

# Computersimulation des Sprachübertragungsindex STI nach DIN EN 60268-16:2003

Thomas Steinbrecher<sup>1</sup>

<sup>1</sup> Bose GmbH, Professional Systems Division, 61381 Friedrichsdorf, Deutschland

Email: thomas\_steinbrecher@bose.com

## Einleitung

Sowohl die Simulation als auch die Messung der Sprachverständlichkeit gewinnen im Zuge der sich verändernden Normung von elektroakustischen Sprachalarmsystemen (SAS) zunehmend an Bedeutung. So wurde im September des Jahres 2007 mit der DIN VDE 0833-4 eine neue Deutsche Norm für SAS innerhalb von Brand- oder Gefahrenmeldeanlagen veröffentlicht, die detaillierte Anforderungen an die Messung der Sprachverständlichkeit beschreibt. Auch auf internationaler Ebene ist die diesbezügliche Standardisierung im Wandel, so erschien gerade die ISO 7240-19 und andere Anwendungsrichtlinien, wie der British Standard BS 5839-8, befinden sich in Überarbeitung.

Als Methode zur Bestimmung der Sprachverständlichkeit von SAS hat sich in den vergangenen Jahren der Sprachübertragungsindex (engl.: Speech Transmission Index - STI) durchgesetzt, der zusammen mit seinen Unterformen (das STIPA-Verfahren wird mittlerweile häufig angewendet) in DIN EN IEC 60268-16, letztmalig überarbeitet im Jahr 2003, spezifiziert ist. Der STI basiert auf den Ergebnissen von Steeneken und Houtgast (siehe z.B. [1]) und erlaubt die objektive Evaluierung eines Übertragungskanals bei Beeinträchtigung durch unterschiedliche Störungen im Zeit- oder Frequenzbereich. Er ist mit Hilfe subjektiver Verständlichkeitstests für eine Vielzahl von Testkonditionen validiert worden. Die Berechnung des STI erfolgt durch die Bestimmung und Nachbearbeitung der Modulationsübertragungsfunktion MTF des Übertragungskanals. Messtechnisch lässt sich die MTF entweder auf analogem Weg (unter Verwendung modulierter Rauschsignale) oder durch Fouriertransformation der quadrierten Impulsantwort bestimmen („Schroeder-Methode“, [2]). Bei der Simulation des STI ist bevorzugt die letztere Methode anzuwenden, siehe z.B. [3].

In den vergangenen Jahren wurden häufig zwei Probleme mit dem STI assoziiert, nämlich die geringe Empfindlichkeit gegenüber den Einflussfaktoren Amplitudengang und Echos, welche deshalb nachfolgend genauer analysiert werden.

## Einfluss des Amplitudenganges

Gemäß Norm ist die MTF als Grundlage des STI in den Oktavbändern von 125 Hz bis 8 kHz zu bestimmen. Dabei darf nicht außer Acht gelassen werden, dass die Berechnung des Modulationsverlustes weitestgehend bandgetrennt erfolgt. Einzig die Modellierung der pegelabhängig wirkenden Aufwärtsmaskierung bildet hier eine Ausnahme. Da bei linearen Systemen der Einfluss anderer typischer Parameter wie Nachhall oder Echos unabhängig vom Amplitudengang ist, verbleibt als einzige signalpegelabhängige Einflussgröße der Signal-Stör-Abstand (SNR). Wird also unter sehr geräuscharmen Randbedingungen gemessen oder simuliert, so erge-

ben sich auch trotz stark eingeschränkter Wiedergabespektren noch sehr hohe STI-Werte.

Um die Empfindlichkeit des STI in Bezug auf Verzerrungen des Amplitudenganges zu untersuchen, wurde für sechs unterschiedliche Wiedergabespektren der STI in Abhängigkeit vom SNR berechnet. Die untersuchten Konditionen umfassen ausgewählte und teils praxisnahe Einschränkungen der Übertragungsbandbreite. Sie sind zusammen mit dem normgerechten Sprachspektrum (männliche Sprache) als Oktavbandpegel in Tabelle 1 angegeben.

Tabelle 1: Untersuchte Amplitudengänge

EQ	Oktavbandpegel [dB rel]						
	125	250	500	1k	2k	4k	8k
Flat	0	0	0	0	0	0	0
Knee 4k	0	0	0	0	0	-6	-12
TNO BP2	-40	-8	0	0	0	-4	-40
Dip 2k	0	0	0	0	-15	0	0
LP 500	0	0	0	-18	-36	-40	-40
BP 1k	-40	-40	-24	0	-24	-40	-40
Sprache m	+2.9	+2.9	-0.8	-6.8	-12.8	-18.8	-24.8

Das (relative) Frequenzspektrum des Störgeräusches wurde auf das Sprachspektrum eingestellt, so dass sich für die lineare Bedingung („Flat“) ein frequenzunabhängiger SNR ergibt, der dem nominellen Wert entspricht. Dieses nominelle Geräuschkennlinien wurde nun für die Untersuchungen der anderen Übertragungskanäle jeweils konstant gehalten. Für sie wird der SNR also in den Oktavbändern mit abgesenktem Signalpegel entsprechend verringert. Abb. 1 zeigt die Ergebnisse für Signal-Stör-Abstände von -50 bis +50 dB (nom.).

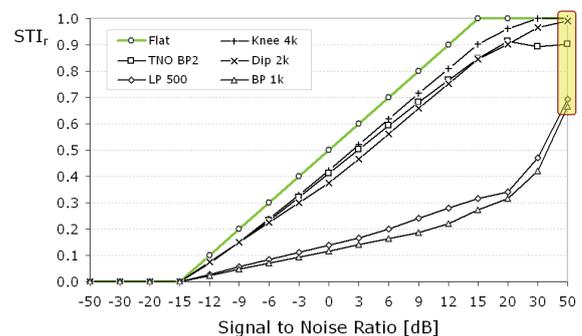


Abbildung 1: STI in Abhängigkeit vom nominellen Signal-Stör-Abstand (SNR, siehe Text) bei unterschiedlichen Verzerrungen des Amplitudenganges gemäß Tabelle 1.

Für ein ungestörtes Übertragungsspektrum ergibt sich die grüne Kurve, welche normgemäß bei  $\pm 15$  dB SNR begrenzt wird. Die eingeschränkten Spektren verursachen bei sehr hohen Störabständen (50 dB, s. Markierung) nur unwesentliche Änderungen des STI. Bei allen für Sprachalarmanlagen relevanten SNR's im Bereich von etwa 0 bis 20 dB, führen

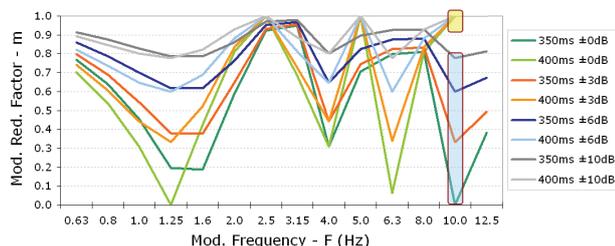
jedoch bereits moderate Einschränkungen wie ein Tiefpassverhalten ab 4 kHz zu einer Verminderung des STI um bis zu 0,1, was einer deutlichen Verschlechterung entspricht (ca. 3x JND). Für die Konditionen mit extremen Einschränkungen (Tiefpass ab 500 Hz bzw. Beschränkung auf das Oktavband 1 kHz) ergeben sich dramatische Einbrüche, welche die Untauglichkeit solcher Systeme unter realen Bedingungen belegen. Es sei betont, dass sich bei einem elektroakustischen System der hier beschriebene Amplitudengang aus der Summe aller Komponenten - vom Mikrofon bis zum Abhörpunkt im Schallfeld - zusammensetzt.

## Einfluss von Echos

Löst man die „Schroeder-Gleichung“ für ein System, das durch einen Impuls und eine verzögerte und abgeschwächte Wiederholung dieses Impulses nach einer Zeit  $\Delta T$  gekennzeichnet ist, so ergibt sich nach Betragsbildung die folgende Abhängigkeit zwischen dem Erhalt an Modulationstiefe  $m$  und der Modulationsfrequenz  $F$ , der Verzögerungszeit  $\Delta T$  und dem Verhältnis  $\delta$  der Intensitäten von Direktschall (1st arrival) und Echo (Gleichung 1):

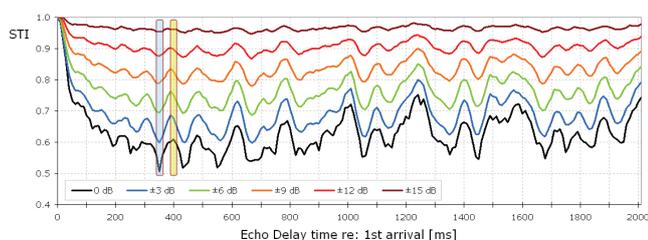
$$m(F) = \frac{\sqrt{1 + 2\delta \cos(2\pi F \Delta T) + \delta^2}}{1 + \delta} \quad [-] \quad (1)$$

Die typische Form einer durch ein Echo gestörten MTF ist durch eine oder mehrere Kerben gekennzeichnet, deren Tiefe durch das Pegelverhältnis bestimmt wird, siehe Abb. 2.



**Abbildung 2:** Durch Echo mit 350 bzw. 400 ms Verzögerung gestörte MTF bei unterschiedlichen Echopegeln.

Die Position der Kerbe im Modulationsfrequenzspektrum wiederholt sich periodisch bei ungeraden Vielfachen des Kehrwertes der doppelten Verzögerungszeit. Hier wird der Cos-Term in (1) minimal, also gleich -1. Prinzipiell hat diese Periodizität auch einen entsprechenden Einfluss auf den STI, wie in Abb. 3 gezeigt wird.



**Abbildung 3:** STI in Abhängigkeit von der Verzögerungszeit eines einzelnen Echos bei unterschiedlichen Pegeln. Die markierten Zeiten beziehen sich auf Abb. 2.

Deutlich im Verlauf zu erkennen sind die wiederkehrenden lokalen Minima und Maxima, von denen zwei benachbarte

Regionen exemplarisch gekennzeichnet sind: Alle 50 ms alterniert der Modulationsindex bei  $F = 10$  Hz zwischen Minimum (blau) und Maximum (gelb). Für andere Modulationsfrequenzen wiederholt sich dieser Wechsel in anderen Zeitabständen. Eine weitere Problematik ergibt sich aus der Symmetrie der Cos-Funktion, die für negative Zeitverzögerungen identische Ergebnisse liefert. Dies gilt auch für das Intensitätsverhältnis  $\delta$ : Wird der Pegel von Direktschall und Echo vertauscht, ändert sich die MTF nicht.

Selbstverständlich hat die Tatsache, dass sich die Modulation für größere Verzögerungszeiten periodisch erhöht, keine Entsprechung im menschlichen Hörprozess: Eine um mehrere Perioden verschobene Modulation dient nicht der Verbesserung der Verständlichkeit. Auch die Vertauschung von Echo und Direktschall bewirkt subjektiv deutlich größere Unterschiede. In diesem Zusammenhang sei angemerkt, dass die von Steeneken und Houtgast durchgeführte Validierung des STI für viele Testbedingungen unter Einfluss von Echos sogar eine Unterbewertung der Sprachverständlichkeit durch den STI ergab. Dieses Ergebnis ist möglicherweise der Tatsache geschuldet, dass die Parametrierung des STI auf Basis der Übereinstimmung mit der Verständlichkeit von in Sätzen eingebetteten CVC-Nonsens-Wörtern erfolgte und nicht mit Hilfe von Satzverständlichkeiten.

Im Allgemeinen ist zu beachten, dass sich die Periodizität des STI erst durch die diskrete Abtastung der MTF besonders bemerkbar macht. Dabei kommt zusätzlich zum Tragen, dass die normgemäßen Modulationsfrequenzen (s. Abb. 2) teils Vielfache voneinander sind. Zur Verdeutlichung dieses Effektes wurde Abb. 3 deshalb unter Zugrundelegung nicht-bevorzugter Frequenzen  $F$  (0,625 Hz, 0,8 Hz, 1,0 Hz und ihren exakten Vielfachen) berechnet. Bei einer DFT wie im vorliegenden Fall wiederholt sich deshalb das Modulationsspektrum nach einer Verzögerungszeit, die dem kleinsten gemeinsamen Vielfachen der Periodendauer aller Modulationsfrequenzen entspricht. Für die hier verwendeten Frequenzen sind dies 40 Sekunden.

## Zusammenfassung

Zwei bekannte Schwachstellen des STI-Verfahrens und der zugrunde liegenden Modulationsübertragungsfunktion MTF wurden analysiert. Die Empfindlichkeit des STI gegenüber Schwankungen des Amplitudenganges kann durch Berücksichtigung eines geeigneten Grundgeräusches deutlich verbessert werden. Die mangelnde Empfindlichkeit des STI gegenüber Echostörungen ist systembedingt und kann durch Nachbearbeitung nicht angemessen verbessert werden. Zur Analyse von Echostörungen sollten deshalb bevorzugt weitere Methoden oder ggf. eine Auralisation verwendet werden.

## Literatur

- [1] Houtgast, T., Steeneken, HJM. "A physical method for measuring speech transmission quality". JASA, Vol. 67 (1) (1980).
- [2] Schroeder, M.R. "Modulation Transfer Functions: Definitions and Measurement". Acustica 49 (1981), 179-182.
- [3] Steinbrecher, T. "Computersimulation der Sprachverständlichkeit". Fortschritte der Akustik (DAGA 2006).