

T. Lenarz (Hrsg.)

Cochlea-Implantat

Ein praktischer Leitfaden
für die Versorgung
von Kindern und Erwachsenen

Mit 128 Abbildungen und 14 Tabellen



Springer

Professor Dr. med. Thomas Lenarz
Medizinische Hochschule Hannover
Klinik für Hals-Nasen-Ohrenheilkunde
Carl-Neuberg-Straße 1
30625 Hannover

ISBN 3-540-60388-3 Springer-Verlag Berlin Heidelberg New York

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Cochlea-Implantat : ein praktischer Leitfaden für die Versorgung
von Kindern und Erwachsenen / Hrsg.: Thomas Lenarz. – Berlin ;
Heidelberg ; New York ; Barcelona ; Budapest ; Hongkong ; London ;
Mailand ; Paris ; Singapur ; Tokio : Springer, 1998
ISBN 3-540-60388-3

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

© Springer-Verlag Berlin Heidelberg 1998
Printed in Germany

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Produkthaftung: Für Angaben über Dosierungsanweisungen und Applikationsformen kann vom Verlag keine Gewähr übernommen werden. Derartige Angaben müssen vom jeweiligen Anwender im Einzelfall anhand anderer Literaturstellen auf ihre Richtigkeit überprüft werden.

Einbandgestaltung: de'blik, Graphic Design, Berlin
Satz: Fotosatz-Service Köhler OHG, Würzburg

SPIN: 10494659 26/3134 – 5 4 3 2 1 0 – Gedruckt auf säurefreiem Papier

Codierungsstrategien – Grundlagen und Evaluation

N. Dillier

3.1	Sprachkommunikation und Sprachcodierung	53
3.1.1	Sprachproduktion	53
3.1.2	Sprachperzeption	54
3.1.3	Informationstransmissionsanalyse	55
3.1.4	Vocoder	55
3.2	Evaluationsverfahren	57
3.2.1	Testanforderungen	57
3.2.2	Beschreibung mit Wahrscheinlichkeitsmodellen	57
3.3	Physiologische und technische Rahmenbedingungen	58
3.3.1	Reizformen	58
3.3.2	Anordnung der Elektrodenpaare	59
3.3.3	Bestimmung der Wahrnehmungsschwelle	59
3.3.4	Modelle der Hörnervstimulation	60
3.4	Typen von Cochlea-Implantaten – Stand der Technologie	61
3.4.1	Sprachdiskrimination	61
3.4.2	Elektrodenanordnung	62
3.4.3	Sprachcodierung	62
3.4.4	Signalübertragung	62
3.4.5	Zuverlässigkeit	62
3.5	Experimentelle Untersuchungen verschiedener Sprachcodierungsstrategien	62
3.5.1	Analyse der Verfahren	64
3.5.2	Evaluationsexperimente mit Patienten	65

EINLEITUNG

Weltweit wurden bis heute schätzungsweise über 10 000 Cochlea-Implantate (CI) eingesetzt, in zunehmendem Maß bei Kleinkindern. Obwohl noch nicht alle Aspekte dieser noch jungen medizintechnischen Errungenschaft vollständig geklärt sind, wurde mittlerweile ein technologischer Reifegrad erreicht, welcher gekennzeichnet ist durch hohe Zuverlässigkeit, große Erfolgschancen und geringe Komplikationsraten. Die aktuelle Forschung und Entwicklung zielt auf die weitere Verbesserung und Optimierung der Sprachverständlichkeit sowie vereinfachtere und objektivere Methoden der Prozessoranpassung.

Anregungen für verbesserte Signalcodierung und -transformation können vom Studium der Sprachproduktion und -perzeption sowie der technischen Kommunikationssysteme zur Übertragung, Analyse und Synthese von Sprache gewonnen werden. Deshalb werden im folgenden zuerst einige relevante Grundlagen und Anwendungen der akustischen Phonetik und Informationstheorie zusammengetragen.

Zur Weiterentwicklung und Evaluation neuer Sprachverarbeitungsverfahren und -prozessoren werden differenziertere Analysewerkzeuge benötigt. Ansätze verfeinerter Sprachtestmethoden werden im Abschnitt 3.2 kurz dargestellt.

Der Rückblick auf bereits abgeschlossene Entwicklungsschritte und grundsätzliche Überlegungen des Abschnitts 3.3 vermag unter Umständen neue Anstöße für weitere Forschungen zu vermitteln. Viele der heute modernen Konzepte wurden bereits vor vielen Jahren schon einmal erarbeitet und mangels technologischer Möglichkeiten oder aufgrund anderer Faktoren wieder fallengelassen.

Die physiologischen und technologischen Rahmenbedingungen, welche im Abschnitt 3.4 erläutert werden, schränken die Menge der möglichen Codierungsoptionen in der Praxis ein und erfordern für die Verarbeitung der akustischen Signale eine Informationsauslese bzw. Redundanzreduktion (Filterung, Merkmalsextraktion). Einschränkungen ergeben sich durch die limitierte Anzahl von Kanälen (Frequenzauflösung), die begrenzte Selektivität der elektrischen Stimulation und die maximale Kapazität des Übertragungskanals. Der Sprachprozessor führt deshalb eine Verarbeitung des Eingangssignals durch, welche die wichtigsten Sprachmerkmale zu extrahieren und als Stimulationssignal so zu codieren sucht, daß das Sprachverständnis des CI-Trägers optimiert wird.

Eine Übersicht über die heute wichtigsten Systeme wird in Abschnitt 3.5 gegeben. Die Hauptstoßrichtung der aktuellen experimentellen Forschungsarbeit ist die theoretische und praktische Evaluation verschiedener neuer Sprachverarbeitungsstrategien mit mehrkanaligen Cochlea-Implantaten. Abschnitt 3.6 stellt einige Beispiele zur Implementation verschiedener Strategien mit leistungsfähigen digitalen Signalprozessoren dar, welche im Labor und Feldversuch mit CI-Trägern getestet werden konnten. Mit einigen dieser neuen Verfahren konnten trotz fehlenden Trainings beträchtliche Verbesserungen mit Sprachverständlichkeit erreicht werden. Die Umsetzung dieser Ergebnisse auf breiter Ebene ist Aufgabe für die weitere Forschung und Entwicklung.

3.1 Sprachkommunikation und Sprachcodierung

Eine der wesentlichen Funktionen des menschlichen Gehörs ist die Gewährleistung der lautsprachlichen, zwischenmenschlichen Kommunikation. Ein einfaches Schema der Sprachkommunikation umfaßt die drei Elemente Sender, Empfänger und Übertragungsstrecke (s. Abb. 3.1). Die Beziehungen dieser Elemente sind in der Regel mehrdeutig und komplex. Beispielsweise ist ein Sender oft auch Empfänger einer sprachlichen Mitteilung, die Art der Übertragungsstrecke (z.B. eine laute Umgebung, eine Telefonleitung, ein hallender Raum) kann Rückwirkungen sowohl auf Form wie auch Inhalt des sprachlichen Austausches haben. Eine gestörte Telefonverbindung kann z.B. den Sprecher dazu anregen, lauter, langsamer oder mit anderer Wortwahl oder Syntax zu sprechen. C.E. Shannon hat in seiner „Mathematischen Kommunikationstheorie“ (1948) die Grundlagen zur Beschreibung verschiedenartiger Informationsübermittlungssysteme gelegt. Er hat insbesondere auch Grenzwerte und Formeln zur Berechnung der Kapazität von gestörten Informationskanälen hergeleitet. Eine der wichtigsten Formeln ist in Abb. 3.1 angeführt. Sie besagt, daß die maximal zu übertragene Information proportional zur Bandbreite und zum Signal-Rausch-Verhältnis ist.

Ein teilweiser oder vollständiger Verlust des Gehörs kann sowohl eine Reduktion der ausnutzbaren Bandbreite als auch eine Verschlechterung des Signal-Rausch-Verhältnisses (oder anders ausgedrückt: der Anzahl unterscheidbarer Intensitätsstufen bzw. des Dynamikbereichs) beinhalten. Signalverarbeitungsstrategien für Cochlea-Implantate zielen auf eine effiziente Ausnutzung sprachlicher Information durch Kompression, Transformation und Codierung. Deshalb werden im folgenden einige Grundlagen der akustischen Phonetik kurz zusammengefaßt.

3.1.1 Sprachproduktion

Vom informationstheoretischen Standpunkt kann das Ohr als Empfänger der von einem oder mehreren Sendern ausgesandten Nachrichten verstanden werden. Nachrichten müssen dafür in einer geeigneten Art verschlüsselt oder codiert sein. Der Übertragungscode kann auf mehreren Zwischenstufen transformiert, erweitert oder reduziert werden. Damit auch bei einer Störung der Übertragungsstrecke wichtige Information erhalten bleibt, wird bei technischen Kommunikationskanälen Redundanz, d.h. zusätzliche Information hinzugefügt, so daß die Nachricht mehrfach übertragen werden kann. Der Nachrichtenempfänger sollte dann wieder in der Lage sein, die Störung vom Nutzsignal zu trennen und die ursprüngliche Information zu rekonstruieren.

Die menschliche Sprachproduktion und -perzeption sind evolutionsbiologisch optimal aufeinander abgestimmt und können durch Modelle mit mehreren hierarchisch organisierten Ebenen beschrieben werden. Der Inhalt einer sprachlichen Mitteilung (die semantische Komponente) wird zuerst mit syntaktischen und lexikalischen Komponenten (Satz- und Wortbildungsregeln, Grammatik) sowie Betonungs- und Sprachmelodieattributen in eine phonologische Komponente abgebildet. Aus den abstrakten phonologischen Sprachmerkmalen werden durch phonetische Implementationsregeln konkrete Lautrepräsentationen (Phoneme) realisiert, welche alsdann zeitliche Muster zur Ansteuerung artikulatorischer Vorgänge erzeugen. Die Artikulationsorgane (Kehlkopf, Vokaltrakt, Zunge, Lippen, Nasaltrakt) schließlich generieren die eigentlichen akustischen Sprachsignale, welche als Schallereignisse durch Luftschwingungen übertragen werden.

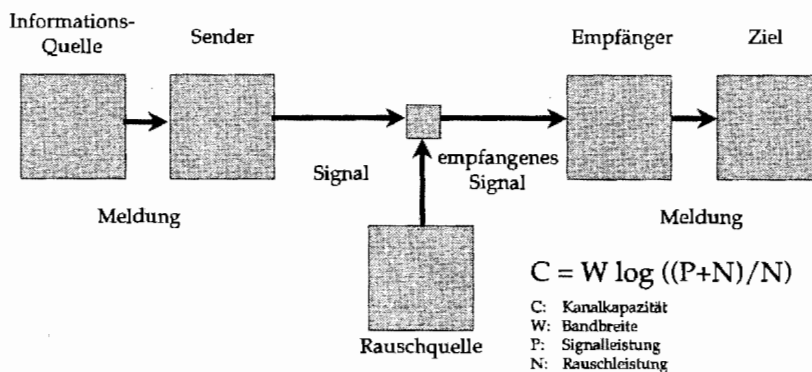


Abb. 3.1. Schematische Darstellung eines gestörten Informationskanals

Sprache ist also zuerst, wenn sie vom Sprecher gesendet wird, eine physikalische Erscheinung. Wenn sie dann vom Hörer empfangen wird, wird sie zuerst als akustisches Signal wahrgenommen. Auf beiden Seiten ist die Sprache von einem komplexen Regelwissen überlagert, dessen Existenz Voraussetzung für die sprachliche Kommunikation ist.

Fehlen einzelne Stufen dieses Regelsystems oder werden die Regeln falsch angewendet, ist die resultierende Sprache verstümmelt, unnatürlich oder unverständlich. Das läßt sich am Beispiel von Text-zu-Sprache-Synthesesystemen anschaulich demonstrieren.

Für die Umsetzung von phonetisch-artikulatorischen Steuerparametern in akustische Laute wurden diverse an der physiologischen Lautproduktion orientierte Modelle entwickelt. Die menschliche Stimme kann vereinfacht mit einem Blasinstrument verglichen werden. Von der Lunge wird ein Luftstrom durch den Kehlkopf geblasen, wobei die Stimmbänder (Stimmklappen) in Schwingungen versetzt werden. Diese stark obertonhaltigen Schwingungen können durch variierendes Anspannen der Stimmklappen frequenzmäßig beeinflußt werden. Die Schwingungsgrundfrequenz (F_0 , Pitch) liegt bei Männern etwa im Bereich von 80 bis 180 Hz und bei Frauen zwischen 150 bis 380 Hz.

Ein weitverbreitetes mathematisches Modell der Sprachproduktion (Klatt 1980) besteht aus einer Quellenanregungsfunktion $S(f)$, welche im einfachsten Fall entweder eine Pulsfolge (stimmhafte Anregung) oder ein Rauschsignal (stimmlose aperiodische Anregung) oder Kombinationen davon bezeichnet. Die Vokaltraktübertragungsfunktion $T(f)$ beschreibt die Resonanzen (Formanten) und die spektrale Umhüllende, welche durch die Bewegungen von Zunge, Gaumen, Zähnen und Lippen verändert

werden. Die Abstrahlungscharakteristik $R(f)$ schließlich beschreibt die Ausgangsimpedanz von Lippen und Nase zum Übertragungsmedium Luft.

3.1.2

Sprachperzeption

Analog zur schematischen Darstellung der Sprachproduktion kann auch für die Spracherkennung ein vereinfachtes Blockscha die wesentlichen Teilprozesse zusammenfassen. Das Modell von Abb. 3.2 (Pisoni u. Sawusch 1975) beschreibt die frühen Stufen der Sprachanalyse. Ein auditorisches Eingangssignal wird in einer Voranalyse als Frequenz-, Intensitäts- und Zeitmuster aufgeschlüsselt und in einer Art sensorischem Informationsspeicher für die weitere Analyse durch ein vierstufiges Erkennungssystem bereitgehalten. Information aus allen Verarbeitungsstufen wird zudem laufend in Kurz- und Langzeitspeicher abgelegt, welche wiederum den Verarbeitungsprozeß beeinflussen können.

Voraussetzung für die Wahrnehmung von Sprache als Sprache ist die Fähigkeit und das Regelwissen zur Analyse eines akustischen Signals auf höheren Abstraktionsebenen.

Die verschiedenen Regelsysteme, die bei der Erfassung und Auswertung aktiviert sind, greifen jedoch stark ineinander, so daß eine Trennung der verschiedenen Ebenen nur als methodologisches Vorgehen und nicht als exakte Beschreibung der sprachlichen Realität angesehen werden darf. Die Tatsache, daß die sprachliche Mitteilung als eine Folge von Ebenen betrachtet wird, deren Abstraktion ständig zunimmt, stellt eine Datenreduktion dar, bei welcher die Komplexität der Sprache nur unzureichend widerspiegelt wird.

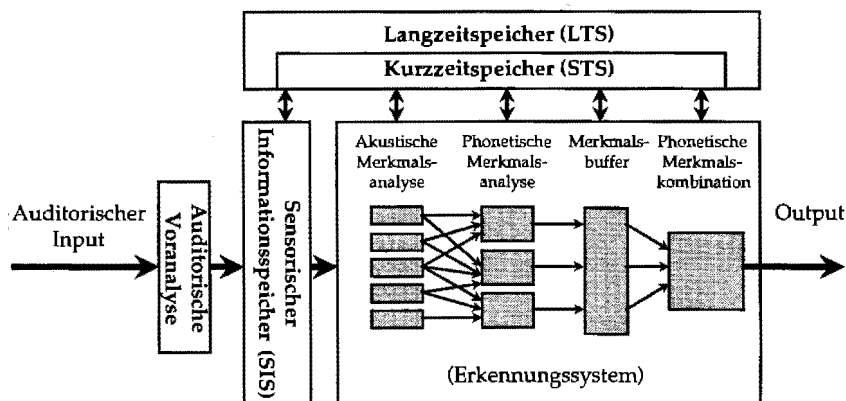


Abb. 3.2. Vereinfachtes Blockscha der Sprachwahrnehmung

3.1.3

Informationstransmissionsanalyse

Für die Analyse von Konsonantenverwechslungen schlugen Miller und Nicely (1955) mathematische Berechnungen vor, welche spezifische Verwechslungsmuster aufzeigen sollten. Diese Informations-transmissionsanalyse findet seither in zahlreichen audiologischen Untersuchungen Verwendung und wurde zuerst von Dowell et al. (1982) für die Verwendung in der CI-Forschung vorgeschlagen. Wird eine Verwechslungsmatrix gemäß bestimmten Gruppenmerkmalen (z.B. linguistische Klassifikationen) in kleinere Untermatrizen aufgeteilt, so sind die Meßergebnisse der Übertragung jeder einzelnen dieser Matrizen gleichbedeutend mit den Meßergebnissen, die man durch simultane Messung der Kommunikationskanäle erzielen würde. Diese Annahme ist allerdings nur dann zutreffend, falls die Merkmale unabhängig voneinander sind (orthogonal), was in den wenigsten Fällen möglich ist.

Eine Methode, Redundanzen in Merkmalsystemen zu korrigieren, schlagen Wang und Bilger vor (1973). Ihre Methode der Analyse sequentieller Informationstransmission eliminiert schrittweise die Merkmale mit der meisten Information, um die Analyse mit den jeweils reduzierten Merkmalen – die eliminierten Merkmale werden dabei konstant gehalten – zu wiederholen, bis noch verbleibende Information unterhalb eines kritischen Wertes ist bzw. alle Merkmale einbezogen wurden. Dieses Verfahren entspricht mathematisch dem der schrittweisen multiplen linearen Regressionsanalyse und erlaubt genauere Aussagen über Stellenwert und gegenseitige Beziehungen bestimmter Merkmale innerhalb eines Systems.

Beispiele von Konsonantenmerkmalen sind

- Stimmhaftigkeit (voicing),
- Nasalität (nasality),
- Frikation (frication),
- Artikulationsstelle (place of articulation) und
- Artikulationsart (manner of articulation).

Beispiele von Vokalmerkmalen sind

- die Zungenposition im Vokaltrakt (vorn, hinten, hoch, mittel, tief),
- die Lippenöffnung (gerundet und geschlossen),
- Vokaldauer und
- Formantfrequenzen.

Solche Merkmalsysteme erlauben eine systematische Datenreduktion bei Minimalpaarverwechslungstests. Die Interpretation der Ergebnisse von Experimenten mit CI-Codierungsstrategien ist jedoch durch die eher artikulatorische und weniger akustische Fundierung dieser Merkmale oft schwierig. Zu-

dem muß auch vor einer Überinterpretation solcher Merkmalsanalysen gewarnt werden, da sie von der Annahme isolierter Wahrnehmungseinheiten (Phoneme) ausgehen, welche zu Worten und Sätzen zusammengesetzt würden. Diese Vorstellung vereinfacht den Sprachperzeptionsmechanismus stark und ist deshalb nur begrenzt brauchbar.

3.1.4

Vocoder

Das Interesse an Sprachcodierungs- und Kompressionsverfahren war seit Beginn der Entwicklung moderner Kommunikationstechniken einerseits durch die Begrenzungen der Übertragungskanäle, andererseits durch die ständig wachsenden technischen Möglichkeiten bestimmt. Neben dem Ziel, beispielsweise über eine Telefonleitung möglichst viele Gespräche gleichzeitig übertragen zu können, waren auch Fragen der Störungsunempfindlichkeit sowie der sicheren (verschlüsselten) Übertragung motivierende Triebkräfte bei diesen Entwicklungen. Die Fortschritte der Multimediatechnologie wären ohne effiziente Audiokompression zur Übertragung und Speicherung digitaler Signale nicht möglich gewesen. Die Qualität komprimierter und codierter Sprache konnte dank Einbezug von psychophysischen Modellen der menschlichen Wahrnehmung beträchtlich gesteigert werden (Jayant 1993).

Die Informationsrate eines Audiosignals in CD-Qualität (Compact Disc, Stereo 44,1 kHz, 16 Bit) beträgt 1,41 Mb/s (MegaBits pro Sekunde). Das nutzbare Frequenzband beträgt dabei 20 kHz. Wird die zu übertragende Bandbreite auf 4 kHz reduziert (Telefonqualität), kann das Signal mit 128 kB/s codiert werden bzw. mit 8 Bit und einer logarithmischen Wandlerkennlinie mit 64 kB/s (PCM, „pulse code modulation standard“). Eine weitere Reduktion der Informationsrate ist praktisch nur noch mit Vocoder-technik und Sprachanalyse- und Synthesetechniken möglich. Qualitativ akzeptable Sprachcodierung kann heute mit etwa 2,5 kB/s durch Einsatz von sehr leistungsfähigen digitalen Signalprozessoren (DSP) erreicht werden. Das Kompressionsverhältnis bezogen auf 16 Bit-PCM beträgt dabei über 50.

Arten von Vocodern

Die ersten Vocoder waren sogenannte Kanalvocoder, die auf der Analyse- (Sender) und Syntheseseite (Empfänger) aus typischerweise 10 Bandfiltern bestanden, welche das Sprachsignal in Grundfrequenz und spektrale Komponenten aufschlüsseln und nach der Übertragung rekonstruieren konnten (Dudley 1939). Varianten mit und ohne Grundfrequenzextraktion sowie Vocoder mit Extraktion von Formanten oder spektralen Maxima wurden entwickelt. For-

mantvocoder erlaubten besonders effiziente Datenreduktion, indem sie nur die für die Spracherkennung wichtigsten Elemente, nämlich die spektralen Maxima sowie die Grundfrequenz, zu extrahieren und übertragen suchten (Flanagan 1972). Qualitativ waren jedoch die in den 70er Jahren entwickelten LPC-Vocoder („linear prediction coding“) den Formantvocodern überlegen. Bei der LPC-Analyse wird aus einem kurzen Signalabschnitt ein digitales Filtermodell des Vokaltrakts bei Anregung durch ein periodisches oder aperiodisches Signal berechnet. Die Parameter dieses Modells werden übertragen und bei der Synthese als Steuergrößen für ein periodisch oder aperiodisch angeregtes digitales Filter verwendet. Für die Datenreduktion macht man sich die Tatsache zu Nutze, daß die Vokaltrakteigenschaften sich innerhalb von etwa 5 bis 15 msec nur unwesentlich ändern und daß nicht jede Periode eines Vokalabschnitts, sondern innerhalb dieses Zeitfensters nur die mittlere Grundfrequenz übertragen werden muß.

Hörprothetische Anwendungen

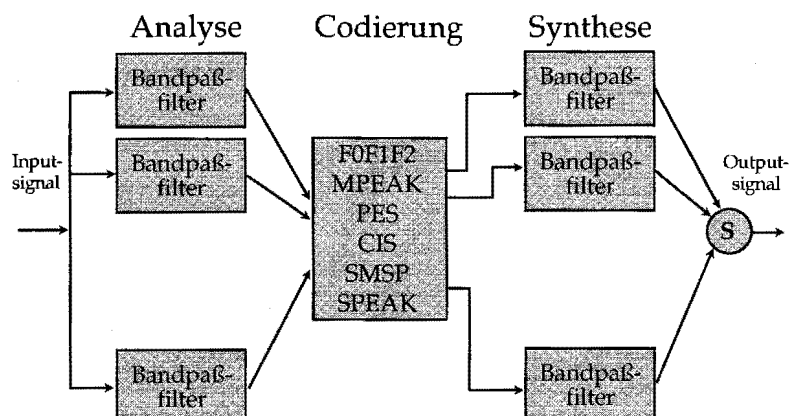
Die Anwendung dieser Vocoder-Technik auf die reduzierte Kanalkapazität bei elektrischer Hörnervstimulation ist nicht trivial. Voraussetzung für den Vocoder-Einsatz zur optimierten Nutzung beschränkter Kanäle ist natürlich, daß der Empfänger das codierte Signal wieder entsprechend decodieren kann. Bei technischen Übertragungskanälen wie Telefon- und Datenleitungen können Sender und Empfänger aufeinander abgestimmt werden. Bei hörprothetischen Anwendungen ist der Empfänger, das menschliche (Rest-)Gehör, vorgegeben. Allenfalls können durch geschickte Wahl von Elektrodenkonfigurationen (Cochlea-Implantaten) die Empfängercharakteristiken verbessert werden. Mit aller Phantasie kann jedoch das menschliche Hörsystem niemals soweit modifiziert werden, daß beispielsweise eines der der-

zeit erfolgreichsten technischen Vocoderverfahren (CELP, „codebook excited linear prediction vocoder“) angewendet werden könnte. Das Studium von Vocoder-Techniken kann jedoch für die praktische Verbesserung von Codierungsstrategien für Hörprothesen durchaus sinnvoll sein.

Anwendung bei Cochlea-Implantaten

Abbildung 3.3 stellt ein Beispiel eines Vocoder für Cochlea-Implantat-Signalverarbeitungsexperimente dar. Der Vocoder wurde auf einem Signalprozessor implementiert und kann in Echtzeit verschiedene Algorithmen für CI-Sprachprozessoren für Normalhörende demonstrieren. Der Analyseteil wurde durch eine Fourieranalyse mit anschließender Aufteilung in verschiedene Frequenzbänder realisiert. Die Synthese beruht auf einem einfachen Modell der auditorischen Wahrnehmung bei elektrischer Stimulation. Gemäß diesem Modell erregt eine Folge von Stimulationspulsen auf einer Elektrode in einem begrenzten Bereich der Basilarmembran eine Gruppe zugehöriger Neurone. Die Erregung entspricht der Anregung eines schmalbandigen Filters mit ebendieser Pulsfolge. Wahrgenommen wird dabei einerseits die charakteristische Frequenz des Filters, andererseits die Reizrate der Pulsfolge. Ein Einzelpuls erzeugt eine abnehmende Oszillation, deren Dauer von der Bandbreite des Filters abhängt. Die Summe aller parallel geschalteten Bandfilter entspricht dem komplexen Klangeindruck bei gleichzeitiger Anregung aller Stimulationskanäle. Zur Simulation der Lage der Reizelektroden können die Mittenfrequenzen von Analyse- und Synthesefiltern unterschiedlich gewählt werden. Die gegenseitige Beeinflussung benachbarter Stimulationskanäle kann ebenfalls durch Wahl eines Übersprechparameters variiert werden. Die Codierungsverfahren werden im Abschnitt über Cochlea-Implantate noch genauer beschrieben.

Abb. 3.3.
Cochlea-Implantat-Vocoder für unterschiedliche Sprachverarbeitungsstrategien



Das akustische Vocoder-Modell hat sich bei der Entwicklung dieser Verfahren und deren Evaluation mit Patienten als sehr nützlich erwiesen. Experimente mit Normalhörenden zum Vergleich der verschiedenen Codierungsstrategien erbrachten ähnliche Ergebnisse wie die Experimente mit CI-Trägern. Obwohl diese Vocoder-Simulationen den Höreindruck mit einem CI nicht vollständig wiedergeben können, sind sie zur Demonstration der Schwierigkeiten, mit denen ein implantierter Patient fertig werden muß, sowie als Richtschnur für das Hörtraining ein brauchbares Hilfsmittel. Obwohl Sprache nach der Reduktion auf einige hundert Bit/s qualitativ nicht mehr besonders gut klingt, ist sie nach etwas Übung und Gewöhnung durchaus verständlich.

3.2 Evaluationsverfahren

Für den analytischen Vergleich unterschiedlicher Codierungsstrategien eignen sich die traditionellen sprachaudiometrischen Tests nur bedingt. Beispielsweise sind die verschiedenen Wortgruppen des Freiburger Einsilbertests nicht äquivalent (Alich 1985). Die Mehrdeutigkeit und die fehlende phonologische Auswertungsmöglichkeit wurde von verschiedenen Autoren bemängelt (von Wedel 1985). Ansätze zu anderen Formen eines Sprachverständlichkeitstests sind schon seit langem bekannt. Unter anderen wurden bereits von Bárány (1910) und Lampert (1923) (zitiert in Feldmann 1976) sog. „Wechsel lautwörter“ vorgeschlagen, entsprechend phonologischen Minimalpaaren. Zur Einschränkung der Antwortmöglichkeiten wurden von amerikanischen Autoren Multiple-choice-Tests entwickelt (Owens u. Schubert 1977; Tyler et al. 1985). Für die deutsche Sprache entwickelte Sotschek einen Reimtest für Anwendungen in der Nachrichtentechnik (1982), welcher später für die Audiologie überarbeitet wurde (von Wallenberg u. Kollmeier 1989). Eine an die Arbeiten von Owens angelehnte deutsche Version von MAC-Batterie (Sprachtests für minimales Restgehör) wurde in Zürich entwickelt (Dillier u. Spillmann 1992). Weiterführende Information findet sich in (Kollmeier 1992).

3.2.1 Testanforderungen

Das Sprachmaterial soll zum einen in seiner Komplexität reduziert werden, zum anderen aber so aufgebaut sein, daß möglichst wirklichkeitsgetreue kommunikative Situationen damit beschrieben werden können. Ferner soll in einem sprachaudiometrischen Test die akustische Seite stärker betont werden als die

inhaltliche. Dies kann am besten erreicht werden, wenn phonetisch-phonologisches Material ausgewählt wird. Nun müssen auch hier Entscheidungen getroffen werden, die die Grundlage der Sprachaudiometrie betreffen. Der tragende Begriff der phonologischen Theorie, die Opposition zweier Phoneme in einem Minimalpaar (ein durch gleichzeitige akustische, artikulatorische und semantische Opposition entstehendes Wortpaar) bietet hier die Möglichkeit, gezielt Lautsegmente zu untersuchen und gleichzeitig Wörter als Träger dieser Information zu benutzen.

Zum Problem der phonetischen Ausgewogenheit von Testlisten sei hier nur kurz angemerkt, daß die Voraussetzung für gleiche oder annähernd gleiche Lautverteilung innerhalb der Wortlisten gemäß der anhand von Stichproben errechneten Häufigkeit der Laute in der betreffenden Sprache eine Statistik der gesprochenen Sprache (und nicht der Schriftsprache) voraussetzen würde. Die Sprachbeobachtung zeigt zudem, daß Phänomene wie Koartikulation, gegenseitige Beeinflussung und Störung die Norm in der natürlichen Sprache darstellen und somit die Effekte der Phonemhäufigkeit weitgehend überdecken.

Mit zunehmender Komplexität des verwendeten Sprachmaterials in Wort- und Satztests erhöht sich die Anzahl der Faktoren, die auf die Wahrnehmung dieses Prüfmateri als wirken. Die Sätze selbst können nach sehr unterschiedlichen Gesichtspunkten zusammengestellt werden, seien diese linguistischer (syntaktischer Aufbau, morphologische Wortstruktur, lexikalische Auswahl) oder quantitativer Art (Satzlänge und Dauer der gesprochenen Sätze, statistische Kennwerte des Materials). Somit verunmöglicht die linguistische Komplexität des Materials eine vorzugsweise akustisch orientierte Auswertung. Außerdem steigen die Anforderungen an eine systematische Analyse der Resultate. Andererseits wären Satztests insbesondere für Untersuchungen im Störlärm sehr hilfreich.

3.2.2 Beschreibung mit Wahrscheinlichkeitsmodellen

Der zunehmende Kontexteinfluss beim Erkennen von Einzellaute, Wörtern und Sätzen kann mit einfachen Wahrscheinlichkeitsmodellen beschrieben werden, welche experimentell verifiziert wurden (Boothroyd u. Nitttrouer 1988). Der Einfluß des lexikalischen Kontexts läßt sich ausdrücken als:

$$k = \frac{\log(1-p_c)}{\log(1-p_i)},$$

wobei k das Verhältnis der Erkennungswahrscheinlichkeit mit zu der Wahrscheinlichkeit ohne Kontext bedeutet. Für CVC-Kombinationen („consonant-vowel-consonant“) wurde k als etwa 1,3 ermittelt.

Auf Satzmaterial angewandt, resultierte ein Kontextfaktor von 1,38 für Sätze mit geringer Voraussagbarkeit (LP, „zero predictability“) bis 2,72 für Sätze mit großer Voraussagbarkeit (HP, „high predictability“).

Die Wahrscheinlichkeit für die Erkennung von aus mehreren Lauten (Phonemen) zusammengesetzten Silben lautet:

$$p_w = p_p^j,$$

wobei p_p die Wahrscheinlichkeit für die Erkennung eines isoliert gesprochenen Einzellauts bedeutet und j zwischen 1 und n (Anzahl der Phoneme) liegt. Für sinnlose CVC-Logtome beträgt j etwa 3, während für sinnvolle CVC-Worte ein Wert von etwa 2,5 typisch ist. Diese Modelle sind im Zusammenhang mit der Entwicklung von Sprachtests von Bedeutung, die Anwendung dieser Formeln im Einzelfall ist jedoch aufgrund der großen Variabilität nur beschränkt möglich.

3.3 Physiologische und technische Rahmenbedingungen

Der Hörnerv kann mit unterschiedlichen elektrischen Reizen angeregt werden. Für die Wahl der Reizart und der Stimulationsparameter sind sowohl elektrochemische Prozesse der Elektroden-Gewebe-Grenzschichten als auch physiologische und kommunikationstechnische Randbedingungen maßgebend. Als Material kommt für dauerhaft implantierte Elektroden praktisch nur Platin bzw. aus mechanischen Gründen eine Legierung aus Platin und Iridium in Frage.

Die für einen effektiven Reiz wirksame elektrische Größe ist der Strom, welcher während einer gewissen

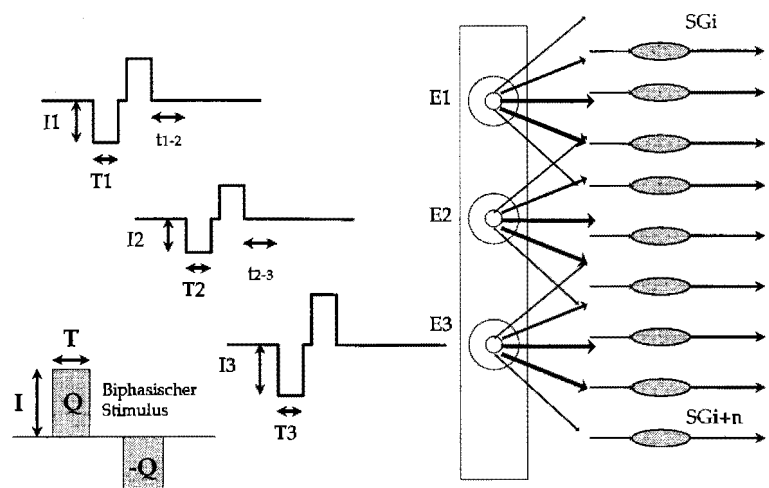
Zeit in das Nervengewebe hinein- oder herausfließt. Das Integral des elektrischen Stroms (I) über eine Zeitperiode (T) ist die elektrische Ladung (Q). Bei rechteckförmiger Stromform (pulsatile Reizung, s. Abb. 3.4) errechnet sich die Ladung als Produkt von Strom und Pulsdauer ($Q = I \cdot T$). Zur Vermeidung irreversibler elektrochemischer Prozesse am Elektroden-Gewebe-Übergang (Erosion von Platin, Gasbildung) ist unbedingt dafür zu sorgen, daß die Ladung, welche in das Gewebe hineingebracht wurde, auch wieder herausgeführt wird. Das elektrische Stimulationssignal muß also gleichstromfrei übertragen werden. Auf einen positiven Strompuls muß unmittelbar ein negativer Puls gleicher Ladung folgen. Da sich die Impedanz (das Verhältnis von Spannung und Strom) implantierter Elektroden mit der Zeit ändern kann, werden wenn möglich Konstantstromquellen verwendet, welche eine von der Elektrodenimpedanz unabhängige definierte Reizstärke ermöglichen.

3.3.1 Reizformen

Als Reizformen kommen sinusoidale (analoge) oder pulsatile Signale zur Anwendung. Die Vor- und Nachteile der unterschiedlichen Verfahren haben in der Vergangenheit Anlaß zu vielen Diskussionen geliefert, welche teilweise durch unklare oder unglücklich gewählte Begriffe erschwert wurden. Wenn von analoger und pulsatiler Reizung die Rede ist, wird oft der Gegensatz analog – digital damit assoziiert, welcher in der Audio- und Meßtechnik eine besondere Bedeutung hat.

Bei der analogen Stimulation wird meist ein gefiltertes und komprimiertes Signal kapazitiv auf die Elektroden gelegt. Der Vorteil bei diesem Verfahren

Abb. 3.4.
Kanalinteraktion (Übersprechen) bei Mehrkanalstimulation (E Elektrode, SG Spiralganglion)



$$\text{Ladung (Q, nC)} = \text{Amplitude (I, mA)} : \text{Pulsbreite (T, } \mu\text{sec)}$$

ist die relativ einfache Verarbeitung, da die Signale ähnlich wie bei einem konventionellen Hörgerät nach Mikrofon, Verstärker und Filter dem Implantat direkt zugeführt werden können. Dieses Verfahren eignet sich recht gut für eine einkanalige Prothese. Sobald allerdings mehrere unabhängige Kanäle vorhanden sind, ergeben sich Lautheitssumationen durch Stromüberlagerungen.

Biphasische Pulse erlauben eine zeitlich besser kontrollierbare Reizung als analoge Sinussignale. Die Ladung pro Puls, welche für die empfundene Lautheit in erster Linie maßgebend ist, kann jederzeit genau definiert werden. Insbesondere bei der Stimulierung von mehreren Kanälen können Interferenzen vermindert werden, indem immer nur ein Kanal gleichzeitig aktiviert wird. Nachteilig wirkt sich unter Umständen bei der sequentiellen Stimulation die verminderte Zeitaufösung aus, da pro Reizpuls immer nur die Information eines Frequenzbandes übertragen wird, die anderen also inaktiv bleiben.

3.3.2

Anordnung der Elektrodenpaare

Voraussetzung für die elektrische Stimulation ist eine genügende Anzahl elektrisch erregbarer Nervenzellen. Die Lautheitsempfindung hängt vermutlich von der Anzahl gleichzeitig aktivierter Zellen ab. Durch die Lage der Stimulations- und Referenzelektroden kann der aktivierte Bereich beeinflusst werden. Bei weit auseinanderliegenden Elektroden werden bei gleicher Reizstärke mehr Neuronen aktiviert als bei nahe beisammenliegenden. Dadurch reduzieren sich die Hörschwellen mit zunehmender Distanz der Elektrodenpaare.

Die Lage der Elektrodenpaare bezüglich des Trägerkörpers ist für die Hörempfindung ebenfalls bedeutsam. Radial angeordnete Elektroden erlauben eine engere Lokalisation der neuronalen Erregung als longitudinale Elektrodenpaare (van den Honert u. Stypulkowski 1987b). Eine lokal begrenztere Erregung ist wünschenswert, weil damit die tonotopie Frequenzselektivität besser erhalten bleibt, bzw. die Unabhängigkeit der verschiedenen Kanäle besser gewährleistet ist.

3.3.3

Bestimmung der Wahrnehmungsschwelle

In der Praxis muß für jede Elektrodenkonfiguration die Wahrnehmungsschwelle („threshold“, T-level) sowie die Reizstärke für eine angenehme, erträgliche Lautheitsempfindung („comfort“, C-level, manchmal auch MCL, „most comfortable level“ genannt, bzw. durch UCL, „upper comfortable level“, ersetzt) bestimmt werden (s. Abb. 3.5). Der Bereich zwischen

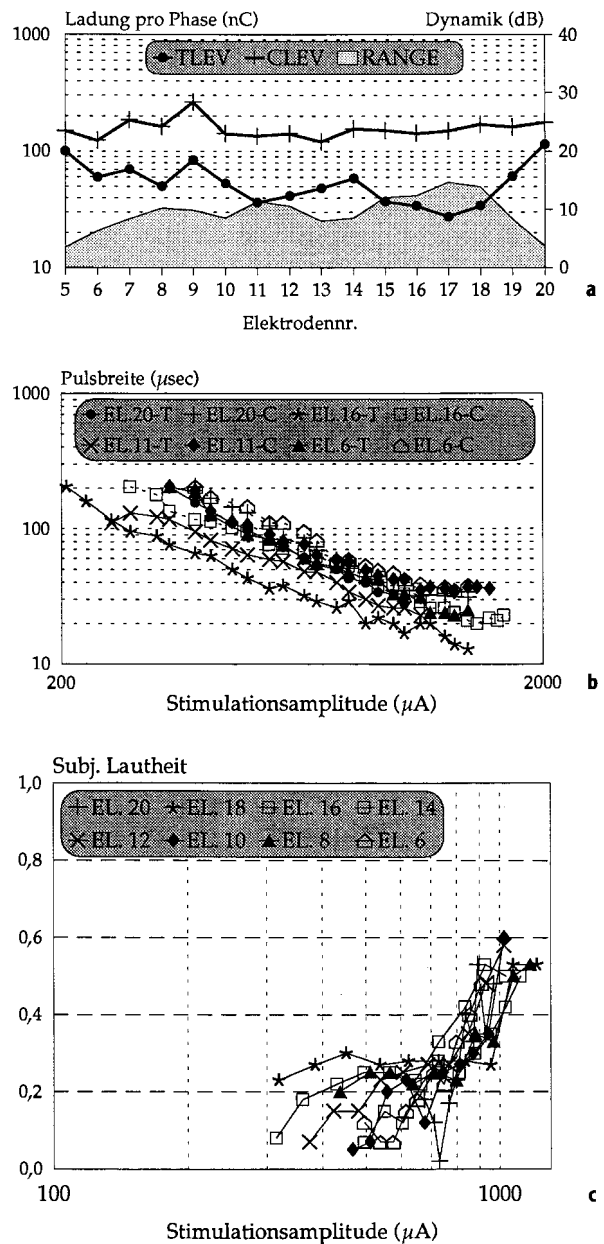


Abb. 3.5 a–c. Psychophysische Messungen mit Cochlea-Implantaten. **a** Bestimmung von Schwellen (TLEV „threshold level“, CLEV „comfortable level“) und Dynamik (RANGE ausnutzbarer Hörbereich), **b** Verhältnis von Stimulationsamplitude (in μA) und Pulsbreite (in μsec) bei gleicher subjektiver Lautheit und verschiedenen Stimulationselektroden, **c** Lautheitsfunktionen: Subjektive skalierte Lautheit als Bereich zwischen unhörbar (0,0) und unerträglich laut (1,0) für verschiedene Elektroden; hier am Beispiel der Patientin U. T. mit bipolarer Stimulation und einer Pulsbreite von 150 μsec /Phase (Ausnahme: Elektrode 20 mit einer Pulsbreite von 200 μsec /Phase)

diesen Werten (dynamischer Bereich) liegt im Schnitt bei 6 bis 12 dB (bezogen auf 1 nC Ladung/Phase) und ist damit wesentlich kleiner als die 120 dB SPL („sound pressure level“, Schalldruckpegel) bei der akustischen Stimulation. Die Schwellwerte sind abhängig von der Stimulationsrate. Bei höheren Pulsraten sinkt die Hörschwelle etwas ab, während der C-Level mehr oder weniger konstant bleibt, so daß insgesamt die Dynamik zunimmt. Durch Modellbetrachtungen der Erregungsmechanismen an der Nervenmembran (Finley et al. 1989) läßt sich dieses Phänomen erklären. Das Ausmaß der Schwellenabhängigkeit von der Reizrate ist jedoch individuell und muß insbesondere für die neuen vielversprechenden Codierungsstrategien mit hohen Reizraten berücksichtigt werden.

Die Lautheitsempfindung bei der elektrischen Stimulation ist eine nichtlineare Funktion der applizierten Ladung/Phase, welche sich mit einer Potenzfunktion annähern läßt. Abbildung 3.5 zeigt Beispiele für Lautheitsskalierungsmessungen bei Stimulation unterschiedlicher Elektroden. Die maximal mögliche Ladung/Phase sollte bei den verwendeten Elektrodengeometrien auf ca. $30 \mu\text{C}/\text{cm}^2$ beschränkt werden, um Schäden an den Elektroden und der biologischen Umgebung zu verhindern (Shepherd et al. 1990).

Die elektrische Stimulation mit pulsatilem Signalen löst meistens zwei unterschiedliche Arten von Empfindungen aus. Die repetitive Anlegung der Stimuli an einer Elektrode (Stimulationsfrequenz, -rate, Pitch) wird als eine Empfindung von Tönen bzw. von der Grundfrequenz der Stimme beschrieben. Die Ratendiskrimination ist allerdings nur bis etwa 300 Pulse/s mit der von Normalhörenden vergleichbar. Bei höheren Raten fällt die Unterscheidung zunehmend schwieriger.

Die Stimulation an verschiedenen Orten innerhalb der Cochlea ruft Empfindungen hervor, die mit der tonotopen Frequenzempfindung bei Normalhörenden verglichen werden kann. Die wahrgenommene charakteristische Frequenz hängt dabei von der Eindringtiefe des Elektrodenträgers ab. Mit den über 17 mm verteilten 22 Elektroden des Nucleus-Implantats beispielsweise kann bei einer Eindringtiefe von ca. 20 mm der Frequenzbereich zwischen 700 und 11 000 Hz angesprochen werden, bei einer Eindringtiefe von 17 mm (alle aktiven Elektroden innerhalb der Cochlea) jedoch nur noch der Bereich von 1000 bis 20 000 Hz. Die Frequenzselektivität ist durch die Fähigkeit des Implantats bestimmt, Neuronen örtlich möglichst begrenzt zu erregen. Da die Hörschnecke ein dreidimensionales Gebilde ist, kann eine elektrische Stimulation auch zu Erregungen außerhalb der tonotopen Frequenzskala führen, was selbstverständlich die Frequenzwahrnehmung für den Implantatträger erschwert.

3.3.4

Modelle der Hörnervstimulation

Die elektrische Hörnervreizung bezweckt die Erzeugung von Aktivierungsmustern im Hörnerv, welche eine Decodierung der übermittelten Information durch den zentralen Auswertungs- und Wahrnehmungsmechanismus ermöglichen. Eine von vielen Forschern auf diesem Gebiet postulierte Arbeitshypothese besagt, daß ein dem natürlichen Aktivierungsmuster bei akustischen Reizen möglichst ähnliches Muster durch elektrische Reize optimale Höreindrücke vermitteln würde und deshalb anzustreben sei. Somit ist bei der Suche nach geeigneten Verfahren zur Hörnervstimulation zuerst einmal von Modellen der normalen Hörnervaktivität auszugehen und aus vergleichenden Experimenten mit akustischer und elektrischer Reizung nach Gemeinsamkeiten und Unterschieden zu suchen.

Der normale Vorgang der Nervenregung im Innenohr kann vereinfacht mit dem durch verschiedene Autoren erweiterten Davis-Modell (MacMullen u. Mountain 1985) beschrieben werden. Dieses Modell läßt sich auch sehr schön zur Veranschaulichung der elektroanatomischen Verhältnisse in der Cochlea verwenden, welche sowohl bei intra- wie auch bei extracochleären Elektrodenplatzierungen nicht zu vernachlässigen sind – Elektroden direkt im Hörnerv bzw. im Hirnstamm haben derzeit praktisch wenig Bedeutung (Spelman et al. 1982). Modelle von Nervenfaserpopulationen wurden von verschiedenen Autoren soweit entwickelt, daß sie die Hörnervaktivität auch bei komplexen akustischen Reizen unter Störlärmbedingungen beschreiben können (Colombo u. Parkins 1987; Deng u. Geisler 1987; Ghitza 1987).

Bei elektrischer Reizung spielen Modelle der elektrischen Feldkonfiguration eine große Rolle, da zwischen Reizelektrode und Nervenfasern unterschiedlich leitende Medien liegen. Vermehrt werden zur Modellierung zwei- und dreidimensionale Rekonstruktionen der Cochlea mit finiten Elementen oder Differenzen eingesetzt. Die modellmäßig charakterisierten elektrischen Felder werden dann in einem weiteren Schritt als Eingangsgrößen für neurale Membranmodelle verwendet (Finley et al. 1989).

Verschiedene Hypothesen der neuronalen Erregung mit analogen und pulsatilem Signalen wurden auch in psychophysischen Experimenten mit CI-Trägern verifiziert (Dobie u. Dillier 1985; Zeng u. Shannon 1992), um beispielsweise die erstaunlich guten Sprachverständlichkeitsleistungen mit einkanaliger Stimulation erklären zu können. Obwohl Modellsimulationen von Einzelfaseraktivitätsmustern teilweise sehr gute Übereinstimmungen mit den psychophysischen Resultaten lieferten (Motz u. Rattay 1986), zeigten elektrophysiologische Untersuchungen klar auf, daß die

beobachteten Phänomene mit der Aktivität einzelner Neurone nicht zu erklären waren, sondern nur mit Betrachtungen, welche ganze Arrays von Nervenfasern gleichzeitig in Betracht ziehen (Hartmann et al. 1984; van den Honert u. Stypulkowski 1987a).

3.4 Typen von Cochlea-Implantaten – Stand der Technologie

Die Zahl der CI-Träger dürfte 1995 weltweit bereits 10 000 übersteigen. Die meisten bisher entwickelten Systeme bestehen aus den folgenden Komponenten:

- Einem Stimulator (Sprachprozessor) von der Größe eines Taschenhörgeräts, welcher die Reizcodierungselektronik und eine Batterie enthält. Er wird am Körper getragen.
- Einem über ein Kabel mit dem Stimulator verbundenen Sender oder (nur noch selten) Hautstecker hinter dem Ohr.
- Dem via Mastoidknochen und Mittelohr chirurgisch in das Innenohr eingesetzten Implantat (Empfänger und Elektroden). Als Zugang zum Innenohr und zum Hörnerv hat sich das runde Fenster der Schnecke (Cochlea) bzw. eine separate Öffnung in der ersten Schneckenwindung als am besten geeignet erwiesen. Bei extracochleärer Reizung werden die Elektroden in Kontakt mit dem runden Fenster gebracht, bei intracochleärer Platzierung vollständig in die Cochlea eingelegt.

Die derzeit bekanntesten kommerziell erhältlichen Implantate sind in der Tabelle 3.1 aufgelistet. Aus der Zusammenstellung wird deutlich, daß praktisch alle Systeme mittlerweile für eine intracochleäre Platzierung ausgelegt sind. Die ursprünglichen Bedenken wegen eventueller Schädigungen verbleibender Innenohrstrukturen konnten glücklicherweise durch die mittlerweile bereits mehrere Jahrzehnte währende praktische Erfahrung mit intracochleären Elektroden sowie die bereits mehrfach erprobte Möglichkeit eines Implantat-Ersatzes ohne nachteilige Folgen zerstreut werden.

3.4.1 Sprachdiskrimination

Ein Vergleich bezüglich der erzielten Sprachdiskrimination ist recht schwierig aus Gründen mangelnder Standardisierung von Sprachtests und unterschiedlichen sprachlichen und biologischen Hintergründen der Patienten. Mit extracochleären und mit intracochleären Systemen werden Umweltergüsse wahrgenommen, Stimmen erkannt und unterschieden und einfache Sprachinhalte verstanden. Die grundsätzliche Überlegenheit mehrkanaliger Geräte bei korrekter Elektrodenlage und adäquater Reizcodierung wird jedoch heute kaum mehr ernsthaft bestritten. Durch die gleichzeitige Anwendung des Periodizitäts- und Ortsprinzips für die Informationsübertragung ist die Kanalkapazität bei

Tabelle 3.1. Die wichtigsten Implantstypen

Typ	Aktive Kanäle	Intra Extra	Art der Stimulation	Signalübertragung	Kodierungsstrategie
Digisonic (früher Chormac, Chouard, 1985)	12	I	BP, pulsatil, sequentiell	3,2 MHz (PM)	Kanalvocoder, digital
AllEar (früher 3M/House, Fretz, 1985)	1	I	analog 340–2700 Hz	16 kHz (AM)	Amplitudenmodulation, analog
Ineraid (Richards, früher Symbion, Eddington, 1980)	6	I	MP, analog, simultan	Stecker	Filterbank, analog
Laura (Antwerp Bionic Systems, Peeters, 1989)	8 BP 15 MP	I	BP, MP, pulsatil, simultan	10 MHz (PM)	CIS, digital
Nucleus (Cochlear Pty, Clark, 1990)	22 CG 21 BP	I	BP, Bp+n, MP, CG, puls, sequentiell	2,5 MHz (PM)	Vocoder mit „Feature Extraction“, Filterbank, digital
Clarion (Advanced Bionics, früher UCSF, Storz, Merzenich, 1985)	4	I	BP, MP, CG, analog, simultan	10 MHz (FM)	CIS und Analogstimulation, digital
Combi-40 (MedEl, früher 3M/Vienna, Hochmair, 1989)	1 bis 8	E oder I	analog/pulsatil 100–6000 Hz	12 MHz (AM)	CIS und Analogstimulation, digital

BP bipolar, MP monopolar; CG „common ground“;
PM Pulsmodulation; AM Amplitudenmodulation;
FM Frequenzmodulation; CIS „continuous interleaved sampler“.

Mehrkanalsystemen potentiell bedeutend größer als bei Einkanalsystemen, so daß viele CI-Träger mittlerweile sogar ohne Lippenablesen kommunizieren und Telefongespräche führen können. Das Fehlen von nennenswerten Risiken bei korrekter Operationstechnik hat sich in bereits über 20 Jahren Benutzungsdauer bestätigt.

3.4.2

Elektrodenanordnung

Die optimale Anordnung der Reizelektroden auf dem Elektrodenträger und deren Ausrichtung innerhalb der Schneckenwindungen ist ein aktuelles und noch offenes Forschungsthema. Inwiefern Unterschiede mit den einzelnen Implantatmodellen auf Variationen der Elektrodengeometrie und -konfiguration zurückzuführen sind oder von Patientenvariablen oder der Art der Signalcodierung abhängen, wird die weitere wissenschaftliche Arbeit zeigen.

3.4.3

Sprachcodierung

Bezüglich der Sprachcodierung haben die Forschungsarbeiten der letzten zehn Jahre zu einer beträchtlichen qualitativen Verbesserung und ebenso zu einer Vereinheitlichung der zugrundeliegenden Konzepte und Strategien geführt. Die reine Merkmalsextraktion gemäß dem Schema eines Formantvocoders, wie sie im ersten Nucleus-Sprachprozessor realisiert war (WSP, „wearable speech processor“) wurde zugunsten immer detaillierterer zeitlich besser aufgelöster Spektralinformation aufgegeben. Die reine Breitband-Analog-Stimulation, wie sie etwa im 3M/Vienna-System für einkanalige oder im Symbion/Ineraid-System für mehrkanalige Konfigurationen realisiert wurde, wurde von mehreren Forschungsgruppen durch pulsatile Reizung möglichst hoher Raten (CIS, „continuous interleaved sampler“) ersetzt. Besonders bei Elektrodenkonfigurationen mit geringem Abstand zwischen benachbarten Elektroden bietet die nichtsimultane Pulsreizung gegenüber simultaner Analogreizung gewichtige Vorteile.

3.4.4

Signalübertragung

Mit einer Ausnahme sind alle derzeit gebräuchlichen Implantate mit transkutaner Signalübertragung (also mit Hochfrequenzmodulation unterschiedlicher Art) ausgestattet, da der heutige technologische Stand den Einsatz von Hautsteckern außer für rein wissenschaftliche Zwecke obsolet erscheinen läßt.

3.4.5

Zuverlässigkeit

In den frühen 80er Jahren wurden erste Versuche mit Cochlea-Implantaten bei Kindern durchgeführt (Eisenberg u. House 1982). Nachdem mit den damaligen einkanaligen Implantaten auch nach vielen Versuchen bei Erwachsenen keine überzeugenden Ergebnisse erzielt werden konnten, war es für die Fachleute keine große Überraschung, daß die Erfolge bei den implantierten Kindern sich in Grenzen hielten (Popelka u. Gittelman 1984). Zudem ließ die Zuverlässigkeit der Systeme sehr zu wünschen übrig. Wie unzuverlässig die Implantate der ersten Generation funktionierten, zeigte sich erst nachträglich, als die Ausfallhäufigkeiten mit unterschiedlichen Implantatmodellen veröffentlicht wurden (von Wallenburg et al. 1993). Daß nach zwei Jahren beinahe ein Fünftel aller 3M/House-Implantate mit Epoxyverkapselung ausgefallen war, ist ein deutlicher Hinweis darauf, daß die Technologie damals noch nicht genügend ausgereift war. Obwohl sich glücklicherweise in mehreren Fällen ein defektes Implantat problemlos durch ein neues ersetzen ließ, ist ein solches System für die Anwendung bei Kindern kaum zu verantworten. Ähnlich problematische Erfahrungen wurden auch bezüglich der Zuverlässigkeit von Sprachprozessoren berichtet (Biesalski et al. 1985).

Daß die Situation durch technische Verbesserungen heute grundsätzlich anders ist, zeigt die Ausfallrate von 1,33% für das Nucleus Mini System 22, welche auf einer Anzahl von 4623 implantierten Geräten basiert (von Wallenburg et al. 1993). Es ist zu hoffen, daß diese Rate künftig noch weiter reduziert werden kann.

3.5

Experimentelle Untersuchungen verschiedener Sprachcodierungsstrategien

Die aktuellste Frage im Bereich der Forschung und Entwicklung für Cochlea-Implantate ist die Suche nach geeigneten Codierungsstrategien zur Unterscheidung bedeutungstragender sprachlicher Merkmale und die Optimierung entsprechender Anpassungs- und Evaluationsverfahren.

Das Grundproblem der Sprachcodierung ist die beschränkte Kanalkapazität für die Übertragung der für das Sprachverständnis notwendigen Informationseinheiten.

Durch eine geeignete Transformation muß sichergestellt werden, daß die akustischen Schallwellen in die adäquaten elektroneuralen Erregungsmuster über-

setzt werden. Dabei müssen elektrophysiologische Maskierungseffekte und Interaktionen der einzelnen Reizelektroden so weit als möglich vermieden werden. Die je nach Pathologie und Lage der Reizelektroden unterschiedliche elektroanatomische Reizsituation macht eine individualisierte Intensitäts- und Frequenzzuordnung erforderlich.

Bei der analogen Signalverarbeitung wird das Eingangssignal durch Bandfilter in mehrere Frequenzbänder aufgeteilt, welche mit Verstärkerbausteinen in adäquate Stimulationssignale transformiert werden. Vor allem bei eng zusammenliegenden Elektroden lassen sich elektrophysiologische Maskierungseffekte sowie elektrische Summation der Erregungsfelder kaum vermeiden. Durch die Verwendung nichtsimultaner pulsartiger Reizformen kann ein Großteil dieser negativen Phänomene vermieden werden. Wie von Shannon und Mitarbeitern und anderen experimentell gezeigt wurde, findet zwar auch eine gewisse zeitliche Integration und Nachverdeckung bei Reizung mehrerer Elektroden statt, welche jedoch gegenüber den Interaktionseffekten bei simultaner Reizung praktisch vernachlässigt werden kann. Durch geschickte Wahl der sequentiell aktivierten Elektroden kann zudem der Integrationseffekt zusätzlich verringert werden. In den letzten Jahren wurden insbesondere durch Verwendung höherer Reizarten beträchtliche Fortschritte bezüglich Sprachverständlichkeit erzielt.

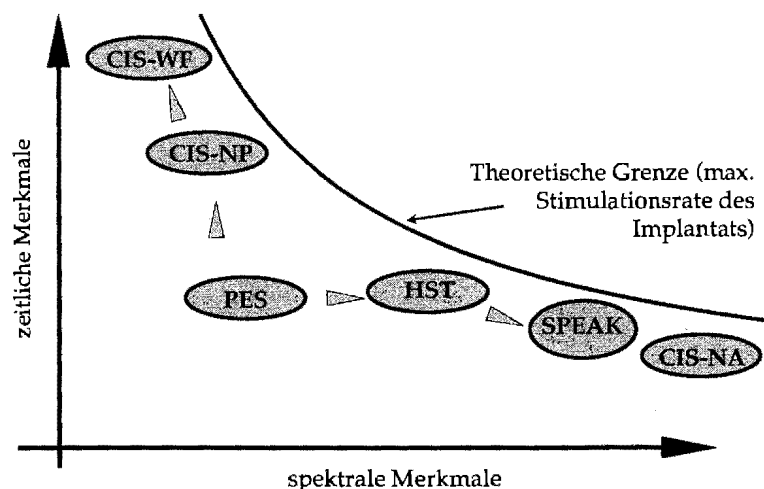
Die technischen Grundlagen des heute am weitesten verbreiteten Nucleus-Implantats wurden in verschiedenen Publikationen ausführlich beschrieben (Clark et al. 1990). Die implantierte Elektrodenansteuerung ist genügend flexibel, um auch komplexere Reizschemata als das im derzeitigen Sprachprozessor realisierte zu verarbeiten. Somit kann durch entsprechende Signalverarbeitung in einem Laborsystem eine Vielzahl von Sprachcodierungsstrategien reali-

siert und die bezüglich Sprachqualität und -verständlichkeit optimale Lösung in interaktiven Tests mit den Patienten gefunden und in einen tragbaren Sprachprozessor einprogrammiert werden. Für Feldversuche über mehrere Wochen wurde deshalb auch ein eigener tragbarer digitaler CI-Prozessor entwickelt (Dillier et al. 1990).

Auf einen für die Signalverarbeitung wichtigen Aspekt soll noch kurz hingewiesen werden: Der eingebaute Empfänger des derzeitigen Implantats wandelt Sequenzen, bestehend aus sechs RF-Pulsen (Radiofrequenzpulsen) variabler Länge in biphasische Elektrodenstimuli. Die effektive Stimulationsdauer ist durch die positive und negative Pulsweite (biphasischer Puls) gegeben. Die Datenübertragung zum Implantat dauert aber wesentlich länger, da pro Sequenz von 6 RF-Pulsen jeweils nur ein Stimulus codiert ist. Daher ist die maximale Stimulationsrate durch die Datenübertragung zwischen Sender (Spule) und Empfänger (Implantat) begrenzt. Bereits angekündigte künftige Implantate sind bezüglich maximaler Stimulationsrate jedoch nicht mehr eingeschränkt (Patrick 1995). Die in Abb. 3.6 aufgeführten Sprachcodierungsstrategien können im Sinne einer Optimierungsaufgabe gemäß vorwiegend zeitlicher oder spektraler Auflösung klassifiziert werden. Die asymptotische Grenzlinie ist durch die maximale Datenübertragungsrate sowie die patientenspezifischen physiologischen und anatomischen Eigenschaften bedingt, welche mittels Testmessungen ermittelt werden können.

Bei der Suche nach verbesserten Codierungsstrategien für Cochlea-Implantate ist die Reduktion der übertragenen Sprachinformation auf die für die Verständlichkeit wesentlichen Merkmale ein Arbeitsziel, welches durch die begrenzte Kanalkapazität aller bisher bekannten Neurostimulationssysteme vorgegeben ist.

Abb. 3.6.
Codierungsstrategien: Optimierung der zeitlichen und spektralen Auflösung



Neben Strategien, welche vor allem spektrale Merkmale wie Formanten bzw. spektrale Peaks zur Sprachcodierung verwenden, wurden auch Strategien untersucht, welche mehr die zeitliche Feinstruktur der Sprachsignale in den Vordergrund stellen. Wie das bekannte Beispiel synthetisch generierter stimmhafter Plosivlaute /ba/, /da/, /ga/ zeigt, ist der einzige Unterschied zwischen diesen drei Silben die etwa 30 bis 40 ms dauernde Frequenzänderung des zweiten und dritten Formanten. Wird diese Information durch diskrete Pulsreize im Takt der Grundfrequenz auf die entsprechenden Elektroden übertragen, steht dem Patienten zur Diskrimination dieser Silben nur eine äußerst minimale Informationsmenge in der Form von 3 bis 4 Reizpulsen zur Verfügung. Wenn hingegen während der Formanttransitionen entweder mehrere benachbarte Elektrodenkanäle gleichzeitig (z.B. SPEAK-Strategie) oder aber die gleichen Elektroden mit höherer Pulsrate (z.B. CIS-Strategie) angeregt werden, ergibt sich eine größere statistische Wahrscheinlichkeit zur richtigen Diskrimination der Konsonant-Vokal-Übergänge.

3.5.1

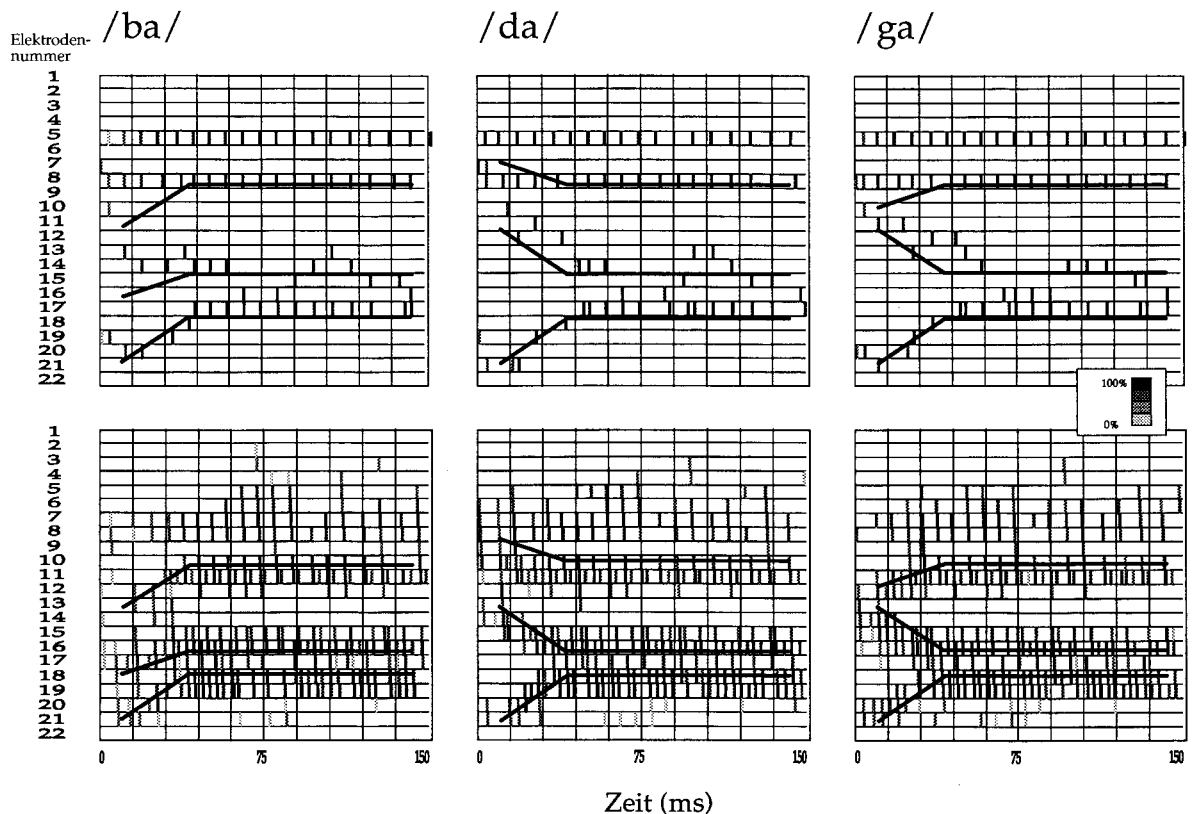
Analyse der Verfahren

Mit einem speziell entwickelten Programm können die auf ein Implantat übertragenen Reizparameter

zwei- und dreidimensional visualisiert und systematisch analysiert werden. Abbildung 3.7 zeigt Elektrodogramme mit MPEAK (Multipeak-Strategie) und SPEAK (Spectral-peak-Strategie). Die Schwärzung der einzelnen kleinen Rechtecke bezeichnet in etwa die mit dem Stimulus verbundene subjektive Lautheit. Aus den Darstellungen wird deutlich, daß das MPEAK-Verfahren durch die pro Zeitabschnitt beschränkte Anzahl aktivierter Elektroden sowie die tiefe, pitch-synchrone Reizrate die Formantübergänge relativ schlecht und fehlerhaft überträgt, während die SPEAK-Strategie durch die höheren Reizraten und die Möglichkeit, benachbarte Elektroden praktisch gleichzeitig anzuregen, eine naturgetreuere Nachbildung der im Sprachsignal enthaltenen Information erlaubt.

Die Eigenschaften der Sprachprozessorstrategien wurden für verschiedene Klassen von Sprachlauten und bei zusätzlichem Störlärm vor allem mit Hilfe dieses Visualisierungsprogramms analysiert. Auf-

Abb. 3.7. Codierung der stimmhaften Plosive mit zwei verschiedenen Strategien (*oben* MPEAK, *unten* SPEAK). Elektrode 1 ist die äußerste Elektrode entsprechend der höchsten wahrgenommenen Ortsfrequenz, Elektrode 22 die innerste Elektrode mit der niedrigsten tonotopen Frequenzempfindung



grund dieser Analysen und theoretischer Überlegungen wurden die Verarbeitungsalgorithmen verfeinert, indem zum Beispiel spezifische Eigenschaften von Sprachsignalen wie spektraler Schwerpunkt hochfrequenter Phoneme oder Formantprofil von Vokalsegmenten zur Verschärfung phonologischer Kontraste eingesetzt wurden. Virtuelle Stimulationselektroden wurden durch Variationen der Anzahl und zeitlichen Abfolge von Stimulationspulsen generiert und führten in der Analyse zu einer potentiell verbesserten Unterscheidung von Vokalen und Konsonanten (Bögli 1993).

Untersuchungen von Wilson et al. (1991) wiesen nach, daß mit hohen Reizraten auf mehreren Stimulationskanälen gegenüber tiefen Pulsraten und vor allem gegenüber Analogsignalen eine deutliche Verbesserung des Sprachverständnisses möglich ist. Diese Untersuchungen konnten bisher nur an Patienten mit Hautsteckern durchgeführt werden, da die technischen Voraussetzungen bei den Patienten mit Nucleus-Implantaten fehlten. Unser Ansatz mittels eigens entwickelter Interface-Hardware und einem experimentellen DSP-Sprachprozessor ermöglichte nun auch für diese Patienten einen Vergleich unterschiedlicher Strategien. Abbildung 3.8 illustriert fünf mögliche Codierungsarten (PES: „pitch excited sampler“, CIS-WF/NA: „continuous interleav-

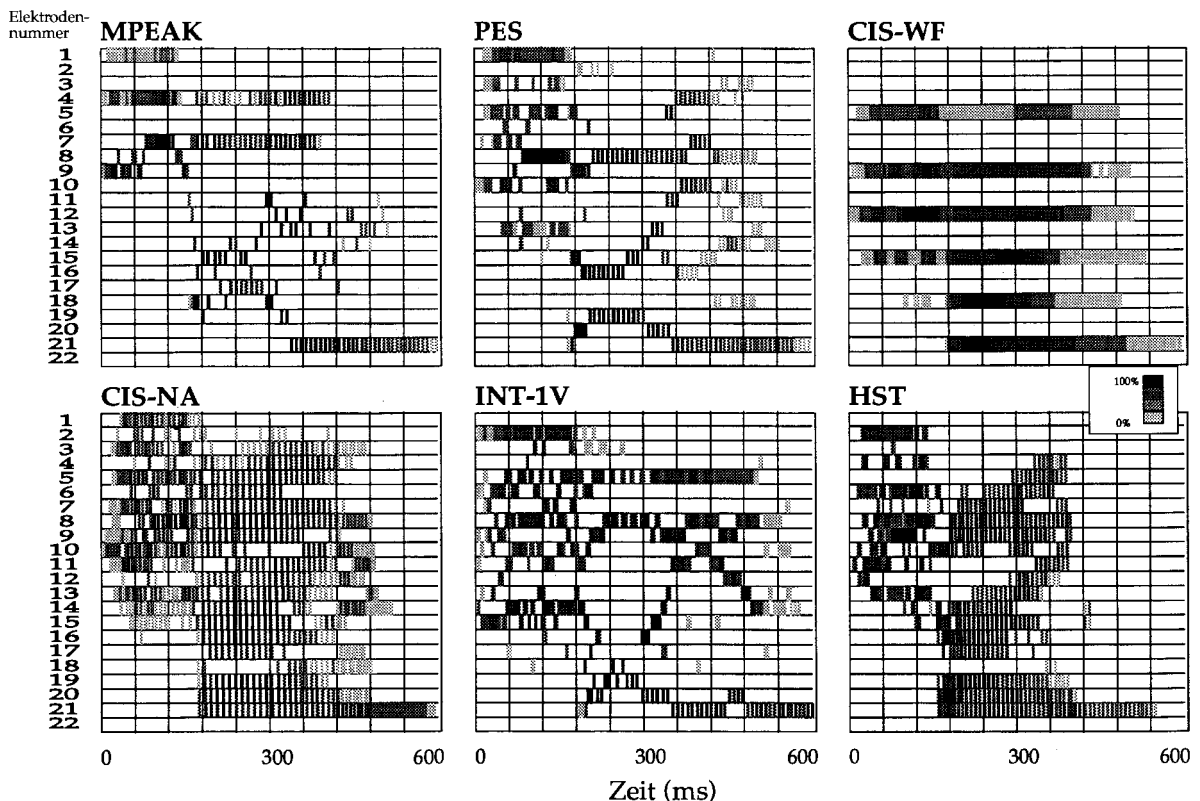
ed sampler wide frequency band analysis/narrow band analysis“, INT-1V: integrated hybrid strategy with one voice excited stimulation channel“, HST: „high spectral transmission“), welche mit CI-Trägern im Labor mit der Codierung des Nucleus-MSP-Sprachprozessors (MPEAK) verglichen wurden. Zusätzlich zu diesen Grundstrategien wurden weitere Verarbeitungsstufen zur Kontrastverschärfung von Vokal- und Konsonantenunterschieden eingebaut.

3.5.2

Evaluationsexperimente mit Patienten

Wie gut die durch die Signalcodierung reduzierte und auf die implantierten Elektroden übertragene Sprachinformation im Einzelfall zur Diskrimination von Wörtern und Sätzen ausreicht, kann nur durch Evaluationsexperimente mit Patienten ermittelt werden. Solche Experimente sind naturgemäß aufwendig und von vielen Variablen abhängig. Deshalb wurden nur die erfolgversprechendsten Strategien mittels Sprachtests in der Laborsituation untersucht, um systematisch die wichtigsten Parameter variieren zu

Abb. 3.8. Codierung des Wortes „Schein“ mit 6 verschiedenen Strategien

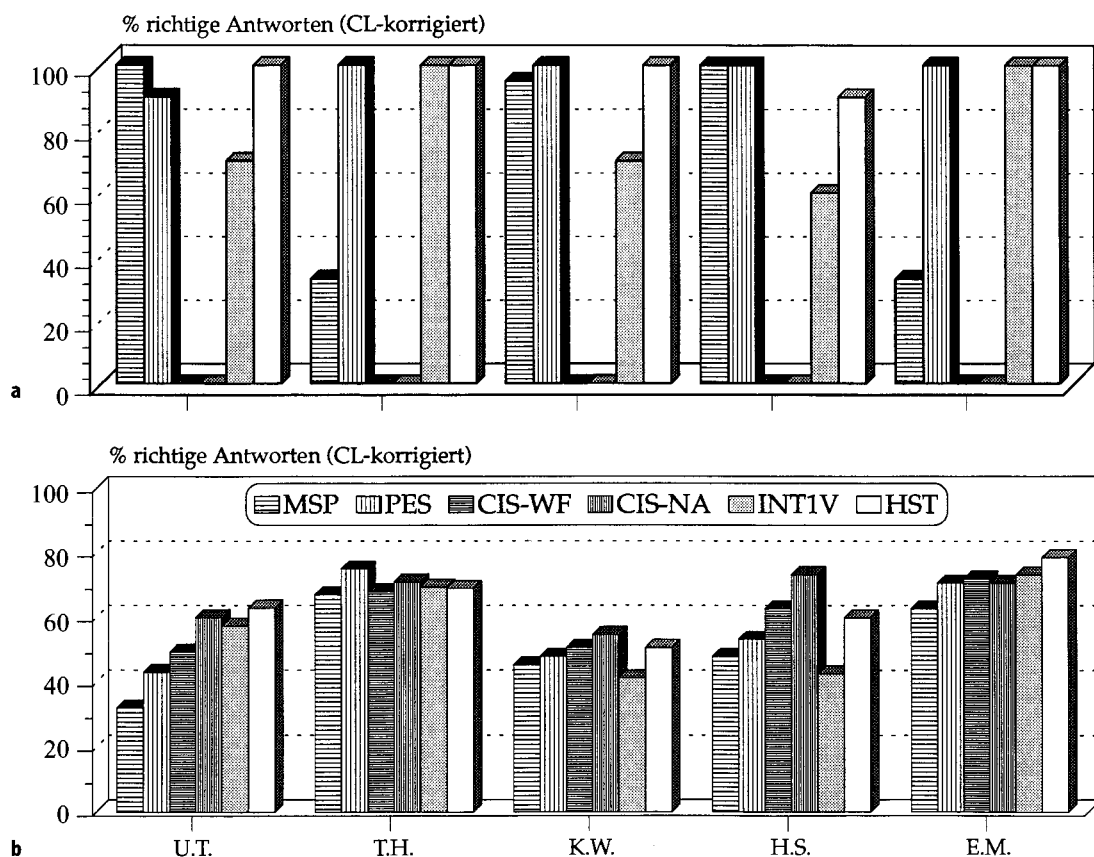


können. In einzelnen Sprachtests konnten mit den neuen Algorithmen deutlich bessere Ergebnisse erreicht werden als mit dem bisherigen Verfahren, in anderen Tests waren die Ergebnisse etwa gleich wie bisher. Als besonders erfolversprechend erwiesen sich Reizmuster mit möglichst hoher Stimulationsrate und einer maximalen Anzahl von Reizelektroden. Die Erkennung von Konsonant-Vokal-Übergängen sowie von Zisch- und Reibelauten konnte dadurch signifikant gesteigert werden. Die neuen Verfahren erwiesen sich auch unter Störlärmbedingungen als erfreulich robust. Die Variationen der Reizparameter ergaben eine relativ hohe Sensitivität der Lautheitsabbildung und zeigten mögliche Wege zur Optimierung künftiger Prozessoranpassungen auf.

Die CIS-Strategien, welche kontinuierliche Reizmuster ohne explizite Berücksichtigung der Grundfrequenz generieren, waren in diesen Experimenten der PES-Strategie, welche Pulsgruppen im Takt der Grundfrequenz erzeugt, meist überlegen. Hingegen wurde von vielen Patienten die unnatürlichere Sprachqualität mit CIS gegenüber PES bemängelt. Deshalb wurde nach Strategien gesucht, welche die Vorteile beider Verfahren kombinieren, also bei guter Sprachverständlichkeit auch qualitativ angenehme Sprache liefern.

Diese sog. Hybridstrategien (HST und INT1V) haben diese Anforderungen tatsächlich erfüllt, wie die Abb. 3.9 zeigt. Im ersten Experiment ging es um die Unterscheidung der Stimmlage von Frauen- und Männerstimmen. Mit den Fo-basierten Strategien (MSP, PES) war dies für die meisten Patienten problemlos zu erreichen, und die zufallskorrigierten Werte liegen meistens bei 100%. Mit der CIS-Strategie hingegen erreicht keine der fünf Versuchspersonen ein von der Ratewahrscheinlichkeit abweichendes Ergebnis. Mit den beiden Hybridstrategien wurden schließlich wieder sehr gute Diskriminationswerte erreicht, wobei die HST-Strategie der INT1V-Strategie etwas überlegen war. In den Konsonantentests konnten die guten Ergebnisse der CIS-Strategien mit den Hybridstrategien ebenfalls erreicht und teilweise sogar übertroffen werden. Die Informationsanalyse der Konsonantenverwechslungsmatrizen zeigt, daß es vor allem die Merkmale Sibilanz und Frikativität waren, welche mit den Strategien mit höheren Pulsraten besser übermittelt wurden. Erstaunlich ist auch die Verbesserung beim Merkmal

Abb. 3.9a, b. Experimente mit Hybridcodierungsstrategien. a Ergebnisse bei einer Frauen- vs. Männerstimme, b Ergebnisse des Konsonantentests CM2 für verschiedene Patienten



Stimmhaftigkeit, welche vermutlich auf die besser erkannten Plosivlaute (b-p, d-t, g-k) und die damit gekoppelte Erkennung des Artikulationsorts zurückzuführen ist.

Mit den neu entwickelten digitalen Strategien wurde spontan ein besseres Konsonantenverständnis erzielt als mit den von den Patienten bisher benutzten Sprachprozessoren. Die besten Ergebnisse wurden mit Verfahren erzielt, welche höhere Reizraten und differenziertere Abbildungen spektraler Feinstrukturen enthielten. Weitere Experimente sowie zusätzliche Verbesserungen sowohl der Anpassungsmethodik als auch der einzelnen Verarbeitungsschritte laufen in Richtung Hybridstrategien, wobei einzelnen Elektroden unterschiedliche Funktionen der Sprachcodierung zukommen. Damit soll insbesondere das Erkennen von Konsonant-Vokal-Übergängen noch weiter verbessert werden.

Seit kurzem ist für das Nucleus-Implantat ein neuer Sprachprozessor mit der Bezeichnung „Spectra-22“ erhältlich, dessen Signalverarbeitung auf einem Filterbankverfahren mit 20 über den Hauptsprachbereich verteilten Bandpaßfiltern und einer Auswahl der maximalen Energiewerte pro Zeitabschnitt beruht (McKay et al. 1991). Zur Evaluation dieses neuen Codierungsverfahrens wurde eine weltweite multi-

zentrische Studie mit Sprachtests und Fragebogen zur subjektiven Bewertung durchgeführt. Die Tests wurden über etwa 3 Monate in 4 Phasen mit dem früheren Standardprogramm MPEAK und dem neuen SPEAK-Programm nach festgelegtem randomisierten Schema durchgeführt. Ergebnisse mit englisch sprechenden Patienten (Skinner et al. 1994), sowie die Ergebnisse der 4 an der Studie beteiligten CI-Zentren im deutschsprachigen Raum wurden kürzlich publiziert (Dillier et al. 1995).

Wie Tabelle 3.2 für die in der Studie verwendeten Sprachtests zeigt, waren mit dem SPEAK-Verfahren deutliche Verbesserungen vor allem in Wort- und Satztests festzustellen. Zwischen den beiden Testwiederholungen mit MPEAK waren für die Gruppe von 20 Versuchspersonen statistisch keine signifikanten Unterschiede zu finden. Hingegen waren die Unterschiede mit dem SPEAK-Prozessor zwischen erster und zweiter Session statistisch signifikant, was als Lerneffekt interpretiert wurde und bedeuten kann, daß die Ergebnisse nach einem halben Jahr unter Umständen noch deutlicher ausgefallen wären.

Für den Innsbrucker Satztest im Störgeräusch von 10 dB SNR („signal to noise ratio“, Verhältnis von Sprache zu Störgeräusch) sind die Ergebnisse für alle Versuchspersonen in Abb. 3.10 zusammengestellt. Verbesserungen von bis zu 60 % (im Mittel 30 %) wurden von einzelnen Personen erreicht. Die subjektiven Bewertungen waren mehrheitlich sehr positiv.

Tabelle 3.2. Gruppenmittelwertvergleich für alle Sprachsubtests. Mittelwerte der Differenzen (SPEAK-MPEAK) in Prozent mit t-Werten, Anzahl Freiheitsgraden (df) und Signifikanzniveaus (n. s. nichtsignifikant); Anzahl Versuchspersonen mit verbesserter (+), unveränderter (0) und verschlechterter (-) Leistung

Subtests	SPEAK-MPEAK (gepaarte Differenzen)				(+)	(0)	(-)
	% Diff.	t-Test	df	2-seit. Sign.			
VO8 ^a	4,3	0,90	19	0,378 (n. s.)	8	6	6
C12 ^b	8,8	4,13	19	0,001 (***)	13	6	1
1-SYL ^c	8,4	3,15	19	0,005 (**)	14	2	4
In-Q ^d	10,2	5,11	19	0,000 (***)	12	8	0
In-15 ^e	34,1	5,20	8	0,001 (***)	7	1	0
In-10	32,5	8,21	19	0,000 (***)	17	3	0
In-5	34,3	5,49	10	0,000 (***)	10	2	0
Gö-Q ^f	10,0	2,95	19	0,008 (**)	12	4	4
Gö-15 ^g	18,7	3,09	13	0,009 (**)	10	2	2
Gö-10	17,5	4,72	19	0,000 (***)	13	6	1
Gö-5	28,2	3,20	5	0,024 (*)	5	0	1

^a Vokaltest, 8 Vokale in dV-Format.

^b Konsonantentest, 12 Konsonanten im aCA-Format.

^c Einsilberwortverständnistest (Freiburger Test).

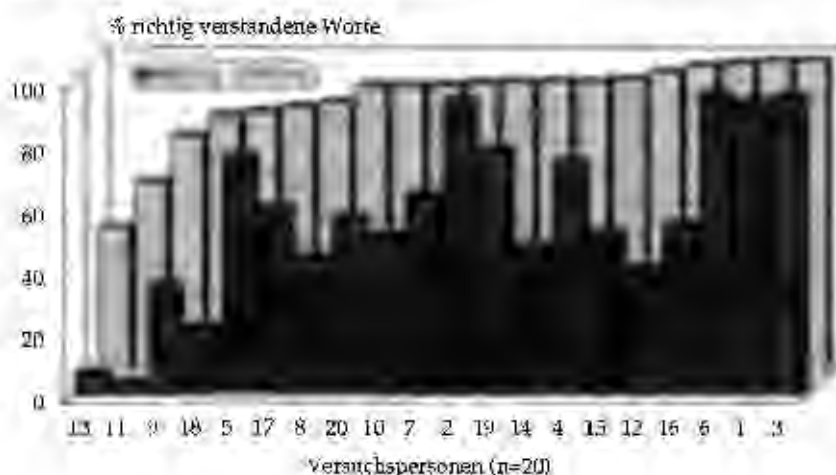
^d Innsbrucker Satztest in Ruhe.

^e Innsbrucker Satztest in Störlärm (15; 10; 5 dB Signal-Rausch-Verhältnis).

^f Göttinger Satztest in Ruhe.

^g Göttinger Satztest in Störlärm (15; 10; 5 dB Signal-Rausch-Verhältnis).

Abb. 3.10.
Vergleich MPEAK mit SPEAK
für 20 Versuchspersonen.
Innsbrucker Satztest im Stör-
geräusch (10 dB)



Die Experimente mit leistungsfähigeren digitalen Signalprozessoren haben gezeigt, daß die Sprachverständlichkeit bestehender Implantate mit verfeinerten Signalcodierungsstrategien und moderner Mikroelektronik noch verbessert werden kann. Die Umsetzung dieser Laborergebnisse in die Alltagssituation des CI-Trägers ist eine Aufgabe für die weitere Forschung und Entwicklung.

Künftige Sprachprozessoren sowie das Implantat selbst werden mit großer Wahrscheinlichkeit trotz gesteigerter Leistungsfähigkeit noch weiter miniaturisiert werden können.

Eine weitere Steigerung der Anzahl der Reizelektroden ist in nächster Zukunft nicht zu erwarten, da die elektroanatomischen Verhältnisse im Innenohr physikalische Grenzen setzen. Falls die in Entwicklung befindlichen enganliegenden Elektrodenträger die erhofften Verbesserungen in der Kanaltrennung ermöglichen, wäre allenfalls eine Verdoppelung oder Verdreifachung der maximalen Elektrodenzahl realisierbar. Zur Erzeugung von Aktivierungsmustern, die auch bezüglich stochasti-

scher Feinstruktur dem natürlichen neuronalen Aktivitätsmuster entsprechen würden, bedürfte es allerdings einer noch größeren Anzahl Elektroden, welche zudem in direktem Kontakt mit den erregbaren Neuronen (z. B. im Modiolus) stehen müßten. Um eine solche Prothese mit hunderten von Elektroden unterschiedlicher Sensitivität bei einem Patienten vernünftig anzupassen, wäre die Entwicklung völlig neuer Fittingprozeduren unter Einbezug elektro-physiologischer Methoden nötig.

Bereits beim heutigen Stand der Ergebnisse ist jedoch eine Erweiterung der Indikation auf resthörige Patienten mit minimalem Sprachverständnis absehbar. Die Vorteile der Cochlea-Implantate gegenüber akustischen Hörgeräten vor allem bei der Erkennung silbanter und frikativ-konsonanten sowie die oft größere maximale Lautheit bei Frequenzen oberhalb 1 kHz dürften die Cochlea-Implantate für eine immer größere Zahl von Hörbehinderten als sinnvolle Alternative zu einem Hochleistungshörgerät erscheinen lassen.

Literatur

- Alich G (1985) Anmerkungen zum Erlanger Sprachverständnistest (EST). Sprache Stimme Gehör 9:1–6
- Biesalski F, Lippert KL, Bohmert A (1985) Erfahrungen in der audiologisch-rehabilitativen Betreuung von gehörlosen Schulkindern mit einer Cochlea-Implanttherapie. Sprache Stimme Gehör 9:29–33
- Boothroyd L, Nittrouer S (1988) Mathematical treatment of context effects in phoneme and word recognition. J Acoust Soc Am 84:101–114
- Bögl F H (1993) Sprachverarbeitungsverfahren für ein melkanaliges Cochlear Implant. Dissertation, ETH Zürich
- Clark GM, Long JC, Patrick JB (1990) Cochlear Prostheses. Churchill Livingstone, Edinburgh London Melbourne New York
- Columbo J, Parkins CW (1987) A model of electrical excitation of the mammalian auditory-nerve neuron. Hear Res 31:287–312
- Deng L, Gaster CD (1987) A composite auditory model for processing speech sounds. J Acoust Soc Am 82/5:2901–2912
- Dillier N, Spillmann T (1992) Deutsche Version der Minimal Auditory Capability (MAC)-Test-Batterie. Anwendungen bei Hörgeräten und CI-Trägern mit und ohne Störärm. In: Kollmeier B (Hrsg) Moderne Verfahren der Sprachaudiometrie. Median-Verlag, Heidelberg, S 238–263

- Dillier N, Senn C, Schlatter T, Stöckli M, Utzinger U (1990) Wearable digital speech processor for cochlear implants using a TMS320C25. *Acta Otolaryngol Suppl* (Stockh) 469:120–127
- Dillier N, Battmer RD, Döring WH, Müller-Deile J (1995) Multi-centric field evaluation of a new spectral peak (SPEAK) speech coding strategy for cochlear implants. *Audiology* 34:145–149
- Dobie RA, Dillier N (1985) Some aspects of temporal coding for single-channel electrical stimulation of the cochlea. *Hear Res* 41–55
- Dowell RC, Martin LFA, Tong YC (1982) A 12-consonant confusion study on a multiple-channel cochlear implant patient. *J Speech Hear Res* 25:509–516
- Dudley H (1939) The Vocoder. *Bell Labs Rec* 18:122–126
- Eisenberg LS, House WF (1982) Initial experience with the cochlear implant in children. *Ann Otol Rhinol Laryngol Suppl* 91:67–73
- Feldmann H (1976) *Das Gutachten des Hals-Nasen-Ohren-Arztes*. Thieme, Stuttgart
- Finley CC, Wilson BS, White MW (1989) Models of the neural responsiveness to electrical stimulation. In: Miller MJ, Spelman FA (eds) *Cochlear implants. Models of the electrically stimulated ear*. Springer Verlag, Berlin Heidelberg New York Tokyo
- Flanagan JL (1972) *Speech analysis, synthesis and perception*. Springer Verlag, Berlin Heidelberg New York Tokyo
- Ghitza O (1987) Auditory nerve representation criteria for speech analysis/synthesis. *IEEE Trans ASSP* 35/6:736–740
- Hartmann R, Topp G, Klinke R (1984) Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea. *Hear Res* 13:47–62
- Hochmair-Desoyer JJ, Hochmair ES, Stiglbrunner HK (1985) Psychoacoustical temporal processing and speech understanding in cochlear implant patients. In: Schindler RA, Merzenich MM (eds) *Cochlear implants*. Raven Press, New York
- Honert C van den, Stypulkowski PH (1987a) Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hear Res* 29:207–222
- Honert C van den, Stypulkowski PH (1987b) Single fiber mapping of spatial excitation patterns in the electrically stimulated auditory nerve. *Hear Res* 29:195–206
- Jayant N (1993) Signal compression based on models of human perception. *Proc IEEE* 81:1385–1422
- Klatt DH (1980) Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am* 67:971–995
- Kollmeier B (1992) *Moderne Verfahren der Sprachaudiometrie*. *Audiol Akust Suppl*. 1–348
- McKay C, McDermott HJ, Vandali A, Clark GM (1991) Preliminary results with a six spectral maxima sound processor for the University of Melbourne/Nucleus multiple-electrode cochlear implant. *J Otolaryng Soc Austral* 6/5:354–359
- McMullen TA, Mountain DC (1985) Model of d.c. potential in the cochlea: Effects of voltage-dependent cilia stiffness. *Hear Res* 17:127–141
- Merzenich MM (1985) UCSF cochlear implant device. In: Merzenich MM, Schindler RA (eds) *Cochlear implants*. Raven Press, New York, pp 121–129
- Miller GA, Nicely PE (1955) An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 27:338–352
- Motz H, Rattay F (1986) A study of the application of the Hodgkin-Huxley and the Frankenhaeuser-Huxley model for electrostimulation of the acoustic nerve. *Neuroscience* 18/3:699–712
- Owens E, Schubert ED (1977) Development of the California Consonant Test. *J Speech Hear Res* 20:463–474
- Patrick JF (1995) Implant designs for future coding strategies. *Ann Otol Rhinol Laryngol Suppl* 104/166:137–138
- Pisoni DB, Sawusch JR (1975) Some stages of processing in speech perception. (Proceedings of the Symposium on Dynamic Aspects of Speech Perception, pp 16–35)
- Popelka GR, Gittelman DA (1984) Audiologic findings in a child with a single-channel cochlear implant. *J Speech Hear Disord* 49:254–261
- Schindler RA, Merzenich MM (1975) Chronic intracochlear electrode implantation: cochlear pathology and acoustic nerve survival. *Ann Otol Rhinol Laryngol* 202–215
- Shannon CD (1948) A mathematical theory of communication. *Bell Syst Tech Journal* 27:623–656
- Shepherd RK, Franz BK, Clark GM (1990) The biocompatibility and safety of cochlear prostheses. In: Clark GM et al. (eds) *Cochlear Prostheses*. Churchill Livingstone, Edinburgh London Melbourne New York, pp 69–98
- Skinner MW, Clark GM, Whitford LA et al. (1994) Evaluation of a new spectral peak coding strategy for the nucleus 22 channel cochlear implant system. *Am J Otol Suppl* 15:15–27
- Sotschek J (1982) Ein Reimtest für Verständlichkeitsmessungen mit deutscher Sprache als ein verbessertes Verfahren zur Bestimmung der Sprachübertragungsgüte. *Fernmelde Ing* 36 (4/5):1–84
- Spelman FA, Clopton BM, Pfungst BE (1982) Tissue impedance and current flow in the implanted ear – implications for the cochlear prosthesis. *Ann Otol Rhinol Laryngol* 92/Suppl 97:3–8
- Tyler RS, Gantz BJ, Lowder MW et al. (1985) Audiological results with two single channel cochlear implants. *Ann Otol Rhinol Laryngol* 94/2:133–139
- Wallenberg EL von, Kollmeier B (1989) Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten. *Audiol Akust* 28:50–65
- Wallenberg EL von, Brinch J, Money DK, West R, Avunduk K (1993) Comparative Reliability of Cochlear Implants. In: Fraysse B, Deguine O (eds) *Cochlear Implants: New Perspectives*. Karger, Basel (Ader ORL, vol 48, pp 79–84)
- Wang MD, Bilger RC (1973) Consonant confusions in noise: A study of perceptual features. *J Acoust Soc Am* 54:1248–1266
- Wedel H von (1985) Reichen die heute verfügbaren sprachaudiometrischen Verfahren zur Hörgeräte-Anpassung? Teil 2. *Audiol Akust* 24:102–120
- Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991) Better speech recognition with cochlear implants. *Nature* 352:236–238
- Zeng FG, Shannon RV (1992) Loudness balance between electric and acoustic stimulation. *Hear Res* 60:231–235