Beeinflusst Nachhall die bimodalen oder audiovisuellen Sprachverständlichkeitsgewinne in Cochlea-Implantat-Trägern?

Peter Ihly[1], Hartmut Meister[2], Daniela Hollfelder[3], Karl-Ludwig-Bruchhage[3], Tim Jürgens[1]

[1]Institut für Akustik, Technische Hochschule Lübeck, peter.ihly@th-luebeck.de

[2]Jean-Uhrmacher-Institut für klinische HNO-Forschung, Universität zu Köln

[3]Klinik für HNO-Heilkunde, Universitätsklinikum Schleswig-Holstein, Lübeck

Schlüsselwörter: Cochlea-Implantat, bimodales Hören, audiovisuelle Integration

Zusammenfassung

Cochlea-Implantat (CI)-Träger mit Zugang zu kontralateral akustischem Hören weisen in störgeräuschvollen Situationen ein besseres Sprachverstehen auf, wenn beide Ohren am Hörvorgang beteiligt sind (bimodaler Sprachverständlichkeitsgewinn, SVG). Zusätzlich können visuelle, synchron dargebotene Artikulationsbewegungen des Sprechers die Sprachverständlichkeit verbessern (audiovisueller SVG). Diese Studie untersucht den Einfluss von Nachhall sowohl auf bimodale als auch auf audiovisuelle SVG in akustisch schwierigen Situationen, sowie deren Interaktion. Es zeigten sich signifikante audiovisuelle SVG von durchschnittlich 13,4% (± 5,1 %) in CI-Trägern. Bimodale SVG betrugen durchschnittlich 31,6% (± 2,6 %) durch zusätzliches akustisches Hören und 31,5% (± 3,2 %) durch zusätzliches elektrisches Hören. Im Vergleich zu verhallten Konditionen zeigten die CI-Träger in unverhallten Konditionen ein um durchschnittlich 26,9% (± 1,5 %) besseres Sprachverstehen. Eine signifikante Interaktion konnte zwischen den Faktoren Nachhall und visueller Stimulus, jedoch nicht zwischen den Faktoren Nachhall und Modalität festgestellt werden, was auf einen größeren audiovisuellen SVG in unverhallten Konditionen hindeutet.

1. Einleitung

Während der COVID-19 Pandemie gehörte das Tragen von medizinischen Masken zum Alltag vieler Menschen. Da diese für einen ausreichenden Schutz Mund und Nase verdecken müssen, verhindern solche Masken folglich den Zugang zum Lippenbild und den Artikulationsbewegungen des Sprechers. Zusätzlich zur gesprochenen Sprache dargebotene visuelle Reize in Form von Artikulationsbewegungen können allerdings das Sprachverstehen für Normalhörende und für CI-Träger deutlich verbessern [1], [2]. Da für CI-Träger mit kontralateralem akustischem Hören zusätzlich noch in einigen Situationen ein bimodaler SVG [3],[4] auftritt, stellt sich die Frage, inwieweit diese einzelnen Gewinne miteinander interagieren und inwieweit diese Gewinne von der akustischen Situation abhängen, also anders in nachhallbehafteten oder trockenen Räumen sind.

2. Material und Methoden

2.1. Probanden

Sieben postlingual ertaubte CI-Träger (Alter 27-70 Jahre, Ø 55 Jahre, vier weiblich, drei männlich) nahmen an der Studie teil. Fünf Probanden trugen zusätzlich zum CI ein Hörgerät auf dem kontralateralen Ohr (bimodale Probanden), der Hörverlust des kontralateralen Ohres reichte von leicht bis hochgradig. Zwei Probanden waren auf dem kontralateralen Ohr normalhörend (single-sided deafness, SSD). Als Kontrollgruppe nahmen sechs normalhörende Probanden (Alter 24-41 Jahre, Ø 31 Jahre, zwei weiblich, vier männlich) teil. Alle Probanden waren Deutsch-Muttersprachler und wiesen keine unausgeglichene Sehschwäche auf.

2.2. Stimuli

Als Sprachmaterial dienten in dieser Studie die Sätze des Hochmair-Schulz-Moser (HSM) Tests [5]. Mit Hilfe des MASSY (articulation model for audiovisual speech synthesis) wurde für jeden Satz der 15 HSM-Testlisten ein virtueller, computergenerierter Sprecher erzeugt als Darstellung für den Bildschirm erzeugt [2]. Die visuellen Reize wurden zeitsynchron (Genauigkeit < 5ms) zu den HSM-Sätzen dargeboten. Als Störgeräusch wurde simultan zum HSM-Satztest das CCITT-Rauschen (Comité Consultatif International Téléphonique et Télégraphique) in allen Konditionen verwendet. In der Hälfte der Konditionen wurden die akustischen Stimuli (Störgeräusch wie auch HSM-Sätze) verhallt durch Faltung mit einer Raumimpulsantwort mit einer Nachhallzeit von 1 s. Die Raumimpulsantwort wurde mit Hilfe des MatLab-Skripts "RIR-Generator" [6] berechnet, wobei eine Raumgröße von 5x5x2,5m mit schallharten Wänden angenommen wurde. Der Signal-zu-Rausch-Abstand (SNR) wurde dabei für jeden Probanden individuell so eingestellt, dass 50% der Wörter in der verhallten, bimodalen Situation ohne virtuellen Sprecher (Bimodal-verhallt-A) verstanden wurden.

2.3. Messaufbau und Kalibrierung

Alle Messungen wurden in einer doppelwandig schallisolierten Hörkabine durchgeführt. Akustische Reize wurden über einen Lautsprecher (Genelec 2080A), visuelle Reize über einen 24″ Monitor in einem Abstand von 100 cm frontal dargeboten. Die D/A Wandlung wurde mittels RME Fireface UC durchgeführt. Mittel seines Pegelmessers (B&K Typ 2250) wurden die akustischen Signale auf 65 dB(A) kalibriert.

2.4. Messkonditionen

Für das Experiment wurden insgesamt drei Faktoren variiert. Der Faktor Modalität mit den Ausprägungen bimodal und monaural (entweder mit CI, mit Hörgerät/normalhörendem Ohr oder bimodal), der Faktor Nachhall mit den Ausprägungen unverhallt und verhallt (Nachhall von 1 s) und der Faktor visueller Stimulus mit den Ausprägungen rein akustisch (kein virtueller Sprecher) und audiovisuell (mit virtuellem Sprecher). Für die CI-Träger bestand der Sprachverständlichkeitstest dieses Experiments somit aus insgesamt 12 Konditionen mit jeweils acht monauralen und vier bimodalen Konditionen. Die normalhörende Kontrollgruppe durchlief den gleichen Messablauf, jedoch bestand der Faktor Modalität hier nicht aus drei, sondern aus zwei Leveln (monaural und binaural).

In monauralen Konditionen, in denen die CI-Träger angewiesen wurden, nur ihr CI zu benutzen, wurde das kontralaterale Ohr mittels eines In-Ear-Kopfhörers mit CCITT-Rauschen bei einem Pegel von 65 dB(A) vertäubt. Bei der normalhörenden Kontrollgruppe wurde in monauralen Konditionen immer das linke Ohr mit CCITT-Rauschen über In-Ear-Kopfhörer vertäubt und rechts gemessen.

2.5. Statistische Analyse

Zur Umwandlung der begrenzten Prozentwerte in eine offene Skala, wurden alle gemessenen Werte in die Einheit "rationalized arcsine units" (RAU) transformiert [7]. Anschließend wurden die Ergebnisse mit Hilfe des Shapiro-Wilk-Tests auf Normalverteilung, und mittels ANOVA mit Messwiederholung (RM-ANOVA) auf signifikante Unterschiede zwischen den einzelnen Faktoren untersucht ($p < 0,05$).

3. Ergebnisse

Abbildung 1 zeigt die über alle CI-Träger (links) und normalhörenden (rechts) gemittelte Sprachverständlichkeit (in Prozent) für alle Konditionen. RAU-Werte aller Konditionen der CI-Träger außer Elektrisch-unverhallt-AV ($p = 0,042$) zeigten sich normalverteilt. Die RM-ANOVA lieferte signifikante Einflüsse der Faktoren visueller Stimulus ($p = 0,003$, $F = 23,09$) und Nachhall ($p < 0,001$, $F = 127,49$) auf das Sprachverstehen. Zwischen den Faktoren visueller Stimulus und Nachhall konnte eine signifikante Interaktion nachgewiesen werden ($p = 0,017$, $F = 10,63$), zwischen den Faktoren Modalität und visueller Stimulus ($p = 0,260$, $F = 1,51$) sowie den Faktoren Modalität und Nachhall ($p = 0,892$, $F = 0,12$) konnte keine signifikante Interaktion nachgewiesen werden. Die Daten der

normalhörenden Probanden waren in allen Konditionen normalverteilt. Es konnte keine signifikante Interaktion zwischen den beiden Faktoren Modalität und visueller Stimulus ($p = 0{,}260$, $F = 1{,}51$), sowie den Faktoren Modalität und Nachhall ($p = 0{,}892$, $F = 0{,}12$) nachgewiesen werden. Zwischen den Faktoren visueller Stimulus und Nachhall zeigte sich dagegen eine signifikante Interaktion ($p = 0{,}017$, $F = 10{,}63$).

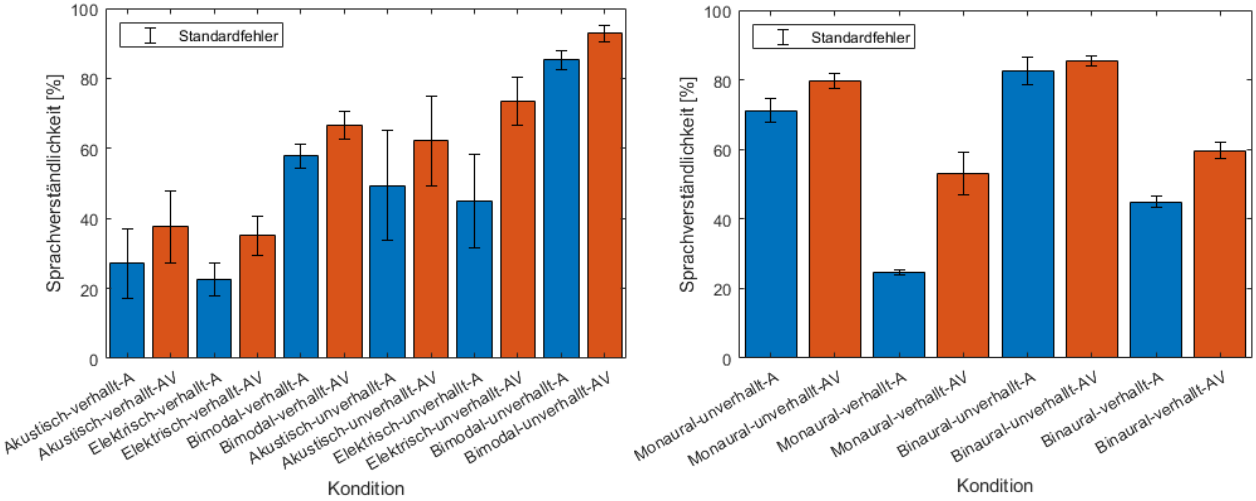

Abbildung 1: Mittlere Sprachverständlichkeit und Standardfehler der CI-Träger (links) und der normalhörenden Probanden (rechts) für alle Konditionen (A: akustisch, AV: audiovisuell)

3.1 Audiovisueller Sprachverständlichkeitsgewinn

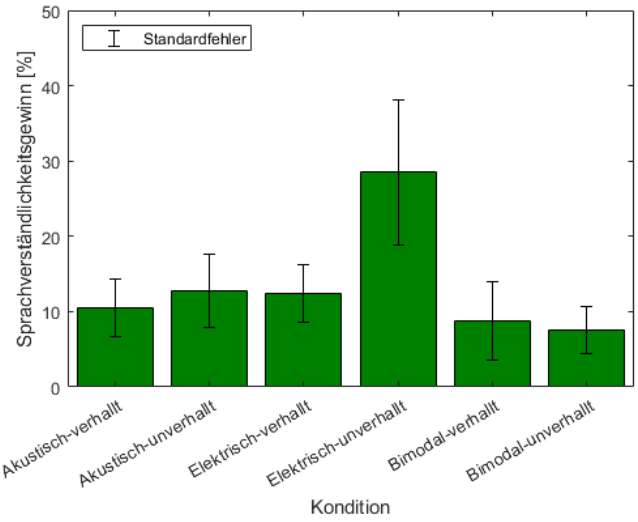

Abbildung 2 zeigt den mittleren audiovisuellen SVG aller CI-Träger gegenüber den rein akustischen Konditionen. Dieser ist in allen Konditionen deutlich positiv und beträgt im Mittel 13,4% (± 5,1 %). Gepaarte T-Tests zwischen den rein auditiven und audiovisuellen Konditionen zeigten signifikante audiovisuelle SVG in den Konditionen Akustisch-verhallt, Akustisch-unverhallt und Elektrisch-verhallt (alle $p < 0{,}05$).

Abbildung 2: Mittlerer audiovisueller SVG der CI-Träger für alle Konditionen

3.1 Bimodaler Sprachverständlichkeitsgewinn

Abbildung 3 zeigt den mittleren SVG durch zusätzliches akustisches Hören (links) und elektrisches Hören (rechts) aller CI-Träger für alle Konditionen. Es sind in allen Konditionen (links) akustische Sprachverständlichkeitsgewinne zu erkennen, mit durchschnittlich 31,6 % (± 2,6 %) über alle Konditionen. Gepaarte T-Tests zeigten signifikante akustische SVG in den Konditionen Verhallt-A, Verhallt-AV und Unverhallt-A (alle $p < 0{,}05$). In allen Konditionen (rechts) sind elektrische SVG zu erkennen, mit durchschnittlich 31,5 % (± 3,2 %) über alle Konditionen. gepaarte T-Tests zeigten signifikante elektrische SVG in den Konditionen Verhallt-A und Unverhallt-AV (alle $p < 0{,}05$).

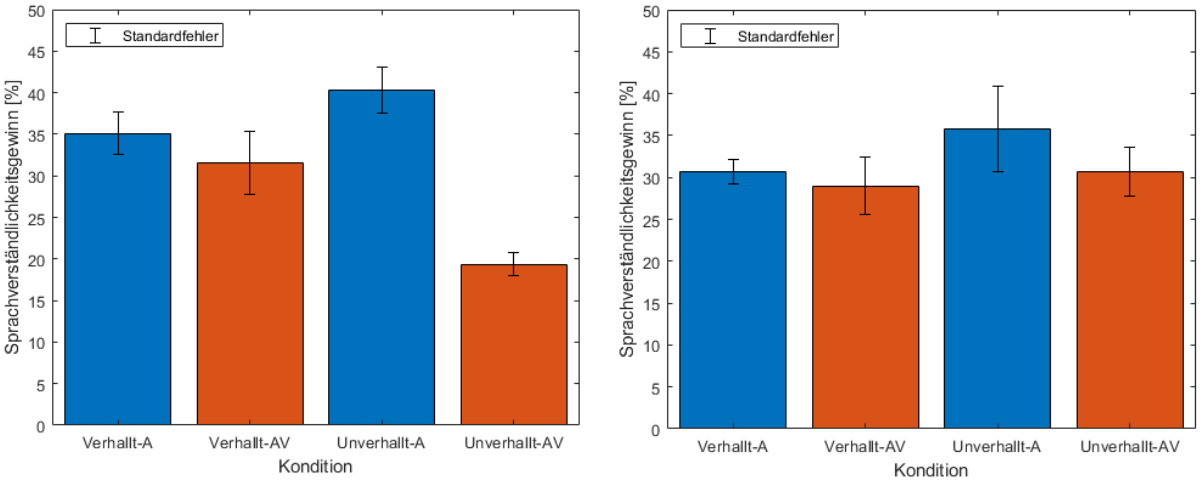

Abbildung 3: Mittlerer Sprachverständlichkeitsgewinn der CI-Träger für alle Konditionen durch zusätzliches akustisches Hören (links) und zusätzliches elektrisches Hören (rechts)

4. Diskussion

Die gemessenen audiovisuellen SVG bestätigen die Arbeiten von [1] und [2], welche ebenfalls signifikante audiovisuelle SVG unter Verwendung eines computergenerierten virtuellen Sprechers nachweisen konnten. Ein Vergleich zwischen dem durchschnittlichen audiovisuellen SVG der CI-Träger (13,4 % ± 5,1 %) mit dem der normalhörenden Kontrollgruppe (13,6 % ± 3,6 %) zeigt, dass beide Gruppen in akustisch schwierigen Situationen von zusätzlichen visuellen Reizen profitieren. Da der individuell verwendete SNR bei den CI-Trägern jedoch deutlich höher lag ist davon auszugehen, dass diese im Alltag eher auf die Unterstützung durch visuelle Reize angewiesen sind.

Zwischen den Faktoren Modalität und visueller Stimulus konnte keine Interaktion festgestellt werden, was darauf hindeutet, dass audiovisuelle SVG unabhängig von der Modalität (akustisch, elektrisch oder bimodal) auftreten. Ebenfalls konnte keine signifikante Interaktion zwischen den Faktoren Modalität und Nachhall nachgewiesen werden. Dieses Ergebnis legt nahe, dass der festgestellte negative Einfluss von Nachhall auf das Sprachverstehen sowohl bei bimodalem Hören wie auch beim Hören mit einem Hörgerät oder einem CI allein auftritt. Eine signifikante Interaktion konnte dagegen zwischen den Faktoren visueller Stimulus und Nachhall festgestellt werden. Diese Ergebnisse deuten darauf hin, dass Nachhall keinen Einfluss auf den bimodalen SVG, wohl aber auf den audiovisuellen SVG hat.

5. Literatur

[1] Desai S., Stickney, G., Zeng, F.-G. “Auditory-visual speech perception in normal-hearing and cochlear-implant listeners”, J. Acoust. Soc. Am., 123, pp. 428-440, 2008.

[2] Schreitmüller, S., Frenken, M., Bentz, L., Ortmann, M., Wagner, M., Meister, H. “Validating a method to assess lipreading, audiovisual gain and integration during speech reception with cochlear-implanted and normal-hearing subjects using a talking head”, Ear and hearing, 39, pp. 503–516, 2017.

[3] Hoppe, U., Hocke, T., Digeser, F., „Bimodal benefit for cochlear implant listeners with different grades of hearing loss in the opposite ear“, Acta Oto-laryngologica 138, 713-721, 2018.

[4] Williges, B., Wesarg, T., Jung, L., Geven, L., Radeloff, A., Jürgens, T. “Spatial Speech-in-Noise Performance in Bimodal and Single-Sided Deaf Cochlear Implant Users”, Trends in Hearing, 23, pp. 1–16, 2019.

[5] Hochmair-Desoyer, I., Schulz, E., Moser, L., Schmidt, M. “The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users”, *Am. J. Otol.*, 6, p. 83, 1997.

[6] Habets, E.A.P. “Room Impulse Response Generator”, Internal Report, pp. 1-17., 2006.

[7] Studebaker, G.A. “A “rationalized” arcsine transform”, *J. Speech Hear. Res.*, 28, pp. 455–462, 1985.