

Anwendbarkeit des SII für nichtlinear dynamik-komprimierte Sprache

Tobias Sankowsky, Martin Hansen

Institut für Hörtechnik und Audiologie, FH Oldenburg/Ostfriesland/Wilhelmshaven, 26121 Oldenburg

Einleitung

Der SII (speech intelligibility index) [1] ist ein seit langer Zeit verwendetes Modell zur Vorhersage der realisierbaren Sprachverständlichkeit von Übertragungssystemen. Er wird angewendet um Einflüsse von linearer Frequenzgangverzerrung, additiv überlagertem Störgeräusch, linearer Dämpfung oder Verstärkung und Nachhall auf die Sprachverständlichkeit zu bestimmen. Darüber hinaus lässt der SII auch eine Aussage über die Auswirkung eines frequenzabhängigen Hörverlustes auf die Sprachverständlichkeit zu. Eine Forderung des SII ist die Linearität des Übertragungssystems. Eine Reihe von Übertragungssystemen weist jedoch eine nicht unerhebliche, bewusst eingeführte Nichtlinearität auf, wie z.B. die Amplitudenkompression in modernen Hörgeräten. Amplitudenkompression wird dort zur Kompensation von Recruitment bei sensorineuralem Hörverlust eingesetzt. In diesem Beitrag wurde der Einfluss von Amplitudenkompression auf die Sprachverständlichkeit von normalhörenden Probanden untersucht. Hierfür wurden verschiedenen Kompressionsverfahren, wie sie auch in Hörgeräten zum Einsatz kommen, verwendet. Neben breitbandiger Kompression wurde auch Kompression von Sprache in mehreren separaten spektralen Bändern verwendet.

Sprachverständlichkeit bei Kompression

Um den Einfluss von Amplitudenkompression auf die Sprachverständlichkeit zu untersuchen, wurde der Oldenburger Satztest [2] verwendet, um die Sprachverständlichkeit in Störgeräusch zu bestimmen. Neben dem im Oldenburger Satztest standardmäßig verwendeten Störgeräusch (N1) wurden alle Messungen auch mit einem zweiten Störgeräusch (N2) durchgeführt, dessen Spektrum relativ zu N1 proportional zur Frequenz um den Faktor 1,3781 verschoben wurde. Die Messungen wurden an 9 normalhörenden, trainierten Probanden im Alter von 24 bis 37 Jahren durchgeführt.

Das Sprachmaterial des Oldenburger Satztestes wurde mit einem in Matlab implementierten Kompressor amplitudenkomprimiert. Die Parameter des Kompressors wurden so gewählt, dass der Dynamikbereich von Sprache in Ruhe effektiv deutlich verringert wurde: Dazu wurden Ein- und Ausschwingzeit gemäß der für Hörgeräte üblichen Definition auf den sehr kurzen Wert von jeweils 5 ms eingestellt. Die Kompressionsschwelle lag bei -80 dB FS, so dass ein sehr breiter Pegelbereich komprimiert wurde. Die lineare Verstärkung unterhalb der Kompressionsschwelle betrug 30 dB. Als statisches Kompressionsverhältnis CR wurden die Werte CR=2 und 4 verwendet. Mit diesen Kompressionseinstellungen

wurde die Sprachverständlichkeit für breitbandige („1-kanalige“) Kompression, Kompression in 4 und in 19 spektralen Bändern untersucht, im folgenden mit B1, B4 und B19 bezeichnet. Zudem wurde auch die Sprachverständlichkeit bei unkomprimierter Sprache gemessen. Die Ergebnisse bei Verwendung des Störgeräusch N1 sind in Abb. 1 dargestellt. Es zeigt sich, dass die Sprach-

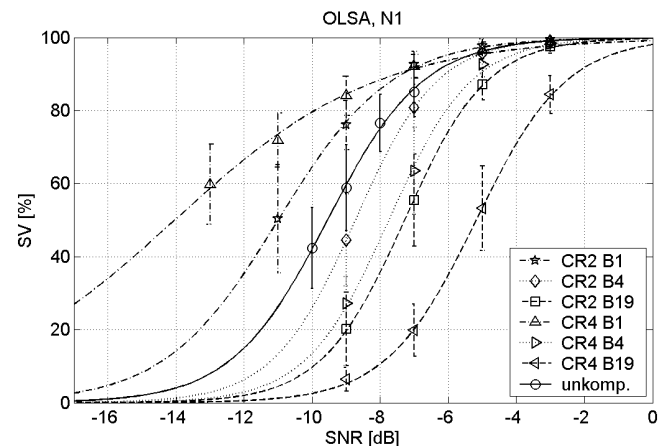


Abbildung 1: Ergebnis der Sprachverständlichkeit als Funktion des S/N bei Störgeräusch N1. Dargestellt sind Mittelwerte und Standardabweichungen für alle Probanden, mit daran angepassten Diskriminationsfunktionen.

verständlichkeit bei B1-Kompression besser wird, wobei der Verlauf der Diskriminationsfunktion deutlich flacher ist. Diese Verbesserung der Sprachverständlichkeit lässt sich durch ein bei hohen Frequenzen erhöhtes Langzeitspektrum aufgrund der breitbandigen Kompression relativ zur unkomprimierten Sprache erklären. Bei Kompression in 4 und in 19 spektralen Bändern wurde die Sprachverständlichkeit schlechter als bei unkomprimierter Sprache, wobei die Diskriminationsfunktion geringfügig steiler war.

Vorhersage der Sprachverständlichkeit durch den SII

Anhand der Leistungsdichtespektren von Sprache und Störgeräusch lässt sich der SII berechnen [1]. In Abb. 2 ist die gemessene Sprachverständlichkeit für unkomprimierte Sprache als Funktion des SII für beide Störgeräusche dargestellt. Es zeigt sich bei gleichem berechneten SII eine Differenz in der gemessenen Sprachverständlichkeit von bis zu 32%. Für den SII – in diesem Fall angewandt auf ein lineares Übertragungssystem – bedeutet dies eine wesentlich ungenauere Vorhersage der Sprachverständlichkeit als bei der einfachen spektralen Veränderung des Störgeräusches zu erwarten war.

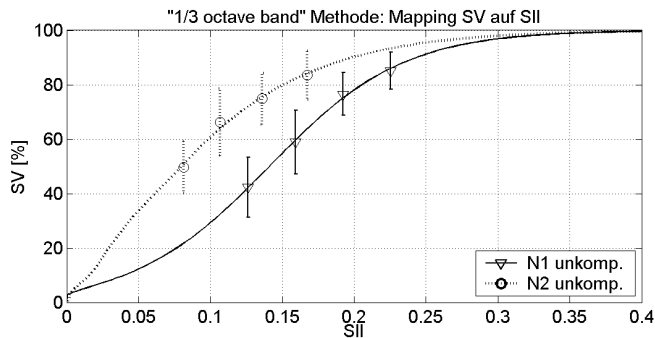


Abbildung 2: Gemessene Sprachverständlichkeit bei unkomprimierter Sprache als Funktion des berechneten SII für zwei verschiedene Störgeräusche.

Der Zusammenhang von gemessener Sprachverständlichkeit und dem SII bei Kompression ist im folgenden durch die Differenz in der Sprachverständlichkeit relativ zu unkomprimierter Sprache dargestellt. Die untersuchte Amplitudenkompression bewirkt insbesondere bei breitbandiger Kompression eine nicht unerhebliche Änderung des Langzeitspektrums der Sprache. Dies wurde bei der Berechnung des SII in Form einer Übertragungsfunktion eines hypothetischen linearen Systems berücksichtigt. In Abb. 3 ist der Einfluss von Kompression auf die Vorhersage der Sprachverständlichkeit durch den SII bei Verwendung des Störgeräusches N1 dargestellt. Es ist zu sehen,

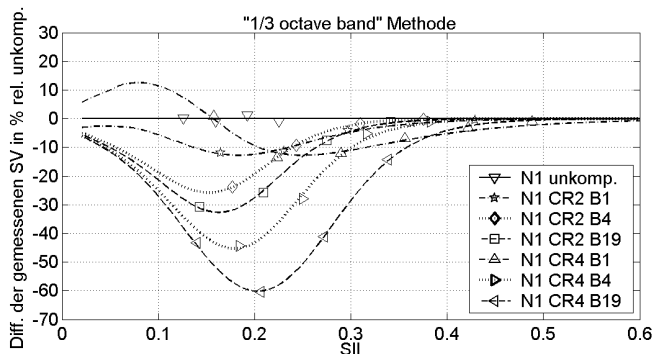


Abbildung 3: Differenz aus gemessener Sprachverständlichkeit und vorhergesagter Sprachverständlichkeit für unkomprimierte Sprache als Funktion des SII bei Störgeräusch N1.

dass der SII die schlechtere Sprachverständlichkeit bei Kompression in 4 und in 19 spektralen Bändern nicht vorhersagen kann. Für B1-Kompression zeigten sich geringere Abweichungen relativ zu unkomprimierter Sprache als für B4 und B19.

Optimierung des SII

Zur Optimierung des SII wurde das effektive Kompressionsverhältnis CR_{eff} des Übertragungssystems bei Sprache für jedes Band des SII bestimmt [3]. Bei der Berechnung des SII nach [1] wird der Hilfsfaktor K_i für jedes Band i des SII nach Gleichung 1 bestimmt durch

$$K_i = \frac{E'_i - D_i + 15}{30} \quad (1)$$

mit dem äquivalenten spektralen Leistungsdichtepiegel der Sprache E'_i und dem äquivalenten Störpektrum D_i . Der Faktor K_i wird auf den Bereich von 0 bis 1 limitiert. Hier wird der Dynamikbereich von Sprache auf 30 dB festgelegt. Mit dem effektiven Kompressionsverhältnis wurde nun der im SII angenommene Dynamikbereich von Sprache nach Gleichung 2 reduziert.

$$k_i = \frac{E'_i - D_i + 15 - CR_{\text{eff},i} \cdot 3}{30 - CR_{\text{eff},i} \cdot 6} \quad (2)$$

Außerdem wurde die Annahme getroffen, dass Kompression den maximal erreichbaren Wert des SII verringert. Dies wurde durch Gleichung 3 umgesetzt.

$$K_i = k_i - \frac{CR_{\text{eff},i} - 1}{28} \quad \text{mit } 0 \leq k_i \leq 1 \quad (3)$$

Der so optimierte SII führt in der Vorhersage der Sprachverständlichkeit bei komprimierter Sprache zu deutlich geringeren Abweichungen relativ zu unkomprimierter Sprache, wie in Abb. 4 zu sehen ist. Lediglich bei

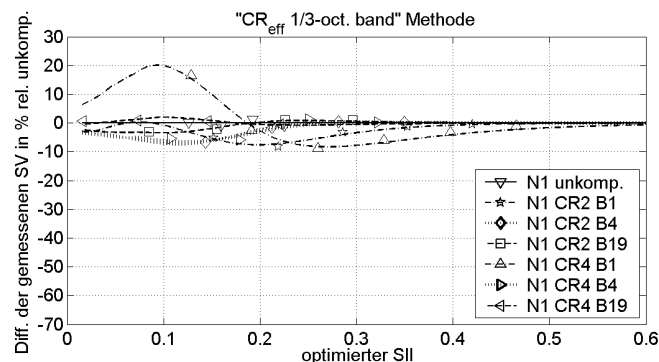


Abbildung 4: Differenz aus gemessener Sprachverständlichkeit und vorhergesagter Sprachverständlichkeit für unkomprimierte Sprache als Funktion des optimierten SII bei Störgeräusch N1.

breitbandiger Kompression mit $CR=4$ zeigt sich bei gleichem SII eine große Differenz in der gemessenen Sprachverständlichkeit relativ zu unkomprimierter Sprache. Hier treten durch die Kompression jedoch auch deutliche Veränderungen im Langzeitspektrum der Sprache auf und wie in Abb. 2 gezeigt wurde, ist die Vorhersage der Sprachverständlichkeit bei spektralen Änderungen ungenau.

Literatur

- [1] Methods for calculation of the Speech Intelligibility Index. American National Standards Institute, S3.5 1997
- [2] Wagener K, Kühnel V, Kollmeier B, „Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests“ Z. Audiol. 38(1), 4–15, 1999.
- [3] Elberling C, Hansen KV, „Hearing instruments – interaction with user preference“, In *Auditory Models and Non-linear Hearing Instruments, Proc. 18th Danavox Symposium*, 341–357, 1999.