

Anpassbarer Fernsehon für Hörgeschädigte über HbbTV 2.0

Michael Weitnauer, Sebastian Goossens, Theresa Liebl

Institut für Rundfunktechnik, 80939 München, E-Mail: weitnauer@irt.de; goossens@irt.de; liebl@irt.de

Einführung

Besonders für hörgeschädigte Personen ist die gute Verständlichkeit des gesprochenen Dialogs im Fernsehen ein wichtiger Punkt. Aus vielerlei Gründen wird die Sprachverständlichkeit aktueller TV-Produktionen vor allem von älteren Personen und Hörgeschädigten oft als unzureichend beurteilt. Das EU-Projekt HBB4ALL¹ (Hybrid Broadcast Broadband for All) [1] beschäftigt sich mit dem barrierefreien Zugang zu Medien für alle Bürger über den HbbTV-Standard (Hybrid Broadcast Broadband TV). Eine der Herausforderungen dabei ist die Bereitstellung von Inhalten über verschiedene Kanäle hinweg mit einer Spezifizierung für unterschiedliche Zielgruppen. Die Personalisierung von interaktiven Fernsehinhalten und Dienstleistungen, ist eine der zentralen Aufgabenstellungen von HBB4ALL.

Ziel des Projektes ist es, unter anderem die Sprachverständlichkeit von TV-Produktionen insbesondere für Hörgeschädigte zu verbessern. Es soll dabei ein automatisierbarer Dienst realisiert werden, welcher einfach in bestehende Workflows zu integrieren ist. Dabei soll sowohl der aktuelle Produktions-Workflow, als auch die Bereitstellung eines Clean Audio Services für den Endnutzer unter Nutzung spezieller Funktionen des HbbTV 2.0 Standards berücksichtigt werden. Für die Übertragung des Clean-Audio-Dienstes zum Endnutzer soll das Internet (IP) genutzt werden. Das Institut für Rundfunktechnik (IRT) beschäftigt sich mit der automatisierten Bearbeitung und Übertragung von individuell auswählbaren Clean-Audio-Signalen. Clean Audio bedeutet in diesem Zusammenhang, dass die bereits zusammengemischten Signale bezüglich ihrer Sprachverständlichkeit aufgebessert werden. Nachdem mittels HbbTV eine Bearbeitung im Endgerät nicht möglich ist, muss hierbei ein Kompromiss zwischen Personalisierung und zur Verfügung zu stellenden Streams getroffen werden. Es zeigte sich bereits sehr früh im Projekt, dass es keine allgemeingültige Bearbeitung gibt, welche alle Anforderungen von Hörbeeinträchtigten an ein solches Signal abdeckt. Dementsprechend ist es das Ziel, verschiedene Bearbeitungen anzubieten.

Online-Umfrage

Im Zuge der Bestrebungen den Fernsehon speziell für Hörgeschädigte zu verbessern wurde vom IRT für diese Zielgruppe eine Online-Umfrage erstellt. Dabei sollte ermittelt werden, ob es spezielle Faktoren gibt, welche die Sprachverständlichkeit im aktuellen TV-Programm besonders beeinflussen und welche Maßnahmen zur Verbesserung ergriffen werden sollten.

An der Umfrage haben zum aktuellen Stand 116 Personen teilgenommen, wobei der Großteil der Teilnehmer eine an Taubheit grenzende Hörschädigung aufweist (siehe Abbildung 1) und 87% aller Teilnehmer bereits eine Hörhilfe trägt.

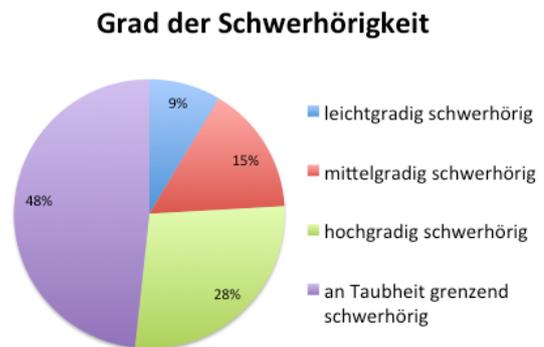


Abbildung 1: Prozentuale Verteilung des Grades der Schwerhörigkeit der 116 Umfrage-Teilnehmer

Bei der Frage, welche Sendungen im TV besonders problematisch im Bezug auf die Sprachverständlichkeit sind, wurde vom Großteil der Teilnehmer der Spielfilm genannt, dicht gefolgt von Live-Fernsehsows (siehe Abbildung 2). Auch Sport-Sendungen wurden von fast 50% als problematisch eingestuft. Dokumentationen und Nachrichten schnitten im Vergleich zum Rest relativ gut ab.

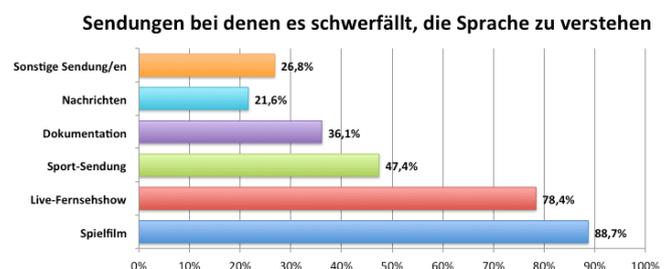


Abbildung 2: Prozentuale Verteilung der Antworten aller 116 Teilnehmer an der Online-Umfrage auf die Frage bei welchen Sendungen es besonders schwer fällt die Sprache zu verstehen

Im Weiteren wurde in der Umfrage nach Störfaktoren gefragt, die beim Fernsehon für Probleme bei der Sprachverständlichkeit sorgen. Dabei zeigt sich in der Auswertung der Antworten (Abbildung 3), dass nahezu alle genannten Faktoren mehrheitlich als *sehr störend* eingestuft wurden. Eine Ausnahme bildet dabei der Faktor Dialekt. Dieser wurde als einziger etwas schwächer eingestuft. Als größter Störfaktor wurde mit 68% der Bewertungen bei *sehr störend* die undeutliche Aussprache genannt. Zu schnelles Sprechen wurde mit 57% bei *sehr störend* als zweithäufigster Faktor genannt. Daran zeigt sich, dass Faktoren wie Nuscheln, zu schnelle oder undeutliche Aussprache, oft mehr Einfluss auf die

¹ Das Projekt HBB4ALL wird gefördert im Rahmen des 'Competitiveness and innovation framework programme 2007-2013' der Europäischen Union

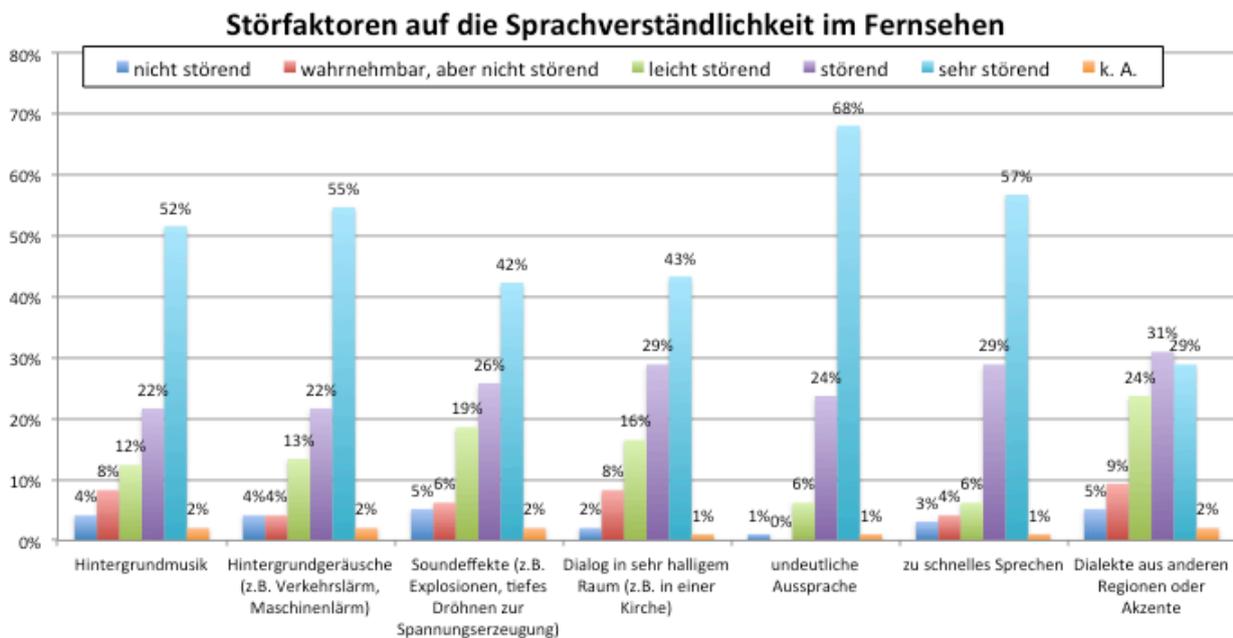


Abbildung 3: Prozentuale Verteilung der Antworten aller 116 Teilnehmer an der Online-Umfrage auf die Frage welche Störfaktoren die Sprachverständlichkeit im Fernsehen beeinflussen

Sprachverständlichkeit haben als zu laute Hintergrundmusik oder -geräusche. Da allerdings auch Faktoren wie Hintergrundmusik, Hintergrundgeräusche und Soundeffekte einen großen Störeinfluss auf die Verständlichkeit haben, würde eine Anpassung aller genannten Einflüsse sicherlich die bestmögliche Verbesserung bewirken.

Diese Schlussfolgerung spiegelt sich auch in den Ergebnissen der Frage wieder, mit welchen Maßnahmen eine Verbesserung der Sprachverständlichkeit im Fernsehen am besten erreicht werden sollte (siehe Abbildung 4). Mit 67% wurde die Maßnahme „Klangfarbe der Stimme betonen/verbessern“

am häufigsten mit sehr wichtig bewertet. Auch der Wunsch nach einer eigenständigen Lautstärkeregelung der Sprache, mit 56% der Bewertungen bei *sehr wichtig*, zeigt sich an diesen Ergebnissen sehr deutlich.

Einige dieser Erkenntnisse aus der Online-Umfrage wurden Ende 2014 schon von einer Arbeitsgruppe der ARD in einer Produktionsempfehlung für Programm und Technik [2] festgehalten. Darin werden vielerlei Probleme entlang der Produktionskette aufgezeigt und Empfehlungen gegeben, um Sprachverständlichkeit im Fernsehen von Grund auf zu verbessern.

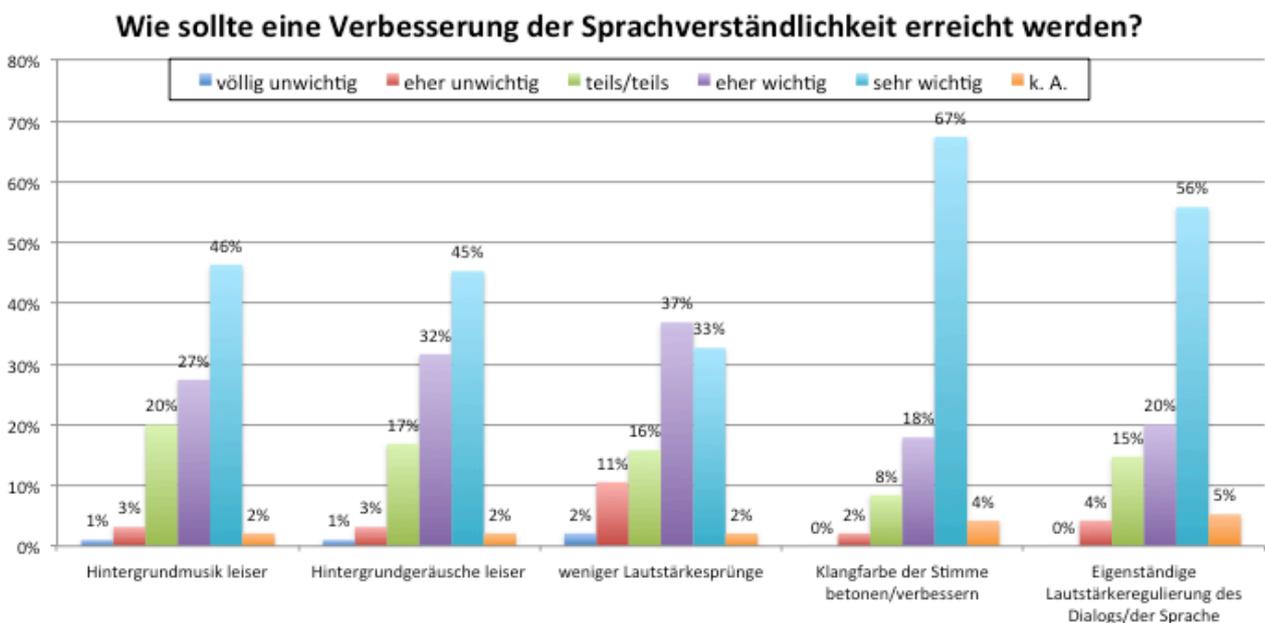


Abbildung 4: Prozentuale Verteilung der Antworten aller 116 Teilnehmer an der Online-Umfrage auf die Frage wie eine Verbesserung der Sprachverständlichkeit im Fernsehen erreicht werden sollte

HbbTV 2.0

Der weit verbreitete HbbTV-Standard („Hybrid Broadcast Broadband TV“) ermöglicht über kombinierte TV- und Breitbandempfangswege, neue Informationen und Dienste am Fernseher anzubieten. Dabei wird neben klassischem Fernsehempfang via DVB auch der Broadband-Kanal über das Internet genutzt. Aktuell stehen über diese Anwendung Dienste wie senderspezifische Mediatheken (VoD), elektronische Programmführer (EPG), News, Zugang zu sozialen Netzwerken und vielerlei weitere Information zur Verfügung.

Die HbbTV-Version 1.0 [3] oder 1.5 [4] ist bereits in nahezu jedem neuen Fernsehgerät integriert. Die neue Version HbbTV 2.0 wurde im Februar 2015 veröffentlicht [8], wobei Geräte mit diesem Standard vermutlich erst Mitte 2016 verfügbar sein werden.

Das wichtigste Features für den in HBB4ALL geplanten Clean-Audio-Dienst ist die Synchronisation von IP-Streams mit DVB-Streams. Dies ermöglicht eine zeitsynchrone Wiedergabe von zusätzlichen über IP verbreiteten Audiostreams mit dem DVB-Video-Stream. Darüber hinaus eröffnet dies viele Möglichkeiten für zusätzliche A/V-Anwendungen wie beispielsweise die Bereitstellung von Gebärdensprache, Originalton, Audio-Deskription und Clean Audio.

Pilot „HbbTV 2.0 live“ – Workflow

Mittels des Clean-Audio-Dienstes sollen dem Endnutzer mehrere verschiedene Clean-Audio-Spuren zur Verfügung gestellt werden. Ein Überblick über den geplanten Workflow der Clean-Audio-Anwendung ist in Abbildung 5 zu sehen.

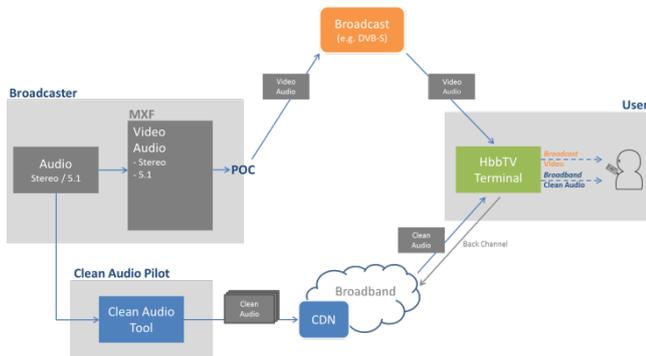


Abbildung 5: Workflow des Clean-Audio-Piloten für die Live-Anwendung im Standard HbbTV 2.0

Zuerst wird das produzierte Stereo-, oder 5.1-Ausgangsmaterial (ohne Video-Spur) vom Sender an das Clean-Audio-Tool übergeben. Dieses erstellt automatisch die verschiedenen Clean-Audio-Spuren. Diese verschiedenen Audiospuren werden dann in einem MPEG-DASH-Container übertragen. Dieser Container wird zum CDN (Content Delivery Network) gesendet, wo er dann zum Abruf bereit steht.

Die DVB-Video-, sowie Audio-Spur wird über den normalen Broadcastkanal (z.B. DVB-S) ausgestrahlt. Falls Clean-Audio-Spuren zur Verfügung stehen, erkennt die HbbTV-Anwendung dies und bietet dem Nutzer den Dienst zur Auswahl an. Falls sich der Nutzer dafür entscheidet, eine Clean-Audio-Spur anstelle der normalen DVB-Ton-Spur zu

hören, wird die HbbTV-Anwendung den passenden bzw. angeforderten Stream aus dem DASH-Container vom CDN-Server anfordern. Die normale DVB-Ton-Spur wird darauf durch eine Clean-Audio-Spur ersetzt und zeitsynchron zum DVB-Video wiedergegeben.

Clean Audio Tool

Das Clean Audio Tool kann grundsätzlich auf zwei verschiedene Weisen arbeiten, nämlich im Basic Mode und im Advanced Mode. Der einzige Unterschied zwischen den beiden Varianten ist, dass der Basic im Gegensatz zum Advanced Mode ohne Frequenzbearbeitung arbeitet.

Basic Mode

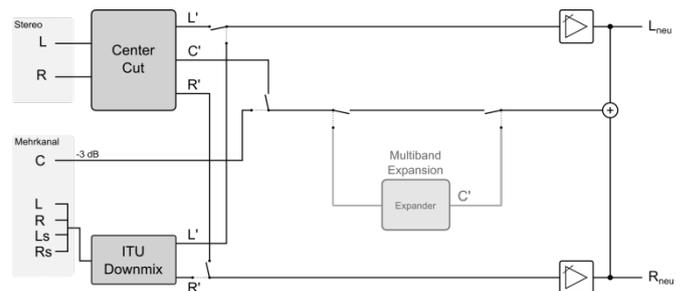


Abbildung 6: Signalkette des Clean Audio Tools im Basic Mode

Die grundlegende Signalkette des Basic Modes ist in Abbildung 6 zu sehen. Als erstes muss der Sprachanteil im Signal vom Rest getrennt werden. Falls nur Stereo-Material zur Verfügung steht, wird der sogenannte Center-Cut-Algorithmus [5] angewandt, um die Sprachanteile zu extrahieren. Dieser Algorithmus nutzt eine Korrelationsanalyse, um die Stereo-Mitte von den restlichen Signalanteilen zu trennen. Da die Sprache bei einem typischen Stereosignal in die Mitte des Stereopanoramas gemischt sein sollte, ist der Center-Cut-Algorithmus eine effektive Methode um die Sprachanteile bei einer entsprechenden Stereomischung zu extrahieren. So erhält man aus den beiden Stereo-Kanälen einen Center (C'), sowie einen neuen linken (L') und rechten (R') Kanal.

Für 5.1-Material ist keine zusätzliche Extraktion der Sprachanteile notwendig, da der Center-Kanal bereits vorhanden ist und normalerweise nur Sprache enthalten sollte. Der neue linke sowie rechte Kanal (L' und R') wird über einen ITU-Downmix aus den ursprünglichen Kanälen L, R, L_s und R_s erzeugt.

Danach erfolgt die weitere Bearbeitung für Stereo- und 5.1-Input synchron. Die Lautstärke der erzeugten Kanäle L' und R' wird um einen Faktor, beispielsweise -9 dB verringert. Zuletzt wird das Center-Signal wieder mit den im Pegel reduzierten Kanälen L_{neu} und R_{neu} zusammengeführt.

Optional kann auf den Center-Kanal vor der Zusammenführung noch eine Multiband-Expansion angewandt werden. Darauf wird in der Beschreibung des Advanced Mode noch näher eingegangen.

Advanced Mode

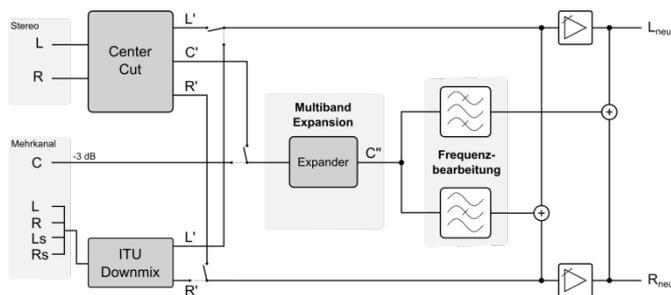


Abbildung 7: Signalkette des Clean Audio Tools im Advanced Mode

Die Signalkette des Advanced Mode ist in Abbildung 7 dargestellt. Der erste Teil in dieser Kette entspricht genau dem Vorgehen des Basic Mode des Clean Audio Tools. Der Unterschied ist, dass hier vor der Pegelreduktion der erzeugten linken und rechten Kanäle (L' und R') der Center-Kanal C' zwei zusätzlichen Bearbeitungen unterzogen wird.

Zuerst wird auf die sprachrelevanten Frequenzen eine Multiband Expansion angewandt, um die Sprachanteile im Center-Kanal zu verstärken. Gleichzeitig werden die Geräuschteile im Signal verringert. Erste Untersuchungen zeigten, dass dieses Vorgehen die Sprachverständlichkeit zumindest leicht verbessert [6].

Im zweiten Schritt wird jeweils ein Bandpass und ein Bandstop Filter auf das Center-Signal angewendet. Diese beiden Filter werden durch eine Kombination aus einem Hochpass und einem Tiefpass mit den gleichen Grenzfrequenzen (1,6 kHz und 8 kHz) realisiert. Der per Bandstop gefilterte Teil des Center-Signals wird zu L' und R' addiert und anschließend der Pegel der beiden Kanäle, wie schon im Basic Mode, reduziert. Der Bandpass gefilterte Teil wird mit den bereits im Pegel reduzierten L' und R' zusammengeführt. So ergeben sich wieder die fertig bearbeiteten Stereo-Kanäle L_{neu} und R_{neu} . Diese zusätzliche Frequenzbearbeitung dient dazu, besonders die für die Sprache relevanten Frequenzanteile noch hervorzuheben.

Hörversuche

Um die durchgeführten Bearbeitungsschritte zur Verbesserung der Sprachverständlichkeit zu testen, wurden bisher in zwei Testrunden Hörversuche durchgeführt. Bei beiden Runden wurde mit zwei unterschiedlichen Probanden-Gruppen getestet. Eine Gruppe bestand aus Hörbeeinträchtigten, die auf keine Hörhilfe angewiesen sind, die andere Gruppe aus Hörgeräteträgern.

Die Versuchsbeschreibung, Durchführung und Auswertung der ersten Testrunde vom Juli 2014 können in [7] nachgelesen werden.

Bei der zweiten Testrunde im Februar/März 2015 bestanden beide Probanden-Gruppen, mit und ohne Hörhilfe, aus je 27 Teilnehmern. Der Versuch mit den Nicht-Hörgeräteträgern wurde am IRT und der mit den Hörgeräteträgern in Zusammenarbeit mit dem rbb (Radio Berlin Brandenburg) in Berlin durchgeführt. Untersucht wurden lediglich Stereo-Signale als Ausgangsmaterial, da sich in der ersten Testrunde schon

gezeigt hat, dass mit 5.1-Material deutlich einfacher eine verbesserte Sprachverständlichkeit erreicht werden kann.

Das untersuchte Material war für beide Test-Gruppen unterschiedlich, da sich in der ersten Testrunde auch gezeigt hat, dass Teilnehmer mit Hörhilfe den Basic Mode (ohne Frequenz- und Dynamikbearbeitung) bevorzugen, Teilnehmer ohne Hörhilfe hingegen den Advanced Mode. Dies kann vermutlich darauf zurückgeführt werden, dass eine Kompensation des Hörverlustes bereits durch Hörgeräte angestrebt wird. Dabei wird in den meisten Fällen auch eine Anhebung von Frequenzbändern vorgenommen. Eine zusätzliche Anhebung durch den Advanced Mode könnte dadurch deutlich extremer wirken als für Nicht-Hörgeräteträger.

Da die zweite Testrunde erst Anfang 2015 durchgeführt wurde, sind die Ergebnisse noch nicht vollständig ausgewertet.

Zusammenfassung und Ausblick

Ein Ansatz, verschiedene Clean-Audio-Versionen automatisch zu erzeugen und mithilfe von HbbTV 2.0 optional zu übertragen, wurde vorgestellt. Die Ergebnisse der Online-Umfrage zeigen, dass bereits Faktoren wie undeutliche Aussprache und zu schnelles Sprechen einen erheblichen Einfluss auf die Sprachverständlichkeit darstellen und eine grundlegende Verbesserung bereits bei der Produktion umgesetzt werden muss.

Eine nachträgliche Verbesserung der Sprachverständlichkeit speziell für Hörgeschädigte ist über die Clean-Audio-Anwendung gut möglich. Im nächsten Schritt gilt es nun die passende Varianten für den geplanten Pilotdienst festzulegen.

Literatur

- [1] HBB4ALL: <http://www.hbb4all.eu/>
- [2] Sprachverständlichkeit im Fernsehen - Empfehlung für Programm und Technik: <https://www.irt.de/de/publikationen/technische-richtlinien.html>
- [3] HbbTV 1.0 ETSI standard document: https://www.hbbtv.org/pages/about_hbbtv/HbbTV-errata-2.pdf
- [4] HbbTV 1.5 ETSI standard document: http://www.etsi.org/deliver/etsi_ts/102700_102799/102796/01.02.01_60/ts_102796v010201p.pdf
- [5] VirtualDub: The "center cut" algorithm. <http://www.virtualdub.org/blog/pivot/entry.php?id=102>
- [6] F. Matzura, S. Goossens, J. Groh und E. Wilk: Verbesserung der Sprachverständlichkeit von fertig gemischtem Stereo-Fernsehton. DAGA, Oldenburg, 2014.
- [7] Weitnauer M., Bohne M.: Automatic and customizable improvement of the speech intelligibility from TV signals for hearing impaired people. TMT, Köln, 2014
- [8] HbbTV 2.0 specifications document: https://www.hbbtv.org/pages/about_hbbtv/HbbTV_specification_2_0.pdf