

Modellierung der Sprachperzeption bei Cochlea-Implantat-Nutzern mit rein elektrischer und elektrisch-akustischer Stimulation (EAS)

Tobias Rader^{1,2}, Youssef Adel², Uwe Baumann¹ und Hugo Fastl²

¹ *Audiologische Akustik, Klinik für HNO-Heilkunde, Goethe-Universität Frankfurt am Main
e-Mail: tobias.rader@kgu.de*

² *AG Technische Akustik, MMK, TU München*

Einleitung

Schwerhörige Patienten mit einem Hochtonsteilabfall können durch die Kombination von Hörgerät und Cochlea-Implantat (CI) am selben Ohr wieder ein gutes Sprachverstehen erlangen. Die simultane Erregung des Hörnervs mit einem CI für Frequenzen oberhalb von etwa 500 Hz und andererseits der Nutzung des akustischen Restgehörs unterhalb von etwa 500 Hz am selben Ohr mit einem Hörgerät wird als elektrisch-akustische Stimulation (EAS) bezeichnet. Patienten mit einer EAS-Versorgung zeigen ein signifikant besseres Sprachverstehen im Störgeräusch (Rader et al. 2009) gegenüber konventionell versorgten CI-Patienten. Dieser Effekt wird an einem Modell nachvollzogen und mit den Daten von EAS-versorgten Patienten verglichen.

Material und Methoden

Das EAS-Modell basiert auf dem Sprachmaterial des Oldenburger Satztests (OLSA, Wagener et al. 1999). Die Sätze wurden spektral transformiert und durch Signalverarbeitung in eine CI/EAS-Simulation übergeführt (Rader et al. 2010). Zur Simulation der über die Elektroden übertragenen Sprachanteile wurde das Sprachsignal oberhalb der Übergabefrequenz entsprechend der zwölf Mittenfrequenzen des Med-El DUET Sprachprozessors bandpassgefiltert. Der Übertragungsbereich der „elektrischen“ CI-Stimulation beträgt 500 Hz bis 8500 Hz.

Die Resynthese des Sprachmaterials erfolgte mit einem 12-Band-Sinusvocoder, bei dem die Vocoder-Frequenzen mit den 12 Mittenfrequenzen der CI-Bandpasskanäle übereinstimmen. Die Störgeräusche (Fastl-noise und OL-noise) wurden gleichermaßen in einer CI-Simulation übergeführt.

Die Berechnung der Schalle erfolgte in der Teilzeitmusterdomäne nach Heinbach (1988). Durch Variation der Grenzfrequenz des tiefpass-gefilterten Anteils wird das Signal von einer reinen CI-Simulation hin zu einer EAS-Simulation mit variablen akustischen tieffrequentem Restgehör variiert. Dies entspricht dem in gewissen Grenzen individuell variierenden tieffrequentem Restgehör von EAS-Trägern.

Um den Einfluss des Tiefpass-Anteils eines EAS-Systems auf die Sprachdiskrimination zu untersuchen, wurde ein auf dem Hidden-Markov Modell basierendes, automatisiertes Spracherkennungssystem (HTK-Toolkit, Young et al. 2006) implementiert. Mit Hilfe einer Beschreibungssprache wurde die Grammatik der OLSA-Sätze beschrieben und das Spracherkennungssystem mit den unverarbeiteten OLSA-Sätzen trainiert. Der Frequenzbereich des Spracherkennungssystems wurde auf 8 kHz - entsprechend dem MED-EL DUET-Sprachprozessor - beschränkt.

Als Störgeräusche kamen das quasi unmodulierte OL-noise (30 verschobene und überlagerte Sätze des OLSA) und das sprachmodulierte Fastl-noise (mit stochastisch verteilten Pausen) zum Einsatz. Beide Rauschsignale beinhalten keine verwertbaren Sprachinformationen (d.h. kein „informational masking“).

Die Sprachverständlichkeitsschwelle (SVS), also der SNR-Wert, bei dem 50 % des Satzes korrekt erkannt wird, wurde mit Hilfe des Oldenburger Satztestes in Verbindung mit dem HTK ermittelt. Die variable Steuerung des Störgeräusches erfolgte adaptiv entsprechend der aktuellen Erkennungsrate. Die Sprachsignalwiedergabe wurde mit einem konstanten Pegel präsentiert.

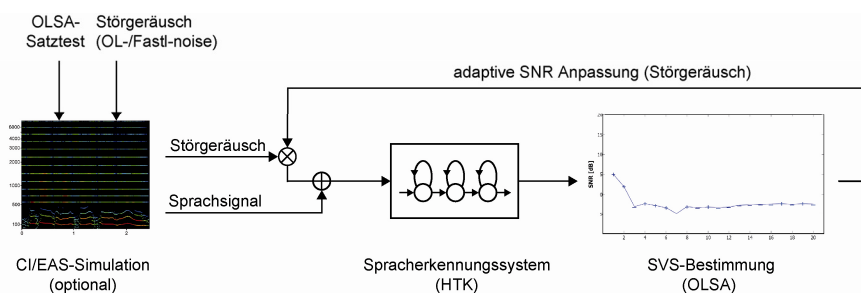


Abbildung 1: Modell zum Sprachverstehen mit Hörimplantaten im Störgeräusch. Je nach Erkennungsrate des Spracherkennungssystems wird das Störgeräusch adaptiv angepasst, um die Sprachverständlichkeitsschwelle (SVS: Signal-Rausch-Abstand bei 50 % korrekter Erkennungsrate) zu ermitteln.

Ergebnisse

Abbildung 2 zeigt die Erkennungsraten des Spracherkennungssystems für rein elektrische Simulation (CI), EAS mit verschiedenen tieffrequenten Anteilen und ohne CI/EAS-Simulation (unbearbeiteten Schalldateien) im Störgeräusch bei 0 dB SNR. Während bei einer reinen CI-Simulation und bei einer EAS-Simulation mit OL-noise die Erkennungsleistungen sehr gering sind, steigen sie für zunehmenden tieffrequentem Anteil im Fastl-noise stetig an, was durch die erhöhte Erkennungsleistung in den Pausen des fluktuierenden Fastl-noise erklärbar ist. Ohne CI/EAS-Simulation erreicht das Spracherkennungssystem 91.8 % im Fastl-noise.

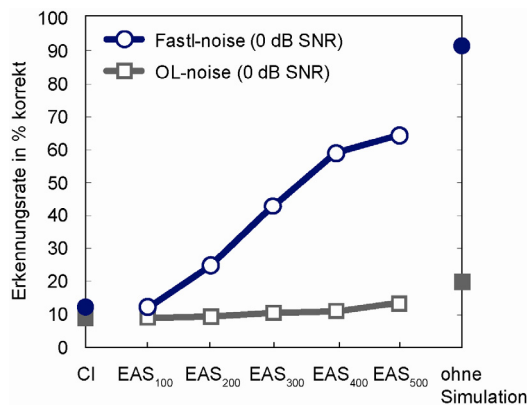


Abbildung 2: Erkennungsraten für Fastl- und OL-noise bei 0 dB SNR mit folgenden Modellkonditionen: CI (rein elektrische Simulation), EAS mit zunehmenden akustischen tieffrequenten Anteil von 100 bis 500 Hz und ohne CI/EAS-Simulation.

Abbildung 3 zeigt die SVS des EAS-Modells in OL- und Fastl-noise (farbige Symbole) im Vergleich zu den Boxplots (Median, Interquartile und Maximalwerte) von realen EAS-Trägern (n=12) und einer normalhörenden Kontrollgruppe (n=20).

Während im quasi stationären OL-noise die reale Diskriminationsleistung der EAS-Patienten besser ist als der im Modell vorhergesagte Wert, passen Modell- und Realdaten für das sprachmodulierte Fastl-noise gut überein: Der Medianwert der EAS-Patienten (1.6 dB SNR) liegt fast identisch zu dem mithilfe des Modells ermittelten Werte von 1.4 dB SNR bei 300 Hz Tiefpassfrequenzanteil.

Zusammenfassung

Das Modell zeigt einen Ansatz zur Abschätzung des möglichen Sprachverstehens in Abhängigkeit vom nutzbaren tieffrequenten Restgehör bei Patienten mit einer EAS-Versorgung. Die mit dem Spracherkenner-Modell gewonnenen Ergebnisse verdeutlichen den größten Nutzen tieffrequenter akustischer Stimulation bei sprachmoduliertem Störgeräusch ab einer Tiefpassfilter-frequenz größer als 300 Hz. Eine gute Übereinstimmung zwischen Modell und gemessenen Patientendaten zeigt sich für das sprachmodulierte Fastl-noise.

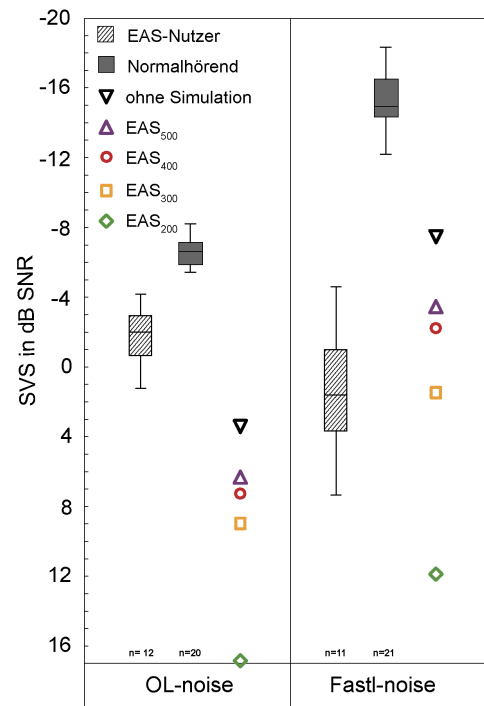


Abbildung 3: Sprachverständlichkeitsschwellen (SVS: Boxplot mit Median, Quartilen und Maximalwerten) für die EAS-Nutzer, Normalhörende und zugehörige Modelldaten im OL-noise und Fastl-noise. Die Modelldaten beinhalten akustischen tieffrequenten Anteil von 100 bis 500 Hz sowie eine Messreihe ohne CI/EAS-Simulation.

Danksagung

Die Autoren danken Felix Weninger und Martin Wöllmer vom Lehrstuhl für Mensch-Maschine-Kommunikation der TU-München für die Bereitstellung der Sprachkorpora und die Unterstützung bei der Adaption des HTK-Toolkit.

Literatur

- Rader T, Baumann U and Fastl H. Comparison of Speech Intelligibility by EAS, Bimodal, Uni- and Bilateral Cochlear Implant Patients in a 'Multi-Source Noise Field' (MSNF). In: M. Boone, editor, Proc. Intern. Conf. on Acoustics (NAG/DAGA 2009), Rotterdam, The Netherlands, pp. 1288–1290, 2009.
- Rader T, Baumann U, und Fastl H.: Simulationen zum Sprachverstehen im Störgeräusch bei elektrisch-akustischer Stimulation (EAS). In: Tagungsband der 13. Jahrestagung der Deutschen Gesellschaft für Audiologie, Frankfurt. DGA, 2010.
- Heinbach W. Aurally adequate signal representation: The Part-Tone-Time-Pattern. *Acustica* 67, 113-121, 1988.
- S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, P. Woodland: *The HTK Book (for HTK version 3.4)*, Cambridge University Engineering Department, 2006.
- Wagener K, Brand T, Kuehnel V, Kollmeier B. Entwicklung und Evaluation eines Satztests in deutscher Sprache I – III: Design, Optimierung und Evaluation des Oldenburger Satztests. *Z Audiol*; 38 (1-3): 4 – 15, 44 – 56, 86 – 95, 1999.