Lütters, Holger

**Article**

# Text versus Speech versus Video: Möglichkeiten und Grenzen der Verwendung gesprochener Sprache im digitalen Marktforschungsinterview

PraxisWISSEN Marketing

**Provided in Cooperation with:**
AfM – Arbeitsgemeinschaft für Marketing

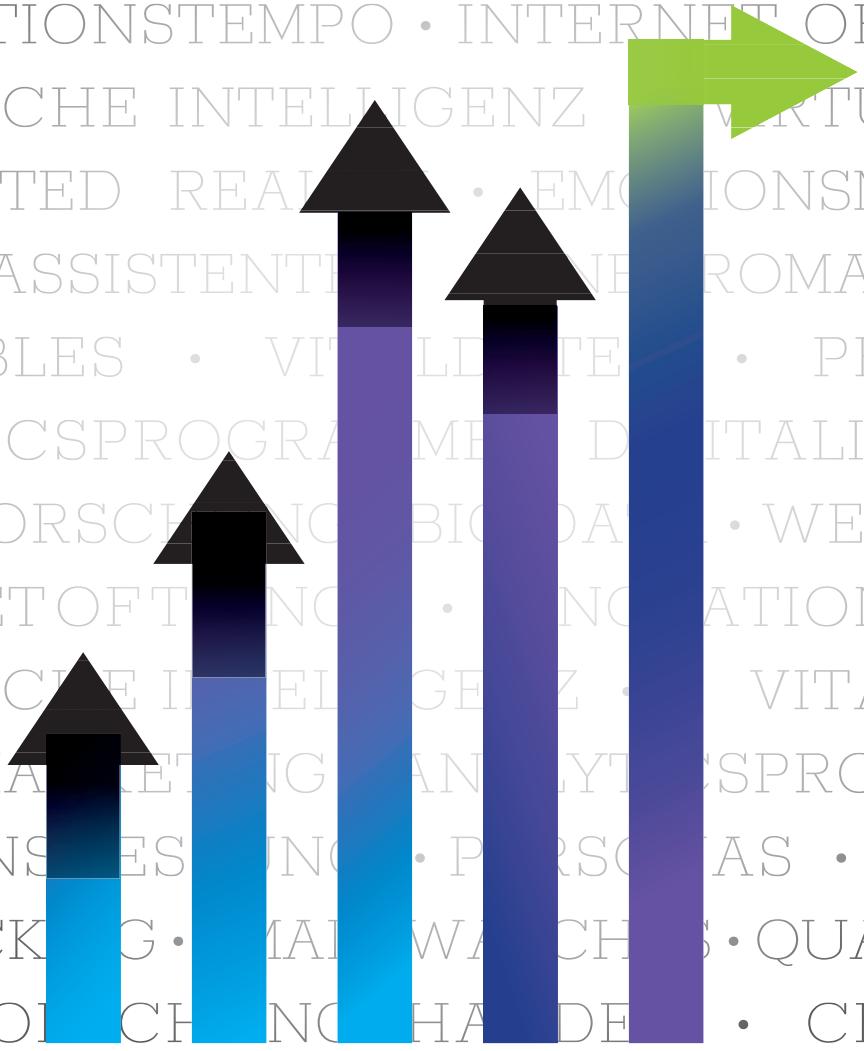
*Suggested Citation:* Lütters, Holger (2020) : Text versus Speech versus Video: Möglichkeiten und Grenzen der Verwendung gesprochener Sprache im digitalen Marktforschungsinterview, PraxisWISSEN Marketing, ISSN 2509-3029, Arbeitsgemeinschaft für Marketing (AfM), Berlin, Vol. 5, Iss. 01/2020, pp. 69-85, https://doi.org/10.15459/95451.42

This Version is available at:
https://hdl.handle.net/10419/261153

# PraxisWisser

## GERMAN JOURNAL OF MARKETING ®

MARKTFORSCHUNG · DIGITALISIERUNG
TECHNISCHER FORTSCHRITT · BIG DATA
INNOVATIONSTEMPO · INTERNET OF THINGS
KÜNSTLICHE INTELLIGENZ · VIRTUAL UND
AUGMENTED REALITY · EMOTIONSMESSUNG
SPRACHASSISTENTEN · NEUROMARKETING
WEARABLES · VITALDATEN · PERSONAS
ANALYTICSPROGRAMME · DIGITALISIERUNG
MARKTFORSCHUNG · BIG DATA · WEARABLES
INTERNETOFTHINGS · INNOVATIONSTEMPO
KÜNSTLICHE INTELLIGENZ · VITALDATEN
NEUROMARKETING · ANALYTICSPROGRAMME
EMOTIONSMESSUNG · PERSONAS · GOOGLE
EYETRACKING · SMARTWATCHES · QUALITATIVE
MARKTFORSCHUNG · CHATBOTS



## Innovation in der Marktforschung

Heft 01/ 2020
ISSN 2509-3029

AfM
Arbeitsgemeinschaft
für Marketing

PraxisWisser
GERMAN JOURNAL OF MARKETING ®

# Innovation in der Marktforschung

# Impressum

PraxisWisser
GERMAN JOURNAL OF MARKETING

**Organ der Arbeitsgemeinschaft für Marketing (AfM)**
**http://arbeitsgemeinschaft.marketing/praxiswissen-marketing**
**ISSN 2509-3029 Heft 1/2020**

# Vorwort

Die **Marktforschung** ist ein vergleichsweise **junges Fachgebiet**, das in seiner Entwicklung bereits eine **Vielzahl von Veränderungen** erfahren hat. Kaum eine Disziplin verändert den eigenen Methodenkanon aufgrund technischen Fortschritts so häufig wie das Handwerk der Marktforschung. Seit dem Aufkommen des Internets hat sich dort das **Innovationstempo**, wie in anderen Marketingdisziplinen auch, **deutlich erhöht**.

In den vergangenen Jahren waren die **Digitalisierung** sowie **Big Data** wichtige Themen. Technische Innovationen wie **Chatbots** werden zumindest testweise zunehmend eingesetzt. **Künstliche Intelligenz, Virtual** und **Augmented Reality** sind weitere Techniken, die das Potenzial haben, die Marktforschung nachhaltig zu wandeln. Die Vernetzung im **Internet of Things** kann der klassischen Marktforschung Konkurrenz machen, indem auch ohne klassische Marktforschung Nutzerdaten gesammelt werden. Auch **Sprachassistenten** können dazu eingesetzt werden.

Die **qualitative Marktforschung** profitiert ebenfalls von der Digitalisierung. So können **Smartphones** mit ihren integrierten Kameras dazu eingesetzt werden. Der technische Fortschritt beflügelt die Forschung unter dem Schlagwort **Neuromarketing**. **Eyetracking und Emotionsmessung** wird **via Webcam** möglich und bringt das Marktforschungslabor in nahezu jeden Haushalt. Einfache Hirnstrommessungen finden über Kopfhörer statt und mit Hilfe von **Smartwatches** und **Wearables** werden Vitaldaten von Menschen zum festen Bestandteil der Forschung. Last but not least sind **Google und Co.** zu nennen, die mit ihren **Analyticsprogrammen** der etablierten Marktforschung Konkurrenz machen.

Diese und weitere Veränderungen wollen wir in dieser Ausgabe von „PraxisWissen Marketing – German Journal of Marketing“ unter dem Titel **„Innovation in der Marktforschung“** analysieren. In acht Beiträgen werden der **Einsatz humanoider Roboter** in der Marktforschung, **qualitative Forschungsmethoden** wie etwa der Einsatz von **Gesichtserkennung** sowie des **Eye Trackings** näher untersucht. Es gibt ein Fallbeispiel aus dem **Handel**, in dem Erkenntnisse des **Neuromarketings** berücksichtigt werden sowie eines aus dem **Tourismus**, in dem **Personas für das nachhaltige Reisen** vorgestellt werden.

Wir bedanken uns ganz herzlich bei allen Autorinnen und Autoren, den Mitgliedern des Herausgeberbeirats und allen anderen Personen, die an der Entstehung dieses Werks beteiligt waren.

Berlin im Oktober 2020

Andrea Bookhagen Andrea Rumler

# Inhalt


7 **Einsatzpotenziale humanoider Roboter in der Marktforschung – eine explorative Analyse unter besonderer Berücksichtigung des Fallbeispiels Pepper**

Kathrin Reger-Wagner
Günter Buerke

21 **Die Anwendung von Gesichtserkennung im stationären Einzelhandel und ihre Auswirkungen auf die Kaufbereitschaft**

Christina Koch
Marcus Simon
Klaus Mühlbäck

41 **Developing ethical consumer personas for the tourism industry: a means-end approach**

Steffen Sahm

53 **Neuromarketing – Grundlagen, Best-Practice-Beispiele aus dem Handel und kritische Würdigung**

Gerd Nufer

69 **Text versus Speech versus Video – Möglichkeiten und Grenzen der Verwendung gesprochener Sprache im digitalen Marktforschungsinterview**

Holger Lütters

87 **The photo-based qualitative interview – potential applications to market research and current challenges**

Anne-Katrin Kleih
Mira Lehberger
Kai Sparke

99 **Empathic market research: The added value of eye tracking data for affective computing UX research**

Alexander Hahn
Katharina Klug
Florian Riedmüller

111 **Automatisierung qualitativer Marktforschung mit Künstlicher Intelligenz**

Annette Hoxtell

PraxisWISSEN Marketing

AfM
Arbeitsgemeinschaft
für Marketing

eingereicht am: 15.11.2019
überarbeitete Version: 05.03.2020

# Text versus Speech versus Video – Möglichkeiten und Grenzen der Verwendung gesprochener Sprache im digitalen Marktforschungsinterview

**Holger Lütters**

Der Beitrag beschreibt die technischen Möglichkeiten und Grenzen des Einsatzes gesprochener Sprache im Marktforschungsinterview. Kern des Beitrags ist die Darstellung der empirischen Studie “Text vs. Speech vs. Video”. Hier wird ein klassisches Interview mit Texteingabe der offenen Antworten mit zwei Experimentalgruppen verglichen, bei denen die Antworten als Sprach- bzw. Videoantwort zu geben waren. Die Art der Interviewführung in der Marktforschung wird sich durch diese neuen technischen Interaktionsformen verändern. Insbesondere die qualitativen Interviewformen werden damit eine digitale Renaissance erleben.

The article describes the technical possibilities and limitations of the use of spoken language in a market research interview. The core of the article is the presentation of the empirical study "Text versus Speech versus Video". A classic interview with text input of open answers is compared with two experimental groups, which had to give the answers as voice response or as video response. The nature of interviewing in market research will change as a result of these new forms of technical interaction. Especially the qualitative interview forms will experience a digital renaissance.

**Prof. Dr. Holger Lütters** ist Professor für International Marketing an der Hochschule für Technik und Wirtschaft Berlin. Seit der Promotion beschäftigt er sich mit dem Internet als Befragungskanal. Sein Forschungsschwerpunkt ist seither die digitale Marktforschung mit allen Facetten innovativer Entwicklungen von der Skalenverwendung über Location Based Research zu neuartigen Formen der Interaktion. Die Studie, die diesem Beitrag zugrunde liegt, wurde 2019 mit dem Best Practice Award der Deutschen Gesellschaft für Onlineforschung ausgezeichnet.
Holger@Luetters.com

# 1. Erhebung von Sprache im digitalen Interview

Die letzten Jahre haben erhebliche Veränderungen im Umgang des Menschen mit Technologien über neue Interfaces gebracht. So sind gegenwärtig Geräte aus dem Hause Amazon unter dem Namen Alexa, Geräte der Marke Google Home, die Assistentin Siri von Apple und auch das System Cortana von Microsoft geläufig. Aber auch andere Anbieter arbeiten derzeit an Geräten, bei denen Sprache im Fokus der Interaktion steht.

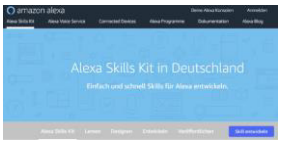
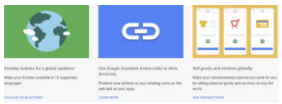
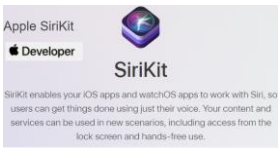
Während einige Anbieter auf eigene Hardwareangebote setzen, ist der Kampf um die Marktführung bereits als Softwareimplementierung gestartet. Abbildung 1 zeigt eine Auswahl der wichtigen Anbieter im Jahr 2019. Die neueren Versionen werden zusätzlich teilweise mit Bildschirmen und Kameras ausgestattet.


**Abb. 1** Markt der Systemanbieter von Sprachassistenzsystemen (Auswahl)

Aus Sicht der Marktforschung stellt sich die Frage, wie derartige Technologien zu Zwecken der Marktforschung verwendet werden könnten.

## 1.1 Marktforschung mit Sprachassistenzsystemen

Die Anbieter der bisherigen Assistenzsysteme bieten zur Entwicklung von Anwendungen eigene Funktionsumwelten an. Entwickler können also als Drittanbieter von Applikationen in der jeweiligen technischen Welt des Anbieters in Erscheinung treten. Die Entwicklung von Skills zu Zwecken der Befragung ist grundsätzlich denkbar. Die technische Entwicklung für verschiedene geschlossene Systeme widerspricht jedoch der – in der Marktforschung wichtigen – Idee von Repräsentativität. Tab. 1 zeigt eine kurze Übersicht der Entwicklungsumgebungen führender Anbieter, die sich prinzipiell für eine Realisierung in der Marktforschung eignen könnten.

| Amazon Skills | Google ACTIONS | Microsoft Cortana Skills | Apple SirKit |
|---|---|---|---|
|  |  |  |  |
| Amazon 2019 | Google 2019a | Microsoft 2019a | Apple 2019 |

**Tab. 1** App Markets für Voice Skills

Nutzer von Amazon, Google oder Apple unterscheiden sich bereits aufgrund der getroffenen Geräteentscheidung soziodemographisch voneinander. Im Vergleich zu NICHT-Verwendern derartiger Geräte sind die Unterschiede erheblich. Wer jemals eine mobile App Softwareentwicklung aufgrund der Komplexität der Beschickung von “nur” zwei mobilen Betriebssystemen (Android und iOS) erlebt hat, wird die parallele Entwicklung von vier oder mehr Plattformen eher scheuen. Kapitel 1.4. beschreibt, wie ein Unternehmen diese Problematik im Bereich der Marktforschung umschifft hat, um eine einsatzfähige Lösung zu Zwecken der Marktforschung zu realisieren. Zuvor wird die prinzipielle Arbeitsweise von Sprachsynthese und Spracherkennung erklärt, um darauf aufbauend die Idee der Unterstützung digitaler Interviews durch Sprache zu verdeutlichen.

## 1.2 Synthetisierung von Stimmen (Text2Speech)

Die ersten Ideen der künstlichen Erzeugung von Sprache gehen zurück ins 18. Jahrhundert, als verschiedene Entwicklungen einer “Sprachorgel” konkurrierten. Die erste – seinerzeit unbeachtete – Entwicklung einer “Sprechmaschine” wird Wolfgang von Kempelen zugeschrieben (Hoxbergen 2005, S. 13 f.), der als Vordenker der heutigen Systeme angesehen werden kann. 1939 gelang dem Erfinder Homer Dudley mit dem „Voice Operation Demonstrator“ (Voder) die erste elektronische Umsetzung von Sprache aus einer Maschine, die jedoch von einer Person in Form einer Klaviatur bedient wurde (Dudley 1940). 1976 beschreibt Flanagan sprechende Maschinen als Interaktionspartner, die „beinahe auf Augenhöhe, aber als Diener des Menschen“ bereits einsatzfähig sind (Flanagan 1976). Größere Konsumentengruppen werden in den 1980er Jahren erstmals mit der künstlichen Erzeugung von Sprache konfrontiert, als der Commodore C64 das “Magic Voice Speech Module” einsetzt, welches einen Wortschatz von 235 Äußerungen umfasste (Uhlmann 2018). Die ersten Personal Computer kamen noch ohne Soundausgabe aus und konnten lediglich Pieptöne erzeugen. Erst mit dem Synthetisieren von Sprache durch zusätzliche Sound-Hardware seit Anfang der 1990er Jahre können komplexere Töne erzeugt werden, zu denen auch die Ausgabe von Sprache gehört. Inzwischen sind diese Technologien fester Bestandteil jedes modernen Digitalgerätes und ermöglichen neben Telefonaten auch den Konsum von Musik und Film. Diese Systeme haben schnell Einzug gehalten in die Unterstützung von

Menschen mit visuellen Handicaps, welche solche Vorlesesysteme in unterschiedlichen Kontexten nutzen (vgl. zur Technologie Lütters 2018).

In den letzten Jahren haben sich verschiedene Unternehmen daran gemacht, Sprache zu synthetisieren, indem ein vorgegebener Text durch eine artifizielle Stimme vorgelesen werden kann. Das Projekt „tacotron“ aus dem Hause Google kann derzeit als technische Benchmark-Referenz angesehen werden. Unter dem Begriff „WaveNet“ werden künstliche Stimmen angeboten, die Sprache auf natürliche Art und Weise wiedergeben und nur noch schwer von der Stimme echter Menschen unterschieden werden können. Basis dieser Sprachausgabesysteme ist ein Training eines Neuronalen Netzwerkes, welches durch weitere gesprochene Inhalte immer weiter verbessert werden kann. Google bietet derzeit 17 Sprachen an, die in unterschiedlicher technischer Qualität von einer männlichen oder weiblichen Kunststimme vorgelesen werden können. Die aktuelle Liste ist einsehbar unter https://cloud.google.com/text-to-speech/docs/ voices. Hierbei fällt auf, dass die Entwicklung tendenziell wirtschaftlichen Interessen folgt und die neueste Technologie immer in den wirtschaftlich relevanten Märkten eingesetzt wird. Um dieses wichtige Entwicklungsfeld nicht den Internetgiganten zu überlassen, hat sich die Mozilla Foundation die Aufgabe gestellt, alle Sprachen des Planeten zu erfassen, um, dem Open-Source Gedanken folgend, jede Art von Sprache zur Verwendung in Technologien zu ermöglichen. Es ist im Projekt https://voice.mozilla. org/ möglich, auf zwei Ebenen an der allgemeinen Verfügbarkeit von Speechsystemen mitzuwirken: Zum einen kann die eigene Stimme “gespendet” werden, indem Texte vorgelesen werden, die das System zum Training der eigenen Stimme verwendet. Zum anderen werden eingesprochene Texte von einer Community bewertet und gelangen erst nach erfolgreicher Prüfung in die Trainingsdatenbank des Systems. Langfristig können damit Fachtexte, aber auch nicht klassische Texte von einem lernenden System besser verstanden werden. Das Projekt wird damit in der Lage sein, irgendwann auch Sprachen zur Verfügung zu stellen, die für die großen Anbieter wirtschaftlich weniger interessant erscheinen. Derzeit sind bereits erste Samples in den seltenen Sprachen Tschuwaschisch oder Kinyarwanda verfügbar. Es ist beim heutigen Entwicklungsstand davon auszugehen, dass bereits in wenigen Jahren Bots in der Lage sein werden, die menschliche Stimme in Annäherung an Perfektion zu imitieren. Unter dem Begriff “Voice Cloning” wird das Synthetisieren von persönlichen Stimmen zur Forschungsdisziplin gemacht, wobei die Gefahren häufig hervorgehoben werden (vgl. Lorenzo-Trueba et al. 2019). Insgesamt ergeben sich allerdings durch neue Stimmen eine Vielzahl von Optionen zur Durchführung gesprochener Interviews.

## 1.3 Spracherkennung und Transkription (Speech2Text)

Die heute noch herausfordernde Aufgabe ist die Erkennung des gesprochenen Wortes bei Umwandlung in geschriebenen Text. Während noch vor wenigen Jahren Sprecher mühevoll und individuell den Systemen die eigene Stimme vermitteln mussten, können heute bereits ohne persönliche Trainings Geräte angesprochen werden, die im Hintergrund die Sprache transkribieren, um daraus einen Befehl zu erkennen, der dann ggf. abgearbeitet werden kann. Neuere Systeme können inzwischen verschiedene Sprecher treffsicher differenzieren.

### 1.3.1 Verfügbare Sprachen bei Speech2Text

Als Anbieter für Spracherkennung stellen sich verschiedene internationale Unternehmen zur Verfügung. Ein Pionier in diesem Bereich ist das Unternehmen IBM, das einige Entwicklungen im Bereich der Spracherkennung wesentlich vorangetrieben hat. Mit dem Aufkommen von Sprache als Interaktionsmöglichkeit treten die Konzerne Microsoft, Amazon, Google, Facebook, Samsung und neuerdings Huawei auf den Plan. Sie alle bieten in unterschiedlich vielen Sprachen Angebote zur Erkennung und Transkription an. Während IBM derzeit sieben Sprachen als API anbietet (vgl. Rakuten 2019), umfasst das Angebot von Google im November 2019 ca. 120 Sprachen in der Spracherkennung (vgl. Google 2019b). Unter den als “Sprache” gezählten Einheiten sind jedoch viele Varianten der Sprachen Englisch, Spanisch und Französisch zu finden sind. So erlaubt Google ein südafrikanisches Englisch und ein US-amerikanisches Spanisch, neben afrikanischen Dialekten von Französisch. Das Angebot enthält auch vermeintlich exotische Sprachen, wobei Google einen klaren Schwerpunkt nach wirtschaftlichen Interessen zu setzen scheint. Eine Sprache muss demnach anscheinend in einem wirtschaftlich relevanten Land oder aber von einer großen Anzahl Menschen gesprochen werden, um eine lokale Anpassung zu rechtfertigen. In Summe wird damit die Erste Welt abgedeckt und Länder mit einer Millionenbevölkerung, für deren Bedarf sich eine technische Entwicklung aus Sicht von Google zu lohnen scheint.

Das Mozilla Projekt hingegen verschiebt auch hier die Grenzen und kümmert sich mit weitaus bescheideneren Mitteln ohne wirtschaftliche Vorbehalte um möglichst viele Varianten unterschiedlicher Sprachen. In dem Projekt engagieren sich Menschen z.B. um die Rettung der sorbischen Sprache in die Digitalisierung etc. Eine Liste weiterer Anbieter, die Sprache via API ermöglichen, findet sich auf Rakuten 2019.

## 1.4 Entwicklung eines Erhebungssystems mit BYOD

Aufgrund der dargestellten marktlichen Komplexität hat das Unternehmen pangea labs einen eigenen Ansatz geschaffen, der auf möglichst vielen Geräten unter Verwendung von Voice zum Einsatz gelangen soll (pangea 2019). Das Fragebogensystem questfox soll hierbei “Device agnostisch” operieren und auf möglichst vielen Plattformen ohne Installation von Zusatzsoftware als Browseranwendung einsatzbereit sein.

1. Der Grundgedanke ist, dass nahezu jeder aktive Konsument in Europa über ein eigenes Gerät mit Mikrofon verfügt, das zu Zwecken der Marktforschung eingesetzt werden könnte (BYOD – Bring your own Device). Damit werden etwaige Repräsentativitätsprobleme nicht ausgeschaltet, aber zumindest durch einen erheblich größeren Pool potenzieller Teilnehmer verringert.
2. Aus Sicht der Marktforschung besteht der große Vorteil bei diesem Ansatz in der Möglichkeit, auf die bestehende Infrastruktur der Marktforschung zugreifen zu können. Der Zugriff auf Online-Access-Panels mit den etablierten Mechanismen von Quotierung und Filterung erlauben den sofortigen Einsatz dieser Technologie in bevölkerungsrepräsentativen Stichproben. Jeder Felddienstleister kann sofort mit seinen Probanden derartige Studien umsetzen.

3. Das Tool kann neben Standardfragen der Marktforschung auch Voice- und Videofragen verarbeiten und diese kombinieren. Auf diese Weise entsteht ein neuer Typus des Hybridfragebogens mit unterschiedlichen Gestaltungsmöglichkeiten. Durch die Möglichkeit, auch visuelle Stimuli verarbeiten zu können, fällt das System nicht hinter andere Fragebogenarten zurück. Es ist also keine Entweder-oder Entscheidung zwischen Text oder Speech, sondern Studien in allen Ausprägungen sind denkbar.
4. Das System erlaubt das Vorlesen von geschriebenen Texten (Text-to-Speech/ TTS). Je nach gewählter Sprache stehen hier unterschiedliche Sprecher zur Verfügung. Während die Stimmen in englischer Sprache bereits hoch entwickelt sind, ist die Auswahl in deutscher Sprache auf zwei Sprecher beschränkt.
5. Von Audio- bzw. Videodateien können aus der gesprochenen Sprache automatisch Transkripte erstellt werden (Speech-to-Text/STT). In dem System sind verschiedene Fragetypen verfügbar, die entweder im Hintergrund transkribieren oder aber eine Live-Korrektur der Transkripte erlauben. Die Erweiterung der Transkriptionsfunktionalität der gesprochenen Sprache aus einem Video ist eine weitere interessante Möglichkeit zur Digitalisierung qualitativer Forschung.
6. Die auf diese Weise gewonnenen Datenformate ermöglichen eine Auswertung der Texte, der Audios oder gar der Videos in Applikationen Dritter. Derzeit ist die Textanalyse sicherlich das etablierteste Verfahren der Auswertung. Die Transkripte können in allen gängigen Tools weiterverarbeitet werden. In Zukunft werden aber über sogenannte “Cognitive Services” auch Videoanalysen und Stimmanalysen zum Alltag des Marktforschers gehören.

Ab. 2 stellt diese neue Maschinerie der Marktforschung als Schaubild entsprechend der vorgestellten Schritte dar.

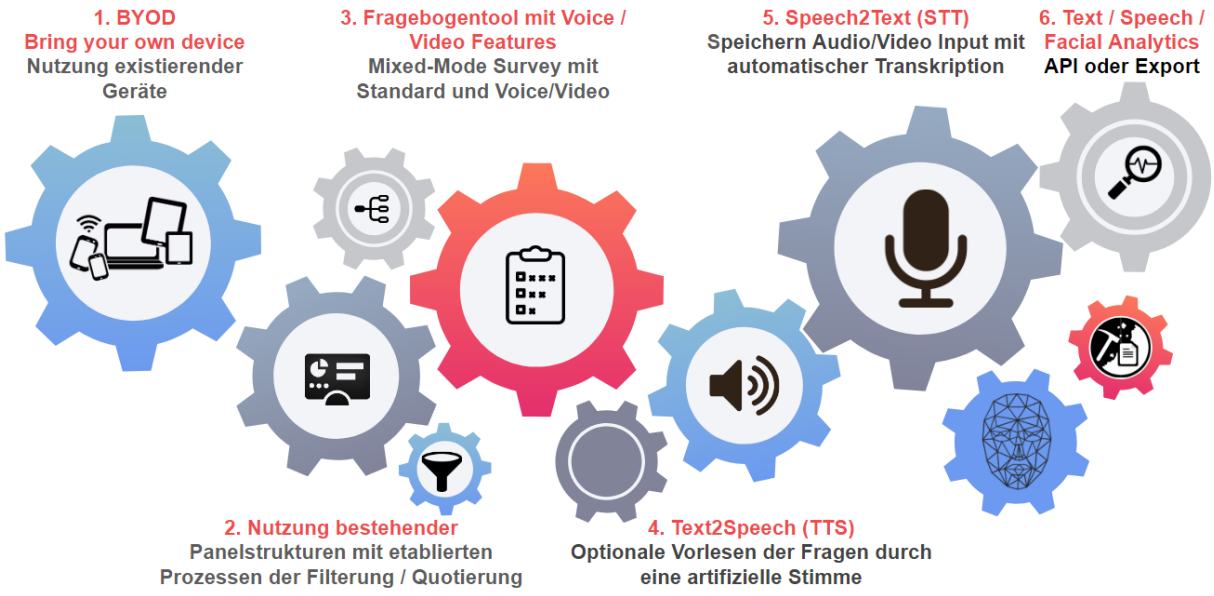



**Abb. 2** Übersicht über das Gesamtsystem zur Integration von Sprache in Interviews

# 2. Grenzen des Ansatzes

Die Idee BYOD erweitert das Spektrum möglicher Teilnehmer erheblich. Dennoch sind hier auch technische Grenzen zu erwähnen, die einen problemlosen Einsatz verhindern können. Folgende Voraussetzungen sind derzeit zu nennen, wenn es um einen großflächigen Einsatz der Erhebung per Sprache geht.

## 2.1 Moderner Browser erforderlich

Voice Technologien sind nur einsatzfähig auf Geräten jüngeren Entwicklungsdatums. Es ist schlichtweg technisch ausgeschlossen, dass ein zehn Jahre alter Browser, wie z.B. der Internet Explorer, mit einer drei Jahre alten Technologie kompatibel ist. Da jedoch der genannte Browser in der Praxis immer noch anzutreffen ist und sowohl im privaten als auch im geschäftlichen Bereich verbreitet ist, stellt dies ein Problem dar. Zum jetzigen Entwicklungsstand erlaubt der Anbieter Apple keine Verwendung von Mikrofonen in seinen Browser Technologien (Safari). Dies bedeutet, dass viele Apple-Nutzer momentan von dieser Idee der Erhebung ausgeschlossen sind.

Weltweit sollten derzeit theoretisch 73 Prozent der Internetuser an einer derartigen Studie teilnehmen können (vgl. die aktuelle Verbreitung auf canIuse). Dies bedeutet aber, dass ein Viertel potenzieller Teilnehmer per Definition durch Studienansätze mit Speech ausgegrenzt wird.

## 2.2 Dauerhafte Internetverbindung mit hohem Datendurchsatz

Das System ist nur funktionsfähig, wenn eine permanente Internetverbindung den Austausch zwischen Client und Server ermöglicht. Im Bereich der Video-Befragung sind höhere Bandbreiten erforderlich als bei der Übertragung von Sprache. Diese, von einigen Marktforschern als Einschränkung empfundene Voraussetzung wird bereits in wenigen Jahren keine spürbare Restriktion mehr darstellen, da mit dem Ausbau des 5G Netzes mobiles Internet in jedem Winkel eines Industrielandes zur Verfügung stehen wird. Eine asynchrone Aufzeichnung von Sprache und anschließende Transkriptionen bei bestehender Internetverbindung ist grundsätzlich denkbar, verhindert jedoch die Automatisierung dieser Prozesse.

## 2.3 Funktionsfähiges Mikrofon und Erlaubnis der Nutzung

Viele moderne Geräte sind bereits hardwareseitig mit Mikrofonen ausgestattet. In den Vorstudien erwies sich insbesondere der Desktop-PC als kritisches Gerät, da hier meistens keine feste Mikrofon-Installation vorzufinden ist. In konkreten Tests gaben die Teilnehmer zwar an, über ein Mikrofon zu verfügen, dieses war jedoch oft nicht einsatzfähig.

Weitaus gravierender als die mangelnde technische Funktionsfähigkeit ist die mangelnde Bereitschaft der Menschen, an den neuartigen Sprachinteraktionen teilzunehmen. Diese, derzeit noch ungewohnte, Interaktionsform wird jedoch zunehmend akzeptierter und findet immer mehr Nutzer auch in anderen Bereichen der Softwareentwicklung. Für Marktforscher besteht jedoch die Notwendigkeit, bei jeder Aufzeichnung der Stimme die explizite Genehmigung der Teilnehmer zu erfragen. Dies führt immer wieder zu Ausfällen bei der Teilnehmerschaft.

In Vorstudien haben sich zwei gangbare Wege im Umgang mit den technischen Restriktionen herausgestellt:

1. Direkte Konfrontation des Befragten mit einer Audio Frage und Selbsteinschätzung der Transkriptionsqualität als Filterfrage.
2. Vollständige technische Eingrenzung der Machbarkeit durch Abfrage von Betriebssystem, Browser, technischer Ausstattung und Teilnahmebereitschaft. Danach erst Schritt eins der Konfrontation mit einer Testfrage.

Die nachfolgend dargestellte Studie geht den zweiten Weg der aufwendigen Erhebung aller technischen Details zur Sicherstellung der technischen Qualität.

## 2.4 Umfang und technische Qualität der Audioantworten

In den ersten Studien mit Voice ergeben sich Antwortmuster, die ohne besondere Anweisungen im Schnitt gesprochene Antworten von durchschnittlich sechs bis acht Sekunden ergeben. Da einige Teilnehmer die Möglichkeit der Sprache weitaus intensiver nutzen, sollte der Median der Antwortzeit hier als Eckwert dienen. Die kürzesten Antworten liegen bei ca. zwei Sekunden und unterscheiden sich inhaltlich meist nicht von einem getippten kurzen Statement.

Die Qualität der Transkriptionen von Audio-Dateien hängt wesentlich vom technischen Setting und den gegebenen Umweltstörfaktoren in der Befragungssituation ab. Viele Devices enthalten heutzutage bereits Mikrofone von guter Qualität. Auch günstige Smartphones sind zumeist mit hinreichenden Technologien ausgestattet, um Transkriptionen fehlerfrei zu ermöglichen.

Das jeweilige Umfeld in Form einer Geräuschkulisse ist für die korrekte Umwandlung kritischer zu bewerten. So ist eine Erkennung während einer Busfahrt aufgrund von Störgeräuschen erheblich fehleranfälliger als in einem Büro, wenn mit Headset gearbeitet wird. In ersten Studien mit wiederholten Aufzeichnungen waren auch Effekte der Probandenermüdung zu konstatieren. Während in den ersten Minuten einer Studie noch perfekte Transkripte entstanden, sank die Qualität im Interviewverlauf.

Diese Qualitätseinbußen sind auch in einigen Tools quantifizierbar. So erzeugen einige Tools einen ***Transcription Quality Score***, der zwischen null und eins (für perfekte Qualität) liegt. Die Erkennungsraten sind bei Werten oberhalb von 0,8 bereits sehr gut. Weitere Studien müssen versuchen, die tolerierbare Grenze der Transkriptionsqualität zu ermitteln. Tab. 2 zeigt die Ergebnisse eines online repräsentativen Samples für eine Voice-Studie mit vier getrennten Befragungsgruppen.

| | | Antwortzeit | | | | |
|---|---|---|---|---|---|---|
| n | Transcription Confidence Score | Mean | Std Dev. | Median | Min | Max |
| 126 | 0,84 | 8,1 | 6,0 | **6,4** | 2,1 | 37,1 |
| 132 | 0,84 | 9,2 | 6,4 | **7,1** | 2,1 | 33,1 |
| 157 | 0,82 | 8,8 | 6,1 | **6,9** | 2,1 | 32,3 |
| 139 | 0,83 | 9,9 | 6,9 | **8,1** | 1,9 | 37,0 |

**Tab. 2** Quantitative Ergebnisse von vier Voice Gruppen im Vergleich (Quelle: Freksa et al. 2019**)**

## 2.5 Neue Wertkette der Forschung mit Systemen Dritter

Der vorgestellte Ansatz enthält an verschiedensten Stellen Technologien unterschiedlicher Unternehmen. Dies bedeutet, dass der Forscher sich auf deren technische Fortentwicklung verlässt. Etwaige Änderungen, z.B. im Transkriptionsmodus, hätten direkte Auswirkungen auf die eigene Arbeit, ohne dass der Forscher darauf Einfluss nehmen könnte. Entstehende Abhängigkeiten sind daher kritisch zu betrachten.

Gleichzeitig betritt der Forscher in dieser API-Landschaft ein neues Gebiet, bei dem Bausteine der Forschung auch relativ unkompliziert durch andere Anbieter übernommen werden können. Es entsteht demzufolge eine neue Art Wertkette der Forschung im eigenen Projekt. Der Forscher wird zukünftig die Wahl haben zwischen konkurrierenden Ansätzen auf den verschiedenen Ebenen. Dies darf als vorteilhaft angesehen werden, da damit immer mehr neue Module in die Forschungsprojekte Einzug halten können. Die Rüstzeiten für Forschung werden damit steigen, da Forscher sich in neue Themengebiete einfinden müssen.

## 2.6 Veränderungen der Budgetplanung

Während Unternehmen heute oft mit einem Tool für Befragungen nach einem Lizenzkostenmodell arbeiten, wird sich diese Welt wahrscheinlich fragmentieren. Diese Art der Forschungsplanung erfordert ein Umdenken bei Fragen der Budgetierung von Forschung. Viele der genutzten Anwendungen werden nach dem Pay-per-Use Prinzip abgerechnet, was Unternehmen, aber insbesondere auch Forschungseinrichtungen vor Probleme der Budgetierung stellen wird. Die Marktforschung kennt bisher Preismodelle, bei denen nur vollständige Interviews abgerechnet werden. Die API-Welt kennt diese Gedankenwelt nicht und rechnet die Nutzung pro Minute ab – unabhängig davon, ob aus der erbrachten Leistung im Interview ein abgebrochenes oder ein vollständiges Interview entsteht. Diese zu erwartenden höheren Kosten sind bei der Planung zu berücksichtigen. Der Umgang mit nicht genau planbaren variablen Kosten ist institutionell zu erlernen.

### 2.7 Motivationale Aspekte der Teilnahmebereitschaft

Die eingesetzte Technologie ist so jung, dass selbst erfahrene Respondenten noch keinen Kontakt zu dieser Art von Interaktionsmöglichkeit hatten. Die Ablehnungsquote der Verwendung des Mikrofons als Antwortmöglichkeit ist dementsprechend ungewohnt und führt zu verstärkter Verweigerung.

## 3. Studie Text vs. Speech vs. Video

Um die Leistungsfähigkeit der neuartigen Erhebungsmethode zu überprüfen, wurde eine empirische Studie realisiert, bei der eine Kontrollgruppe eine klassische Befragung mit Texteingabe erhielt (A), während in zwei Experimentalgruppen die Varianten Voice (B) und Video (C) im Vergleich zur klassischen Dateneingabe verglichen wurden.

Insgesamt entsteht ein erheblicher Aufwand zur Durchführung der Interviews. In einem online repräsentativen Sample wurden 9.818 Personen kontaktiert. Während in der normalen Textversion mehr als jeder Dritte das Ende des Fragebogens erreicht, liegt dieser Wert sowohl bei Voice als auch bei Video unterhalb von 10%. Auf Inputseite werden also dreimal so viele Probanden benötigt wie in klassischen Studien. Dies liegt zum einen in der Technologie begründet, die von vielen Menschen (noch) nicht beherrscht wird. Jedoch ist auch ein erheblicher Anteil der Menschen derzeit nicht bereit, eine solche Befragung per Sprache oder Video durchzuführen. Tabelle 3 zeigt die Eckdaten der Studie im Überblick.

| Version | **Kontrollgruppe A Text** | **Experimentalgruppe B Voice** | **Experimentalgruppe C Video** |
|---|---|---|---|
| |  |  |  |
| Starter | 1218 | 4422 | 4178 |
| Screen-out* | 781 | 2765 | 2499 |
| Quota-Full | 435 | 108 | 70 |
| Drop-out | 215 | 1142 | 1202 |
| Complete Interviews (%) | 437 (35,9%) | 407 (9,2%) | 407 (9,7%) |
| * Ausschluss vom Interview aufgrund nicht erteilter Erlaubnis, das Mikrofon zu nutzen oder mangelnder Transkriptionsqualität | | | |

**Tab. 3** Antwortverhalten in den drei Gruppen

Diese Werte werden sich in der Zukunft wahrscheinlich verbessern, jedoch ist davon auszugehen, dass ein Voice- oder Videointerview eine andere Art von Aufmerksamkeit erfordert. Auch ist die mangelnde Bereitschaft der Teilnehmer insbesondere auf das jeweilige situative Umfeld zurückzuführen, in welchem es der Teilnehmer als unangemessen empfindet, an einer Voice Studie teilzunehmen. Fragebögen, die bisher in Großraumbüros per Tastatur und Maus ausgefüllt wurden, werden nicht unproblematisch in einen neuen Modus überführbar sein, da Menschen unter Beobachtung eine andere Teilnahmebereitschaft zeigen werden.

Auf technischer Ebene wird dadurch ein Bias erzeugt, der sich nicht mathematisch korrigieren lässt. Derzeitige Samplings sind in der Regel etwas zu technikaffin ausgeprägt, was sich soziodemographisch mit zu jung und zu männlich übersetzen lässt. In Panelstudien lässt sich dieser Effekt durch Quotierungen zwar ausgleichen, jedoch entstehen dadurch in der Regel höhere Kosten des Samplings.

# 4. Auswertung gesprochener Sprache

## 4.1 Auswertung der transkribierten Texte

Die gesprochenen Inhalte der Video- als auch der Voice-Route liegen als transkribierter Text vor. An Paradaten entsteht z. B. Antwortzeit sowie in neueren Systemen ein Transkriptionsqualitätsscore, der die Wahrscheinlichkeit für ein korrektes Transkript auf Basis der vorliegenden Audio-Information schätzt. Auf Basis dieser Informationen können klassische textanalytische Auswertungen vorgenommen werden.

Abb. 3 zeigt einige zentrale Eckwerte der Studie. So ist zu konstatieren, dass in der Voice-Route in etwas weniger Zeit deutlich mehr Inhalte hinterlassen werden. Die Video-Route hingegen erzeugt nur geringfügig mehr Inhalt, benötigt dafür aber erheblich mehr Antwortzeit.

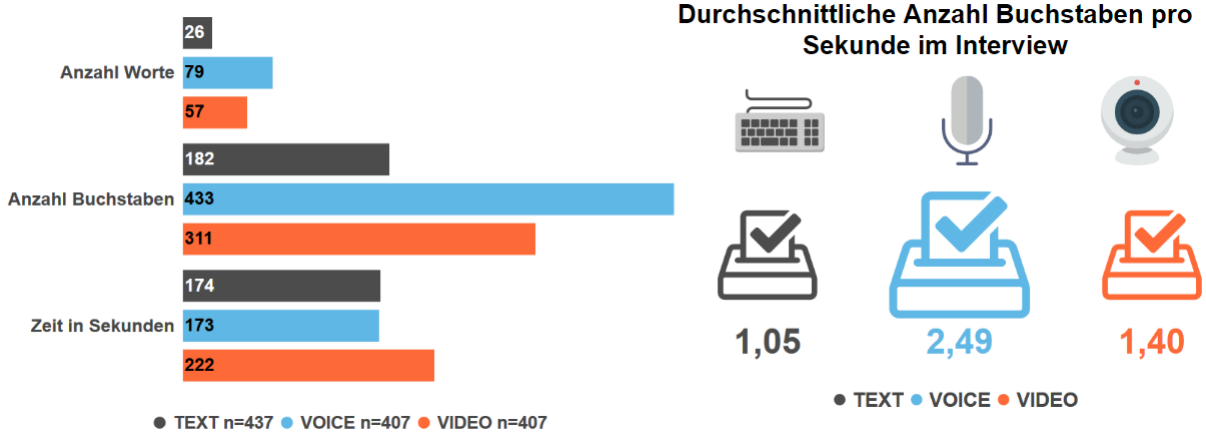



**Abb. 3** Nicht-reaktive Daten Text vs. Voice vs. Speech

Dieses Ergebnis kann mit der vermuteten höheren kognitiven Belastung bei der Auseinandersetzung im Videoformat entstehen. Der Antwortende scheint zu einem gewissen Grad mit sich selbst beschäftigt zu sein, was wiederum seine sprachliche Ausdrucksweise zu hemmen scheint.

Der in dieser Studie vorgefundene Multiplikator Faktor 2,5 auf offene Antworten für Voice darf als durchaus realistisch erreichbares Ergebnis angesehen werden. In weiteren Experimenten wird nun versucht, durch Neugestaltung der Fragestellungen diesen Wert sogar noch zu erhöhen.

Die etablierten Textanalyse-Tools bieten sich auch in diesem Kontext an. In der konkreten Studie wurde ein ausgefeiltes Textanalyse-System der Unternehmung insius zum Einsatz gebracht. Dieses dem Social Media Research entstammende Tool kann inhaltsanalytische Auswertungen auf Basis großer Textmengen erstellen.

Im direkten Vergleich der drei Erhebungsmodi fällt insbesondere auf, dass die offene Textantwort letztendlich zu Kurzphrasen führt (z.B. “war ok”). Die Monolog-Analyse der gesprochenen Anteile zeigt hingegen auf, dass Menschen in einer Antwort gleichzeitig mehrere Themenkreise adressieren. Abb. 4 zeigt die textanalytische Aufbereitung der Studie mit dem semantischen Analysetool insius (vgl. zur Funktionsweise dieser Analyse Lütters/Egger 2013).

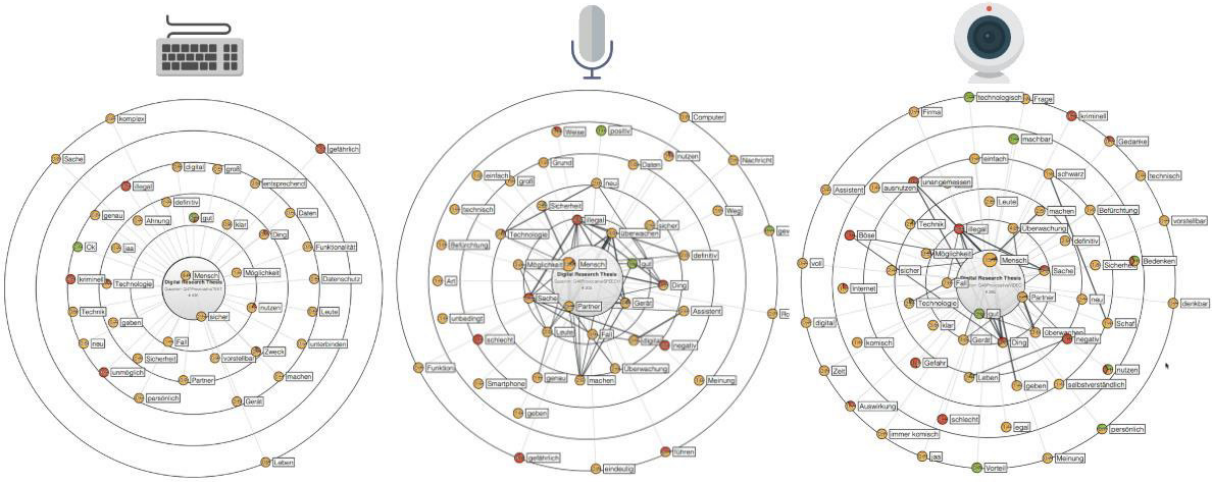


**Abb. 4** Textanalyse der Daten aller drei Routen

Die Auswertung derartiger Formate lässt sich nur unzureichend in statischer Form darstellen. Die zukünftige Entwicklung wird höchstwahrscheinlich als serverseitige Auswertung mit Machine Learning Algorithmen geschehen. Derartige Systeme verbessern die Auswertung durch jede durchgeführte Studie. Dies stellt einen Wendepunkt in der Verwendung von Textanalyse-Tools dar, die bisher mühsam von jedem Forscher in Einzelarbeit in seinem Spezialgebiet codiert und trainiert werden mussten.

## 4.2 Auswertung der Audio- und Video-Daten

In der Studie sind über 7.000 anonymisierte Audio und Video-Dateien entstanden, die mit klassischen Methoden der Analyse nicht auswertbar erscheinen. Die Zukunft wird

verschiedene Systematiken bringen, bei denen sowohl Sprachdateien als auch Videodateien besonderen Analyseverfahren zugeführt werden können. So gilt z. B. die Voice Stress Analysis CVSA als etabliertes Verfahren, welches in Zukunft auch mit großen Datenmengen problemlos arbeiten können wird (vgl. Damphousse et al. 2007). Hierbei werden verschiedene Charakteristika aus der Sprache extrahiert und in die entsprechenden Analyse Engines überführt. Die zuweilen als Lügendetektoren bezeichneten Verfahren können aus der Art des Stimmeinsatzes kognitiven Stress erkennen, der bei der Formulierung von Sprache entsteht. Auch wenn diese Ideen derzeit noch nicht konkurrenzfähig erscheinen im Vergleich zum Polygraphen, so ist der Gedanke der Digitalisierung hierbei in jedem Fall ein Argument, um eine Methode aus der Laborsituation in die Feldforschung zu überführen.

Im Bereich der Videoanalyse sind derzeit gewaltige Entwicklungen im Gange. Verschiedene Anbieter können aus Bildinhalten unterschiedliche Analysefunktionen auslesen. Abb. 5 zeigt das Ergebnis der Verwendung der sogenannten cognitive API von Microsoft, welche eine maschinell erstellte Analyse des Bildinhalts ermöglicht (Microsoft 2019b). Hierbei werden Bildinhalte per Algorithmus in auswertbare Daten überführt. Von der soziodemographischen Einordnung über Alter und Geschlecht bis hin zum emotionalen Zustand unter Einbeziehung persönlicher Merkmale wie Haarwuchs-Levels sind vielfältige Ergebnisse auf Basis der Bild-Informationen bereits zu erreichen.


**Abb. 5** Analyse Bild des Autors mit Microsoft cognitive services (Quelle: Microsoft 2019b**)**

Derartige Auswertungen werden zukünftig auch auf Basis von Bewegtbildern angeboten.

## 4.3 Nutzerbewertung der User-Experience im Speech Interview

Für alle Panelteilnehmer der Studie war es die erste Teilnahme mit den Optionen Speech und Video. Es war daher von besonderer Bedeutung, die Nutzererfahrung auch quantitativ bewerten zu lassen. Hierzu wurde der Net Promoter Score (NPS) eingesetzt (vgl. Reichheld 2011). Diese Frage wurde den marktforschungserfahrenen Panelisten am Ende der Befragung präsentiert.

Während der NPS bei Standardbefragungen fast immer negativ ausfällt, wird sowohl die Voice- als auch Video-Route positiv bewertet. Die abgefragte Antwortzeit für die Bearbeitung der Befragung liegt in allen drei Routen unterhalb der tatsächlich benötigten Antwortzeit (vgl. Vargas Quintana 2018 S. 53).

| Version | **A Text n=437** | **B Voice n=407** | **C Video n=407** |
|---|---|---|---|
| |  |  |  |
| Promotoren | 144 (32,95%) | 155 (38,08%) | 163 (40,05%) |
| Passive | 111 (25,4%) | 105 (25,08%) | 96 (23,59%) |
| Detraktoren | 182 (41,63%) | 147 (36,11%) | 148 (36,37%) |
| **NPS** | **-8,68%** | **1,97%** | **3,68%** |
| Benötigte Antwortzeit (Min.) | 13,53 | 14,09 | 16,91 |
| Geschätzte Antwortzeit (Min.) | 10,08 | 10,05 | 11,74 |
| Bereitschaft digitales Interview* | 5,34 | 7,12 | 7,08 |
| * Wären Sie bereit, in Zukunft ein vollständiges Audio-Interview mit einem digitalen Assistenten zu führen? Skala 0=auf gar keinen Fall … 10=auf jeden Fall | | | |

**Tab. 4** Quantitative Bewertung der Befragung in den drei Routen

Auf die Frage, ob sich die Menschen ein vollständig sprachgesteuertes Interview vorstellen können, antworten diejenigen Menschen mehrheitlich positiv, die gerade ihre erste Erfahrung mit dieser Interaktionsform hatten sammeln können.

# 5. Ausblick

Die verschiedenen Varianten der neuen Interaktionsform Sprache zum Einsatz in der Marktforschung funktionieren. Auch wenn zum heutigen Zeitpunkt noch keine bevölkerungsrepräsentativen Stichproben möglich sind, so wird dieses Hemmnis bereits in wenigen Jahren aufgrund weiterer Verbreitung der Geräte in der Bevölkerung ausgeräumt sein. Gleichzeitig interagieren mehr und mehr Menschen mit Sprach-Devices, was zu einer höheren Akzeptanz der Nutzung auch in der Marktforschung führen wird.

Die Marktforschung ist im Thema Voice nicht Vorreiter, sondern wird letztendlich von den Entwicklungen großer Anbieter im Markt getrieben. Es ist dringend geboten, Standesregeln für den Berufsstand der Marktforscher im Umgang mit Sprache und Bewegtbildaufnahmen zu definieren. Die aufgezeigten Entwicklungen zur Synthetisierung von Sprache sollten Anlass genug sein, sich intensiv mit einem Regelwerk für den heutigen aber auch zukünftigen Umgang mit den Technologien und deren Analysemöglichkeiten zu beschäftigen.

Der Austausch offener Fragetypen durch Speechfragen ist bereits heute problemlos möglich. Die Idee greift aber zu kurz und wird dem disruptiven Potenzial nur unzureichend gerecht. Eine völlig neue Art zu befragen steht bevor und diese sollte nun mit der Erprobung neuartiger Konzepte qualitativer und quantitativer Forschung sukzessive erforscht werden.

Die kleine historische Rückschau zu Beginn des Beitrags macht deutlich, dass bisher jede Generation der Entwicklung von Sprechmaschinen für sich in Anspruch genommen hat, sehr nah am menschlichen Imitat operieren zu können. In der Rückschau stellen sich diese ersten Gehversuche als sehr bescheiden dar. Vielleicht wird man aber auch die Zukunft milde über unsere heutigen Versuche lächeln, die nicht mehr als der Anfang einer bahnbrechenden Entwicklung sein können.

# Literatur

Amazon (2019): Alexa Skills Kit in Deutschland. https://developer.amazon.com/de/alexa-skills-kit, Zugriff: 15.11.2019.

Apple (2019): Siri for Developers. https://developer.apple.com/sirikit/, Zugriff: 15.11.2019.

canIUse: Statistiken über die Verbreitung von Browsertechnologien und deren Funktionsweisen, https://caniuse.com/#search=getusermedia, Zugriff: 05.03.2020.

Damphousse, K.; Pointon, L.; Upchurch, D.; Moore, R.; Winscher, T. (2007): Assessing the Validity of Voice Stress Analysis Tools in a Jail Setting. U.S. Department of Justice; Document No. 219031, https://www.researchgate.net/publication/251785785_Assessing_the_Validity_of_Voice_Stress_Analysis_Tools_in_a_Jail_Setting, Zugriff: 05.03.2020.

Dudley, H. (1940): The Carrier Nature of Speech. In: The Bell System Technical Journal. Vol. XIX, October 1940, No. 4, S. 495-515; https://archive.org/stream/bellsystemtechni19amerrich/bellsystemtechni19amerrich, Zugriff: 05.03.2020.

Flanagan, J. L. (1976): Computers that talk and listen: Man-machine communication by voice. Proceedings of the IEEE, 64(4), 405–415. doi:10.1109/proc.1976.10150.

Freksa, M.; Vitt, S.; Lütters, H. (2019: Emotionale KI in der Werbeforschung. Beyond the Real Voice of the Customer. Studie präsentiert auf der Research&Results 2019.

Google (2019a): Integrate with the Google Assistant, https://developers.google.com/assistant, Zugriff: 11.11.2019.

Google (2019b): Cloud Speech-to-Text. Durch maschinelles Lernen unterstützte Umwandlung von Sprache in Text für kurz- und langformatige Audioinhalte, https://cloud.google.com/speech-to-text, Zugriff: 11.11.2019.

Google Research Blog (2018): Expressive Speech Synthesis with Tacotron. https://research.googleblog.com/2018/03/expressive-speech-synthesis-with.html, Zugriff: 05.03.2020.

Horxbergen, A. (2005): Die Geschichte der Sprachsynthese anhand einiger ausgewählter Beispiele. Studienarbeit am Institut für Informatik der Humboldt Universität zu Berlin, http://waste.informatik.hu-berlin.de/Diplom/studienarbeit_hoxbergen.pdf, Zugriff: 05.03.2020.

Lorenzo-Trueba, J. Fang, F.; Wang, X; Echizen, I.; Yamagishi, J; Kinnunen, T. (2019): Can we steal your vocal identity from the internet? Initial investigation of cloning Obama's voice using gan, wavenet and low-quality found data," arXiv preprint arXiv: 1803.00860, 2018.

Lütters, H. / Egger, M. (2013): "Listening is the new asking": Social Media-Analyse in der Marktforschung. In: Transfer Werbeforschung & Praxis Zeitschrift für Werbung, Kommunikation und Markenführung, S. 34-41, 2013.

Lütters, H. (2018): Digitale Sprachassistenz in der Forschung. Von der Technologie zur Anwendung. In: Knaut, M. (Hrsg.): Kreativität + X = Innovation, S. 34-41, BWV Berliner Wissenschafts-Verlag, Berlin, 2018, ISBN 978-3-8305-3844-8, 978-3-8305-4014-4 (E-Book).

Microsoft (2019a): Cortana's got skills, https://developer.microsoft.com/en-us/cortana, Zugriff: 05.03.2020.

Microsoft (2019b): Cognitive Services. A comprehensive family of AI services and cognitive APIs to help you build intelligent apps, https://azure.microsoft.com/de-de/services/cognitive-services, Zugriff: 05.03.2020.

Pangea (2019): Voices of pangea: 120 languages supported in voice and video question types in questfox for real-time transcription, https://questfox.wordpress.com/2019/11/04/voices-of-pangea-120-languages-supported-in-voice-and-video-question-types-in-questfox-for-real-time-transcription, Zugriff: 05.03.2020.

Rakuten (2019): Top 10 Best Speech Recognition APIs: Google Speech, IBM Watson, SpeechAPI, and others, https://blog.api.rakuten.net/top-10-best-speech-recognition-apis-google-speech-ibm-watson-speechapi-and-others, Zugriff: 05.03.2020.

Reichheld, F. F. (2011) The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-Driven World. Boston, Mass.: Harvard Business Press

Revilla, M. / Couper, M. (2019): Improving the Use of Voice Recording in a Smartphone Survey. In: Social Science Computer Review S. 1-20; DOI: 10.1177/0894439319888708.

Uhlmann, S. (2018): Magic Voice Ein Sprachausgabemodul für den C64/C128, http://www.stefan-uhlmann.de/cbm/MVM/index.html, Zugriff: 05.03.2020.

Vargas Quintana, G. (2018): Effects of voice assistance in market research interviews: A comparison of classic vs. speech vs. video interaction. Masterarbeit an der Hochschule für Wirtschaft und Recht Berlin.

## Schlüsselwörter

Marktforschung, Speech in Research, Video in Research, Real Voice of the Customer, Audio, Transcription, API Research

MARKTFORSCHUNG · DIGITALISIERUNG
TECHNISCHER FORTSCHRITT · BIG DATA
INNOVATIONSTEMPO · INTERNET OF THINGS
KÜNSTLICHE INTELLIGENZ · VIRTUAL UND
AUGMENTED REALITY · EMOTIONSMESSUNG
SPRACHASSISTENTEN · NEUROMARKETING
WEARABLES · VITALDATEN · PERSONAS
ANALYTICSPROGRAMME · DIGITALISIERUNG
MARKTFORSCHUNG · BIG DATA · WEARABLES
INTERNET OF THINGS · INNOVATIONSTEMPO
KÜNSTLICHE INTELLIGENZ · VITALDATEN
NEUROMARKETING · ANALYTICSPROGRAMME
EMOTIONSMESSUNG · PERSONAS · GOOGLE
EYETRACKING · SMARTWATCHES · QUALITATIVE
MARKTFORSCHUNG · HANDEL · CHATBOTS

AfM

Arbeitsgemeinschaft
für Marketing