

ELiS_e

[e'li:zə]

<Essener Linguistische Skripte_elektronisch>

Ute Jekosch

Auditive Sprachqualitätsmessung:

Ein modernes Anwendungsgebiet von Sprach- und Kommunikationswissenschaften im Zusammenhang von Informations- und Kommunikationstechniken

Faszination

Sprache

elise@uni-essen.de

<http://www.elise.uni-essen.de>

Auditive Sprachqualitätsmessung: Ein modernes Anwendungsgebiet von Sprach- und Kommunikationswissenschaften im Zusammenhang von Informations- und Kommunikationstechniken

Ute Jekosch (Bochum)

„Sprich nicht so undeutlich und etwas lauter, wenn's geht!“ „Ich verstehe Dich nicht! Stell' bitte das Radio leiser!“ „Könnten Sie bitte etwas langsamer sprechen? Ich lerne zwar seit einiger Zeit deutsch, kann Ihnen aber nicht so recht folgen!“ „Was hast Du gesagt? Ich habe gerade nicht aufgepasst!“ „Da steht wohl jemand auf der Leitung! Deine Stimme klingt so, als seiest Du unendlich weit entfernt!“ ...

Gesprächssituationen, die uns allen wohl bekannt sind. Das Gesprochene wird vom Hörer aus unterschiedlichen Gründen nicht oder nur schwer verstanden, Stimmen klingen unnatürlich und verzerrt. Damit der Gedanken- und Informationsaustausch besser gelingt, wird der Sprecher vom Hörer über dessen Verstehensschwierigkeiten informiert. Normalerweise reagiert der Sprecher auf solche Rückmeldungen, indem er beispielsweise seine Sprechweise den Bedingungen der Kommunikationssituation anpasst oder Quellen, die störende Umgebungsgeräusche verursachen, ausschaltet. Durch seine Rückmeldungen, die auch gestischer oder mimischer Art sein können, strebt der Hörer eine Optimierung des zum Kommunikationserfolg notwendigen Hör- und Konzentrationsaufwandes an.

Wenn Sprech- und Höraufwand in einem ausgewogenen Verhältnis stehen, ist das Gesprochene vorwiegend reiner Informationsträger und somit pures Mittel zum Zweck. Der Hörer wird sich keineswegs in jedem einzelnen Fall der inhärenten Beschaffenheit und Wirksamkeit von lautsprachlichen Äußerungsformen bewusst, nicht zuletzt deshalb, weil der normal hörende Mensch im sozialen Umfeld täglich von sprechenden Menschen umgeben ist. Gesprochene Sprache ist in ihrer Materialität stark verwurzelt, sie gilt als etwas Selbstverständliches, etwas Normales. Was den Hörer primär interessiert, ist der Inhalt des Gesprochenen, das, was durch die lautsprachliche Äußerung bezeichnet wird, weniger die Materialität des Bezeichnenden, die Form des lautsprachlichen Zeichens selbst.

Das ändert sich jedoch, wenn die lautsprachliche Äußerung ihren Zweck nicht mehr erfüllt, wenn das Gesprochene nicht eindeutig verstehbar bzw. verständlich ist. Die Aufmerksamkeit des Hörers richtet sich dann auch auf die Form des Gesprochenen, d. h. der Informationsträger als solcher rückt plötzlich in den Vordergrund. Wortwahl, also Wahl des Informationsträgers, lautsprachliche Realisierung, situative Angemessenheit des Gesprochenen, Kontamination des Sprachschallsignals durch Störschalle, beeinträchtigt Hörvermögen oder mangelnde Aufmerksamkeit können herausragende Gründe für Verstehens- und Verständnisprobleme sein. Solche Schwachstellen im gesprochen-sprachlichen Kommunikationsablauf kann der Hörer in der

Regel ohne Probleme identifizieren. Das bedeutet mit anderen Worten, dass der Hörer eine – zumindest grobe – Vorstellung davon hat, welche lautsprachlichen Äußerungsformen tolerabel bzw. akzeptabel sind und wann er den Sprecher besser bittet, seine Sprechweise den situativen Gegebenheiten anzupassen.

Es ist für alle am Kommunikationsprozess Beteiligten selbstverständlich wünschenswert, dass die sprachliche Verständigung direkt und ohne korrigierende Rückmeldung gelingt. Sprecher und Hörer haben gewöhnlich ein sehr feines Gespür dafür, wann und wie sie sich den jeweiligen Kommunikationsbedingungen bestmöglich anpassen.

Anpassung ist jedoch ein aktiver Prozess. Er setzt voraus, dass Sprecher und Hörer kooperativ sind, dass sie willens sind, sich mit der Kommunikationssituation aktiv auseinander zu setzen, um aktuelle Voraussetzungen und Bedingungen im sprachlichen Verständigungsbemühen zu berücksichtigen. Das allein reicht jedoch keineswegs aus, denn eine Anpassung wird nur dann optimal gelingen, wenn die Gesprächspartner auch alle Störfaktoren kennen, die die sprachliche Verständigung beeinflussen. Befinden sich z. B. Sprecher und Hörer in einem Raum, in dem die sprachliche Verständigung durch Baulärm beeinträchtigt ist, so erfahren beide Kommunikationspartner den Störfaktor „Baulärm“ in derselben Störschallumgebung in nahezu gleicher Weise. Beide können sich auf die Umgebungsbedingungen einstellen.

Anders verhält es sich, wenn Sprecher und Hörer räumlich voneinander getrennt sind und elektronische Kommunikationsmittler als Übertragungsmedium verwenden:

Wird z. B. ein Autofahrer über ein Mobilfunktelefon von seinem Gesprächspartner erreicht, so können Sprachqualitätseinbußen durch Feldstärkechwankungen auftreten. Feldstärkechwankungen ergeben sich u. a. aus wechselnden Übertragungsbedingungen in der landschaftlich-baulichen Umgebung, durch die sich der Mobilfunkteilnehmer bewegt. Der Anrufer kann, während er spricht, nicht voraussehen, wann Übertragungsstörungen auftreten, und er weiß auch nicht, wie sich diese im übertragenen Sprachsignal äußern. Folglich kann er sich den spezifischen Merkmalen von Funkübertragungsstörungen nicht oder zumindest nur bedingt anpassen. Es kommt erschwerend hinzu, dass selbst eine Bitte des Hörers, deutlicher oder lauter zu sprechen, im nächsten Moment bereits hinfällig sein kann, wenn sich die Feldstärkebedingungen in der Zwischenzeit kommunikationsvorteilhaft geändert haben.

Es gibt folglich Störfaktoren in der lautsprachlichen Kommunikation, die außerhalb der Kontrolle der am Gespräch Beteiligten liegen. Sie treten u. a. dann auf, wenn sich Sprecher und Hörer elektronischer Informations- und Kommunikationstechniken bedienen. So wird z. B. durch das Funktelefon die Möglichkeit geschaffen, zu jeder Zeit an jedem Ort zu telefonieren, ohne an einen bestimmten Standort gebunden zu sein. Im Vergleich zum stationären Telefon bedeutet das zwar eine Steigerung der Systemleistung in Bezug auf die Benutzerfreundlichkeit, die reaktive Anpassungsmöglichkeit an die spezifischen Kommunikationsbedingungen ist jedoch begrenzt (zeitvariantes System).

Gegenwärtig sind vielfältige Kommunikationstechniken verfügbar, durch die der sprachlich-kommunikative Handlungsbereich faktisch erweitert wird. Obwohl Wissenschaft und Technik bereits vielfältige Lösungen für moderne Informations- und Kommunikationsaufgaben bereitstellen, ist die Diversifikation der Angebotsvielfalt noch immer nicht erschöpft. Im Bereich der modernen Telekommunikationstechnik und Telematik ist es nicht mehr allein der ungestillte Bedarf des Kunden, der lautstark zu Innovationen aufruft und die Technologieentwicklung antreibt. Das bestehende Warenangebot ist bereits äußerst vielseitig und umfangreich. Das Ziel des

nach Erfolg strebenden Kommunikationstechnologien ist es nicht mehr allein, möglichst schnell und flexibel auf gesellschaftliche Ansprüche zu reagieren, sondern er ist bestrebt, seiner Zeit immer voraus zu sein, daran mitzuwirken, dass gesellschaftlicher Bedarf entsteht, sich entwickelt und festigt. Er ist bemüht, immer wieder Anlässe für weitere Innovationen, für gezielte Systemoptimierungen und neue Anwendungen zu schaffen.

Die Suche nach neuen Märkten hat dazu geführt, dass moderne Informations- und Kommunikationstechnologien eine neue Richtung einschlagen. Moderne Forschungs- und Entwicklungsziele dieser Technologien sind nicht weiter vordringlich auf den Bereich des Realen ausgerichtet, in dem sie dazu dienen, den Informationsaustausch zu beschleunigen und zu vereinfachen, um ungehindert Nachrichten für den Menschen direkt abrufbar und praktisch nutzbar zu machen. Die Technologie von Gegenwart und Zukunft konzentriert sich mehr und mehr auf den Bereich des Virtuellen, wird deshalb u. a. auch als Phänoteknik bezeichnet. Ziel der Phänoteknik ist es, die wahrgenommene Welt zu simulieren und neue Welten zu erzeugen, zu speichern und unabhängig von Raum und Zeit zu bestehen.

Entsprechend wird die Vision der Möglichkeit, sich in künstliche Welten hineinversetzen zu lassen und in ihnen zu kommunizieren, heute bei weitem nicht mehr als reine Phantasterei abgetan. Es gilt z. B. durchaus als machbar, dass der multi-sensorische, kognitive Apparat des Menschen durch Simulationstechniken so gelenkt wird, dass der Wahrnehmende eine andere Welt erlebt. Die Vorstellung, sich an einem virtuellen Ort zu einer Besprechung zu treffen, ohne sein heimisches Büro verlassen zu müssen, trotzdem aber auditiv und visuell kommunizieren und gleichzeitig die virtuelle Umgebung taktil und olfaktorisch wirklichkeitsnah erfahren zu können, erscheint keineswegs mehr unrealistisch. War es bislang das vorwiegende Ziel von Informations- und Kommunikationstechnologen, Tatsachen zu erforschen, das Gegebene zu erklären, darzustellen und so zu verändern, dass ein Erkenntnisfortschritt mit einem möglichst direkten gesellschaftlichen Nutzen und Zugewinn entsteht, so wird nun angestrebt, die Basis für neue Tatsachen zu schaffen, damit zeitliche und räumliche Strukturen überwunden werden können.

Das setzt selbstverständlich voraus, dass die Verlässlichkeitsbedingungen bekannt sind, aufgrund derer der Wahrnehmende die ihm vertraute materielle Basis tatsächlich vergessen und für nachfolgende Bewusstseinstätigkeiten zu Bezugspunkten umlenken kann, auf die sich seine Wahrnehmung dann stützt. Die Strukturierung und Erforschung dieser Fragen ist programmatisch für zukünftige Aktivitäten, durch die die Informationstechnik Schritt für Schritt ihr traditionelles Terrain verlässt und mehr und mehr die Herausforderungen der Gestaltung der intellektuellen Realitätsvorstellung annimmt. Es wird angestrebt, Inhalte natürlicher Wahrnehmung und Intelligenz durch Phänotekniken und durch die elektronische Informationsverarbeitung zu erweitern.

Dabei ist es u. a. ein vordringliches Forschungsziel, Teilaspekte der Fähigkeit des Denkens, Erkennens, Sprechens und Kommunizierens von der menschlichen Intelligenz zu entkoppeln, so dass umfangreiche Kommunikationsleistungen durch Kommunikationstechnologien erbracht werden können. Die sprachproduktive und -rezeptive Leistung ist eine dieser Komponenten.

Wunsch und Wirklichkeit klaffen jedoch (wie zu erwarten) sehr weit auseinander. So sind zwar manche z. Zt. verfügbaren Informations- und Kommunikationstechniken dem Menschen in der Erfassung, Speicherung und Verarbeitung von Informationen infolge elektronisch gesteuerter Datenverwaltung und elektronisch beschleunigtem Datenflusses überlegen. Verglichen mit

den umfassenden Fähigkeiten des kommunizierenden Menschen hinken Kommunikationstechnologien den menschlichen Leistungen aber weit hinterher.

Typisches Beispiel hierfür ist der Stand von Forschung und Technik im Bestreben, eine sprechende und eine hörende Maschine zu entwickeln. Zur Erreichung dieser Ziele werden recht einschneidende Einschränkungen bezüglich der gestellten Aufgabe vorgenommen: So wird Sprechen weitgehend auf den Vorgang des Text-Vorlesens eingeschränkt, wobei der Prozess des Denkens, der Konzeptualisierung und Umsetzung in Sprache zumeist ausgeklammert wird. Hören wird dann auf das Sprache-Hören unter genau definierten Randbedingungen reduziert. Lautsprachliche Äußerungen, deren Strukturen grammatisch regelwidrig oder die genuschelt sind, sind häufig für die Maschinenbenutzung nicht gerecht, und maschinelle Spracherkennungsfehler werden nicht selten mit einem unangemessenen Sprecherverhalten entschuldigt. Ähnliches gilt für Spracherkennungsfehler, die auf die Störgeräuschsituation zurückgeführt werden.

Das bedeutet, dass einem maschinellen Sprecher und einem maschinellen Hörer (also einem Sprachsynthese- und einem Spracherkennungssystem) nicht in Ansätzen eine natürliche Kommunikationsfähigkeit zugeschrieben werden kann. ‚Kommunikation‘ zwischen Maschine und Mensch (Sprachsynthese) oder Mensch und Maschine (Spracherkennung) hat nicht im Entferntesten etwas mit dem Ideen- und Gedankenaustausch (wohl aber Informationsaustausch!) in einer natürlichsprachigen Kommunikationssituation zu tun. Strenggenommen ist sie nichts weiter als eine normierte Datenübertragung in einem festgelegten, durch Regeln gesteuerten System. Gegenwärtige Sprachsynthese- und Spracherkennungssysteme sind kaum mehr als Umsetzer, die eine in einem Standardformat vorliegende Information am Eingang aufnehmen, nach festgelegten Regeln verarbeiten, umcodieren und am Ausgang wieder abgeben. Sobald jedoch schon die Eingangsinformation nicht dem Format entspricht, wie es das Sprachverarbeitungssystem verlangt, ist noch nicht einmal eine rudimentäre Kommunikation in diesem eingeschränkten Sinn möglich.

Auch wenn man lediglich die spezifischen Teilleistungen von Sprachverarbeitungssystemen genauer analysiert, fällt auf, dass die erreichte Systemleistung größtenteils noch recht weit von der entsprechenden Teilleistung des Menschen entfernt ist. Dies wird direkt erfahrbar, wenn man einer künstlichen Stimme zuhört, die einen geschriebenen Text vorliest. Als Hörer weiß man nach relativ kurzer Zeit, dass man es mit einer Maschine zu tun hat. Man merkt, dass die Maschine den Inhalt des Gesprochenen nicht versteht und hat teilweise den Eindruck, dass sie sich ähnlich verhält wie ein 6-jähriges Kind, das lesen lernt. Ähnliches gilt für Spracherkennungssysteme: So wie der Diktatstoff vom Lehrer gezielt auf den Schreibenlernenden abgestimmt wird und sich der Diktierende lautsprachlich auf die Kompetenz des Lerners einstellt, so sind Wortwahl, Syntax und Sprechweise auf die Leistungsfähigkeit der Maschine ausgerichtet. Je besser sich nun der menschliche Hörer der synthetischen Stimme bzw. der menschliche Sprecher dem maschinellen Spracherkennungssystem anpasst, desto höhere Systemleistungen können erzielt werden.

Nicht immer sind aber Sprecher und Hörer dazu bereit, sich auf die Erfordernisse von Maschinen einzustellen und sich ihnen anzupassen. Ein Grund hierfür ist sicherlich, dass sich viele Menschen von Informationstechniken überladen fühlen, was zur Folge hat, dass der Alltag unüberschaubar, intransparent und sinnentleert empfunden wird. Es gibt durchaus Kommunikationstechniken, die zwar einzelne, qualitativ hochwertige Teilleistungen erbringen, die aber in ihrer Gesamtheit noch unfertig wirken und den Alltag des Benutzers nicht entscheidend ent-

lasten. Solche Systeme werden – sofern sie nicht einen systemexternen Qualitätsschub wie etwa gesellschaftliches Prestige erhalten – sehr schnell zurückgewiesen und stoßen nur schwerlich auf Akzeptanz. Hinzu kommt, dass die Fähigkeiten des menschlichen Sprechers bzw. Hörers das Vergleichsnorm bilden. Es ist ein sehr hoch gestecktes und anspruchsvolles Ziel, die Flexibilität eines natürlichen Kommunikationspartners nachzubilden.

Entwicklung, Planung, Einführung und Wartung von Kommunikationstechnologien ist deshalb untrennbar mit Bedürfnissen, Erwartungshaltungen und Forderungen der anvisierten Benutzergruppe verbunden. Dies gilt in ganz besonderem Maße für die Entwicklung der Phänotechnologien, denn die Frage, ob eine Technologie auf Akzeptanz oder auf Ablehnung stößt, entscheidet letztendlich allein der an ihrer Anwendung Interessierte. Dies kann am Beispiel des Mobilfunktelefons verdeutlicht werden: Das Mobilfunktelefon ermöglicht dem Fernsprechteilnehmer räumliche Unabhängigkeit. Der Gewinn an räumlicher Unabhängigkeit geht jedoch u. a. auf Kosten der Sprachübertragungsqualität. Sprachübertragung ist eine Teilleistungsaufgabe des Systems, für die in anderen Technologien (z. B. beim stationären Telefon) bereits ein hohes Qualitätsniveau erreicht wurde. Dieses Qualitätsniveau ist zunächst das Vergleichsnorm für die vom Mobilfunktelefon erwartete Sprachübertragungsqualität. Die Sprachübertragungsqualität im Mobilfunktelefonverkehr ist aber meist schlechter als beim Festnetz erwartet wird. Damit der primäre Eindruck beim Gebrauch der neuen Technologie nicht zu negativ ausfällt, werden schon vor der eigentlichen Technikeinführung Erwartungshaltungen beim potentiellen Kunden gelenkt und korrigiert. Noch bestehende Systemmängel werden durch gezielte Marketingstrategien und Werbeaktionen, die den innovativen Charakter der Technologie und den mit ihr verbundenen gesellschaftlichen bzw. individuellen Zugewinn herausstellen, verdeckt. Von den Kunden wird erwartet, dass sie den jeweiligen technologischen Zugewinn höher einstufen als die mit ihm einhergehenden Qualitätseinbußen. Man hofft, dass sich der Kunde der noch nicht voll ausgereiften Technologie anpasst.

Heute wird das Mobilfunktelefon vielfach verkauft. Obwohl zum Zeitpunkt der Technikeinführung ein funktionsoptimaler Endzustand von Teilleistungsaufgaben und eine Integration dieser Systemteilleistungen in ein Endprodukt mit einer bestmöglichen Gesamtfunktionalität noch nicht eingetreten war, das vorstellbare Qualitätsoptimum des Produkts also noch nicht erreicht war, ist die Einführung des Mobilfunktelefons weitgehend erfolgreich verlaufen. Der Zugewinn an Mobilität wurde von vielen Benutzern offenbar höher eingestuft als eine gute Sprachübertragungsqualität.

Nicht immer füllen innovative sprachtechnologische Systeme so deutliche Marktlücken aus wie die, in die das Mobilfunktelefon vorgestoßen ist. Es gibt durchaus auch verdeckte Marktnischen, bei denen der Funktionsgewinn, der durch den Einsatz einer neuen Technologie erzielt werden soll, noch nicht für alle angesprochenen Benutzergruppen klar ersichtlich ist. Hat sich das Interesse an einer neuen Technologie noch nicht herausgebildet, ist Technikeinführung nur schwer durchzuführen. Beispiel hierfür ist die Sprachsynthese: Obwohl gegenwärtig bereits funktionstüchtige Sprachsynthesysteme verfügbar sind, hat sich ihre eigentliche Rolle als Werkzeug im Kommunikationswesen noch immer nicht klar herausgeschält. Abgesehen vom Einsatz der Sprachsynthese als Vorleseautomat für Sehbehinderte fehlt die Einbindung in zentrale Aufgabenabläufe, die Hilfe und Erleichterung, d. h. persönlichen oder gesellschaftlichen Zugewinn bedeutet. Sprachsynthesysteme stoßen in der heutigen Gesellschaft eher auf Skepsis als auf Zustimmung.

Um eine rigorose, unreflektierte Technikablehnung zu vermeiden, ist man darum bemüht, die Technologie Schritt für Schritt in neue Funktions- und Aufgabenzusammenhänge einzuführen, also die Möglichkeit zur Bedarfsentwicklung zu schaffen und zu nutzen. Da sich Technologie nicht verordnen lässt, muss sichergestellt sein, dass es einen geeigneten Bezugspunkt in der Benutzerwelt gibt, an dem die neuartige Technologie andocken kann. Wird dafür keine Sorge getragen, wird die Bereitschaft beim Kunden, sich einer neuartigen Technologie zu öffnen und diese zu akzeptieren, nicht gezielt geweckt werden können. Die Frage, ob Technologieeinführung erfolgreich oder erfolglos verläuft, bleibt dann eher dem Zufall überlassen. Unter Berücksichtigung der internationalen Wettbewerbssituation und der ohnehin schon vorhandenen Marktenge kann sich dies aber niemand leisten. Wirtschaftlichkeit und Risikobegrenzung gehören zu den Haupt Gesichtspunkten in der Technologiefolgenabschätzung.

Folglich ist es das Ziel aller an der Sprachtechnologieforchung, -entwicklung, -einführung und -begleitung Beteiligten, unter den gegebenen Voraussetzungen das höchst erreichbare Ziel zu erkennen, um dann unter Berücksichtigung aller damit einhergehenden Notwendigkeiten geeignete Produkte zu entwickeln. *Qualitätsgestaltung* und *Qualitätssicherung* sind dabei die erfolgweisenden Strategien, im Zusammenhang moderner Sprachkommunikationstechnologien ist *Sprachqualität* ein Aspekt der Gesamtqualität von Technologien.

Was aber ist Qualität? Was ist Qualität im Rahmen der Sprachtechnologie allgemein? Welche Qualitätsanforderungen werden an Informations- und Kommunikationstechnologien gestellt? Wie lässt sich Qualität gestalten? Reicht es bereits aus, wenn Produkte einwandfrei funktionieren und keine Fehler aufweisen? Was ist Qualität gesprochener Sprache im Besonderen? Was ist Qualität künstlich erzeugter Lautsprache? Kann man Sprachqualität auditiv messen? Wie lässt sich die Qualität natürlich gesprochener oder künstlich erzeugter Sprache bestimmen? Welche Qualitätsanforderung bildet das Vergleichsnormale? Welche Aspekte der Qualität der Lautsprache sind identifizierbar? Nach welchen Kriterien wird Lautsprachqualität im gegebenen Fall beurteilt? Wird sie immer gleich beurteilt? Von welchen Einflussgrößen hängt das Urteil ab? Ist die Qualität von Lautsprache instrumentell messbar? Gibt es Vorgaben, nach denen die Qualität von sprachtechnologischen Systemen beurteilt werden soll?

Diese Fragen umreißen den Aufgabenbereich der auditiven Sprachqualitätsmessung, ein Forschungsgebiet, das sich in jüngster Zeit zu einem Schwerpunktthema im Zusammenhang von Informations- und Kommunikationstechniken entwickelt hat. Das Ziel der auditiven Sprachqualitätsmessung ist es, die wahrgenommene Beschaffenheit von lautsprachlichen Äußerungen zu ermitteln. In die Betrachtung werden alle wesentlichen Aspekte eingeschlossen, die mit dem Messvorgang verknüpft sind. Dabei bilden Sprach- und Kommunikationswissenschaft einen Hauptschwerpunkt. Für ein globales Verständnis des Themengebietes ist es jedoch unumgänglich, auch Fragestellungen aus anderen Wissensbereichen anzusprechen. Zu diesen Wissensbereichen zählen insbesondere die Psychoakustik und die Metrologie. Diese finden immer dort Berücksichtigung, wo sprachwissenschaftliche Betrachtungen sie entweder erforderlich machen oder wo sie deren Ergebnisse beeinflussen bzw. auf diese zurückwirken. Die Sprachwissenschaft wird hier funktionalisiert. Es werden keine Themen diskutiert, die sich aus einem isolierten Forschungs- und Erkenntnisinteresse der Sprachwissenschaft selbst ergeben, sondern es werden Fragen behandelt, die sich an die Sprachwissenschaft in der interdisziplinären Zusammenarbeit richten. Hier stellt sich die Sprachwissenschaft gerade auch neuen Herausforderungen assoziierter Wissens- und Aufgabenbereiche, bringt dort Wissen ein und ist zugleich offen für aktuelle

Problemstellungen, für die sie heute noch keine befriedigenden Lösungen bereitstellt.

Im Zuge der Behandlung dieses Forschungsbereichs werden im interdisziplinären Sinn momentan vorwiegend Grundsatzfragen der Sprachwissenschaft im Umfeld der Psychoakustik behandelt. Die angewandte Sprachwissenschaft bildet den wissenschaftlichen Kern, die Psychoakustik den wissenschaftlichen Rahmen. Aus dem Bereich der Psychoakustik werden hauptsächlich Aspekte der sprachbezogenen Psychoakustik und der auditiven Sprachqualitätsbeurteilung herangezogen. Die Beziehung zwischen Psychoakustik, sprachbezogener Psychoakustik und auditiver Sprachqualitätsbeurteilung kann dabei folgendermaßen beschrieben werden:

- Die *Psychoakustik* untersucht physikalische Ereignisse und die mit ihnen korrelierten Wahrnehmungen.
- Gegenstand der *sprachbezogenen Psychoakustik* sind speziell Sprechschalle, die sich als Angebot an ein auditives System richten, sowie der jeweils individuelle Hörer, der Sprechschalle selektiert und kognitiv verarbeitet.
- Die *auditive Sprachqualitätsbeurteilung* erforscht hauptsächlich die Vorgänge, die mit der Beurteilung gehörter Sprache zusammenhängen, insbesondere alle wesentlichen Aspekte, die für die Ermittlung der Sprachqualität von Bedeutung sind. Ergebnisse auditiver Beurteilung werden u. a. für die Entwicklung von instrumentellen Verfahren zur Berechnung der Sprachqualität benutzt.

Beurteilung von Lautsprache setzt i. a. auditive Wahrnehmung voraus. Dementsprechend sind die beiden Forschungsbereiche „Sprachbezogene Psychoakustik“ und „Auditive Sprachqualitätsbeurteilung“ sehr eng miteinander verzahnt. Beide Bereiche haben zum Ziel, Voraussetzungen, Bedingungen und allgemeine Abläufe von sprachbezogenen auditiven Wahrnehmungsleistungen zu erfassen und zu verstehen. Sie unterscheiden sich jedoch hinsichtlich des gewählten Schwerpunkts: Die sprachbezogene Psychoakustik bezieht sich hauptsächlich auf die auditive Sprachwahrnehmung, die Sprachqualitätsbeurteilung auf den Vorgang der Reflexion des auditiv Wahrgenommenen im Bezug auf das Erwartete.

Der explizite Vorgang der Beurteilung gesprochener Sprache findet in natürlichen Kommunikationssituationen meist in ungelenkter Weise statt. In natürlichen Gesprächssituationen wird die Qualität des Gehörten immer dann zum Thema, wenn die lautsprachliche Äußerung durch bestimmte Besonderheiten auffällig wird. Die Beurteilung der Sprachqualität kann jedoch auch durch die Inszenierung künstlicher Situationen (z. B. durch Labortests mit Sprache) gezielt provoziert werden und somit gelenkt ablaufen.

Das Wissenschaftsgebiet „Auditive Sprachqualitätsbeurteilung“ ist auf die Vorgänge beim Bestimmen der Sprachqualität sowohl in natürlichen als auch in künstlichen Situationen ausgerichtet. Vordringliches Ziel ist es, Prozesse der Sprachwahrnehmung und -beurteilung zu erfassen, um vor dem Hintergrund von Sprachwissenschaft, Wahrnehmungstheorie, Psychoakustik und Metrologie Wissen zur Bestimmung und Prognose der wahrgenommenen Sprachqualität bereitzustellen.

Weiterführende Literatur

- Blauert, Jens, ed. (2004), *Communication Acoustics*. Berlin : Springer, in prep.
- Jekosch, Ute (2000), *Sprache hören und beurteilen. Ein Ansatz zur Grundlegung der Sprachqualitätsbeurteilung*. Habilitationsschrift Universität/GH Essen, FB 3, Prof. Bunting
- Jekosch, Ute, Möller, Sebastian, eds. (2004), *Auditory Quality of Systems*. Special Issue of *acta acustica*, in print
- Möller, Sebastian (2000), *Assessment and Prediction of Speech Quality in Telecommunications*. Boston: Kluwer
- Möller, Sebastian (2004), *Quality of Telephone-Based Spoken Dialogue Systems*. Habilitationsschrift Ruhr-Universität Bochum, Institut für Kommunikationsakustik, Prof. Blauert
- Wilson, Stephen (2002), *Information Arts. Intersections of Art, Science, and Technology*. London: The MIT Press