

Untersuchung eines neuronengebasierten Vocoder hinsichtlich der Fähigkeit virtuelle Kanäle zu simulieren

Sean-Patrik Cretti, Lilia Lajmi, Kai Michael Blum

Ostfalia Hochschule für angewandte Wissenschaften – Fakultät Elektrotechnik, 38302 Wolfenbüttel

Einleitung

Ein Cochlea-Implantat (CI) ist ein elektronisches medizinisches Hörsystem für Menschen mit starker bis hochgradiger Hörminderung. Ein CI übernimmt die Funktion des beschädigten Innenohrs, um Audiosignale in Form von modulierten Impulsen an die Hörnerven zu übertragen und ermöglicht so für viele Menschen, die an Hörverlust leiden wieder das Hören.

Forscher im Bereich der Sprachkodierung für CI interessieren sich dafür, wie ein Audiosignal bei einem Patienten mit hochgradiger Schwerhörigkeit durch ein CI wahrgenommen wird. Zahlreiche Untersuchungen und subjektive Messungen wurden seit dem letzten Jahrhundert diesbezüglich bei Probanden durchgeführt. Das Ziel solcher Messungen ist dabei stets ein natürliches Klangbild zu erreichen. Dabei ist jedoch zu beachten, dass die Ergebnisse stark auf den Patienten bezogen und unter Umständen nur den klinischen Organisationen zugänglich sind.

Für die Forschung und Entwicklung von CI ist es sehr hilfreich zuerst die neuartigen Sprachkodierungsstrategien bei Normalhörende (NH) zu testen [1], bevor sie von echten CI-Trägern ausgewertet werden. Die subjektiven Ergebnisse der CI-Probanden weisen in der Regel eine große Abweichung auf [2] [3].

Eine übliche und schnelle Simulationsmethode zur Beurteilung der Qualität der elektrischen Reize als Ergebnis der Kodierung ist der Vocoder [4]. Häufig werden für die Simulation entweder Sinus- oder Rausch-Vocoder angewendet [5] [6]. Mit dieser Syntheseart wird die Hüllkurveninformation aus dem Audiosignal durch Modulieren der akustischen Träger repräsentiert. Die synthetisierten Audiodaten werden dann von Normalhörenden subjektiv bewertet.

In [7] [8] wurde berichtet, dass sich die Klangqualität dieser Vocoder vom ausgewerteten Klang eines CI unterscheiden. Sinuswellen induzieren vermutlich eine engere Ausbreitung der Anregung über die tonotopische Achse, als die monopolare Stimulation einer Intracochlea-Elektrode und können daher die Kanalwechselwirkungen, die CI-Träger erfahren, nicht nachahmen [9] [10].

In Karoui et al. [11] wurden Sinus-, Rausch- und PSHC-Vocoder (Pulse Spreading Harmonic Complex) [10] verglichen. Dabei kam heraus, dass das Hörerlebnis von NH am besten mit dem Hörerlebnis von CI-Trägern korreliert, wenn der PSHC-Vocoder verwendet wird. Dennoch gibt es Unterschiede, sodass Untersuchungen mit NH nicht auf CI-Träger übertragbar sind. Ähnliche Ergebnisse wurden in [7] berichtet.

Damit die Auswertung der Tests mit NH eine höhere Aussagekraft über die Qualität der Algorithmen bieten kann, müssen Audiosignale erzeugt werden, die bei NH zu einem Hörerlebnis führen, das dem von CI-Trägern ähnelt. Das erfordert Kenntnisse der biologischen Vorgänge im Gehör auf neuronaler Ebene. Die aktuellen Forschungen und Entwicklungen gehen dahin, ein künstliches Nervensystem für die Erstellung der Audiosignale nachzubilden.

In der Universität Zürich wurde ein neuronengebasiertes Vocoder entwickelt, mit dem die biomedizinischen Effekte auf der neuronalen Ebene und die physikalischen Auswirkungen der eingesetzten CI-Elektroden berücksichtigt werden [12]. Beispiele hierfür sind die Refraktärzeit (Erholungszeit) der Neuronen und das Übersprechen der Stromimpulse auf benachbarte Kanäle.

Virtuelle Kanäle werden in Sprachkodierungsstrategien für Cochlea-Implantate eingesetzt, um die wahrnehmbaren Töne für CI-Träger zu erhöhen [13]. Es werden zwei Elektroden gleichzeitig aktiviert, wodurch es zu einer Überlagerung der elektrischen Felder kommt.

In dieser Publikation wird untersucht, wie die neuronengebasierte Synthese auf Stromimpulse reagiert, die zur Erzeugung virtueller Kanäle dienen.

Verwandte Arbeiten

In diesem Kapitel wird auf die zugrundeliegenden Arbeiten eingegangen. Zunächst wird die Technik der virtuellen Kanäle beschrieben. Anschließend wird auf die Funktionsweise des in [12] beschriebenen neuronalen Modells eingegangen.

Virtuelle Kanäle

Durch die Anzahl an Elektroden ist die spektrale Auflösung stark limitiert. Trotz dieser Limitierung ist es Cochlea-Implantat-Trägern möglich Sprache in ruhiger Umgebung zu verstehen. Gibt es allerdings viele Nebengeräusche, wie z.B. in einer Bar, ist eine höhere spektrale Auflösung hilfreich. Instrumente zu erkennen und Melodien wahrzunehmen ist ebenfalls nur mit einer erhöhten spektralen Auflösung möglich [13]. Townshend et al. [14] berichteten 1987, dass zusätzliche Tonhöhen von Patienten wahrgenommen werden, wenn zwei benachbarte Elektroden gleichzeitig angesteuert werden. Welche Tonhöhe wahrgenommen wird, kann durch den Strom, der durch die Elektroden fließt, gesteuert werden. Die wahrgenommene Lautstärke bleibt dabei gleich, solange die Summe, der durch die beiden Elektroden fließenden Ströme, konstant ist [13]. Um diese Stromsteuerung zu realisieren, musste jedoch eine neue Generation von Cochlea-Implantaten entwickelt werden. Denn bei den bisherigen Modellen war es nicht möglich zwei Elektroden gleichzeitig mit Strom zu versorgen. Die neu entwickelten Cochlea-Implantate versorgen jede Elektrode mit einer eigenen

Stromquelle. So ist es möglich zwei Elektroden gleichzeitig anzusteuern und die Stromstärke zu variieren. In Abbildung 1 ist das Prinzip der Stromsteuerung dargestellt.

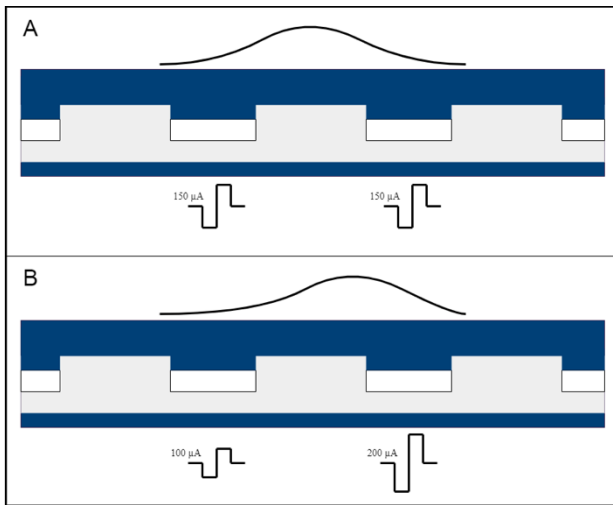


Abbildung 1: Prinzip der Stromsteuerung beim Einsatz virtueller Kanäle

In Teil A der Abbildung ist zu sehen, wie zwei benachbarte Elektroden jeweils mit $150\mu\text{A}$ angesteuert werden. Durch die Stromausbreitung kommt es zu einer Überlagerung der beiden Stromimpulse. Dadurch kommt es zu einem Strommaximum zwischen den Elektroden und die Hörnervenzellen werden dort am stärksten angeregt. So nimmt der Patient einen anderen Ton wahr. Wenn nun das Stromstärkeverhältnis zwischen den Elektroden verändert wird, ändert sich auch der Ort, an dem die Hörnervenzellen aktiviert werden. Dies ist in Teil B der Abbildung zu sehen. Nun wird die erste Elektrode mit $200\mu\text{A}$ angesteuert und die zweite Elektrode mit $100\mu\text{A}$. Dadurch verschiebt sich die wahrgenommene Tonhöhe in Richtung der ersten Elektrode.

Neuronenbasierter Vocoder

In den Studien [12], [15] und [16] wurden drei Simulationsmodelle zum Aufbau eines Neuronenmodells vorgeschlagen. In dieser Arbeit ist das Simulationsmodell in Anlehnung an diese drei Modelle in Abbildung 2 dargestellt.

Im Folgenden werden die Funktionen jedes Blocks aus Abbildung 2 erklärt.

Sprachkodierungsstrategie

Die Kodierungsstrategie verarbeitet das akustische Signal und wandelt die Informationen des Signals in elektrische Impulse um. Diese Impulse stimulieren die Hörnervenzellen in der Cochlea des CI-Trägers. Dadurch wird dem CI-Träger das Hören wieder ermöglicht. Es gibt unterschiedliche Arten von Sprachkodierungsstrategien, wie beispielsweise Continuous Interleaved Sampling [17] oder Psychoacoustic Advanced Combinational Encoder [18]. Die Information über die Amplitude der elektrischen Impulse dient als Eingabe für den neuronengebasierten Vocoder.

Stromausbreitung

Die von den Elektroden abgegebenen elektrischen Impulse stimulieren die Hörnervenzellen nicht punktuell an den Stellen, an denen die Elektroden in der Cochlea platziert sind.

Stattdessen breitet sich der Strom in der Cochlea aus, sodass eine Gruppe von Hörnervenzellen stimuliert wird. Wie weit sich der Strom in der Cochlea ausbreitet hängt von der Amplitude und Form des Impulses ab. [19]

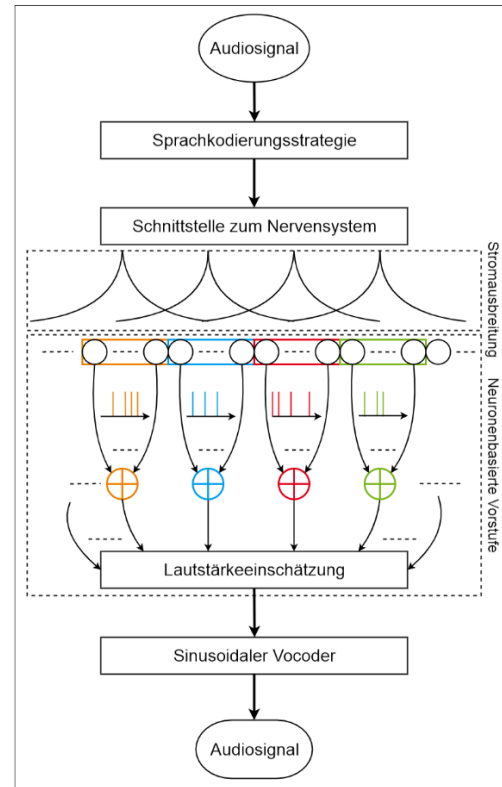


Abbildung 2: Simulationsmodell des Nervensystems

Bei der Berechnung der Stromausbreitung wurde zur Vereinfachung von einer gaußförmigen Kurvenform der Stromausbreitung ausgegangen. In Abbildung 3 ist der Verlauf der Stromausbreitung in Abhängigkeit der elektrischen Amplitude dargestellt. Es zeigt sich, dass durch hohe Stromamplituden Neuronen angeregt werden, die anderen Elektroden zugeordnet sind.

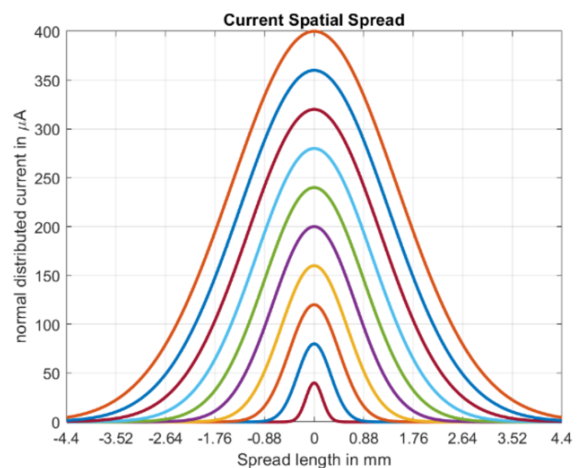


Abbildung 3: Stromausbreitung bei verschiedenen Eingangsströmen

Diese Überschneidung wird als Kanalinteraktion bezeichnet und führt zur Beeinträchtigung der Sprachverständlichkeit [20].

Neuronenbasierte Vorstufe und Synthese

Die neuronenbasierte Vorstufe verarbeitet die Stromimpulse der Elektroden. Sie simuliert das Verhalten der Hörnervenzellen. In Abbildung 4 wird die Anordnung der Neuronen sowie die Zuordnung zu den Elektroden veranschaulicht.

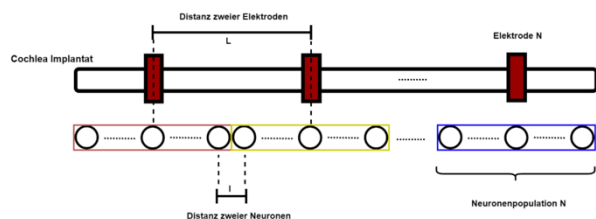


Abbildung 4: Anordnung der Neuronen entlang des Elektrodenarrays

Für die Simulation der Neuronen wurde das in [21] vorgestellte Modell verwendet. Dieses berücksichtigt die absolute Refraktärzeit der Neuronen. Außerdem wurde es um die relative Refraktärzeit nach [22] erweitert. Die neuronale Informationsweiterleitung basiert auf dem „Alles-oder-Nichts-Prinzip“ [23]. Darunter ist ein einfacher Mechanismus eines Neurons zu verstehen. Entweder wird das Neuron aufgrund eines genügend großen elektrischen Reizes feuern oder aufgrund zu geringer Energie nicht feuern. Aus der Aktivität der einzelnen Neuronen wird die Feuerrate für eine Population von Neuronen berechnet. Anhand der Feuerrate wird nach [12] die empfundene Lautstärke geschätzt. Diese Lautstärke dient als Eingangssignal für den sinusoidalen Vocoder.

Untersuchung des neuronengebasierten Vocoders

Um das Verhalten des Neuronenmodells bei der simultanen Ansteuerung zweier Elektroden zu testen, wurden Eingabedaten erzeugt, bei denen zwei benachbarte Elektroden mit einer bestimmten Amplitude angesteuert werden. Alle anderen Elektroden geben keinen Stromimpuls ab. In diesem Modell überlagert sich die Stromausbreitung der beiden elektrischen Impulse additiv. Die resultierende Stromausbreitungskurve ist in Abbildung 5 beispielhaft dargestellt.

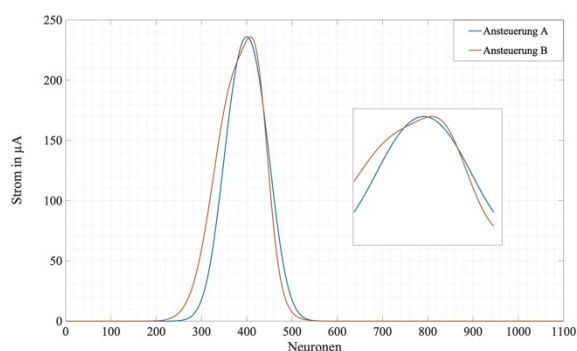


Abbildung 5: Beispielhafte Darstellung der Überlagerung zweier Stromausbreitungskurven bei gleichzeitiger Ansteuerung der Elektroden (blau: jeweils $150\mu\text{A}$, rot: Elektrode A $100\mu\text{A}$ / Elektrode B $200\mu\text{A}$)

Die blaue Kurve ergibt sich durch die gleichzeitige Ansteuerung der Elektroden mit jeweils $150\mu\text{A}$. Für die rote

Kurve wurde eine Elektrode mit $200\mu\text{A}$ und die zweite mit $100\mu\text{A}$ angesteuert.

Abbildung 6 und 7 zeigen die, bezüglich des Maximums amplitudennormierte, durchschnittliche Feuerrate der Neuronen für die beiden oben genannten Beispiele. In Blau ist die durchschnittliche Neuronenfeuerrate aufgetragen. In Rot ist noch einmal die dazu gehörige Stromausbreitungskurve dargestellt. Es ist zu sehen, dass die Neuronenfeuerrate mit steigender Stromamplitude steigt. Auch die Verschiebung in Richtung Elektrode B bei unterschiedlichen Stromamplituden ist zu sehen. Dass die Kurvenform der Neuronenfeuerrate nicht genau der Form der Stromausbreitung entspricht, hängt mit der Variation der Simulationsparameter für die Neuronen zusammen [12].

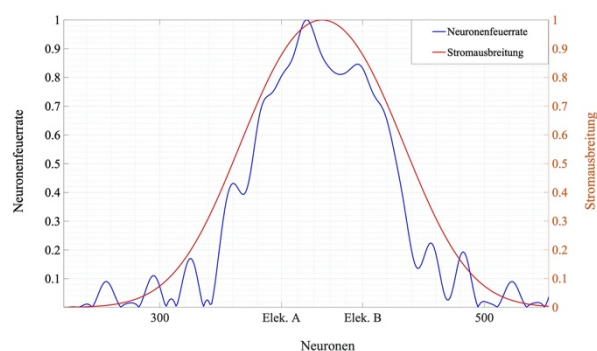


Abbildung 6: Beispielhafte Darstellung der Neuronenfeuerrate bei gleichzeitiger Ansteuerung der Elektroden mit jeweils $150\mu\text{A}$

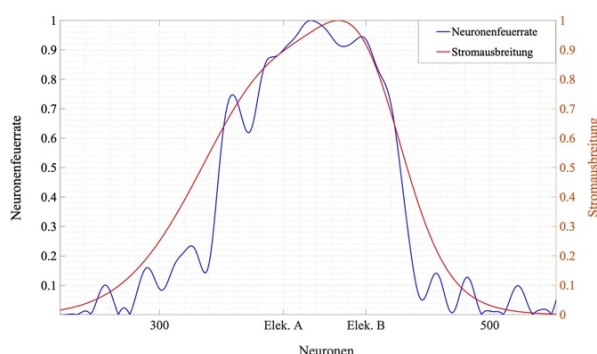


Abbildung 7: Beispielhafte Darstellung der Neuronenfeuerrate bei gleichzeitiger Ansteuerung der Elektroden mit $200\mu\text{A}$ bzw. $100\mu\text{A}$

Zusammenfassung und Ausblick

In dieser Arbeit wurde der in [12] präsentierte neuronengebasierte Vocoder hinsichtlich der Verarbeitung virtueller Kanäle untersucht. Es zeigt sich, dass die additive Überlagerung der Stromausbreitungskurven in Abhängigkeit des Stromstärkeverhältnisses der beiden Elektroden zu einem Maximum der Neuronenfeuerrate zwischen den Elektroden führt. Die genaue Position des Maximums ist allerdings von den Modellierungsparametern der Neuronen abhängig. Wie präzise spektrale Komponenten durch die Verwendung virtueller Kanäle repräsentiert werden können ist somit von diesen Parametern abhängig. Das entspricht den Ergebnissen in [24], denen zufolge die Wahrnehmung weiterer Frequenzkomponenten individuell stark verschieden sein kann.

Im weiteren Verlauf dieser Arbeit soll die Verarbeitung virtueller Kanäle mit dem neuronengebasierten Vocoder hinsichtlich der Audioqualität der erzeugten Audiosignale untersucht werden. Außerdem soll untersucht werden, ob die in [25] präsentierte Methode zur Erzeugung virtueller Kanäle zu ähnlichen Ergebnissen führt, wie die hier vorgestellte.

Literatur

- [1] Dorman, M. F., Loizou, P. C.: Speech intelligibility as a function of the number of channels of stimulation for normal hearing listeners and patients with cochlear implants. *The American Journal of Otology* (1997), 113-114
- [2] Lazard, D., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterker, O., Blamey, P. J.: Pre-, per- and postoperative factors affecting performance of postlingually deaf adults using cochlear implants: A new conceptual model over time. *PLoS One* (2012)
- [3] James, C. J., Laborde, M.-L., Lepage, B., Molinier, C.-E., Tartayre, M., Fraysse, B.: Early sentence recognition in adult cochlear implant users. *Ear and Hearing* (2018), 905-917
- [4] Shannon, R. V., Zeng, F., Kamath, V., Wygonski J., Ekelid, M.: Speech recognition with primarily temporal cues. *Science* 270 (1995), 303-304
- [5] Baskent, D., Shannon, R. V.: Frequency-place compression and expansion in cochlear implant listeners. *J. Acoust. Soc. Am.* 116 (2004), 3130-3140,
- [6] Rosen, R., Faulkner A., Wilkinson, L.: Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *J. Acoust. Soc. Am.* 106 (1999), 3629-3636
- [7] Dorman, M. F., Natale, S. C., Butts, A. M., Zeitler, D. M., Carlson, M. L.: The sound quality of cochlear implants: Studies with single-sided deaf patients. *Otology & Neurotology* 38 (2017), 268-273
- [8] Peters, J. P. M., Wendrich, A. W., van Eijl, R. H. M., Rhebergen, K. S., Versnel, H., Grolman, W.: The sound of a cochlear implant investigated in patients with single-sided deafness and a cochlear implant. *Otology & Neurotology* 39 (2018), 707-714
- [9] Snyder, R. L., Bierer, J. A., Middlebrooks, J.: Topographic spread of inferior colliculus activation in response to acoustic and intracochlear electric stimulation. *J. As. Res. Otolaryngology* 5 (2004), 305-322
- [10] Mesnildrey, Q., Hilkuysen, G., Macherey, O.: Pulse spreading harmonic complex as an alternative carrier for vocoder simulations of cochlear implants. *J. Acoust. Soc. Am.* 139 (2016), 986-991
- [11] Karoui, C., James, C., Barone, P., Bakhos, D., Mathieu, M., Macherey, O.: Searching for the Sound of a Cochlear Implant: Evaluation of Different Vocoder Parameters by Cochlear Implant Users With Single-Sided Deafness. *Trends in Hearing* (2019), 1-14
- [12] El Boghdady, N.: *Multichannel EEG Analysis of Sound Coding Strategies for Cochlear Implants*. Universität Zürich. 2014.
- [13] Advanced Bionics, URL: https://advancedbionics.com/content/dam/advancedbionics/Documents/Global/en_cc/Professional/Technical-Reports/Sound-Processing/AB_HiRes_Fidelity_120_Sound_Processing_report.pdf.
- [14] Townshend, B., Cotter, N., Van Compornolle, D., White, R. L.: Pitch perception by cochlear implant subject. *J. Acoust. Soc. Am.* (1987), 106-115
- [15] Bruce, I. C., White, M. W., Irlicht, L. S., O'Leary, S. J., Clark, G. M.: The effects of stochastic neural activity in a model predicting intensity perception with cochlear implants: low-rate stimulation. *IEEE Transactions on Biomedical Engineering* (1999), 1393-1404
- [16] Krause, T. U. *Rate Coding and Temporal Coding in a Neural Network*. Ruhr Universität Bochum, 2014
- [17] Wilson, B., Finley, C., Lawson, D., Wolford, R., Eddington, D., Rabinowitz, W.: Better speech recognition with cochlear implants. *Nature* 352 (1991), 236-238
- [18] Nogueira, W., Büchner, A., Lenarz, T., Edler, B.: A Psychoacoustic "NofM"-Type Speech Coding Strategy for Cochlear Implants. *EURASIP Journal on Applied Signal Processing* (2005), 3044-3059
- [19] Cohen, L. T., Richardson, L. M., Saunders, E., Cowan, R. S. C.: Spatial spread of neural excitation in cochlear implant recipients: comparison of improved ECAP method and psychophysical forward masking. *Hearing Research* 179 (2003), 72-87
- [20] Noble, J. H., Gifford, R. H., Hedley-Williams, A. J., Dawant, B. M., Labadie, R. F.: Clinical evaluation of an image-guided cochlear implant programming strategy. *Audiology & Neurotology* (2014), 400-411
- [21] Chen, F., Zhang, Y.-T.: An integrate-and-fire-based auditory nerve model and its response to high-rate pulse train. *Neurocomputing* (2007), 1051-1055
- [22] Avissar, M., Wittig, J. H., Saunders, J. C., Parsons, T. D.: Refractoriness Enhances Temporal Coding by Auditory Nerve Fibers. *Journal of Neuroscience* (2013), 7681-7690
- [23] Euler, M. *Kreative Uhren in unserem Kopf: Experimente zum Hören und Nachdenken*. 2004
- [24] Firszt, J. B., Koch, D. B., Downing, M., Litvak, L.: Current steering creates additional pitch percepts in adult cochlear implant recipients. *Otology & Neurotology* (2007), 629-636
- [25] *Pitch Steering with Sequential Stimulation of Intracochlear Electrodes* URL: <https://www.semanticscholar.org/paper/Pitch-Steering-with-Sequential-Stimulation-of/08432dbc0c54eac9445fe6abf578ea02f37c1897>