

Modellierung der Sprachverständlichkeit in schwerhörenden Probanden

Christoph Scheidiger¹, Torsten Dau²

Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Kgs.

Lyngby ¹E-Mail: csche@elektro.dtu.dk

²E-Mail: tdau@elektro.dtu.dk

Einleitung

Menschen kommunizieren vorwiegend über Sprache. Schwerhörigkeit beeinträchtigt die Fähigkeit, sich mit anderen Menschen zu unterhalten. Trotz modernster Hörgerätetechnik kann die Sprachverständlichkeit (SV) in schwerhörenden Menschen zur Zeit nur teilweise wiederhergestellt werden. Ein wichtiger Schritt in Richtung besserer Hörgeräte sind Erkenntnisse darüber, wie das menschliche Hörsystem durch einen Hörverlust beeinträchtigt ist. Solche Kenntnisse können helfen, effizientere Kompensationsstrategien zu entwickeln. Im folgenden beschreiben wir, wie ein Sprachverständlichkeitsmodell (SVM) für Schwerhörige für diesen Zweck eingesetzt werden kann.

Sprachverständlichkeit (SV) ist stark vom Störgeräusch abhängig, in dem das Sprachsignal verstanden werden soll sowie vom Sprachsignal selbst. SV gibt an wie viele Spracheinheiten von dem präsentierten Sprachmaterial verstanden wurden. Spracheinheiten können in diesem Zusammenhang beispielsweise Wörter eines Satzes sein. Seit mehreren Jahrzehnten versuchen Forscher, SV vorherzusagen [1, 2, 3, 4]. Die Modelle werten das Sprach- und Hintergrundrauschsignal aus und kommen auf Grund der temporalen und spektralen Eigenschaften dieser Signale auf eine SV-Vorhersage.

Jørgensen und Dau haben das „speech-based envelope power spectrum“ Modell (sEPSM) vorgestellt, welches SV für normalhörende Personen in verschiedenen akustischen Konditionen vorhersagen kann [3, 4]. Das Modell basiert seine Voraussage auf dem Signal-Rausch-Verhältnis in der Domäne der Einhüllenden (SNR_{env}). Es wurde gezeigt, dass dieses Modell sowohl in stationärem als auch in fluktuierendem Störgeräusch SV präzise voraussagen kann. Aufgrund dieser Allgemeinheit des Modells wird es in der vorliegenden Arbeit als Grundlage für ein SVM für Schwerhörige verwendet.

Schwerhörigkeit und ihr Einfluss auf Sprachverständlichkeit

Ein Hörverlust kann durch eine Abschwächungs- und eine Verzerrungskomponente beschrieben werden [5]. Die Abschwächungskomponente wird durch eine reduzierte Sensitivität für Töne verschiedener Frequenzen im Vergleich zu Normalhörenden gemessen. Die Verzerrungskomponente wird beispielsweise durch eine verringerte Frequenz- oder Zeitauflösung beschrieben.

Die Abschwächungskomponente kann die Unterschiede der SV in Ruhe zwischen verschiedenen schwerhörenden

Probanden gut erklären. Im Störgeräusch hingegen wird oft die Verzerrungskomponente als entscheidend betrachtet [5, 6, 12, 20].

Normalhörende zeigen einen markanten Unterschied in der SV in zeitlich fluktuierendem Störgeräusch im Vergleich zu stationärem Rauschen [6]. In fluktuierendem Störgeräusch erreichen sie die gleiche SV wie im stationären Hintergrundrauschen bei ca. 6dB verringerter Rausch-Energie. Dieser Effekt wird als „Hören in den Lücken“ oder „Masking Release“ (MR) bezeichnet. Bei schwerhörenden Probanden ist dieser Effekt deutlich verringert [6, 10, 11].

Der Grund für den Verlust der Fähigkeit, Sprache in den zeitlichen Lücken des Störgeräusches wahrzunehmen, ist unklar. Es wurde argumentiert dass der reduzierte MR in Schwerhörenden auf die normalerweise verringerte Frequenzselektivität zurückzuführen ist. Eine Korrelation zwischen Frequenzselektivität und MR konnte jedoch nicht klar nachgewiesen werden [6, 10]. Hingegen wurde in verschiedenen Studien eine Korrelation zwischen SV in fluktuierendem Hintergrundrauschen und Nachverdeckung, also dem zeitlichen Auflösungsvermögen des Gehörs, gezeigt [13, 14, 15]. Eine andere Studie konnte den verringerten MR mit der Abschwächungskomponente des Hörverlustes erklären [16].

Mit einem SVM können die verschiedenen Hypothesen, die zu einem reduzierten MR bei schwerhörenden Probanden führen können, getestet und quantifiziert werden. In diesem Beitrag werden erste Simulationsergebnisse mit einem auf Schwerhörigkeit angepassten sEPSM vorgestellt.

Methoden

Die SV von 13 schwerhörenden Probanden (fünf männlich und acht weiblich) im Alter von 51 bis 72 Jahren [20] wurden mit einem modifizierten sEPSM simuliert. Die Simulationen wurden für stationäres Rauschen mit einem sprachähnlichen Spektrum (SSN) und einem 8-Hz sinusförmig modulierten Rauschen (SAM) durchgeführt [20].

Die Hörschwellen des sEPSM wurden dem individuellen Hörverlust entsprechend angepasst. Dies hat typischerweise zur Folge, dass im schwerhörenden Modell weniger Filterbänder verarbeitet werden als im normalhörenden Modell.

Resultate

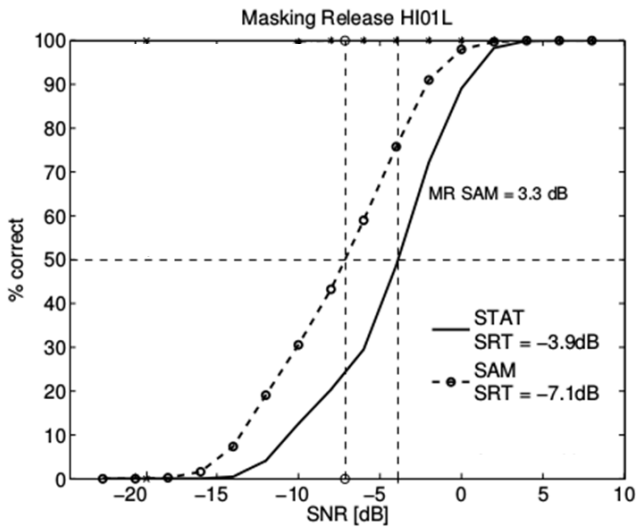


Abbildung 1: Simulation der Sprachverständlichkeit von der schwerhörenden Versuchsperson HI01L in zwei Hintergrundsrauschen: in stationärem sprachgewichtetem Rauschen (-) und in 8-Hz sinusförmig moduliertem Rauschen (--).

Abbildung 2 zeigt die SV Simulationsergebnisse für Versuchsperson HI01L. Für ein Signal-Rausch Verhältnis (S/R) im Bereich von -20 bis 10 dB wurde die SV mit beiden Hintergrundsrauschen simuliert. Die erhaltenen SV-Werte wurden mit einer logistischen Funktion angenähert. Dadurch konnte der S/R bestimmt werden, bei dem 50% des präsentierten Satzes verstanden wurden (auch bekannt als „speech reception threshold“, SRT). Die Differenz zwischen dem SRT bei moduliertem Hintergrundsrauschen und dem SRT bei stationärem Rauschen kennzeichnet den simulierten MR. Abbildung 1 zeigt die SV-Vorhersagen für die zwei Hintergrundsrauschen für Versuchsperson HI01L. Der simulierte MR betrug 3.3 dB simuliert.

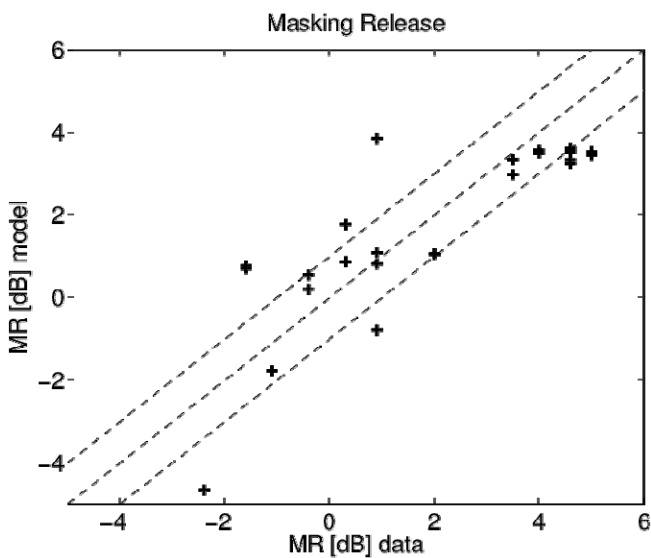


Abbildung 2: Simulierter MR (Ordinate) als Funktion des gemessenen MR (Abszisse) für beide Ohren der 13 schwerhörenden Probanden.

Um die MR-Simulationen von allen Probanden mit den gemessenen MR zu vergleichen, zeigt Abbildung 2 die Simulationsergebnisse als Funktion der gemessenen Werte. Wären die Simulationen alle korrekt, würden alle Punkte auf der Diagonalen liegen. Interessanterweise kann das Modell den MR erstaunlich gut voraussagen und nur fünf Simulationsergebnisse liegen mehr als ±1dB von den gemessenen Werten entfernt.

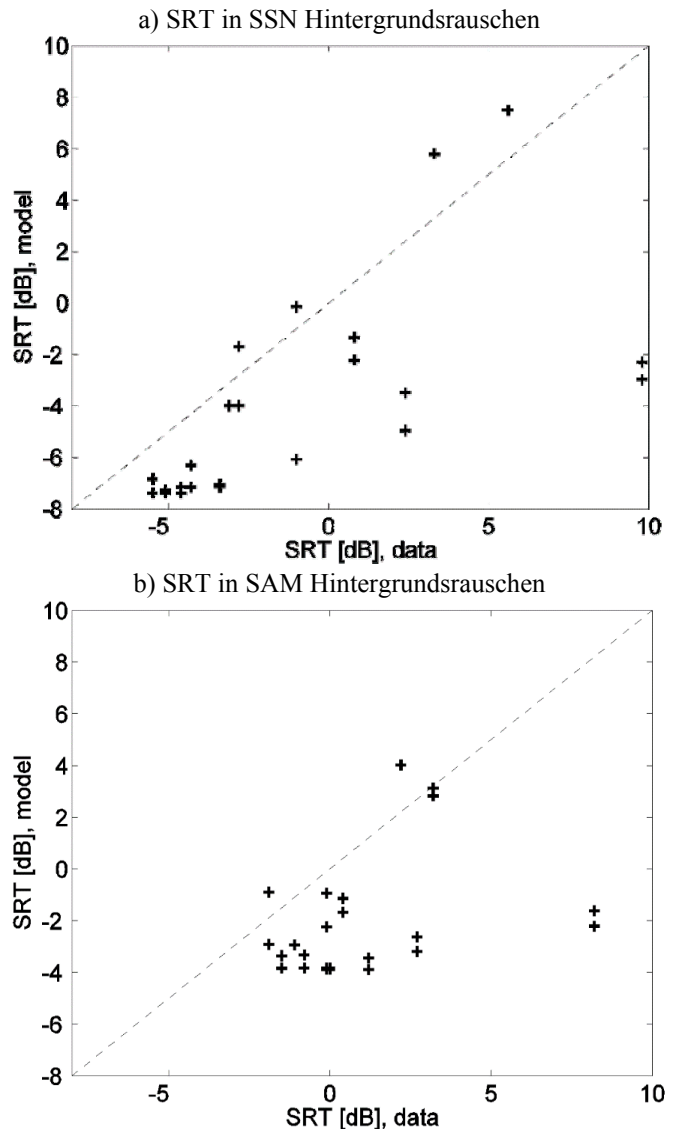


Abbildung 3: SRTs für die beiden Rausch-Typen. Die Simulationsergebnisse der stationären sowie der fluktuierenden SRTs sind tiefer (besser) als die gemessenen Werte.

Abbildung 3 zeigt die simulierten SRTs in beiden Hintergrundsrauschen (SSN oben, SAM unten). Hierbei wird deutlich, dass die SRTs in beiden Fällen als zu gut (zu tief) simuliert werden.

Diskussion

Obwohl die SRTs vom Modell in stationärem und in fluktuierendem Rauschen unterschätzt wurden, waren die MR Simulationen stark mit den Daten korreliert. Dies deutet darauf hin, dass die Abschwächungskomponente alleine gegebenenfalls den verringerten MR erklären könnte,

während eine oder mehrere Verzerrungskomponenten für die Verschiebung der beiden SRTs verantwortlich sein könnten.

Realistischere SRT-Vorhersagen könnten durch eine detailliertere Simulation des Hörverlustes erreicht werden. Zum Beispiel könnte die reduzierte Frequenzauflösung des geschädigten Gehörs simuliert werden.

Ein Schwachpunkt des Modelles könnte in der Integration von Information über verschiedene Frequenzen hinweg liegen. Im Modell werden die quadrierten SNR_{env} Werten in den einzelnen Kanälen aufsummiert, basierend auf der Annahme unabhängiger Beobachtungen in der Signaldetektionstheorie [17]. Allerdings sind die peripheren Filter des Modells überlappend, s.d. benachbarte Filter redundante Information enthalten. Redundante Information in den Frequenzkanälen wurde in Studien mit Nicht-Sprachsignalen zwar als vernachlässigbar bewertet [18] und scheint auch im normalhörenden Modell keinen negativen Effekt zu zeigen [3, 4]. Es wurde aber gezeigt, dass spektral weitentfernte Bänder zusammen mehr Informationen beinhalten als im Falle der Annahme unabhängiger Verarbeitung. Das bisherige Integrationsmodell zieht solche Synergieeffekte im Sprachsignal nicht in Betracht. Eine signal-basierte Gewichtung der spektralen Integration, die auf Korrelationen zwischen verschiedenen Frequenzbändern beruht, könnte das Integrationsmodell verbessern.

Zusammenfassung

In der Literatur wurden verschiedene Hypothesen diskutiert, warum der MR bei schwerhörenden Personen kleiner ist als bei normalhörenden. In der vorliegenden Studie wurde ein SVM verwendet, um den MR in schwerhörenden Probanden zu simulieren. Es wurde gezeigt, dass basierend auf der Abschwächungskomponente des Hörverlustes die unterschiedlichen MR in 13 Probanden erklärt werden konnten. Jedoch wurde auch gezeigt, dass das präsentierte Modell Schwierigkeiten bei der Voraussage von SRTs in stationärem und fluktuierendem Rauschen hat. Das auf Signaldetektionstheorie basierende Integrationsmodell wurde untersucht und die Einführung einer signal-basierten Gewichtung der einzelnen Kanäle wurde diskutiert.

Literatur

- [1] French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *The Journal of the Acoustical Society of America*, 19(1), 90–119.
- [2] Houtgast, T., Steeneken, H. J. M., & Plomp, R. (1980). Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I. General Room Acoustics. *Acta Acustica United with Acustica*, 46(1), 60–72.
- [3] Jørgensen, S., & Dau, T. (2011). Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing. *The Journal of the Acoustical Society of America*, 130(3), 1475–87. doi:10.1121/1.3621502
- [4] Jørgensen, S., Ewert, S. D., & Dau, T. (2013). A multi-resolution envelope-power based model for speech intelligibility. *The Journal of the Acoustical Society of America*, 134, 436.
- [5] Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, 63(2), 533–549.
- [6] Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, 88(4), 1725.
- [7] Jørgensen, S., Ewert, S. D., & Dau, T. (2013). A multi-resolution envelope-power based model for speech intelligibility. *The Journal of the Acoustical Society of America*, 134, 436.
- [8] Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, 63(2), 533–549.
- [10] Lorenzi, C., Husson, M., Ardoit, M., & Debrulle, X. (2006). Speech masking release in listeners with flat hearing loss: Effects of masker fluctuation rate on identification scores and phonetic feature reception. *International Journal of Audiology*, 45(9), 487–495.
- [11] Bernstein, J. G. W., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 125(5), 3358–3372.
- [12] Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *The Journal of the Acoustical Society of America*, 125(5), 3328–3345.
- [13] Hou, Z., & Pavlovic, C. V. (1994). Effects of temporal smearing on temporal resolution, frequency selectivity, and speech intelligibility. *The Journal of the Acoustical Society of America*, 96(3), 1325–1340.
- [14] Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2003). Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *The Journal of the Acoustical Society of America*, 113(4), 2084–2094.
- [15] George, E. L. J., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 120(4), 2295–2311.
- [16] Desloge, J. G., Reed, C. M., Braid, L. D., Perez, Z. D., & Delhorne, L. A. (2010). Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise. *The Journal of the Acoustical Society of America*, 128(1), 342–59.

- [17] Green, D. M., Swets, J. A., & others. (1966). Signal detection theory and psychophysics (Vol. 1). Wiley New York.
- [18] Ewert, S. D., & Dau, T. (2000). Characterizing frequency selectivity for envelope fluctuations. *The Journal of the Acoustical Society of America*, 108(3 Pt 1), 1181–96.
- [19] Warren, R. M., Bashford Jr, J. A., & Lenz, P. W. (2005). Intelligibilities of 1-octave rectangular bands spanning the speech spectrum when heard separately and paired. *The Journal of the Acoustical Society of America*, 118(5), 3261–3266.
- [20] Christiansen, C., & Dau, T. (2012). Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise. *The Journal of the Acoustical Society of America*, 132(3), 1655–1666.