt im Spektrogramm kaum noch Formanten, und dies ist beabsichtigt, denn man flüstert etwas in ein Ohr um zu vermeiden dass es ein Dritter verstehen kann. Das Kinderspiel „stille Post“ beweist: Geflüstert übertragene Information ist ziemlich unsicher. Eine Übermittlung von Nonsenssilben gelingt kaum. Elemente geflüsterter Sprache werden offensichtlich durch Identifikation mit im Gedächtnis vorhandenen ähnlichen Mustern erkannt, also in einem aktiven Prozess. Wie Warren [4] gezeigt hat, überbrückt das Gehör Lücken in spektralen und zeitlichen Mustern dann, wenn sie mit Rauschen ausgefüllt sind. Akustisch zeichnet sich geflüsterte Sprache dadurch aus, dass ihre Übertragung infolge höherer Frequenz gerichteter aber stärker gedämpft erfolgt. Entscheidend ist aber zweifellos die drastische Einschränkung aller stimmhaften Laute, nicht nur der Vokale. Angesichts dieses Informationsverlusts verblüffen die verbleibende Verständlichkeit und der nahtlose Übergang zwischen geflüsterter und normaler Sprache. Ein Schlüssel zum Verständnis dieser Phänomene ist die in [5] aufgezeigte Präzision der scheinbaren Entfernung zwischen Ohr und Flüsterquelle. Während man bei normalem Sprechen oder Schreien mit wachsendem Pegel den falschen Eindruck einer wesentlich größeren Entfernung hat, gestattet die bei geflüsterter Sprache weiterhin vorhandene aber stärker hervortretende Zeitstruktur stets eine genaue Lokalisation. Das auditorische System passt sich also bei der Verwertung der Signale aus dem Innenohr an die verbliebene Kodierung an. Stevens [6] verglich PST-Histogramme von Reaktionen des Hörnervs miteinander. Während er für normal gesprochene Silben mit globaler Mittelung bei niedrigen und mittleren CFs gleiche Muster sah, dominierten beim Flüstern die CF-Ensembles über 2,4 bzw. 4,8 kHz. Er folgerte, dass Hörnerv-Fasern hoher CF wesentlich zur Kodierung geflüsterter Sprache beitragen. Sie vermitteln nicht nur rein spektrale Information sondern vor allem den von Steinschneider u. a. [7] sogar noch in der Großhirnrinde nachgewiesenen nicht spektralen Kode. Quer über alle CFs wird dabei mit niedriger Folgefrequenz die Zeitstruktur übertragen. Der Begriff zeitliche Feinstruktur ist insofern fragwürdig als das Gehirn für eine präzise Demodulation die mit kleiner CF erhebliche cochleare Latenz berücksichtigen müsste, wofür es im Mittelhirn keine Anzeichen gibt. Im Großhirn erfolgt die Demodulation dann vermutlich nur durch simple Summation über die Signale von allen Fasern. Sie ist also nur gemessen an der Zeitauflösung im Spektrogramm fein, verglichen mit dem ursprünglichen Signal aber grob. Um tiefe Frequenzen hören zu können brauchen wir paradoxerweise vor

alle jene den hohen CF zugehörigen Haarzellen im basalen Teil der Hörschnecke, die durch Schädigung und Alterung gewöhnlich zuerst ausfallen. Wir brauchen sie zum Erkennen transienter, nicht spektraler und rein spektraler Merkmale von Sprachinhalten und akustischen Signaturen sowie zur Ortung von Schallquellen.

3 Unterscheidung von Explosion und Implosion

Damit beim Flüstern die Unterscheidung der harten (explosivem) Konsonanten k, p und t von den weichen (implosivem) g, b und d erhalten bleibt, dürfen die nachfolgenden Vokale nicht völlig verschwinden. Zwischen tss und dss hört man keinen Unterschied. Wie kommt das? Bekanntlich geht bei den harten die Explosion dem Onset des Vokals um etwa 60 ms voraus. Konsonanten mit weniger als 10 ms Vorlauf werden als weich erkannt. Zischlaute entstehen ohne viel Vorlauf, denn sie wirken kaum auf die Stimmritze zurück. Ein leises „Psst!“ klingt meist eher wie bssd. Wäre der Vorlauf das einzige Merkmal, dann könnte man einen knallenden Sektkorken schwerlich von einem schmatzenden Kuss unterscheiden. Beide akustischen Ereignisse unterscheiden sich vor allem in Laufrichtung und Dispersion der Welle, also spektral. Für korrekte Wiedergabe müssen Lautsprecher auch richtig gepolt sein. Haarzellen im Innenohr wirken zunächst wie einfache Dioden, vgl. M180. Folglich unterscheidet sich die Latenz bei Reaktion auf einen positiven Druckimpuls (condensation click) von der auf einen Saugimpuls (rarefaction click) um eine Halbperiode. Saugimpulse dagegen depolarisieren die Haarzellen unmittelbar. Diese Zeitverschiebung hört man jedoch bekanntlich normalerweise nicht. Lediglich bei Klickfrequenzen um 100 bis 150 Hz hören sich Übergänge von positiver zu negativer Polarität deutlich anders an als in umgekehrter Richtung (M130). Die auffällige Trägheit solcher Übergänge ist ein Indiz ihrer Wahrnehmung im zentralen Nervensystem über den zeitlichen Aspekt des spektralen Codes. Rein zeitliche Merkmale findet bereits der Nucleus cochlearis. Die rein zeitliche (nicht spektrale) Wahrnehmung ist wie ein Zweiweg-Gleichrichter generell nicht phasentreu, und dies wäre bei niedriger Frequenz wegen mehrfachen Feuerns der Neuronen innerhalb einer Periode auch nicht sinnvoll. Die wahrgenommene Frequenz verdoppelt sich [4], wenn die Periode den Wert von etwa 12 ms überschreitet.

4 Wahrnehmung von Mikrowellen und Brummtton

[8] beschreibt als „Brummtton“ eine vermutlich durch Radarwellen verursachte, also nicht-akustische Geräusch-Wahrnehmung. Die Hörbarkeit von Mikrowellen-Impulsen ist längst bekannt [9], und Radars senden scharf fokussierte Megawatt-Impulsleistungen aus. Experten wundern sich dennoch, dass sich einzelne Menschen erheblich belästigt fühlen. Zunächst hätten sie erwartet, dass man kein Brummen sondern höchstens den der Impulsfrequenz von einigen Kilohertz entsprechenden Pfeifton hört, welcher jenen Menschen vertraut ist, die unmittelbar am Radar arbeiten. Außerdem war die rechnerische Dichte der diffus reflektierten Leistung mit etwa 0.01 bis 10 mW/cm² deutlich kleiner als die für Menschen zu 80 mW/cm² bei 70 µs [10] bzw. 300 mW/cm² bei 15µs Impulsdauer [11] ungenau bestimmten Schwellwerte, ungenau deshalb, weil man teilweise die Leistung einfach durch den Querschnitt des Wellenleiters dividierte. Die Diskrepanz wird kleiner wenn man berücksichtigt, dass die Flüssigkeit des menschlichen Körpers eine gute Antenne darstellt und ein Teil eines „pencil“- Radarstrahls direkt auftreffen kann.

Das Brummen wird unterhalb der Hörschwelle nur von einem kleinen Teil der Bevölkerung (in Taos 2%, örtlich 50%) gehört und zunächst auch nur nachts in Häusern, also in sehr stiller Umgebung, bis es die Betroffenen mit ihrer fortschreitenden Sensibilisierung auch im Freien und schließlich ständig wahrnehmen. Auch eine ertaubte Frau klagt über das Brummen. Die Wahrnehmung von Mikrowellen stellte sich die Mehrheit der Forscher so vor, dass die Haarzellen von einer thermisch bewirkten Welle mechanisch stimuliert werden. White [12] berechnete elastische Wellen, die über nur um 0,001grd differierenden Temperaturen piezo-elektrisch Spannungen von 1 bis 60mV je kW/cm² erzeugen. Während der zitierten Untersuchungen galt das spektrale Hören als einzig vorstellbares. Der Autor schließt jedoch für den Brummtton weder nicht

thermische noch nicht spektrale Mechanismen vorschnell aus. Er behauptet, dass die direkte Einwirkung des elektromagnetischen Felds auf Ionen-Dipole vor den Spitzen der Cilia mit weniger Energie auskommt und sieht die Deformation der Öffnungen an den Spitzen der Cilia durch veränderten Zug an sie verbindenden Bändern (M119) beim Verbiegen der Cilia nicht als die einzig mögliche Steuerung des Einstroms in die Ionenkanäle an.

5 Unterschwellige Brummttonwahrnehmung

Übliche Mechanismen der allmählichen Erregung von äußeren Haarzellen benötigen an der Hörschwelle mehrere Perioden zum Aufbau einer an der Basilmembran messbaren Schwingung. Kurze Radarimpulse sind dazu zu schwach. Bei Versuchen, mit kräftigen, 70 µs langen Mikrowellen-Impulsen gleiche Reaktionen in einzelnen Fasern des Hörnervs von Katzen zu erzielen wie mit akustischen Klicks hat man Latenzen zwischen 1,5 und 5 ms beobachtet [13]. Andere Untersuchungen vermerken: Radar hört nur wer oberhalb 5 (bzw. 8) kHz gut hört [9, 11, 14]. Der Autor sieht darin Anzeichen der nicht spektralen Wahrnehmung. Einen lokalen Mechanismus würde man an einem tonalen Höreindruck erkennen. Das Geräusch von einer 400 Hz-Radar-Impulsfolge ließ sich aber keinem Sinuston zuordnen [9].

Das Brummen hat man als 72-Hz-Ton nachgestaltet. Es hört unruhig an. Der Autor sieht die Ursache dafür in von Umlauf zu Umlauf veränderten Strahlpositionen und in Besonderheiten der Wahrnehmung. Sie stützt sich vermutlich auf wenige 14ms lange Intervalle alle 5 Sekunden.

Viele Beobachtungen zeigen unabhängige voneinander einen Zusammenhang zwischen Brumm-Phänomen und Radar an, der sich als unterschwelliges Hören einer 36-Hz-Periodizität deuten lässt. Der Autor geht davon aus, dass derart niedrige Frequenzen kaum noch spektral wahrgenommen werden. Er vermutet, dass zur Demodulation der Radarechos die Haarzellen beitragen. Ihre Nichtlinearität mag Einfluss auf die Rate spontanen Feuerns jener besonders empfindlichen Neuronen des Hörnervs nehmen, deren Refraktärzeit besonders groß ist. Angenommen, diese Taubzeit betrüge 10 ms, dann wäre eine Modulation der Spontanrate nur unterhalb 100 Hz möglich. Begrenzt ist die scheinbare Frequenz der vermuteten intervallbasierten Wahrnehmung des Brummens auch vom Gehirn (so wie die des Flimmerns). Die Beschränkung des Phänomens auf bisher wenige Betroffene beweist nicht, dass deren Innenohren abnorm sind. Wahrscheinlicher sind Efferenz-Probleme wie für Tinnitus vermutet und die unwillkürlich erlernte Fähigkeit des Gehirns, Intervalle ohne Tonotopie zu erkennen.

Literatur

- [1] E. Blumschein: Do ears really just perform frequency analysis of crackle sound? Forum Acusticum 1999.
- [2] E. Blumschein: Was leistet unser Gehör wirklich? DAGA 01.
- [3] A.R. Møller: Hearing. San Diego: Academic Press 2000.
- [4] R. Warren: Auditory Perception. Cambridge: Univ. Press 1999.
- [5] D. Mershon, J. Philbeck: Auditory perceived distance of familiar speech sounds. *Proc. Psychonomic Soc.* (32nd) 1991.
- [6] H.E. Stevens: Ensemble response of the aud. nerve to normal and whispered stop consonants. *Hearing Research* (1999) 47-62.
- [7] M. Steinschneider et al.: Click train encoding in primary auditory cortex of awake monkey: Evidence for two mechanisms subserving pitch perception. *JASA* **104** (1998) 5, 2935-2955.
- [8] Blumschein: Prävention subjektiver Störgeräusche. DAGA 02.
- [9] A.H. Frey: Auditory system response to radio-frequency energy. *Aerospace Med.* **32** (1961) 1140-1142.
- [10] A.H. Frey, R. Messenger, Jr.: Human perception of illumination with pulsed ultrahigh-frequency electromagnetic energy. *Science* **181** (27 July 1973) 356-358.
- [11] C. Cain, W. Rissman: Mammalian audit. response to 3.0 GHz microw. pulses. *IEEE Trans. Biomed. Eng.*, **25** (1978) 3, 288-293.
- [12] R.M. White: Generation of elastic waves by transient surface heating. *J. Appl. Phys.* **34** (1963) 12, 3559-3567.
- [13] R.M. Lebovitz, R.L. Seaman: Microwave hearing: Response of single auditory neurons. *Radio Sci.* **12** (1977) 6S, 229-236.
- [14] V.V. Tyazhelov, R.E. Tigranian, et al.: Some peculiarities of auditory sensations evoked by pulsed microwave fields. *Radio Sci.* **14** (1979) 6S, 259-263.