

# Der Einfluss von Bedeutung auf die Wahrnehmung des Klangs von Sprache

Alexander Raake, Institut für Kommunikationsakustik, Ruhr-Universität Bochum, [raake@ika.ruhr-uni-bochum.de](mailto:raake@ika.ruhr-uni-bochum.de)

## Einleitung

Sprache lässt sich mittels moderner Telekommunikationsnetze in Zukunft breitbandig (50-7000 Hz) übertragen. Daneben werden bereits heute zunehmend Signalverarbeitungs-komponenten wie Low-Bitrate-Kodierer oder vom Handapparat abweichende Endgeräte eingesetzt. Im Vergleich zur klassischen Schmalbandtelefonie ergeben sich daraus für den Benutzer deutlich veränderte Klangeigenschaften. Um moderne Telekommunikationssysteme im Hinblick auf hohe Sprachqualität planen, aufbauen und überwachen zu können, muss die Qualitätsdimension 'Klang' in instrumentellen Modellen zur Qualitätsprädiktion, -schätzung und -überwachung zukünftig berücksichtigt werden (vgl. z.B. [7]).

Zur Erfassung des Zusammenhangs zwischen Sprachqualität und Klangeigenschaften sind Untersuchungen notwendig, die meist in Form reiner Hörtests durchgeführt werden. Abhängig vom Versuchslabor kommen dabei unterschiedliche Sprachmaterialien zum Einsatz. Als ausschlaggebend für die Bewertung der Klangeigenschaften natürlicher Sprache, die über einen Kanal mit bestimmter Frequenzcharakteristik übertragen wird, werden a priori nur die Eigenschaften des Telefonkanals, des Sprechers bzw. des Hörers sowie die phonologischen Eigenschaften des Textmaterials angenommen.

Die absolute Klangqualität eines Übertragungssystems existiert aber ebenso wenig wie die absolute Gesamtqualität. Stattdessen wird sie dem Gehörten durch den Hörer in Abhängigkeit des Funktionszusammenhangs zugewiesen (vgl. [5]). Bei der Entwicklung einer geeigneten Methode zur auditiven Klangqualitätsmessung unterschiedlicher Systeme stellt sich demnach beispielsweise die Frage, in wie weit auch die *Bedeutung* bzw. der Sinn des Sprachmaterials Einfluss auf die Beurteilung nimmt.

Der vorliegende Beitrag soll diese Frage anhand der Ergebnisse zweier auditiver Untersuchungen vor einem semiotischen Hintergrund erörtern. Die Hörversuche wurden exemplarisch für drei verschiedene prototypische Übertragungsbandbreiten durchgeführt, um unterschiedliche Klangeigenschaften des Sprachmaterials zu generieren. Es wurden vier Arten von Textmaterial verwendet: Auszüge aus zwei Quellen klassischer Literatur, Alltagssprache und sinnleere Sätze. Es zeigt sich eine deutliche Abhängigkeit der Urteile vom Sinngehalt des ausgewählten Textmaterials.

## 1. Klangqualität: Auditive Tests

Der Begriff 'Klang' beschreibt – im umgangssprachlichen Gebrauch – vermutlich am ehesten das Ergebnis der Qualitätsbeurteilung der wahrgenommenen, mehrdimensionalen Klangfarbe auf einer eindimensionalen Skala. Die hier untersuchte Eigenschaft von Sprache – die Klangqualität – ist bisher nicht eindeutig definiert. Der aus der Produktgeräuschqualitätsforschung stammende englische Begriff "Sound Quality" kann im Deutschen treffend mit *Geräuschqualität* übersetzt werden (vgl. Bednarzyk [2]). Bednarzyk unterscheidet dabei zwei Beschreibungsmerkmale der Geräuschqualität: Die *Klangqualität* zur Identifikation der Quelle und spezifischer ästhetischer Merkmale sowie die *Geräuschwirkungsqualität*, welche die Funktion des Geräuschlauts betrifft, oder vielmehr "...die emotionale Wirkung des Produktgeräusches auf den Menschen sowie die Auswirkung des Produktgeräusches auf das Verhalten des Menschen".

Für die Sprachqualitätsforschung ist noch zu klären, in wie weit die Bewertung der Klangfarbe des Sprachsignals auf einer eindimensionalen Beurteilungsskala als *Klangqualität* bezeichnet werden kann. Dieser Begriff soll aber dennoch im Folgenden zur Beschreibung des Untersuchungsgegenstandes verwendet werden: Die Beurteilung der Qualität der Stimme des Gesprächspartners, (möglichst) nicht im Hinblick auf individuelle Eigenschaften des Sprechers, sondern auf (z.B. spektrale) Eigenschaften der Telefonverbindung.

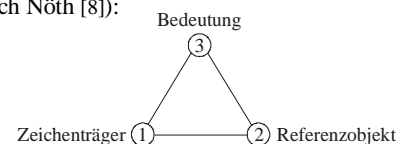
Ungeklärt bleibt zunächst die Eindeutigkeit dieses Begriffs für naive Versuchspersonen, die Klangqualität als ausschließliches Kriterium zur Bewertung einer Sprachprobe beurteilen sollen. Dabei sollen, überträgt man die Definition der Klangqualität von Bednarzyk, möglichst nur jene Merkmale eines Sprachübertragungssystems beurteilt

werden, die a priori keine direkte Funktion der Sprache beeinflussen, also nur die, die Informationen über den Sprecher modifizieren und insbesondere ästhetische Merkmale von Sprache betreffen. Dies ist bei Sprache offenbar eine reduzierte Betrachtungsweise, da sich Sprache von Geräuschen insofern unterscheidet, als sie gerade zur Übertragung von Information erzeugt wird (als Geräuschschall wird laut DIN 1320 der "...Schall bezeichnet, welcher nicht vorwiegend zur Übertragung von Information erzeugt wurde"). Auf diesen Sachverhalt wird im nachfolgenden Abschnitt zur Semiotik noch näher eingegangen. Dabei geht es um die Beantwortung der Frage, welche Rolle der Inhalt oder die *Bedeutung* des verwendeten Text- bzw. Sprachmaterials – also letztlich die Funktion des Gesagten – hinsichtlich der Klangqualität spielt.

Zur Bestimmung der Qualität bzw. Identifikation und Messung ihrer unterschiedlichen perceptiven Dimensionen sind in der Literatur diverse Untersuchungen beschrieben worden. Für die verschiedenen Untersuchungsziele wurde jeweils spezifisch ausgewähltes Textmaterial eingesetzt (vgl. z.B. [1], [4], [5]). Keine dieser Studien befasst sich aber explizit mit dem Einfluss von *Bedeutung* auf die Beurteilung von *Klang*. Als fundamentale weitgehend orthogonale Qualitätsdimensionen übertragener Sprache konnten die Sprachverständlichkeit ("intelligibility") und Natürlichkeit ("fidelity" bzw. "naturalness", vgl. [6]) gefunden werden. Der Begriff Klangqualität lässt sich zunächst nicht eindeutig in dieses Konzept einfügen, bezieht sich aber vermutlich weniger auf Sprachverständlichkeit denn auf Natürlichkeit.

## 2. Semiotik: Klangqualität und Bedeutung

Sprache ist ein Zeichensystem, dessen Zeichen u.a. in geschriebener oder akustischer Form vorliegen können. Triadische Zeichenmodelle der Semiotik unterscheiden drei Konstituenten (Korrelate) eines Zeichens, die das sogenannte semiotische Dreieck bilden, hier in allgemeiner Form (nach Nöth [8]):



Die drei Konstituenten eines Zeichens stehen in engem Bezug zueinander. Zeichenträger sind bei Sprache meist das geschriebene Wort oder das akustische Sprachsignal, mithin also die Form, in der es vorliegt. Referenzobjekt ist das Objekt, der Sachverhalt, welcher durch das Zeichen bezeichnet, referenziert wird. *Bedeutung* ist von unterschiedlichen Autoren auf verschiedene Weise definiert worden. Als Korrelat des Zeichens verwendet kann Bedeutung als *Sinn* des Zeichens aufgefasst werden. Bedeutung als Sinn beinhaltet demnach unter anderem die Rolle, die das Zeichen für einen Sender oder Empfänger spielt. Die Diskussion des Begriffes *Bedeutung* soll hier nicht weiter vertieft werden (vgl. z.B. Nöth [8]).

Gesprochene Sprache hat sowohl für einen Sprecher bzw. Sender als auch einen Hörer bzw. Empfänger Zeichencharakter. Das Zeichen stellt sich aus Sender- und Empfängersicht allerdings unterschiedlich dar. Für Hörtests – beispielsweise zur Messung der Klangqualität einer Sprachübertragung – müssen zunächst Sprachproben aufgezeichnet werden. Der Text liegt in geschriebener Form bereits als Zeichen vor und wird vom Sender als solches perzipiert. Der Sender liest den Text vor und setzt ihn damit in ein akustisches Zeichen um. Der Empfänger fasst das akustische Sprachsignal wiederum als ein Zeichen auf und soll nun die Klangqualität beurteilen, den Zeichenträger betreffend (Form). Handelt es sich bei dem verwendeten Sprachmaterial nun z.B. um sinnleere Sätze, wie SUS-Sätze (syntaktisch korrekte, *Semantically Unpredictable Sentences* [3]), gibt es für Zeichensender (hier ein vorlesender Sprecher) und Zeichenempfänger mehrere Möglichkeiten der Verarbeitung, beispielsweise:

- (a) Der Sender gibt dem Gelesenen die Form einer inhaltlich sinnvollen Aussage, obwohl der Text sinnleer ist. Der Empfänger soll die Form bewerten, kann sich aber möglicherweise nicht von der Sinnleere bzw. Unvollständigkeit des Zeichens lösen. Die Klangqualität wird dann negativer bewertet als für sinnvolle Sätze.

(b) Der Sender reflektiert die Sinnleere des Textmaterials durch eine entsprechende Form (z.B. überartikuliertes Sprechen). Der Empfänger kann diese Form unterschiedlich verarbeiten: Entweder, er ist in der Lage hinsichtlich der zu bewertenden Form von der Bedeutung zu abstrahieren, d.h. die Sinnleere hat keinen Einfluss auf sein Urteil. Oder aber er wird auf die Sinnlosigkeit aufmerksam gemacht, d.h. die Klangqualität wird negativer bewertet. Eine dritte Möglichkeit besteht darin, dass der Empfänger die auffällige Form unabhängig von der Bedeutung als schlechter bewertet.

### 3. Versuchsergebnisse und Schlussfolgerung

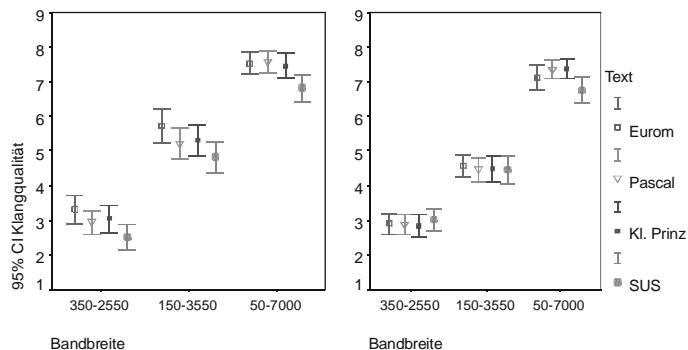
Auf Grundlage dieser Betrachtungsweise sollen die Ergebnisse zweier auditiver Tests erörtert werden, in denen naive Versuchspersonen die Klangqualität von unterschiedlich bandbegrenzten Sprachproben beurteilen sollten. Die verwendeten Sprachproben wurden in französischer Sprache aufgezeichnet: Es wurden Sätze unterschiedlichen Sinngehalts eingesetzt, die von sechs französischen Muttersprachlern (4w, 2m) vorgelesen wurden:

Gekürzte *EUROM-Sätze*: Normalsprache [3]; *SUS-Sätze*: Sinnleer; Auszüge aus *Der kleine Prinz*, Antoine de Saint-Exupéry: Einfache, klassische französische Literatur; Auszüge aus *Pensées*, Blaise Pascal und *Dieu et l'Etat*, Bakounine: Komplexer formulierte philosophische Gedanken. (Das Sprachmaterial wurde nicht explizit auf phonemische Ausgewogenheit hin untersucht. Allerdings sind SUS-Sätze aus phonemisch ausgewogenen Wortlisten gebildet, ebenso wie bei den EUROM-Sätzen auf phonemische Ausgewogenheit geachtet wurde.)

Die Sätze wurden mit Bandpässen von 350-2550 Hz, 150-3550 Hz und 50-7000 Hz gefiltert. Die Sprachproben wurden randomisiert über einen speziell konstruierten Breitband-Telefonhörer präsentiert, um möglichst den Eindruck einer Telefonsituation zu erzeugen. Die Versuchspersonen wurden dazu angeleitet, die Klangqualität der Sprachproben im Sinne der in Abschnitt 1 angegebenen Beschreibung des Untersuchungsgegenstandes auf einer mit fünf Attributen beschrifteten kontinuierlichen 11-Punkte-Skala zu bewerten (vgl. [9]).

An der ersten Versuchsreihe nahmen nur französische Muttersprachler teil, die somit das gesamte sprachliche Zeichen bewerteten (20 Personen, 10w, 10m, Bewertung der *qualité sonore*). Sie wurden im Vorfeld des Versuchs nicht speziell auf den Inhalt des Gesprochenen aufmerksam gemacht, aber darauf hingewiesen, dass auch sinnleere Sätze vorkommen werden. Dadurch sollte verhindert werden, dass diese Sätze aufgrund überraschenden Auftretens wegen ihrer schlechteren Verständlichkeit negativer bewertet würden. An der zweiten Versuchsreihe nahmen nur deutschsprachige Versuchspersonen teil, die nach eigenen Angaben keine Französischkenntnisse besaßen (15 Personen, 3w, 12m). Ihre Beurteilungen betreffen demnach eher den Zeichenträger (Form), während bezüglich Referenzobjekt und Bedeutung vermutlich auf Analogien zur deutschen Sprache zurückgegriffen wurde, oder diese ganz ausgeblendet wurden.

Die Ergebnisse zeigen, dass die *Bedeutung* des Textmaterials offenbar in die Bewertung der Klangqualität einfließt:



In der linken Darstellung ist zu erkennen, dass die Klangqualität der SUS-Sätze von den französischsprachigen Hörern unabhängig von der Bandbreite signifikant schlechter als die der Eurom-Sätze bewertet wurde. Die Beurteilungen der Auszüge aus Pascals "Gedanken" und dem "kleinen Prinzen" liegen dazwischen. Die deutschsprachigen Versuchspersonen dagegen bewerteten die unterschiedlichen Texte für jeweils feste Bandbreite im Mittel über alle Sprecher weitgehend identisch (rechte Darstellung). Die Form der Sprachproben wird demnach als ähnlich bewertet, wenn aufgrund der Fremdheit der Sprache

kein Bezug zur transportierten Bedeutung besteht. Lediglich im Falle der Breitbandübertragung wurden die SUS-Sätze von den deutschsprachigen Versuchspersonen insgesamt etwas schlechter beurteilt. Eine Analyse der Urteile in Abhängigkeit der einzelnen Sprecher zeigt, dass die von einer bestimmten Sprecherin gesprochenen SUS-Sätze sowohl von den französischsprachigen als auch von den deutschsprachigen Hörern besonders negativ bewertet wurden.

Wird diese Sprecherin aus der Wertung ausgenommen, so sind die verschiedenen Texte anhand der Urteile der deutschsprachigen Versuchspersonen nicht mehr zu unterscheiden. Auf die durchschnittlichen Urteile der französischsprachigen Versuchspersonen hat das Streichen der Sprecherin nur einen geringen Einfluss, die Urteile für SUS-Sätze und Eurom-Sätze sind für die jeweilige Bandbreite dann immer noch statistisch signifikant verschieden. Daraus kann geschlossen werden, dass die Sprecherin die sinnleeren Sätze in einer selbst für Nicht-Muttersprachler auffälligen Weise vorgelesen hat, was durch informelles Abhören der Sprachproben bestätigt wurde (monoton mit erhöhter Stimmlage). Demnach hat die Sprecherin als Sender des sprachlichen Zeichens offenbar die Sinnleere durch eine entsprechend veränderte Form ausgedrückt, die vom Hörer negativer bewertet wird (vgl. Abschnitt 2, Fall (b)).

Für die abhängig vom Sprachmaterial unterschiedlichen Urteile der französischsprachigen Hörer gibt es verschiedene Erklärungsansätze: Entweder, sie sind zusätzlich durch die Bedeutung des Sprachmaterials beeinflusst, so dass das sprachliche Zeichen bei Sinnleere als unvollkommen wahrgenommen wird (vgl. Abschnitt 2, Fall (a)): D.h., im Gegenzug, dass ein der Bandbreite und der Telefonsituation angemessener, erwarteter Sinn durch eine positive Beurteilung der Form widerspiegelt wird. Oder aber die möglicherweise auch bei SUS-Sätzen von anderen Sprechern veränderte Form betrifft Merkmale, die von den deutschsprachigen Versuchspersonen nicht wahrgenommen wurden (Fall (b)).

Die Abstände zwischen den Urteilen der französischsprachigen Hörer bei verschiedenen Bandbreiten sind aber unabhängig vom Textmaterial. Wenn das Ziel einer Klangqualitätsmessung lediglich der subjektive Vergleich unterschiedlicher Kanaleigenschaften ist, dann sind prinzipiell alle verwendeten Textmaterialien dafür geeignet.

Bereits die hier dargestellte einfache Untersuchung zeigt, dass der Einfluss der *Bedeutung* des verwendeten Sprachmaterials bei der Entwicklung eines Verfahrens zur auditiven, absoluten Messung formeller Eigenschaften von Sprache (hier: Klangqualität) berücksichtigt werden muss. Es existiert offenbar kein Sprachmaterial, welches so prototypisch ist, dass es wirklich unabhängig von Inhalt und Bedeutung Klangeigenschaften besitzt, die das Testverfahren dann nur misst. Bei gesprochener Sprache kann demnach nicht von einer simplifizierten Form ausgegangen werden. Form und Bedeutung von Sprache konstituieren sich für Sender und Empfänger immer wieder neu. Es ist somit notwendig, die in der Sprachqualitätsforschung häufig vorgenommene Trennung von akustischem Signal und Gesamtheit des Zeichens mit Bezug auf das Messziel stets neu zu hinterfragen.

**Danksagung:** Die vorliegende Arbeit entstand im Rahmen einer PROCOPE Kooperation zwischen dem IKA (Prof. J. Blauert, PD. U. Jekosch) und dem LMA, CNRS, Marseille (G. Canévet).

#### Literatur:

- [1] Bappert, V., Blauert, J. (1994). Auditory Quality Evaluation of Speech Coding Systems, *Acta Acustica*, 2, 49-58.
- [2] Bednarzyk, M. (1999). Qualitätsbeurteilung der Geräusche industrieller Produkte, VDI-Verlag, D-Düsseldorf.
- [3] Gibbon, D., Moore, R., Winski, R., eds. (1997). *Handbook of Standards and Resources for Spoken Language Systems*, De Gruyter, D-Berlin.
- [4] Huggins, A.W.F., Nickerson, R.S. (1985). Speech quality evaluation using "phoneme-specific" sentences, *J. Acoust. Soc. Am.*, 77 (5), 1896-1906.
- [5] Jekosch, U. (2000). *Sprache hören und beurteilen: Ein Ansatz zur Grundlegung der Sprachqualitätsbeurteilung*, Habilitationsschrift, Universität Essen, D-Essen.
- [6] McGee, V.E. (1964). Semantic Components of the Quality of Processed Speech. In: *J. of Speech and Hearing Research*, 7, 310-323.
- [7] Möller, S., Raake, A. (2000). Planung von Telefon-Sprachqualität: Unterschiedliche Modellansätze, in: *Fortschr. Akust. DAGA 2000*, 208-209, Deutsche Ges. Akust., D-Oldenburg.
- [8] Nöth, W. (2000). *Handbuch der Semiotik*, J.B. Metzler, D-Stuttgart.
- [9] Raake, A. (2000). Perceptual Dimensions of Speech Sound Quality in Modern Transmission Systems, *Proc. ICSLP 2000*, 4, 744-747, CHN-Beijing.