

Musik-Mixing Präferenz bei bimodalen CI-Trägern

Leonard Imhäuser^{1,2}, Tim Jürgens¹, Karl-Ludwig Bruchhage², Daniela Hollfelder²

¹ Institut für Akustik, Technische Hochschule Lübeck, Mönkhofer Weg 239, 23562 Lübeck

² Sektion für Hals- Nasen- und Ohrenheilkunde, Universitätsklinikum Schleswig-Holstein, Ratzeburger Allee 160, 23562 Lübeck

Schlüsselwörter: Cochlea-Implantat, Bimodal, Musik.

Zusammenfassung:

Aktuelle Forschungen legen nahe, dass Musikstücke mit betontem Gesang von Cochlea-Implantat (CI)-Trägern bevorzugt werden. In der vorliegenden Studie wird untersucht, inwieweit sich die Präferenz von betontem Gesang auch bei bimodalen CI-Trägern wiederfindet und welche Rolle hierbei das kontralaterale Ohr spielt.

Um diese Frage zu beantworten, wurden zwei Experimente durchgeführt. Bei Experiment 1 wurden die Probanden gebeten mittels digitaler Schieberegler das bevorzugte Lautstärkeverhältnis zwischen einer Gesangsspur und einer Instrumentalspur einzustellen. In Experiment 2 wurde untersucht, ob vorverarbeitete Musik (mit hervorgehobenem Gesang) in einem paarweisen Vergleich gegenüber der Originalversion bevorzugt wurde. Beide Experimente wurden von 8 bimodalen CI-Trägern und 15 normalhörenden Probanden durchgeführt. Die bimodale CI-Gruppe wurde über Freifeld-Lautsprecher und über direktes Streaming in das CI getestet, um die Rolle des akustisch hörenden kontralateralen Ohres untersuchen zu können.

In der direkten Streaming-Bedingung von Experiment 1 stellte die bimodale CI-Gruppe die Gesangsspur signifikant höher im Pegel ein als die normalhörende Kontrollgruppe (bis zu 2,3 dB). In der Freifeld-Bedingung von Experiment 1 wurde die Gesangsspur ebenfalls höher eingepgelt (bis zu 2,6 dB), jedoch waren die Ergebnisse stark von dem präsentierten Audiomaterial abhängig. Experiment 2 zeigte stets einen signifikanten Unterschied in der angegebenen Präferenz zwischen der bimodalen CI-Gruppe und der NH-Kontrollgruppe. Die normalhörende Kontrollgruppe bevorzugte stark die Originalversionen der dargebotenen Musikstücke (bis zu 95 % Präferenz für die Originalversion). Die Präferenz der CI Gruppe orientiert sich bei der Freifeld Darbietung ebenfalls in Richtung Original bzw. unterscheidet sich über direktes Streaming in das CI nicht von der Ratewahrscheinlichkeit.

1. Einleitung:

Das Cochlea-Implantat ist die erfolgreichste sensorische Neuroprothese, die es ermöglicht, sehr hochgradige Hörminderungen bis hin zur Taubheit zu versorgen. Während die Rehabilitation des Sprachverstehens häufig mit Erfolg gekrönt ist, bleibt die Wahrnehmung und der Genuss von Musik häufig mangelhaft. „Ich habe Sprache zurück aber nicht die Musik“ wird häufig von CI-Trägern berichtet (Drennan & Rubinstein, 2008). In Kang et al. (2009) wurden verschiedene Wahrnehmungsebenen von Musik beleuchtet. Diese beinhalteten unter anderem Tonhöhen- und Klangfarbendiskrimination.

Die Ergebnisse zeigen für CI-Träger stark erhöhte Diskriminationsschwellen und liefern damit Anhaltspunkte für erschwerte Musikwahrnehmung bei CI-Trägern. Studien von Buyens et al. (2014) und Pons et al. (2016) zeigten, dass sich CI-Träger die Gesangsspur eines Musikstücks im Vergleich zum Instrumental höher einpegeln. Zusätzlich wurde in der Studie von Buyens die Präferenz von betontem Gesang mit einem Forced-Choice Experiment geprüft, bei welchem zwei Versionen des gleichen Musikstücks zur Auswahl standen.

Die folgende Studie bezieht sich auf die Pilotstudie von Buyens et al. und beinhaltet ebenfalls zwei Experimente. Das Erste ist eine Mixing-Aufgabe (das Abmischen von 2 Spuren) und das Zweite erfasst die Präferenz von paarweise dargebotenen Musikstücken mit unterschiedlicher Abschwächung des Instrumentals. Zusätzlich untersucht die vorliegende Studie den Unterschied zwischen Wahrnehmung bei bimodalem Hören (Hörgerät und CI) und monauralem Hören (direkt in das CI gestreamt).

Es ergeben sich folgende Fragestellungen:

Fragestellung 1: Wird die Gesangspur von bimodalen CI-Trägern im Vergleich zur normalhörenden Kontrollgruppe mit höherem Pegel abgemischt bzw. Musikstücke mit betonter Gesangspur häufiger präferiert?

Fragestellung 2: Gibt es einen signifikanten Unterschied zwischen der Musikdarbietung im Freifeld und Direktstreaming über das Cochlea-Implantat? Welche Rolle spielt das kontralaterale Ohr?

1.1. Material und Methodik

An der Studie nahmen 8 bimodale, postlingual ertaubte CI-Träger (im Alter zwischen 35 und 70 Jahren) teil. Alle Probanden der CI Gruppe trugen Prozessoren der Firmen MED-EL oder Cochlear. Kontralateral hatten die Probanden eine leicht- bis hochgradige mit Hörgerät versorgte Schwerhörigkeit. Die normalhörende Kontrollgruppe beinhaltete 15 Probanden.

Um die Audiospuren für beide Experimente der Studie zu isolieren (Experiment 1) bzw. abzuschwächen (Experiment 2), wurde ein Tool namens RX7 von der Firma Izotope benutzt. Dieses macht es möglich aus einer einzigen Gesamspur verschiedene Elemente wie Gesang, Bass, Percussion zu extrahieren. Die Signalverarbeitung ist dabei sehr artefaktarm. Auch die in der Software (RX7) eingestellten Pegelangaben haben sich bei der Validierung mittels Perzentilanalyse bestätigt (Imhäuser, 2021). Eine ähnliche Spurentrennung durch Signalverarbeitungsalgorithmen wurde auch in der Studie von Pons et al. (2016) angewendet.

Experiment 1 untersucht, wie der Proband sein Mischungsverhältnis zwischen Gesangspur und Instrumental einstellt. Dafür wurde eine GUI mit Schieberegler in MATLAB programmiert. Je nachdem, wie die jeweilige Spur eingepegelt wurde, ergibt sich ein Verhältnis zwischen beiden Pegeln ähnlich wie bei der Berechnung eines Signal-Rausch Verhältnisses (SNR). Hier wird von einem Vocal-to-Instrument Ratio (VIR) gesprochen, der sich wie folgt berechnet. Je negativer der VIR-Wert, desto höher wurde der Pegel der Gesangspur gegenüber dem Instrumental eingestellt.

$$VIR = L_{BackgroundInstrumental} - L_{Vocals}$$

Bei Experiment 2 werden unterschiedliche Versionen desselben Musikstücks paarweise dargeboten. Verglichen werden drei Versionen: Das Original (0 dB Abschwächung), mit abgeschwächtem Instrumental von -9 dB und mit einem abgeschwächten Instrumental von -18 dB. Durch die paarweise Darbietung ergeben sich folgende Kombinationsmöglichkeiten: 0/-9 dB, 0/-18 dB und -9/-18 dB. Bei der Präferenz der „dem Original näheren“ Version (0 dB bzw. -9 dB) wird einen logische 1 hinterlegt. Somit ergibt sich bei einer stetigen Präferenz des Originals ein Prozentwert von 100%. Für alle Versionen wurde ein Lautheitsausgleich durchgeführt. Insgesamt wurden hier 6 unterschiedliche Musikstücke verschiedener Musikstile dargeboten.

Für beide Experimente wurde auf lizenzfreie Musik zurückgegriffen, die auf der Webseite www.soundstripe.com zu finden sind. Außerdem wurden jedes Experiment sowohl im Freifeld über Lautsprecher (bimodales Hören), als auch über direktes Streaming in das CI durchgeführt. Über das direkte Streamen ist es möglich die Wahrnehmung über das kontralaterale Ohr auszuklammern.

2. Ergebnisse:

Die Ergebnisse von Experiment 1 zeigen den gemessenen VIR-Wert (Vocal-to-Instrument Ratio) der jeweiligen Versuchsgruppen für 2 unterschiedliche Musikstücke, wobei *Musikstück 1* eine mehr legato gespielte Instrumentierung aufweist, und *Musikstück 2* deutlich rhythmischere Elemente beinhaltet.

Abbildung 1 links zeigt die VIR-Werte der entsprechenden Musikstücke und vergleicht dabei CI-Gruppe und normalhörende Kontrollgruppe im Freifeld. Für beide Musikstücke zeigt sich eine Tendenz zu negativeren VIR-Werten bei der CI-Gruppe. Diese zeigt sich für *Musikstück 2* als hoch signifikant im t-Test ($p < 0,001$). Je negativer der VIR-Wert, desto höher wurde der Gesangspegel gegenüber dem Instrumental eingestellt. Im Mittel hat sich die CI-Gruppe für *Musikstück 2* den Gesangspegel um 2.6 dB höher eingestellt als die Kontrollgruppe.

In Abbildung 1 rechts werden analog beide Gruppen verglichen, nur dass hier die CI-Gruppe über direktes Streaming getestet wurde. Für beide Musikstücke zeigen sich hier hoch signifikante Unterschiede ($p < 0,001$) zwischen den VIR-Werten beider Gruppen. Auch hier wurden von der CI-Gruppe die Gesangspegel höher eingestellt als das Instrumental.

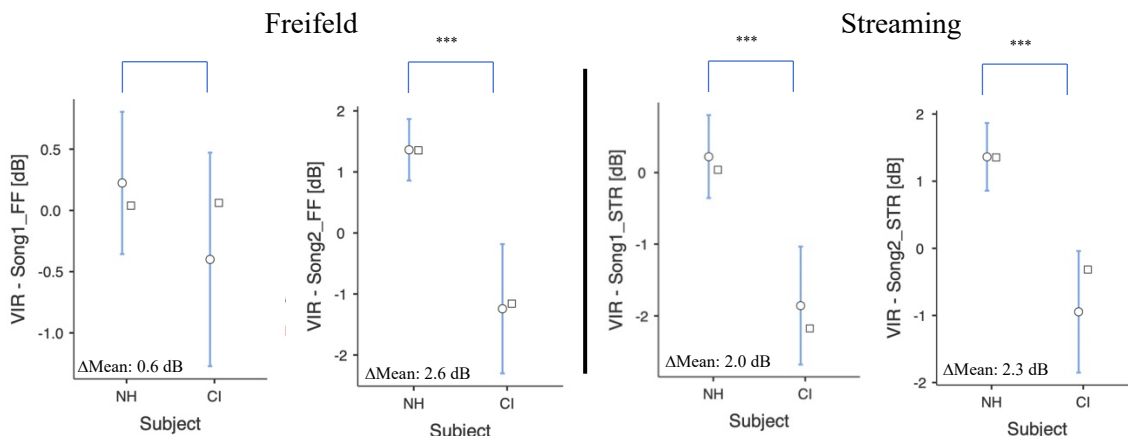


Abbildung 1: Vergleich des eingestellten Vocal-to-Instrument Ratio (VIR) im Freifeld (links) und mittels Streaming (rechts). Hierbei zeigt das jeweils linke Diagramm die VIR-Werte für Musikstück 1 und das jeweils rechte für Musikstück 2.

Abbildung 2 bezieht sich auf den Vergleich von Freifeld-Bedingung (bimodales Hören) und direktem Streaming in das Implantat (monaurales/elektrisches Hören) innerhalb der CI-Gruppe. Hierbei zeigt sich für Musikstück 1 ein signifikanter Unterschied zwischen den VIR-Werten bei entsprechender Darbietung. Anders zeigt sich für Musikstück 2 kein signifikanter Unterschied. Die Ergebnisse von Experiment 2 in Abbildung 3 zeigen die Präferenz für eine Version eines dargebotenen Musikstücks. Wenn tendenziell das originale(re) Musikstück gegenüber der Version mit betontem Gesang bevorzugt wurde, ergeben sich in den Grafiken von Abbildung 3 Prozentwerte in Richtung 100%.

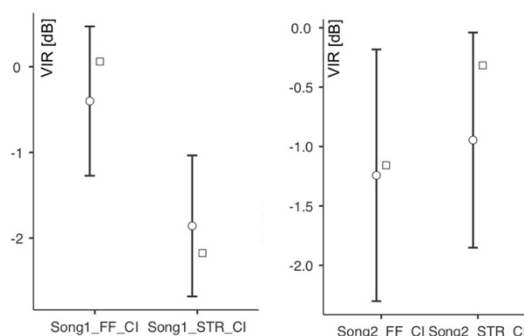


Abbildung 2: Vergleich der VIR innerhalb der Gruppe der CI-Träger zwischen Freifeld und Streaming-Kondition für Musikstück 1 (links) und Musikstück 2 (rechts)

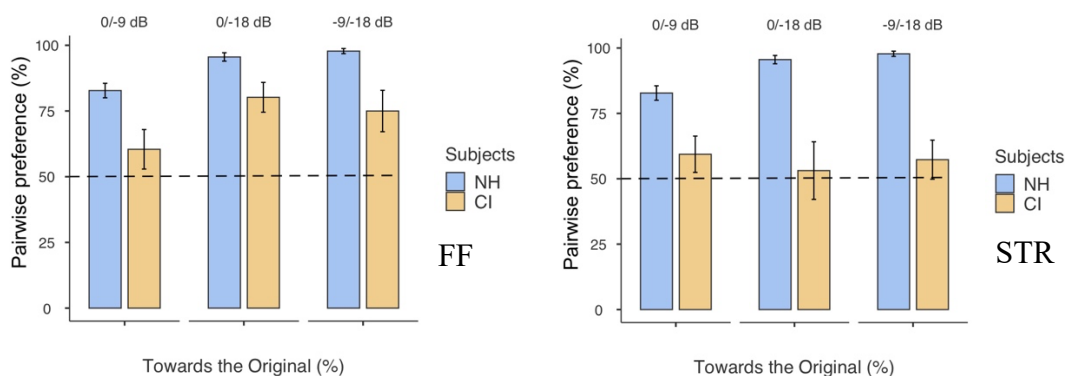


Abbildung 3: Präferenz im Paarvergleich (Paare siehe Überschriften) im Freifeld (FF, links) und in der Streaming Kondition (STR, rechts). NH blau, CI gelb; Ein höherer Prozentwert bedeutet, dass die originalere Version gegenüber der Version mit betonter Gesangsspur bevorzugt wurde. Die gestrichelte Linie repräsentiert die Ratewahrscheinlichkeit von 50%

Die normalhörende Kontrollgruppe (blaue Balken) bevorzugte stets die originalere Version des jeweiligen Musikstückes. Die CI Gruppe zeigte in der Freifeld-Kondition ebenfalls eine Tendenz zur Präferenz der originalen Musikstücke, im Gegensatz zur Streaming-Kondition, in welcher sich die Angaben nahe der Ratewahrscheinlichkeit von 50% bewegen, d.h. es gibt keine eindeutige Präferenz für eine der beiden präsentierten Versionen.

3. Diskussion:

Die Ergebnisse von Experiment 1 zeigen VIR- Werte von bis zu 2,6 dB Differenz im Mittelwert für die Freifeld-Kondition. Es wurde also die Gesangsspur im Mittel bis zu 2,6 dB (*Musikstück 2*) höher eingepegelt als das Instrumental. Dabei sind Ergebnisse auch stark vom dargebotenen Musikstück abhängig. Ein Grund dafür könnte sein, dass die Gesangsspur in *Musikstück 2* eine wesentlich rhythmischere Struktur aufweist als *Musikstück 1* und dadurch eine zentralere Rolle einnimmt, die von der CI-Gruppe tendenziell höher eingepegelt wurde. Im Gegensatz dazu hat die Kontrollgruppe die Gesangsspur von *Musikstück 2* im Mittel sogar etwas niedriger eingepegelt als das Original. Obwohl dieser Effekt bei dem Vergleich beider Gruppen für *Musikstück 1* nicht signifikant ausfällt, lässt sich auch die Tendenz ablesen, dass sich die CI-Gruppe die Gesangsspur höher eingepegelt hat (Δ Mean: 0.6 dB).

In der Streaming-Kondition sind die unterschiedlich eingestellten Verhältnisse von Gesangsspur und Instrumental (VIR-Werte) hoch signifikant (bis zu 2,3 dB) zwischen beiden Gruppen. Somit ist mit dem Verweis auf Fragestellung 1 zu bejahen, dass sich hier die beiden Probandengruppen unterschiedliche Mischverhältnisse zwischen Gesang und Instrumental einstellen.

In Bezug auf die zweite Fragestellung, ob es innerhalb der CI-Gruppe Unterschiede zwischen der Darbietungsart (Freifeld oder Streaming) gibt, kann gezeigt werden, dass vor allem die Gesangsspur von *Musikstück 1* in der Streaming-Kondition höher eingepegelt wurde als bei der Freifeld Darbietung. Auch hier ist anscheinend die Struktur des Musikmaterials ausschlaggebend.

Experiment 2 zeigt für die normalhörende Gruppe eine starke Präferenz für die „originaleren“ Versionen der dargebotenen Musikstücke (bis zu 95%). In der Freifeld-Kondition liegt die Präferenz der CI Gruppe ebenfalls bei den originalen Versionen (bis zu 78%), was den Einfluss des kontralateral, akustisch hörenden Seite widerspiegelt. Demgegenüber zeigt sich für die Streaming-Kondition, in Einklang mit den Daten von Pons et al. (2016), kein signifikanter Unterschied zu einer Ratewahrscheinlichkeit von 50%. Demnach ist hier keine eindeutige Präferenz ablesbar.

Die Ergebnisse der Studie deuten darauf hin, dass bimodale CI-Träger, insbesondere wenn Musik ins CI gestreamt wird, sich selbst eine betonte Gesangsspur einstellen würden (Experiment 1). Dies könnte zu einem erhöhten Hörerlebnis beitragen. Außerdem suggerieren die Ergebnisse, dass das konterlaterale Ohr eine dominierende Rolle bei der Freifeld Darbietung einnimmt. Ein Echtzeit-Algorithmus zur Spurentrennung stellt ein vielversprechendes Werkzeug für eine Implementierung in modernen Hörsystemen dar.

4. Literatur:

Buyens, W., van Dijk, B., Moonen, M., & Wouters, J. (2014). Music mixing preferences of cochlear implant recipients: A pilot study. *International journal of audiology*, 53 (5), 294–301.

Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear im- plant users and its relationship with psychophysical capabilities. *Journal of rehabilitation research and development*, 45(5), 779-790.

Imhäuser, L. (2021). Music perception of bimodal Cochlea Implant Users. [Bachelorarbeit, Technische Hochschule Lübeck].

Kang, R., Nimmons, G. L., Drennan, W., Longnion, J., Run, C., Nie, K., Won, J., Worman, T., Yueh, B. & Rubinstein, J. (2009). Development and validation of the University of Washington clinical assessment of music perception test. *Ear and hearing*, 30 (4), 411-418.

Pons, J., Janer, J., Rode, T., & Nogueira, W. (2016). Remixing music using source separation algorithms to improve the musical experience of cochlear implant users. *The Journal of the Acoustical Society of America*, 140 (6), 4338–4349.

The jamovi project (2021). (2013, October). *jamovi. (version 2.0)*. [Computer Software]. Retrieved from <https://www.jamovi.org>.