

KONZEPT UND EVALUATION EINES SOFTWARESYSTEMS ZUR UNTERSTÜTZUNG DER CRM-BASIERTEN SPRECHWIRKUNGSUNTERSUCHUNG

Thorben Frank Jahnke¹, Corinna Sonnen², Mathias Walther³

¹Pengulabs, ²amt – Gesundheitsakademie im Vest, ³Technische Hochschule Wildau
corinna.sonnen@gmail.com

Kurzfassung: Das Continuous-Response-Measurement-Verfahren bildet durch die kontinuierliche Bewertungsmöglichkeit eine wichtige Ergänzung zu den gängigen Methoden im Repertoire der Wirkungsforschung. Um diesen Mehrwert voll ausschöpfen zu können, wird als Verfahrensoptimierung die Entwicklung einer Softwarelösung vorgestellt. Die Überprüfung des optimierten CRM-Verfahrens erfolgt mittels eines Anwendungsfalls aus der sprechwissenschaftlichen Telekommunikationsforschung im Rahmen eines User-Acceptance-Tests. Dabei wird die Funktionalität und Bedienerfreundlichkeit der entwickelten CRM-Softwarelösung unter Beachtung der für die Sprechwirkungsforschung relevanten Kriterien in Form einer A-B-Studie getestet.

Das Gesamtergebnis des User-Acceptance-Tests fällt für die Software *Evalve* positiv aus. Mit Hilfe der Verfahrensoptimierung des CRM-Verfahrens ist eine variabel einsetzbare und damit vielfältig nutzbare CRM-Softwarelösung entstanden.

1 Hintergrund

Wirkungsuntersuchungen stammen ursprünglich aus der Sozialpsychologie und verfolgen das Ziel, die Wirkung von Programmen im Radio, Produkten im Handel oder von Personen, wie z. B. Politikern, zu beobachten. Die Sprechwissenschaft begann Ende der 1960er Jahre, sich systematisch mit dem Begriff der Wirkung und damit auch der Erfassung der Wirkung von Sprache und Sprechen auf den Menschen zu beschäftigen [1].

Eine Problematik bei der Erhebung repräsentativer Daten in diesem Bereich resultiert aus der starken Subjektivität bei Wirkungseinschätzungen. Eine zusätzliche Schwierigkeit stellt der teilweise große Bewertungsunterschied zwischen Experten- und Laiengruppen dar. Zudem ist es für die Hörenden oft nicht leicht, die im Verlauf eines Gesprächs auftretenden Irritationen direkt in Worte zu fassen oder später einer konkreten Stelle zuzuordnen. Die Bewertung erfolgt aus der Retrospektive und spiegelt den Gesamteindruck wider, welcher zwangsläufig vom Ergebnis beeinflusst ist.

Bei dem CRM (Continuous Response Measurement)-Verfahren handelt es sich um eine Möglichkeit, die Wirkung eines kontinuierlichen Stimulus, in der Regel ein audiovisuelles oder rein auditives Signal, auf einen Probanden in Echtzeit zu erfassen [2]. Es geht um subjektive Prozesse, wie z. B. die Wirkung in Bezug auf die Beurteilung des Materials, ausgelöste Emotionen oder bestimmte Wahrnehmungen des Probanden, welche begleitend zur Rezeption erfasst werden sollen. Das Verfahren bildet durch die kontinuierliche Bewertungsmöglichkeit eine wichtige Ergänzung zu gängigen Methoden im Repertoire der Wirkungsforschung.

Die zeitlich parallele und kontinuierliche Vorgehensweise ermöglicht dabei dank der intuitiven Bedienung ein Erfassen der noch unreflektierten und damit affektiven Bewertung. Zeitgleich können wirkungsauslösende Passagen zuverlässig extrahiert werden [3].

Forschungsarbeiten bestätigen eine interne Validität und Reliabilität des Verfahrens (u. a. [2]) und es konnte ebenfalls bereits eine externe Validität nachgewiesen werden (u. a. [4]).

Aktueller Stand der Technik ist das CRDI (Continuous Response Digital Interface), ein Hardware-Software-Bundle mit elektromechanischen Bedienelementen [5], welches in *Abbildung 1* zu sehen ist.

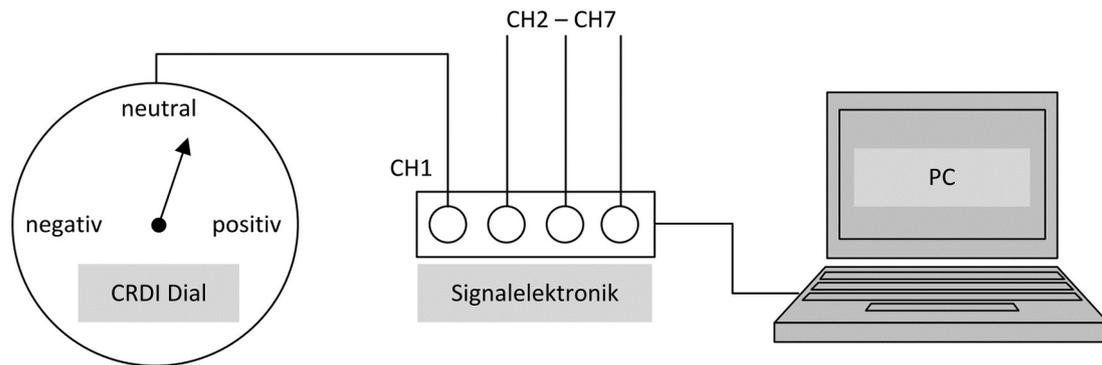


Abbildung 1 - Schematische Versuchsanordnung mit CRDI-Dial [5]

Durch Probleme in der Handhabung für Probanden und Untersuchungsleiter ist es schwierig, den Mehrwert für Wirkungsforschungen voll auszuschöpfen. So müssen z. B. Audiodatei und Bewertungssoftware separat synchron gestartet werden und die Erfassung von lediglich acht Eingabegeräten gleichzeitig führt zu Problemen bei quantitativ angelegten Studien [6]. Auch eine Nutzung der aktuell existierenden Softwarelösungen ist für die weiterführende Sprechwirkungsforschung nur sehr eingeschränkt möglich, da sie z. T. in Handhabung und Design veraltet sind. Damit entsprechen sie nicht den aktuellen Nutzungsgewohnheiten der Gesellschaft. Zudem bieten die Softwarelösungen keine ausreichende Möglichkeit, variable Forschungsinteressen zu bedienen und große Probandengruppen online zu erreichen. Viele existierende Systeme benötigen vorkonfigurierte Endgeräte, beispielsweise speziell präparierte Handys, was den Aufwand zur Durchführung quantitativer Studien sowohl von den Kosten als auch von der benötigten Zeit für Mehrfachtestungen wegen der begrenzten Geräteausgabe stark erhöht [3].

Die notwendige Konsequenz ist die Erstellung einer eigenen Software, welche spezifische Anforderungen erfüllt, die im folgenden vorgestellt werden.

2 Entwicklung der CRM-Online-Softwarelösung *Evalue*

2.1 Zielsetzung der Entwicklung

Die zugrundeliegende Idee der CRM-Softwarelösung basiert auf der Anforderung, dass die Probanden die Durchführung der Echtzeit-Bewertung einer Audioaufnahme, in Zukunft eventuell erweiterbar auf Videoaufnahmen, selbstständig von einem heimischen System, in erster Programmausführung einem Desktop-PC oder Laptop, aus starten können. Dabei muss die Audiodatei bzw. später eventuell die Videodatei komplett vorgeladen werden, so dass ein Abspielen ohne Verzögerung gewährleistet ist. Parallel zum Abspielen des Audios wird dann die Benutzereingabe aufgezeichnet.

Weiterhin soll die Software modular aufgebaut sein, so dass es möglich ist, über dieses Online-Portal allgemeine Studien zu erstellen und die Audio-Bewertung als Teil einer komplexen Umfrage zu integrieren. Dies ist nötig, da zwecks statistischer Auswertung zusätzliche Daten, wie z. B. demographische Angaben, erhoben werden müssen.

Insgesamt wurde basierend auf bisheriger Forschung ein Lastenheft mit den Anforderungen erarbeitet, die eine CRM-Lösung erfüllen muss, um einen Nutzen für die sprechwissenschaftlichen Wirkungsuntersuchungen zu generieren [3].

2.2 Anforderungen

Die Grundanforderung ist, dass es sich um eine Webanwendung handeln soll, die für Desktopgeräte und alle gängigen mobilen Endgeräte geeignet ist. Dies soll eine große und unkomplizierte Erreichbarkeit für qualitative und quantitative Studien ermöglichen sowie die Durchführung mit Hilfe einer intuitiven Bedienung für die Probanden und die Versuchsleitung erleichtern. Mit der Erstellung einer reinen Webanwendung ist es für die Versuchsleitung nicht nötig, vorher an alle Probanden ein bestimmtes Endgerät auszugeben. Dies führt zu einer Reduzierung des Aufwands und der Kosten. Eine Webanwendung kann im Gegensatz zu einer nativen App von allen gängigen Endgeräten aus bedient werden, egal ob es sich um Desktopgeräte oder Smartphones mit Apple- oder Android-Betriebssystem handelt. *Tabelle 1* gibt eine Übersicht über die Anforderungen.

Tabelle 1 – Auszug aus dem Lastenheft für eine CRM-Softwarelösung [3]

| <i>Allgemeine Betriebsanforderungen</i> | <i>Inhalt/Design</i> |
|---|--|
| <p>Erhalt aussagekräftiger Messdaten</p> <ul style="list-style-type: none"> • Einstellbare Abtastrate zur Beurteilung des Audiomaterials <p>Stabilität der Anwendung</p> <ul style="list-style-type: none"> • Keine Abbrüche, z. B. wenn ein einzelner Messwert nicht erfasst werden kann • Audiodatei muss durchlaufen und darf nicht durch Ladezeiten blockiert werden <p>Implementiertes Audioabspielprogramm</p> <ul style="list-style-type: none"> • Audio- und Mess-Spur verlaufen synchron • Referenzmedium: Audiodateien im mp3-Format <p>Exportfunktion der Daten in ein gängiges Format, welches z. B. mit MS Excel geöffnet werden kann</p> <p>Software ist einfach zu warten und weiterentwickelbar</p> <p>Systemübergreifend geeignet</p> <ul style="list-style-type: none"> • Firefox, Chrome, Edge • Browser müssen HTML5-Standard und JavaScript unterstützen <p>Einbettung in übergreifendes Datenerfassungssystem</p> <ul style="list-style-type: none"> • Kombination mit Fragebögen/Anleitung • Erstellung mehrerer Studien gleichzeitig • Webanwendung ist von überall her zugänglich • Daten werden auf dem Server gespeichert, damit sie für die Studien abrufbar sind | <p>Benutzeroberfläche</p> <ul style="list-style-type: none"> • Auswahl zwischen ein- und zweidimensionalen Bewertungsskalen durch die Umsetzung mit Hilfe eines digitalen Sliders oder eines Koordinatensystems • Bedienung mittels Touchscreens, Touchpads oder Maus, ggf. Tastatur • Beschriftung der Koordinaten/Skalenendpunkte frei wählbar • Benutzerbestimmbares Starten der jeweiligen Audiodatei und Messung <p>Moderne UX</p> <ul style="list-style-type: none"> • Keine Installation externer Dienste/Anwendungen/Treiber auf dem Endgerät • Webanwendung soll möglichst intuitiv, zeitgemäß und übersichtlich sein • Visuelle Rückmeldung an den Anwender, an welcher Position er sich befindet, z. B. über hinterlegtes Farbsystem |

2.3 Entwicklung des funktionsfähigen Prototyps der Software

Basierend auf dem Lastenheft und einem von der Firma *Pengulabs* realisierten Klick-Dummy wurde von Studierenden der Technischen Hochschule Wildau die Software *Evalue* entwickelt, welche die neuen Anforderungen umsetzt. Das Projekt wurde im Januar 2021 durch die Fertigstellung der Webapplikation *Evalue* als Minimum-Viable-Product (MVP) für die Nutzung zur Erstellung von Experimenten erfolgreich beendet. Zusätzlich wurde eine detaillierte Anleitung der Software in Form eines Benutzerhandbuchs erstellt [7]. Eine Übersicht des Designs gibt der Screenshot in *Abbildung 2*.

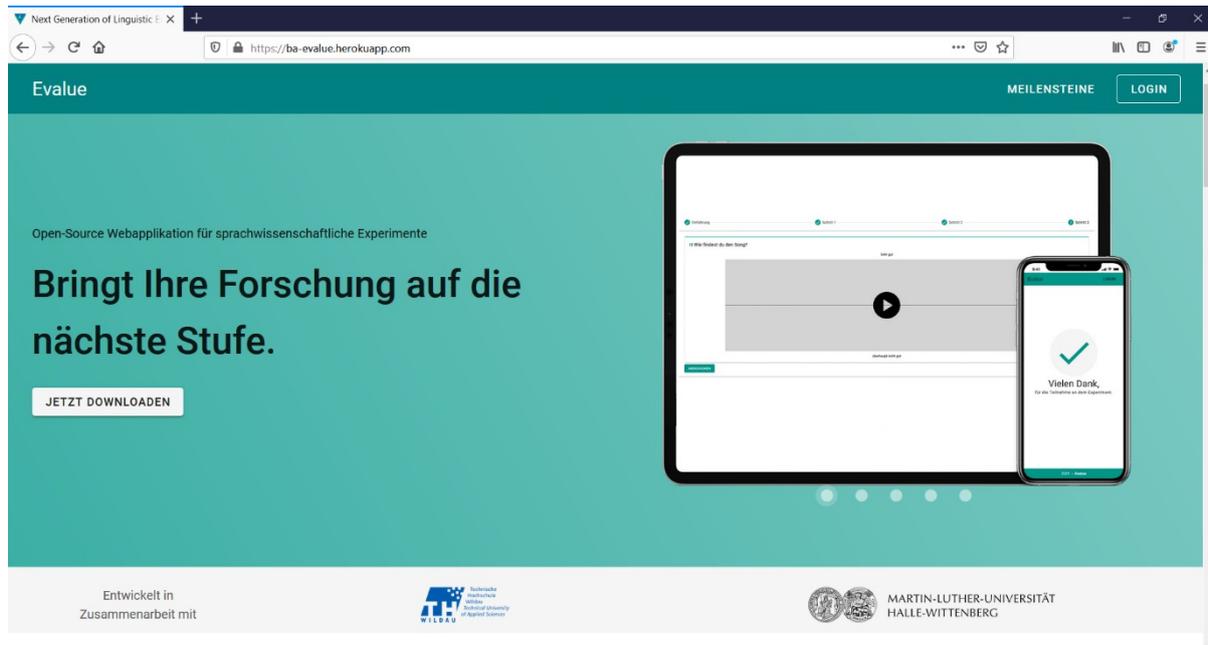


Abbildung 2 - Screenshot der Software *Evalue* [8]

2.4 Abgleich der Anforderungen mit deren Umsetzung bei *Evalue*

2.4.1 Erfüllte allgemeine Betriebsanforderungen

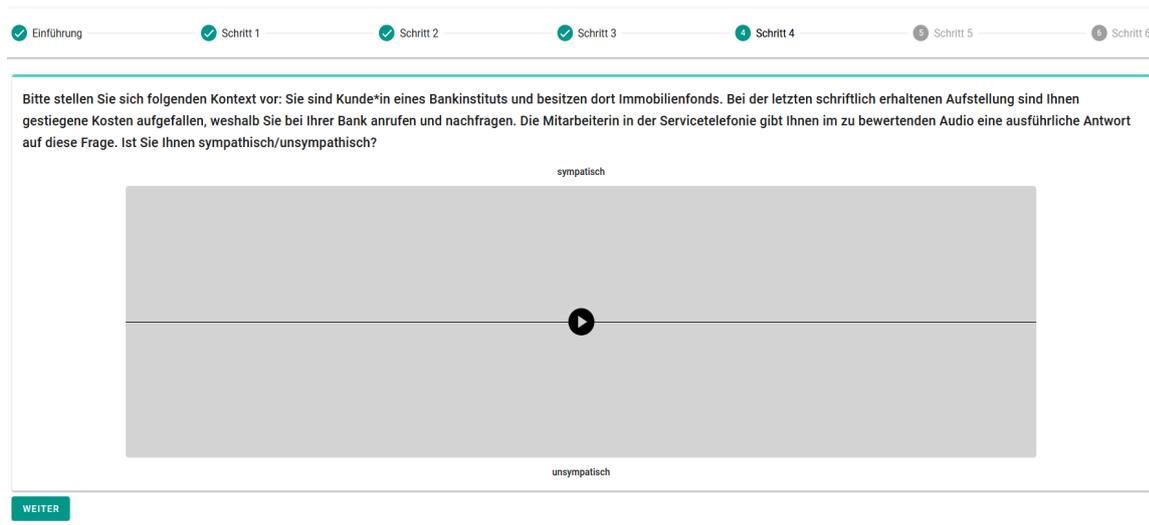
Evalue wurde mit modernen Bibliotheken und Frameworks entwickelt. Für die Darstellung im Browser kommt *vue.js* zum Einsatz. Auf der Serverseite ist die Logik mit *node.js* umgesetzt und wird durch eine PostgreSQL-Datenbank unterstützt. Daher ist die Anwendung bei Nutzung verschiedener Browser funktionsfähig. Audiospur und Bewertung starten synchron, wenn die Probanden den Play-Button in der Mitte des Bewertungsfeldes drücken (siehe *Abb. 3*). Zum Erhalt aussagekräftiger Messdaten ist die empfohlene Abtastrate von 1 pro Sekunde umgesetzt. Als Referenzmedium können im implementierten Audioabspielprogramm Audiodateien im mp3-Format mit bis zu 10 MB hochgeladen und eingebunden werden.

Die Einbettung in ein übergreifendes Datenerfassungssystem wurde durch die individuelle Kombinationsmöglichkeit einzelner Module realisiert. Dazu gehören neben dem CRM-Tool Module mit Anleitung und weiteren klassischen Umfrage-Tools.

Es existiert keine Beschränkung der Experimente pro registriertem Experimentleitenden. Die Teilnahme über den Link ist für jeden Probanden im Internet möglich und das Datenbanksystem ist auf dem jeweiligen Server hinterlegt.

Die Exportfunktion ist im Administrationsbereich enthalten und über CSV-Dateien realisiert, welche z. B. mittels MS Excel problemlos weiterverarbeitet werden können.

35. Konferenz Elektronische Sprachsignalverarbeitung



2022 – Evalue

Abbildung 3 - Screenshot-Ausschnitt des Audiobewertungsfensters aus der Probandensicht [9]

2.4.2 Erfüllte Anforderungen in Bezug auf Inhalt und Design

Bei der Bewertung der Audiodatei steht die Fragestellung über dem Bewertungsfeld und die beiden Skalen sind mit den passenden gegenüberstehenden Polaritäten gekennzeichnet (s. *Abb. 3*). Durch das Klicken des Play-Buttons starten die Audiobewertung und die Messung. Die Probanden können die Maus nun in dem Feld frei nach oben und unten zu einer der beiden gekennzeichneten Polaritäten bewegen und damit ihre Bewertung für die Audiodatei abgeben, wobei die farblich variable Hinterlegung des Feldes der Visualisierung dient. Der Start-Button ist so platziert, dass die Probanden direkt im neutralen Bereich beginnen.

Die Beschriftung der Y-Achsen-Koordinatenpunkte ist frei wählbar. Die Bedienung erfolgt mittels Maus oder Touchpad. Nach dem Start werden die Cursor-Bewegungen des Probanden auf der Y-Achse aufgezeichnet. Die Aufzeichnung basiert auf der reinen Mausbewegung, so dass die Maustaste nicht durchgängig betätigt werden muss. Die Messwerte ergeben sich aus dem Abstand des Mauszeigers von der Null-Geraden (siehe *Abb. 3*). Die Abweichungen werden mit 1 Hz Abtastrate ermittelt und auf die Skaleneinteilung abgebildet, z. B. auf das Intervall $[-1;1]$ oder $[0;1]$. Sobald der Cursor das Feld zu weit verlässt, stoppt das Abspielen und beginnt dann ggf. wieder von vorne. Dieser Mechanismus ist sinnvoll, damit sichergestellt ist, dass die Probanden aufmerksam mitmachen und die Anwendung nicht im Hintergrund laufen lassen können. Nach Beenden der Audiodatei schließt sich das Bewertungsfenster automatisch. Die Messwerte sind dann als Zeitreihe gespeichert und können dementsprechend ausgewertet werden.

Die Webanwendung ist übersichtlich und zeitgemäß durch ansprechendes Farbdesign, angezeigte Schritte für die Probanden und Experimentleitenden sowie die variable Anpassung der Größenverhältnisse und des Seitenaufbaus ans Endgerät. Die Bedienung ist intuitiv umgesetzt und die farbliche Umrandung des Bewertungsfeldes sorgt durch eine visuelle Rückmeldung an den Nutzer für eine bessere UX. Der Administrationsbereich für die Experimentleitenden ist ebenfalls ansprechend und übersichtlich gestaltet, was in *Abbildung 4* zu sehen ist.

3 User-Acceptance-Test

3.1 Vorstellung des Use-Case

Damit die Software bezüglich aller gestellten Anforderungen auf ihre Funktionalität überprüft werden kann, ist es notwendig, einen konkreten Anwendungsfall zur Verfügung zu stellen. Um für die CRM-Bewertungen der Probanden einen sinnvollen Kontext zu schaffen, wird als Fallbeispiel eine Beschwerdesituation aus der Inbound-Servicetelefonie des Finanzwesens ausgewählt. Da es sich bei Telefonaten um rein auditive Daten handelt, kann hierbei besonders gut die Wirkungsweise des Sprechausdrucks untersucht werden.

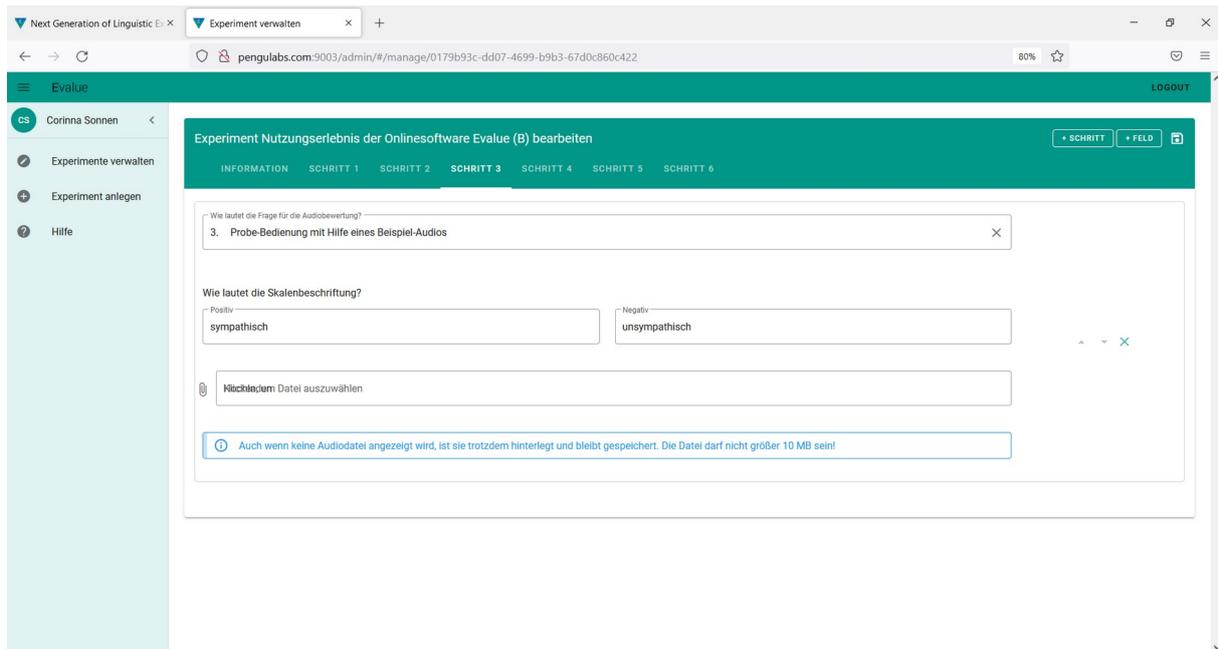


Abbildung 4 - Konfiguration der Skalen für Experimentleitende [9]

Aus dem Bereich der Servicetelefonie muss ein Gebiet gewählt werden, mit dem möglichst viele Menschen aus verschiedenen Altersgruppen Berührungspunkte haben. Es ist davon auszugehen, dass die Nutzung einer Bank-Hotline für Rückfragen oder in Beschwerdesituationen ein realistisches Szenario für den Durchschnitt der Bevölkerung darstellt.

Die Probanden hören eine Audiodatei mit der ausführlichen Antwort der Servicekraft auf eine fiktive Beschwerde. Sie sollen sich vorstellen, dass sie als Kunde auf ihre Rückfrage eine Antwort erhalten. Während des Hörens bewerten die Probanden die Servicekraft aus der Kundenperspektive durch das Bewegen des Mauszeigers zwischen den Skalenpunkten *sympathisch* – *unsympathisch* [3].

Eine alltagsnahe Bezeichnung der Skalenwerte ist wichtig [6] und Sympathiezuschreibungen sind im Alltag ein gängiger, ständig ablaufender Vorgang, welcher bei Experten und Laien regelmäßig mit Stimme und Sprechweise in Verbindung gebracht wird [10]. Zusätzlich hat sich das Gegensatzpaar *sympathisch* - *unsympathisch* als valide und reliabel für konsistente Hörerbewertungen herausgestellt [11].

Damit sich die Probanden persönlich angesprochen fühlen, wurde im Use-Case bewusst auf einen Dialog verzichtet und damit die Interaktivität eines Gesprächs nicht berücksichtigt.

3.2 Methodik des User-Acceptance-Tests

Der User-Acceptance-Test umfasst die Untersuchung der Sprechwirkung mittels des CRM-Bewertungstools der Online-Software *Evaluate* anhand einer Sprachaufzeichnung.

Im Rahmen dessen wird eine A-B-Studie mit zwei verschiedenen Audiodateien durchgeführt.

Die Studie A enthält die sachlich, neutral gesprochene Audiodatei, während die Studie B die Audiodatei mit Variationen ausgewählter Sprech- und Stimmausdrucksmerkmale durch Manipulation mittels Software enthält.

Neben der CRM-Bewertung beinhaltet die Studie zusätzliche Fragen zum Erfassen der personenbezogenen Daten, der Abfrage der retrospektiven Wirkung des Audiomaterials auf den Probanden sowie Fragen zum Nutzungserlebnis der Software.

Zusätzlich wurde eine Video-Anleitung, welche die Bedienung des Bewertungstools erklärt, und eine Probe-Bedienung mit Hilfe eines Beispiel-Audios integriert, damit die Anwender sich mit der Nutzung der Software vertraut machen können [3].

4 Ergebnisse

Insgesamt haben 54 Probanden am User-Acceptance-Test teilgenommen. Das Alter der Teilnehmenden verteilt sich schwerpunktmäßig auf die Altersgruppen der 20-29-, 30-39- und 50-59-Jährigen. Insgesamt haben 39 weibliche und 15 männliche Probanden teilgenommen.

Das Gesamtergebnis der Befragung zur Bedienfreundlichkeit und dem Nutzungserlebnis von *Evalue* fällt positiv aus:

Rund 98% der Teilnehmenden finden das Anleitungsvideo verständlich gestaltet und das Beispiel zum Ausprobieren der Software ausreichend, um die Bedienung korrekt zu erfassen. Dabei haben 77,78% die Frage zum Anleitungsvideo und 87,04% die Frage zur Beispielbewertung mit *trifft sehr zu* bewertet. In der Kategorie *Design & Erfassung* wird überprüft, ob die Online-Software ein ansprechendes Design hat, die Beschriftung der Achsen deutlich zu erfassen und die visuelle Rückmeldung über das Farbsystem hilfreich bei der Orientierung ist. Insgesamt finden über 90% der Teilnehmenden alle drei Punkte zutreffend.

In der Kategorie *Technik & Bedienung* ist aufgeschlüsselt, ob die Bedienung als intuitiv empfunden wird, die Software den aktuellen Nutzungsgewohnheiten entspricht und das Abspielen ohne Verzögerung oder Abbrüche erfolgt. *Abbildung 5* gibt dazu eine Übersicht [3].

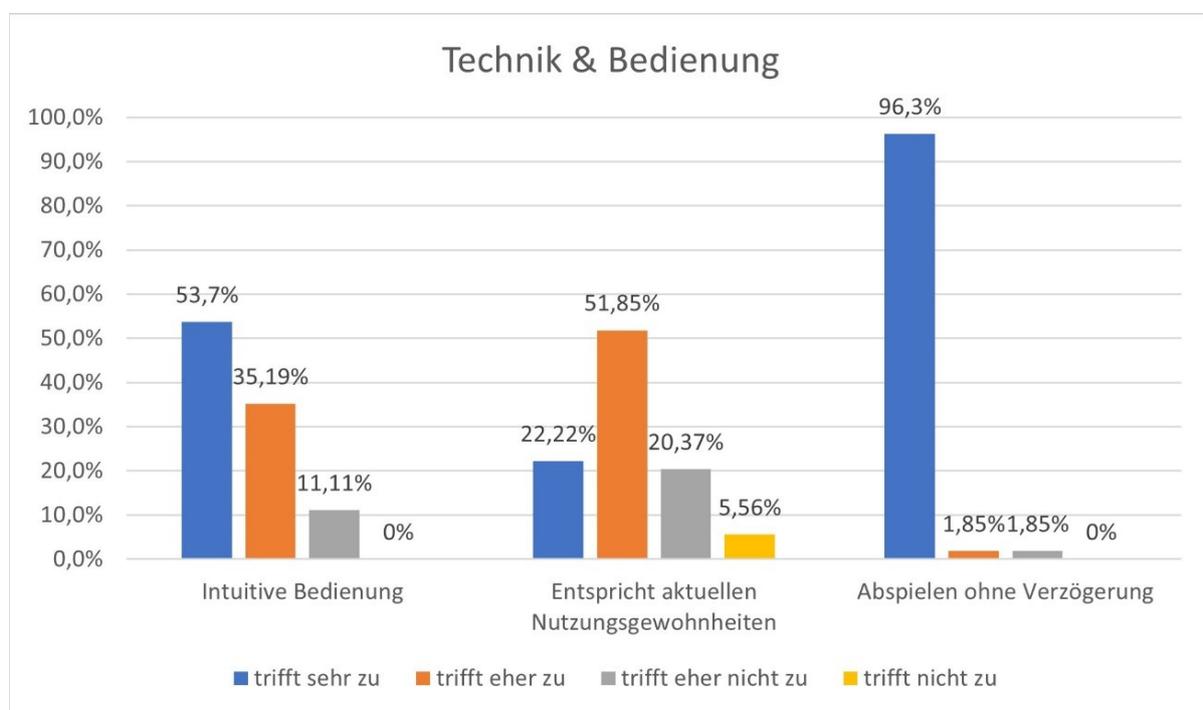


Abbildung 5 - Ergebniszusammenfassung der Bewertungskategorie Technik & Bedienung im UAT zur Software *Evalue* [3]

Aus Sicht des Experimentleitenden erweist sich die Software ebenfalls als praxistauglich und valide. Die erhobenen Daten wurden ohne Probleme exportiert. Auf ihrer Basis konnte dann sowohl ein graphischer Vergleich der entstandenen Bewertungskurven als auch eine statistische Auswertung erfolgen. Mit Hilfe der Verfahrensoptimierung des CRM-Verfahrens ist eine variabel einsetzbare und damit vielfältig nutzbare CRM-Softwarelösung entstanden.

Die Software bietet durch die Kombination mit Fragebögen und die freie Wahl der Fragestellung und Achsenbeschriftung beim CRM-Tool vielfältige Einsatzmöglichkeiten. Damit ist die praktische Grundlage für weitere großflächig angelegte Wirkungsforschungen in der Sprechwissenschaft, aber auch in allen verwandten Wissenschaftsdisziplinen, wie z. B. Musikwissenschaft und Medienforschung, geschaffen.

Danksagung

Vielen Dank an das studentische Team der Technischen Hochschule Wildau, bestehend aus Christopher Dethloff, Nico Richter und Niklas Schwanitz, das *Evalue* im Rahmen eines einjährigen Studienprojekts entwickelt hat.

Literatur

- [1] HIRSCHFELD, U., STOCK, E.: *Zur Entwicklung der Sprechwissenschaftlichen Phonetik an der Universität Halle*. In: U. HIRSCHFELD, H.-J. SOLMS, E. STOCK (Hrsg.): *Otto Bremer – Wegbereiter der sprechwissenschaftlichen Phonetik an der Universität Halle*, S. 11 – 46, Frank & Timme, Berlin, 2016.
- [2] BACHL, M.: *Analyse rezeptionsbegleitend gemessener Kandidatenbewertungen in TV-Duellen. Erweiterung etablierter Verfahren und Vorschlag einer Mehrebenenmodellierung*. epubli GmbH, Berlin, 2014.
- [3] SONNEN, C.: *Verfahrensoptimierung der CRM-basierten Sprechwirkungsuntersuchung am Beispiel der Sprechausdrucksmerkmale in der Servicetelefonie*. Dissertation, Martin-Luther-Universität Halle-Wittenberg, 2022.
- [4] MAIER, J., HAMPE, J. F., JAHN, N.: *Breaking Out Of The Lab. Measuring Real-Time Responses To Televised Political Content In Real-World Settings*. In: *Public Opinion Quarterly*, 80(2), S. 542 – 553, 2016.
- [5] WALTHER, M., BEIER, F., MELLOULLI, T., NEUBER, B.: *Automatische Klassifikation des Ersteindrucks von Sprechereigenschaften bei ungeschulten Hörern*. In: O. JOKISCH (Hrsg.): *Elektronische Signalverarbeitung (ESSV). Tagungsband der 27. Konferenz*. S. 93 – 100, TUDpress, Leipzig, 2016.
- [6] UNGER, A., ALTSTADT, M., GEBAUER, M. L.: *Messung kontinuierlicher Hörerbewertungen mit dem CRDI*. In: I. BOSE (Hrsg.): *Radio, Sprache, Klang. Forschungen zur Radioästhetik und Radioidentität*. S. 239-262, Peter Lang Verlag, Frankfurt a. M. u. a., 2015.
- [7] RICHTER, N., DETHLOFF, C., SCHWANITZ, N.: *Benutzerhandbuch - Erklärung und Darstellung des Einsatzgebietes von Evalue*. 2021.
<<https://git.th-wildau.de/niri6164/evalue/-/blob/master/Benutzerhandbuch.pdf>>
(31.01.2021)
- [8] <<https://ba-evalue.herokuapp.com/>> (30.01.2021)
- [9] <<http://pengulabs.com:9003/proband/#/?id=0179b93c-dd07-4699-b9b3-67d0c860c422>>
(29.01.2022)
- [10] WEISS, B., MÖLLER, S., POLZEHL, T.: *Zur Wirkung menschlicher Stimme auf die wahrgenommene Sympathie - Einfluss der Stimmanregung*. In: H. MIXDORFF (Hrsg.): *Proc. Elektronische Sprachsignalverarbeitung (ESSV)*. S. 56 – 63, TUDpress, Dresden, 2010.
- [11] <<https://gepris.dfg.de/gepris/projekt/211929770/ergebnisse?context=projekt&task=showDetail&id=211929770&selectedSubTab=2&>> (21.03.2020)