

# Private Telefonie im Fahrzeuginnenraum durch Sprachmaskierung

René Skowronek<sup>1</sup>, Gerhard Krump<sup>1</sup>

<sup>1</sup>Technische Hochschule Deggendorf, 94469 Deggendorf, E-Mail: gerhard.krump@th-deg.de

## Einleitung

Die Zeit, die täglich im Fahrzeug verbracht wird, steigt seit Jahren kontinuierlich an. Um bei Telefongesprächen hierbei die gesetzlichen Rahmenbedingungen einzuhalten, steht eine Freisprecheinrichtung zur Verfügung, die allerdings dem Wunsch, einen vertrauensvollen Dialog zu führen, widerspricht. Um auch in Chauffeurfahrzeugen mit unbekanntem Fahrer eine Privatsphäre sicherstellen zu können, sollte in Zusammenarbeit mit der BMW Forschung und Technik GmbH eine Funktion zur Maskierung der Sprache entwickelt werden. Mit Hilfe von Maskiersignalen, die über Kopfstützenlautsprecher ausgegeben werden, sollte der Sprachinhalt telefonierender Passagiere für den Fahrer unverständlich gemacht werden. Bei der Erstellung der Maskiersignale wurde auf zwei Methoden zurückgegriffen:

1. Spektrale Verdeckung: Zwei Versuchssignale (spektrale Maskierer a und b) sollten durch spektrale Maskierung zu einer Verdeckung der Information führen. Die Maskiersignale basieren auf frequenzgruppengefiltertem Rosa Rauschen. Während Signal a durch die Filterung durch Bandpässe realisiert wurde, wurden bei Signal b Tiefpässe verwendet. In beiden Fällen entsprachen die Grenzfrequenzen denen der Frequenzgruppen. Das Sprachsignal wurde jeweils auf gleiche Weise gefiltert und der jeweilige Spitzenwert pro Frequenzband ermittelt. Über diese Spitzenwerte wurden die Frequenzbänder des Rosa Rauschens gesteuert. Auf diese Art und Weise sollte das Sprachsignal in Abhängigkeit der Frequenzbandpegel bandspezifisch verdeckt werden. [1]

2. Informationelle Maskierung: Während bei der spektralen Maskierung die Verdeckung durch die simultane Überlagerung im Frequenzbereich erreicht wird, wird hier der Sprecher durch zusätzlich auftretende Maskiersprecher unverständlich gemacht, so dass eine Separation zwischen Maskier- und Zielsprecher verhindert wird. Hierzu wurden modifizierte Stimmversionen der Sprecherstimme erzeugt, die dann als Maskiersprecher dienen. [2]

Ein Hörversuch sollte Aufschluss darüber geben, welche Pegel bei den verschiedenen Signalen nötig sind, um eine ausreichende Verdeckung der Sprache zu erzielen.

## Versuchsdurchführung

Als Versuchsfahrzeug wurde ein 7er BMW mit modifizierten Kopfstützen verwendet, in denen sich zwei Lautsprecher für die Ausgabe der Maskiersignale befanden. Als Hörposition der Testpersonen wurde der Beifahrersitz gewählt. Als Sprecher diente ein sprechender Kunstkopf auf der Rückbank hinter dem Beifahrer auf Kopfhöhe. Zur Sprachaufnahme wurde gemäß Abb. 1 ein 1/2"-Druck-Mikrofon am Fahrzeughimmel mittig über dem Kunstkopf angebracht.

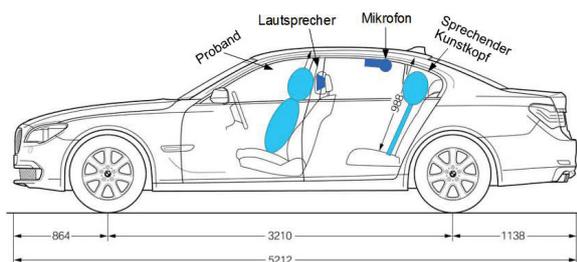


Abb. 1: Skizze des Versuchsaufbaus im Fahrzeug.

Als Sprachsignale wurden 40 Zufallszahlen im drei- bis vierstelligen Bereich gewählt, die durch einen Sprachsimulator erzeugt wurden. Diese wurden in zufälliger Reihenfolge wiedergegeben und im Pegel erhöht.

Die Probanden sollten die vom Kunstkopf gesprochenen Zahlen wiederholen, so dass permanent deren Sprachverständlichkeit kontrolliert wurde. Der Versuchsleiter steuerte den Kunstkopfsprecher über eine grafische Benutzeroberfläche. Dabei wurde der Pegel-Schwellwert, bei dem nur ein Teil der Zahlen richtig verstanden wurde, als „gelber“ Pegel abgespeichert. Der Wert, bei dem die Zahlen eindeutig nicht mehr verstanden wurden, wurde als „roter“ Pegel markiert. Solange die Zahlen korrekt wiedergegeben werden konnten, wurde der Pegel pro gesprochener Zahl um 3 dB erhöht, sobald der „gelbe“ Pegel erreicht wurde, wurde der Pegel um 1 dB pro Zahl erhöht, bis die Sprache unverständlich war. Zur Überprüfung der Unverständlichkeit wurden drei weitere Zahlen mit gleichem Pegel dargeboten, bevor der Test beendet wurde. Zudem wurden die Testpersonen gebeten, ihre subjektiven Empfindungen zu den Signalen anzugeben. Um den Einfluss von Musik als Grundmaskierer zu berücksichtigen, wurde der Versuch jeweils mit und ohne Musik durchgeführt. Dafür wurde das Lied „Ruby“ von den „Kaiser Chiefs“ verwendet. Der Pegel der Musik wurde von den Versuchsleitern einmalig, subjektiv auf einen angemessenen Pegel für Hintergrundmusik eingestellt und konnte von den Testpersonen nicht verändert werden. Somit entstanden bei drei Maskiersignalen insgesamt sechs Maskierszenarien. Um eine effizientere Maskierung zu realisieren, wurde das Musiksignal ebenfalls gefiltert und die Spitzenwerte der einzelnen Bänder ermittelt. Die Ansteuerung der Maskiersignalbänder erfolgte nur, wenn der Spitzenpegel des Sprachsignals höher war, als der Spitzenpegel des Musiksignals des jeweiligen Bandes.

Szenario 1-2: spektrale Maskierer ohne Musik

Szenario 3 : informationeller Maskierer ohne Musik

Szenario 4-5: spektrale Maskierer mit Musik

Szenario 6: : informationeller Maskierer mit Musik

Um die Pegeldifferenzen zwischen Nebengeräuschen und Maskiersignalen vergleichen zu können, wurden wie in Tabelle 1 dargestellt jeweils einzeln das Fahrgeräusch, der Sprecher am Hör- und Sprechplatz und das Musiksignal gemessen. Alle Signale, bis auf den Sprecher am Sprechplatz, wurden auf dem Beifahrersitz, also an der Position der Versuchsperson, gemessen. Dafür wurde ein Mikrofon-Array mit acht Mikrofonen verwendet, das an der Kopfstütze angebracht wurde. Es wurde mit der Zeitbewertung „fast“ (125 ms) gemittelt und die Signale A-bewertet, um tieffrequente Nebengeräusche auszuschließen.

Tabelle 1: Pegel der verschiedenen Geräuschquellen im Fahrzeug

Signal	dB(A)
Fahrgeräusch	56
Sprecher am Hörplatz	60
Sprecher am Sprechplatz	70
Musik	51

## Messergebnis

Um die Verteilung der "roten" Pegel aller 16 Testpersonen (13 Männer und 3 Frauen, im Alter zwischen 16 und 35 Jahren) zu betrachten, wurde auf Kastendiagramme zurückgegriffen. Sie werden durch den maximalen und den minimalen Wert (Querstriche) beschrieben sowie dem Median und dem oberen und unteren Quartil (Kasten). Zwischen dem oberen und unteren Quartil befinden sich somit 50% der eingestellten Werte, wobei der Median diese halbiert. Von den restlichen 50% befinden sich 25% oberhalb des oberen Quartils und 25% unterhalb des unteren Quartils. Hierbei wurde der Pegel des einzelnen Maskiersignals ohne Nebengeräusche betrachtet (Abb. 2 und 3). In dieser Darstellungsmethode wird sichtbar, dass die Boxen der Szenarien 3 und 6 unterhalb der anderen Szenarien liegen. Allerdings lassen sich hier die stärkeren Schwankungen der Szenarien mit Musik deutlich erkennen. Aufgrund der Steuerung der Maskiersignale in Abhängigkeit der Musik ist der Abstand des Minimum und des Maximum größer als bei den Maskierern ohne Musik. Auffällig ist die schmale Box bei Szenario 3. Die Hälfte der eingestellten Werte befinden sich innerhalb von 1 dB. Dieses Szenario ist mit einem Median-Pegel von ca. 70 dB(A) auch das Signal mit dem niedrigsten notwendigen Maskierungspegel. Das reine Maskiersignal, bei dem 75% (oberes Quartil) der Testpersonen nichts mehr verstehen, hat einen Pegel von 70,4 dB(A) (siehe Abb. 2). Es besitzt 10,4 dB(A) mehr Pegel als der Sprecher am Hörplatz.

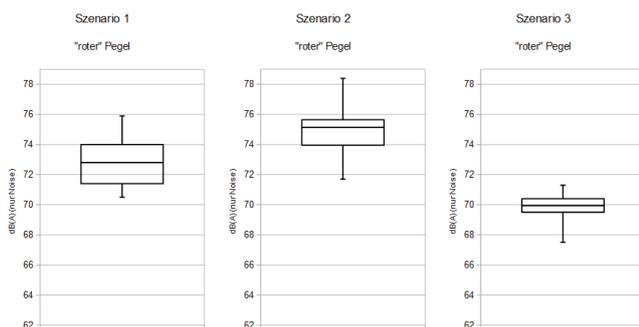


Abb. 2: Pegel des Maskiersignals für Szenario 1-3.

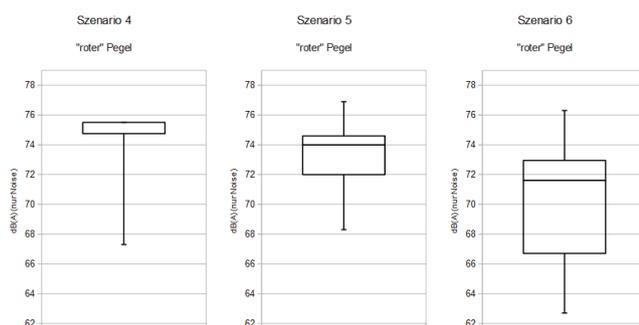


Abb. 3: Pegel des Maskiersignals für Szenario 4-6.

## Subjektive Empfindungen der Signale

Subjektiv wurde Szenario 3 von der Mehrheit als relativ angenehmes Signal empfunden. Einige Testpersonen gaben an, größere Probleme bei der Filterung des Zielsprechers zu haben. Im Vergleich zu den anderen Szenarien wurde die empfundene Gesamtlautstärke positiver empfunden.

Allerdings gaben einige der Testpersonen an, den Anfang der Zahlen besser zu verstehen als bei Szenario 1 und 2. Dies könnte daran liegen, dass Effekte im Zeitbereich eine Rolle spielen.

Das tiefpassgefilterte Rauschen (Signal b) wurde weitestgehend als ein gewohntes Rauschsignal wahrgenommen. Das frequenzgruppendefinierte Rauschen hingegen wurde häufig als unangenehm bezeichnet, was an der veränderten Stimmcharakteristik liegen kann, die eine ungewohnte Geräuschkulisse erzeugt.

Die Verwendung der Musik als Grundmaskierer wurde von den Probanden unterschiedlich wahrgenommen. Obwohl die messtechnische Auswertung keinen positiven Effekt ergab, wurde das Musiksignal von einigen Testpersonen positiv bewertet. Das Hörempfinden scheint sich subjektiv zu verbessern. Die Personen schienen sich in diesem Fall, mehr auf das Musiksignal zu fokussieren, als auf das Maskiersignal. Somit wurde das Maskiersignal weniger stark wahrgenommen. Außerdem gaben einige Testpersonen an, sich mehr konzentrieren zu müssen, da die Aufmerksamkeit reduziert wurde. Dies führte auch zu einem besseren Maskiereffekt, da die Personen Probleme dabei hatten den Sprecher zu verstehen. Der andere Teil der Versuchspersonen schien sich von der Musik nicht ablenken zu lassen. Sie waren der Meinung, dass die Musik keinen Einfluss auf die Sprachverständlichkeit hat. Möglicherweise spielt auch der persönliche Musikgeschmack, und somit der ausgewählte Musiktitel, eine entscheidende Rolle. Bei Musik, die einer Person gefällt, ist es wahrscheinlicher, dass sie sich hauptsächlich darauf konzentriert und den Sprecher vernachlässigt.

## Zusammenfassung

Nach Auswertung der Messergebnisse wird deutlich, dass das Maskiergeräusch mit informationeller Maskierung weniger Pegel benötigt um Sprache zu maskieren als die Maskiergeräusche, die durch Rauschen erzeugt werden. Die informationelle Maskierung ohne Musik (Szenario 3) wurde dabei mit dem niedrigsten Pegel bewertet. Bei einem Pegel von ca. 70 dB(A) war es 75% der Versuchspersonen nicht mehr möglich, den Sprecher zu verstehen. Der Pegel ist somit um ca. 10 dB(A) höher als der Sprecherpegel am Hörplatz.

Auch das Hörempfinden bewerteten die Testpersonen bei der informationellen Maskierung weitestgehend positiv. Die entstehenden Maskiersprecher erschweren es zudem einigen der Probanden, sich auf den Sprecher zu konzentrieren. Die Musik als Grundmaskierer zu verwenden, hat in Bezug auf den nötigen Pegel keinen Erfolg gebracht. Allerdings wurde es von einigen Testpersonen dennoch als angenehmer empfunden, wenn Musik verwendet wurde. Auch der Einfluss auf den Maskiereffekt wurde teilweise als positiv bezeichnet.

## Literatur

- [1] Zwicker E., Psychoakustik, Springer Verlag, 1982
- [2] Brungart D. u. Simpson B., Informational and energetic masking effects in the perception of multiple simultaneous talkers, JASA, 2001